universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Ligand and Structure based Studies on Methylphenidate Analogues

angestrebter akademischer Grad

Magister der Pharmazie (Mag.pharm.)

Verfasserin / Verfasser:	Martin Wurzer
Matrikel-Nummer:	0501210
Studienrichtung/Studienzweig (lt. Studienblatt):	Pharmazie
Betreuerin / Betreuer:	O. Univ. Prof. Mag. Pharm. Dr. Gerhard Ecker

Wien, im Mai 2010

Danksagung

Zu aller erst möchte ich Herrn Univ. Professor Dr. Gerhard F. Ecker für seine Betreuung und Unterstützung danken die er mir im Rahmen dieser Diplomarbeit entgegengebracht hat.

Ich möchte meiner gesamten Arbeitsgruppe danken. Sie haben mich in vielen Fragen tatkräftig unterstützt. Oft haben sie mir durch ihre größere praktische Erfahrung dabei geholfen Probleme meiner Arbeit zu lösen.

Danke Andrea, Andreas, Christoph, Daniela, Freya, Ishrat, Johanna, Lars, Rene und Yogesh!

Außerdem möchte ich meiner Freundin Lisi und meiner Familie danken. Danke für euer Verständnis und eure Unterstützung!

Table of Contents

1 Biological Background

1.1 Biology of the Dopamine Transporter

1.1.1 Dopamine

Dopamine is a neurotransmitter belonging to the monoamine family. It is biosynthesized in various locations of the nervous system in two steps (Figure 1) out of the amino acid Tyrosine and furthermore it is also the precursor for Epinephrine and Norepinephrine.

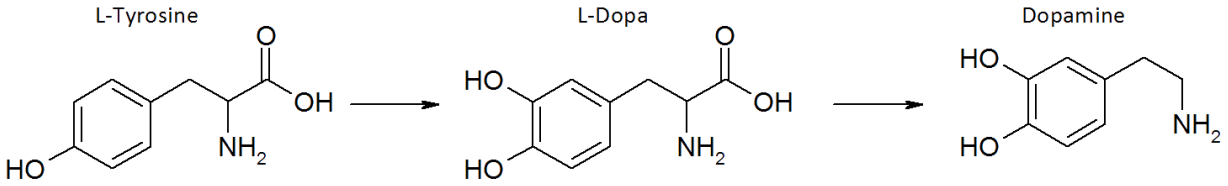


Figure 1

On the one hand Dopamine has several functions in the periphery but also in the central nervous system. Two groups containing five subtypes of dopamine receptors (D_1-D_5) are known. In the periphery it is involved in the regulation of the blood pressure and kidney blood flow. The effects in the CNS are diverse. It plays an important role in regulating the lactation (via prolactin), the motor function and several affective and cognitive processes. Therefore many diseases (for example M. Parkinson and schizophrenia) are associated with the malregulation of dopamine concentration in some parts of the human brain. (Steinhilber et al. 2005)

1.1.2 The Dopamine Transporter

The dopamine transporter (DAT) has the function to pump dopamine out of the synaptic cleft into the neuron after its release. It is a transmembrane protein consisting of 12 α-Helices and one bigger extracellular loop. The N- and the C-terminus both face the inside of the cell (Figure 2). Until today no crystal structure has been obtained and therefore the exact 3D structure is still unknown. (Vaughan et al. 2005)

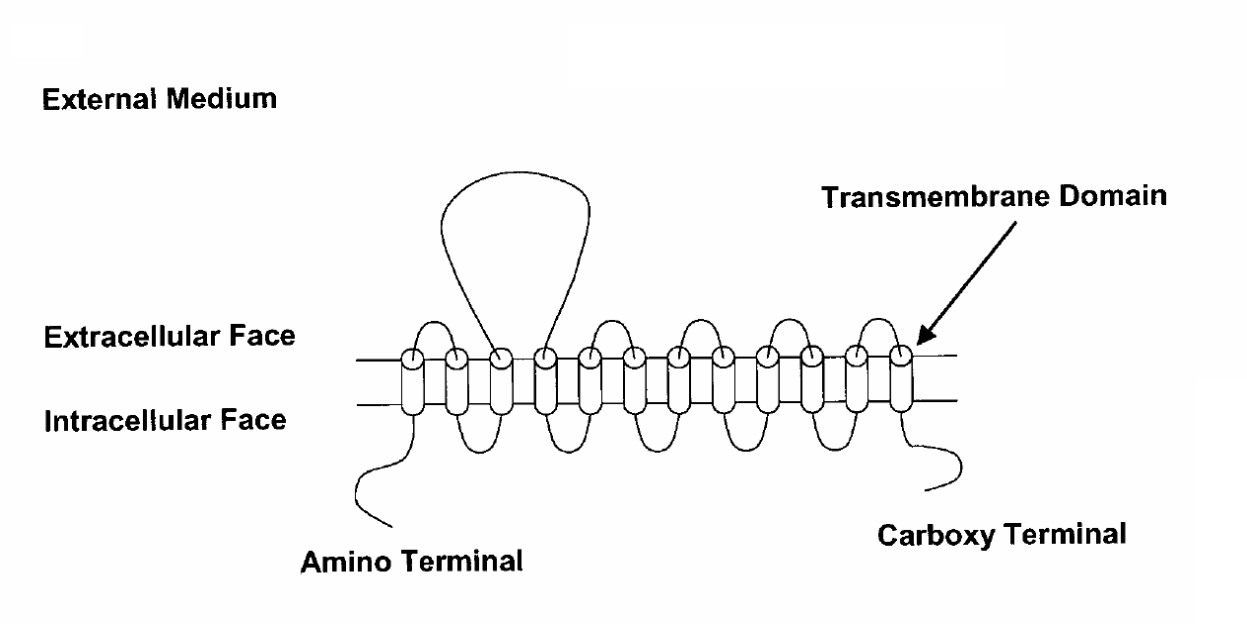


Figure 2 taken from (Volz and Schenk 2005)

DAT is a member of the sodium/monoamine neuro-transmitter co-transporter family (other members: The serotonin transporter (SERT) and the norepinephrine transporter (NET)). Dopamine transport is driven by a Na^+ gradient inwards the cell which is maintained by the membrane Na^+/K^+-ATPase. One molecule of dopamine is accompanied by two Na^+ and one Cl^- Ions.

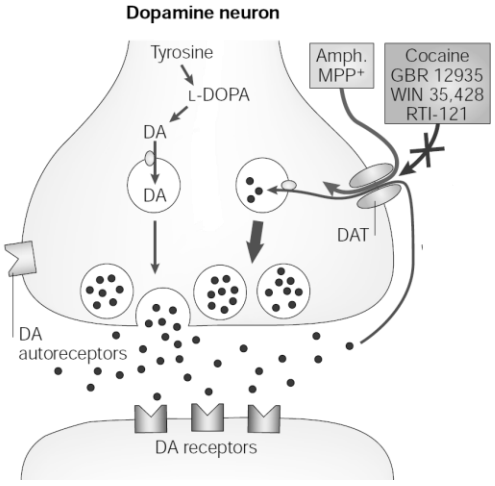


Figure 3 taken from (Torres et al. 2003)

Many drugs interact with the monoamine transporters targeting them with different binding affinities. Some of

them like cocaine and methylphenidate block (Figure 3) the transporter which results in a higher concentration of the transmitters in the synaptic cleft. Amphetamine is even able to invert the direction of the dopamine transport (besides it also increases the neuronal excretion of dopamine vesicles). (Torres et al. 2003)

1.1.3 Important interacting Drugs

- Cocaine

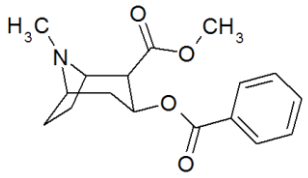


Figure 4

The coca plant Erythroxylum coca, which is native in South America, has been known for its stimulant hunger-suppressant effects for centuries. The drug is applied by chewing the leaves of the plant and people still use it this way today.(Goldstein et al. 2009)

In the 19th century cocaine was extracted as an active substance and the drug was first widely used as an anesthetic drug (This effect is related to the block of sensory nerve fibers by blocking Na^+Channels). But it wasn't only used in the medical field: the leaves were even used for the preparation of alcoholic and non alcoholic beverages (the best known example is coca cola)(Goldstein et al. 2009).

Today cocaine is one of the heaviest abused drugs even starting with adolescence.

> *"The National Institute on Drug Abuse (NIDA)- funded 2007 Monitoring the Future Study showed that 2.0% of 8th graders, 3.4% of 10th graders, and 5.2% of 12th graders had abused cocaine in any form and 1.3% of 8th graders, 1.3% of 10th graders, and 1.9% of 12th graders had abused crack at least once in the year prior to being surveyed."*
>
> (Goldstein et al. 2009)

Besides being one of the drugs with the highest addictive potential (Nutt et al. 2007) it can also cause severe side effects especially concerning the cardiovascular and the CN system. Due to its effects on the vascular system it increases the risk of a stroke or heart attacks

and arrhythmia (Qureshi et al. 2001). The psychic disorders caused by sustained cocaine abuse are wide ranging.

As already mentioned before the stimulant effects of cocaine are the result of the unselective blockade of the three monoamine transporters DAT, SERT and NET.

- Methylphenidate

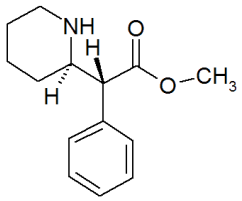


Figure 5 (Threo MP)

Methylphenidate (Figure 5) is an artificial drug which was first synthesized in the early 1940s. It appeared on the market in the middle of the 1950s in the USA and Germany. At the end of that decade it was used for the treatment of "behavioral problems" in children for the first time. (Schmutz 2004)

Today Methylphenidate is one of the most commonly used medicinal drugs to treat ADHD (**A**ttention **D**eficit **H**yperactivity **D**isorder) especially in children (Kollins et al. 2001). Furthermore it has also been under discussion as a substitution therapy for cocaine or amphetamine addicts. Until now no significant positive effects have been shown for cocaine dependence but there have been first positive outcomes regarding the treatment of amphetamine and met-amphetamine addicts. Nevertheless both indications need further investigation. (Grabowski et al. 1997; Elkashef et al. 2008; Konstenius et al. 2009)

Like cocaine methylphenidate has a high affinity for blocking DAT and the NET while the affinity for SERT is significantly decreased. (Threo) Dexmethylphenidate HCl was identified as the more active one but until now also the racemate has still been in use. (Gatley et al. 1996)

Due to the rather similar type of pharmacokinetic and its stimulant effects it is no surprise that methylphenidate is also a drug that is abused. In therapeutic doses used for the

treatment of ADHD its abusive potential is rather low. However, this risk shouldn't be disregarded. (Kollins et al. 2001)

Methylphenidate can also cause a number of side effects: Tachycardia, headache, appetite loss, increase of blood pressure, allergic reaction and others have been reported.(Mutschler et al. 2008)

- Amphetamin

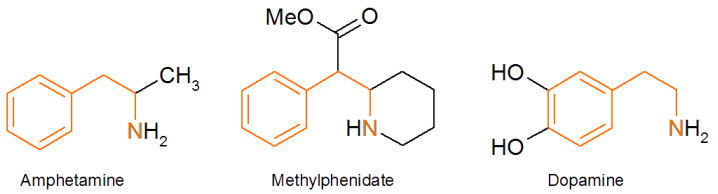


Figure 6

Amphetamine (Figure 6) first came up in the late 90s of the 19th century. It was synthesized in Germany at the 'Berliner Humboldt-University' during a dissertation. It took almost 50 years until the stimulating properties of the compound were brought to light and documented by Gordon Alles and the rest of his group.(Fleckenstein et al. 2007; Tauss 2008)

Since then Amphetamine and its derivatives have been used for the treatment of many different diseases, including narcolepsy, ADHD and obesity (Seiden et al. 1993). But the treatment of illnesses wasn't the only purpose the drug has been used for. Not only during the Second World War it was used to 'improve the performance' of soldiers during the fight (Tauss 2008). Lately it was proposed in a Nature article that the use of amphetamines could be a benefit in a university environment to enhance the cognitive abilities of the users:

> "*Based on our considerations, we call for a presumption that mentally competent adults should be able to engage in cognitive enhancement using drugs."*(Greely et al. 2008)

Last but not least this substance family is like cocaine one of the most widely abused drugs. (Tauss 2008) Besides the high risk of dependence amphetamines can also cause many side

effects. They are mostly congruent with the ones which were already mentioned in the methylphenidate section. (Mutschler et al. 2008)

If you look at the molecules (Figure 6) you can see easily that methylphenidate and amphetamine are quite closely related according to their structure. Interestingly, the mechanism of action is not the same. Methylphenidate like cocaine blocks the dopamine transporter. This is not the case with amphetamine. Anyway the effect is the same: the concentration of dopamine in the synaptic cleft increases significantly. There are different hypotheses for its mechanism of action. Studies suggest that the compound is transported actively into the cell by the DAT just like dopamine itself. There it inverts the direction of transport and the transporter pumps dopamine out of the neuron. Another hypothesis indicates that amphetamine boosts the excretion of dopamine vesicles. Most likely more than one mechanism plays a role. (Pifl et al. 1995; Fleckenstein et al. 2007)

Amphetamine also has a positive effect on the release of serotonin and noradrenalin (via NET and SERT). The quantity of released serotonin is much lower than that of the other neurotransmitters (Rothman and Baumann 2002).

- Amphetamine derivatives

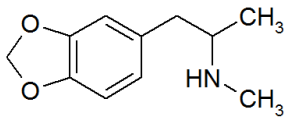


Figure 7 MDMA

One of the most prominent derivatives of amphetamine is MDMA (3,4- **M**ethylen**D**ioxy-N-**M**ethyl**A**mphetamin) which was already synthesized in the early 20th century. It became very popular in the 1980's and 90's under the name Ecstasy, an illicit 'club drug' (Benzenhofer and Passie 2006; Senn et al. 2007). However, you can also find other amphetamine derivatives and other stimulating drugs in ecstasy tablets. Consumers can never be sure which - or how much - active substance is included (Parrott 2004).

In contrast to amphetamine MDMA is about tenfold more active at SERT and NET than at DAT transporters (Rothman and Baumann 2002). It was the first member of a new group of drugs:

the entactogens. MDMA has not only stimulating but also hallucinogenic attributes which differentiates this family from the classical amphetamines.

1.1.4 Attention Deficit Hyperactivity Disorder (ADHD)

As ADHD is currently a very important indication for methylphenidate (and also for amphetamine) a few facts about the disease should be given.

Normally the disorder arises in children before the age of six. Persons with ADHD are not able to keep up attention and focus on something. They are easily detracted and because of their permanent agitation these people are often referred to as 'Fidgety Philip'. It's hard for them to plan their activities and think foresightedly. By definition you only speak of ADHD when the symptoms show over a period longer than 6 months and occur in at least two different areas of life (for example at school and at home with family). Also the disease is quite often accompanied by other psychic syndromes like depression or anxiety disorders.

In general the therapy is very important. When the disease is not treated children can easily have deficits in their social and intellectual development. Different factors are important for a successful therapy. First, it is very important to make the child's environment aware of the circumstances and also give them further advice how to interact with the kid. Secondly the patient should undergo psychotherapy. Finally, this can also be combined with pharmacotherapy. (Gerlach et al. 2009)

For the medicinal treatment methylphenidate and amphetamine are the first choice. Methylphenidate is favored as long as the effect is strong enough. If that is not the case amphetamine is used. Both drugs are well capable of reducing the symptoms of ADHD. In the last years retarded versions of the drug were more often used. They show the benefit of better patient's compliance. (Gerlach et al. 2009)

Other drugs used for the treatment: Atomoxetin can be considered as second choice (only for special indications as a first choice) and other stimulants and antidepressant drugs as third choice. (Gerlach et al. 2009)

2 Computational Background

There are different ways to analyze interactions between a compound and a biological target (protein). You can separate them into two larger groups: Ligand and target based methods.

