



universität  
wien

# DIPLOMARBEIT

Titel der Diplomarbeit

Molecular Modelling Studies for Analysing Interactions of  
Substrates and Inhibitors of the Serotonin Transporter

angestrebter akademischer Grad

Magister der Pharmazie (Mag. pharm.)

Verfasser:	René Weissensteiner
Matrikel-Nummer:	9906644
Studienrichtung (lt. Studienblatt):	Pharmazie
Betreuer:	Ao. Univ.-Prof. Dr. Gerhard F. Ecker

Wien, im November 2008

---

*Im Gedenken an meine geliebte Tochter Sophie Katharina,  
die im 2. Lebensjahr diese Welt wieder verlassen musste.*

*lasst sie sein, was sie sind,  
gebt ihnen rückenwind,  
auf dass sie glücklich sind.*

*manuel wögerer, deKor*



Hiermit möchte ich mich bei den folgenden Personen danken, die mich während meines Studiums auf vielfältige Weise unterstützten und beeinflussten.

An erster Stelle möchte ich mich bei dem Betreuer meiner Diplomarbeit Herrn Ao. Univ.-Prof. Dr. Gerhard F. Ecker für die Möglichkeit sehr herzlich bedanken, dass ich meine Arbeit als Teil seiner Arbeitsgruppe verfassen durfte.

Weiters danken möchte ich Herrn Ao. Univ.-Prof. Dr. Harald H. Sitte für die sehr gute Zusammenarbeit und die Einführung in die Materie.

Meinem geschätzten Kollegen Mag. Michael Demel möchte ich für die umfangreiche Einführung in das Themengebiet und die vielen Diskussionen danken. Auch der gesamten Arbeitsgruppe spreche ich für das produktive Arbeitsklima meinen Dank aus.

Ich danke meiner Verlobten Martina, die mich während des Studiums tatkräftig unterstützte und mich auch in stressigen Zeiten mit viel Rücksicht erträgt.

Meiner Tochter Sophie Katharina für die viel zu kurze aber bisher wunderbarste Zeit meines Lebens, in der sie es schaffte, meinen Blick auf die vielen faszinierenden Kleinigkeiten im täglichen Leben zu lenken und mir mein Herz für die wesentlichen Dinge öffnete.

Meinen Großeltern Elfriede und Karl Timpel möchte ich für die angenehme Atmosphäre während der Prüfungsvorbereitungen und auch für ihre tatkräftige Unterstützung und Motivation sehr herzlich danken.

Ein besonderer Dank gilt meinen Eltern Sigrid und Josef Weissensteiner, die mir die Ausbildung ermöglichten.

Auch ganz besonders danken will ich meinen Freunden der Musikgruppe deKor, die mir beim Musizieren immer wieder das kreative Ventil öffnen.

Abschließend danke ich meiner ganzen Familie und allen Freunden und Kollegen, die auch in schwierigen Zeiten immer ein offenes Ohr für mich haben.

Danke

# Table of Contents

1. Introduction and Aim of the Study.....	1
2. Anatomy and Biology.....	2
2.1. Monoamine Systems in Brain.....	2
2.2. Cell Membrane and Membrane Proteins.....	6
2.3. Neurotransmitters and Receptors.....	7
2.4. Synapse and Neuronal Transmission.....	9
3. Pharmacology.....	13
3.1. Serotonin.....	14
3.2. Dopamine.....	15
3.3. Norepinephrine.....	17
3.4. Amphetamine.....	17
3.5. MDMA.....	19
3.6. Methylphenidate.....	19
3.7. Cocaine.....	20
3.8. CFT.....	21
3.9. MPP+.....	21
3.10. Psilocin.....	22
3.11. DMT.....	23
4. Pharmacoinformatics.....	24
4.1. Structure of Proteins.....	24
4.1.1. From Gene to Protein.....	24
4.1.2. Structure Analysis.....	29
4.2. Homology Modelling.....	32
4.2.1. The Alignment.....	33
4.2.2. The Template.....	33
4.2.3. The Model.....	34
4.2.4. Refinement and Validation.....	35
4.3. Docking.....	35
4.3.1. The Ligand.....	36
4.3.2. The Site.....	36
4.3.3. The Placement.....	37
4.3.4. Scoring and Ranking the Poses.....	37

5. Neurotransmitter:sodium symporter.....	39
5.1. The Leucine Transporter from Aquifex aeolicus.....	40
5.2. Neurotransmittertransporters.....	45
5.2.1. Structure and Function.....	46
5.2.2. The Serotonin Transporter.....	46
5.2.3. The Dopamine Transporter.....	48
6. Materials and Methods.....	49
6.1. Molecular Operating Environment.....	49
6.2. Homology Modelling.....	50
6.2.1. The Model of the Serotonin Transporter.....	53
6.2.2. The Model of the Dopamine Transporter.....	55
6.2.3. Comparison of the Models.....	57
6.3. Molecular Modelling and Simulation.....	58
6.3.1. Ligand Preparation.....	59
6.3.2. Defining the Binding Site.....	59
6.3.3. Running Docking.....	61
6.3.4. Ranking the Poses.....	61
7. Results and Discussion.....	66
7.1. Docking Ligands into the Serotonin Transporter.....	66
7.1.1. Serotonin.....	66
7.1.2. Methylphenidate.....	69
7.1.3. Cocaine and CFT.....	71
7.1.4. MDMA.....	73
7.1.5. DMT and Psilocin.....	73
7.2. Differences and Similarities.....	74
7.3. Conclusion and Outlook.....	76
Abstract.....	77
Zusammenfassung.....	78
References and Links.....	79
Curriculum Vitae.....	86

# 1 Introduction and Aim of the Study

The serotonin transporter (SERT) is a member of the solute carrier gene 6 (SLC6) family as also the dopamine transporter (DAT) and norepinephrine transporter (NET), that code the neurotransmitter sodium symporters (NSS) and other solute carriers. This group of transmembrane proteins is located on the presynaptic end of neurotransmitter emitting neurons. For rapid termination of neuronal transmission, substrate undergoes a reuptake into the emitting nerve cell, which is dependent on a gradient of sodium in the extracellular space. Although these transporters are highly selective, they not only transport the origin substrate, but also drugs and illicit substances, e.g. amphetamines and hallucinogens. Widely used therapeutics of the central nervous system (CNS) and abused drugs target the SERT, DAT and NET to inhibit its function, e.g. selective serotonin reuptake inhibitors (SSRI), tricyclic antidepressants (TCA), selective norepinephrine reuptake inhibitors (SNRI), methylphenidate (MPD), and psychostimulants like cocaine.

Due to the lack of a crystal structure of the neurotransmitter sodium symporters, computational methods such as homology modelling and docking are needed to predict the three dimensional structure of the protein and to reveal protein ligand interactions. Homology modelling is a probate method for investigations on proteins, where no crystal structures are available. Due to the increasing amount of biological and structural data and the increasing performance of computer technology, the precision of in silico methods is rising.

The crystal structure of the bacterial leucine transporter (LeuT) from *Aquifex aeolicus*, in structure and function very homologue to the mammalian NSS, we chose as template for homology modelling. Docking usually is used for high throughput screening of putative new drugs in protein crystal structures. In the academic field, molecular docking is used to elucidate binding modes in either crystal structures or homology models.

The aims of the present study are to create a homology model of SERT, DAT and NET, to run docking with substrate and inhibitors on SERT and DAT, to hypothesize interactions and binding modes and to give suggestions for further mutations.

## 2 Anatomy and Biology

The main tasks of the nervous system are the regulation of complex vegetative functions in body, coordinating movements and mental processes and processing incoming environmental signals from the sense organs. This is basically effected in a binary, all-or-none manner of collecting and forwarding excitatory and inhibitory signals by single neurons over a network of billions of nerve cells.

In this chapter a short overview of the distribution of monoaminergic nerve fibres in the central nervous system (CNS) are given. We also shortly explain processes at the nerve terminals.

### 2.1 Monoamine Systems in Brain

Monoamines are small active molecules, which are acting as neurotransmitters in the brain (Serotonin) and respectively as mediators in different tissues of the body (Histamine). Therefore they are very important for the highly complex physiological conditions. But they are also involved in pathological processes.

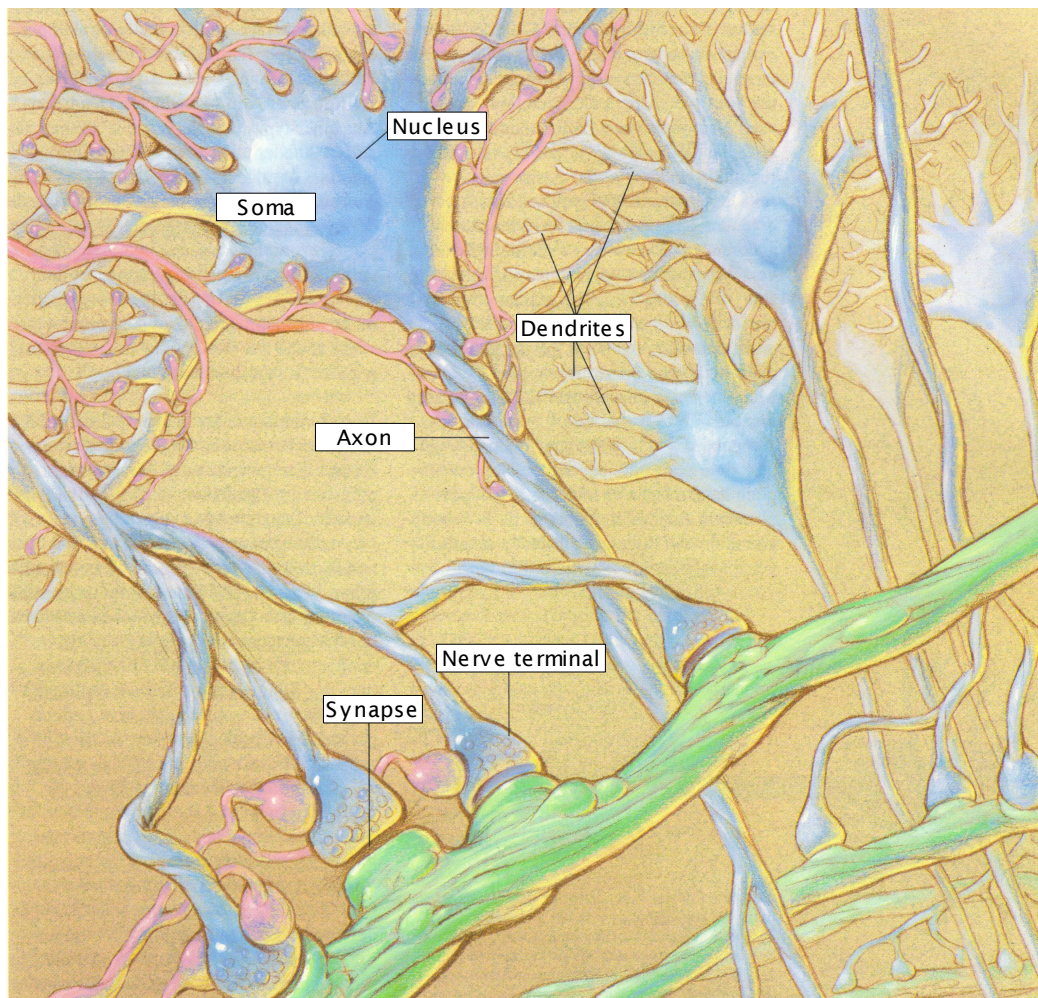


Figure 2.1: Neurones in the CNS (from S.H. Snyder, 1990)



## Serotonin (5HT)

The serotonergic cell bodies are predominantly located in the raphe nuclei in brain stem. They project to almost every region in the brain, to cerebral cortex, hippocampus, thalamus, basal ganglia, cerebellum and amygdala (Howell and Kimmel, 2008). Serotonin occurs in very low concentrations in the brain, but adopts multiple functions due to the high degree of differentiated receptors. The highest concentration of serotonin is found in enterochromaffinic cells in the intestine. In the periphery it acts as mediator of inflammation by sensitising the nociceptors and takes also part in blood coagulation by activating platelets. Serotonin evinces cardiovascular activity. Depending on tissue and receptor subtype it effects vascular constriction or dilatation and positive inotropic effects (Siegel, 1994).

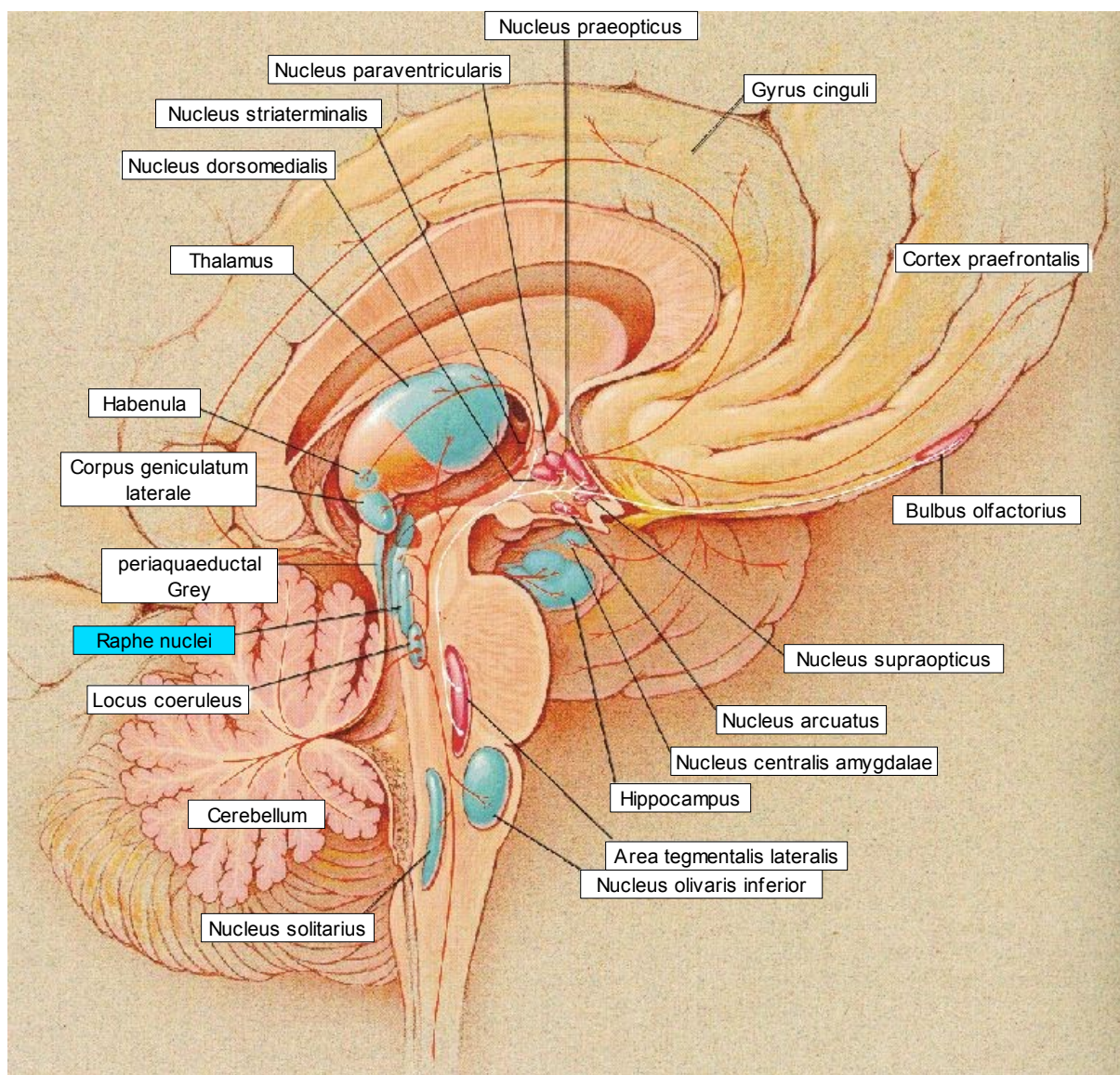


Figure 2.2 Serotonergic cell bodies are tightly located in the Raphe nuclei (blue). Their projections reach nearly each region in the human brain (from S. H. Snyder, 1990)



## Dopamine (DA)

The soma of dopaminergic neurons are located in the centre of the brain stem, called substantia nigra, which is part of the extrapyramidal system, ventral tegmental area and hypothalamus. A strong nerve tract projects to corpus striatum, which is coordinating arbitrary movements. Other projections draw to nucleus accumbens, nucleus caudatus, putamen, to mesolimbic system, prefrontal cortex and amygdala. Paracrine excreted dopamin influences the hypophysis by inhibiting prolactine excretion (Howell and Kimmel, 2008). In PNS, dopamin regulates blood supply of abdominal organs, especially of kidney and is also found in postganglionic sympathetic fibres in common with norepinephrine. It has vasodilatative and sympathomimetic effects in peripheral vessels (Siegel, 1994).

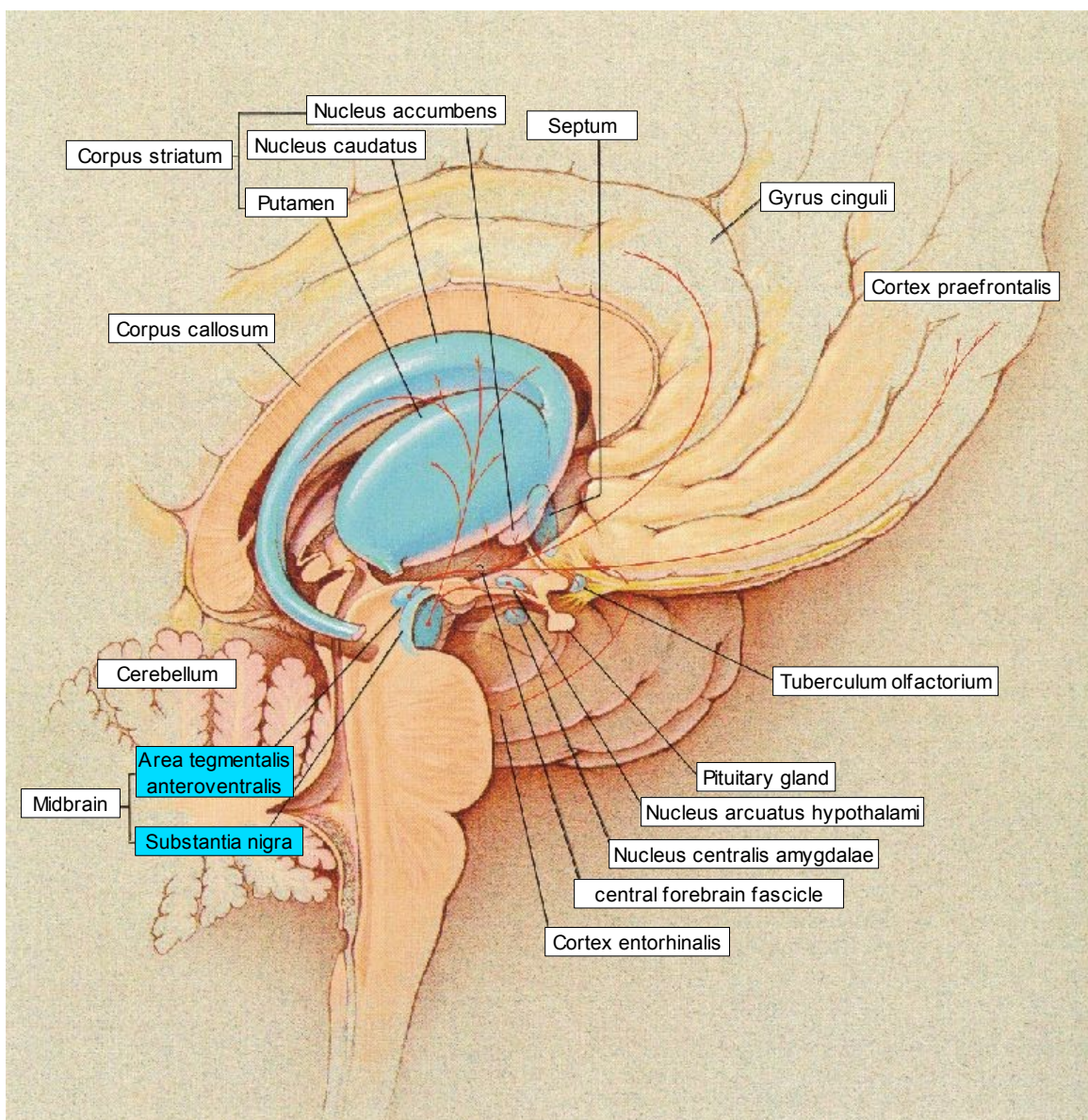


Figure 2.3: Dopaminergic cell bodies are tightly located in Substantia nigra and Ventral tegmental area (blue), where they project to corpus striatum (from S. H. Snyder, 1990)



## Norepinephrine (NE)

The norepinephrinic cell bodies are concentrated in the locus coeruleus and in the area tegmentalis. Similar to the projections of the serotonergic neurons, the axons project to nearly each region in brain, as to the cerebral cortex, ventral tegmental area, hippocampus, thalamus, hypothalamus, nucleus accumbens, limbic system, cerebellum and amygdala. (Howell and Kimmel, 2008). Due to the innervation of the sympathetic by norepinephrine and epinephrine, there are fibres to medulla and spinal cord. In this role, it takes part on multiple roles in peripheral nervous system (PNS), such as cardiovascular effects, influencing heart beat and blood pressure, stress reaction and energy balance. High concentrations are found in adrenal medulla, lung and placenta (Siegel, 1994).

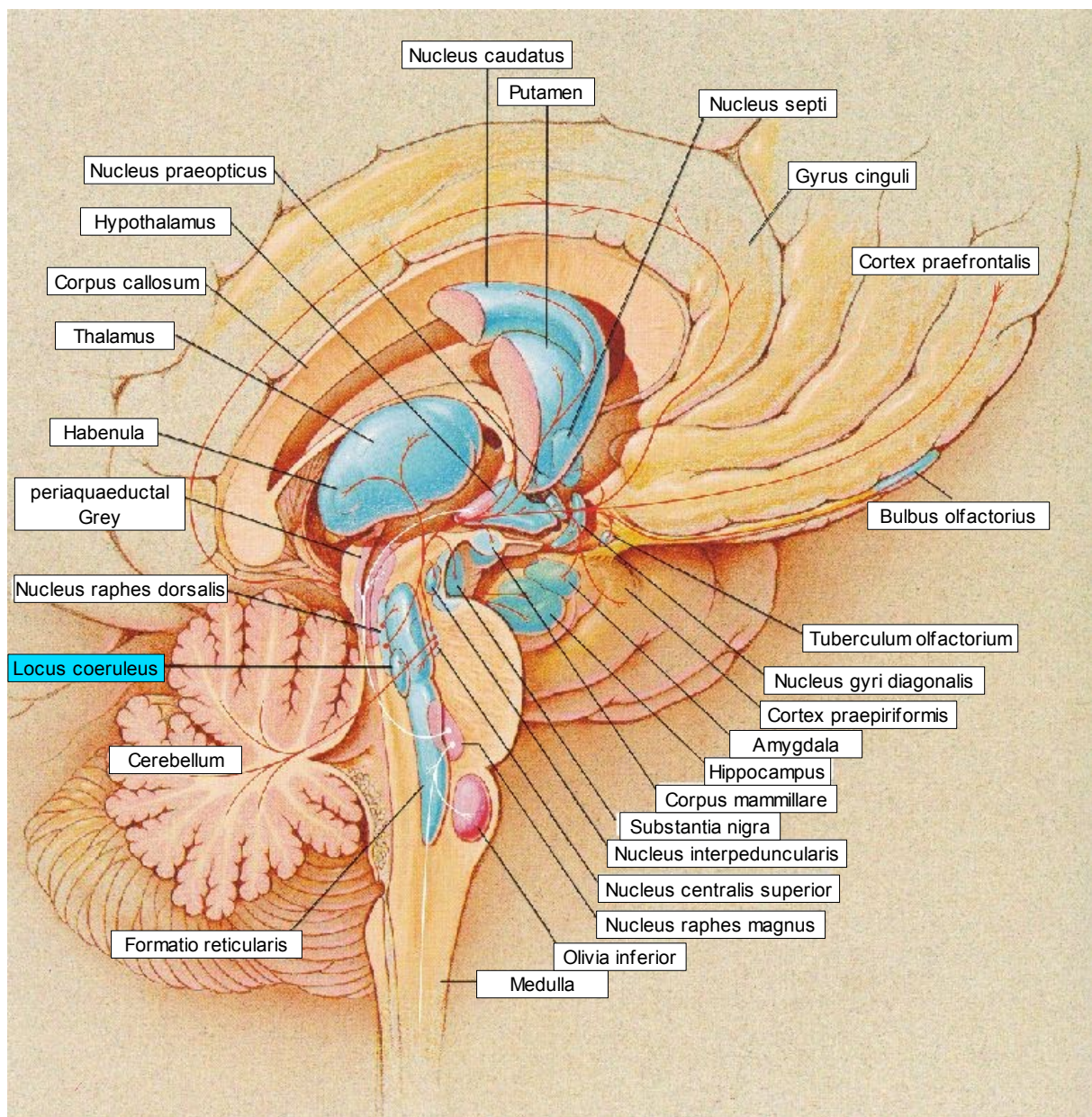


Figure 2.4: Norepinephrinic cell bodies are mainly located in the Locus coeruleus (blue). Their projections also reach nearly each region in the brain (from S. H. Snyder, 1990).



## 2.2 Cell Membrane and Membrane Proteins

The cell membrane is organised as bilayer, built up of amphiphilic phospholipid monomers. The hydrophilic phosphoric acid moieties of the monomers point towards the aqueous environment and the lipophilic tails of the fatty acids are orientated towards themselves. This organisation leads to a separation of the intracellular and extracellular compartment acting as a barrier of diffusion. There are integrated steroids and other lipids, determining the characteristics like flexibility, and proteins, that facilitate the transport through the membrane, which is impermeable for large molecules, charged compounds and ions. Due to the semipermeable property of the cell membrane, it acts as capacitor for the electrochemical potential, built up through the different osmolarities of the solute electrolytes in cytoplasm and liquor. The separation of charges between the extracellular and intracellular space acts as electrochemical force for building up action potentials and facilitating secondary active membrane transport (Siegel, 1994).

Membrane proteins are integrated in the phospholipid bilayer. The common properties, depending on function, are the lipophilic transmembrane domains (TMD), glycosylation and phosphorylations sites G-protein contact regions and ligand recognition sites on the intra- or extracellular domains, i.e. loop regions (IL, EL).

- Ion pumps build up and maintain the gradient of electrolytes by transporting their respective ion against the gradient through the cell membrane by using adenosin triphosphate (ATP). It is the electrochemical driving force for action potentials and secondary active transport. They are also involved in creating and forwarding action potentials.
- Receptors proteins are the lock for the chemical key of forwarding the electrochemical impulse through the synaptic cleft. Ligand binding initiates G-proteins and accordingly enzyme activation.
- Ion channels are divided in ligand gated and voltage gated. Changes in membrane potential due to an action potential lead to an opened or closed channel. Accordingly ligands induce after binding also an open channel.
- Transporter proteins facilitate the passing of solutes, like glucose and neurotransmitter, through the cell membrane. This form of transportation is dependent on energy and called secondary active transport.

(Siegel, 1994)

Transport through the cell membrane for impermeable substances is mediated by

transport proteins. Huge volumes are transferred by endocytosis (phagocytosis) and accordingly exocytosis (liberation of neurotransmitters).

- Active transport is necessary, when solutes are not able to pass the membrane by diffusion due to their volume, charge or an existing gradient. Primary active transport mechanism is coupled to ATP usage. Secondary active transport uses the prior built up gradient of electrolytes, where the substrates are coupled with.
- Passive transport does not need energetic effort. Free and facilitated diffusion are the two modalities. The simplest way crossing the membrane is diffusion, done by small lipophilic molecules. Membrane impermeable substances use carrier proteins, called facilitated diffusion. Transport mechanisms of molecules occur as uniport, symport and/or antiport.

(Fain, 1999; Siegel, 1994)

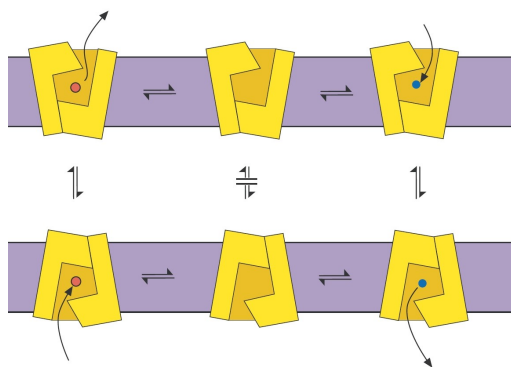


Figure 2.5a: Antiport

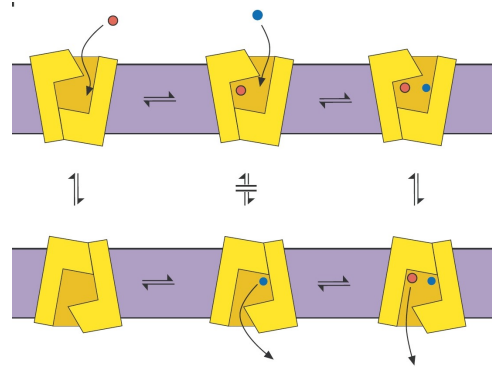


Figure 2.5b: Symport

from Rudnick, 2007

## 2.3 Neurotransmitters and Receptors

Neurotransmitters are the chemical factors that forward signals from one neuron to the next at the synaptic cleft. Synthesised in the cell corpus, they are packed to vesicles and transported along the neurotubuli to the nerve terminals. Release is effected by  $\text{Ca}^{2+}$  mediated exocytosis. Chemically they are derivatives of amino acids, purines and peptides. Functionally there are grouped to excitatory, inhibitory and modulating transmitters. Each transmitter has its particular receptor. In the 1920s Otto Loewi revealed, that the forwarding of the electrical impulse between two neurons is done in a chemical manner. Later, he discovered the first neurotransmitter, acetylcholine (Siegel, 1994).

Neurotransmitters can be divided due to their functional properties:

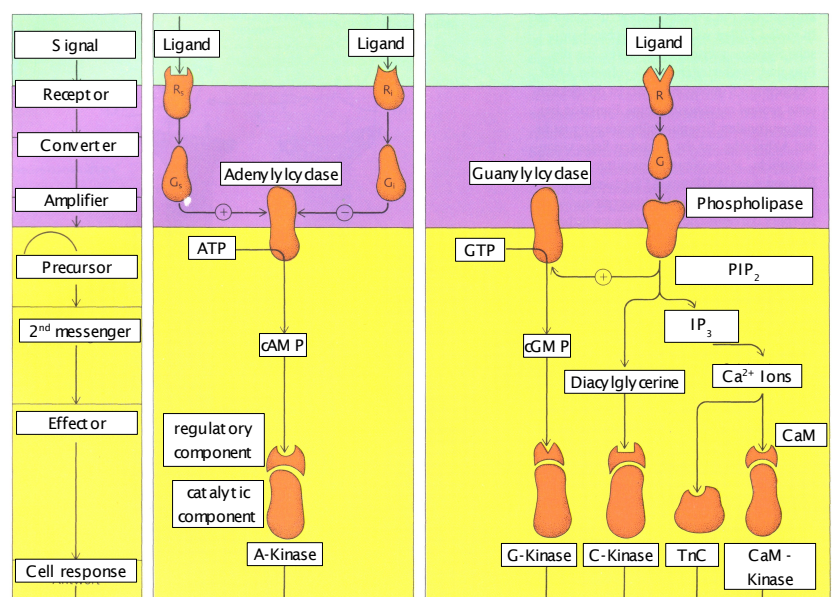
- $\gamma$ -amino butyric acid (GABA) and glycine are the main inhibitory transmitter, by opening  $\text{Cl}^-$  channels.
- Glutamate and aspartate are excitatory transmitters, acting at the NMDA, AMPA and kainat receptors
- Monoamines, i.e. catecholamines (dopamin, norepinephrine, epinephrine), Serotonin and Histamine are modulating transmitters and involved in multiple processes of the personality and regulation of peripheral systems
- Acetylcholine (ACh) coordinating the transmission on motor end plate, as well as in the central nervous system (CNS), where it acts in the cholinergic system, important for synaptic plasticity, learning and reward.
- Endorphins, neuropeptides
- Adenosine, ATP and phosphoinositides

The receptors may appear in subtypes and are divided in three groups due to their biochemical property:

- The majority of the receptors are G-protein coupled. This G-protein is part of the cell signal transduction and activates enzymes, e.g. adenylylcyclase, phospholipase C, in the cytoplasm for creating 2<sup>nd</sup> messengers, like cAMP and  $\text{IP}_3$  to forward the signals.
- Ligand gated ion channels are another type of receptors. These proteins initiate and terminate action potentials by opening or closing the selective pore of their proper electrolyte. Another type of ion channel is the voltage gated ion channel, which is innervated in changes of the membrane potential.
- Enzyme coupled receptors initiate enzymatic activity after ligand binding, e.g. tyrosine kinase.

(Löffler, 1999)

Figure 2.6: Cascade of signaltransduction (S. H. Snyder, 1990)



Transmitter	Receptor	Type	Ion	Amplifier	second messenger
Adenosine	A1-A3	G-Protein coupled			
Epinephrine and Norepinephrine	$\alpha$ 1a-1d	G-Protein coupled		+ phospholipase	DAG/IP <sub>3</sub>
	$\alpha$ 2a-2c	G-Protein coupled		- phospholipase	DAG/IP <sub>3</sub>
	$\beta$ 1-3	G-Protein coupled		adenylyl cyclase	cAMP
Acetylcholine	Nicotinic N <sub>musc</sub> , N <sub>neur</sub>	Ion channel	Na <sup>+</sup>		
	Muscarinic M1-M5	G-Protein coupled			
Dopamine	D1, D5	G-Protein coupled		adenylyl cyclase	
	D2-D4	G-Protein coupled		adenylyl cyclase	
$\gamma$ -amino butyric acid	A	Ion channel	Cl <sup>-</sup>		
	B	G-Protein coupled		ion channels	
Glutamate and Aspartate	NMDA	Ion channel	Na <sup>+</sup>		
	AMPA	Ion channel	Na <sup>+</sup>		
	Kainat	Ion channel	Na <sup>+</sup>		
	mGlu1-mGlu7	G-Protein coupled			
Glycine	GlyRA	Ion channel	Cl <sup>-</sup>		
	GlyRB	Ion channel	Cl <sup>-</sup>		
Serotonin	5HT <sub>1</sub>	G-Protein coupled		- adenylyl cyclase	cAMP
	5HT <sub>2</sub>	G-Protein coupled		+ phospholipase	DAG/IP3
	5HT <sub>3</sub>	Ion channel	Na <sup>+</sup>		
	5HT <sub>4</sub>	G-Protein coupled		+ adenylyl cyclase	cAMP
	5HT <sub>5</sub> - 5HT <sub>7</sub>	G-Protein coupled			
ATP-Receptors	P2X	Ion channel			
	P2Y	G-Protein coupled			
Histamine	H1	G-Protein coupled		+ phospholipase	DAG/IP3
	H2	G-Protein coupled		+ adenylyl cyclase	cAMP
	H3	G-Protein coupled		autoreceptor	
Endorphines and Opiates	$\mu$ 1-2	G-Protein coupled		- adenylyl cyclase	
	$\kappa$ 1-3	G-Protein coupled		- adenylyl cyclase	
	$\delta$ 1-2	G-Protein coupled		- adenylyl cyclase	

Table 2.1 shows the neurotransmitter and their receptors, data not completed; + = stimulation; - = inhibition (Thews, Mutschler, Vaupel, 1999; Siegel 1994).

## 2.4 Synapse and Neuronal Transmission

Synapses are junctions between the terminal of one neuron with another nerve cell for discontinuous forwarding of excitation. There are different types of synapses described:

- Electrical synapses or gap-junctions are channel like connections of vicinal cells for a rapid exchange of information and small substrates, e.g. amino acids, nucleotides, glucose and ions, without innervation of a ligand-receptor-binding.
- Chemical synapses connect either two neurons or neurons with glands or neurons with muscles on the motor end plate.

(Fain, 1999)

The commonly most known synapse is the chemical synapse, where one nerve terminal gets in contact with another nerve cell membrane. On the membrane surface different types of membrane proteins, e.g. receptors, ion channels, ion pumps, transporter are located, both on the presynaptic and the postsynaptic membrane.