2.1 Ligand based Methods: Quantitative Structure Activity Relationship

2.1.1 Theory:

The idea of QSAR is it to differences in the chemical structure with differences in the activity/affinity of a compound. This can be used for different purposes, on the one hand for in silico virtual screening of large databases of compounds to make predictions which of them could be interesting for further testing. Hereby one can save money and time. On the other hand you can use the results of a QSAR analysis to see which substructures and chemical properties of a compound could be of importance for the interaction with the target and which of them are weakening it. The big advantage of this computational method is that you do not need to know the structure of the target to make certain assumptions about the interaction pattern. But even if you do know the structure the combination of structure- and ligand-based methods can be beneficial. (Klebe 2009)

2.1.2 Requirements:

Compounds can only be compared by a QSAR analysis if they are chemically quite similar and of course it's only possible to compare substances in regard to their affinity if they all interact with

the same target. It is necessary to have a series of compounds with experimentally determined affinity data for the target protein.

2.1.3 Descriptors:

In order to be able to compare molecules you need to be able to describe their chemical properties. A QSAR model is a purely mathematical model, so you need to convert chemical features into numeric values. This is what descriptors do: they define certain chemical features numerically and thus make them quantifiable.

"The molecular descriptor is the final result of a logical and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment." (Todeschini and Consonni 2000)

A few examples:

logP (o/w):	Logarithmic partition coefficient octanol/water
weight:	Molecular weight of a molecule
a_acc:	Number of H-bond acceptors in the molecule
Pc+:	Total positive partial charge

(Chemical Computing Group 2009)

Descriptors can be calculated quite rapidly. The calculation of a small number of descriptors is based on experimental determined data, but most of them are calculated right away if the molecule is given (for example the number of atoms).

2.1.4 Hansch-Analysis:

In 1964 Hansch and Fujita developed the so called “Hansch-Analysis”. This was the first real “QSAR Equation”. It is a multiple linear regression model which gives an equation as folows:

$$\text{Log } 1/C = k_1*(\text{DesA}) + k2*(\text{DesB}) - k_3*(\text{DesC})\ldots$$

This equation is normally accompanied by the following statistical terms:

“n” is the number of compounds

“R” is the correlation coefficient

“S” is the standard deviation of the model

“Log1/C” is the logarithmized reciprocal of the calculated activity (The EC_{50} or IC_{50} value). “DesX” is a certain descriptor and “k_x” is the regression coefficient which shows the amount of contribution to the model of each descriptor.

2.1.5 HQSAR:

1. **The Method**

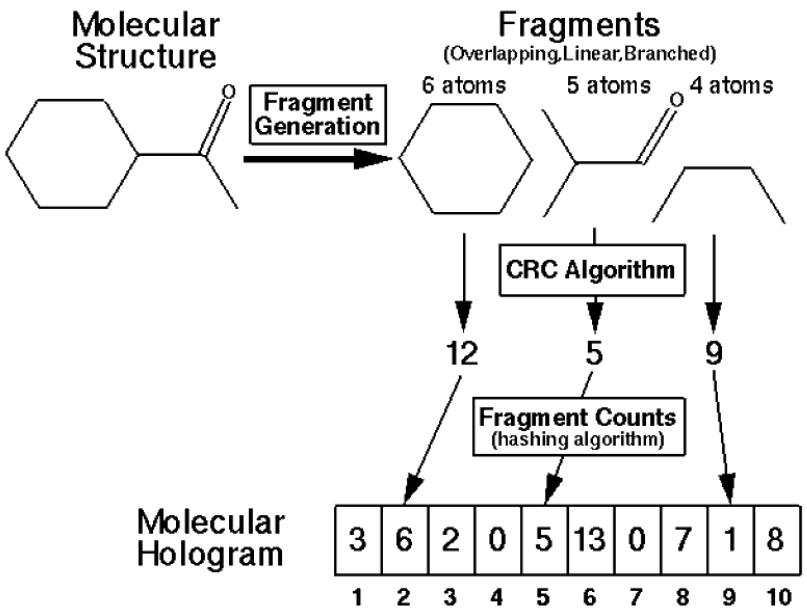
HQSAR is a fragment based QSAR method available in the Sybyl software package by (Tripos™Inc. 2007). It correlates features of substructures of a molecule with biological activity. It is a fast and easily applicable method which only requires 2D structures as input format and the activity data of these molecules. In contrast to normal 3D QSAR no alignment is necessary.

The analysis consists of two steps:

1. Creation of the molecular holograms

2. Correlation of the hologram with the biological activity via PLS

2. **Details**



First, molecules are divided into all possible fragments (of connected atoms) of a certain length. Subsequently they are transformed into molecular holograms, strings of integers, via two different algorithms (Figure 8). (Lowis 1997)

Figure 8 taken from (Lowis 1997)

The hologram encodes all unique fragments, also branched and cyclic and overlapping ones.
The fragment and the hologram length can both be set by the user which affects the quality of the model.
The user can choose six different features which define the fragment as unique. These featured are listed as follows:

1. Atom: Based on the element types
2. Bond: Based on the bond types
3. Connection: Based on the hybridization state
4. Hydrogen: Hydrogen Atoms are included or not

5. Chirality: Based on chirality
6. Donor & Acceptor: Searches for predefined donor or acceptor atoms

Stereochemistry can be considered as well. By taking these features into account one can see that 3D structural information is included to a certain extent, which differentiates the method from 2D QSAR analyses. (Tripos^{TM}Inc. 2007)

The holograms of the different fragments are then correlated to the given activity data via PLS (Partial Least Squares)

2.2 Structure based Methods: Small ligand docking

Small ligand docking is a method used in structure based drug design. The idea of structure based design first came up in the early 1970s in London in the group of Chris Bedell and Peter Goodford. They tried to investigate new agents for the DPG (Bisphosphoglycerate) binding site of Hemoglobin, which at that time was the only known protein structure known to be responsible for a disease, namely the Sickle-cell disease.(Klebe 2009) Though the methods and the possibilities have changed greatly since that time, their work laid the cornerstone for structure based in silico methods.

One of today's aims of small ligand docking is to deduce predictions from the properties of the protein surface concerning the interaction mode of a ligand with the target structure. Basically in SLD a small molecule is positioned in the binding pocket via a placement method and after that the docking program calculates if the interactions are energetically preferred or not. These calculations are done by a scoring algorithm which uses different energy terms. The given score makes the single ligand poses comparable and analyzable. Another application area for docking is the screening of ligand databases versus one target protein. As the screening approach was not used in this work it will not be described in further detail.

It is mandatory for a docking process that some facts are known about the target's structure.
The characteristics of the protein have to be studied in detail before a docking job can be started. You need a 3D conformation for analysis. In the optimal case the protein structure is known from X-ray diffraction measurements. These can depict the exact structure of the protein in a certain conformation. Often this is not possible as there are no X-Ray structures available. There are alternatives to creating a model structure from closely related protein families via alignment and further processing. These models are called homology models.

If the interacting partner amino acids for the ligand in the receptor have been uncovered, one can try to make predictions for the effect of structural modifications of the ligand. One can try to increase the affinity of the ligand or just design optimal ligands for this binding site. This is called rational structure based design.(Klebe 2009)

2.3 Aim of the Work

The detailed mechanism of interaction of methylphenidate in the dopamine transporter is still unknown. Although many structural derivatives of the compound were tested in terms of affinity to the dopamine transporter no residues of the protein have been investigated to be crucial for this interaction.

It was the aim of this work to investigate the binding mode of methylphenidate in the transporter.

To achieve this, as a first step, a set of methylphenidate derivatives will be analyzed in QSAR experiments. The developed QSAR models should not only be useful tools for the prediction of ligand affinity. Here these models should additionally be used for interpretational purposes.

It will be checked if structural influences on DAT affinity values - observed in the ligand based analyses - can be explained by interactions proposed in structure based models which will be created as a second step. This will be done in docking experiments in a homology model of the Dopamine transporter with flexible side chains.

By this combination of structure and ligand based methods the inevitable uncertainties which are associated with a docking experiment using a homology model as a protein structure, should be compensated to obtain a reliable methylphenidate binding pose in the Dopamine transporter.

3 Methods

3.1 Creation of a Ligand Database

First a database of the methylphenidate derivatives with available affinity data for the dopamine transporter was created. This was done by an intensive screening of the literature, which resulted in a final dataset of 135 derivatives in total. Some of the recorded compounds only differ stericly: the threo or erythro form or different enantiomers. The database was generated in the MOE Package using the builder tool to create the chemical structures.

Each entry got a name which makes it possible to trace its origin back to the appropriate study.

Example: ***P A2*** **RR 10h**

The first three characters indicate the source of the compound. “P” stands for Paper/Study and A1-A7 is an internal numbering. The following letters are the designation of the compound used in the study.

The list of the sources for the database:

P A1	(Schweri et al. 2002)
P A2	(Froimowitz et al. 2007)
P A3	(Davies et al. 2004)
P A4	(Meltzer et al. 2003)
P A5	(Kim et al. 2007)
P A6	(Gatley et al. 1996)
P A7	(Deutsch et al. 1996)

3.2 Creation of 2D-QSAR and H-QSAR Models

3.2.1 2D-QSAR Model with broad structural range

First of all the dataset had to be adjusted to the needs of a 2D-QSAR analysis. As enantiomers cannot be differentiated when 2D descriptors are used, only one enantiomer of each compound including its activity data was kept in the database. Furthermore, the data of threo-diastereomer was preferred to that of the erythro-diastereomer, since the former are described as the active ones in literature (Gatley et al. 1996), which is quite obvious when you look at the activity data of these compounds.

The given dataset was divided into a test and a training set before the creation of the QSAR Model. 20% of the compounds where selected via the MOE tool “Diverse Subset”. This feature ranks the compounds according to their distance to each other in diversity metric. (Chemical Computing Group 2009) So the most diverse agents were selected automatically for the test set. The remaining 80% of the dataset were used to create a training set. 76 methylphenidate derivatives formed the training set and 18 the test set.

Descriptors which describe basic molecular properties where calculated using MOE. Examples of the descriptors used: Number of Rotatable Bonds, Molecular Weight, Molrefraction, Number of Hydrogen Bond Acceptors/Donors, ClogP(o/w) and others.

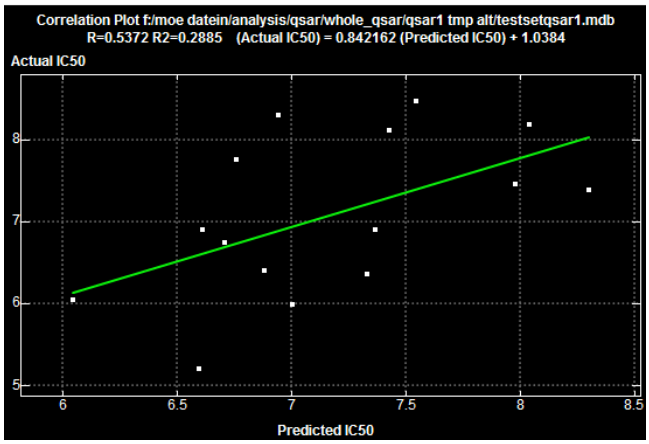


Figure 9 Poor correlation using global 2D-QSAR models

In a second run a series of VSA descriptors were used for the creation of the model. In addition to that an Auto QSAR Script, which evaluated different descriptor combinations automatically via PLS analyses, was used. These automatically created models are not perfect, but they do give you a clear hint to some descriptors which could be useful for

the model of the correspondent series of compounds. Nevertheless, even with extensive optimization of the model no optimal equation could be obtained. The final model was applied to the test set. As can be seen in the correlation plot on the left side (Figure 9) the correlation of the predicted values with the actual ones is not very high.

As it was not possible to create a predictive model for the whole derivative dataset the number of the included compounds was reduced.

3.2.2 2D-QSAR Models with smaller structural range

Selected studies were investigated individually to have a smaller range of structural diversity. So the influence of small structural changes on the biological activity could be analyzed in detail.

In consideration for the size of the datasets it was decided to validate the models via leave one out cross validation, as the size of a test or training set would not have been big enough and so not very significant for testing the model.

Model 1:

First compounds of study A7 (Deutsch et al. 1996) were used to determine how different substituents at various positions of the methylphenidates phenyl ring (Figure 10) affect the affinity.

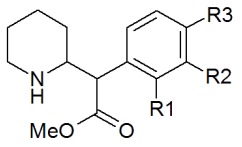


Figure 10

For this QSAR analysis not only descriptors calculated in MOE were used for the equation but also Sigma Hammett, Pi , Es and MR constants from "*Substituent Constants for Correlation Analysis in Chemistry and Biology"(Hansch 1979).*

The Model was created on basis of 24 compounds from study A7 which are listed in section (*G Compounds used for the QSAR Models*)

Model 2:

Compounds of study A2 were analyzed next. In these compounds methylphenidate's carboxy methyl ester group is replaced by different more lipophilic alkyl groups. The compounds differ in size and length of these alkyl groups (Figure 11: R1).

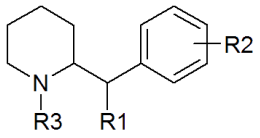


Figure 11

The descriptors used in this model all were calculated within the MOE Software Package.

The 25 Compounds used for the creation of the model are listed in section (*G Compounds used for the QSAR Models*)

3.2.3 Creation of the H-QSAR models:

To reinforce the models also H-QSAR analyses of these datasets were accomplished with the Sybyl 8.0 Software Package (Tripos^{TM}Inc. 2007).

Around 40 primary models were created for each final model using different combinations of 2 to 4 of the 6 available H-QSAR descriptors (Atom, Bond, Connection, Hydrogen, Chirality ,Donor & Acceptor) and the default fragment size of 4-7 atoms. The results were validated via leave one out cross validation. The one with the highest cross validated R^2 value was chosen to be processed further. After that, different models were calculated with the determined descriptors, only the fragment size was systematically changed. Again the optimal one was chosen. This is the final model. The aim was to gain activity predictions with residual values below one order of magnitude.

Combination of successful H-QSAR models

As a last step to verify whether it is possible to create reasonable H-QSAR Models via compounds with a broader structural diversity range, two additional models were created. One included the compounds of study A2 and A7 and one included the ones of study A2, A5 and A7.

Compounds study A5:

The compounds analyzed in this study were more diverse than the ones of the other two studies. Anyway it focuses on restricted rotation analogues with a substituted scaffold as it is shown in (Figure 12).

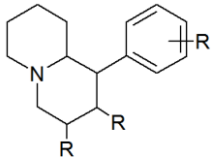


Figure 12

The detailed development of the QSAR models as well as the outcome of these analyses are described in the result section.