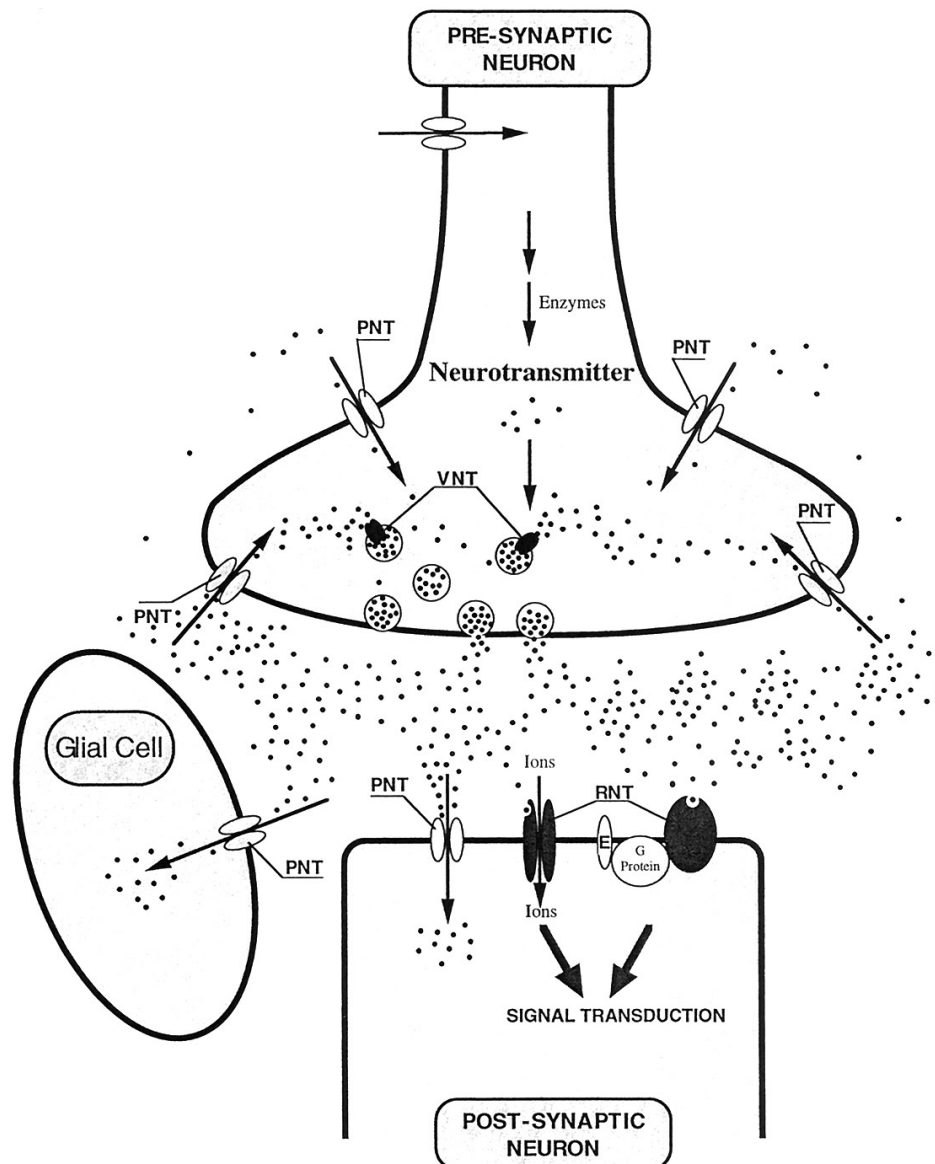


Figure 2.7: Neuronal signal transmission in the synapse during firing:  
PNT = Plasmamembrane Neurotransmitter Transporter  
VNT = Vesicular Neurotransmitter Transporter  
RNT = Receptors of NeuroTransmitters  
E = Effector protein  
(Masson et al., 1999)

Due to the effect on the downstream neuron, synapses are distinguished in excitatory and inhibitory. In excitatory neurons an action potential releases neurotransmitters, that induce a depolarisation in the following neuron by opening ligand gated  $\text{Na}^+$  channels. Inhibitory neurons act similarly, with the difference of the involved ion. The released transmitters open ligand gated  $\text{Cl}^-$  channels that leads to hyperpolarisation. Neurons are permanently adding and subtracting currents of excitatory or inhibitory signals from their dendrites. After exceeding the threshold at the axon-hillock an action potential is generated. This is a temporary change in membrane potential of excitable cells as answer of an incoming stimulus. It is according to the all-or-none law, i.e. only after reaching and exceeding the stimulus threshold  $\text{Na}$  and  $\text{K}$  channels open and the depolarisation occurs completely. The action potential is forwarded over the axon or over electrical synapses and the affected membrane segments are turning into the refractory phase. After this step repolarisation to the resting potential is done and the cell is ready for the next depolarisation (Mutschler, 2001; Fain, 1999; Siegel, 1994).

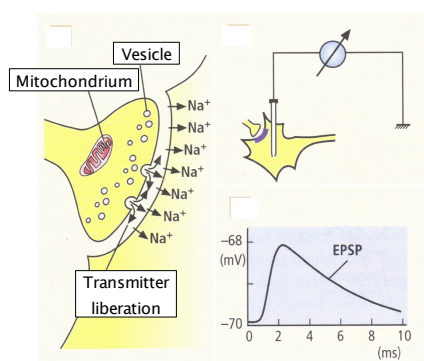


Figure 2.8a: EPSP coupled with  $\text{Na}^+$  inward and  $\text{K}^+$  inward currents

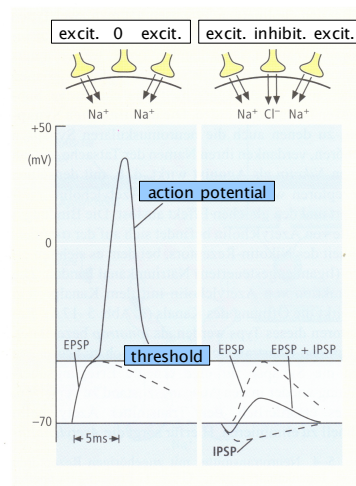


Figure 2.8b shows the reaching of an action potential.

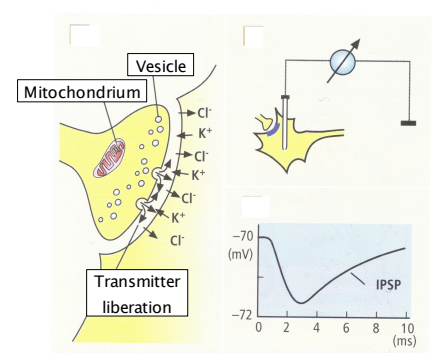


Figure 2.8c: IPSP coupled with  $\text{Cl}^-$  inward and  $\text{K}^+$  outward currents

1. -70mV resting potential
2. opening postsynaptic  $\text{Na}^+/\text{Cl}^-$  and  $\text{Ca}^{2+}$  channels, ligand gated
3. -50mV; threshold, voltage gated  $\text{Na}^+$  channels open, depolarisation,
4. +50mV overshoot, voltage gated  $\text{K}^+$  channels open, repolarisation,  $\text{Na}^+$  channels close
5. -90mV hyperpolarisation by  $\text{Na}/\text{K}$ -ATPase, open  $\text{K}^+$  and  $\text{Ca}^{2+}$  channels - refractory phase
6. returning to resting potential for the next cycle

(Mutschler , 2001)



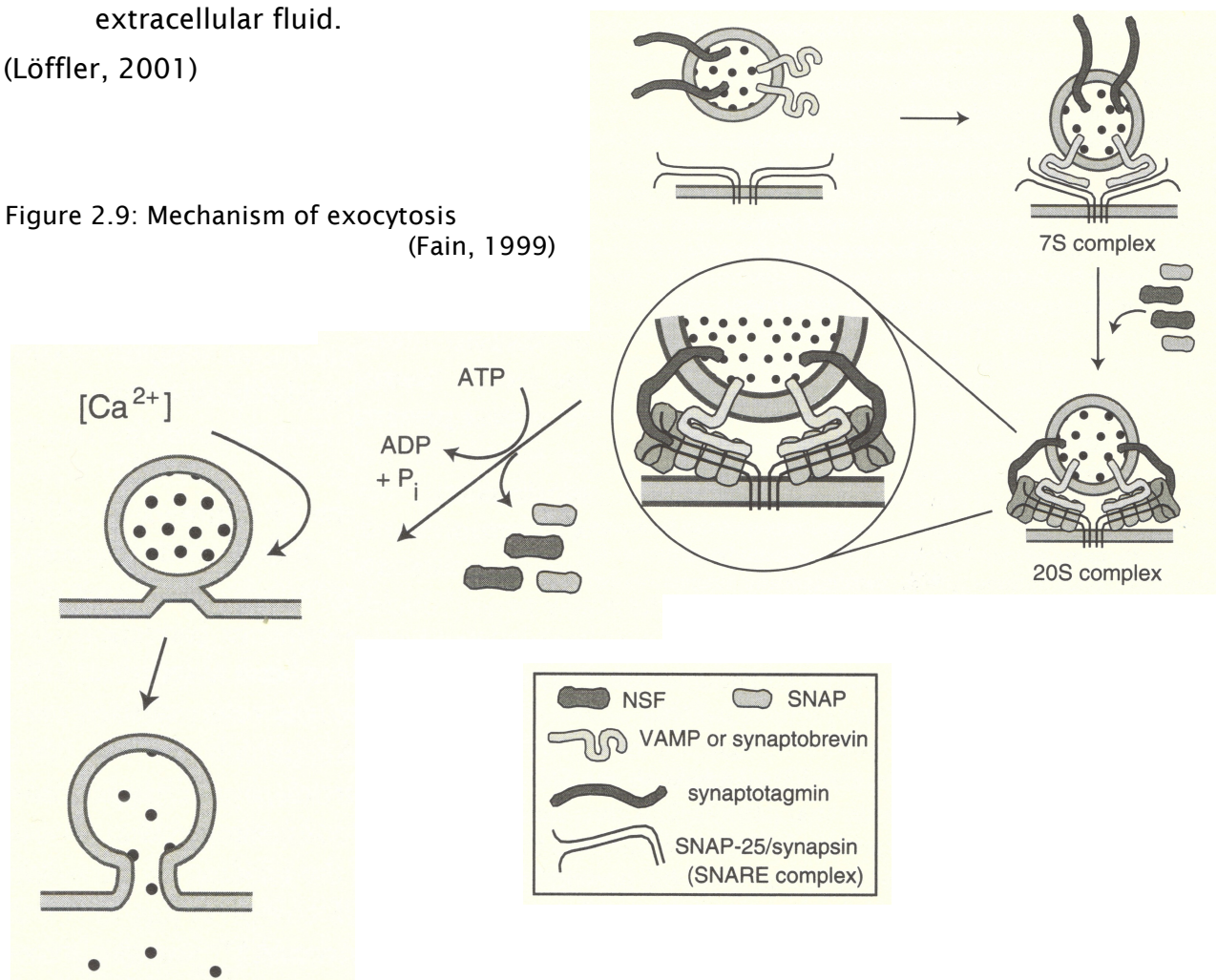
At the synaptic cleft, the electrochemical impulse formed by an action potential is converted to chemical information. In the presynaptic nerve terminal neurotransmitters are stored in vesicles. After an incoming action potential  $\text{Ca}^{2+}$  channels open and initiate the phosphorylation of synapsin, the anchoring protein of vesicles with cytoskeleton. This phosphorylation abolishes the connection of the vesicle with the cytoskeleton and is now able to move to the cell membrane. Vesicles carry the protein synaptotagmin that interacts with syntaxin located on the cell membrane. This interaction induces the fusion of vesicle membrane and axon membrane and further excretion of the vesicle contents. (Siegel, 1994)

After fulfilled transmission, the neurotransmitter must be removed from the synaptic cleft, in order to avoid further activation. The termination of signalling is done by the following ways:

- enzymatic, e.g. ACh-esterase decomposes the transmitter in the synaptic cleft for fast removal
- reuptake, e.g. 5HT is removed by a reuptake into the presynaptic neuron by a specific transporter
- thermodynamic, where the transmitters diffuse from the synaptic cleft in the extracellular fluid.

(Löffler, 2001)

Figure 2.9: Mechanism of exocytosis  
(Fain, 1999)



### 3 Pharmacology

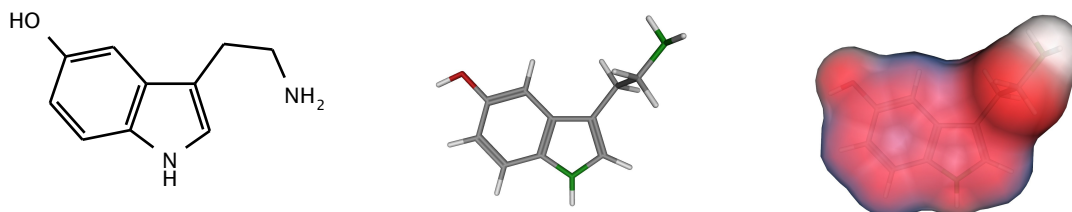
In this chapter, the pharmacological properties of selected ligands interacting with monoamine transporters will be described. Due to their manner of interaction they can globally be divided in two groups: in substrates and inhibitors. The resulting effect is a direct elevation of the transmitter concentration in the synaptic cleft by initiating reversal transport activity and accordingly the indirect increase of transmitter by inhibiting the reuptake activity of the transporter. The NSS are target of a variety of therapeutics, like antidepressants, and illicit drugs, such as psychostimulants.

Monoamines		
5-HT		substrate of SERT
DA		substrate of DAT and NET
NE		substrate of DAT and NET
Amphetamines		
MPD		inhibitor of DAT and NET
MDMA		substrate type releaser of SERT, DAT and NET
d-Amphetamine		substrate type releaser of DAT and NET
Tropane alkaloids		
Cocaine		inhibitor of DAT, NET and SERT
CFT, CIT		inhibitor of DAT, NET and SERT
Benztropine		inhibitor of DAT, NET and SERT
Indole alkaloids		
Psilocin		substrate of SERT
DMT		substrate of SERT
DIPT		inhibitor of SERT
Ibogaine		inhibitor of SERT
Toxines		
MPP+		substrate type releaser of DAT, NET and SERT
Tyramine		substrate of DAT and NET
Antidepressants		
Desipramine		inhibitor of NET
Fluoxetine		inhibitor of SERT
Citalopram		inhibitor of SERT

Table 3.1: Ligands of neurotransmitter:sodium symporter: divided into substrate and inhibitors (Rothman and Bauman, 2003)



### 3.1 Serotonin / 3-(2-aminoethyl)-1H-indole-5-ol / 5-hydroxy tryptamine



Figures 3.1 a - c: Serotonin depicted in (a) 2D, (b) 3D and (c) with surface:  
blue = hydrophilic, red = lipophilic, white = H-bond region  
(The following figures are depicted in a similar modality.)

5-HT is a biogenic amine with an indole base scaffold and a diethyl amino side chain. In the 1930s serotonin was first discovered by Vittorio Erspamer from rabbit gastric mucosa. He called the substance enteramine. Ten years later, Maurice Rapport, in the group of Irving Page, first isolated the vasoconstricting factor in blood serum and determined the chemical structure. In the early 1950s it was recognised, that both substances, enteramine and serotonin, are identical. (Whitaker-Azmitia et al., 1999). During the last 70 years serotonin was found to be involved in a wide range of physiological and pathophysiological processes. The physiological effects of serotonin in CNS effect mood, sleep, appetite, sleep, anxiety, fear, reward and aggression (Howell and Kimmel, 2008). In PNS serotonin acts in blood coagulation, intestinal motility, pain and vascular activity. Serotonin is also implicated in different pathological processes, like migraine, intestinal disorders, depression, anxiety and also Alzheimer disease (Thews, Mutschler, Vaupel, 1999).

The biosynthesis of 5HT is effected by two enzymatic steps from the essential amino acid tryptophane. In the first step the aromatic indole ring is hydroxylated in position 5 and in the second step the decarboxylation of the C $\alpha$  atom takes place. After usage, the decomposition of 5HT is effected by the enzymes MAO and aldehyde dehydrogenase to 5-MeO-indole acetate (Jonnakuty and Gagnoli, 2008). Serotonin is also an intermediate in the biosynthesis of another hormone of the indol type, melatonin (Hardeland et al., 2008).

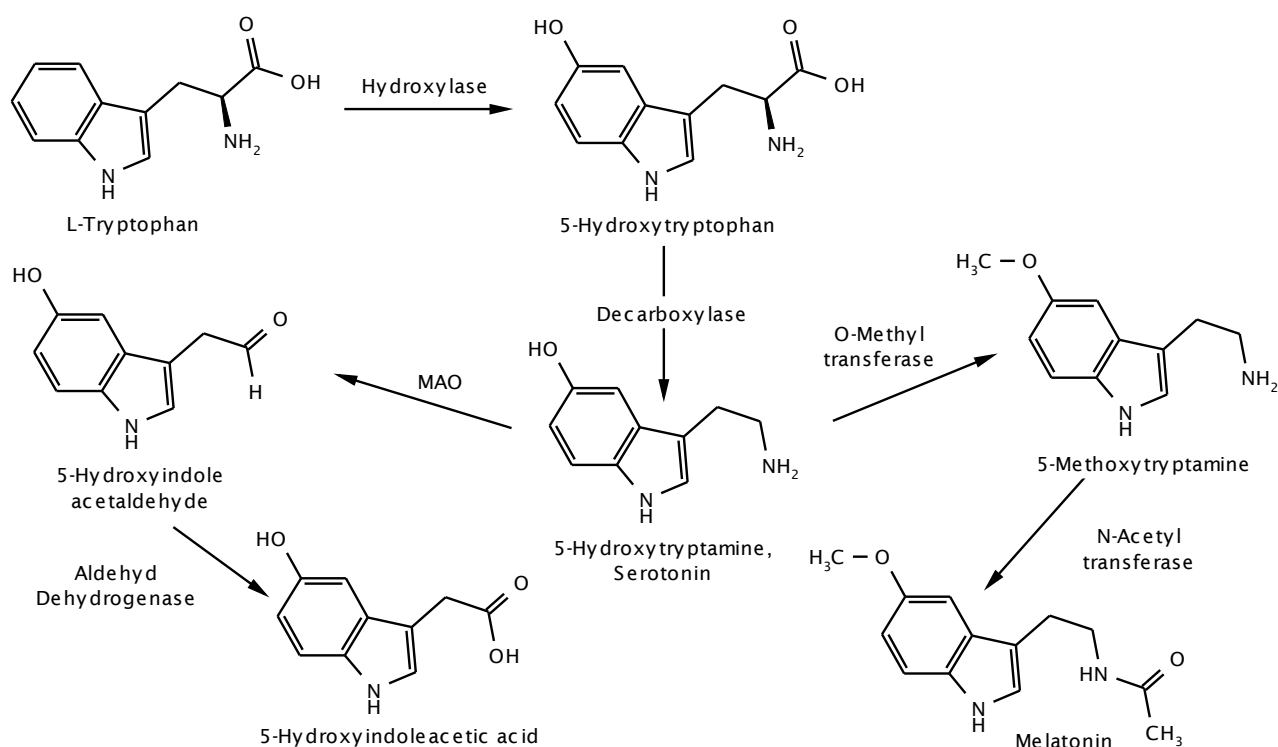
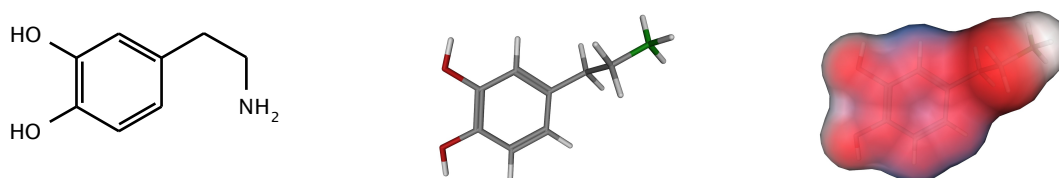


Figure 3.2: Biosynthesis and metabolism of Serotonin

### 3.2 Dopamine / 4-(2-aminoethyl)benzene-1,2-diol / 4-(2-aminoethyl)catechol



Figures 3.3: Dopamine

Dopamine is a member of the catecholamines, as epinephrine and norepinephrine are, consisting of dihydroxylated aromatic base scaffold and a hydroxylated diethylamine sidechain. Arvid Carlsson discovered dopamine as main transmitter involved in movements and Parkinson's disease in the 1950s. Similar to the findings on serotonin, research on this neurotransmitter also elucidated a number of functions in human physiology, such as agonist in peripheral tissues, e.g. vessels in stomach, intestine and kidney, and as well as in CNS. Dopamine is also involved in multiple pathological processes. An important pathological process is the Parkinson's disease, where the dopaminergic neurons in the corpus striatum become poor of the transmitter. It also plays a role in schizophrenia, that is ascribed to an enhanced activity of the dopamine system in brain (Iversen and Iversen, 2006). Due to the role in reward, dopamine plays an important role in drug addiction. The dopamine neurons in the mesolimbic system, which project to nucleus

accumbens and ventral tegmental area are involved (Nestler, from Madras 2006).

Biosynthesis starts at the precursor L-tyrosin, an essential amino acid, that is enzymatically converted to levodopa by hydroxylation of the aromatic ring. After decarboxylation dopamine is an intermediate of the further biosynthesis of norepinephrine and epinephrine. Epinephrine is found in small amounts in CNS. During the degradation two enzyme complexes are involved for hydroxylation and oxidation steps, notably MAO and COMT, converting dopamine to vanillylmandelic, as depicted below (Mutschler, 2001; Siegel, 1994).

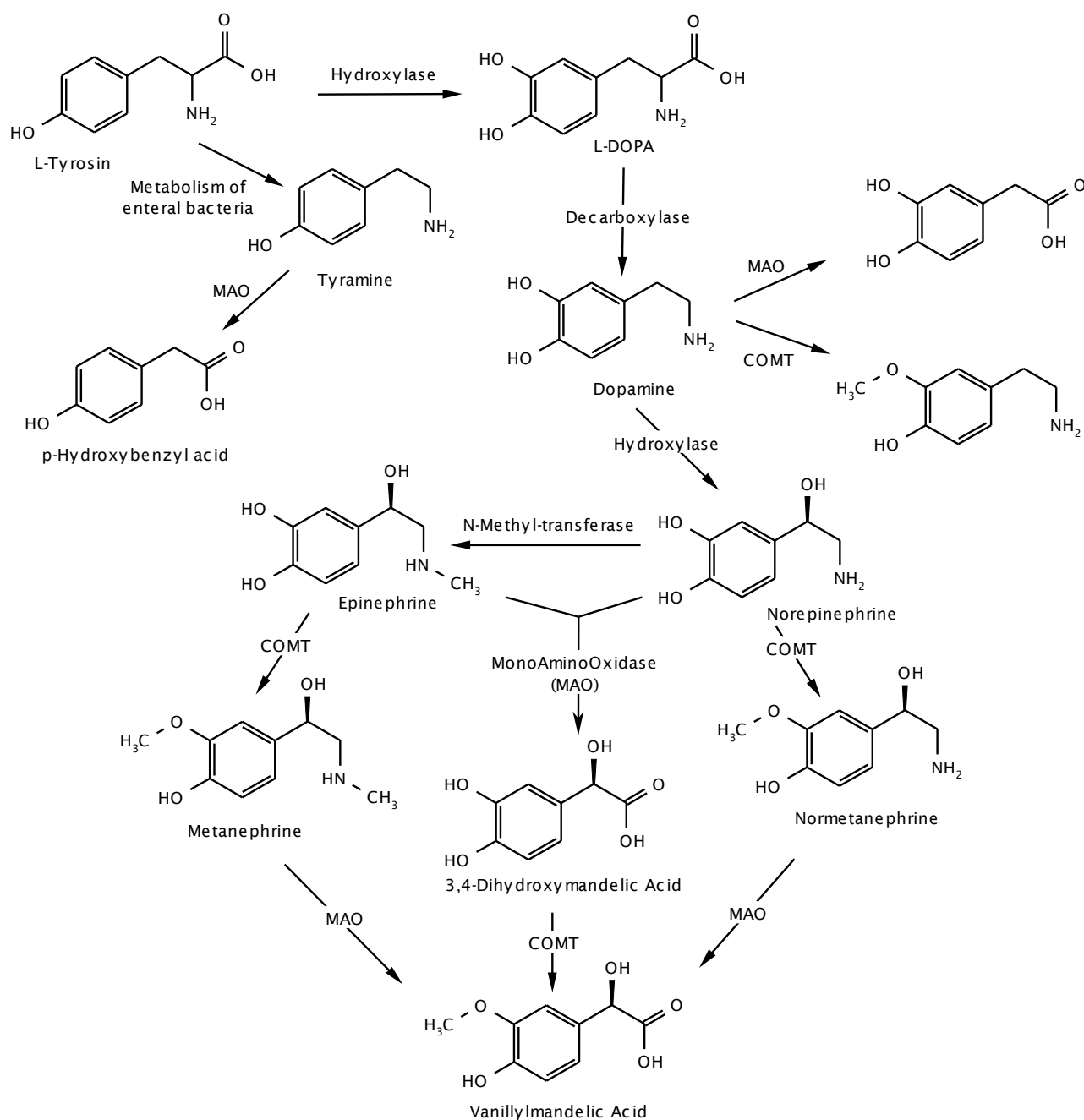


Figure 3.4: Biosynthesis and metabolism of catecholamines

### 3.3 Norepinephrine / (R)-4-(2-amino-1-hydroxyethyl)-1,2-dihydroxybenzene / noradrenaline

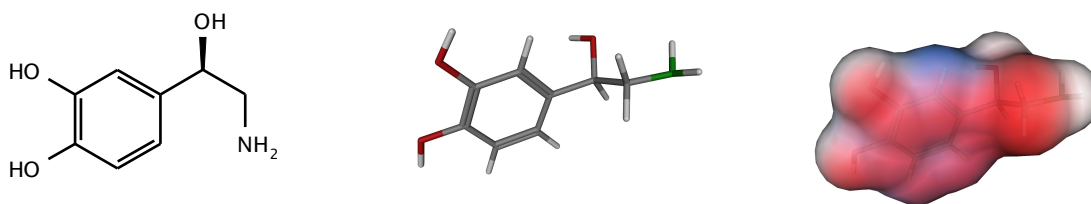


Figure 3.5: R-Norepinephrine

Norepinephrine differs from dopamine in the hydroxylated sidechain, that builds a centre of chirality. So there exist two enantiomeres, the biologically active R-enantiomere and the isoform. In the 1940s Ulf von Euler discovered norepinephrine as neurotransmitter. Ten years later Julius Axelrod elucidated the mechanism of neurotransmitter reuptake on adrenergic and noradrenergic nerves, and also discovered COMT responsible for catecholamine inactivation, see figure 3.4 (Goldstein et al., 2005; Torres et al., 2003).

This neurotransmitter is responsible for multiple effects in human body in the periphery, e.g. innervation of sympathetic fibres, blood pressure, heart frequency, and in CNS, e.g. attention, learning, sleep-wake-rhythm, mood and locomotion. Due to the ubiquitous appearance of this mediator and neurotransmitter, different diseases are related to the norepinephrinic system. It is supposed that norepinephrine is involved in depression, attention deficit hyperactivity disorder (ADHD), post traumatic stress disorder (PTSD), stress, parkinsons disease, schizophrenia and autonomic disorder (Mutschler, 2001). Therapeutically it is used at anaphylactic shock, and also as adjuvant in local anaesthetic preparations (Pschyrembel, 2002).

### 3.4 Amphetamine / (RS)-1-phenylpropane-2-amine / "Speed"

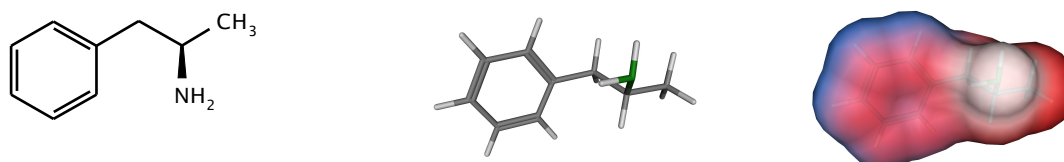


Figure 3.6: R-Amphetamine

Amphetamine is built up of a benzene ring and an amino propion residue, which contains a chirality centre in C $\beta$  that lead to two stereoisomeres. It is a synthetic drug, first synthesized in 1887, and was not found in nature. Due to the sympathomimetic effects, it was used as treatment of common cold, hay fever or asthma after its introduction in the

1930s. In 1940s, and later amphetamine was applied for different diseases, e.g. depression, narcolepsy, anti-obesity agent, until the addictive potential as well as misuse was observed in 1960s. In further consequence it was forbidden by law, and today it is a widely abused drug, as also its derivative methamphetamine. Abuse of these substances shows serious consequences on health and society, because of the addictive potential and damage on CNS. Only few preparations are on the market for treatment of ADHD and narcolepsy (Iversen, 2006).

The name amphetamine is derived from alpha-methylphenylethylamin and is the lead substance and eponymous for a large group of compounds with similar structure and effects, called amphetamines (Shulgin and Shulgin, 1995). Due to the similar structure to dopamine and norepinephrine, the catecholaminergic systems in CNS are primarily influenced by this drug. R-amphetamine is biologically more active than the L enantiomer. It acts as indirect sympathomimetic by stimulating release of norepinephrine at nerve terminals and also elevates the concentration of dopamine in the synaptic cleft of dopaminergic neurons (Howell and Kimmel, 2008). Amphetamine is a substrate-type releaser of neurotransmitters, i.e. the mechanism of increasing catecholamine concentration is due to the inversion of transport direction of the vesicular monoamine transporters and also of DAT and NET. The affinity for SERT is about 100 fold lower (Rothmann and Baumann, 2003).

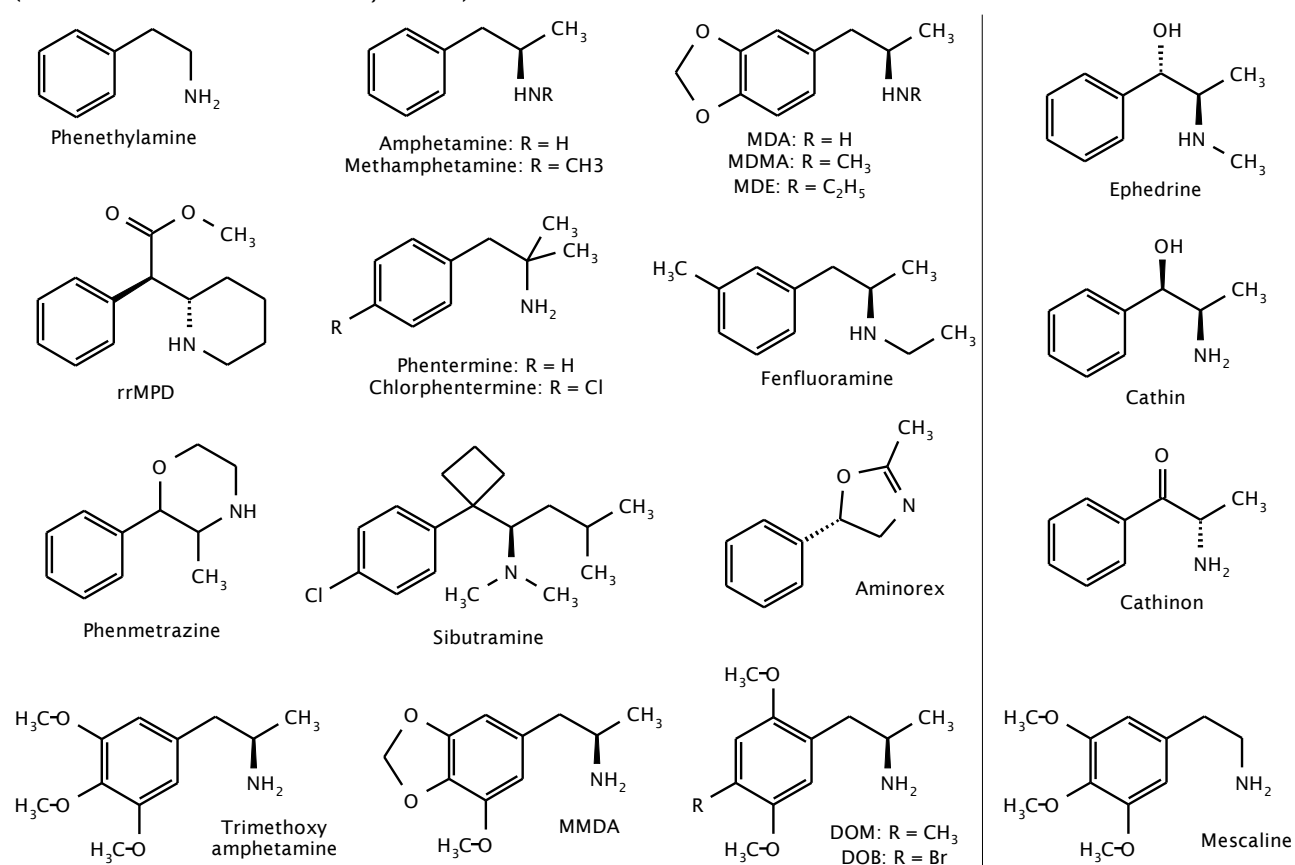


Table 3.2: Synthetic amphetamines and amphetamine like compounds

biogenic analogues

### 3.5 MDMA / 3,4-methylenedioxy-N-methylamphetamine / "Ecstasy"

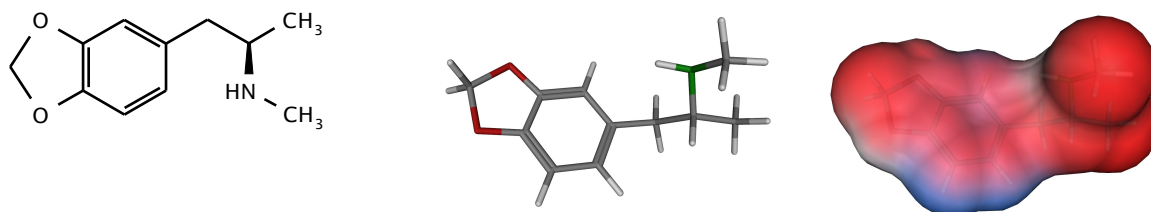


Figure 3.7: R-MDMA

MDMA was first synthesised in the 1910s as intermediate in synthesis of haemostatic compounds. It was first described as psychotropic drug in the 1960s by Alexander Shulgin. Ten years later it was used in common with MMDA by psychoterapists in clinical practice until the 1980s. In the late 1960s MDMA became popular under students and the consumption increased. Due to the increasing occurence abuse of this compound it was restricted by law in the mid 1980s (Iversen, 2006).

In CNS, MDMA acts as substrate for the NSS, similar to amphetamine. It induces the inversion of transport direction of both the membraneous and the vesicular monoamine transporters. In contrast to D-amphetamine MDMA targets the SERT with higher affinity than DAT and NET (Rothman and Bauman, 2003).

### 3.6 Methylphenidate / $\alpha$ -phenyl-2-piperidinacetate methylester / Ritalin

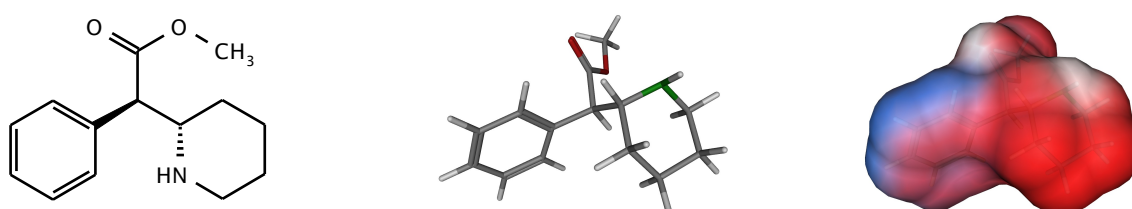


Figure 3.8: R,R-Methylphenidate

Methylphenidate was first synthesized in the 1940s. The basic scaffold is similar to the amphetamines, but on the C $\alpha$  atom there is connected a methylester group. This leads to two optical centres, which four enantiomers. The R,R-D-threo form is the more active isoform and used for therapeutic purposes (Meltzer et al., 2002). The pharmacological effect differs from other amphetamines. In contrast to amphetamine and MDMA, which are substrate type releaser, MPD blocks DAT and NET with similar affinity (Iversen, 2006). On SERT the affinity is more than 200 fold lower (Rothman and Baumann, 2003). The pharmacological activity is similar to that of cocaine, but MPD does not lead to euphoria

and dependency, because of the slower onset (Swanson and Volkow, 2003).