3.3 Docking Basis and Workflow

3.3.1 The Dopamine transporter model

Since currently no crystallographic structure of the Dopamine transporter, determinated by X-Ray diffraction, is available, a homology model had to be created. The model of (Weissensteiner 2008) was used. This model was created on basis of the high resolution (2Å) crystallographic structure of the LeuT of "Aquifex Aeolicus" published by (Singh et al. 2008) (PDB: 3F3A). The X-Ray structure shows the protein in an open to out conformation state. The bacterial Leucine-Transporter is an established model for the mammalian Neurotransmitter-Na^+-Symporters transporter family structurally matching with the Dopamine Transporter. (Weissensteiner 2008)

3.3.2 The Binding Sites

Up to now no binding sites for Methylphenidate in the Dopamine Transporter have been published. Nevertheless, there are two prominent and recent poses for cocaine in the DAT, published by the group of Beuming (Beuming et al. 2008) and one by the group of Huang (Huang et al. 2009), which are depicted in Figure 13. Due to the similar type of interaction of Cocaine and Methylphenidate at the transporter, as already mentioned in the "Biological background" section, the proposed cocaine binding modes where taken as a clue for a possible binding pocket for Methylphenidate. While the binding site proposed by the group of Beuming overlaps with that of Dopamine and Amphetamine (Beuming et al. 2008) the initial one of Huang's group is close to the Dopamine site but does not overlap.(Huang et al. 2009) Huang further states that a later move of cocaine to the Dopamine binding site after a conformational change of the transporter protein, resulting in an expansion of the binding pocket, is possible.

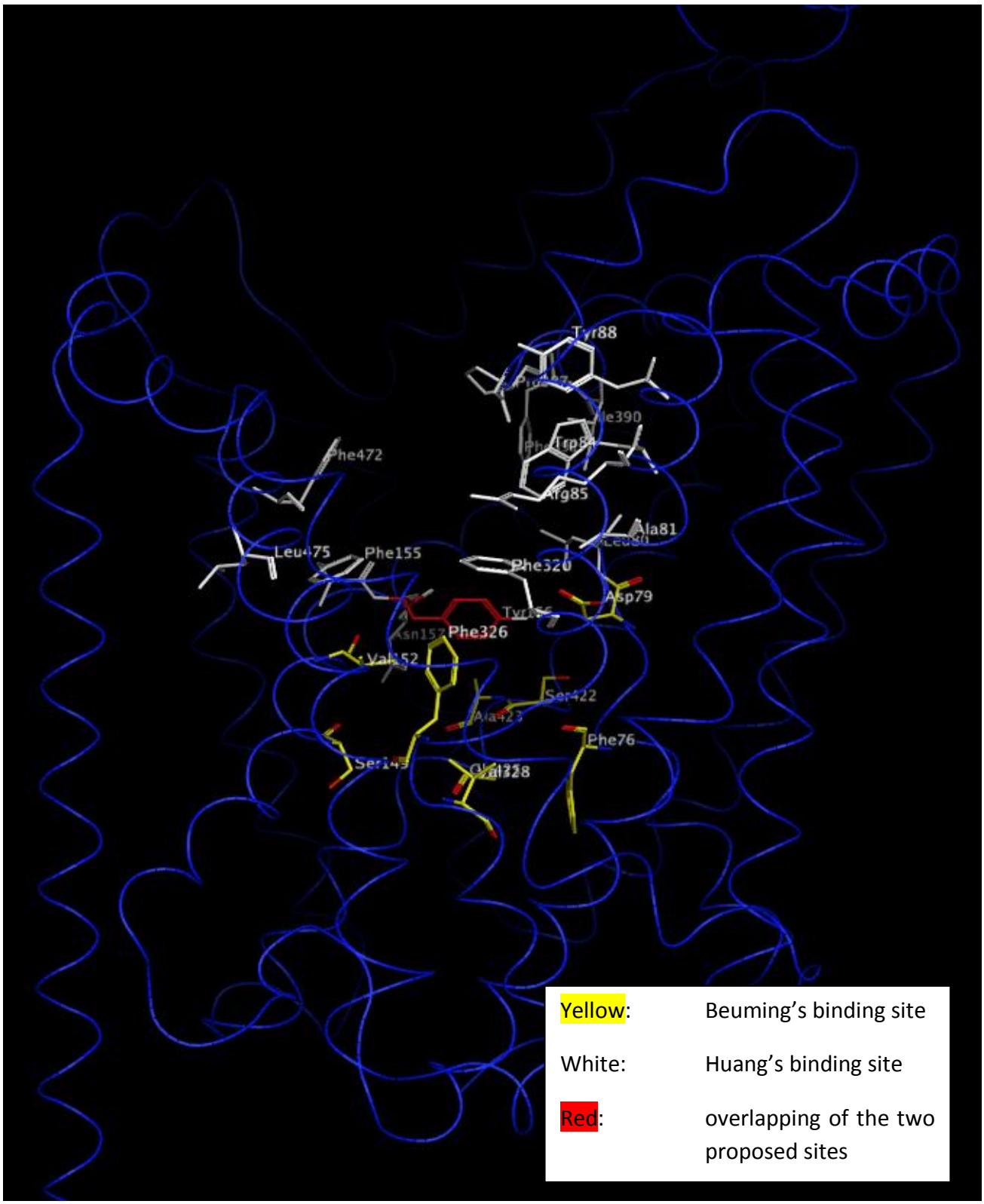

Tyr88
Ile390
Phe472
Arg85
Ala81
Leu475
Phe155
Phe320
Asp79
Phe326
Ser422
Phe76
Yellow: Beuming's binding site
White: Huang's binding site
Red: overlapping of the two proposed sites

Figure 13

Due to extensive experimental validation of the computational model of (Beuming et al. 2008) this suggested binding pocket was preferred over the other one. The model was not only validated in mutagenesis experiments but the authors also trapped the radio labeled cocaine analogue [H^3]-CFT in the binding site over cysteine cross linking.

The following amino acids were identified to form the cocaine pocket:

PHE76 ASP79 SER149 VAL152 TYR156 Phe320 PHE326 VAL328 SER422 ALA423 GLY425

3.3.3 The Compounds

Five highly potent Methylphenidate analogues and Methylphenidate itself were chosen for the docking runs, four alkyl and one ester derivative analogue. (Figure 14)

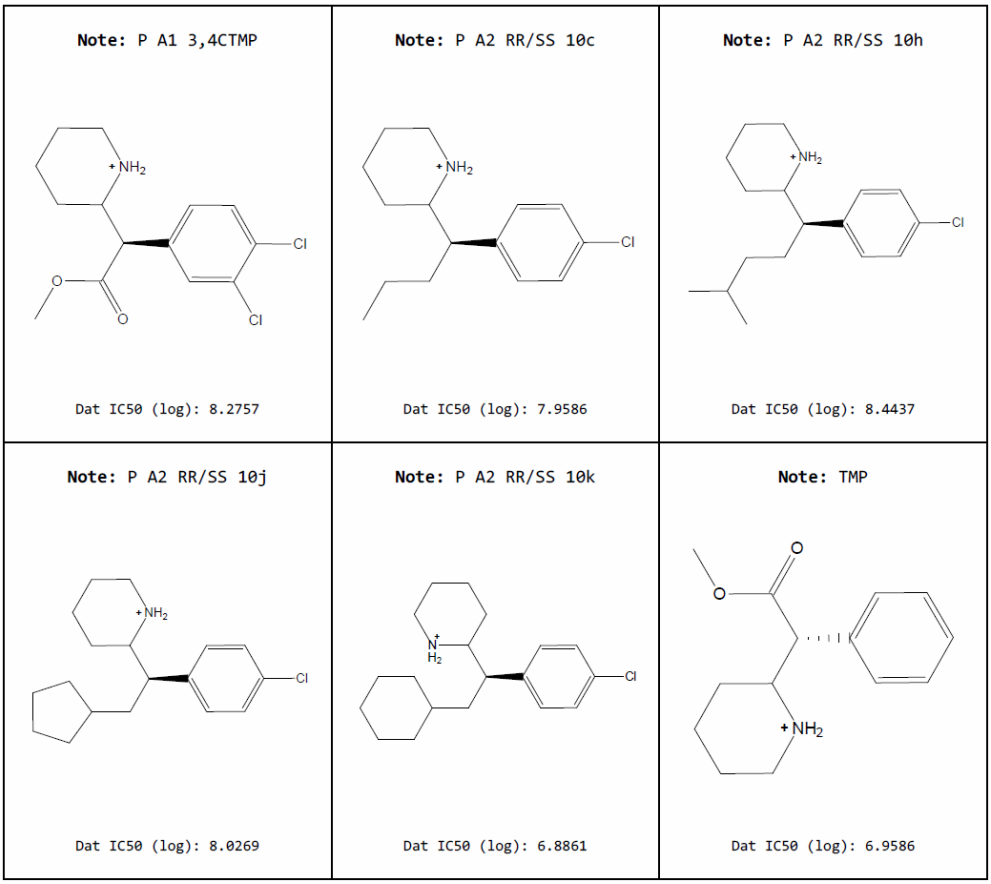


Figure 14

Though only the affinity data for the Threo-racemates (d/l) of these methylphenidate derivatives are available in literature you have to choose a distinct steric conformation of the molecules for the program to work with. As the d-threo form of methylphenidate is described as the more active inhibitor (Gatley et al. 1996) this form was chosen for the docking runs.

3.3.4 Docking with GOLD

The Gold Suite 4.1.1 (Hermes) for Microsoft Windows© was the first program used for the docking experiments. To validate if the program and the method were suited for the following docking experiments a redocking was performed:

The crystal structure complex (3F3A) of LeuT and L-Tryptophan was used for the validation.
As one can see in Table 2 among the first twenty top scored poses 18 are within an RMSD of under 0.5 Å to the reference ligand, which is a good value. So the system was considered to be suited for the following docking experiments.

Nevertheless, Numbers 19 and 20 of the score show that trusting in scoring functions alone is not recommendable. Good scores do not always correlate with right docking poses and so other aspects should be taken into consideration as well when different poses or clusters are evaluated.

Index	Gold Score	RMSD
1	48.508	0.330
2	48.327	0.351
3	48.327	0.413
4	48.234	0.440
5	48.227	0.475
6	48.216	0.315
7	48.021	0.405
8	47.931	0.341
9	47.705	0.330
10	47.621	0.383

Index	Gold Score	RMSD
11	47.604	0.345
12	47.585	0.446
13	47.495	0.284
14	47.446	0.338
15	47.395	0.464
16	47.259	0.392
17	47.040	0.342
18	47.005	0.371
19	46.206	5.023
20	45.201	5.056

Figure 15

3.3.5 Final Workflow: Docking with flexible sidechains in GOLD

1. First the protonated methylphenidate derivatives were docked into the protein.

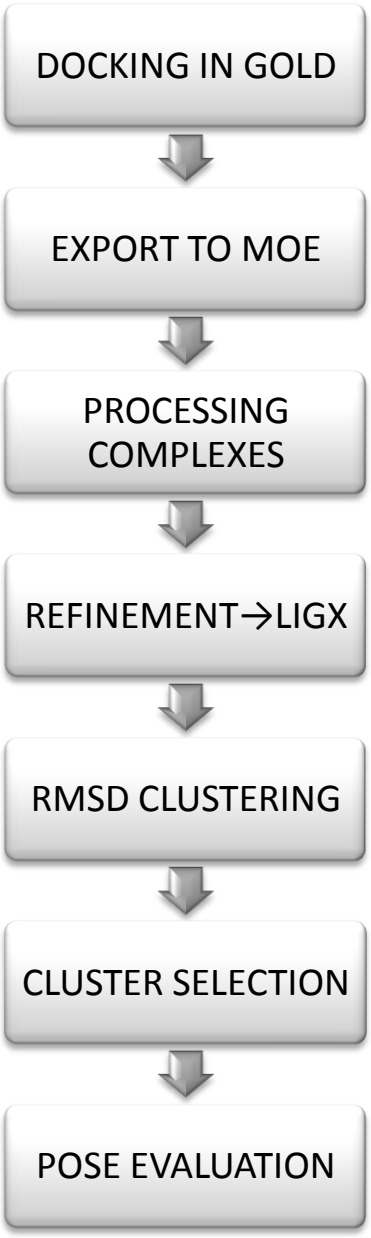


Figure 16

Settings:

First the protein was prepared in GOLD Setup. Hydrogens were added and water molecules deleted. The docking site was defined by a text file, listing the already suggested amino acids.

The side chains were set to be flexible via the creation of different rotamers. The program created the rotamers using a library.

The number of poses was set to 150 for each of the six compounds to cover a wide range of possible positioning.
Furthermore the possibility of early docking determination was cancelled to receive a maximum of poses.
The ligand-input-file (an sdf database file) and also an output file were defined. Then the docking run started.
After the docking run was finished the poses had to be evaluated. The following steps are done in MOE.

2. The data had to be exported and processed so they could be used and interpreted in the Molecular Operating Environment software, which is not a trivial task.

The ligand-protein complexes had to be exported from the GOLD output sd file to an mdb file. While the positioning and the coordinates of the ligands were saved in a molecule database field, the structural changes of the DAT protein resulting from the flexibility

of the side chains were written to a coordinate matrix (Figure 17) by GOLD. The program used them in combination with the primal Protein input file to depict the new altered proteins. MOE cannot interpret the coordinate matrices by itself.

However, a Python script ("apply_rotated_atoms.py") is provided with the GOLD Suite package (CCDC 2009) to apply these matrices to the protein model. This script exports the new conformations into mol2 files, one file per ligand. All these mol2 files had to be transferred to an sdf-database file. Afterwards, the protein field has to be combined with the correspondent ligand field for the interpretation process.

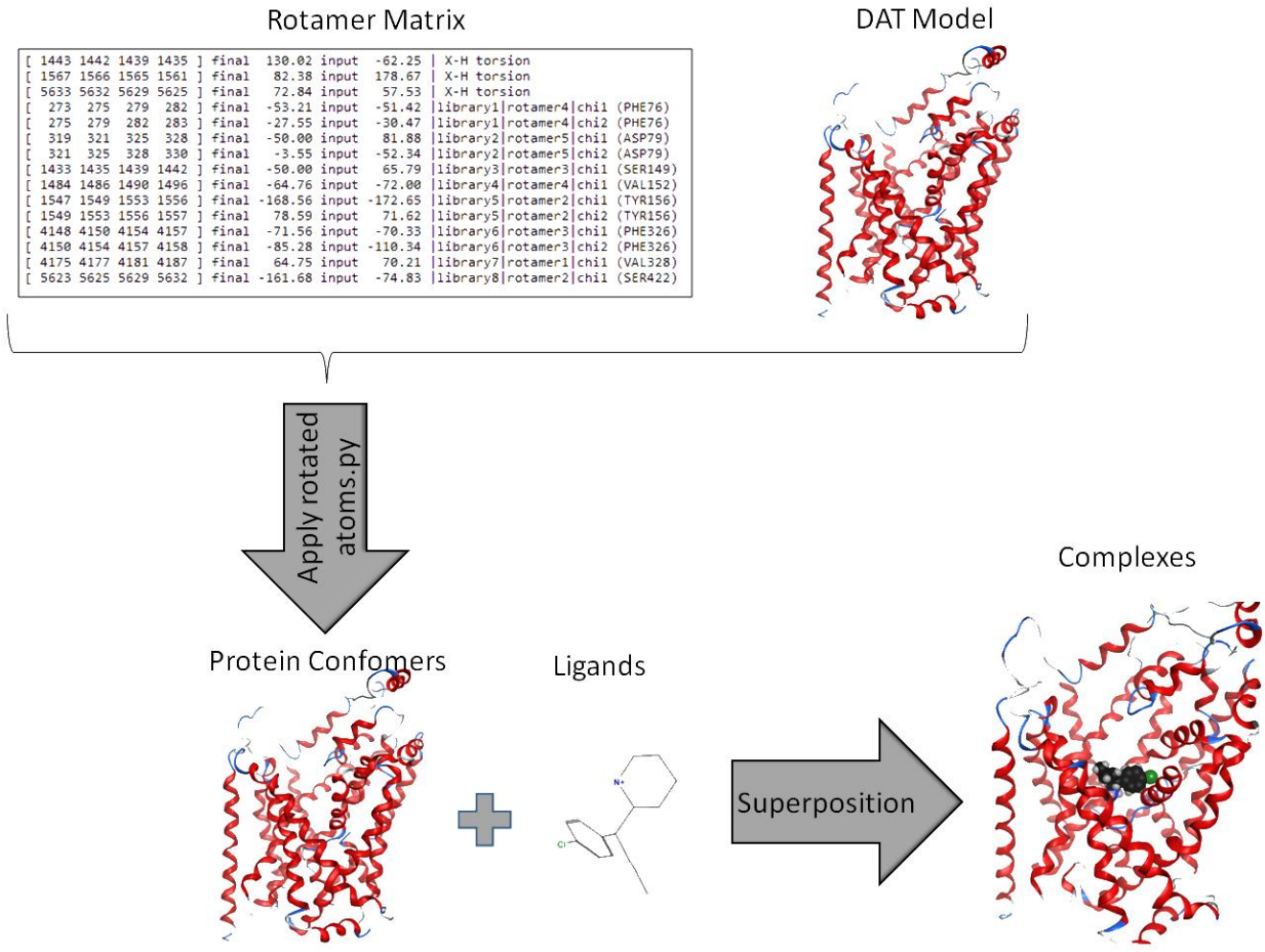


Figure 17

3. Processing the Complexes

As GOLD uses different methods to calculate receptor ligand interaction, lone pairs, which are created by the GOLD during the docking process, had to be deleted and charges had to be recalculated. This was done via the function “Protonate 3D” in MOE. This function does not only add hydrogen atoms but also calculates ionization states and partial charges. The calculations are based on user settings, defining the protein environment with different parameters like temperature, pH, salt concentration and the dielectric constant. Standard settings have been used.

4. Minimization of the Complexes

LigX, a program implemented in MOE, offers a function to minimize the pocket and the ligand using a forcefield.

The three steps of deleting the lone pairs the protonation step, and the complex minimization were scripted to speed up the process.

5. Clustering of the poses

The poses were clustered according to the RMSD of their maximum common substructure.

To do this, first the molecules have to be divided into fragments. The maximum common substructure has to be defined by the user. Then the MSC is copied to a new

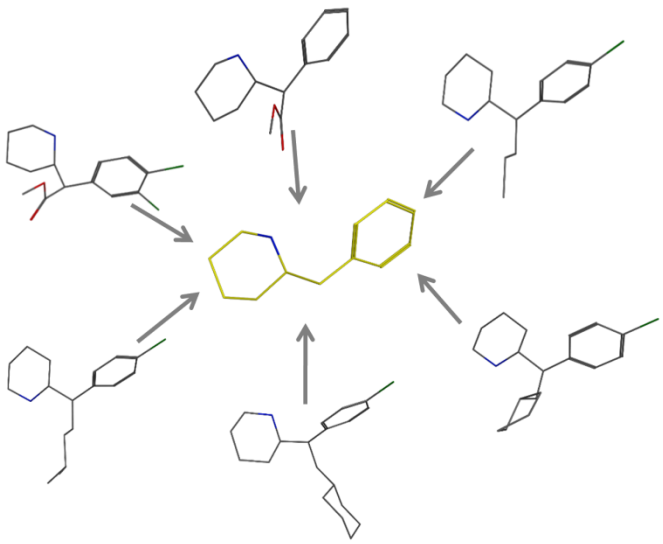


Figure 18

mol field for each ligand maintaining the positional information. This new database field is used by the clustering script during the process.

Then the poses are clustered hierarchically. A threshold can be set by the user, specifying the maximum RMSD deviation for the poses within one cluster. All poses that exceed this limit are assigned to different clusters.

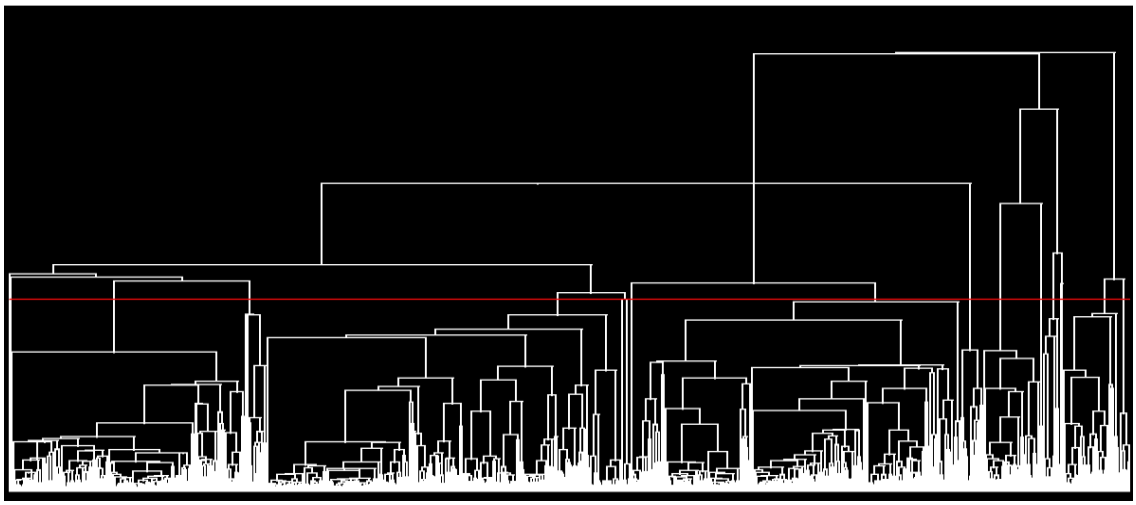


Figure 19

6. Cluster selection

The given pose clusters had to be evaluated and their number had to be limited down to make them interpretable and to find the correct one. As a first criterion, only clusters were considered which contained poses of each of the six compounds. All docked ligands belong to the same structural class and have high affinity to the transporter. Therefore one can assume that they interact with the protein in the same manner.

There are different possibilities to weight the remaining clusters. This can either be done computationally using scoring functions, by the size of the cluster, or experimentally by comparing the protein-ligand interaction fingerprint (PLIF) with mutational data.

Steps until the final workflow was obtained

Three docking runs were performed with GOLD 4.1.1 before the optimal workflow was obtained.

First run:

Unprotonated ligands were used for the first docking run. The output number was set to 20 poses per ligand (120 poses in total). The protein itself was rigid. No minimization was applied after docking. The poses were clustered according to their common scaffold with a threshold of 1.9 Å. This resulted in five different clusters.

Cluster number five was by far the biggest one, containing 63 of the 120 poses and five of the six docked compounds.

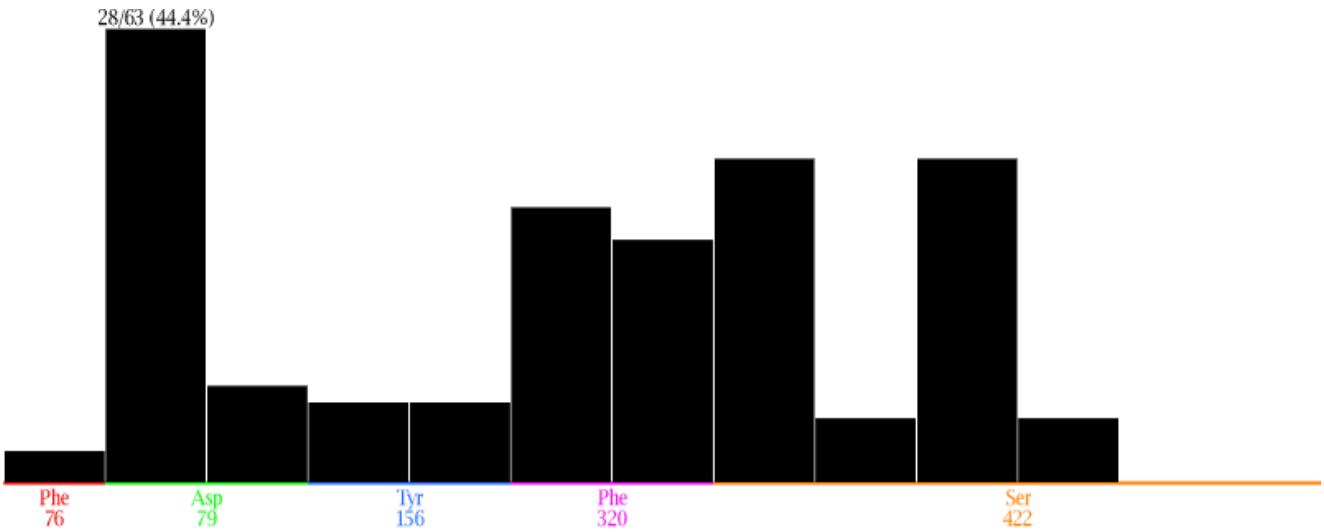


Figure 20

In this model only protein-ligand interactions with the nitrogen atom of the methylphenidate derivative could be observed. The PLIF (Figure 20) shows the amino acids which interact with this basic group. Asp 79, which is also described as the essential interaction partner for the amino group in cocaine (Beuming et al. 2008), is an important interaction partner for the amine nitrogen of methylphenidate and its derivatives, other significant interaction partners are Phe320 and Ser422. The pose of methylphenidate itself differs a little from the other compounds in this cluster (Figure 21 in red). That was also the reason why it is the main interaction partner of Ser422 in this cluster.

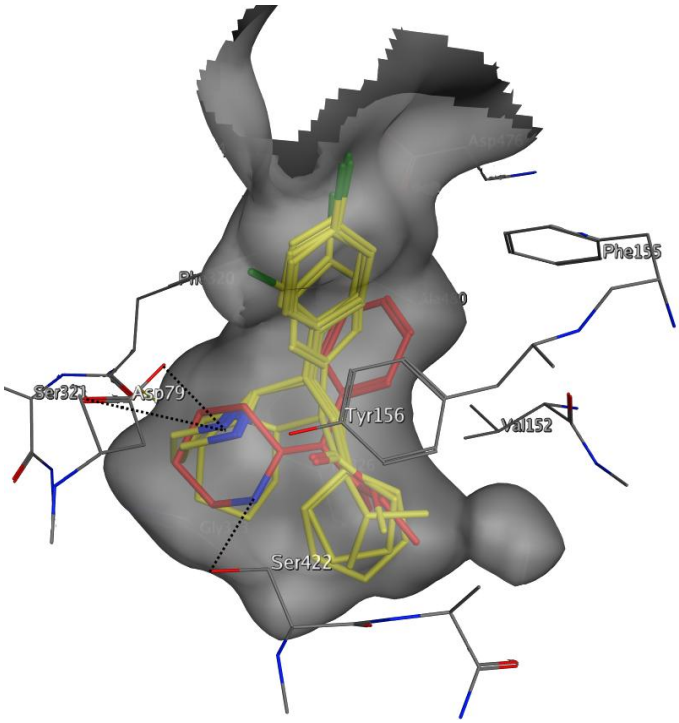


Figure 21

The cluster did not only include the highest number of poses and compounds, it also contained the poses which were ranked best with the GOLD score.

The **GOLD score** is calculated considering four elements:

- Energy of hydrogen bonds between receptor and ligand
- Van der Waals energy of receptor and ligand
- Internal vdw energy of the ligand
- Energy of Torsional strain of the ligand

The score value arises out of the negative value of the sum of the individual energy terms → the larger the fitness score the better the positioning. (CCDC 2010)

"GOLD Fitness = $S_{hb\_ext} + S_{vdw\_ext} + S_{hb\_int} + S_{vdw\_int}$"

(Marcel et al. 2003)

However, the cluster did not contain any pose of the compound “P A2 RR 10c” (Figure 22). As the number of poses per ligand was rather small, namely 20 poses, this was considered to be one possible reason for the absence of this compound.

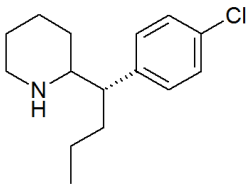


Figure 22
Comp “P A2 RR 10c”

Second run:

To find a cluster which contains all six compounds the pose number per compound was increased to a hundred. Other parameters weren’t altered in this second run.

Clustering the compounds led to eight different clusters (1.7 Å RMSD deviation). One of them contained poses of all six compounds and was analyzed further.

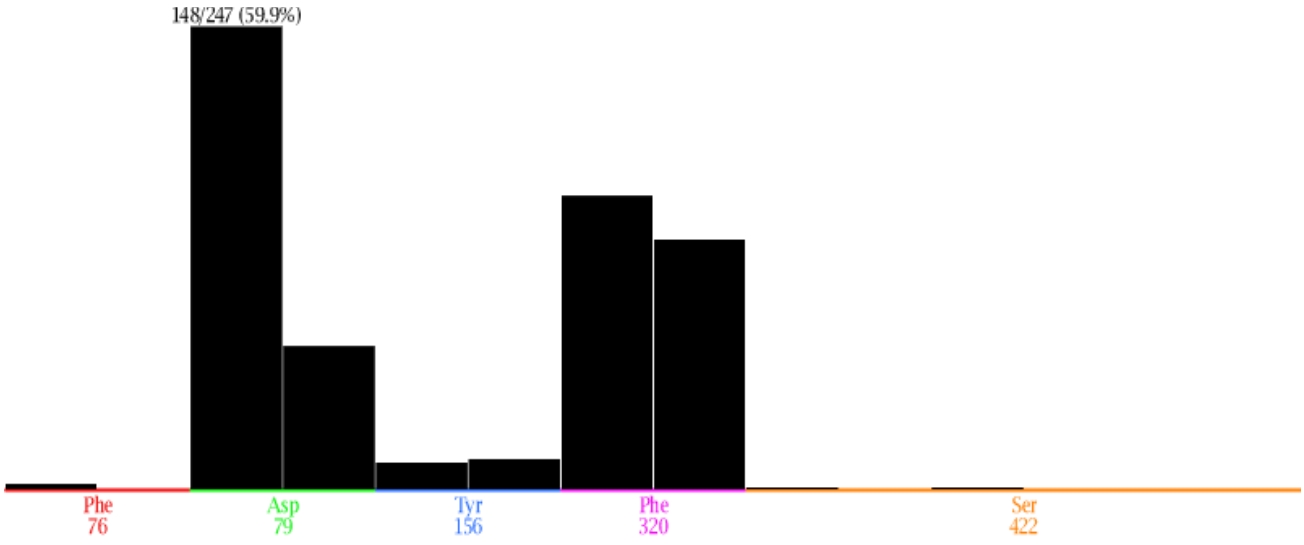


Figure 23

The PLIF of the Cluster (Figure 23) shows rather similar interaction compared to the first docking run with Asp79 and Phe320 interacting with the amino group of the ligands. Interestingly, nearly no interaction is observed with Ser422. This can be explained with the difference of the methylphenidate pose, as the former Methylphenidate’s pose was the main

interaction partner for Ser422 in the GOLD docking run number 2 (see Page 39 Figure 21). In contrast to the last run the threshold was lowered from 1.9 to 1.7Å RMSD deviation. So the other pose of methylphenidate - seen in Figure 21 - was assigned to another cluster. Now the position of the methylphenidate poses overlapped with those of the derivatives.

3.1.4 Docking with MOE

This docking run in the molecular operating environment suite was accomplished to compare its poses with the docking results which were received with the GOLD suite. If these concurred with each other it would be useful for the reinforcement of the GOLD poses and also help to choose the preferred one. As it was mainly of interest if the same poses could be found, MOE's standard docking settings were used without an extensive check of all available placement scoring functions.