The main area of application of MPD is the treatment of attention deficit hyperactivity disorder (ADHD). This disease has an incidence of 3 to 5 % with a multifactorial etiology. It is diagnosed in the first decade of life mainly on boys. Often their parents are also affected. ADHD is a disorder in processing information and attention. The concerned brain regions are the frontal lobe, striatum and cerebellum where the projections of norepinephrinic neurons are ending (Iversen, 2006; Biderman et al., 1999).

### 3.7 Cocaine / 3 $\beta$ -Hydroxy-1 $\alpha$ H,5 $\alpha$ H-tropane-2 $\beta$ -carboxylic acid-methylester-benzoate

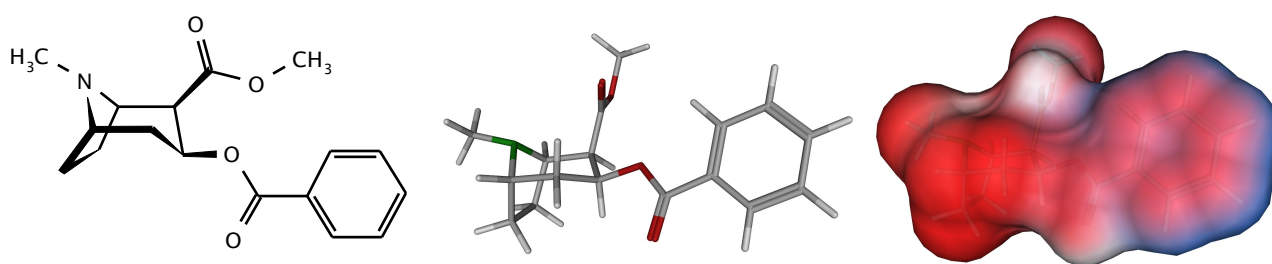


Figure 3.9: Cocaine

Cocaine is the psychostimulant factor in the South American plant *erythroxylum coca*. The leaves are used traditionally to increase work capacity and treat anoxia. The alkaloid was isolated in the 1860s by Albert Neiman. In Europe cocaine was added into beverages until the addictive potential became apparent. Since the 1970s cocaine became a widely abused drug with serious consequences of the consuming individual due to the high addictive potential (Madras and Lin, from Madras, 2006). In the beginning of the 20<sup>th</sup> century, it was first also applied as local anaestheticum, until more specific derivatives have been synthesised (Pschyrembel).

The neurochemical effects are the inhibition of the NSS and for that an indirect increase of serotonin, dopamine and norepinephrine in the synaptic cleft. It inhibits in ascending affinity SERT < DAT < NET (Rothman and Baumann, 2003). Cocaine is also a direct agonist on monoamine receptors, and inhibitor of Na<sup>+</sup> channels in CNS.

### 3.8 CFT / (-)-2-β-Carbomethoxy-3-β-(4-fluorophenyl)tropane (β-CFT, WIN 35,428)

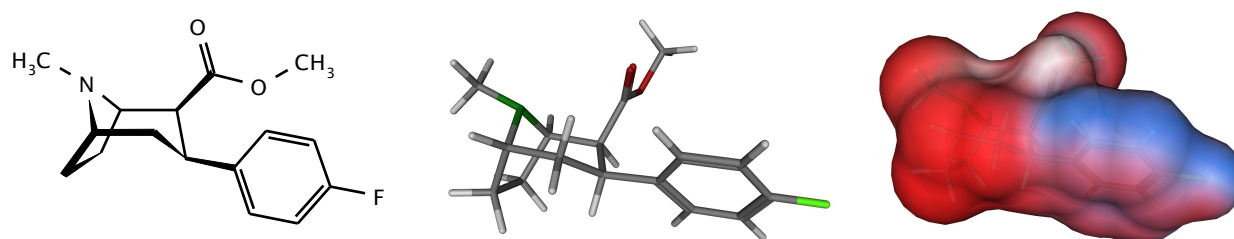


Figure 3.10: CFT

A number of derivatives have been synthesised in the research on DAT. CFT is an analogue of cocaine and also a stimulant drug. It is primarily used for pharmacological experiments and radiolabeling studies due to the chemical stability (Kelkar et al., 1994; Kline et al., 1994). The neurochemical effect is similar to that of cocaine, by blocking the NSS in ascending order of affinity DAT < NET < SERT (Torres et al., 2003).

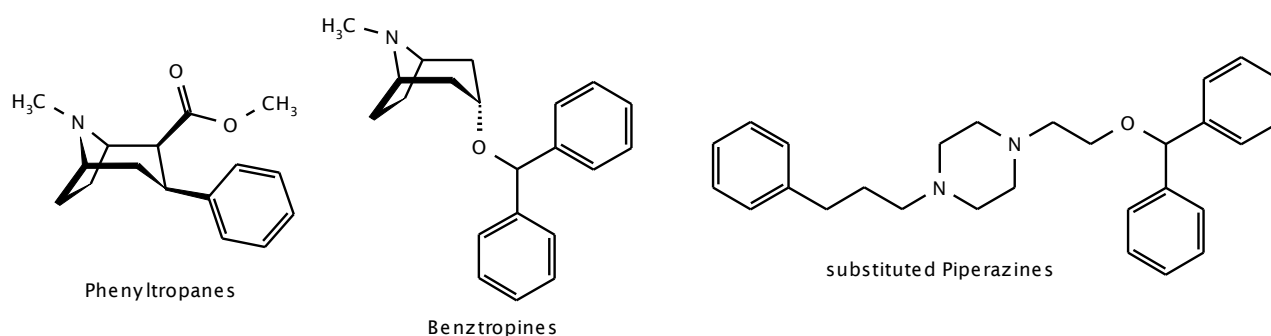


Figure 3.11 shows lead structures of cocaine analogues and DAT inhibitors (Rothman et al., 2008).

### 3.9 MPP<sup>+</sup> / (1-methyl-4-phenylpyridinium)

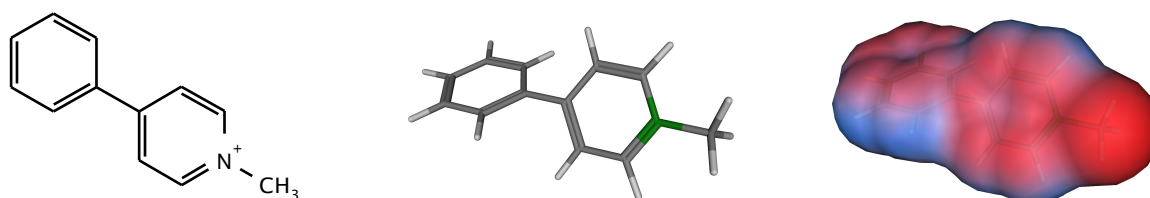


Figure 3.12: MPP<sup>+</sup>

The precursor MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) is converted by MAO B into its active, neurotoxic metabolite MPP<sup>+</sup> (Javitch, 1985). MPTP is also an intermediate of the MPPP (1-methyl-4-phenyl-4-propionoxypiperidin) synthesis, which is used as a synthetic opiate. MPP<sup>+</sup> is transported by the NSS and causes a depletion of dopaminergic neurons in corpus striatum. By interfering with the neuronal mitochondria, the energy



production is disrupted. In further consequence the dopaminergic neurons die off and the Parkinsons disease is induced (Storch et al., 2004). Derivatives of pyridines are also used as "legal ecstasy". These compounds act similar to MDMA. Corrupted synthesis may lead to a precursor of MPP+ and exhibits a health risk of the illicit usage (Baumann et al., 2005).

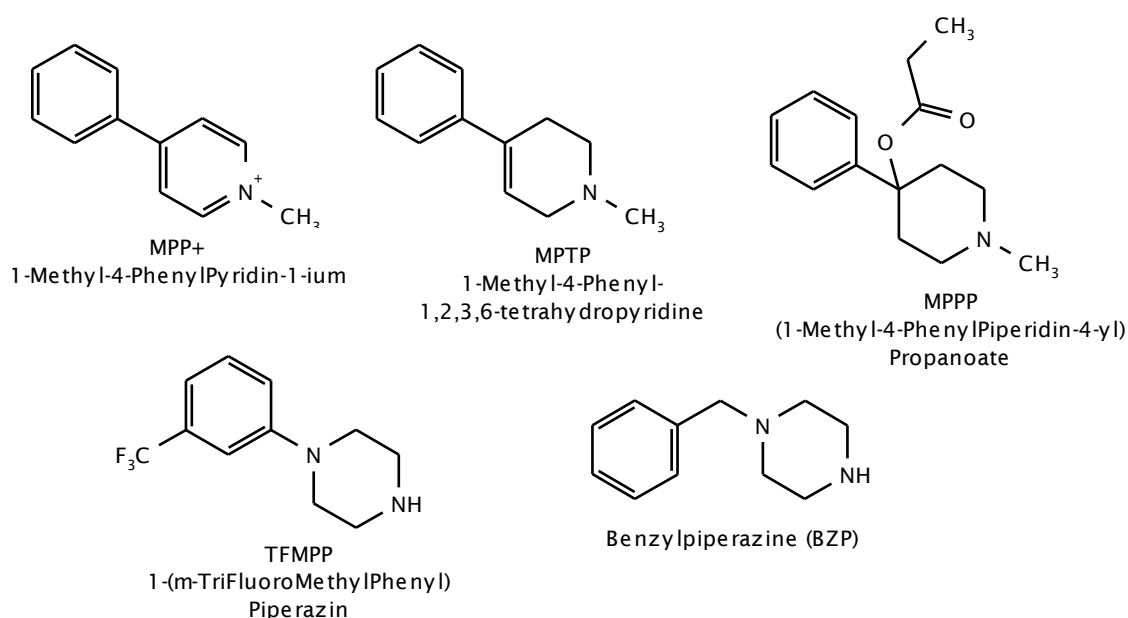
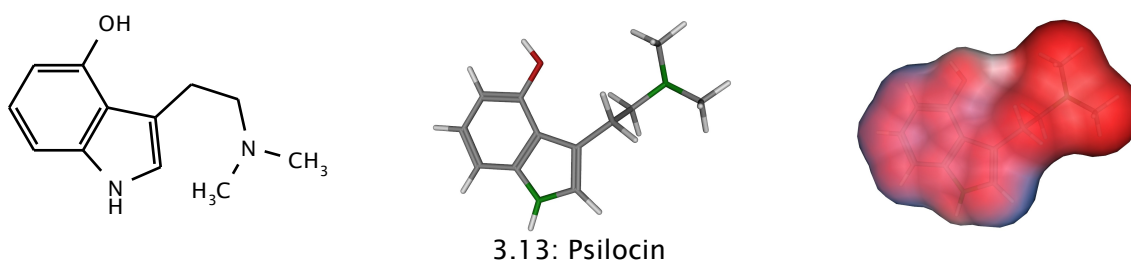


Table 3.3: Pyridines, Piperidines, Piperazines which show activity on NSS

### 3.10 Psilocin / N,N-Dimethyl-4-Hydroxytryptamine



Psilocin is an indole alkaloid, first isolated in 1950s by Albert Hofmann from mushrooms of *Psilocybe* sp., which were used as hallucinogen ritually by indigenous peoples in Mesoamerica. As indole alkaloid, the scaffold is very similar to serotonin. Psilocin acts as agonist on serotonin receptors and as substrate at SERT. It is the active form of the prodrug psilocybin, which is the main alkaloid in the *Psilocybe* species. After intestinal absorption, psilocybin acts potent hallucinogenic drug. The dephosphorylation of psilocybin leads to psilocin, which is also highly active (Schultes and Hofmann, 1998).

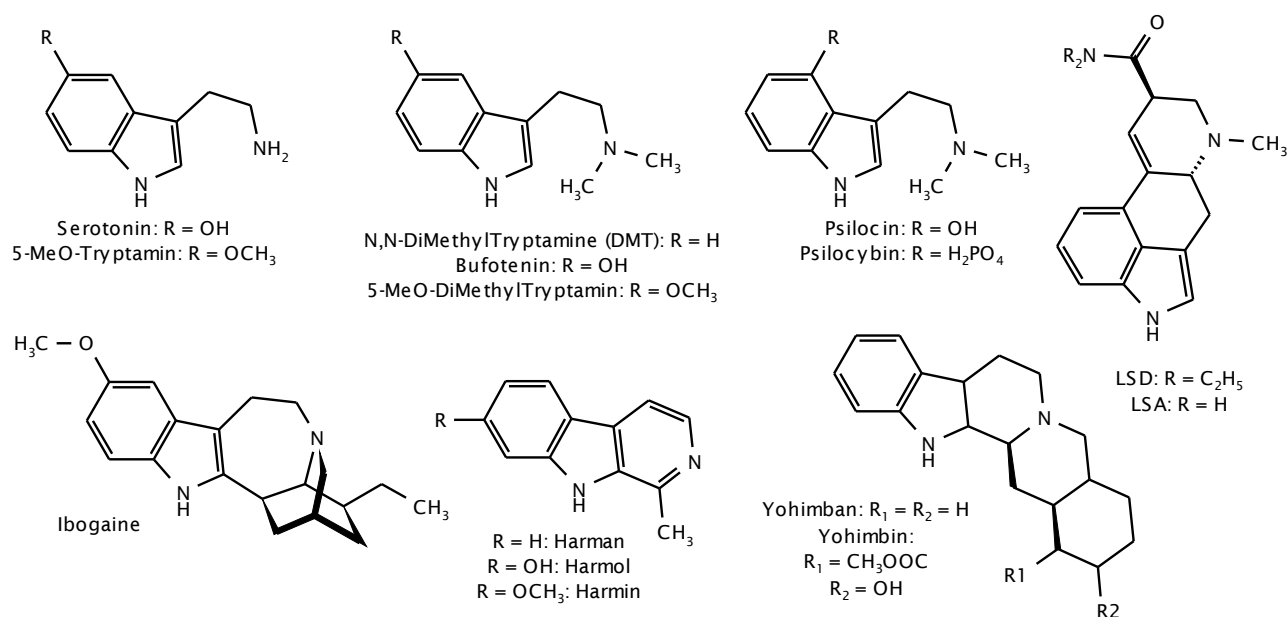


Figure 3.4: Indole alkaloids

### 3.11 DMT/ N,N-Dimethyltryptamine

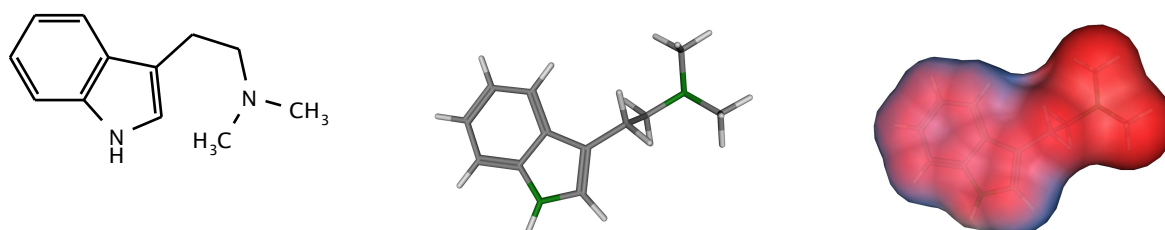


Figure 3.14: DMT

N,N-DMT was discovered in plants growing in the Amazon rainforest, and also in the mucus of the toad, *Bufo bufo*. Orally applied, DMT is rapidly eliminated by MAO in the liver and thereby loses the efficacy. Indigenous peoples in South America use a decoct of DMT containing leaves and the bark of a liana containing beta-carbolines as hallucinogenic drug, which is called ayahuasca. Beta-carbolines act as inhibitor of MAO that makes DMT available orally (Schultes and Hofman, 1998). It is a highly active agonist on 5HT receptors and acts also as substrate on the SERT (Nonaka et al., 2007, Adkins et al., 2001).

## 4 Pharmacoinformatics

In history, modelling of chemical structures and macromolecules was and nowadays still is important for imagination and understanding. As the computational power is rapidly rising, computers are getting more and more important for simulating and predicting in multiple fields (Böhm, Klebe, Kubinyi, 1996). In the field of pharmacoinformatics, modelling of proteins for simulations like docking is useful to elucidate information about the interaction of proteins with ligands. With the also rapidly increasing amount of biological data, such as mutagenesis data, protein sequences and crystal structures, and with the empirical knowledge of the behaviour of proteins, in silico simulations are getting more precise (Rauhut, 2001). Today, pharmacoinformatic methods are used for screening libraries of organic substances to test each compound for their drug-like potential or to find derivatives of known drugs for more specific effects and reducing side effects (Rauhut, 2001; Steger, 2003). It is the first step in the cue of finding hits and leads. There are used ligand based methods, like quantitative structure activity relationship (QSAR), comparative molecular field analysis (CoMFA), and structure based methods, such as homology modelling and docking (Böhm, Klebe, Kubinyi, 1996). In this study we focus on structure based methods.

### 4.1 Structure of Proteins

The structure of a protein determines its proper function. Neurodegenerative diseases, such as Creutzfeldt Jakob disease and also Alzheimer disease, are due to misfolded proteins. It is still not clear, how a protein finds its correct three dimensional structure only from its primary sequence. Nature's algorithm of protein folding is depending on different forces and interactions of the single amino acid and the environment (Lesk, 2001). Until now, it is not possible to predict the three dimensional, tertiary or quaternary structure of a protein only from its primary amino acid sequence (Lesk, 2003; Steger, 2003; Rauhut, 2003).

#### 4.1.1 From Gene to Protein

"Proteins are, where the action is" (Lesk, 2001). They fulfill lots of different structural (cytoskeleton) and functional (enzymes) functions in organisms. The building plan is coded in the DNA in the cell nucleus by using the four nucleotides A, C, G, T. Each of the

20 canonical amino acids, localised in cytoplasm, is coded in a more or less distinct three letter code of nucleotides on the mRNA. At the ribosomes, the primary sequence is synthesised by connecting the carboxyle moiety with the amino group of the following amino acid. The folding takes place immediately and is supported by chaperones. After growing from primary to tertiary or quaternary structure, post translational processes like glycosylation are done in the Golgi apparatus (Löffler, 2001, Lesk, 2001).

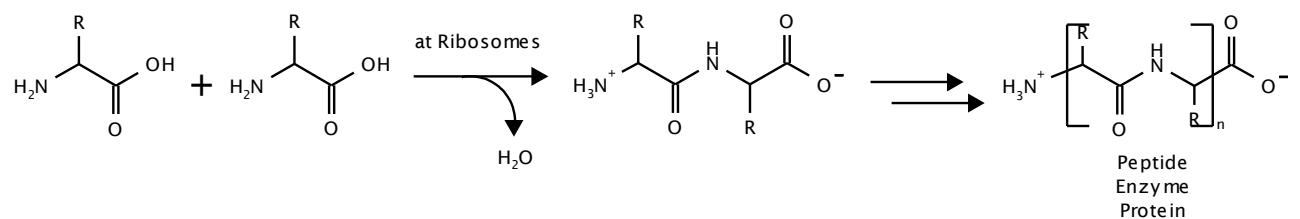


Figure 4.1: Connection of amino acid at the back bone scaffolds peptides

Proteins are macromolecules that are built up of monomers. These chemical bricks are the 20 proteinogenic amino acids. They can be clustered in different ways due to their structure, their physical and chemical properties and their charge (see figure 4.2). All of them are alpha amino carbonic acids with a side chain on the  $\text{C}_\alpha$ , except the simplest member glycine. This property turns the compounds asymmetric, from which nature uses only the L form for building up proteins. Another exceptional amino acid is proline, where the side chain forms a ring from  $\text{C}_\alpha$  to the amino group. Non-canonical amino acids result from modifications after the synthesis of the protein or during metabolism, e.g. selenocysteine, ornithine, pyrrolysine, hydroxyproline (Steger, 2003).

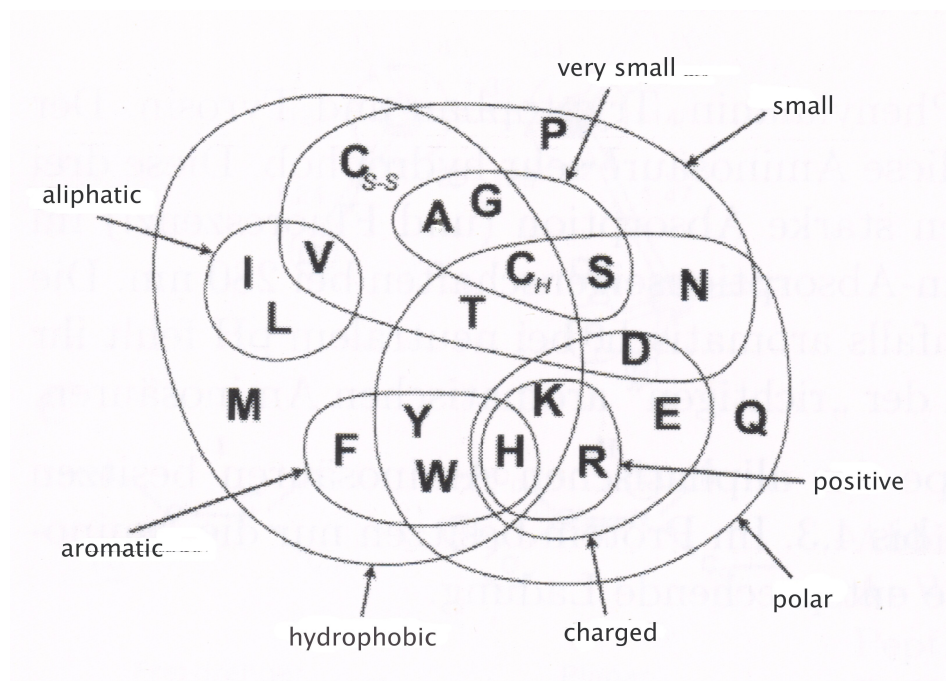


Figure 4.2: Amino acids clustered by physicochemical and steric properties (Steger, 2003).

The hierarchy of the protein architecture starts at the single amino acid and the further resulting linear amino acid sequence. This is called the primary structure of proteins. Each of the residues is characterised by its physicochemical properties. In the collectivity of all concatenated monomers, it leads to distinct properties of the protein. Multiple forces and interactions, such as Van-der-Waals interaction, hydrogen bonds, hydrophobic interaction, salt bridges and disulphide bonds within the conglomerate influence and keep the substructures stable and the protein functionable (Lesk, 2001; Löffler, 2001). The connection between vicinal amino acid is called peptide bond. It is a planar amide, defined by the angle  $\omega$ , with rotatable bonds to  $C^\alpha$  and  $C'$ , defined by the angles  $\Phi$ ,  $\Psi$  and  $\chi$ . This planarity limits the degrees of freedom for the protein backbone. During the folding process, proteins follow a defined geometry concerning the two angles  $\Phi$  and  $\Psi$ . The Ramachandran plot depicts both angles in a two dimensional coordinate system (Lesk, 2001; Steger, 2003).

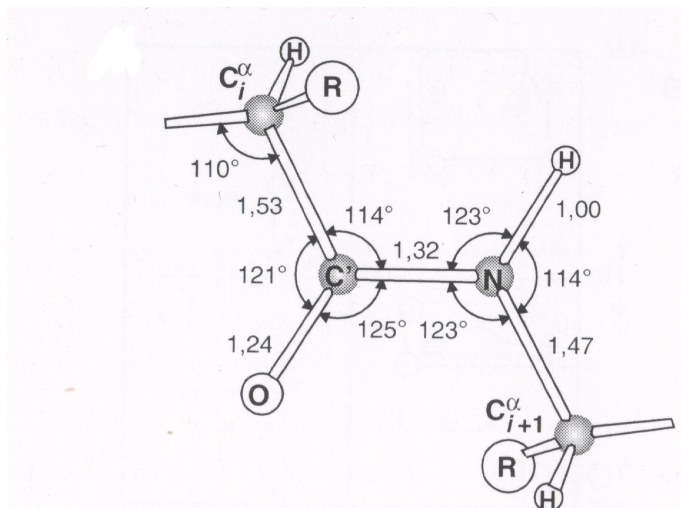


Figure 4.3 shows the characteristic bond lengths and angles of the planar peptide bond (Steger, 2003).

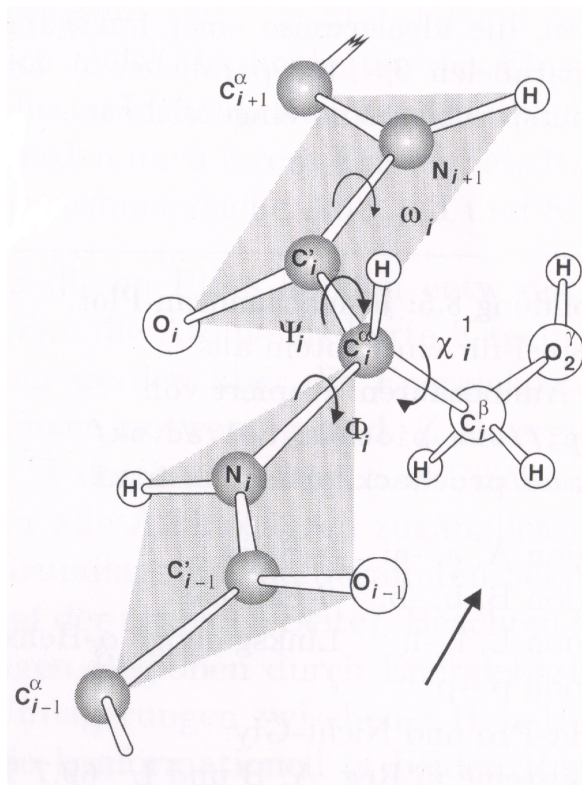


Figure 4.4: The grey surface displays the planarity of the peptide bond and the four torsional angles characterise the geometric properties of the  $C^\alpha$  referred to the neighbours:

$\Phi$ :  $C^\alpha$  to amide N (x-axis in ramachandran plot)

$\Psi$ :  $C^\alpha$  to carbonyl C (y-axis in the phi-psi-plot)

$\chi$ :  $C^\alpha$  to  $C^\beta$  of the residue (angle of the rotamer)

$\omega$ : angle of the amide ( $\omega = 1$ , due to planarity)

(Steger, 2003)



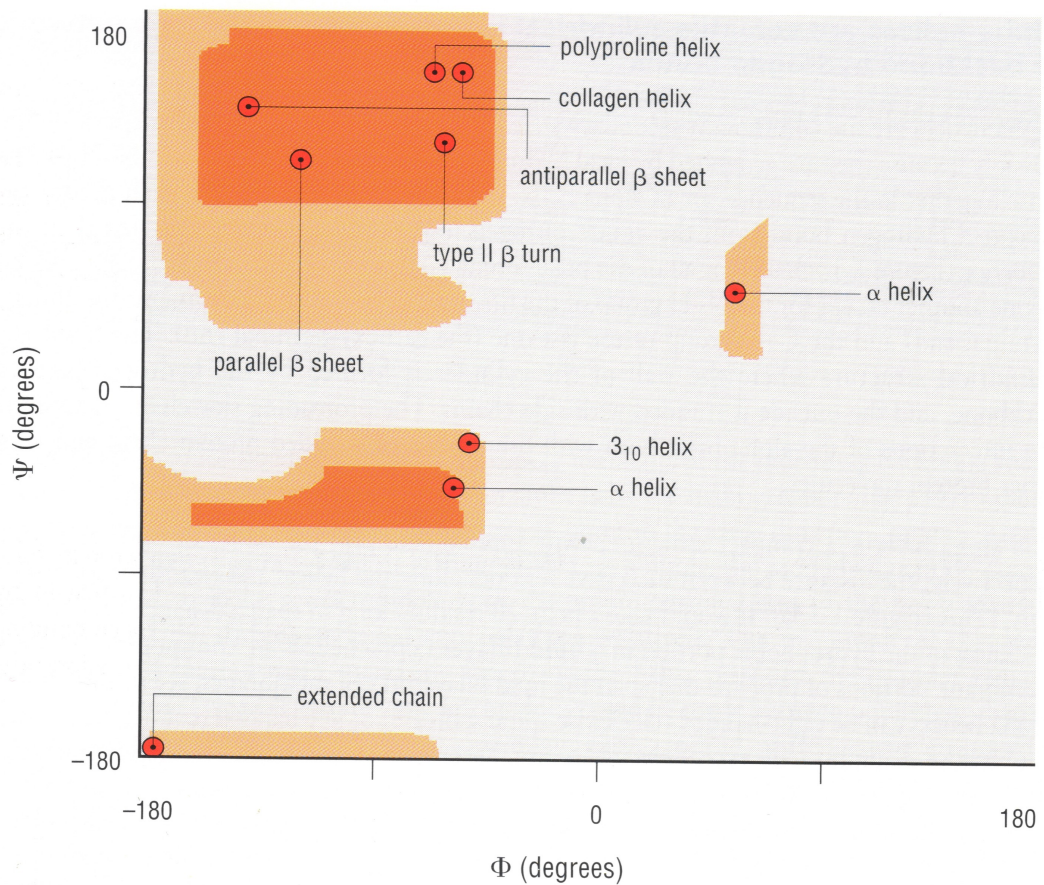


Figure 4.5: Regions of secondary structures in the Ramachandran plot (Petsko and Ringe, 2004)

Structural parameters for protein secondary structures					
Structure	$\phi$	$\psi$	$n$	$d$	$p$
$\alpha$ -helix	-57	-47	3.6	1.5	5.5
$3_{10}$ helix	-49	-26	3.0	2.0	6.0
$\beta$ -helix	-57	-70	4.4	1.1	5.0
Polyproline II helix	-79	+149	3.0	3.1	9.4
Parallel $\beta$ strand	-119	+113	2.0	3.2	6.4
Antiparallel $\beta$ strand	-139	+135	2.0	3.4	6.8

$\phi$  and  $\psi$  are the conformational angles of the mainchain, with  $\omega \sim 180^\circ$  (the trans conformation)  
 $n$  = the number of residues per turn.  
 $d$  = the displacement between successive residues along the helix axis.  
 $p$  = the pitch of the helix, the distance along the helix axis of a complete turn.  
 Note that  $p = n \times d$ . (The equation is exact; the values of  $p$ ,  $n$  and  $d$  in the table have been rounded to two significant figures.)

Figure 4.6: Parameters of secondary structures (Lesk, 2001)

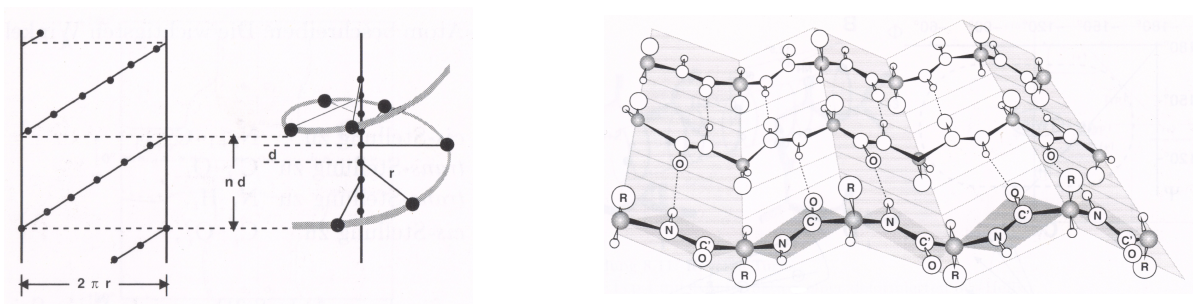


Figure 4.7: (a) Parameters of  $\alpha$  helices;

(b) Antiparallel  $\beta$  strand (Steger, 2003)

Peptides are chains of few amino acids. Small peptides are linear and flexible chains of amino acids. Polypeptides can be organised to repeating structural elements, which are called secondary structure. In a protein, the three dimensional organisation is distinguished in the often appearing  $\alpha$  helix, and the  $\beta$  strand. Depending on the interactions between proximal neighbours helices are subdivided in 3.10 helix,  $\pi$  helix and  $\beta$  helix as well as the  $\beta$  strand appears in distinct structures, i.e. the  $\beta$  bulge and the  $\beta$  barrel. Proteins also contain characteristic three dimensional structures, which are called super secondary structure, e.g.  $\alpha$  and  $\beta$  hairpin, helix-turn-helix motifs, coiled-coil helix (Lesk, 2001; Steger 2003).

The third hierarchy in protein architecture is the tertiary structure, which is the finally folded three dimensional structure of a protein chain or a single protein. Stable and fully functional substructures of proteins are called domains. Forces and interactions mentioned above fix the protein or the domains in the stable and functional morphology, such as covalent disulphide bonds between two cysteines, salt bridges between acidic and basic residues and the weaker interactions like hydrogen bonds, hydrophobic interactions and Van-der-Waals interactions. Proteins and the side chains are moving and interacting with the surrounding medium as well as residues interact among each other.

The quaternary structure is the fully developed protein with its coupled domains and feasible glycosylations. It is the unit which fulfils multiple tasks in cells of living organisms.

### 4.1.2 Structure Analysis

If no crystal structure of a protein is available, biochemical, pharmacological and computational methods are combined to reveal information about the structure. It is hypothesised that identical or homologous sequences give similar three dimensional structure (Rauhut, 2001). Sequenced proteins and known structures are stored in free accessible databases, e.g. the protein data base ([www.rcsb.org](http://www.rcsb.org)). Multiple tools for homology searching and alignments are available, see table 4.1(Lesk, 2001).

Biochemical assays and physical methods are done for elucidating the three dimensional structure and the function of proteins:

- Site directed mutagenesis: By changing one or more codons in the gene and transferring the DNA in an expressing cell by a viral vector, the mutated protein is expressed. The following experiments elucidate, how capacious the mutations change the function of the protein compared to the wild type.
- Cysteine scanning: Single or multiple mutations from the original amino acid to cysteines are done. The cysteines are targeted by reactive agents containing a sulfhydryl moiety (MTS-reagents: MTSET, MTSES, MTSEA). These experiments enable the localisation of the mutation, in particular at membrane proteins and is also called SCAM-method. Cross linking with bifunctional MTS reagents gives information about distances between the reaction partners.
- Alanin scanning: Mutation of residues to alanine elucidate the functional importance of the residue. The workflow is similar to the cysteine scanning method.
- Species scanning mutagenesis: Building chimeras of similar proteins from different species may show differences in functionality and selectivity.
- Engineered Zn and Cd binding sites: Only few amino acids (His) interact with Zn ions. This method is used to elucidate the proximity of specific residues or domains.



Spectroscopic methods deliver precisely the three dimensional structure:

- The X-ray diffraction is the commonly used method to get exact information about the three dimensional structure of a protein. It is highly sophisticated in preparation and analysis. The principle is, to irradiate a crystallised protein with X-rays. From the angles of scattering and interferences, the shape of the atoms can be recalculated. Depending on the method, the resolution ranges from 3.5 to 0.7 Å. At high the resolution it is possible to localize the electronic sheath of each atom, including the often bound water molecules and ions. The crystal structure of a protein is a snap shot of an energetic favourable state for the crystal. Information of domain movements can not be obtained based on the conformation of one crystal.
- With NMR spectroscopy, the disadvantage of only having a single and rigid snap shot, can be abolished. It is predominantly used for peptides and small proteins. In recent years, progress in this field allows to analyse bigger proteins and even membrane proteins. The advantage of this method is to get information about protein domain movements, protein-protein interactions and protein-membrane contacts.