The detailed settings and the results are described in detail in the result section "4.2.2 Docking with MOE".

4 Results

4.1 QSAR – Ligand based methods

4.1.1 2D-QSAR Model phenyl ring substituted analogues (A7)

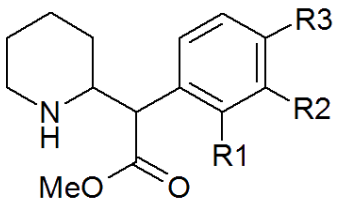


Figure 24

As already mentioned in the method section the 24 compounds used for the creation of this model mainly differ in the substitution of the aromatic ring system. The three positions shown in Figure 24 are occupied by various residues: In most of the cases halogens (Cl,Br,J,F), hydroxyl- or methoxy groups are attached to one or two of them. There are also compounds with an amine- a nitro- and a tert.But rest attached.

Two kinds of descriptors were used for the creation: on the one hand global ones which described the whole molecule and on the other hand substitution constants only describing the contribution of the single substituents to the activity.

Step 1: The descriptive substituent constants (Hansch 1979) Pi Ortho, Pi Meta, Pi Para, Sigma Ortho, Sigma Meta, Sigma Para, MR Ortho, MR Meta, MR Para, Es Ortho, Es Meta, Es Para were used.

RMSE:	0.44515
R2:	0.85876
XRMSE:	2.10569
XR2:	0.17379
Nr. C.	12

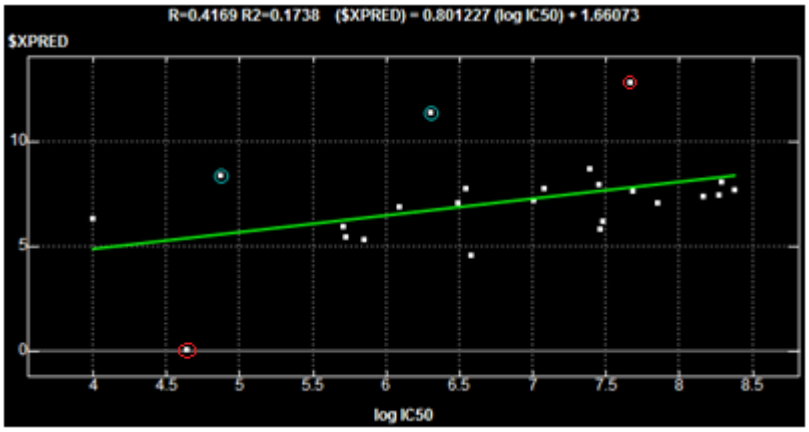


Figure 25: Correlation predicted/actual Step 1

Step 2: After an analysis of the outliers of Step 1 the parameter vdw_vol was calculated and added to describe the influence of substituents' size (e.g. tert. Butyl and -NO_2 → The compounds in Figure 25: Correlation predicted/actual Step 1 marked with white rings) at the aromatic ring systems more accurately which increased predictivity. The equation could describe two additional compounds only poorly: The ones marked in with red rings. One of them was hydroxylated in Ortho- and one methylated in Meta- position.

RMSE:	0.25409
R2:	0.95398
XRMSE:	1.06742
XR2:	0.50378
Nr. C.	13

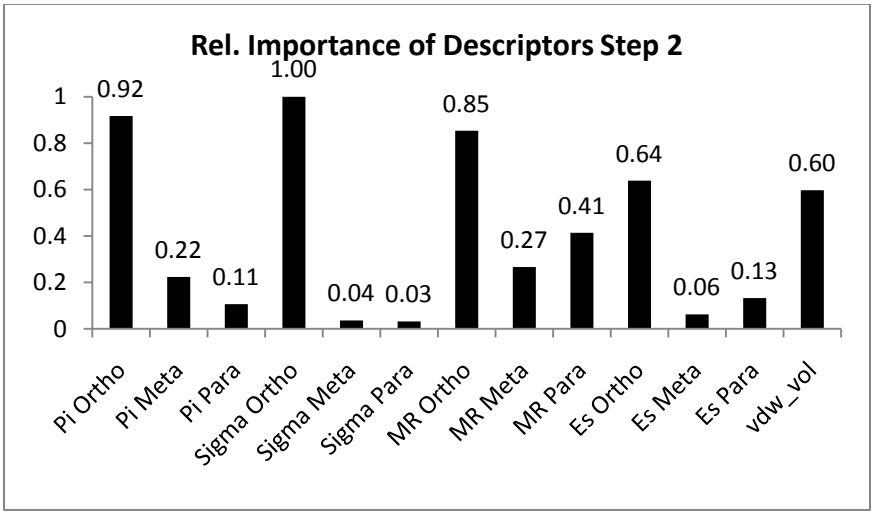


Figure 26 Rel. Importance of Descriptors Step 2

Step 3: 3 statistically unimportant descriptors were excluded via backward selection: Sigma meta, Sigma para and Es meta. The dimensionality was reduced via PLS to 9 components.

RMSE:	0.32619
R2:	0.92416
XRMSE:	0.99356
XR2:	0.54418
Nr. C.	9

Step 4: An automatic QSAR script was used to find other descriptors that could contribute to the model positively. Due to this analysis the Weiner Path Descriptor was added. The Component limit was set to 9.

RMSE:	0.30567
R2:	0.93340
XRMSE:	0.67503
XR2:	0.70596
Nr. C.	9

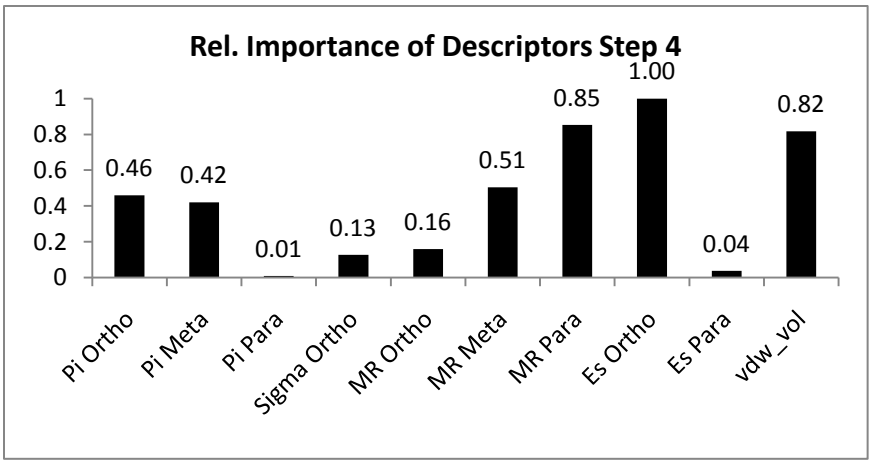


Figure 27 Rel. Importance of Descriptors Step 4

Step 5: Again 2 descriptors were excluded because of their insignificance for the model. Pi para and Es para. The Component Limit was set to 7.

RMSE:	0.30679
R2:	0.93291
XRMSE:	0.55290
XR2:	0.78593
Nr. C.	7

Step 6: A correlation matrix analysis (Figure 28) for the descriptors was performed to identify redundant descriptors. Sigma ortho highly correlated with Pi ortho and Es ortho with MR ortho. MR ortho and Sigma ortho were excluded. So the number of descriptors was reduced further and the component limit was set to 5. Although the Weiner Path and the vdw_vol descriptors show a high correlation too, the quality of the model decreases when one of them is left out. So they were both kept in the model.

RMSE:	0.31353
R2:	0.92993
XRMSE:	0.54889
XR2:	0.79232
Nr. C.	5

	1	2	3	4	5	6	7	8	9
1. Pi Ortho	100	-2	86	57	-11	-11	-59	3	-11
2. Pi Meta	-2	100	0	-4	29	0	5	10	0
3. Sigma Ortho	86	0	100	9	3	3	-20	2	-6
4. MR Ortho	57	-4	9	100	-26	-28	-89	2	-15
5. MR Meta	-11	29	3	-26	100	-23	31	23	15
6. MR Para	-11	0	3	-28	-23	100	33	73	65
7. Es Ortho	-59	5	-20	-89	31	33	100	12	25
8. vdw_vol	3	10	2	2	23	73	12	100	84
9. weinerPath	-11	0	-6	-15	15	65	25	84	100

Figure 28 Descriptor correlation matrix Step 5

Step 7: In a final step MR meta and MR para were combined additively to MR ges. 6 descriptors/ 5 components were used for the final model for this series of compounds. The additive combination of Pi Meta and Pi Ortho did not further improve the equation.

RMSE:	0.32685
R2:	0.92385
XRMSE:	0.51934
XR2:	0.81128
Nr. C.	5

The QSAR equation of the model:

log(1/EC50) = 23.801 +2.093 * Es Ortho -0.042 * vdw_vol -0.004 * weinerPath +1.919 * Pi Ortho +0.944 * Pi Meta +2.000 * MR Ges

Interpretation of the model:

The term "+2.093 * Es Ortho" indicated that big substituents in ortho position of the aromatic ring contributed negatively to the activity values. The Es value is the more negative the bigger the substituent gets.

The term "+1.919 * Pi Ortho" showed that substituents in ortho position contribute more positively to the activity if they increased the lipophilicity and they contribute more negatively if they increased the hydrophilicity. The same was valid for substituents in the para position of the compound.

Thus big and hydrophilic substituents contributed most negatively in ortho position.

Furthermore it seems that substituents which increase the overall molar refractivity of the phenyl ring system also increase the activity.

The descriptors vdw_vol and weinerPath describe size and topology of the molecules. They correlate with each other. Both descriptors correlated negatively with the activity. So it seemed that substituents at the aromatic system which enhanced the size of the compounds in a high degree contributed negatively to the activity as well.

4.1.2 2D-QSAR model of the alkyl substituted MP derivatives (A2)

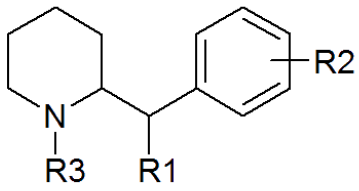


Figure 29: Common Scaffold of the compounds analyzed in model.

The 25 methylphenidate derivatives of this study mainly carry different alkyl- or aralkyl- residues in Position R1 (Figure 29). These residues differ in their size and length as well as in their degree of branching and bulkiness. In methylphenidate there is a carboxy-methyl- ester group in this position. Some of the molecules also carry substituents in the para- or meta- group of the aromatic ring, mainly chlorine atoms. This is depicted by "R2" in Figure 29. Only in one compound the amine group is methylated (R3).

To be able to describe the difference in size and shape of R1 groups the following two descriptors where chosen as a first step. The model was mainly built via forward selection of two additional descriptors.

Step 1: The initial model based on the descriptors "weight" and "Zagreb". Both of the descriptors show a high relative importance:

Weight: 0.99 Zagreb: 1.00

RMSE:	0.60831
R2:	0.52483
XRMSE:	0.68244
XR2:	0.40878
Nr. C.	2

Step 2: To find other suited descriptors for the model an auto QSAR analysis was performed which made it possible to find another suitable descriptor for the model (PEOE_VSA-1).

RMSE:	0.51547
R2:	0.65880
XRMSE:	0.59667
XR2:	0.54766
Nr. C.	3

Step 3: As it was not possible to find another statistically beneficial descriptor using auto QSAR methods, other descriptors were checked manually. Considering the differences in size of the molecules a diameter descriptor was chosen for the model.

RMSE:	0.46143
R2:	0.72660
XRMSE:	0.56166
XR2:	0.60183
Nr. C.	4

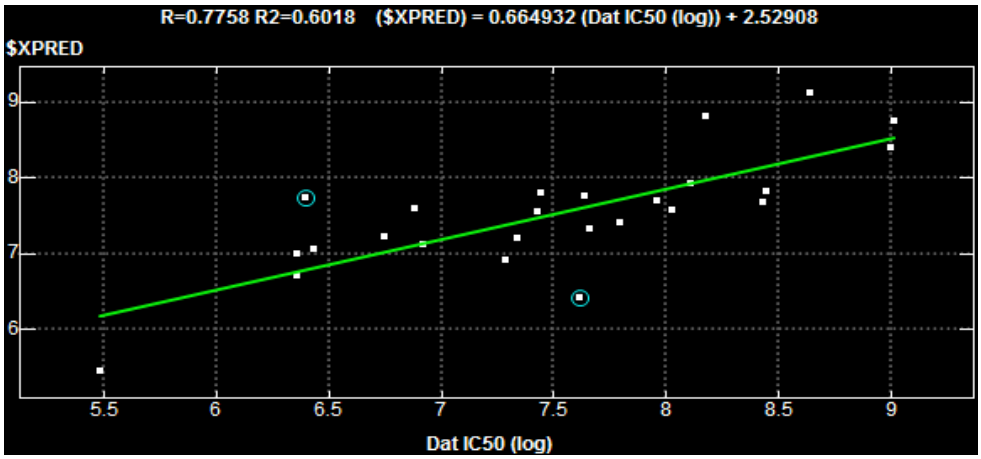


Figure 30

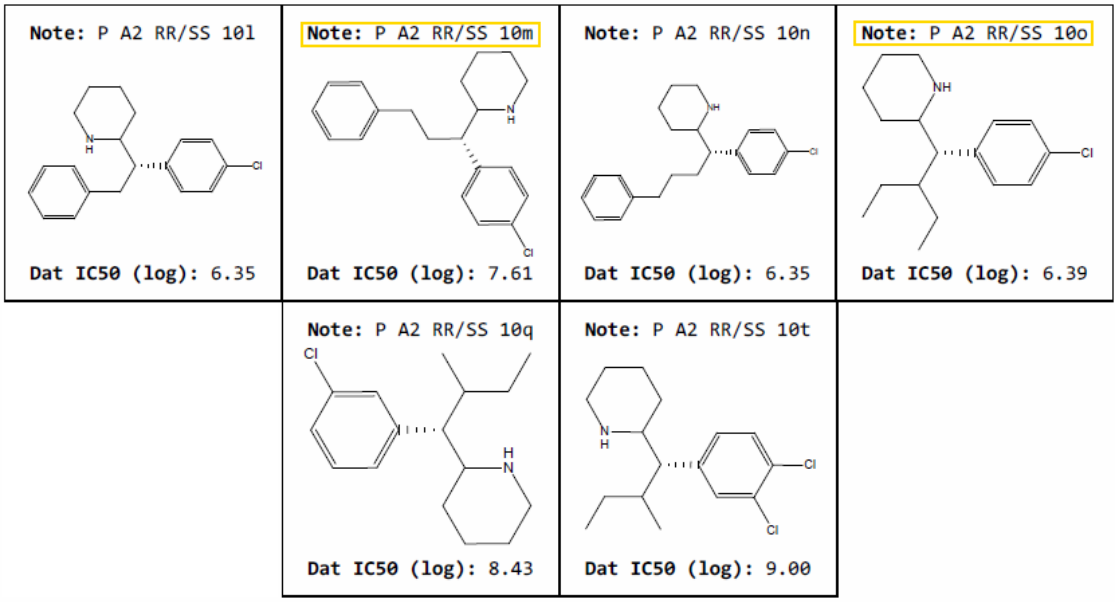


Figure 31

However, if one looks at the outliers it can be seen that there are two activity cliffs. The marked compounds in the correlation plot show these outliers (the corresponding marked structures can be seen in Figure 30). If the structural differences of these two compounds are considered it is difficult to explain the given activity values. Compound “P A2 RR/SS 10m” is a member of a homologues series of compounds in this dataset. If the alkyl residue is one atom longer or shorter the activity is more than one order of magnitude higher. In the case of compound “P A2 RR/SS 10o” the cliff is even more drastical. Considering the small dimension of the structural changes and - especially in the case of compound “P A2 RR/SS 10m” the activity values of the homologues derivatives - the question arises, if there were errors in the measurement. Therefore it was also considered to exclude these two outliers.

Step 4: The mentioned two outliers where excluded.

RMSE:	0.33529
R2:	0.85786
XRMSE:	0.42651
XR2:	0.77405
Nr. C.	4

Dat IC50 (log) = 8.477 -0.519 * diameter -0.022 * PEOE_VSA-1 +0.066 * Weight -0.130 * Zagreb

Interpretation of the model:

The model delivered quite a good output (RX2: 0.77) but it was not easily possible to interpret it. The Zagreb Index descriptor and the molecular weight descriptor correlated positively with each other in the given dataset while they showed antipodal behavior in the QSAR Model. The Zagreb Index is like the Wiener Path a topological index describing the shape of a molecule. While the molecular weight parameter does only consider the weight of the single atoms of a molecule and not their constellation, the Zagreb index also takes into account the type and number of branching of the molecule. So the descriptors Zagreb Index and diameter, which are both contributing negatively to the calculated activity, could be seen as hint that the size of the molecules is an important factor possibly because of space limits in the binding pocket. Maybe the weight parameter is contributing positively to the model because the most active compounds all carry two chlorine atoms at the meta and para positions of their phenyl group, which increases their molecular mass in total. This could explain the antipodal influences of the descriptors which describe the size of the molecules and the one which describes their weight.

	1	2	3	4
1. diameter	100	42	60	71
2. PEOE_VSA-1	42	100	73	60
3. Weight	60	73	100	89
4. zagreb	71	60	89	100

Figure 32

The PEOE_VSA-1 descriptor characterizes the size of the Van der Waals surface of atoms, which feature a slightly negative charge ($-0.1 < q_i < -0.05$). (Chemical Computing Group 2009) This parameter contributed negatively to the activity.

4.1.3 The H-QSAR Models

Model for phenyl ring substituted analogues A7:

All of the H-QSAR models where created following the same procedure which is described in detail in section "3.2.3 Creation of the H-QSAR models:".

The same compounds as in the corresponding 2D-QSAR model were used for the development of the model.

Used Descriptors: **B+Co**

The unique fragments were selected on basis of the types of their bonds and on the hybridization state of their carbon atoms.

Fragment Size	Best CV R2	Best Full R2
4-7 Atoms	0.912	0.973

Best CV StdErr	Best Full StdErr	B Long	Comp
0.418	0.230	97	6

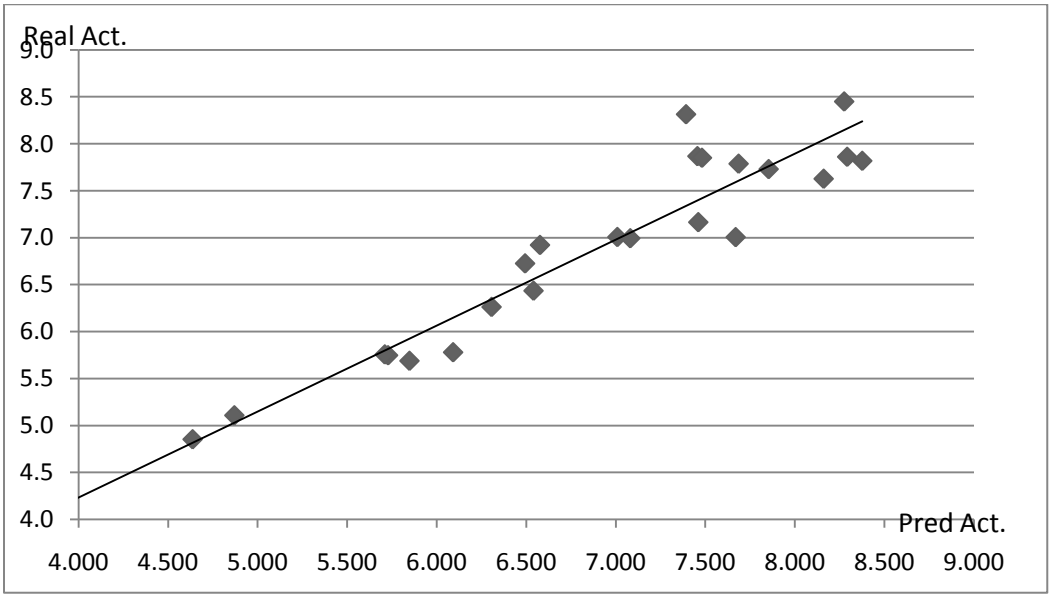


Figure 33

All the predicted activities were within one order of magnitude to the pharmacologically measured activity values. The H-QSAR model even performed better than the corresponding 2D QSAR Model.

Model for alkyl substituted MP analogues (A2):

The same compounds as in the corresponding 2D-QSAR model were used for the model's development.

Used Descriptors: **A +Ch**

In this case unique fragments were selected on basis of their atom types as well as on their chirality.

Fragment Size	Best CV R2	Best Full R2
4-7 Atoms	0.580	0.893

Best CV StdErr	Best Full StdErr	B Long	Comp
0.674	0.340	59	6

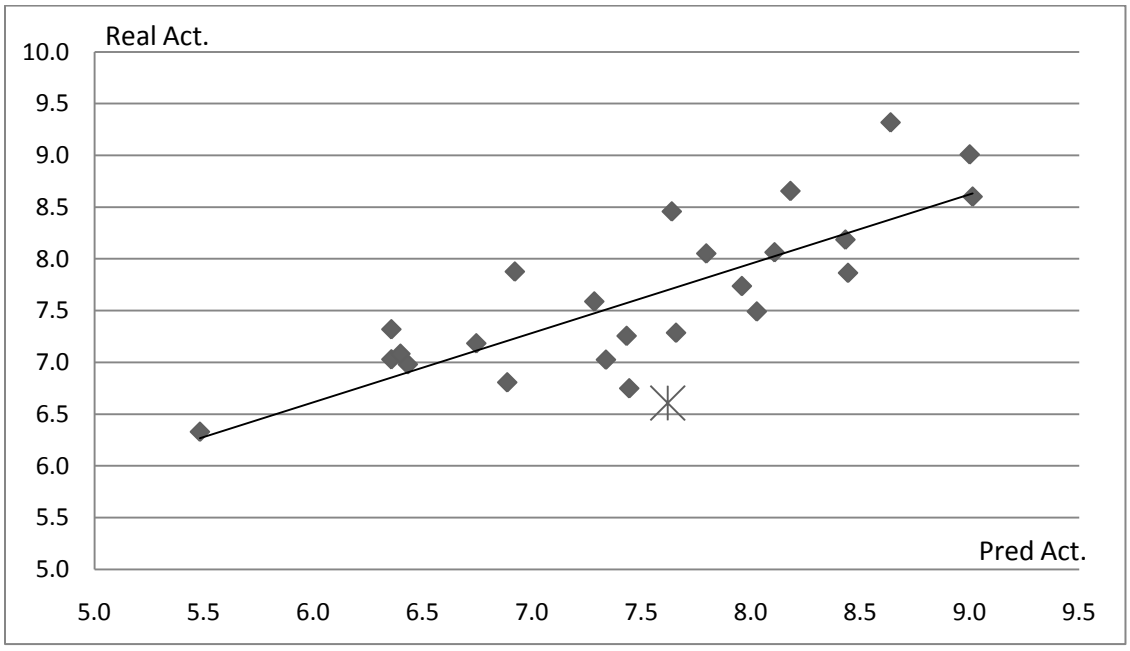


Figure 34

One predicted activity calculated with the cross validated model exceeded the limit of one order of magnitude difference (depicted by a cross in Figure 34). This concerned compound "P A2 RR/SS 10m" (Figure 35), which was also an outlier in the corresponding 2D-QSAR model. This could support the theory of an error in measurement concerning this compound's activity. The rest of the predicted values were within the requested range.

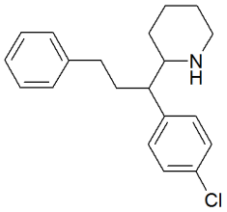


Figure 35
"P A2 RR/SS 10m"

H-QSAR Model of alkyl phenyl ring substituted derivatives together (A2 +A7):

Used descriptors: **B + Ch +D**

Unique Fragments were selected on basis of their bond types, their chirality and their donor/acceptor properties.

Fragment Size	Best CV R2	Best Full R2
5-8 Atoms	0.736	0.901

Best CV StdErr	Best Full StdErr	B Long	Comp
0.604	0.369	151	5

The model's accuracy approximately reaches an average value of the two previously created models which only dealt with the single studies. The residual values of four of the examined 49 compounds exceeded the internally set mark of one order of magnitude.
Even if the compounds of study A7 were not predicted as accurately by this HQSAR-Model as by the first one it could still be considered as a progress, as a prediction of more diverse methylphenidate derivatives is possible within one model.

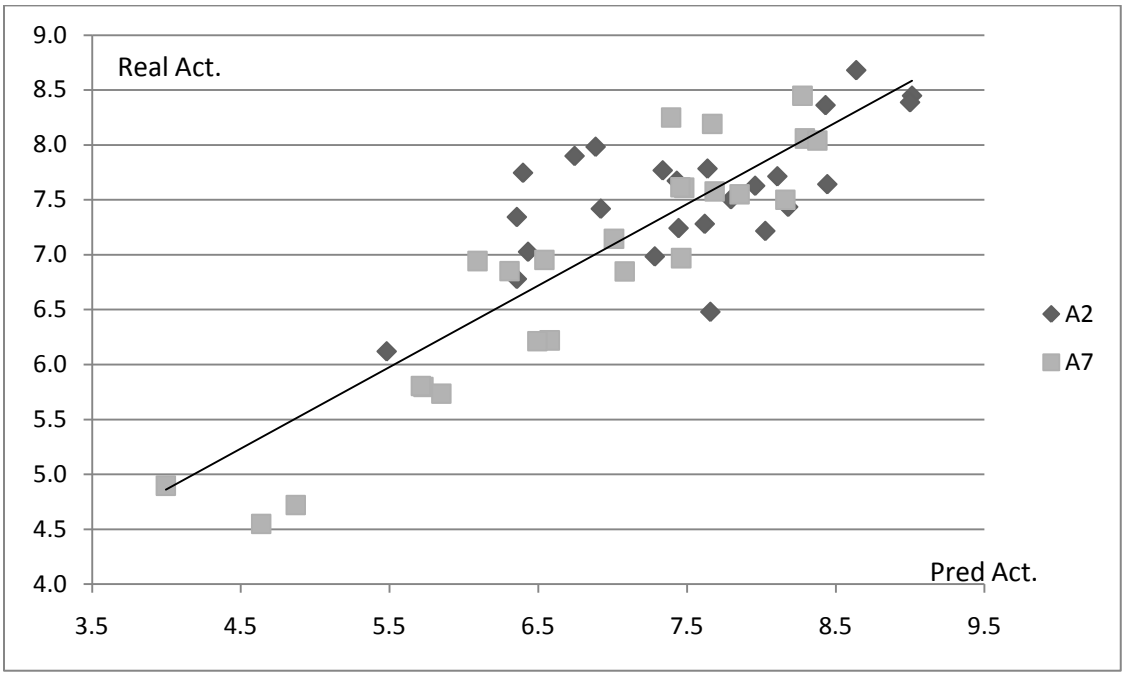


Figure 36 Correlation of the predicted (CV) versus the actual activity of the compounds (A2 + A7)

<u>H-QSAR Model of Study A2,A5 and the rotationally fixed MP analogues (A7) together:</u>

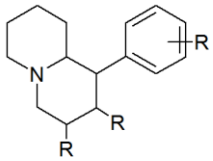


Figure 37 Rigidized Compounds Study A5

Figure 37 on the right shows the rigid scaffold of the compounds of study A5. The authors wanted to bring to light what role methylphenidate's free rotatetability plays for the binding affinity.

Used descriptors: **B+Co+D**

Unique fragments were selected on basis of their bond types, the hybridization state of their carbon atoms and based on their donor/acceptor properties.

Fragment Size	Best CV R2	Best Full R2
4-7 Atoms	0.650	0.854

Best CV StdErr	Best Full StdErr	B Long	Comp
0.673	0.435	53	6

The lower correlation of the predicted versus the actual activity values can also be observed in the correlation graph (Figure 38.) The residual values of nine of the observed 66 compounds (13.6%) exceeded the internally set mark of one order of magnitude.

This model again points out the difficulty of finding a good QSAR model for a rather diverse subset of structures. The predicted activity values of three compounds of each of the three studies exceeded that limit respectively.

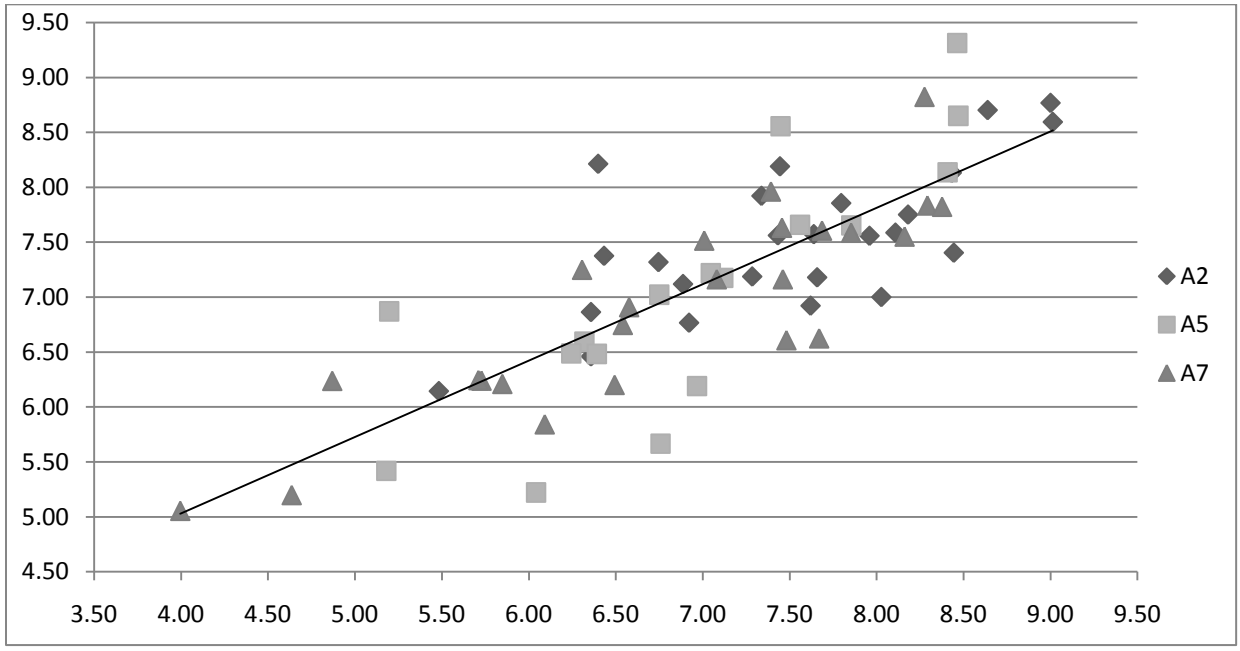


Figure 38

When the H-QSAR models were analyzed visually, it was not possible in all models to draw conclusions concerning the contributions of single fragments via the standard color encoded interpretation tool which is included in the Sybyl Software package.

Normally this tool depicts fragments which contribute negatively or positively to the calculated activity values in different color patterns. This would have made it possible to better understand the properties of the analyzed compounds.