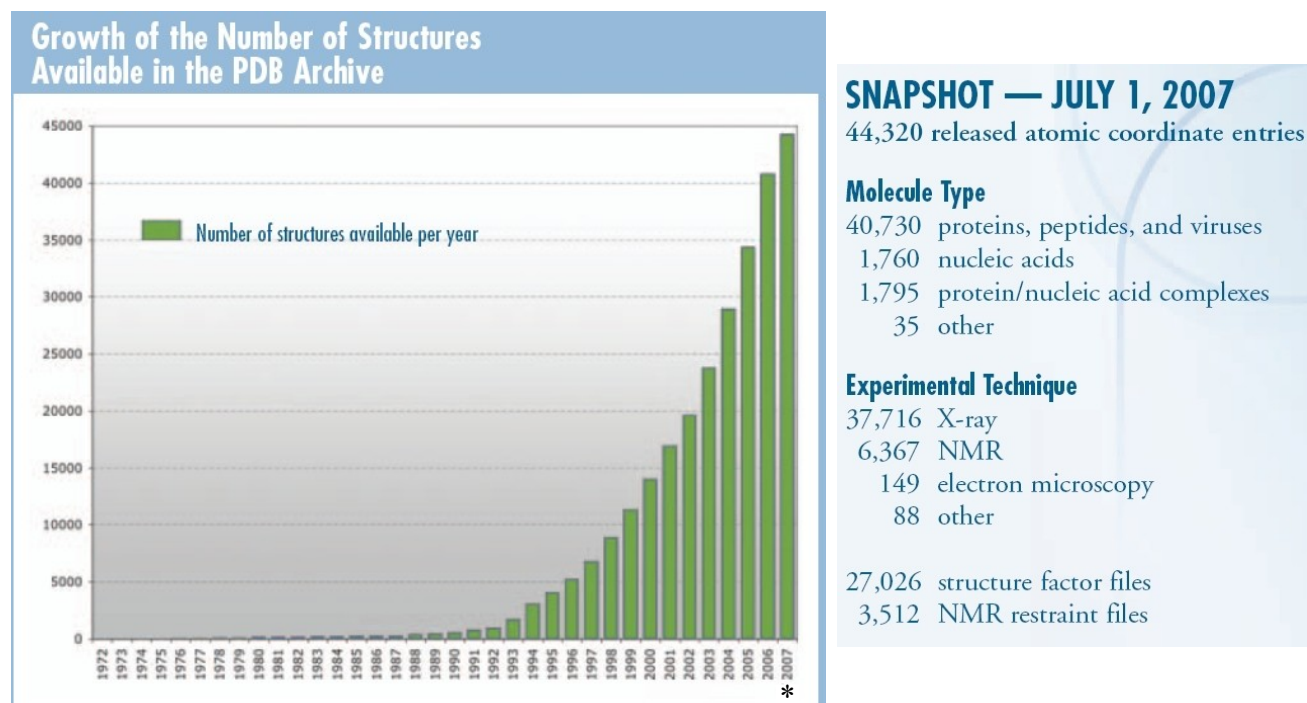


Figure 4.8: Structures in the protein database (annual report of the pdb, 2007);

\* data not completed for 2007 (www.rcsb.org).

Name	Type	WorldWideWeb address
<b>Databases</b>		
CATH	S	<a href="http://www.biochem.ucl.ac.uk/bsm/cath/">www.biochem.ucl.ac.uk/bsm/cath/</a>
GenBank	S	<a href="http://www.ncbi.nlm.nih.gov/GenBank">www.ncbi.nlm.nih.gov/GenBank</a>
GeneCensus	S	<a href="http://bioinfo.mbb.yale.edu/genome">bioinfo.mbb.yale.edu/genome</a>
MODBASE	S	<a href="http://guitar.rockefeller.edu/modbase/">guitar.rockefeller.edu/modbase/</a>
PDB	S	<a href="http://www.rcsb.org/pdb/">www.rcsb.org/pdb/</a>
PRESAGE	S	<a href="http://presage.stanford.edu">presage.stanford.edu</a>
SCOP	S	<a href="http://scop.mrc-lmb.cam.ac.uk/scop/">scop.mrc-lmb.cam.ac.uk/scop/</a>
SWISSPROT+TrEMBL	S	<a href="http://www.ebi.ac.uk/swissprot">www.ebi.ac.uk/swissprot</a>
<b>Template search</b>		
123D	S	<a href="http://www.lmmb.ncifcrf.gov/nicka/123D.html">www.lmmb.ncifcrf.gov/nicka/123D.html</a>
BLAST	S	<a href="http://www.ncbi.nlm.nih.gov/BLAST/">www.ncbi.nlm.nih.gov/BLAST/</a>
DALI	S	<a href="http://www2.ebi.ac.uk/dali/">www2.ebi.ac.uk/dali/</a>
FastA	S	<a href="http://www2.ebi.ac.uk/fasta3">www2.ebi.ac.uk/fasta3</a>
MATCHMAKER	P	<a href="http://bioinformatics.burnham-inst.org">bioinformatics.burnham-inst.org</a>
PHD, TOPITS	S	<a href="http://www.embl-heidelberg.de/predictprotein/predictprotein.html">www.embl-heidelberg.de/predictprotein/predictprotein.html</a>
PROFIT	P	<a href="http://www.came.sbg.ac.at">www.came.sbg.ac.at</a>
THREADER	P	<a href="http://globin.bio.warwick.ac.uk/jones/threader.html">globin.bio.warwick.ac.uk/jones/threader.html</a>
UCLA-DOE, FRYSVR	S	<a href="http://www.doe-mpi.ucla.edu/people/frsvr/frsvr.html">www.doe-mpi.ucla.edu/people/frsvr/frsvr.html</a>
<b>Sequence alignment</b>		
BCM SERVER	S	<a href="http://dot.imgen.bcm.tmc.edu:9331/">dot.imgen.bcm.tmc.edu:9331/</a>
BLAST	S	<a href="http://www.ncbi.nlm.nih.gov/BLAST">www.ncbi.nlm.nih.gov/BLAST</a>
BLOCK MAKER	S	<a href="http://blocks.fhcrc.org/blocks/blockmkr/make_blocks.html">blocks.fhcrc.org/blocks/blockmkr/make_blocks.html</a>
CLUSTAL	S	<a href="http://www2.ebi.ac.uk/clustalw/">www2.ebi.ac.uk/clustalw/</a>
FASTA3	S	<a href="http://www2.ebi.ac.uk/fasta3/">www2.ebi.ac.uk/fasta3/</a>
MULTALIN	S	<a href="http://pbil.ibcp.fr/">pbil.ibcp.fr/</a>
<b>Modelling</b>		
COMPOSER	P	<a href="http://www-cryst.bioc.cam.ac.uk">www-cryst.bioc.cam.ac.uk</a>
CONGEN	P	<a href="http://www.congenomics.com/congen/congen.html">www.congenomics.com/congen/congen.html</a>
CPH models	S	<a href="http://www.cbs.dtu.dk/services/CPHmodels/">www.cbs.dtu.dk/services/CPHmodels/</a>
DRAGON	P	<a href="http://www.nimr.mrc.ac.uk/mathbio/a-aszodi/dragon.html">www.nimr.mrc.ac.uk/mathbio/a-aszodi/dragon.html</a>
ICM	P	<a href="http://www.molsoft.com">www.molsoft.com</a>
InsightII	P	<a href="http://www.csc.fi/english/research/software/insightii">www.csc.fi/english/research/software/insightii</a>
MODELLER	P	<a href="http://guitar.rockefeller.edu/modeller/modeller.html">guitar.rockefeller.edu/modeller/modeller.html</a>
MOE	P	<a href="http://www.chemcomp.com">www.chemcomp.com</a>
LOOK	P	<a href="http://www.mag.com">www.mag.com</a>
QUANTA	P	<a href="http://scv.bu.edu/documentation/software-help/scientific-engineering/quanta.html">scv.bu.edu/documentation/software-help/scientific-engineering/quanta.html</a>
SYBYL	P	<a href="http://www.tripos.com">www.tripos.com</a>
SCWRL	P	<a href="http://www.cmpharm.ucsf.edu/bower/scrwl/scrwl.html">www.cmpharm.ucsf.edu/bower/scrwl/scrwl.html</a>
SWISS-MOD	S	<a href="http://www.expasy.ch/swissmod">www.expasy.ch/swissmod</a>
WHAT IF P		<a href="http://www.sander.embl-heidelberg.de/whatif/">www.sander.embl-heidelberg.de/whatif/</a>
<b>Model evaluation</b>		
ANOLEA	S	<a href="http://www.fundp.ac.be/pub/ANOLEA.html">www.fundp.ac.be/pub/ANOLEA.html</a>
AQUA	P	<a href="http://www-nmr.chem.ruu.nl/users/rull/aqua.html">www-nmr.chem.ruu.nl/users/rull/aqua.html</a>
BIOTECHd	S	<a href="http://biotech.embl-ebi.ac.uk:8400/">biotech.embl-ebi.ac.uk:8400/</a>
ERRAT	S	<a href="http://www.doe-mpi.ucla.edu/erratsrvr.html">www.doe-mpi.ucla.edu/erratsrvr.html</a>
PROCHECK	P	<a href="http://www.biochem.ucl.ac.uk/roman/procheck/procheck.html">www.biochem.ucl.ac.uk/roman/procheck/procheck.html</a>
ProCeryone	P	<a href="http://www.proceryon.com/">www.proceryon.com/</a>
Prosalle	P	<a href="http://www.came.sbg.ac.at">www.came.sbg.ac.at</a>
PROVE	S	<a href="http://www.ucmb.ulb.ac.be/UCMB/PROVE">www.ucmb.ulb.ac.be/UCMB/PROVE</a>
SQUID	P	<a href="http://www.yorvic.york.ac.uk/oldfield/squid">www.yorvic.york.ac.uk/oldfield/squid</a>

Table 4.1: WorldWideWeb addresses for homology search, alignment and modelling tools;  
P = Programm, S = Server (Marti-Renom et al., 2000)

## 4.2 Homology Modelling

Homology modelling is the computational method to predict the tertiary structure of proteins. Two different approaches in homology modelling are available, where the ab-initio or de novo method is done by calculating the global free energy minimum. The three dimensional protein folds are modelled only from the amino acid sequence (Moult, 2005). In contrast, the comparative modelling uses templates of homologous proteins, which is the more accurate method (Baker et al., 2001; Martí-Renom et al., 2000). Due to the enormous effort necessary to get a protein crystallized, homology modelling is an accurate method for research on proteins. The cue of creating homology models is divided in four steps: finding homologous structures, aligning the sequence, creating the model and validating the model (Baker, 2001).

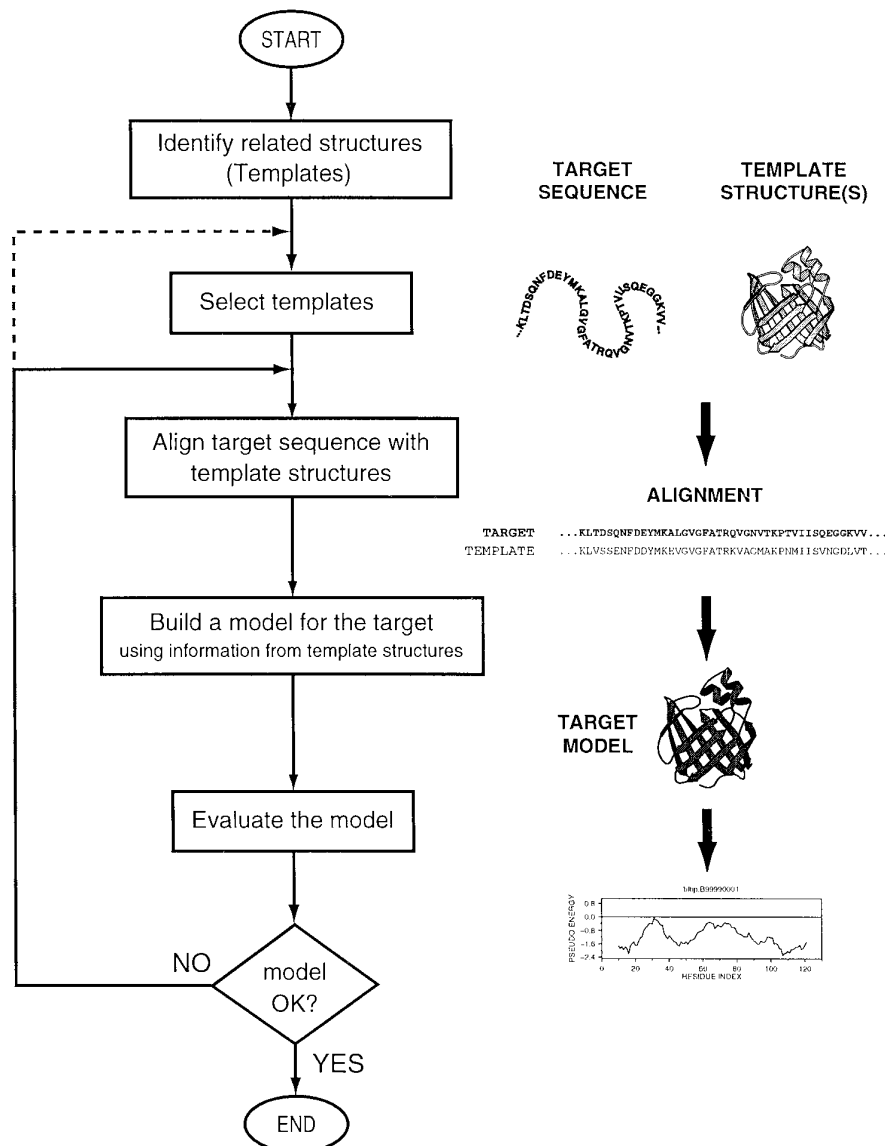


Figure 4.9: Workflow of homology modelling (Martí-Renom et al., 2000)

## 4.2.1 The Alignment

The higher the sequence identity the higher the accuracy of the model is. But also the sequence homology is an important factor in molecular modelling. In this case residues with similar chemical characteristics can be treated as "identical". Different algorithms for sequence alignment are available (Rauhut, 2001; Martí-Renom et al., 2000). The first step in the cue of homology modelling is finding templates for the query structure by using sequence comparison methods, such as PSI-BLAST ([www.ebi.ac.uk/Tools/psiblast](http://www.ebi.ac.uk/Tools/psiblast)). Another way of alignment is a sequence-structure comparison. In this case, the fold of a protein is compared, which is called fold recognition (Baker et al., 2001; Lee et al., 2007).

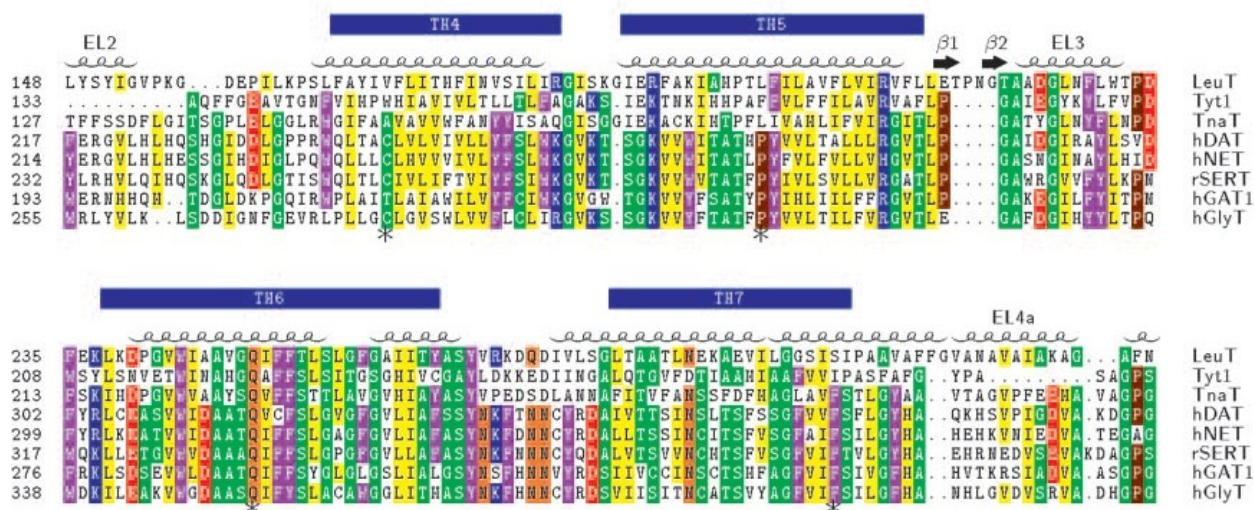


Figure 4.10: Example of an alignment (Beuming et al., 2006)

## 4.2.2 The Template

An important step towards a good homology model is a good template. A high degree of both structural homology and function, a high quality model will result. In many cases the template is a crystal structure taken by X-ray structure analysis. A high resolution of the crystal structure also elevates the quality of the model. There are also NMR-data used for creating models. The advantage of NMR spectra is that movements of the protein and the residues get visible. Molecular dynamic simulation and energy minimization calculations can be done with higher precision, if protein movements are known. But it is still a challenge to get NMR spectra of huge proteins or of membrane proteins (Lee et al., 2007; Martí-Renom et al., 2000)

### 4.2.3 The Model

In principal, comparative homology modelling superposes conserved residues on the coordinates of the template. Differing residues are changed on basis of the C $\alpha$  atoms of the protein backbone. An indicator for the precision is the root mean square deviation (RMSD), which is the spatial deviation of query protein from the template. At a sequence identity of more than 50 %, the RMSD between the model and the template is about 1 Ångström (Å, 1 Å = 10<sup>-10</sup> m = 100 pm), which approximates medium resolution X-ray structures. The error is due to different orientations of the side chains. When the sequence identity ranges from 30 to 50 %, the RMSD increases to 1,5 Å based on the backbone atoms and modelled loops. A sequence identity of lower than 30 % the error as well as the RMSD increases very rapidly (Baker et al., 2001; David et al., 2005).

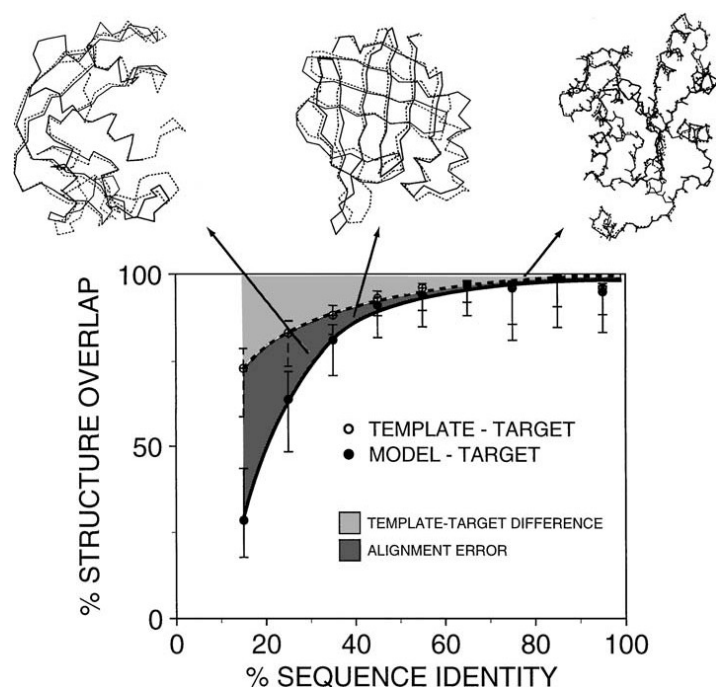


Figure 4.11: Correlation of sequence identity and accuracy of structure prediction (Marti-Renom et al., 2000)

Multiple methods and programs are available for creating homology models (Marti-Renom et al., 2001). The necessary calculations for building the homology model *in silico* are done in forcefields parameterised for proteins. These forcefields are gained empirically due to the increasing knowledge about the behaviour of the proteins. During the modelling process energy minimizations based on a rotamer library are done to avoid clashes of residues. This leads to slight movements in the backbone and also the side chains. Conserved residues or protein regions are calculated exactly, since they act as anchor point during the modelling. The loop folding, due to their high degree of flexibility, still is quite difficult to calculate. They are modelled by using loop folding databases (Al-Lazikani et al., 2001).

## 4.2.4 Refinement and Validation

Up to now, there are no reliable procedures for evaluating protein homology models. Different programs are available, considering steric, geometric and energetic criteria (Petrey and Honig, 2005; Dunbrack et al., 2006). Depending on the software package, energetic functions, such as electrostatic solvation energy and H-bonding are calculated. For evaluation of the protein model, protein geometry tools, such as the Ramachandran plot can be used. A good parameter are the amount of clashes between residues. On the other hand the model can be compared with evident biological data, such as characteristic regions or binding sites (MOE).

## 4.3 Docking

Docking can be done with small ligands and also proteins. In this study, we concern small ligand docking, which is the prediction of ligand orientation in the binding site of a protein. It was pioneered in the early 1980s and now docking and virtual screening is an important method in drug discovery for hit finding and lead optimisation. The aim is to get structural information for describing interactions and for predicting activity. Multiple software packages are now available for this multistep procedure and allow different consideration of the interaction partners (Kitchen et al., 2004; Leach et al., 2006).

docking program	algorithm	ligand treatment	protein flexibility
AutoDock	genetic algorithm	stochastic simulation	systematic
CDOCKER	grid based MD simulation	incremental, simulation	
DOCK		incremental	
FlexE		incremental	
FlexX		database	
FLOG		incremental	systematic
Glide	Monte Carlo	stochastic simulation	
GOLD	genetic algorithm		
Hammerhead			
ICM	Monte Carlo		
MCDOCK			none
MDD	Monte Carlo		
MOE-Dock		stochastic	
PRO_LEADS		stochastic	
QXP	Monte Carlo		
Surflex			

Table 4.2: Docking software and features

### 4.3.1 The Ligand

Chemical structures have a continuous intramolecular movement due to the degrees of freedom of certain chemical bonds. To consider these movements during the docking process, multiple conformers of the ligand must be created. Different approaches are possible to allow ligand flexibility:

- Systematic methods: The systematic search forces to explore all degrees of freedom of a molecule. Another method is the incremental one, where the ligand is divided into rigid and flexible fragments. After that, it is rebuilt based on the rigid core. Conformational databases are also used for rigid docking, i.e. during the docking process the ligand is treated rigid.
- Stochastic methods: This method makes random changes of the ligand by using Monte Carlo methods, genetic algorithms or the tabu search.
- simulation methods: Molecular dynamic simulation is a popular approach for creating conformers of ligands. In this approach, the molecule is virtually heated. A problem is, that this method cannot cross high energy barriers and therefore the compound might accommodate in local minima of the energy surface.

(Kitchen et al., 2004)

### 4.3.2 The Site

As well established, the principle of protein ligand interaction seems like the lock and key principle. In physiological conditions, the ligand approximates its receptor by multiple forces, such as electrostatic energies, H-bonds and Van-der-Waals interactions. When the ligand reaches its receptor, an adaptation of both the receptor and the ligand takes place (Kubinyi, 1996). For this reason not only ligand flexibility, but also the side chain movements of the protein must be considered, which is still challenging. Most of the docking programs treat the receptor nearly rigid. For considering receptor flexibility different approaches are available:

- molecular dynamic simulations
- Monte Carlo calculations
- rotamer libraries and
- dead-end elimination algorithms

(Kitchen et al., 2004).

Before starting the docking run, the binding site should be defined. In the present study we defined the biologically important interaction partners which are known from the literature by both site directed mutagenesis and docking studies.

### **4.3.3 The Placement**

The placement is the positioning of the ligand in the binding site. This is calculated with different methods. Depending on the algorithm, the following methods can be distinguished:

- Forcefield based methods consider ligand flexibility as well as protein, i.e. binding site flexibility by using molecular dynamic simulation or Monte Carlo algorithms. These methods perform most accurately.
- Evolutionary methods create sets of different ligand conformers called chromosome, where the single genes are parameters such as angles between vicinal atoms of the molecule. By mutations of single genes, new chromosome populations are generated. In the group of evolutionary methods are genetic algorithms, evolutionary programming and tabu search.
- Fragment based or incremental methods split the ligand into fragments. The first fragment should have anchoring features on which the ligand is built up in the binding site.
- Shape complementary based methods estimate the shape of both, the ligand and the binding site. After evaluation of both shapes, the best fit is estimated on basis of energetic energy terms, such as Van-der-Waals interaction surface.
- Hybrid methods combine the previously mentioned methods to improve the accuracy of docking.

(David et al., 2005)

### **4.3.4 Scoring and Ranking the Poses**

After the docking run up to thousands of poses are generated. Scoring functions evaluate the docked poses on basis of energetic values as combination of entropic and enthalpic effects or other parameters such as distances. Each docking program can be combined with each scoring function. Finally the most favourable poses should be ranked first by estimating parameters such as binding affinity, binding mode or biological activity. The scoring functions as well can be divided in groups depending on the algorithm that is used:



- Forcefield based scoring functions end up and quantify docking poses in the sum of energy terms, e.g. Van-der-Waals interaction energy, electrostatic energy and geometrical parameters. They consider receptor-ligand interactions as well as internal ligand energies.
- Empirical scoring functions should reproduce experimental data like binding conformations or energies. They are obtained by a sum of uncorrelated terms due to regression analysis. These terms are elucidated from crystal structure information or experimentally determined binding energies.
- Knowledge based scoring functions reproduce experimental structures
- Shape based scoring functions are calculated simply and often act as primary filter in the scoring cascade.
- Consensus scoring as combination of different scoring functions.

(Kitchen et al., 2004; David et al., 2005)

It is still difficult to get the most accurate pose only on basis of scoring functions. In contrast to docking into crystal structures or homology models, the prediction of activity or binding affinity is quite challenging. Because in vivo there is a multitude of forces present, which are highly complicated to calculate (Leach et al., 2006).

## 5 Neurotransmitter:sodium symporter

In this chapter we will give a description of structural elements of SERT and DAT, and also of the template LeuT. These membrane proteins share common features like the sodium dependency of transport action, the similar three dimensional structure of 12 TMDs and conserved regions in the substrate translocation pathway, including the binding site.

This type of membrane proteins, the solute carriers, facilitate the penetration of the cell membrane for charged organic molecules, that are not able to diffuse through. The driving force is an osmotic gradient built up by Na/K-ATPases. They are ubiquitous in prokaryotes and eukaryotes and passably homologous. Recently, investigations have been done on the bacterial amino acid transporters Tyt1, a selective tyrosine transporter expressed in *Escherichia coli* (Quick et al., 2006) and TnaT, a selective tryptophane transporter expressed in *Symbiobacterium thermophilum* (Androutsellis-Theotokis et al., 2003; Zomot et al, 2007), which are also structural and functional homologues of LeuT

Today, more than 300 known transport protein families are classified and accessible in the transport classification database from the University of California, San Diego [www.tcdb.org](http://www.tcdb.org). (Saier et al, 1999). The taxonomy of the transporter classification system (TC-System) is:

- class, determined by the mode of transport and energy coupling mechanism
- subclass, type of transporter and energy coupling mechanism
- family or superfamily
- phylogenetic cluster within a family
- substrate specificity

The neurotransmitter:sodium symporter are classified in:

- class 2 - electrochemical potential-driven transporters
- subclass A - porters (uniporters, symporters, antiporters)
- family 22 - neurotransmitter sodium symporters

coded in a five digit TC-number. Table 5.1 lists the classified members of the NSS family:

TC-Code	Name	Abbreviation	swissprot-code	Organism
2.A.22.1.1	Serotonin transporter	SERT	P31645	Homo sapiens
2.A.22.1.2	Norepinephrine transporter	NET	P23975	Homo sapiens
2.A.22.1.3	Dopamin transporter	DAT	Q01959	Homo sapiens
2.A.22.1.4	Dopamin transporter	T23G5.5	Q03614	Caenorhabditis elegans
2.A.22.2.1	Proline transporter	PROT		Rattus norvegicus
2.A.22.2.2	Glycine transporter	GlyT1c		Rattus norvegicus
2.A.22.2.3	Neutral and cationic amino acid transporter	B0+		Homo sapiens
2.A.22.2.4	Neutral amino acid transporter	CAATCH1		Manduca sexta
2.A.22.2.5	Neutral amino acid transporter	KAAT1		Manduca sexta
2.A.22.2.6	Glycine transporter	GlyT2b		Mus musculus
2.A.22.2.7	Acetylcholin and choline transporter	Snf-6	O76689	Caenorhabditis elegans
2.A.22.2.8	nutrient amino acid transporter	AAT1	AAR08269	Aedes aegypti
2.A.22.2.9	Densovirus type-2 receptor	Nsd-2	BAG39453	Bombyx mori
2.A.22.3.1	Betaine and GABA transporter		P48065	Homo sapiens
2.A.22.3.2	GABA transporter	GAT-1	P50531	Homo sapiens
2.A.22.3.3	Taurine transporter	Taurine	P31641	Homo sapiens
2.A.22.3.4	Creatine transporter	Creatine		Oryctolagus cuniculus
2.A.22.3.5	Creatine transporter	CRT	P28570	Rattus norvegicus
2.A.22.3.6	GABA transporter	GAT-1	AAT02634	Caenorhabditis elegans
2.A.22.4.1	Tryptophan transporter	TnaT		Symbiobacterium thermophilum
2.A.22.4.2	Leucine transporter	LeuTAa	O67854	Aquifex aeolicus
2.A.22.5.1	Hypothetical Na <sup>+</sup> -dependent permease	MJ1319		Methanococcus jannaschii
2.A.22.5.2	Tyrosine transporter	Tyt1	Q8RHM5	Fusobacterium nucleatum
2.A.22.6.1	Proline transporter	SIT1	Q64093	Rattus norvegicus
2.A.22.6.2	Neurotransmitter transporter	NTT4	P31662	Rattus norvegicus
2.A.22.6.3	Neutral amino acid transporter	B0AT1	Q695T7	Homo sapiens

Table 5.1: Members of the Neurotransmitter:Sodium Symporter Family, from [www.tcd.org](http://www.tcd.org) (Quick, 2002)

## 5.1 The Leucine Transporter from Aquifex aeolicus

The Leucine transporter (LeuT) is a homologous transmembrane transport protein of the mammalian NSS. It transports the amino acids leucine and alanin with higher affinity through the bacterial membrane driven by the sodium gradient in the cytoplasm (Shi et al., 2008). In 2005, the crystal structure of the bacterial amino acid transporter was published. The crystallized transporter from Aquifex aeolicus (LeuT, pdb code 2A65, [www.rcsb.org](http://www.rcsb.org), swiss prot ID O67854) has bound its substrate leucine in the centre of the protein, two sodium ions near the substrate binding site, one chloride ion, five detergent molecules and 210 water molecules. The resolution is at 1.65 Å and shows the transporter in the substrate occluded state. For substrate translocation, LeuT must change its conformation. It is supposed that the transporter appears in three positions, standing in equilibrium.

- outward faced, where sodium and substrate can bind
- the occluded state, where substrate is bound (X-ray structure in the pdb)
- inward faced, where leucine and the two sodium ions are being released.

(Yamashita et al., 2005).

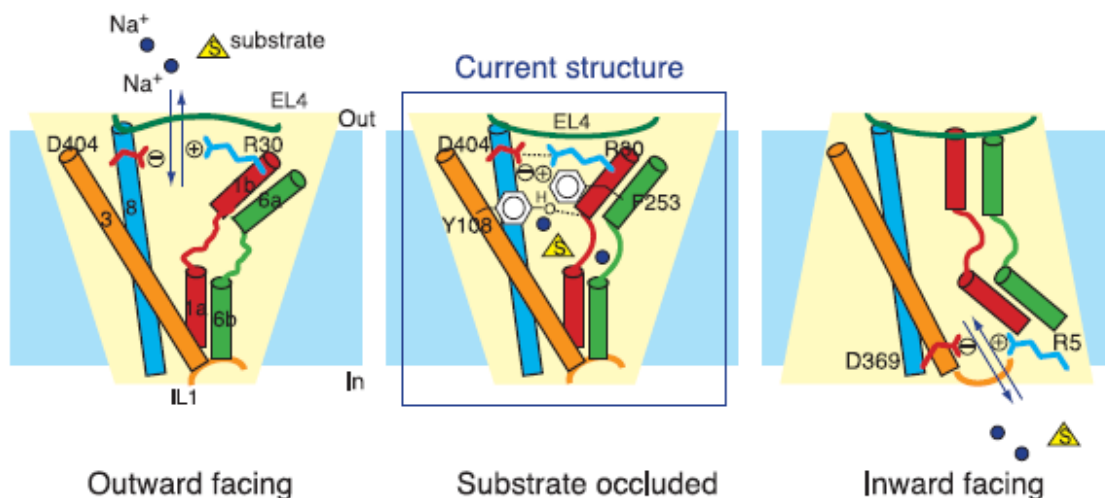


Figure 5.1: The transportation cycle of LeuT, from Yamashita et al., 2005

In 2007, crystal structures of LeuT with bound TCAs were published. Due to the conserved residues near the bound TCA, these structures revealed the putative binding pocket of antidepressants on the SERT, DAT and NET (Singh et al., 2007, Zhou et al., 2007, pdb-codes 2Q6H, 2Q72, 2QB4, 2QE1, 2QJU). The co-crystallized TCAs are clomipramine, imipramine and desipramine. One of the crystals has bound alanine as substrate (pdb-code 2QE1). These structures reveal a second binding pocket, outside of the extracellular gate, where substrate is supposed to be bound (Shi et al., 2008; Rudnick, 2007).

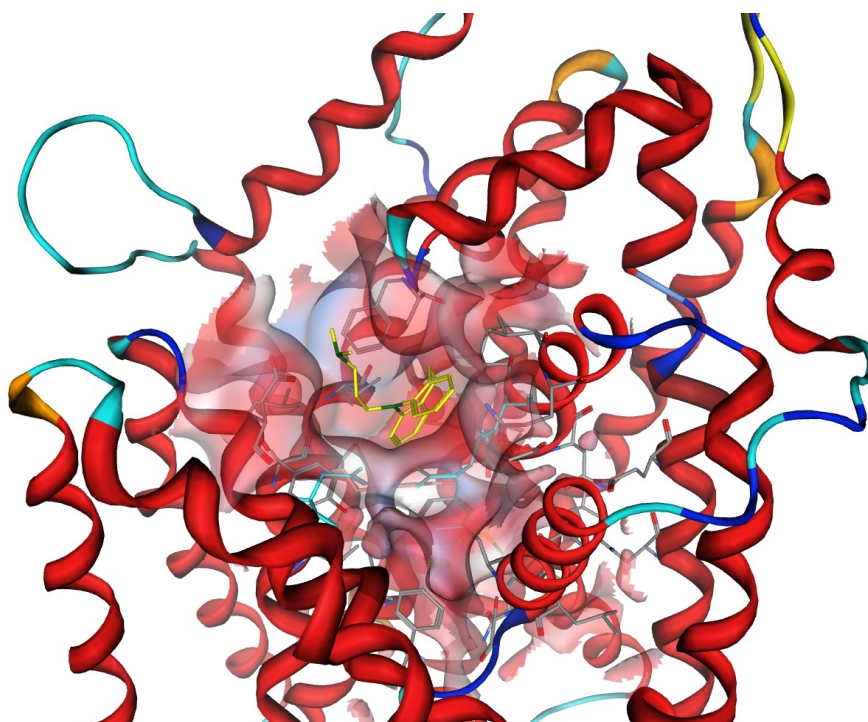


Figure 5.2: The extracellular vestibule of LeuT with the bound TCA desipramine, pdb-code 2Q72

## Structure and Function (from Yamashita et al., 2005)

LeuT, as bacterial homologue of the mammalian NSS, shows a certain degree of identity with the monoamine transporters. Also the length, the three dimensional structure and the transport action is supposed to be similar all over the NSS family. The primary sequence of LeuT consists of 519 amino acids with a formula molecular weight of about 58,1 kDa ([www.rcsb.org](http://www.rcsb.org)). They are organized to 12 TMDs with transporter dimensions of 70 Å through the membrane and 48 Å in diameter.

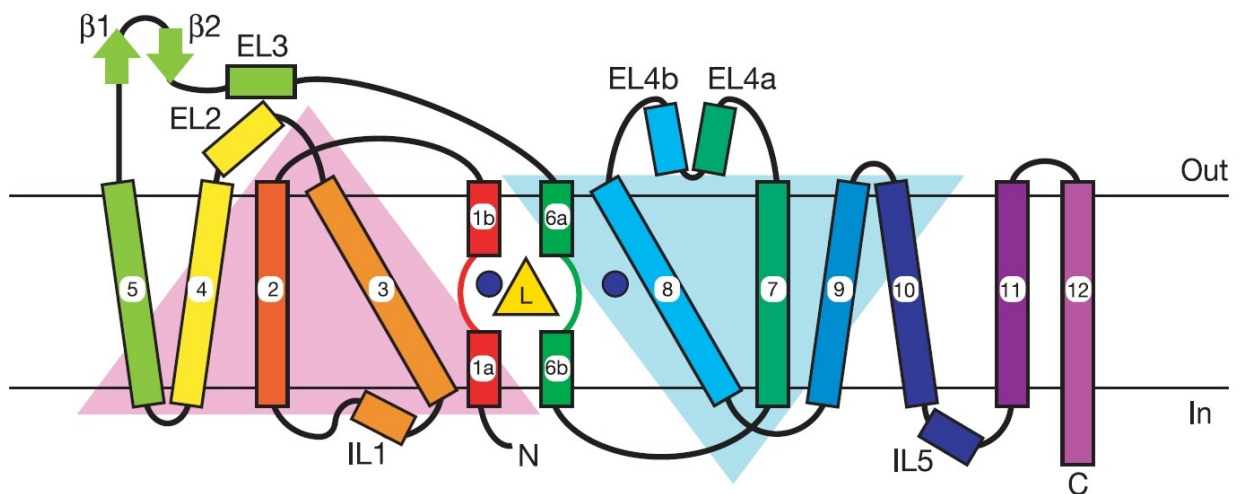


Figure 5.3: LeuT depicted in 2D. The structural repeat of the first ten domains is marked as pink and blue arrows.