Nevertheless, predictive HQSAR models could be created, which could also be used to strengthen results which were obtained by the use of the previously developed 2D-QSAR Models.

4.1.4 Résumé for the Ligand based analyses of methylphenidate derivatives:

The QSAR experiments gave an insight into the properties of the methylphenidate analogues and on influences on the activity of structural modification of the compounds. The set of available synthesized derivatives of this rather small compound is extremely rich. A wide variety of modifications was done and examined for nearly every position of the structure. It seems no element of the structure is irreplaceable. High potent and selective derivatives can be found in every class of the modified compounds.

That is also why it was decided to examine the influence of structural modifications in a smaller breadth.

These insights especially on phenyl substituted analogues could be used for the interpretation of the docking experiments. The combination of ligand und structure based methods is realized in the section (5 Discussion).

4.2 Docking

4.2.1 Docking with GOLD

The detailed workflow of the final docking run is described at the beginning of the “Methods” section.

The analysis of the final docking run produced the following results:

150 poses were calculated for each of the six ligands (Figure 39). So the run produced 900 poses in total. The scaffold clustering of these poses with a threshold of 2.0 Å resulted in 17 clusters. Two of these clusters (Cluster 9 and 12) contained poses of all six compounds.

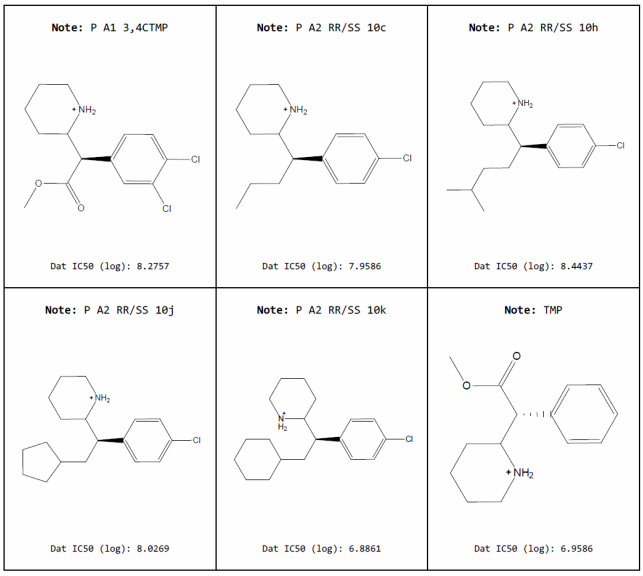


Figure 39

However, when the clusters where checked in detail it was found that there were two different positionings of methylphenidate within one cluster.

Thus, the RMSD range of each cluster had to be decreased. Different thresholds were investigated (1.5Å, 1.6Å, 1.7Å). On the one hand the deviation should not be too low (otherwise no cluster was retrieved which contained all compounds) and on the other hand not too high.

Considering these criteria 1.6Å was found as the optimal value. Only one cluster remained containing all compounds. The remaining cluster carries the label Cluster 25 and contains 158 poses. The omitted cluster (Cluster 21) lacks poses of methylphenidate itself and only contains one pose of the di-chloro substituted analogue. In total it contains 163 poses.

Nevertheless, Cluster 21 was also analyzed here as it was also present in the MOE docking experiment where it contained poses of all derivatives.

Cluster 25:

In this positioning the phenyl ring of all compounds is pointing towards solvent direction. As is observed in each cluster the ionic interaction between ASP79 and the Nitrogen group in the ligands is the most dominant one and occurs in the majority of the poses. Furthermore, Phe320 interacts with the amine group, partly via its backbone's carbonyl group, partly via its -rich in electrons- aromatic side chain.

Interactions with Tyr156 can be observed most often within poses of the two ester compounds. Either aromatic stacking interactions between the two aromatic rings of methyl-phenidate and Tyr156 (as it can also be seen in Figure 40) or a hydrogen bond between the Tyr156's hydroxyl group and the carbonyl oxygen of Methylphenidate/ Di-Cl- Methyl-phenidate can be observed.

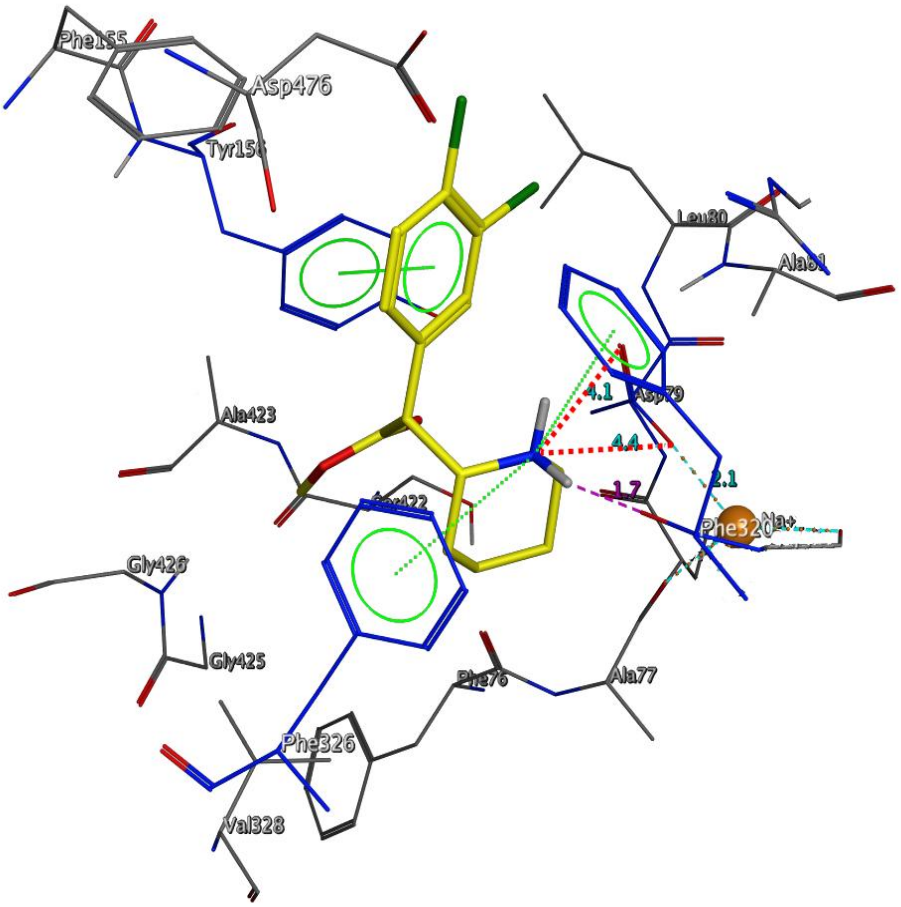


Figure 40

Interactions with the residues Tyr156 and Asp79 are especially interesting as they are also known to be important for cocaine binding to the Dopamine transporter (Beuming et al. 2008). Though the mode of interaction for methylphenidate is not known in detail, its effects on the Dopamine uptake and respectively the Dopamine transporter are very similar to those of

cocaine (see 1.1.3 Important interacting Drugs). Thus it is likely that Tyr156 is of importance for the binding of methylphenidate too.

Unfortunately, interactions with an aromatic π-electron system cannot be depicted or included in the statistical analyses of a PLIF in MOE. These interactions can only be observed when one looks at each complex and its interactions manually. The interaction rate with Tyr156, Phe320 and Phe326, shown in Figure 41, would be much higher if also π-π interactions would be considered by the MOE's PLIF algorithm. That is also why the interactions with Phe326, which can be seen in Figure 40, do not show up at all in the PLIF. An interaction which could only be observed with the ester analogues is the hydrogen bond between the hydroxyl group of Ser422 and the carbonyl group.

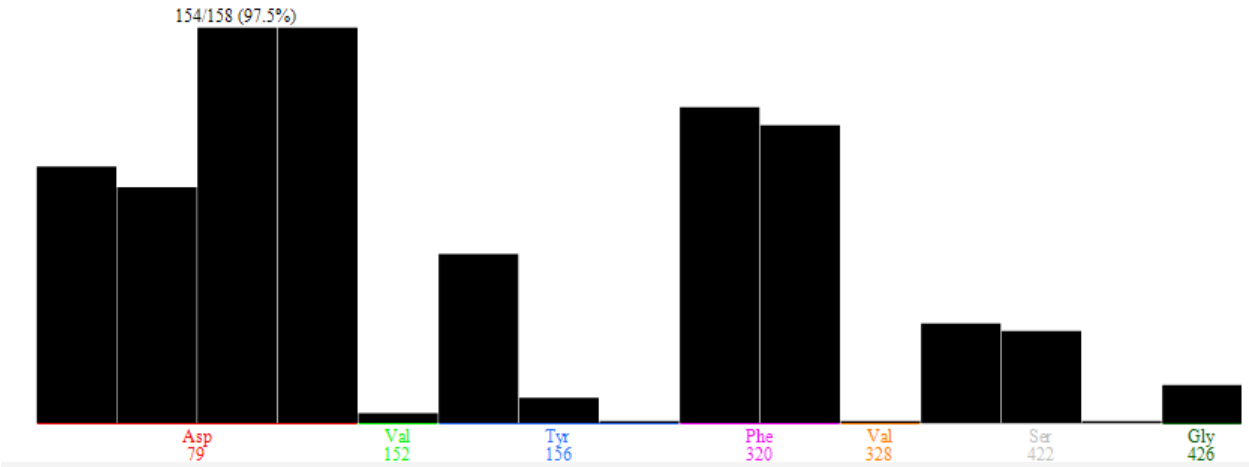


Figure 41 PLIF Cluster 25

Composition of Cluster 25:

P A1 3,4CTMP	103 Poses
P A2 RR 10c	4 Poses
P A2 RR 10j	8 Poses
P A2 RR 10k	15 Poses
P A2 RR 10h	6 Poses
TMP	22 Poses

Composition of Cluster 21:

P A1 3,4CTMP	1
P A2 RR 10c	2
P A2 RR 10j	51
P A2 RR 10k	50
P A2 RR 10h	59
TMP	0

The Clusters were also compared with respect to their GOLD Score Values. To achieve this, the two clusters were merged to one database. Then the poses of both clusters were ranked according to their Gold score values. Cluster 25 is definitely preferred according to these values. The top scored 127 poses -of 321 in total- all belong to cluster 25. This represents more than 80% of the poses of this cluster.

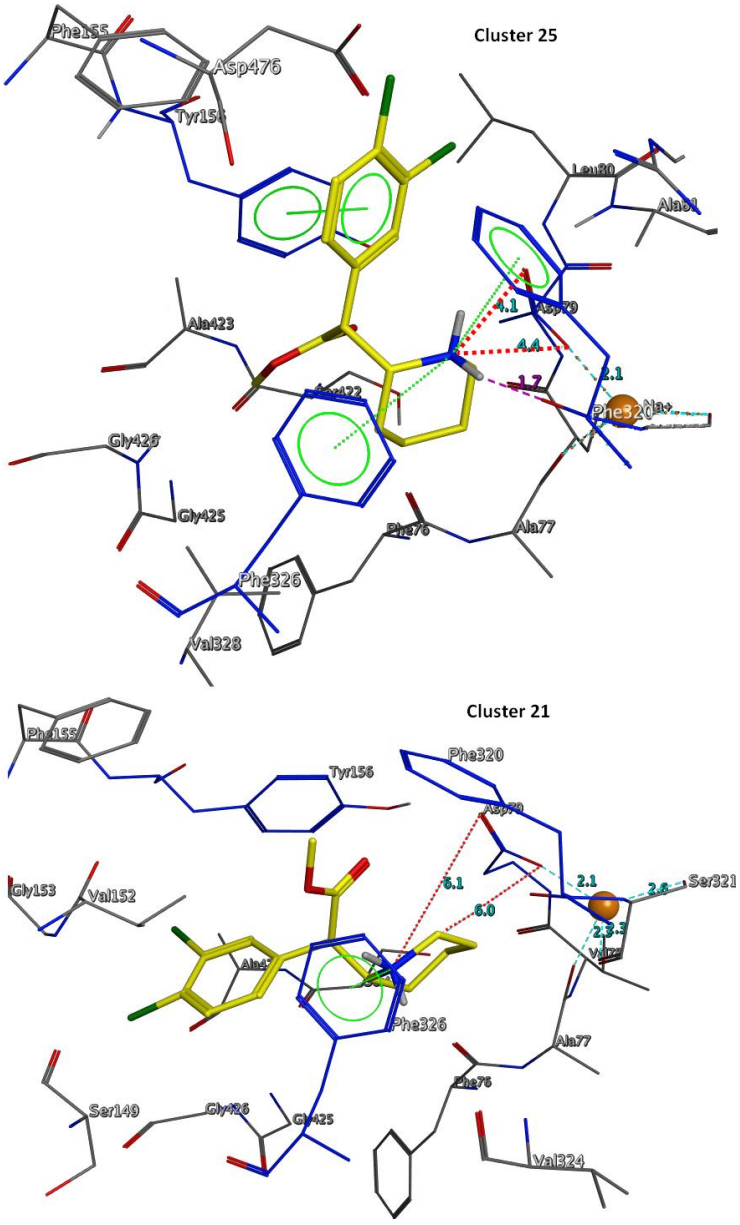


Figure 42

Cluster 21:

In this cluster the alkyl- or ester- side chain is pointing towards a solvent exposed area while the aromatic ring is pointing to the inside of the transporter, facing TMD 3. Figure 42 shows Cluster 21 in comparison to Cluster 25. It is visible that the ester and the aromatic group swapped their position. Hardly any π-π interactions with Tyr156 are possible in this cluster as the compound's aromatic system is too far away from Tyr156's side chain and also fewer interactions between Phe320's π system and methylphenidate's charged nitrogen occur.

4.2.2 Docking with MOE

In this docking run the same pocket in the rigid protein was used in combination with the same 6 ligands used in the GOLD docking run. The “Triangle Matcher” was used as a placement method and the “London dG” scoring function for rescoring purposes.

After the automatic elimination of duplicate poses approximately 60 poses per ligand were retrieved. The complexes were minimized using MOE’s LigX tool. Afterwards the poses were clustered with a threshold of 3 Å according to the RMSD of the ligands, using the same SVL Clustering script (within MOE) which was described in Section “3.3.5.5 Final Workflow: Docking with flexible sidechains in GOLD”.

All clusters which did not contain poses of all of the six ligands where omitted. This brought up three final clusters, Cluster 16, Cluster 29 and Cluster 37.

Two of the three clusters resemble with the two Gold clusters, MOE Cluster 37 with GOLD Cluster 25 and MOE Cluster 29 with GOLD Cluster 21.

The positioning of the third one Cluster 16 is a new one, depicted in Figure 43.