The extracellular face consists of 3 markable loops. EL2 which connects TM3 and TM4, rises furthest into the extracellular space. EL4, connecting TM7 and TM8, and EL2 are composed of  $\alpha$  helical segments. EL3 has a helix and two short  $\beta$  strands and is located between TM5 and TM6. The two helical segments of EL4 point like an arrow into the cavity where bound TCAs were crystallized with the transporter, near the extracellular gate, see figure 5.2.

On the intracellular face there are two noticeable loops forming  $\alpha$  helices. IL1, formed to a short  $\alpha$  helix connects TM2 and TM3 and participates on the substrate translocation pathway by occluding the intracellular gate. The second short  $\alpha$  helix is IL5 connects TM10 and TM11.



ligand	residues	domain
leucine -NH3+	A22(BB)	TM1
	F253(BB)	TM6
	T254(BB)	TM6
	S256(SC)	TM6
leucine -COO-	Na1	-
	L25(BB)	TM1
	G26(BB)	TM1
	Y108(SC)	TM3
leucine -CH2-CH-(CH3)2	V104	TM3
	Y108	TM3
	F253	TM6
	S256	TM6
	F259	TM6
	S355	TM8
	I359	TM8

ligand	residues	domain
Na(1)	leucine	-
	A22(BB)	TM1
	N27(BB,SC)	TM1
	T254(SC)	TM6
	N286(SC)	TM7

ligand	residues	domain
Na(2)	G20	TM1
	V23	TM1
	A351	TM8
	T354	TM8
	S355	TM8

Table 5.2 Residues interacting with sodium and moieties of leucine, BB = back bone, SC = side chain

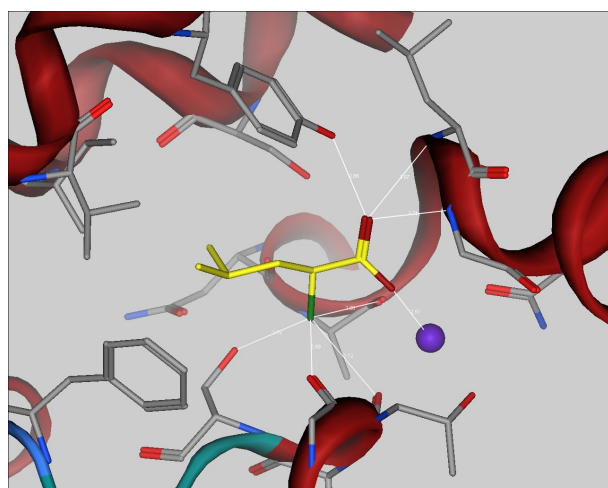


Figure 5.4a: Interaction partners of Leucine in the binding Site of the LeuT

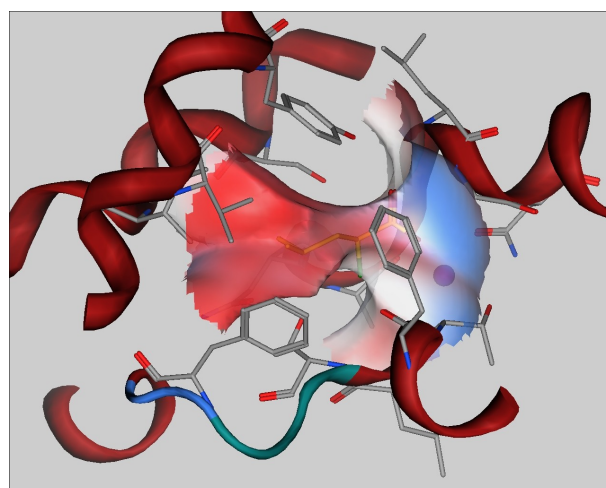


Figure 5.4b: Van-der-Waals surface of the binding pocket. Red lipophilic, blue hydrophilic

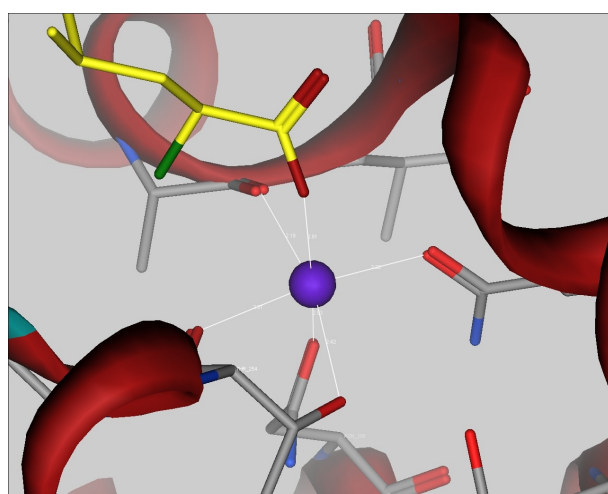


Figure 5.4c: Interaction partners of Na1 in the binding site of LeuT

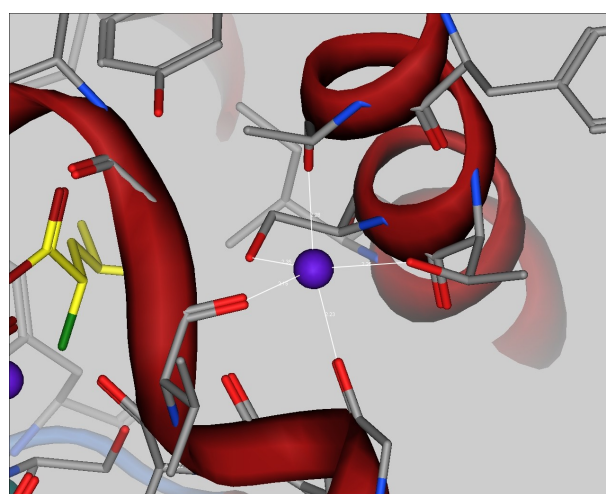
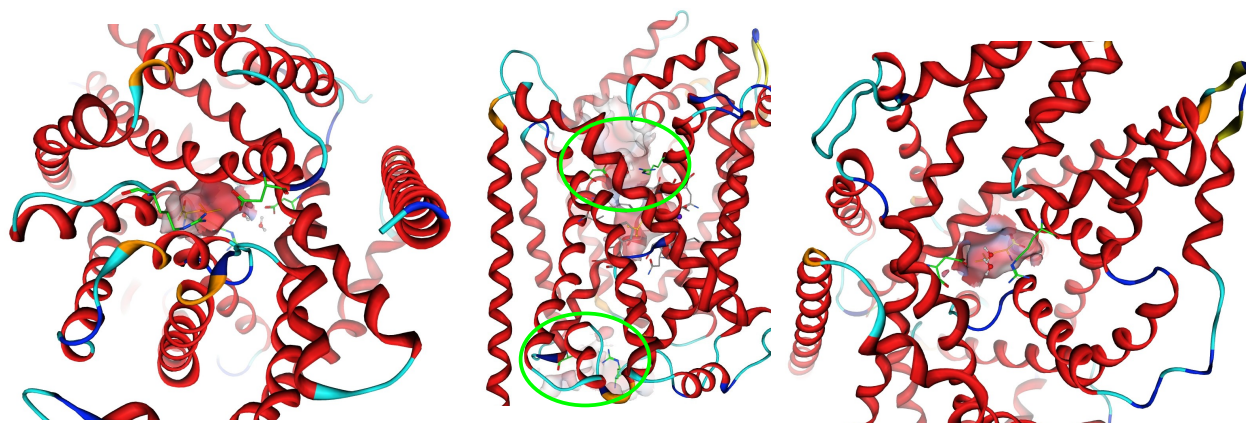


Figure 5.4d: Interaction partners of Na2 in the binding site of LeuT

The binding site of leucine in the crystal contains the substrate and two sodium ions. It is totally free of water. The charged moiety of leucine is coordinated by hydrophilic residues and one sodium. The lipophilic part of leucine protrudes in a lipophilic pocket, surrounded by aliphatic and aromatic residues. Figures 5.4 a - d show the binding sites of leucine and the sodium ions. In table 5.2 the involved residues are listed.

In the first 10 TMDs, there is a structural repeat from TM1 to TM5 and from TM6 to TM10 (see figure 5.3). The alpha carbons of the first 5 TMDs can be superposed over the next 5 TMDs. The involved domains in the substrate permeation are TM1, TM3, TM6 and TM8, the protein core. The highest degree of conserved residues was found in TM1 and TM6, both part of the substrate and ion binding site in common with TM3 and TM8. In the centre of these four TMDs, in the half way through the phospholipid bilayer, distinct unwound regions are forming the binding site for leucine and two sodium ions. TM2, TM4, TM5, TM7, TM9 and TM10 support the protein core in conformational changes and anchoring in the membrane. In TM9 and TM12 are involved in multimerization, as the LeuT was crystallised as dimer.

The extracellular gate consists of R30 and D404, which can form a salt bridge, and Y108 and F253, which allow access to the extracellular solution for the binding pocket. For substrate entering the gate must be opened. After entering of Na1, leucine and Na2, conformational changes occur and the gate closes. Further movements of TM1 and TM6 opens the intracellular gate consisting of R5 and D369, forming a salt bridge in the closed conformation. Substrate and ions are released and the transporter returns to the outward faced conformation for the next cycle. The stoichiometry for LeuT is sodium:chloride:substrate = 2:0:1. In contrast to the mammalian monoamine transporters, LeuT is chloride independent (Yamashita et al., 2005, Singh et al., 2007, Zhou et al., 2007).



Figures 5.5a - c: show the gates of LeuT: (a) intracellular view: R5 and D369 marked in green (b) side view: location (green circles), (c) from extracellular: R30 and D404 also marked in green.

## 5.2 Neurotransmittertransporters

During the neuronal transmission, neurotransmitter are released into the synaptic cleft to forward impulses from one neuron to the next one. It is important to remove the transmitter rapidly from the synaptic cleft to terminate signalling. This mechanism was first elucidated by G. Hertting and J. Axelrod in 1961. The transporter are located at the presynaptic membrane of the transmitter emitting neuron in the peri-synaptic region out of the synaptic cleft (Torres et al., 2003)

The mammalian NSSs are coded on the SLC6 gene. This gene consists of transporters for neurotransmitters and solutes, that carry selectively their respective substrate with the sodium gradient through the cell membrane. The degree of sequence identity in this group is between 40 % and 90 %. In Table 5.3 the members of the SLC6 gene are listed:

gene	swissprot code	abbrev.	substrate	localisation
SLC6A1	P50531	GAT1	gamma-aminobutyric acid transporter	brain
SLC6A2	P23975	NET	norepinephrine transporter	brain, lung
SLC6A3	Q01959	DAT	dopamine transporter	brain
SLC6A4	P31645	SERT	serotonin transporter	brain, platelets, placenta
SLC6A5	Q9Y345	GlyT2	glycine transporter	spinal cord
SLC6A6	P31641	TauT	taurine transporter	thyroid, placenta, retina
SLC6A7	Q99884	PROT	proline transporter	hippocampus, colon
SLC6A8	P48029	CT1	creatine transporter	kidney, brain, eye, testis, muscle
SLC6A9	P48067	GlyT1	glycine transporter	brain
SLC6A10				
SLC6A11	P48066	GAT4	gamma-aminobutyric acid transporter	fetal brain
SLC6A12	P48065		Betaine/GABA transporter	corpus striatum, kidney
SLC6A13	Q9NSD5	GAT2	gamma-aminobutyric acid transporter	liver
SLC6A14	Q9UN76	ATB0+	amino acid transporter	mammary gland
SLC6A15	Q9H2J7	NTT73	orphan Na/Cl dependent NSS	cerebellum, placenta
SLC6A16	Q9GZN6	NTT5	orphan Na/Cl dependent NSS	testis, epithelium
SLC6A17	Q9H1V8	NTT4	orphan Na/Cl dependent NSS	
SLC6A18	Q96N87	XTRP2	Na/Cl dependent transporter	kidney, colon
SLC6A19	Q695T7	B(0)AT1	neutral amino acid transporter	kidney, colon
SLC6A20	Q9NP91	SIT1	proline transporter	

Table 5.3: Members of SLC6 gene family, taken from [www.expasy.org](http://www.expasy.org)

### 5.2.1 Structure and Function

As mentioned above, the bacterial homologue of the mammalian NSS, LeuT, gives insights in their three dimensional structure. Hydropathy analysis reveal, that the transporters are organised to 12 TMDs with intracellular N- and C-termini. Previous studies that TM1, TM3, TM6 and TM8 participate in substrate translocation. The main differences between the bacterial and the mammalian transporters are located in EL2, that is about 25 amino acids longer and contains glycosylation sites and disulphide bonding domains, the of course different binding site and the chloride dependency of the SERT and DAT (Yamashita et al, 2005, Beuming et al., 2006). The glycosylation is necessary for transporter trafficking and function. Also the  $\text{NH}_3^+$  and  $\text{COO}^-$  termini are longer. LeuT forms dimers in the crystal, so it is supposed, that also SERT and DAT appear as dimers and tetramers (Chen et al., 2000; Rudnick, 2007).

The sequence identity of LeuT to SERT and DAT is at 21 and 20 %, respectively. It is a rather low, but the homology ranges from 40 % to 45 %. In the binding site, the amount of identical residues is about 50 % (Zhou et al., 2007, Beuming et al., 2007, Ravna et al, 2007). SERT and DAT share about 50 % of sequence identity (Rudnick et al., 1998; Chen et al., 2000).

### 5.2.2 The Serotonin Transporter, SERT

The primary sequence of the human SERT consists of 630 amino acids with a formula molecular weight of about 70.3 kDa ([www.expasy.org](http://www.expasy.org)). It is expressed across the plasma membrane of neurons and glia, and also on platelets and transports the proper substrate with high selectivity. The SERT is also targeted by other substrates, e.g. MDMA and indoles, and inhibitors, e.g. Cocaine.

Most of the corresponding residues in the binding site of SERT are identical to those in LeuT. A substantial difference is D98, where on this position in LeuT is G24. Serotonin does not have a carboxyl group, so that the acidic moiety of D98 coordinates with the amino group of 5HT. Mutations on this aspartatic acid leads to a loss of transporter function (Henry et al., 2003). Only five out of twelve residues interacting with leucine in the LeuT structure are mutated in the binding site of SERT. Three of them are changed to smaller residues in the hydrophobic part of the binding site, that is rooted to the bigger indole moiety of serotonin (Rudnick, 2007).

Because of the stoichiometry of transport action of SERT (see figure 5.6), it is supposed that only one  $\text{Na}^+$  is bound and co-transported with 5HT. The analysis of the conserved residues around the two  $\text{Na}^+$  in LeuT suggest, that also in SERT could be bound a second  $\text{Na}^+$ , that is not transported (Rudnick, 2007). Recent investigations on the  $\text{Cl}^-$  binding site in NSS members exhibited a binding site for  $\text{Cl}^-$  in NSS. The  $\text{Cl}^-$ -coordinating residues, serines, threonines and tyrosines, are conserved over the chloride dependent mammalian NSS but not in the  $\text{Cl}^-$  independent bacterial members LeuT, TnaT and Tyt1, where these residues are glutamates and aspartates. Mutations of LeuT (E290S) led to a  $\text{Cl}^-$  dependent transporter and analogue mutations in SERT (N368D) led to a  $\text{Cl}^-$  independent transporter. (Forrest et al., 2007; Zomot et al., 2007)

SERT co-transporters its substrate 5HT with  $\text{Na}^+$  and  $\text{Cl}^-$  through the membrane into the cytoplasm and antiports one  $\text{K}^+$  to the extracellular side in a neutral electrochemical manner. For transportation it undergoes a series conformational changes. The permeation pathway is located in TM1, TM3, TM6 and TM8 (Chen and Rudnick, 2000). The order of the binding steps is not evident. In the outward faced conformation, the first ion binds, that is supposed to fix the transporter in open state. After substrate recognition 5HT enters the binding site. The second ion causes a conformational change to the occluded state. A next conformational change opens the intracellular gate and releases the solutes. Unique for the NSS is the following step, where  $\text{K}^+$  binds from the cytoplasmic side and leads the transporter to the outward faced state to release it on the outside of the cell. The outward faced position is the preferred for SERT (Rudnick, 2007).

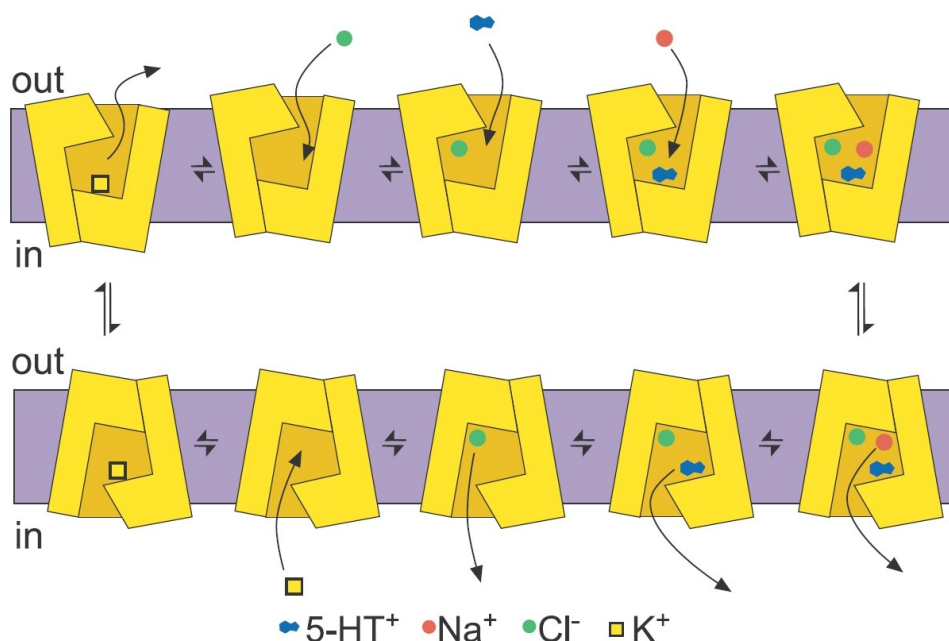


Figure 5.6: transport action of SERT and stoichiometry from Rudnick, 2007. The stoichiometry of the transport is  $5\text{HT}^+:\text{Na}^+:\text{Cl}^- = 1:1:1$  inward and  $\text{K}^+ = 1$  outward. The net charge transport is 0, in contrast to the other members of the NSS.



### 5.2.3 The Dopamine Transporter, DAT

DAT in its primary sequence is a little smaller than SERT and consists of 620 amino acids with a molecular weight of 68,5 kDa based on the primary amino acid sequence ([www.expasy.org](http://www.expasy.org)). Located on the terminals of dopaminergic neurons, it transports the substrate dopamine and also norepinephrine. Among others the DAT is targeted by cocaine and MPD as inhibitor and substrates like amphetamines and the neurotoxin MPP+ (Torres et al., 2003).

The sequence identity to LeuT is a little lower, than the sequence identity from SERT to LeuT (see table 6.4). Crucial for substrate binding and translocation is D79, that coordinates the amine group from dopamine and the Na<sup>+</sup>. This residue corresponds to LeuT G24 and SERT D98. Mutations on this position lead to loss of function. Also in DAT, the region of the highest density of conserved residues is found in the domains participating in substrate translocation. As also the SERT, five out of twelve residues are mutated in the binding site of DAT. And similarly four of them are changed to smaller residues, due to the larger moiety of dopamine in the hydrophobic part of the binding pocket, depicted in table 6.5 (Chen and Reith, 2000).

The transport stoichiometry of DAT is distinct to LeuT, because only one Na<sup>+</sup> is transported and also distinct to SERT, because there is no antiport of K<sup>+</sup>. Dopamine is co-transported with Na<sup>+</sup> and Cl<sup>-</sup>. The sodium gradient is the driving force for the transport action. In contrast to SERT, the transport action is not electrochemical neutral: dopamine<sup>+</sup>:Na<sup>+</sup>:Cl<sup>-</sup> = 1:1:1, so one positive charge is transported into the cell. The mechanism of substrate translocation is supposed to be similar in NSS family. As mentioned above, Na<sup>+</sup> enters the transporter for keeping the extracellular gate open. After entering of dopamine and chloride, conformational changes occur and close the gate between D79 and Y176. When these residues interact, further conformational changes lead to an opening of the intracellular gate and release of substrate and ions (Torres et al., 2003; Chen and Reith, 2000)

## 6 Materials and Methods

As previously mentioned, the NSS terminate neuronal transmission by reuptake in the presynaptic neuron. They are also a target for CNS-modulating drugs like SSRIs and TCAs and as well for abused drugs as amphetamines and cocaine. The mode of action can be distinguished between substrate (5HT), inhibitor (SSRI) and substrate type releaser (MDMA). Due to the lack of a crystal structure of the NSS, we built homology models of SERT, DAT and NET for running docking simulations. In this work we investigated the protein ligand interactions and the binding modes of the substrates and inhibitors of the SERT and the DAT. For all simulations we used the software package Molecular Operating Environment, Version 2007.09.

### 6.1 Molecular Operating Environment

The Molecular Operating Environment (MOE), developed by the Chemical Computing Group (CCG), Canada ([www.chemcomp.com](http://www.chemcomp.com)) is a software package of applications for drug discovery, used by academic and industrial researchers. This package contains multiple features with a graphical user interface:

- Cheminformatics and QSAR: descriptors, similarity search, high throughput search, QSAR-predictive models
- Pharmacophore discovery: pharmacophore search, conformation analysis
- Molecular modelling and simulations: molecule builder, data transfer, molecular dynamics and mechanics
- Protein modelling and bioinformatics: alignment, fold and rotamer libraries, mutations
- Structure-based design: site detection, surface imaging, docking, protein-ligand-interaction-fingerprint
- Methods development and deployment: scientific vector language (SVL)

([www.chemcomp.com](http://www.chemcomp.com))

In the present study we used the protein homology modelling and molecular modelling and simulation features, i.e. docking, which will be described in the following.

## 6.2 Homology Modelling

We started the procedure of creating homology models by using the crystal structure of LeuT as template and the comprehensive alignment published in by Beuming et al., 2006 (<http://icb.med.cornell.edu/trac>). The methodology of homology modelling with MOE is that conserved residues are copied identically over the coordinates of the template. For the differing residues the backbone coordinates are used to place the side chains, based on the rotamer library. For the calculations the Amber99 forcefield was used.

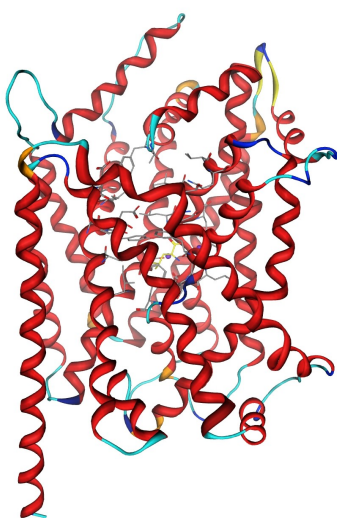


Figure 6.1a: This picture shows our template with bound leucine depicted as surface in the centre of the protein.

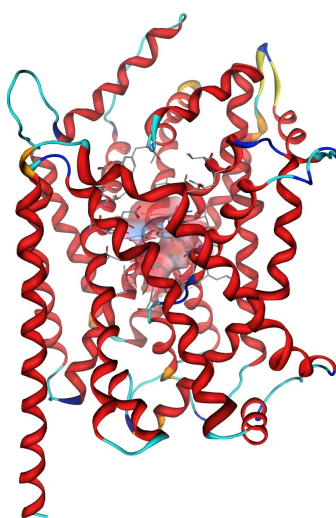


Figure 6.1b: The binding site is

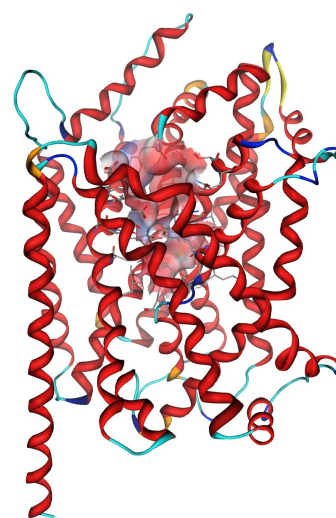


Figure 6.1c: Both the binding site and the external vestibule are displayed as surface.

Figures 6.1a - c are shown in side view

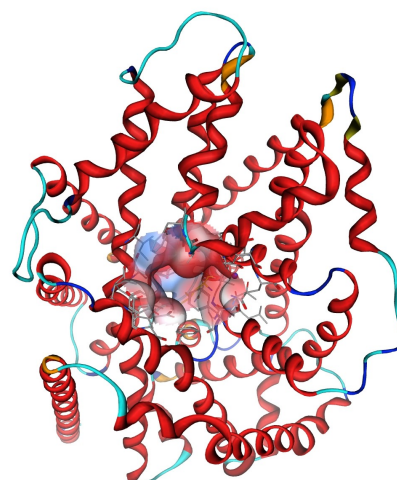
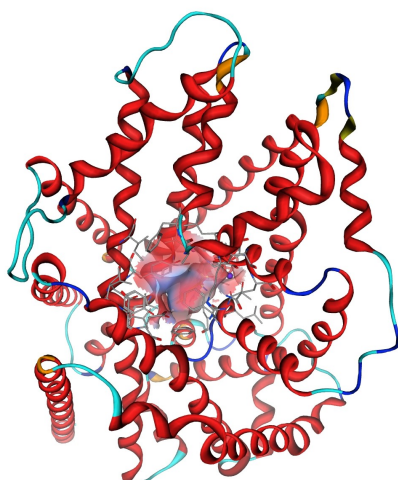
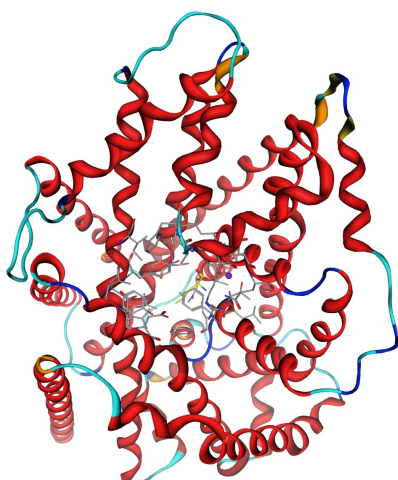


Figure 6.2a-c: The pictures show our template LeuT in the same manner as above, but the transporter rotated 90°. This is the view from the extracellular space (top view).

The precondition of an accurate homology model is a reliable template. Very important criteria for the template are e.g. sequence identity, homology, functional similarity and the resolution of the X-ray structure. The crystal of the LeuT was published with a medium resolution of 1.65 Å. It shows the transporter with bound leucine in the substrate occluded state, two Na<sup>+</sup> ions in the binding site and one Cl<sup>-</sup> ion on the extracellular part of the protein. The sequence identity of about 20 % between LeuT and the monoamine transporters is very low (Yamashita et al., 2005).

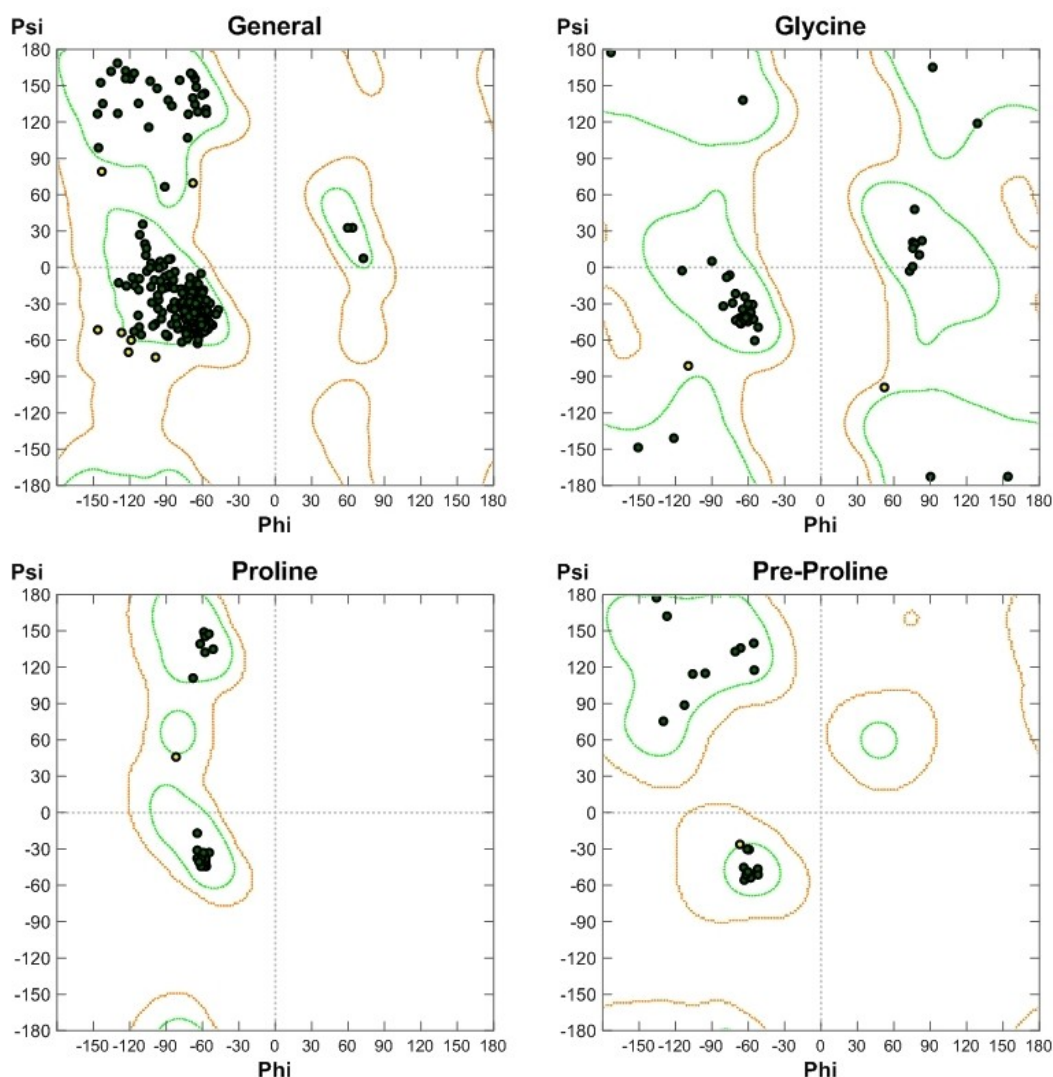


Figure 6.3: Ramachandran Plot of LeuT; the green areas depict the protein core regions and the red areas show allowed regions concerning the  $\Phi$  and  $\Psi$  angles. Glycine and proline have diverse allowed regions because of their different structure.

We started MOE homology modelling with aligning the sequences of the template and the sequences of SERT, DAT and NET manually in the sequence editor of MOE, according to Beuming et al., 2006. The sequences of human DAT, NET and SERT were obtained from the swiss-prot database as fasta-file ([www.expasy.org](http://www.expasy.org)).



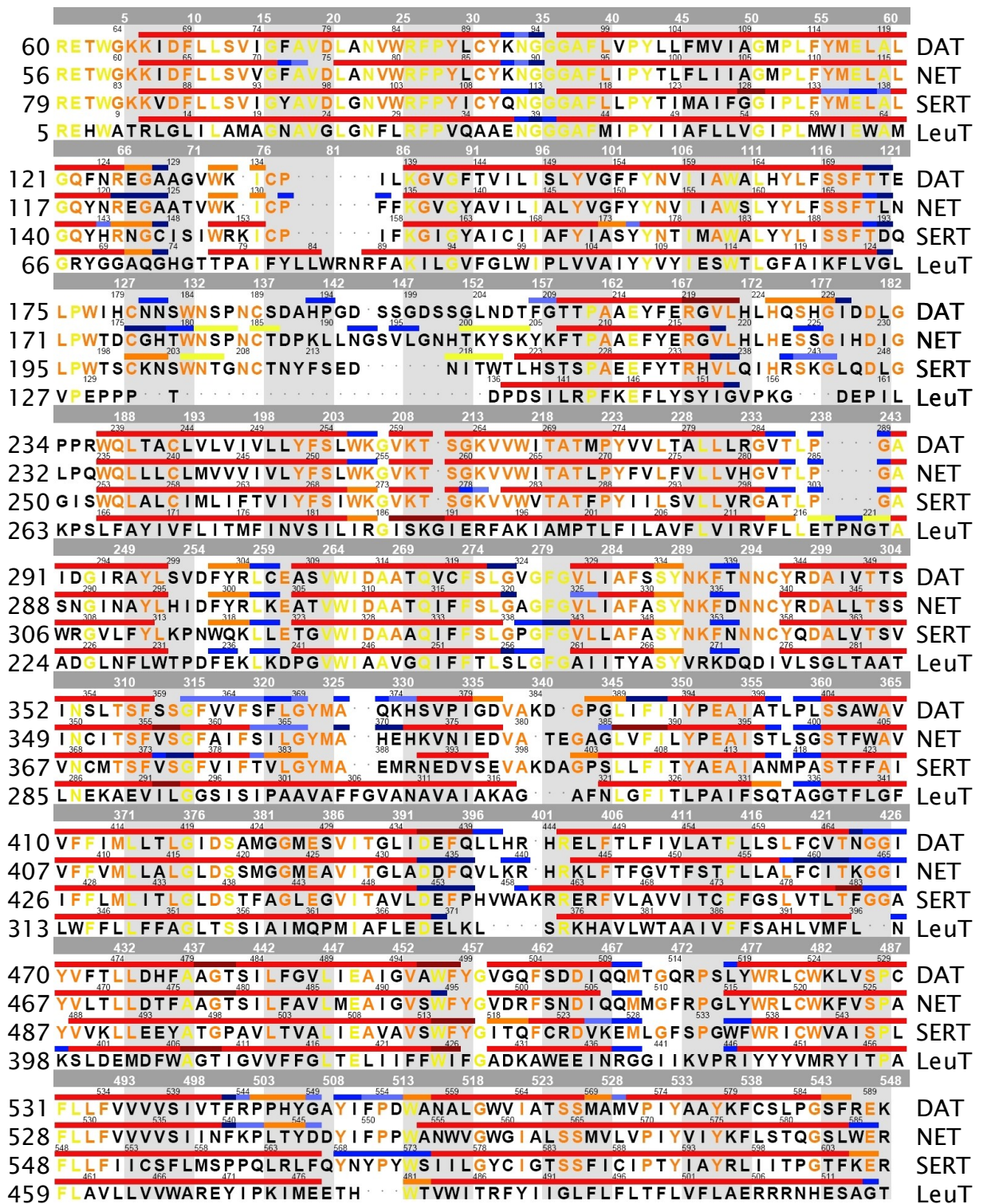


Figure 6.4: Alignment of the template with the sequences of DAT, NET and SERT is based on Beuming et al., 2006. Colour code: yellow = identical residues; orange = identical amino acids in DAT, NET and SERT; red lines = transmembrane helices, blue lines = loops, yellow lines = beta sheets.

After removing water, b-octoglucoside and the Na<sup>+</sup> and Cl<sup>-</sup> ions from the crystal structure, we started the homology modelling. The software calculated the models by using a Boltzmann-weighted randomized modelling procedure and databases for sidechain orientation and loop folding. Scoring and refinement was based on electrostatic solvation energy, calculated with generalized Born/Volume Integral (GB/VI) methodology. The

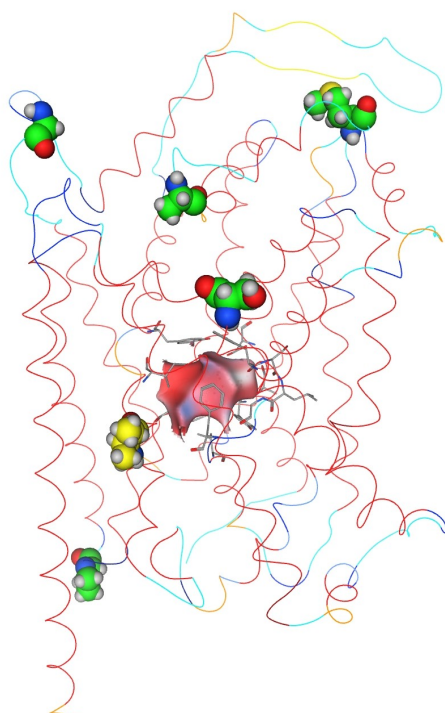
duration of the modelling procedure was approximately 6 hours on a Intel Pentium Core 2 Duo with 2.2 GHz and 2 GB RAM.

## 6.2.1 The Model of the Serotonin Transporter

During the creation of the models, MOE generated ten intermediates. The final model is based on the best scored intermediate of GB/VI. Both termini have been removed from the sequence. As depicted in table 6.1, only few outliers have been reported.

name	GB/VI	E sol	E ele	E vdw	Atom Clashes	BB Bond Outliers	BB Angle Outliers	BB Torsion Outliers	Rotamer Outliers
Model #1	-15654,03	-1270,16	-1428,86	-1977,50	28	1	19	20	3
Model #2	-15665,99	-1275,49	-1426,12	-1912,11	29	2	27	27	1
Model #3	-15654,58	-1163,05	-1459,46	-2030,58	18	0	18	21	4
Model #4	-15630,86	-1308,58	-1411,74	-1927,39	24	1	15	28	3
Model #5	-15635,80	-1297,02	-1409,80	-1983,55	25	0	18	33	5
Model #6	-15662,34	-1152,38	-1442,01	-1947,21	23	0	18	26	4
Model #7	-15634,03	-1292,02	-1399,94	-1981,78	24	1	21	17	4
Model #8	-15612,79	-1193,59	-1444,77	-1971,49	22	1	20	28	6
Model #9	-15639,60	-1260,46	-1371,41	-2063,76	19	1	15	25	2
Model #10	-15607,00	-1230,59	-1413,81	-2012,52	26	1	21	21	5
Final model	-15516,69	-840,41	-1338,23	-2734,49	0	0	2	6	1

Table 6.1: Parameters of the final model and the intermediates. This table shows calculated energetic values and putative outliers in back bone (BB) bond lengths, BB angles, BB torsions and rotamers. As well it shows atom clashes, which are absent in the final model.



Domain	Residue	Outlier	Colour
TM3	I167	Rotamer	yellow
EL2	G244	Ramachandran plot	green
EL4	M386	Ramachandran plot	green
EL4	A401	Ramachandran plot	green
IL4	P455	Ramachandran plot	green
TM11	S559	Ramachandran plot	green

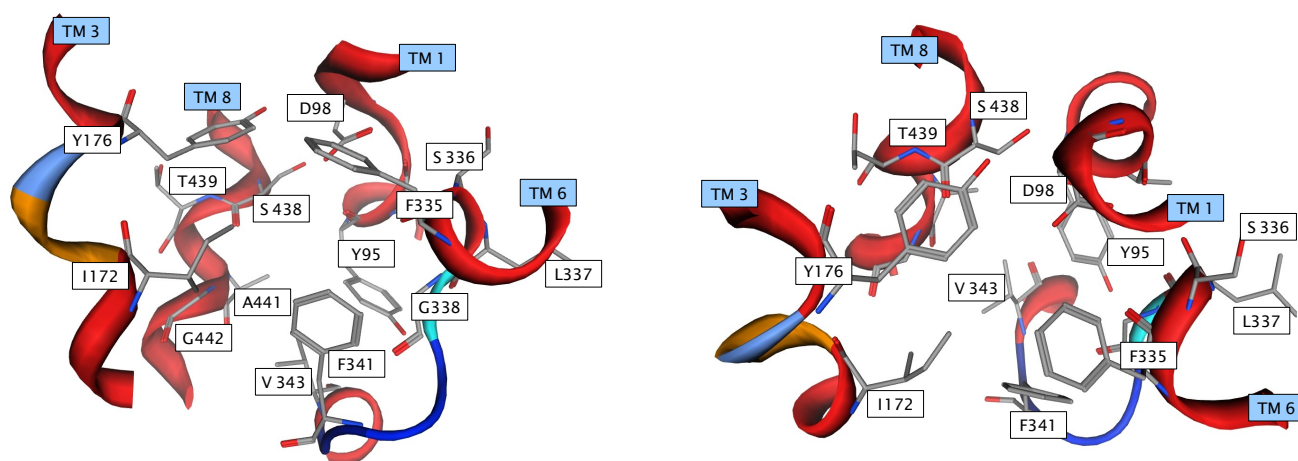
Table 6.2: Involved domains where the residues have been reported as outliers. The colour code refers to Figure 6.5

Figure 6.5: The protein backbone is shown as line for visualising the location of the outliers. In the centre the binding site is depicted as surface.



We evaluated our model first by inspecting the binding site. The TMDs involved in the cavity show a sequence identity of more than 40 % between the LeuT and the SERT accordingly the DAT, see table 6.5. This suggests a highly conserved structural feature of the NSS for substrate translocation (Yamashita et al, 2005; Zhou et al, 2007). The corresponding residues which are directly interacting with leucine we account as binding site in SERT since there is a high degree of identity. In our model, we also found the same residues within a radius of 4.5 Å from the centre of the cavity by using the site finder.

Also the chloride and the sodium binding sites are well defined (Forrest et al., Zomot et al). The extracellular gate formed by the residues R104 and E493 as well as Y176 and D98 which induces the translocation are in proximity (Rudnick, 2007). We further looked for engineered Zn and Cd binding sites, which confirm the findings on sterical proximity (Henry et al., 2006; White et al., 2006; Mitchell et al., 2004). Since our major focus was on the binding site, we did not consider the disulphide bond in EL2. It is supposed due to the fact that C200 and C209 are not accessible with MTS reagents (Kamdar et al., 2001).



Figures 6.6 a, b show the binding site of SERT. The residues within a radius of 3 Å from the centre are depicted. left = side view; right = top view.

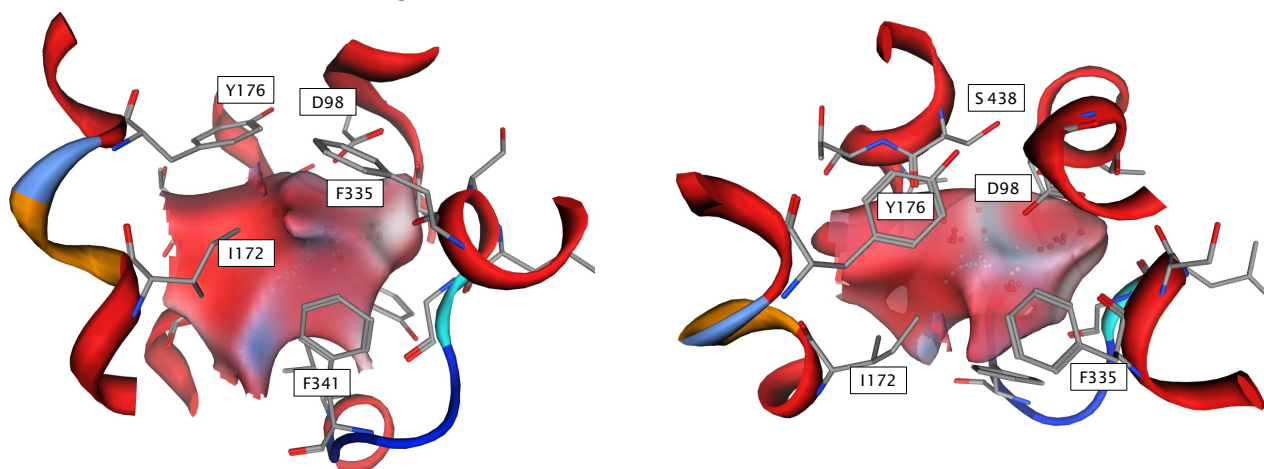


Figure 6.7 a, b: The binding site is depicted as surface: red = lipophilic cavity, blue = hydrophilic interactions, white = H-bond contacts; left = side view; right = top view. Concerning the labelling, the following figures from the binding site have been taken from the same perspective.

The Ramachandran plot reports 5 outliers concerning the two angles  $\Phi$  and  $\Psi$ . None of them is located near the binding site. Our model contains 529 amino acids due to the eliminated N and C terminus. Five residues are located outside of the allowed regions in the  $\Phi$ - $\Psi$ -plot. The amount of outliers is lower than 2 %, which assigns a further quality criterion of our model.

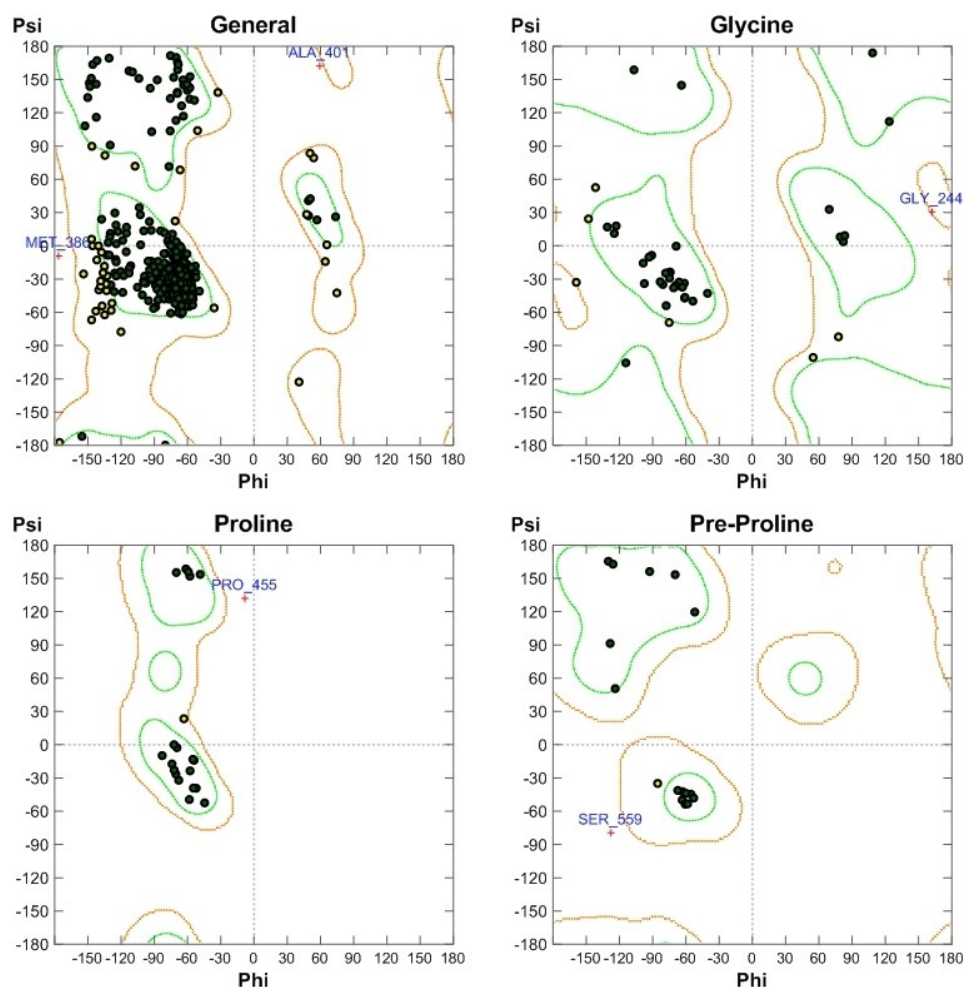


Figure 6.8: The Ramachandran Plot of the SERT. The residues, which are marked in blue are located in not allowed regions due to their angles. None of them is in proximity of the binding site. These uncertainties occur in change-overs of the secondary structure, e.g. from helix to loop.

## 6.2.2 The Model of the Dopamine Transporter

As well as the final model of the SERT also the model of DAT shows few outliers concerning the rotamers, torsions and angles. Here again no atom clashes are reported. Both termini have again been removed from the sequence for modelling.

name	GB/VI	E sol	E ele	E vdw	Atom Clashes	BB Bond Outliers	BB Angle Outliers	BB Torsion Outliers	Rotamer Outliers
Model #1	-14833,06	-1285,35	-1289,16	-2132,15	17	2	14	21	5
Model #2	-14835,38	-1314,88	-1281,15	-2230,52	14	0	16	24	3
Model #3	-14808,46	-1072,73	-1284,67	-2109,99	14	0	9	24	4
Model #4	-14814,07	-1273,75	-1285,3	-2177,41	13	0	14	28	4
Model #5	-14830,31	-1265,66	-1273,85	-2142,26	14	1	13	23	2
Model #6	-14811,53	-1141,37	-1296,34	-2147,61	21	0	21	29	2
Model #7	-14778,66	-1268,79	-1262,96	-2154,22	18	2	13	26	2
Model #8	-14814,91	-1161,11	-1282,55	-2083,73	22	0	12	25	5
Model #9	-14793,96	-1217,63	-1240,55	-2142,9	17	1	18	24	4
Model #10	-14820,6	-1161,4	-1285,8	-2116,04	17	0	18	22	3
Final model	-14665,65	-853,33	-1202,81	-2775,02	0	0	2	9	4

Table 6.3: Parameters of the final model and the intermediates (BB = back bone). As table 8.1, it shows calculated energetic values and putative outliers in the mentioned categories. Also the model of the DAT shows good parameters.

The validation of the model correlates to those in SERT. For this transporter we primarily focused on the binding site to identify differences in the binding mode of MPD between DAT and SERT. Both  $\text{Na}^+$  binding sites and the  $\text{Cl}^-$  binding site, which are located near the substrate binding site, are similarly well detectable in DAT.

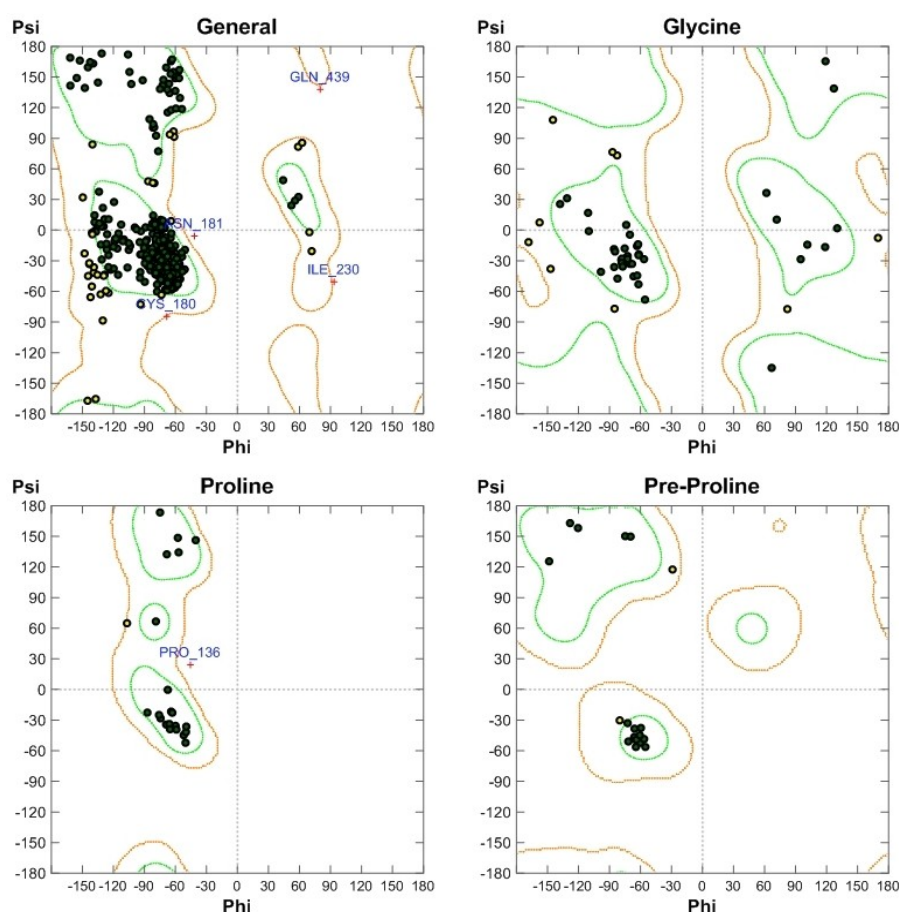


Figure 6.9: Ramachandran Plot of DAT. This model contains 531 amino acids because we left out both termini for modelling. As well as in SERT five residues in DAT are located outside of the allowed regions, also marked in blue. The amount of outliers is again in DAT lower than 1 % which denotes a good model.

Following our aim to elucidate the binding modes of substrates and inhibitors in SERT and to explain the differences in the binding mode of MPD in DAT and in SERT, we left out the evaluation of the homology model of NET. A short view on the binding site shows similar orientations of the residues in this region of NET, which suggests also a reliable model.

### 6.2.3 Comparison of the Models

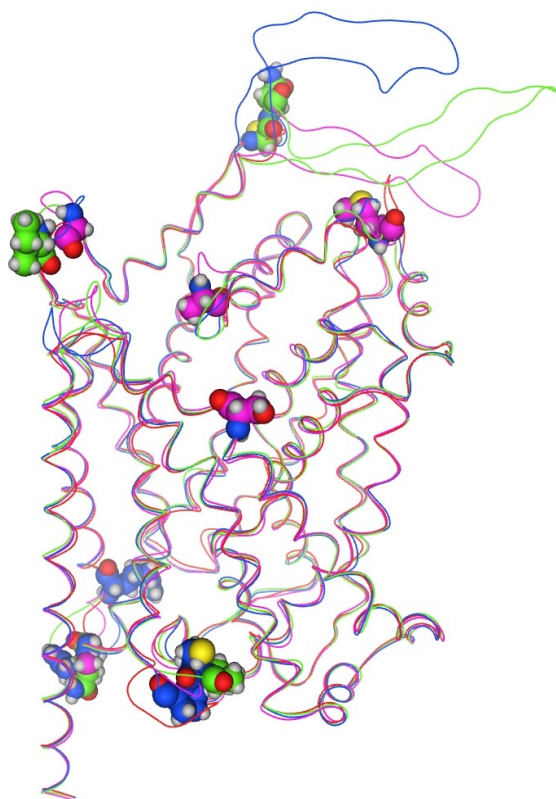


Figure 6.10a: Overlay of the models of SERT, DAT and NET. The backbone is depicted as line. The residues which are outliers are displayed in spacefill mode. None of them is in proximity of the binding site. They rather are located at structural changeovers, e.g. loop-helix. Colour code: SERT = pink, DAT = green, NET = blue.

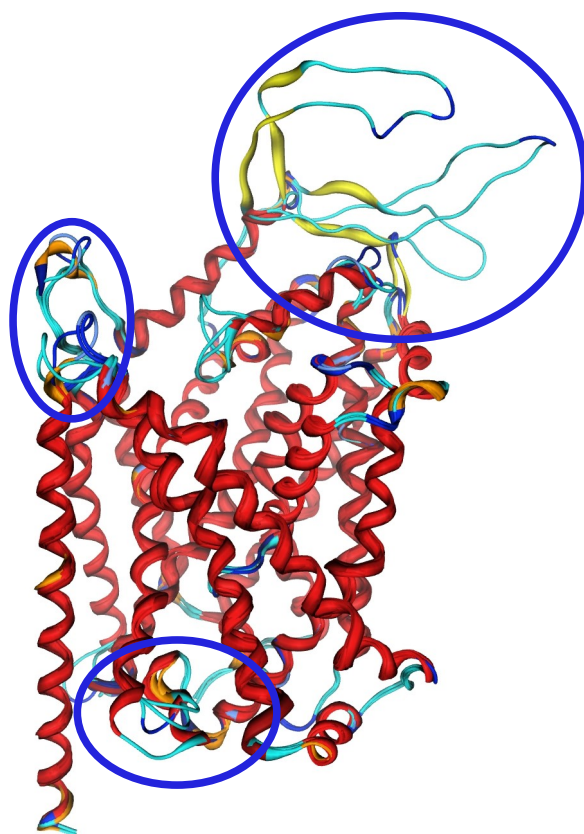


Figure 6.10b: Overlay of the template LeuT with the three models depicted in tube mode. This picture shows the differences due to the loop modelling which effects large RMSD over the whole proteins (see table 6.4).

Table 6.4: The deviations between the template LeuT and the models, measured over the whole structure. Due to a sequence identity of about 20 % between the template and the model, an RMSD of lower than 1.5 Å over the whole structure is very well. The increasing deviation between the three models is due to the differences in the modelled loops but mainly due to EL2 (blue circles in Fig 6.10b)

RMSD	LeuT	DAT	NET	SERT
LeuT		1,2621	1,3812	1,2699
DAT	1,2621		3,4352	2,5364
NET	1,3812	3,4352		3,9220
SERT	1,2699	2,5364	3,9220	



## 6.3 Molecular Modelling and Simulation

The main part of the present study was running docking of selected ligands in the putative binding site of the homology models. Hereafter we describe the procedure of docking into the models of SERT and DAT and ranking the obtained docking poses. Based on the crystal structure of LeuT, we elaborated a method to reduce the amount of poses after redocking for ending up in the pose of leucine near the crystal structure of LeuT. We considered geometric and energetic parameters for this pose shrinking procedure.

Important parameters and program features for docking and ranking are

- Merck forcefield: In contrast to Amber or Charmm, this forcefield is parameterised for small organic ligands. We used this forcefield during the docking run, because the protein was held rigid.
- Dielectric constant: This parameter was set to 4 for the interior of the protein and to 80 for the exterior medium. A dielectric constant of 1 matches vacuum, 2 - 8 are used for protein medium and 80 equals water (Ermondi et al., 2006).
- Protein Ligand Interaction Fingerprint (PLIF): This tool is a method for summarising interactions between the ligand and the protein in a fingerprint scheme. The interactions of the residues with the ligand, such as hydrogen bonding or ionic interactions are divided in single bits
- Placement: Alpha triangle places the molecule by superposing ligand atom triplets and triplet points in the binding site, which are alpha sphere centers. For the determination of the pose a random triplet of the alpha sphere and a random triplet of the ligand atoms with a random conformation are used.
- Scoring: We used London  $\Delta G$  as scoring function. It estimates the free energy of binding between the pose and the protein.
- Refinement: It is an energy minimisation step during the docking run. Only the ligand is considered while the protein is held rigid.

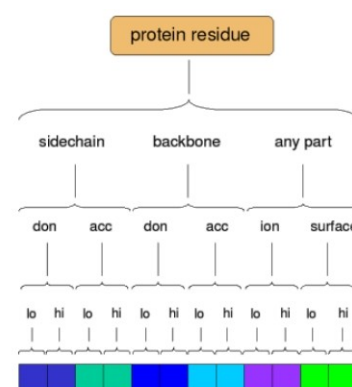


Figure 6.11: Splitting of the interactions between the residue and the ligand in the PLIF

### 6.3.1 Ligand Preparation

The first step before docking was creating the ligands with the molecule builder integrated in MOE. After energy minimisation a conformational database was generated by molecular dynamic simulation. In this simulation the molecules are heated up to 2000 °K for 50 ps in the Merck forcefield (MMFF94x) and under vacuum conditions (dielectric constant = 1). The resulting conformers have been written in a database which was energy minimised. We selected the 200 most different conformations on basis of the potential energy of the ligand and the RMSD to the conformation with the lowest potential energy (Ermondi et al., 2004). For redocking of leucine into the crystal structure we extracted the ligand from the .pdb-file 2A65 and created the conformer database in a similar way.

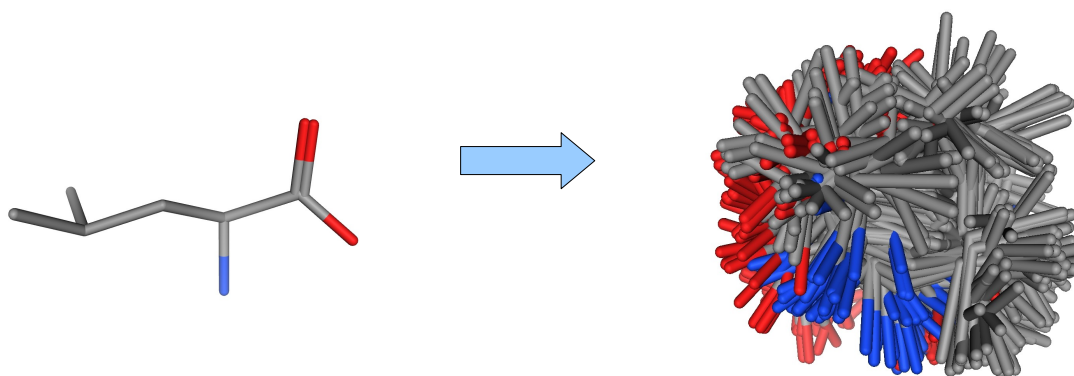


Figure 6.12: On the structure of leucine extracted from the crystal structure (left) the MD simulation was applied, which resulted in 200 conformations. The overlay of all these structures is represented on the right side.

### 6.3.2 Defining the Binding Site

For the homology models, we defined the binding site by using the built site finder, implemented in MOE. It calculates possible active regions in a protein considering the three dimensional coordinates based on alpha shapes. The algorithm also distinguishes between hydrophilic and lipophilic regions in the pocket (see figure 6.14). Each residue within a radius of 6 Å has been part of the binding site during the docking run. These residues are the corresponding amino acids in the binding site of LeuT, see figure 6.14 (Yamashita et al., 2005; Rudnick, 2007; Beuming et al., 2006; Huang et al. 2007). For redocking the binding site was defined by using the bound ligand. Only the directly interacting residues within a radius of 3 Å have been considered.



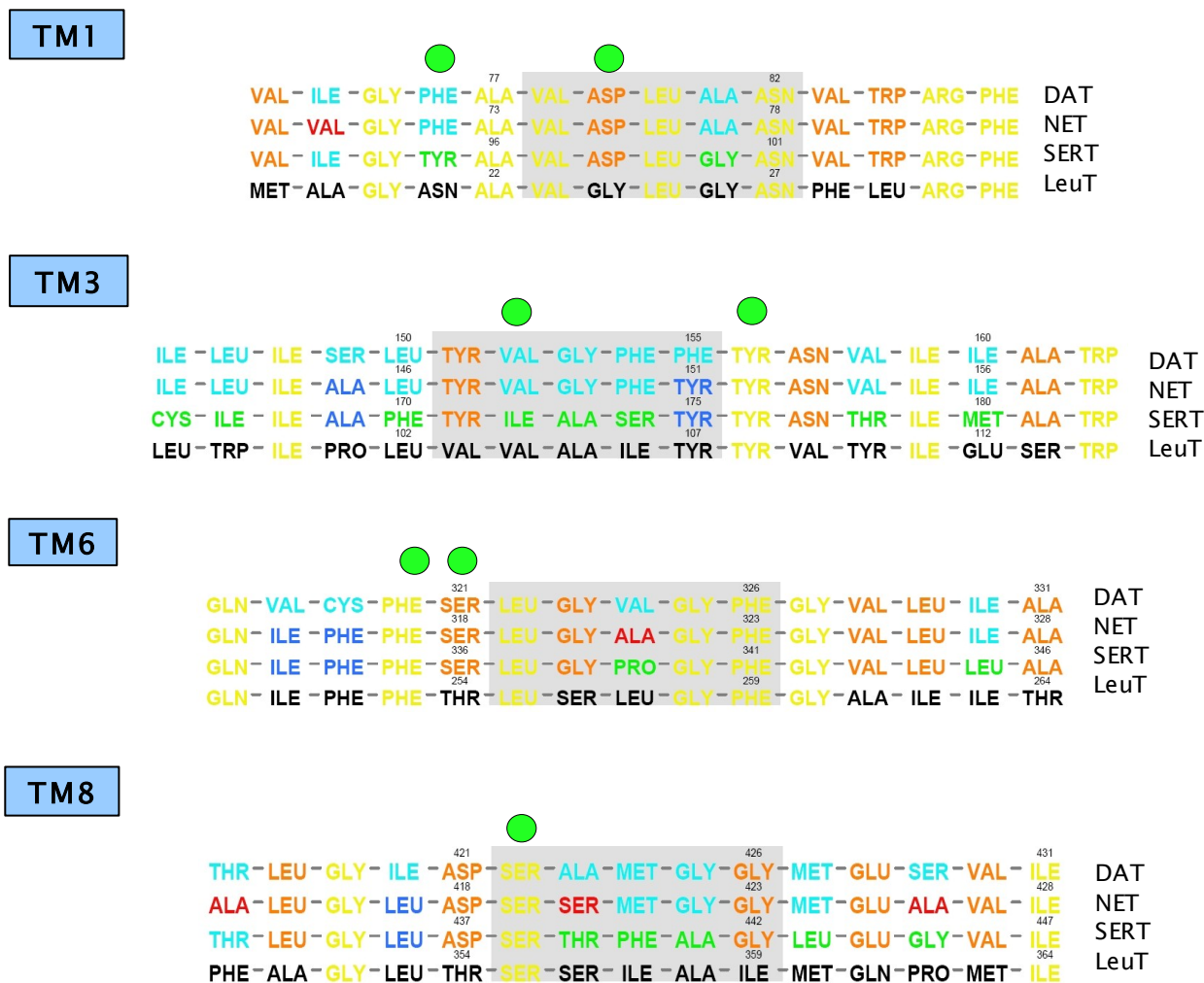


Figure 6.13: This picture shows the residues, which surround the binding site in the template as well as in the models. The depicted sequences are for each the centre of the named transmembrane domains. Colour code: yellow = overall sequence identity, orange = sequence identity between the monoamine transporter, light blue = residual amino acids in DAT and the identical residues in either the NET or the SERT, red = residue only in NET, green = residue only in SERT. Residues important for ligand interactions are marked with a green circle.

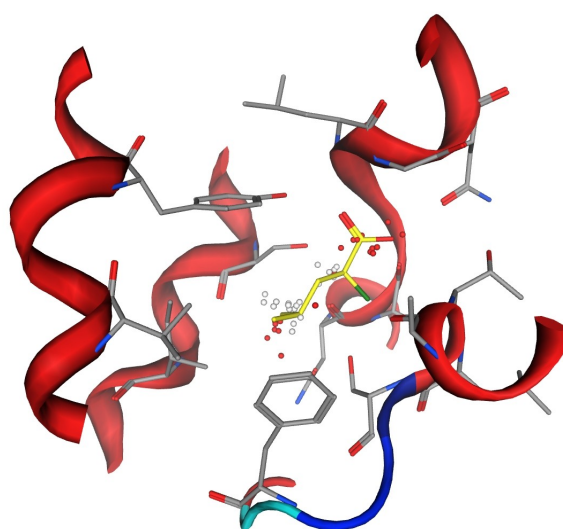


Figure 6.14: The binding site of LeuT with bound leucine (yellow). The dots are the dummy atoms: white = lipophilic regions, red = hydrophilic centres

Domain	LeuT	SERT	DAT
TM1	N21	Y95	F76
	A22	A96	A77
	G24	D98	D79
	L25	L99	L80
	G26	G100	A81
TM3	N27	N101	N82
	V104	I172	V152
	Y108	Y176	Y156
TM6	F253	F335	F320
	T254	S336	S321
	L255	L337	L322
	S256	G338	G323
TM8	F259	F341	F326
	A351	L434	L418
	T354	D437	D421
	S355	S438	S422
	I359	G442	G426

Table 6.5: Alignment of residues lining the binding sites. This shows a sequence identity of about 47 %.

green=identity of all transporters  
 yellow=two transporters share identical residues  
 red=individual residues.

### 6.3.3 Running Docking

We started the docking runs with redocking of leucine in the crystal structure of LeuT for evaluating the reliability of our docking program. The settings were adapted to the Merck forcefield (MMFF94x), an internal dielectric constant of 4 and an external dielectric constant of 80. Since MOE 2007.09 provides an energy minimisation step and two scoring functions during the docking run, the processing time is rather high. We ended up in a duration between 6 and 8 hours per ligand on an AMD Athlon X2 with 2.2 GHz and 2 GB of RAM.

The first analysis of the redocking results was measurement of the deviation between the best pose and the crystallised leucine by RMSD, see figure 8.15 and tables 8.5 a and b. Due to the increase in the amount of crystal structures the docking programs are also improving and tested in benchmarks. A threshold of 2 Å RMSD from the crystallised ligand is well accepted. Referring to David et al., 2005 the docking algorithm of MOE 2007.09 performed very well.

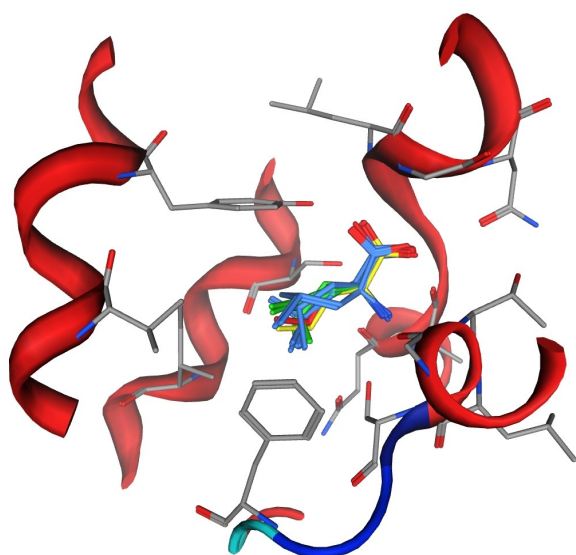


Figure 6.15: The picture shows those poses, with the lowest RMSD compared to the template. In table 6.6b the RMSD values are shown.

best pose 0.3144 Å  
highest deviation 4.6042 Å

poses	number	%
all	1993	100,00
< 1 Å RMSD	661	33,17
< 1.65 Å RMSD	1007	50,53
< 2 Å RMSD	1298	65,13

Table 6.6a: These tables show the results of redocking. Best pose is those with the lowest deviation from the crystallised ligand.

colour	deviation from template	number	%
yellow	Template, 0 Å	0	0
red	0.3144 < 0.3590 Å	122	6,12
green	0.5525 < 0.6244 Å	141	7,07
blue	0.7061 < 0.8196 Å	142	7,12
	sum	405	20,32

Table 6.6b: Referring to figure 6.15, the deviations of the poses after redocking.

Due to the redocking results we elaborated a method to find the pose with the lowest deviation by using calculated descriptors. Following we present the description of our method for ranking and selecting the poses in SERT and DAT based on the analysis of the redocking of leucine into LeuT.

### 6.3.4 Ranking the Poses

In the literature, amounts of data are available concerning mutagenesis experiments on the SERT and the DAT. In the binding site few very important residues are crucial for substrate binding and transporter function, see table 8.5. Mutation of one of these residues, either ligand binding or transporter function or both is reduced. These data reveal important interactions between the ligand and the protein. The exact binding mode of substrate and inhibitors in the DAT and the SERT is still a topic of research. Recent papers give suggestions on the binding mode of serotonin, dopamine, cocaine and CFT (Beuming et al., 2008 revealed pose of dopamine and CFT in DAT similar to our findings in SERT; Celik et al., 2008 published a model of the SERT with bound ligand similar to our top ranked poses; Huang et al., 2007 revealed the binding mode and interactions of dopamine in DAT; Kaufmann et al., 2008 in contrast to Celik et al. and our poses a different orientation of 5HT in the binding site is suggested; Xhaard et al., 2008 suggests differing binding modes of the ligands in the NSS).

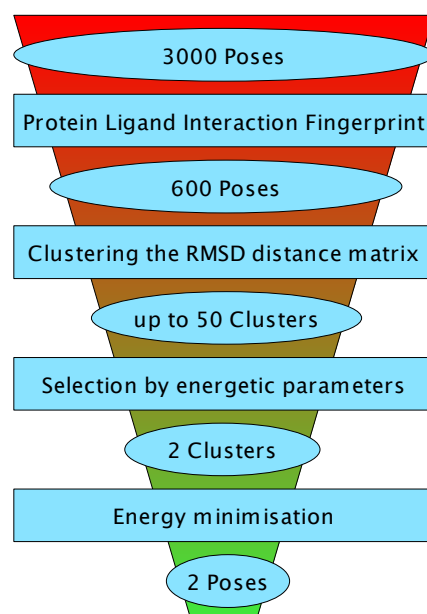
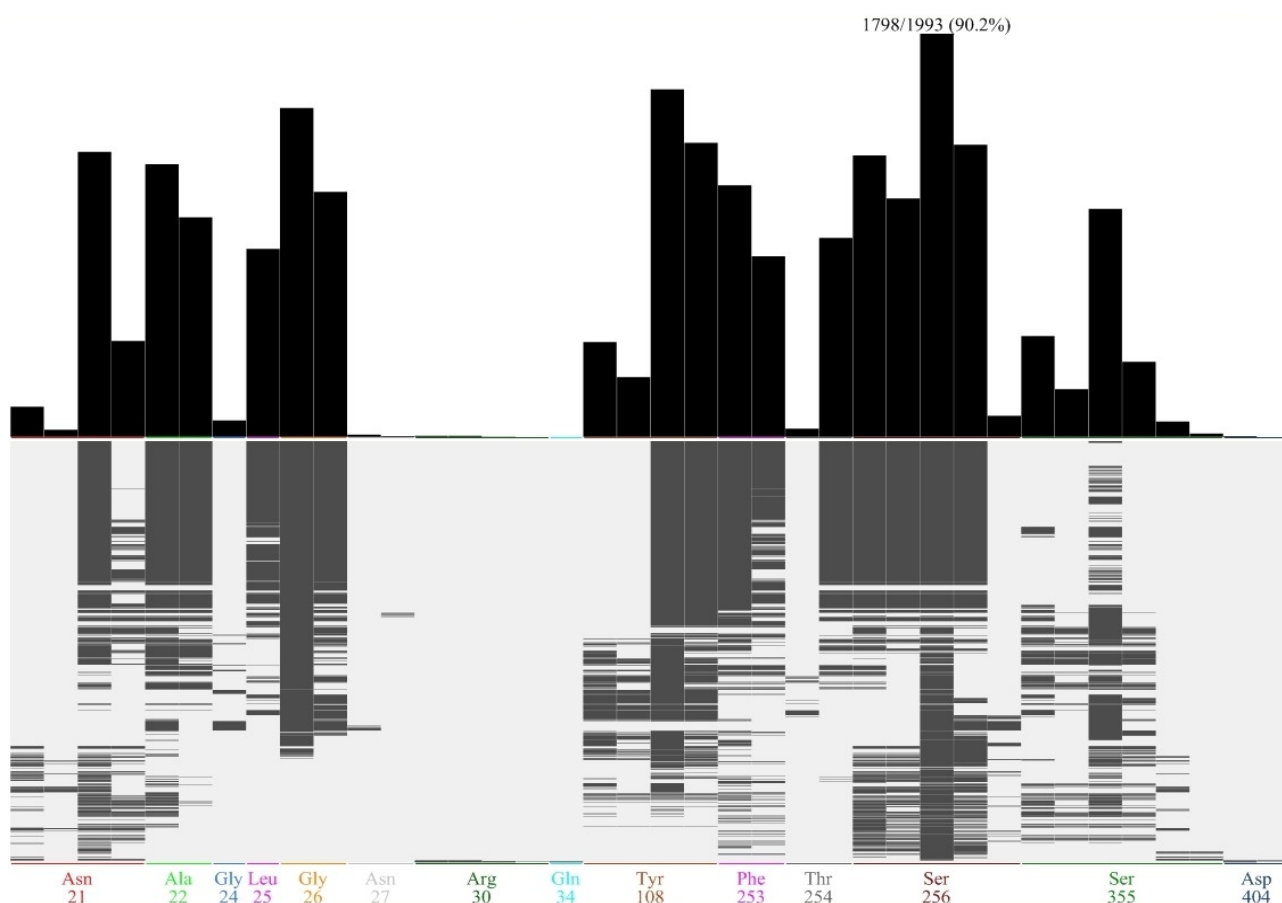


Figure 6.16: Procedure of pose shrinking

In contrast to the high throughput screening of libraries for hit finding, we used docking to elucidate the binding mode of substrates and inhibitors. It is a quite challenging topic since small ligand docking is not capable to find one single and correct pose by using scoring functions or calculated energetic parameters. As mentioned above the sequence identity between the template and the query protein is rather low, docking into such a model has an inherent error and more than one solution. After the docking run we obtained up to 3000 poses for each ligand. In order to select plausible poses, we followed a multi-step procedure, based on geometrical and energetic parameters which describe the ligand or the complex. The first step was creating the protein ligand interaction fingerprint (PLIF). It provides separation of poses by selecting interactions with defined residues. Our first selection criterion was the occurrence of interactions with residues in the binding site. The residues with the highest frequency in the PLIF are Y95, D98, F335 and S438, which have been identified in other docking studies (Celik et al., 2008), and shown to be important, based on mutagenesis data (Henry et al, 2003 and 2006)



Figures 6.17 a and b: The visualisation of the protein ligand interaction fingerprint of the redocking results. The ordinate shows the amount of 1993 docked poses. On the abscissa the interacting residues are depicted, splitted into interaction bits (see Fig. 6.11). The figure above shows the sum of hits per residue. Below, the interaction pattern is dissipated to the single pose.

The selection per PLIF reduced the amount of poses up to 700. For these we built a root mean square deviation (RMSD)-distance-matrix to consider the three dimensional orientations of the poses due to the fact that poses with similar spatial orientation (low RMSD) can have strongly deflected energetic parameters. We calculated the distances between all of them and obtained a  $n \times n$  matrix, which was our input database for the clustering by the statistical Ward-method. The Ward methodology is a hierarchical agglomerative clustering method on basis of the distances between individuals. Depending on the dissimilarity of the generated poses, the threshold was set to 10, 20 and 50 respectively. Each docking pose has been assigned into a cluster of similar poses.

$$RMSD = \sqrt{\frac{\sum_{i=1}^n (a_{ix} - b_{ix})^2 + (a_{iy} - b_{iy})^2 + (a_{iz} - b_{iz})^2}{n}}$$

a, b: compared atoms  
x, y, z: absolute  
coordinate  
i: molecule  
n: number of  
molecules

Formula 6.1: Root Mean Square Deviation (RMSD)

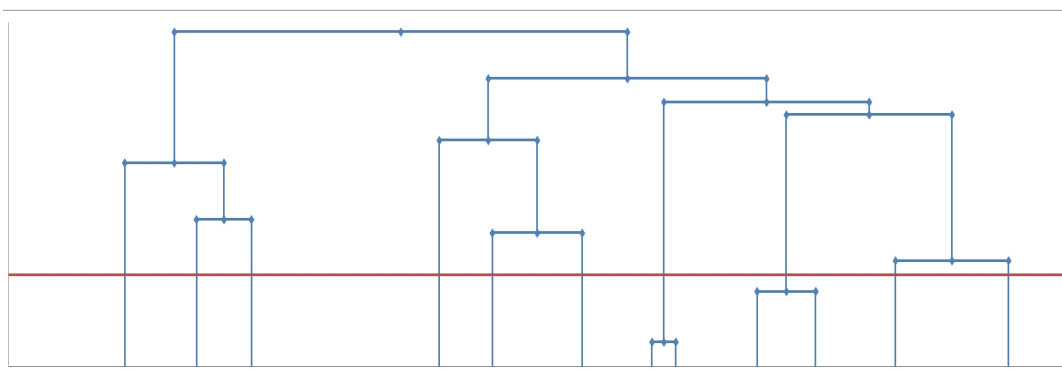


Figure 6.18: Cluster of leucine poses in LeuT; x-axis: number of poses, y-axis: degree of dissimilarity; red line: threshold of 10 clusters.

The further criteria were energetic parameter, calculated during and after docking. Based on the analysis of redocking, we considered the following descriptors:

- Potential energy of the ligand ( $E$ ), which characterises only the ligand. The most favourable redocking poses have been found with this descriptor (cluster 3,  $\text{RMSD} < 0.4 \text{ \AA}$ ).
- Calculated  $\text{pK}_i$  of the complex ( $\text{dock\_pK}_i$ ), that is an estimated value considering both the protein and the ligand. It contains a given degree of uncertainty since we docked into a homology model (cluster 4,  $\text{RMSD} < 0.6 \text{ \AA}$ )
- Docking score ( $E_{\text{ASE2}}$ ), which considers geometrical purposes.
- and the Van-der-Waals energy ( $E_{\text{rvdw}}$ ), a three dimensional descriptor which considers the absolute geometrical configuration as well as the energy of the complex.

The proceeding was to sort the database by each of the mentioned descriptors. We chose those clusters containing the top ranked poses, i.e. the pose with the lowest value, was dedicated to further analysis. The amount of docked poses was reduced down to about 200.

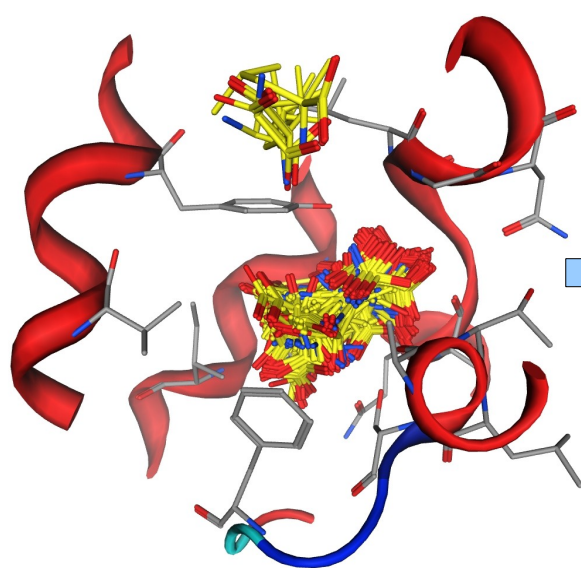


Figure 6.19a: Here the overlay of all poses of the redocking are shown

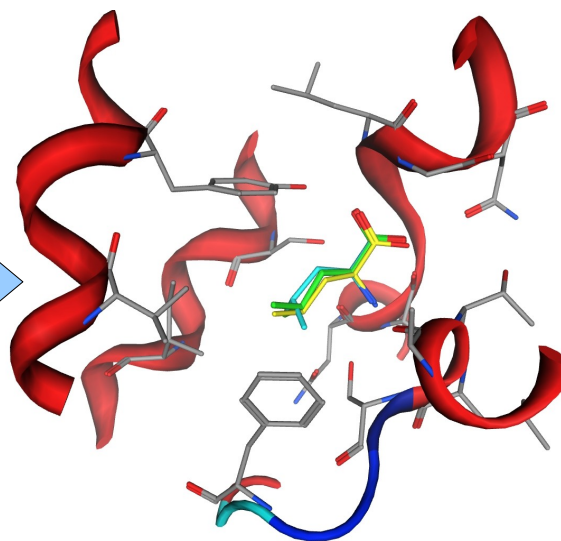


Figure 6.19b: The two best ranked clusters of poses after redocking in LeuT, see table 6.4. Yellow = the original ligand

descriptor	cluster	number of poses	deviation range	best pose ranked on	colour
potential energy	3	122	0,3144 to 0,3590	4	green
dock pKi	4	27	0,5874 to 0,5979	86	turquoise
E_ASE2	3	122	0,3144 to 0,3590	1	green
E_rvdw	3	122	0,3144 to 0,3590	11	green

Table 6.7 shows by which descriptor we chose which cluster and also the number of docked poses in the chosen cluster. The colour code refers to figure 6.19b

For the remaining poses we included a minimisation for each protein-ligand-complex by molecular mechanics to provide slight side chain movements of the binding site. This step we applied only for the docking results of SERT and DAT. Finally we selected two descriptors which were generated during the energy minimisation to select the most favourable complex:

- the total interaction energy of the complex (E\_interaction\_total)
- the Van-der-Waals interaction energy of the complex (E\_interaction\_vdw)



## 7 Results and Discussion

Similarly to the analysis of the redocking results we performed the analysis for the selected ligands in the SERT and the DAT. Finally we ended up with two protein ligand complexes for each ligand.

### 7.1 Docking Ligands into the Serotonin Transporter

For the docking runs we selected few ligands, where a multitude of biological data is available. As docking into homology models is still very risky, we docked substrates and inhibitors regarding to the evidence in the literature. One of our aims was to reveal the differences in the binding side respectively differences in the binding mode of MPD.

#### 7.1.1 Serotonin

Serotonin (5HT) is the proper substrate of the SERT. It binds with high affinity and selectivity due to a rapid removal from the synaptic cleft. Due to the lack of a crystal structure of the SERT, the exact binding mode is still not clear. According to the literature, we found, that the interaction of the dimethylamino sidechain with the D98 is crucial for binding as well as for substrate transportation (Barker et al., 1999; Henry et al. 2003).

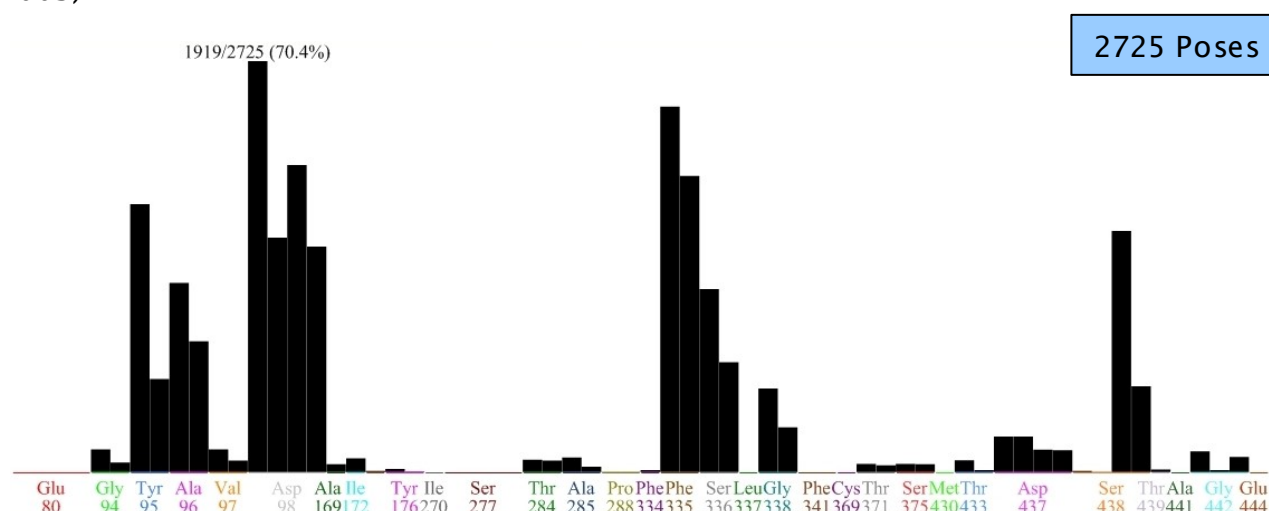


Figure 7.1: Histogram of the distribution of the interactions between the residues (abscissa) and the number of the docking poses (ordinate).

Three further residues have been hit during the docking: Y95, F335 and S438. Interactions of the substrate 5HT with Y95 in TM1 is supposed to be important (Henry et al., 2003 and 2006). Recently published articles on in silico docking suggest a similar binding mode (Kaufmann et al., 2008; Xhaard et al., 2008; Celik et al., 2008). We selected these four residues in the PLIF-tool as first criterion in the cue of reducing the docking poses. After exporting the RMSD-distance-matrix we grouped the poses into 20 clusters.

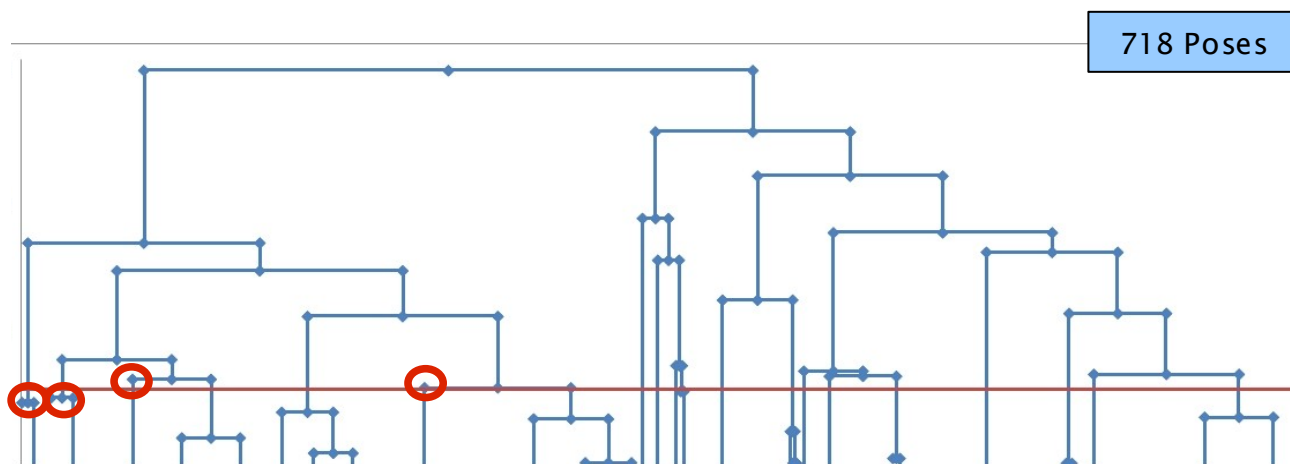
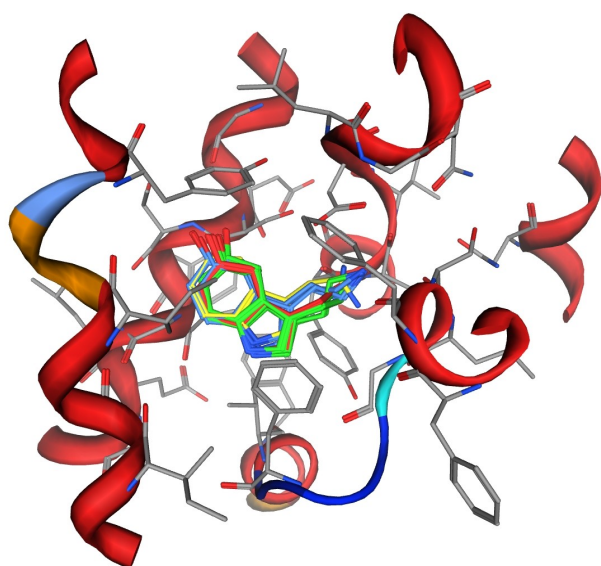


Figure 7.2: Clustering: x-axis shows the number of poses, y-axis depicts the degree of dissimilarity; red line = threshold of 20 cluster; red circles: the cluster we chose for further evaluation.



colour	cluster	descriptor	number
red	1	E	54
green	2	E_rvdw	86
blue	3	E_ASE2	61
yellow	6	dock_pKi	25

Table 7.1 shows the number of docking poses in the respective cluster. We chose the cluster with the top ranked pose referring to the descriptor.

Figure 7.3 depicts the overlay of all poses contained in each cluster. The colour code is written in Table 7.1.

The 20 clusters we sorted first by potential energy of the ligand, second by Van-der-Waals interaction energy, third by E\_ASE2 and dock\_pKi. Next we performed an energy minimisation step of the protein-ligand-complex by molecular mechanic simulations. It allows slight side chain rotations and provides an adaptation of the protein to the ligand. Based on the resulting energetic parameters we selected two complexes with the lowest Van-der-Waals interaction energy and lowest total interaction energy.

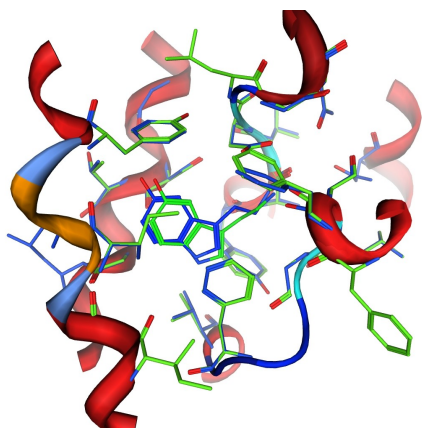


Figure 7.4a shows the two finally chosen poses. Since their orientation is very similar, the interaction pattern differs, see figures; green = total interaction energy, blue = Van-der-Waals interaction energy

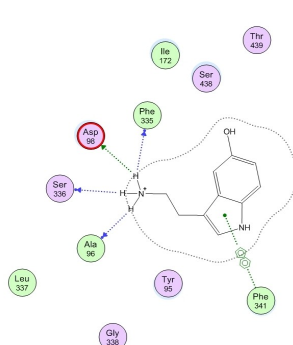


Figure 7.4b: Interaction pattern of serotonin selected by the total interaction energy. The ionic interaction between the nitrogen and D98 is visible.

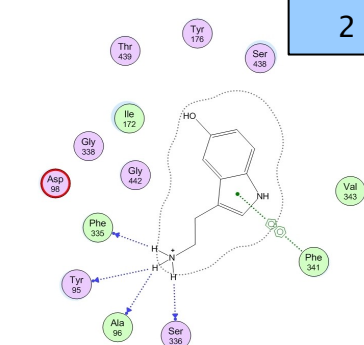
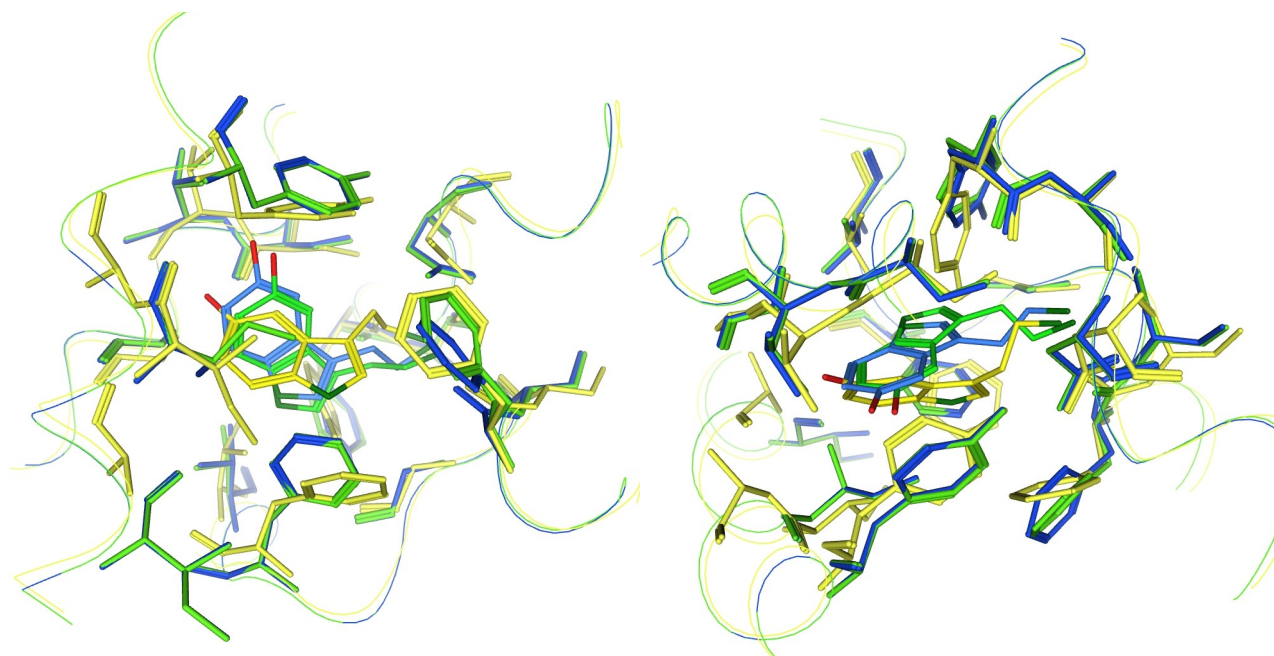


Figure 7.4c: Interaction pattern of the final pose selected by Van-der-Waals interaction energy. Since the important interaction is not present, we favour the pose depicted in figure 9.4b

A recently published paper shows a model of the SERT with bound 5HT. The comparison of the models shows a similar orientation of 5HT in the binding site (Celik et al., 2008). The overall RMSD of the compared transporter is more than 5 Å which is due to the different orientation of the large extracellular loop. Considering the binding site, the RMSD is lower than 1 Å.



Figures 7.5 a and b: These pictures show the residues which are directly lining the binding cavity and in the center the ligand serotonin. It is a direct comparison of the model from Celik et al. and our two most favoured poses. Concerning the protein, the orientation of the residues is quite similar. The bound ligand differs markedly from our selected poses. Yellow = the model of SERT with bound ligand from Celik et al., green = minimised pose selected by total interaction energy, blue = minimised pose selected by Van-der-Waals interaction energy. The backbone is depicted as line.

## 7.1.2 Methylphenidate

The affinity of MPD to SERT is much lower than to DAT. To elucidate the reason of this difference is one of the aims of the present study. Both enantiomers have been docked. Due to the flexibility of the ligand 3568 poses for R,R-MPD and 3620 poses for S,S-MPD were generated. For the PLIF we selected interactions with Y95 and D98 in TM1 and S438 in TM8. The interaction pattern from both enantiomers are quite similar. Compared to the interaction pattern of MPD in DAT, there are no significant differences. In SERT there are interactions with more residues detected.

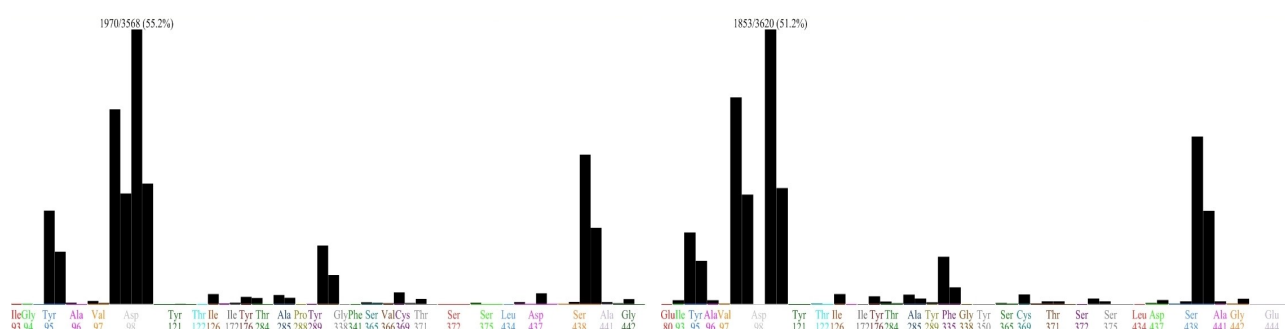


Figure 7.6 a and b: Interaction pattern of R,R-MPD and S,S-MPD in SERT obtained by the PLIF

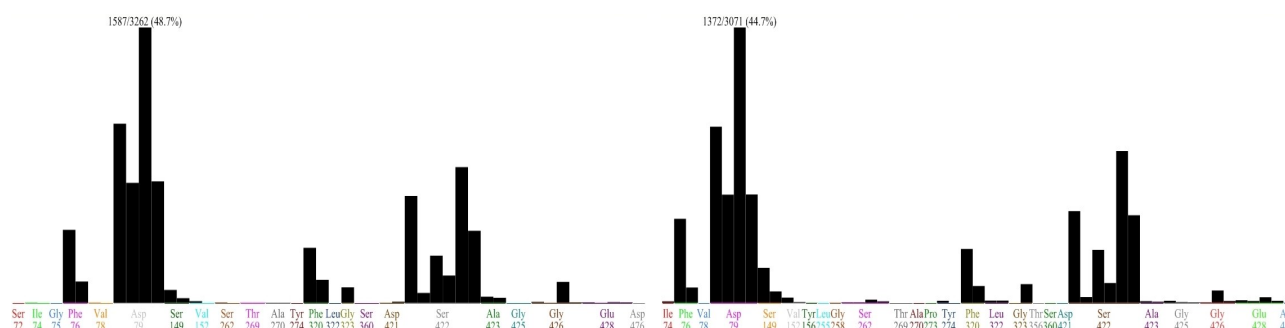
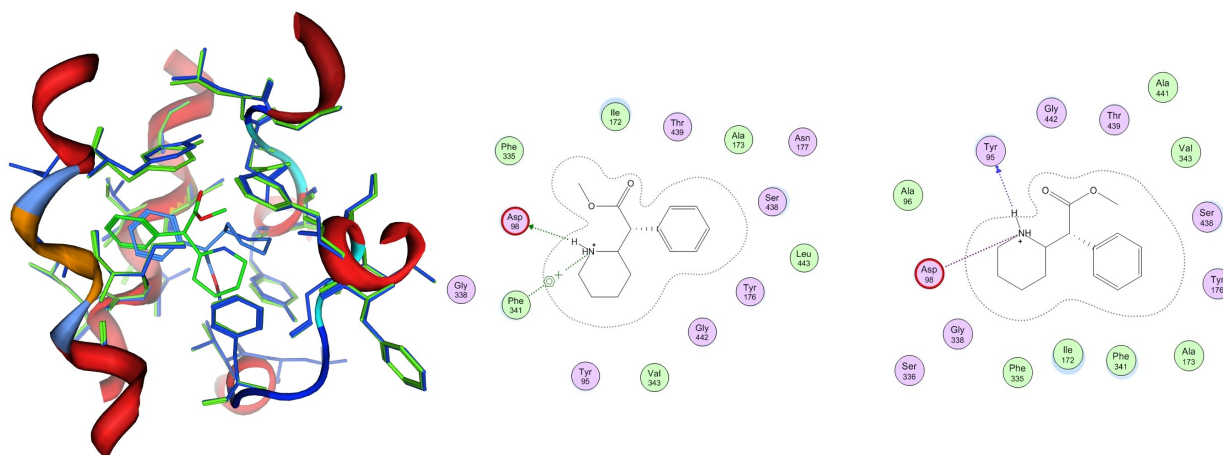
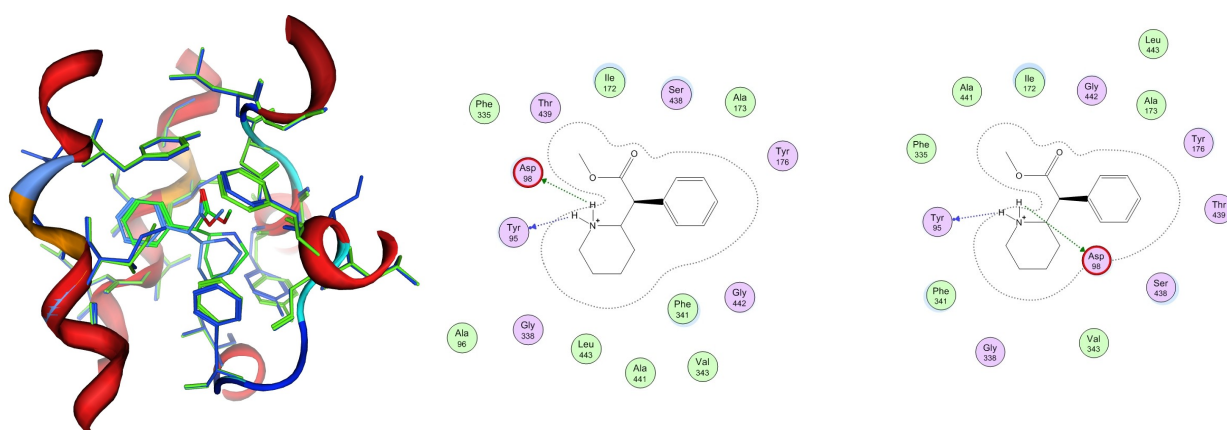


Figure 7.7 a and b: Interaction pattern of R,R-MPD and S,S-MPD in SERT obtained by the PLIF. The comparison of 7.6 and 7.7 reveals similar interactions of both enantiomers in both transporters

1143 poses for R,R-MPD and 1174 for S,S-MPD have been clustered into 50 groups from which we selected the four top ranked clusters for the energy minimisation. The selection of the four cluster for each enantiomer reduced the amount of poses to 107 and 145 respectively.



Figures 7.8 a, b and c show again (a) the orientation of our selected poses for R,R-MPD in the binding site of SERT and the interaction patterns of our top ranked poses due to (b) the total interaction energy (green) and (c) the Van-der-Waals interaction energy (blue). Since the differing orientations of the two poses, both show the interaction with D98.

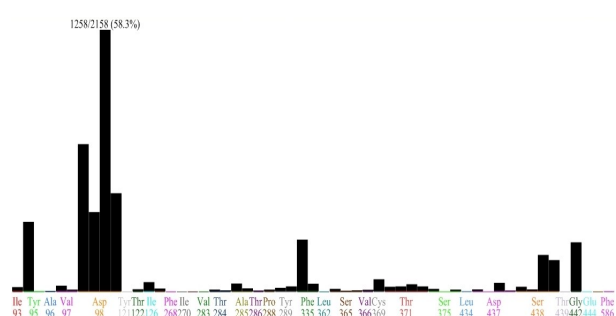


Figures 7.9 a, b and c show a similar orientation and interaction pattern for S,S-MPD. Notably for this enantiomer both top ranked poses are orientated similarly.

After energy minimisation the resulting poses show a characteristic difference concerning the orientation of the methyl ester group. The ester group of the R,R-enantiomer on the one hand points towards the extracellular (Y176 and D98) gate and on the other hand towards the hydrophobic part of the binding site (I172, F320 and V343). In contrast, the finally selected poses of the S,S-enantiomer show ester group orientated towards the extracellular gate. It suggests that the hydrophobic part of the binding site is occupied by the aromatic ring heterocyclic ring.

### 7.1.3 Cocaine and CFT

The affinity of Cocaine to SERT and DAT is on an equal level (Torres, 2003). CFT is used for experiments as cocaine analogue due to the similar pharmacological properties and the higher chemical stability. The PLIF shows a similar interaction pattern for both inhibitors of the transporter.



Figures 7.11a Interaction pattern of cocaine with SERT, obtained from the PLIF

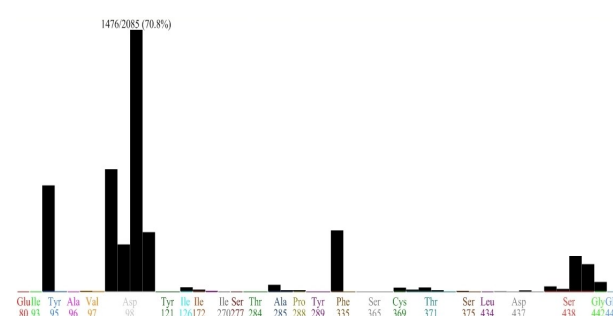


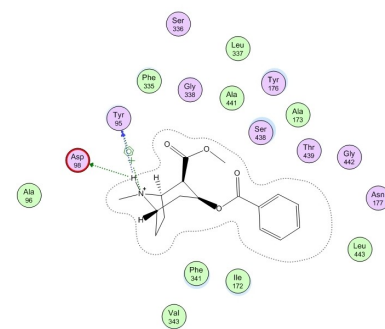
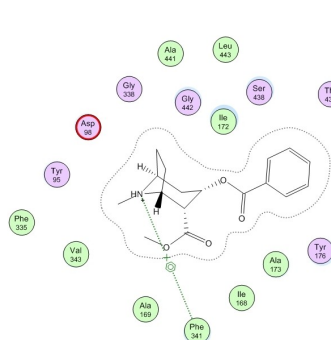
Figure 7.11b: Interaction pattern of CFT with SERT, obtained from the PLIF

In our docking experiments into SERT, cocaine shows two different poses similar to R,R-MPD. The orientation of the methyl ester is again different: on the one hand the ester group points towards the extracellular gate (Y176 and D98) and on the other hand the same moiety is orientated towards the hydrophobic pocket. In contrast, the analysis of the docking results of CFT revealed only one pose (see figure 7.13 a)

As depicted in the ligand interaction graphics (see figure 7.12 b and c), the pose selected by the Van-der-Waals interaction energy shows the important ionic interaction between D98 and the charged nitrogen. This suggests that this pose is more favourable also due to the fact, that the ester moiety points towards the extracellular gate (D98 and Y176).

After evaluation of the poses of CFT, we found only one pose. And this pose corresponds with docked cocaine, which was selected by the Van-der-Waals interaction energy. A recently published paper found a similar pose in DAT. Our finding as well as the results of Beuming et al., 2008 favours this pose as binding mode of tropane alkaloids. It suggests also, that the binding mode of CFT and cocaine is similar in both the DAT and the SERT.





Figures 7.12 a, b and c: As we found in the top ranked poses of R,R-MPD, also for cocaine two different ligand orientations have been selected. The pose selected by Van-der-Waals interaction energy (blue) shows the interaction with D98 (c). For this reason and as well due to the orientation of the ester moiety in the other pose we suggest this pose as favourable.

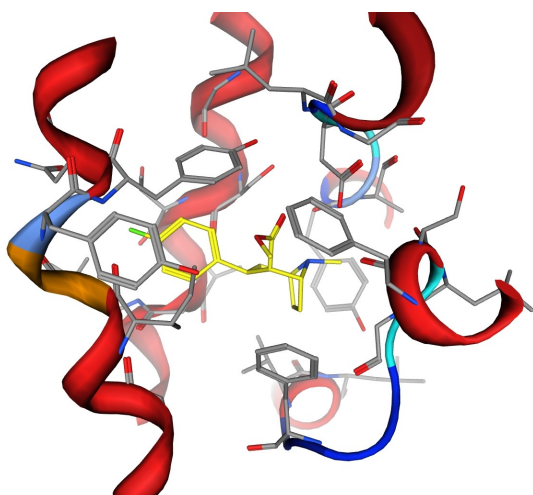


Figure 7.13a: We obtained only one single pose for CFT. The aromatic ring is orientated towards the hydrophobic cavity and the ester is in proximity of D98 and Y176.

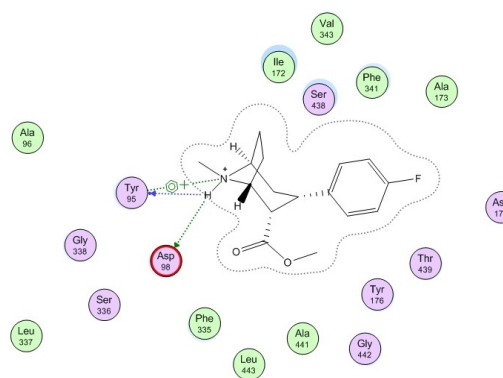


Figure 7.13b: Interaction pattern and surrounding residues of CFT in the binding site of SERT.

#### 7.1.4 MDMA

This widely abused amphetamine has a higher affinity to SERT than to DAT or NET. We obtained for each enantiomer two poses. The orientation of all four poses is quite similar. In comparison to the indoles, we suggest a similar orientation of the scaffold between them.

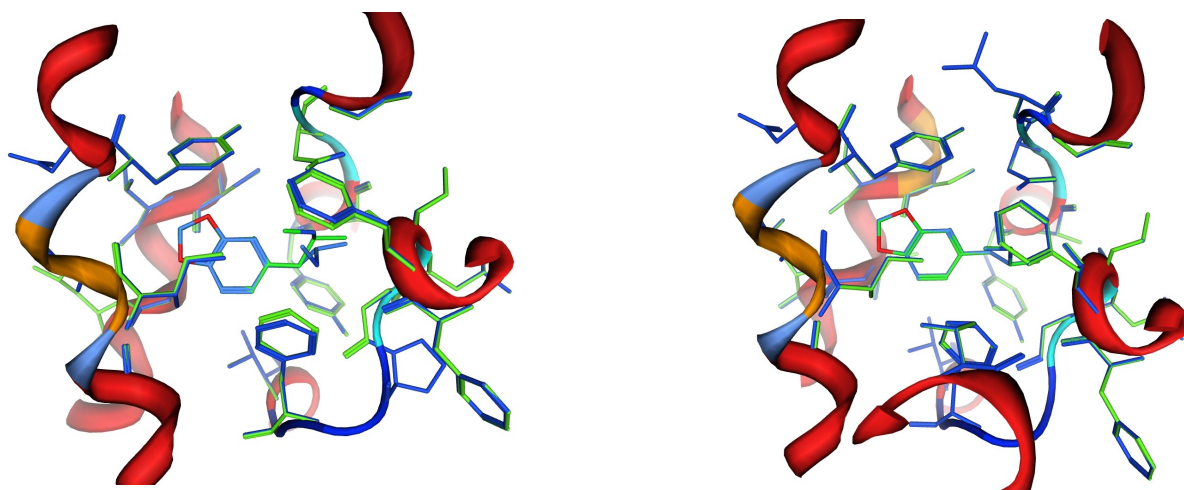
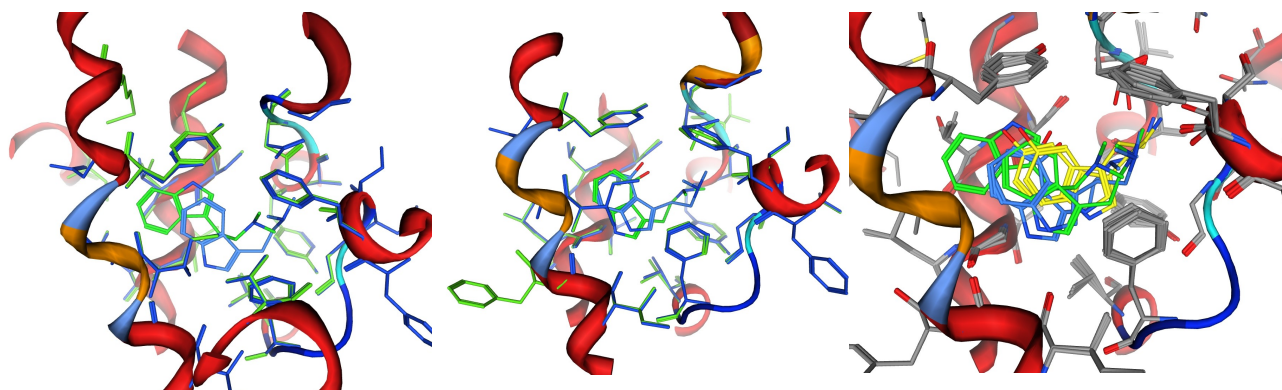


Figure 7.14 a and b show the orientation of our top ranked poses. green = selected by total interaction energy; blue = selected by Van-der-Waals interaction energy

#### 7.1.5 DMT and Psilocin

In order to examine differences in the binding mode of small indole alkaloids, we docked also DMT and psilocin. Our results show in principle a similar binding mode. However the indole moiety of one pose of DMT shows a different orientation in the hydrophobic cavity. Compared to the poses of 5HT, the indole scaffold of the N-methyl-substituted indoles is shifted deeper into the pocket. We suggest, that the reason of this shift is due to the two methyl groups connected with the nitrogen.



Figures 7.15 a, b and c show (a) the top ranked poses of DMT in SERT, (b) the top ranked poses of psilocin and (c) the overlay of 5HT (yellow), DMT (green) and psilocin (blue)

## 7.2 Differences and Similarities

After the procedure of analysing the docking poses, we propose that for the binding of the ligands the amino acid D98 in SERT and its homologue D79 in DAT are very important. This fact is according to the protein ligand interaction fingerprints and also according to the literature. Each of the ligands contain a charged nitrogen, that is placed in the surrounding of the mentioned aspartatic acid. This relatively strong ionic interaction exhibits a sort of ionic anchor point for the ligand binding. The other side of the binding cavity there is surrounded by aliphatic and aromatic residues and built up a hydrophobic pocket. Our data suggest that the aromatic moiety of the docked ligands binds deep in this pocket.

Our data suggest that the substrates are bound in a similar modality. For each of the substrate molecules the orientation of the charged nitrogen is towards the D98 and the aromatic part of the molecules is bound near I172, F320 in the hydrophobic cavity. Unfortunately we were not able to detect exactly definite interactions of the hydroxyl moieties of serotonin, psilocin and the methylene dioxy moiety of MDMA with residues within the binding site. Due to the proximity of these heteroatoms to S438, we suggest a direct interaction.

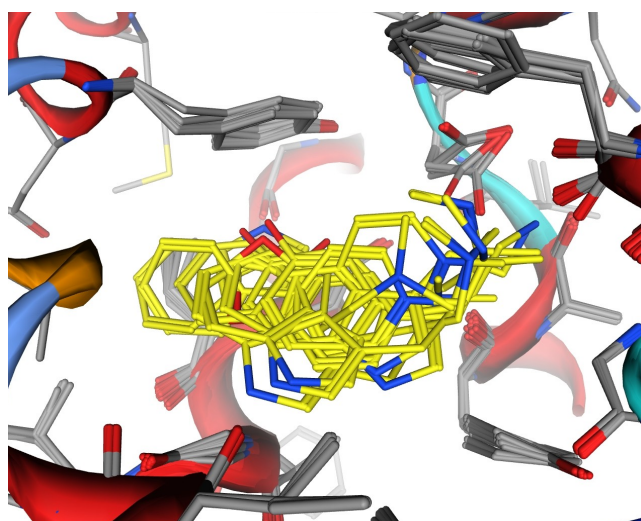


Figure 7.16a: Overlay of all substrates in SERT. It shows a similar binding mode.

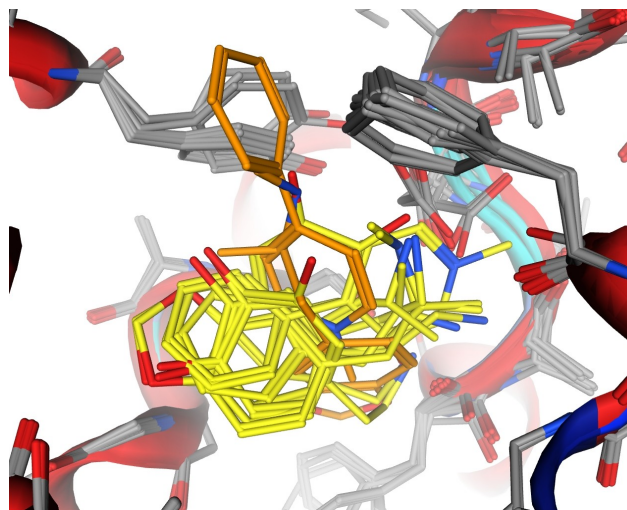


Figure 7.16b: Overlay of all substrates in DAT. The unfavourable poses are marked in orange. But for the other poses, a similar binding mode is also visible.

Concerning the orientation of the inhibitors we found similar binding modes within this group. The ester moiety of the compounds points towards Y176 and D98, which form a part of the extracellular gate and are important for the substrate translocation cycle. In comparison to the recently published article on the cocaine binding site in DAT (Beuming et al., 2008) we found analogous interactions with the respective residues in SERT. This suggests that cocaine shows the same binding mode both in DAT and SERT. Unfortunately our data did not show this binding mode in DAT.

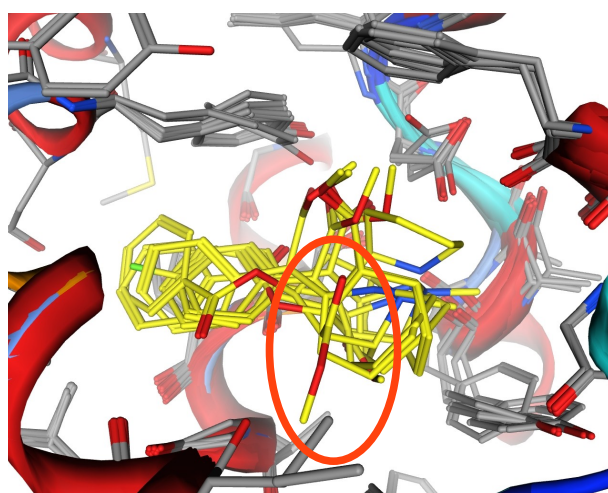


Figure 7.17a shows the docked inhibitors in SERT. The poses, where the ester moiety is turned into the bottom of the pocket we suggest to be unfavourable (orange circle).

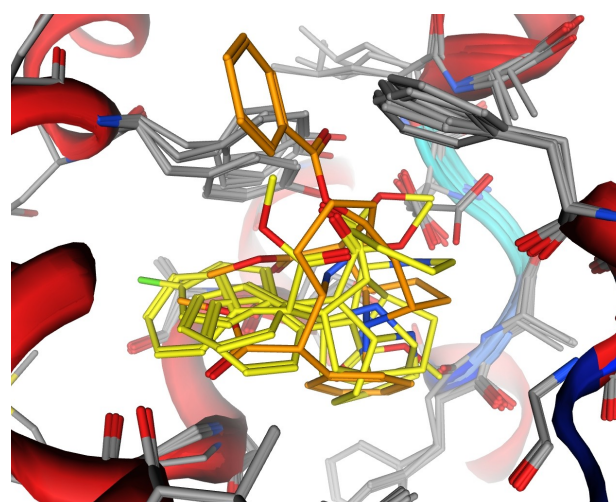


Figure 7.17b shows all docked inhibitors in DAT. We suggest few poses as unfavourable (orange).

At this point we should mention some limitations of our work:

- The template of our model is in the substrate occluded state. Since there is no crystal structure of the DAT, NET and SERT molecular modelling studies will have a certain degree of error.
- For this study we used only one software package.
- The co-transported ions have not been part of the model during the modelling and docking process.
- During the docking run, the receptor was held rigid.

### 7.3 Conclusion and Outlook

Based on our data we finally conclude the present study with the suggestion, that the group of substrates we selected for docking, i.e. 5HT, tryptamines and MDMA bind in a similar mode. The charged nitrogen directly interacts with D98 and the aromatic ring system is orientated towards the hydrophobic cavity. Unfortunately we were not able to detect a direct interaction of S438 with the heteroatoms of 5HT, psilocin and MDMA. But the spatial proximity may suggest an interaction. Since the hydroxyl group of Y95 is not orientated into the binding site, we were also not able to detect interactions.

The group of inhibitors, i.e. the tropane alkaloids, we chose for the docking experiments also bind in a similar mode. As mentioned above, the nitrogen is as well coordinated by a relatively strong ionic interaction with D98. The scaffold of all of the inhibitors show an ester moiety. This lead us propose, that those moieties are the inhibitory factor. The orientation of the ester was mainly placed towards the extracellular gate. But it needs further investigation on this to identify the exact mode of inhibition.

Comparing both the substrates and the inhibitors, we were able to show the coordination of the nitrogen with D98 as anchor point of ligand binding. Both groups bind in a similar mode due to their orientation of the charged nitrogen and the hydrophobic moieties of the ligands. Accordingly to pharmacological data, we suggest that the binding site of serotonin, methylphenidate and cocaine are overlapping.

Regarding the homology model, it will be necessary to create new models by using other software and considering the co-transported ions during the modelling. Structural improvements to include the disulfide bond in EL2 must also be considered. The extracellular vestibule beneath the kink in EL must be taken into account for possible docking into the substrate recognition site. A recently published paper suggests the structural homology of a bacterial glucose transporter with LeuT. The solute:sodium symporter vSGLT from *Vibrio parahaemolyticus* (pdb-code 3DH4) has been crystallised in the inward faced conformation and may be a second template for the NSS (Faham et al., 2008; Forrest et al., 2008)

The next step will be to run a molecular dynamic simulation to elucidate domain movements or the substrate translocation pathway. A recent paper suggests the importance of the extracellular vestibule for the transporter activity and substrate translocation (Shi et al., 2008).



## **Abstract: Molecular Modelling Studies for Analysing Interactions of Substrates and Inhibitors of the Serotonin Transporter**

The serotonin transporter (SERT) is a member of the neurotransmitter:sodium symporter (NSS) family, as well as the dopamine transporter (DAT) and norepinephrine transporter (NET). Serotonin (5-HT) is transported highly selective through the presynaptic membrane. This substrate translocation is effected by a secondary active transport: symport of two sodium and one chloride into the cell and antiport of one potassium ion out of the cell. Widely used CNS-therapeutics, such as tricyclic antidepressants (TCA) and selective serotonin reuptake inhibitors (SSRI), as well as illicit substances like cocaine and indole alkaloids target the SERT.

The ligands have been divided into two groups due to their mode of action: into substrates and inhibitors. For getting more information about the differences in the binding mode between these groups, we created a homology model of SERT and DAT based on the crystal structure of the Leucine Transporter (LeuT), a bacterial homologue of the NSS, published in 2005. The sequence identity between LeuT and SERT is around 20 %, which is rather low for creating homology models. In 2006, a comprehensive alignment of NSS was published. This study shows a homology of more than 40 % between both transporters.

After successful redocking of leucine in LeuT we docked a selection of ligands. The substrates 5-HT, MDMA and DMT and the inhibitors cocaine, WIN 35,428 (CFT = 2 $\beta$ -carbomethoxy-3 $\beta$ -(4-fluorophenyl)-tropane) and methylphenidate (MPD) have been docked to examine differences in the binding mode. We selected the most accurate poses in a four step procedure. In the first step, we used the Protein Ligand Interaction Fingerprint (PLIF), implemented in MOE 2007.09, to select those poses having interactions with residues that have experimentally been shown to be important for the function of the transporter. In the next step we calculated an RMSD-distance-matrix that was used as input for cluster analysis to generate homologous groups of the remaining poses. Next we selected those groups with the most favourable energetic parameters. The last step was an energy minimization of the ligand-receptor-complex. Two energetic descriptors, which describe the interaction energy of the complex were the basis for selecting one pose out of the remaining clusters.

Comparing our results to a recently published model of SERT with bound ligand, we found similar positioning. The analysis of our best poses showed that the different substrates bind in a similar modality. Concerning the inhibitors we found, that the binding mode is also similar within this group. In comparison to the recently published article on the cocaine binding site in DAT we found analogous interactions with the respective residues in SERT. This suggests that cocaine shows the same binding mode both in DAT and SERT. Based on the biological data, we further suggest that the binding sites of 5HT, cocaine and MPD are overlapping.



## **Zusammenfassung: Bindungsstudie von Substraten und Inhibitoren in einem Homologiemodell des Serotonin Transporters**

Der Serotonin-Transporter (SERT) zählt gemeinsam mit dem Dopamin- (DAT) und Noradrenalin- (NET) Transporter zu der Gruppe der Neurotransmitter-Transporter (Neurotransmitter:sodium symporter, NSS), einer Gruppe von membranständigen Transportproteinen, die in einem sekundär aktiven Transportmechanismus ihr Substrat gegen ein Konzentrationsgefälle in das emittierende Neuron transportiert. Serotonin wird gemeinsam mit zwei Natriumionen und einem Chloridion im Austausch gegen ein Kaliumion durch die präsynaptische Membran befördert. Der SERT ist auch Angriffspunkt für sowohl therapeutisch genutzte Wirkstoffe, wie den trizyklischen Antidepressiva oder den selektiven Serotonin Wiederaufnahmehemmern, als auch von psychoaktiven Substanzen, wie zum Beispiel Indolalkaloide oder Kokain.

Diese Liganden werden nach ihrem Wirkungsmechanismus in zwei Gruppen eingeteilt: den Substraten und den Inhibitoren. Um Unterschiede im Bindungsmodus dieser beiden Gruppen aufzudecken, erstellten wir ein Homologiemodell des Serotonin- und Dopamintransporters (DAT) auf Basis der Kristallstruktur eines in Funktion und Struktur dem SERT ähnlichen bakteriellen Leucin Transporters (LeuT). Die Identität der Primärsequenz von etwa 20 % ist für die Erstellung des Modells sehr niedrig. Eine im Jahr 2006 publizierte Studie über die Angleichung der Primärsequenzen von Neurotransmitter-Transportern ergab eine Homologie der beiden Sequenzen von über 40 %.

Wir begannen die Simulationen mit dem Docken von Leuzin in den LeuT. Danach dockten wir verschiedene Liganden des Serotonin Transporters. Als Substrate wählten wir Serotonin, die Indole DMT und Psilocin und das Amphetamin MDMA (Methylendioxyamphetamin). Das Tropanalkaloid Kokain und sein Analogon CFT (2 $\beta$ -karbomethoxy-3 $\beta$ -(4-fluorophenyl)-tropan), sowie das Amphetamin Methylphenidat wurden als Inhibitoren gewählt. Um aus der Vielfalt der erhaltenen Positionierungen der Liganden die plausibelsten auszuwählen, folgten wir einer vierstufigen Auswahlmethode: 1. Einschränkung der möglichen Posen aufgrund des Interaktionsmusters mit Hilfe des in MOE integrierten Protein Ligand Interaction Fingerprint (PLIF), 2. statistisches Gruppieren der Posen auf Basis der räumlichen Lage mit Hilfe einer RMSD-Distanzmatrix, 3. Heranziehen von berechneten, energetischen Parametern für die Auswahl der Gruppen und darauf folgende Energieminimierung und 4. Auswahl von zwei Posen mit den günstigsten energetischen Werten für den Protein-Ligand-Komplex.

Die Ergebnisse zeigen in den jeweiligen Gruppen, Substrate und Inhibitoren, sowohl ähnliche räumliche Orientierung der Posen, als auch ein ähnliches Interaktionsmuster. Aus den vorhandenen pharmakologischen Daten und unseren Ergebnissen können wir schließen, dass die Bindungstaschen von Serotonin, Kokain und Methylphenidat sich überlappen. Weiters lassen die daten schließen, dass der Bindungsmodus von Kokain und Methylphenidat in SERT und DAT sich gleichen.

## References

- Adkins E. M., Barker E. L., Blakely R. D.; Interactions of Tryptamine Derivatives with Serotonin Transporter Species Variants Implicate Transmembrane Domain I in Substrate Recognition; *Mol Pharmacol.*, 2001, 59:514–523
- Al-Lazikani B., Jung J., Zhexin Xiang Z., Honig B.; Protein structure prediction; *Current Opinion in Chemical Biology*, 2001, 5:51–56
- Androutsellis-Theotokis A., Goldberg N. R., Ueda K., Beppu T., Beckman M. L., Das S., Javitch J. A., Rudnick G.; Characterization of a Functional Bacterial Homologue of Sodium-dependent Neurotransmitter Transporters; *J. Biol. Chem.*, 2003, 278:12703–12709
- Baker D., Sali A.; Protein Structure Prediction and Structural Genomics; *Science*, 2001, 294:93-96
- Barker E. L., Moore K. R., Rakhshan F., Blakely R. D.; Transmembrane Domain I Contributes to the Permeation Pathway for Serotonin and Ions in the Serotonin Transporter; *The Journal of Neuroscience*, 1999, 19:4705–4717
- Baumann M. H., Clark R.D., Budzynski A. G., Partilla J. S., Bruce E Blough B. E., Rothman R. B.; N-Substituted Piperazines Abused by Humans Mimic the Molecular Mechanism of 3,4-Methylenedioxymethamphetamine (MDMA, or 'Ecstasy'); *Neuropsychopharmacology*, 2005, 30:550–560
- Beuming T., Shi L., Javitch J. A., Weinstein H.; 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; *Mol Pharmacol*, 2006, 70:1630–1642
- Beuming T., Kniazeff J., Bergmann M. L., Shi L., Gracia L., Raniszewska K., Newman A. H., Javitch J. A., Weinstein H., Gether U., Loland C. J.; The binding sites for cocaine and dopamine in the dopamine transporter overlap; *Nature Neuroscience*, 11:780-789
- Biederman J., Spencer T.; Attention-Deficit/Hyperactivity Disorder (ADHD) as a Noradrenergic Disorder; *Biol Psychiatry*, 1999, 46:1234 – 1242
- Böhm H. J., Klebe G., Kubinyi H.; *Wirkstoffdesign*, Spektrum, 1996
- Celik L., Sinning S., Severinsen K., Hansen C. G., Møller M. S., Bols M., Wiborg O., Schiøtt B.; Binding of Serotonin to the Human Serotonin Transporter. Molecular Modeling and Experimental Validation; *JACS*, 2008, 130:3853-3865

- Chen J., Rudnick G.; Permeation and gating residues in serotonin transporter; PNAS, 2000, 97: 1044–1049
- Chen N., Reith M. E. A.; Structure and function of the dopamine transporter; European Journal of Pharmacology, 2000, 405:329–339
- David L., Nielsen P. A., Hedström M., Norden B.; Scope and Limitation of Ligand Docking: Methods, Scoring Functions and Protein Targets; Current Computer-Aided Drug Design, 2005; 1:275-306
- Dunbrack R. L. Jr; Sequence comparison and protein structure prediction; Current Opinion in Structural Biology, 2006, 16:374–384
- Ermondi G., Carona G., Lawrence R., Longoc D.; Docking studies on NSAID/COX-2 isozyme complexes using Contact Statistics analysis; Journal of Computer-Aided Molecular Design, 2004, 18:683–696
- Ermondi G., Caron G.; Recognition forces in ligand–protein complexes: Blending information from different sources; Biochemical Pharmacology, 2006, 72:1633–1645
- Faham S., Watanabe A., Besserer G. M., Cascio D., Specht A., Hirayama B. A., Wright E. M., Abramson J.; The Crystal Structure of a Sodium Galactose Transporter Reveals Mechanistic Insights into Na<sup>+</sup>/Sugar Symport; Science, 2008, 321:810-814
- Fain G. L.; Molecular and Cellular Physiology of Neurons, Harvard University Press, 1999
- Forrest, L. R., Tavoulari S., Zhang Y., Rudnick G., Honig B.; Identification of a chloride ion binding site in Na/Cl-dependent transporters, PNAS; 2007, 104:12761-12766
- Forrest L. R., Zhang Y, Jacobs M. T., Gesmonde J., Xie L., Honig B. H.; Rudnick G.; Mechanism for alternating access in neurotransmitter transporters, PNAS, 2008, 105:10338–10343
- Goldstein D. S., Eisenhofer G., Kopin I. J.; Clinical Catecholamine Neurochemistry: A Legacy of Julius Axelrod; Cellular and Molecular Neurobiology, 2006, 26:695–702
- Hardeland R.; Melatonin, hormone of darkness and more – occurrence, control mechanisms, actions and bioactive metabolites; Cell. Mol. Life Sci., 2008, 65:2001–2018
- Henry L. K., Adkins E. M., Han Q., Blakely R. D.; Serotonin and Cocaine-sensitive Inactivation of Human Serotonin Transporters by Methanethiosulfonates Targeted to Transmembrane Domain I; J. Biol. Chem., 2003, 278:37052–37063

Henry L. K., Field J. R., Adkins E. M., Parnas M. L., Vaughan R. A., Zou M., Newman A. H., Blakely R. D.; Tyr-95 and Ile-172 in Transmembrane Segments 1 and 3 of Human Serotonin Transporters Interact to Establish High Affinity Recognition of Antidepressants; *J. Biol. Chem.*, 2006, 281:2012–2023

Howell L. L., Kimmel H. L.; Monoamine transporters and psychostimulant addiction; *Biochemical Pharmacology*, 2008, 75:196–217

Huang X., Zhan C.; How Dopamine Transporter Interacts with Dopamine: Insights from Molecular Modeling and Simulation, *Biophysical Journal*, 2007, 93:3627–3639

Iversen Leslie, Speed, Ecstasy, Ritalin, Oxford University Press, 2006

Iversen S. D., Iversen L. L.; Dopamine: 50 years in perspective; *Trends in Neurosciences*, 2006, 30:188–193

Javitch J. A., D'Amato R. J., Strittmatter S. M., Snyder S. H.; Parkinsonism-inducing neurotoxin, N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine: Uptake of the metabolite N-methyl-4-phenylpyridine by dopamine neurons explains selective toxicity; *Proc. Natl. Acad. Sci.*, 1985, 82:2173–2177

Jonnakuty C., Gragnoli C.; What Do We Know About Serotonin?; *J. Cell. Physiol.*, 2008, 217:301–306

Kamdar G., Penado K. M. Y., Rudnick G., Stephan M. M.; Functional Role of Critical Stripe Residues in Transmembrane Span 7 of the Serotonin Transporter; *Journal of Biological Chemistry*, 2001, 276:4038–4045

Kaufmann K. W., Dawson E. S., Henry L. K., Field J. R., Blakely R. D., Meiler J.; Structural determinants of species-selective substrate recognition in human and *Drosophila* serotonin transporters revealed through computational docking studies; published online: 14 Aug 2008

Kelkar S. V., Izenwasser S., Katz J. L., Klein C. L., Zhu N., Trudell M. L.; Synthesis, Cocaine Receptor Affinity, and Dopamine Uptake Inhibition of Several New 2p-Substituted 38-Phenyltropanes; *J. Med. Chem.*, 1994, 37:3875–3877

Kitchen D. B., Decornez H., Furr J. R., Bajorath J.; Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications; *Nature Reviews*, 2004, 3:935–349

Kline, R. H., Jr., Eshleman A. J., Wright J., Eldefrawi M. E.; Synthesis of Substituted 3-Carbamoyllecgonine Methyl Ester Analogues: Irreversible and Photoaffinity Ligands for the Cocaine Receptor/Dopamine Transporter; *J. Med. Chem.*, 1994, 37:2249–2252

Leach A. R., Shoichet B. K., Peishoff C. E.; Docking and Scoring: Perspective - Prediction of Protein-Ligand Interactions. Docking and Scoring: Successes and Gaps; 2006, 49:5851-5855

Lee D., Redfern O., Orengo C.; Predicting protein function from sequence and structure; Nature Reviews, 2007, 8:995-1005

Lesk A. M.; Introduction in Protein Architecture, Oxford University Press, 2001

Löffler G.; Basiswissen Biochemie, Springer Verlag, 2001

Madras B. K. (Editor); Cell Biology of Addiction, Cold Spring Harbour Laboratory Press, 2006

Martí-Renom M. A., Stuart A. C., Fiser A., Sánchez R., Melo F., Sali, A.; Comparative Protein Structure Modeling of Genes and Genomes; Annu. Rev. Biophys. Biomol. Struct., 2000, 29:291-325

Masson J., Sagné C., Hamon M., El Mestikawy S.; Neurotransmitter Transporters in the Central Nervous System; Pharmacological Reviews, 1999, 51:439-464

Meltzer P. C., Wang P., Blundell P., Madras B. K.; Synthesis and Evaluation of Dopamine and Serotonin Transporter Inhibition by Oxacyclic and Carbacyclic Analogues of Methylphenidate; J. Med. Chem., 2003, 46:1538-1545

Moult J.; A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction; Current Opinion in Structural Biology, 2005, 15:285-289

Mutschler E; Arzneimittelwirkungen, WVG, 2001

Nonaka R., Nagai F., Ogata A., Satoh K.; In Vitro Screening of Psychoactive Drugs by [<sup>35</sup>S]GTPγ S Binding in Rat Brain Membranes; Biol. Pharm. Bull., 2007, 30:2328-2333

Petsko G. A., Ringe D. (Editors); Protein Structure and Function, Sinauer Associates, 2004

Petrey D., Honig B.; Protein Structure Prediction: Inroads to Biology; Molecular Cell, 2005, 20: 811-819

Pschyrembel W.; Klinisches Wörterbuch; WdeG, 2001, 259. Auflage

Quick M. W. (Editor); Transmembrane Transporters, Wiley-Liss, 2002

Quick M., Yano H., Goldberg N. R., Duan L., Beuming T., Shi L., Weinstein H., Javitch J. A.; State-dependent Conformations of the Translocation Pathway in the Tyrosine Transporter Tyt1, a Novel Neurotransmitter:Sodium Symporter from *Fusobacterium nucleatum*; J Biol Chem, 2006, 281:26444-26454

Rauhut R.; Bioinformatik, Wiley-Vch, 2001

Rothman R. B., Baumann M. H.; Monoamine transporters and psychostimulant drugs; European Journal of Pharmacology, 2003, 479:23-40

Rothman R. B., Baumann M. H., Prisinzano T. E., Newman A. H.; Dopamine transport inhibitors based on GBR12909 and benztropine as potential medications to treat cocaine addiction; Biochemical Pharmacology, 2008, 75:2-16

Rudnick G.; Bioenergetics of Neurotransmitter Transport; Journal of Bioenergetics and Biomembranes, 1998, 30:173-185

Rudnick G., Serotonin Transporters – Structure and Function; J. Membrane Biol., 2006, 213:101-110

Saier M. H. Jr; Genome archeology leading to the characterization and classification of transport proteins; Current Opinion in Microbiology, 1999, 2:555-561

Schultes E., Hofmann A.; Pflanzen der Götter, AV, 1998

Shi L., Quick M., Zhao Y, Weinstein H., Javitch J. A.; The Mechanism of a Neurotransmitter:Sodium Symporter - Inward Release of Na<sup>+</sup> and Substrate Is Triggered by Substrate in a Second Binding Site; Molecular Cell, 2008, 30:667-677

Shulgin A., Shulgin A.; PIHKAL - a chemical love story, Transform Press, 1995

Siegel, G.J., Agranoff, B.W., Albers, R.W., Molinoff, P.B. (Editors); Basic Neurochemistry, Raven Press, 1994

Singh S. K., Yamashita A., Gouaux E.; Antidepressant binding site in a bacterial homologue of neurotransmitter transporters; Nature Letters, 2007, 448:952-956

Snyder S. H.; Chemie der Psyche - Drogenwirkung im Gehirn, Spektrum, 1990

Steger G.; Bioinformatik, Birkhäuser, 2003



Storch A., Ludolph A. C., Schwarz J.; Dopamine transporter: involvement in selective dopaminergic neurotoxicity and degeneration; J Neural Transm, 2004, 111:1267-1286

Swanson J. M., Volkow N. D.; Serum and brain concentrations of methylphenidate: implications for use and abuse; Neuroscience and Biobehavioral Reviews, 2003, 27:615-621

Thews, G., Mutschler, E., Vaupel, P. (Editors); Anatomie, Physiologie, Pathophysiologie des Menschen, WVG, 1999

Torres G. E., Gainetdinov R. R., Caron M. G.; Plasma Membrane Monoamine Transporters: Structure, Regulation and Function; Nature Reviews Neuroscience, 2003, 4:13-25

Whitaker-Azmitia P. M.; The Discovery of Serotonin and its Role in Neuroscience; Neuropsychopharmacology, 1999, 21

White K. L., Kiser P. D., Nichols D. E., Barker E. L.; Engineered zinc-binding sites confirm proximity and orientation of transmembrane helices I and III in the human serotonin transporter; Protein Sci., 2006, 15:2411-2422

Xhaard H., Backstrom V., Denessiouk K., Johnson M. S.; Coordination of Na<sup>+</sup> by Monoamine Ligands in Dopamine, Norepinephrine, and Serotonin Transporters; J. Chem. Inf. Model advanced online

Yamashita A., Singh S. K., Kawate T., Jin Y., Gouaux E.; Crystal structure of a bacterial homologue of Na/Cl-dependent neurotransmitter transporters; Nature, 2005, 437:215-223

Zhou Z., Zhen J., Karpowich N. K., Goetz R. M., Law C. J., Reith M. E. A., Wang D.; LeuT-Desipramine Structure Reveals How Antidepressants Block Neurotransmitter Reuptake; Science, 2007, 317:1390-1393

Zomot E., Bendahan A., Quick M., Zhao Y, Javitch J. A., Kanner B. I; Mechanism of chloride interaction with neurotransmitter:sodium symporters; nature letters, 2007, 1-6

## Links

The Protein Database	<a href="http://www.rcsb.org">http://www.rcsb.org</a>
Swissprot	<a href="http://www.expasy.org">http://www.expasy.org</a>
The Chemical Computing Group	<a href="http://www.chemcomp.com">http://www.chemcomp.com</a>
Isis Draw	<a href="http://www.mdli.com">http://www.mdli.com</a>

### Cooperation partners:

homepage of Dr. Gerhard F. Ecker	<a href="http://homepage.univie.ac.at/Gerhard.F.Ecker">http://homepage.univie.ac.at/Gerhard.F.Ecker</a>
Medical University of Vienna	<a href="http://www.meduniwien.ac.at/pharmakologie">http://www.meduniwien.ac.at/pharmakologie</a>
SFB35	<a href="http://www.sfb35.at">http://www.sfb35.at</a>

### Contact:

All files are available from René Weissensteiner ([rene@dekor.cc](mailto:rene@dekor.cc)) and Gerhard F. Ecker ([gerhard.f.ecker@univie.ac.at](mailto:gerhard.f.ecker@univie.ac.at)).

# Curriculum vitae

## Persönliche Daten:

Name René Weissensteiner  
Geburtstag 2. November 1978  
Familienstand ledig, verlobt mit Martina Gall,  
gemeinsame Tochter Sophie Katharina (\* 8.10.2006, † 12.2.2008)

## Ausbildung:

1985 – 1989: Volksschule II in Tulln  
1989 – 1993: Hauptschule III in Tulln  
1993 – 1994: HTBLuVA St. Pölten, Abteilung EDV und Organisation  
1994 – 1999: Handelsakademie St. Pölten, Matura im Juni 99  
seit Oktober 1999: Studium der Pharmazie an der Universität Wien,  
Februar 2002: Sprachreise México  
Jänner 2004: 1. Diplomprüfung  
Februar 2005: Botanische Exkursion Costa Rica  
seit Oktober 2007: Diplomarbeit Pharmakoinformatik

## Arbeitstätigkeit

Juli 1994: Ferialpraxis, Leitzinger GmbH  
Sommer 95, 96, 98, 2000: Ferialarbeit, Brucha GmbH  
Okt. 2000 – Okt. 2001: Zivildienst, Emmausgemeinschaft St. Pölten  
September 2002: Ferialpraxis, Apotheke Pöggstall, NÖ  
Nov. 2002 – Jan. 2004: geringfügig beschäftigt, Apotheke A. Moll, Wien 1  
Juli/August 2003: Ferialarbeit, Brucha GmbH  
September 2004: Ferialpraxis, in Apotheke Pöggstall, NÖ  
Okt. 2005 – Jun. 2006: geringfügig beschäftigt, Hellrein GmbH  
Okt. 2006 - Feb. 2007: Karenz  
seit Oktober 2007: Tätigkeit als Schlagzeuglehrer

## Sonstiges / Hobbies:

Musik, erlerne Schlagzeug und Perkussion seit 1993, Gitarre und Klavier  
autodidaktisch, musikalische Tätigkeit in verschiedenen Gruppen und Musikvereinen  
Fotografie und digitale Medien  
Sport und Reisen, ausgedehnte Aufenthalte in der Natur  
Wissenschaftliche Literatur