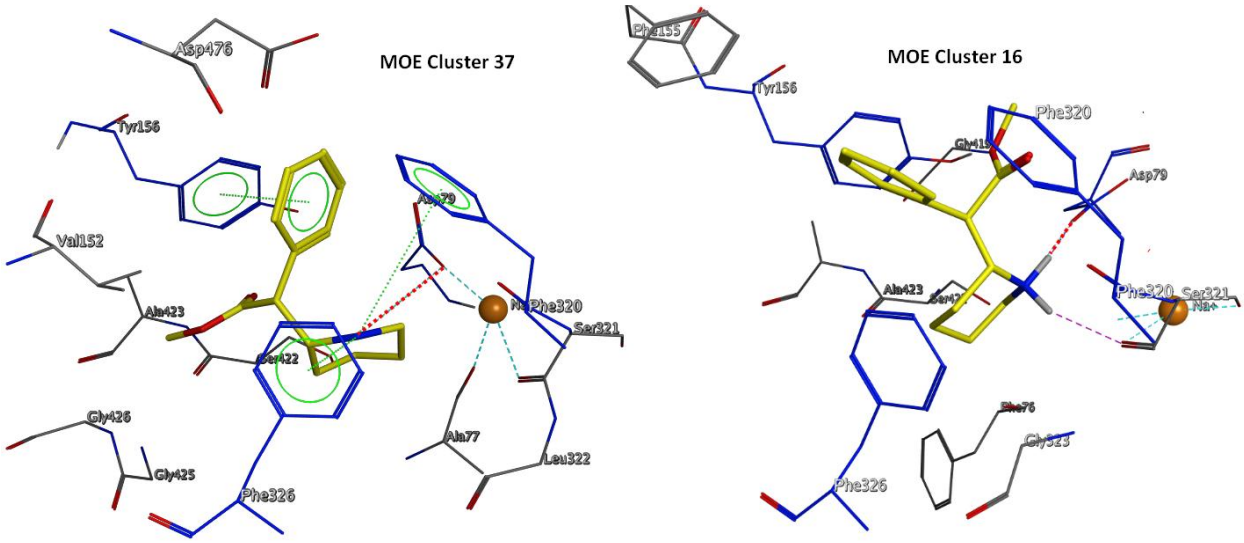


Figure 43: In Cluster 16 the carboxy-methyl-ester and the phenyl ring are oriented towards the outside of the pocket. The aromatic ring is closer to Tyr156’s sidechain. This allows interactions with this residue in some of the cluster’s poses.

Different approaches were tried to rank the clusters.

- First, the results were rescored using four different scoring functions: ASE, Alpha HB, Affinity dG and London dG. The proportion of each pose under the 15 top scored poses calculated with the different scoring methods was determined. The mean over all of the different methods was taken as a criterion. The result can be seen in Figure 44. Cluster 29 was curtly ranked best. Nevertheless, the difference between the clusters is rather small especially if the large deviation between the scoring functions is considered. Cluster 29 shows highest and Cluster 37 the lowest standard deviation between the scoring functions.

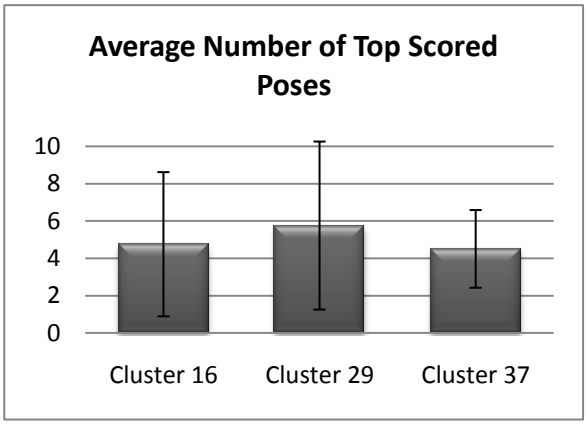


Figure 44

- As a second criterion the assumption was made that the amino acid residues Tyr156 and Asp79 are important for the methylphenidate binding, based on the experimental cocaine data which was already explained in section "4.2.1 Docking with GOLD". While in all clusters nearly in 100% of the poses there is an interaction between Asp 79 and the amine group of the methylphenidate analogues, the interaction rate with Tyr156 varies. Figure 45 depicts the interaction rate in the three clusters in percent. π-π- aromatic stacking interactions of Tyr156 with the methylphenidate's phenyl group are considered. In cluster 16 you also see hydrogen

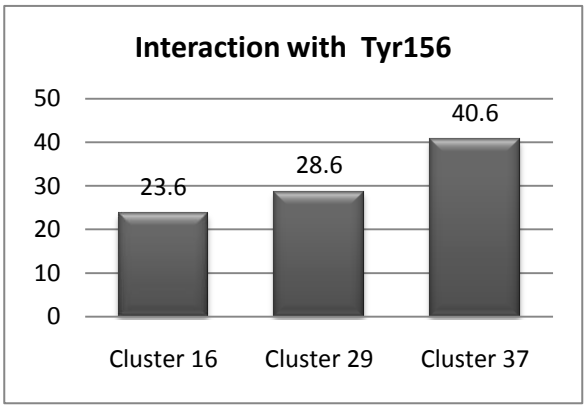


Figure 45

bond interactions between Tyr156 phenolic hydroxyl group and the carbonyl group of the ester compounds (TMP, P A1 3,4CTMP).

Cluster 37 is the only cluster in which there is at least one pose showing interactions with Tyr156 as well as Asp79 at the same time.

- Also the size of the clusters was considered. Figure 46 shows the number of poses within each cluster.

  Cluster 37 was the biggest one.

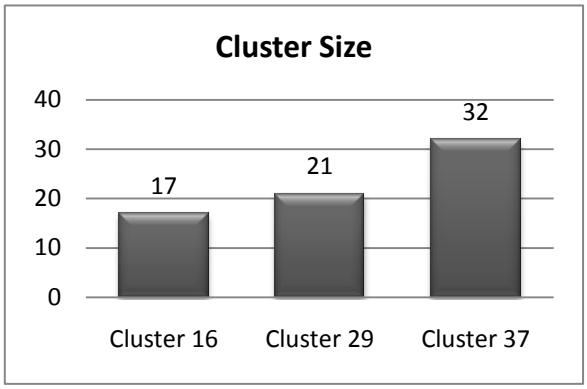


Figure 46

- As already mentioned only two of the clusters were also obtained in the GOLD docking experiments. That was taken as a final criterion.

Taken these criterions (especially the size and the interaction pattern) into account and also considering the results of the GOLD docking run, Cluster 37 (shown in Figure 47 next page) was chosen as the preferred one and the other two were omitted.

4.2.3 Comparison of MOE and GOLD Results

When the PLIFs of the final clusters are compared (see page 65, Figure 48) one can see quite a similar interaction pattern which is obvious as the positionings within the two clusters concur closely. However, the interaction rate with each residue is not equal throughout the clusters, which can be explained when one keeps in mind that in the GOLD run side chains were kept flexible while in the MOE run it was docked into a rigid protein. There are more contact interactions with the amino acids Ser149, Val152, Ser422 and Gly426 in the MOE Cluster. The lack of sidechain flexibility could be an explanation for that. Furthermore, some intereactions

(with similar partner-residues in the protein) differ in their nature: In the final GOLD cluster mainly the ester compounds interact with Ser422 via a hydrogen bond to its hydroxyl group, in the MOE Cluster preferentially the lipophilic alkyl analogues interact with this residue via contact interactions. That is mainly the case because Ser422 hydroxyl group is pointing out of the pocket in the rigid model while it is pointing towards the molecule in some poses in the flexible run. Neither the PLIF of the MOE Cluster nor the one describing the GOLD cluster show any π-ion interactions of Phe326 and the amine nor the π-π- Interactions between Tyr156 and Methylphenidate's Phenyl ring although they occur in both analyses.

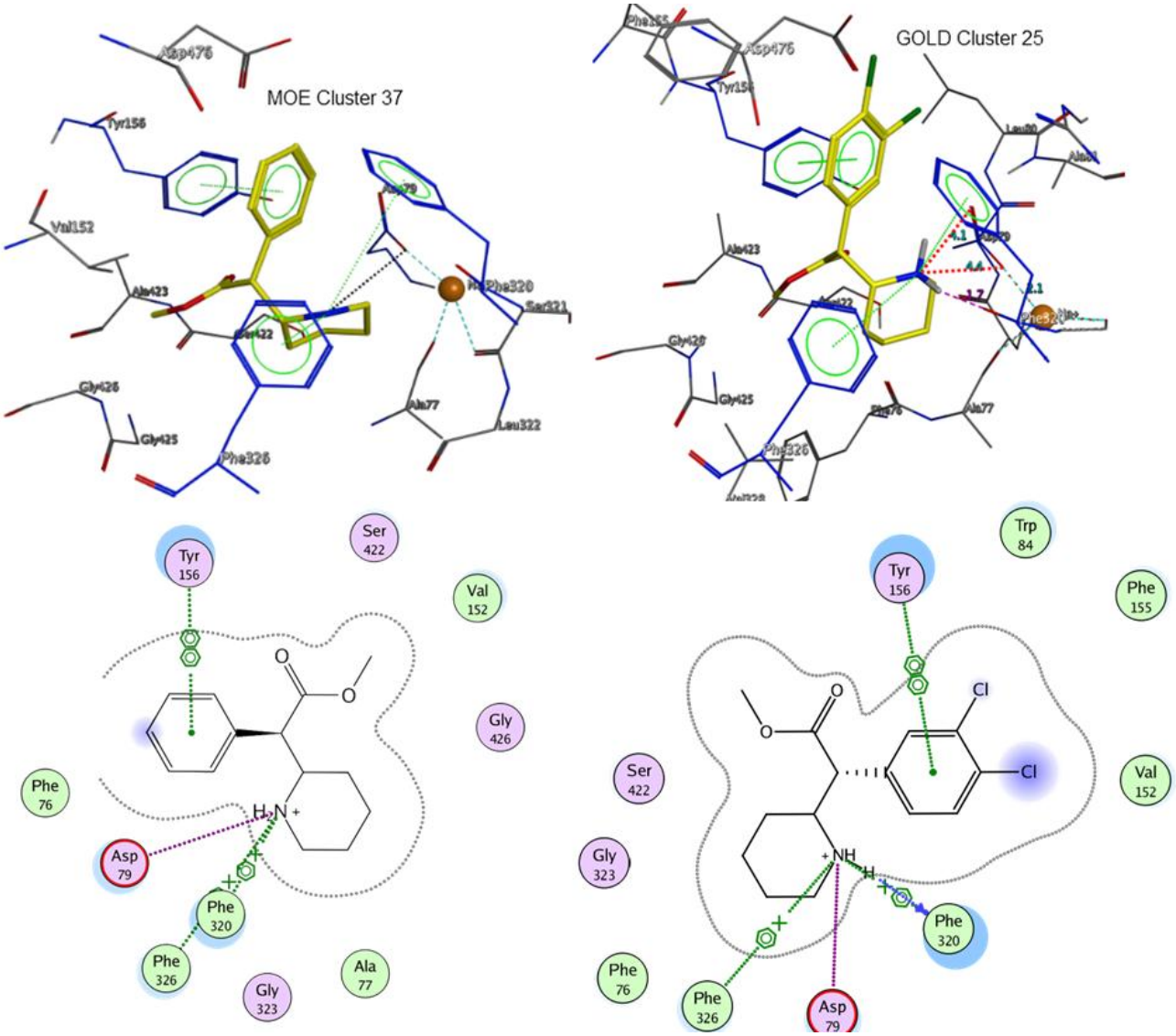


Figure 47: Comparison of the final clusters.
For a better overview a 2D interaction map of the poses is shown at the bottom.
The Atoms marked in purple show areas with exposure to the solvent

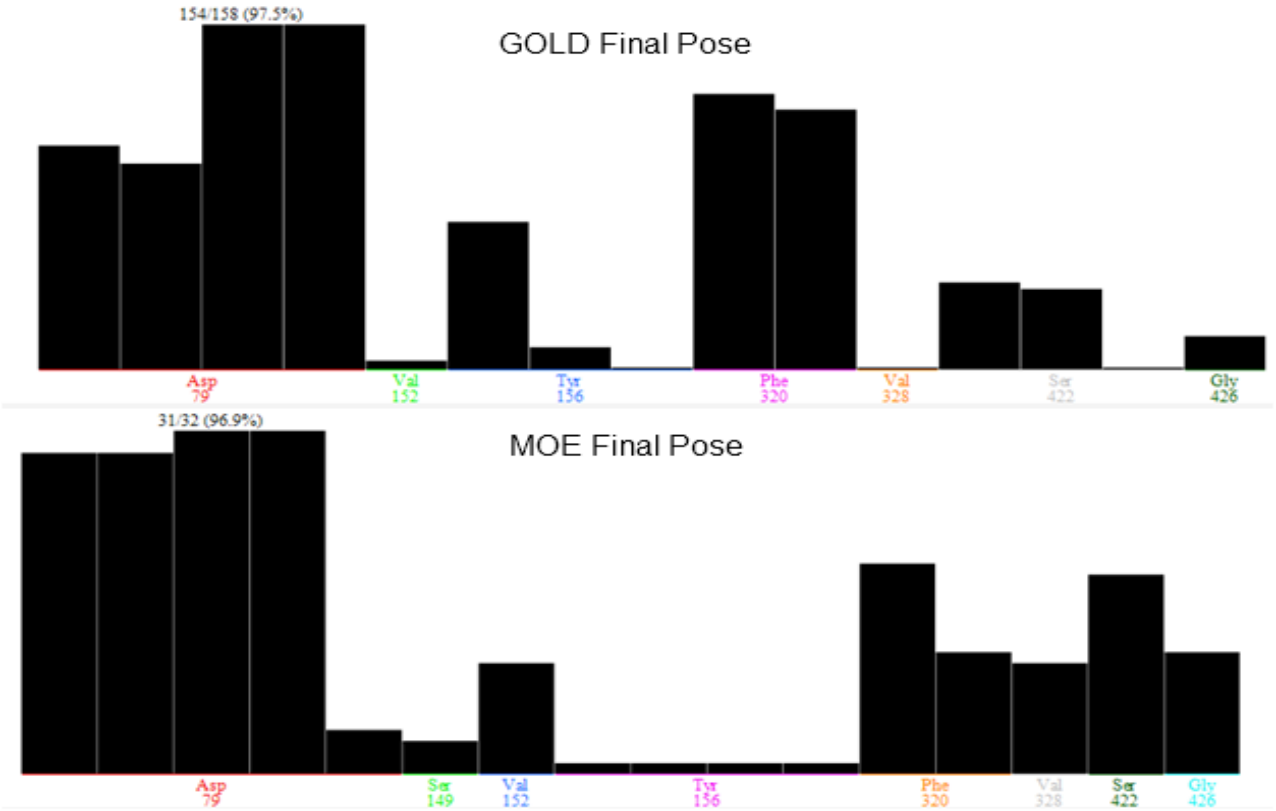


Figure 48

Also only the Backbone – Ion Interactions between Phe320 and the amine are considered in the PLIF, not the π-Ion interactions.

The problem that the PLIF does not show interactions involving π electron systems was already mentioned in "4.2.1 Docking with GOLD". In Figure 47 comparative poses from both clusters and also the interactions involving π electron systems are depicted.

5 Discussion

The docking run was based on the proposed cocaine binding pose of (Beuming et al. 2008). This proposed binding pocket which, as already mentioned, overlaps with that of dopamine, was extensively validated by mutational and cross linking experiments (Beuming et al. 2008). Even if the pharmacological mechanism of methylphenidate and cocaine is similar - they are both blocking the dopamine transporter - one cannot be sure if they do so by binding in an analogous manner. It is risky to make assumptions for the binding mode of one compound based on experimental data of another, related agent. However, as there is no mutational analysis data available for methylphenidate itself, this was taken as a basis.
The same is true for assumptions for single interaction partners (for example in this work: Asp79).

Proposed Interactions for methylphenidate and its analogues:

Asp79 is a strong interaction partner for methylphenidate as it forms an ionic bond with the compound's protonated amine. This positive charge within methylphenidate is also the interaction partner for the residues Phe320 and Phe326.
The interactions with Tyr156 are present throughout all the docked compounds, either via a π-π- Interaction or via a hydrogen bond, formed between the hydroxyl group and the carbonyl group of methylphenidate (the latter is only concerning the ester analogues and methylphenidate). Furthermore, this carbonyl group is a possible hydrogen bond acceptor for Ser422.
It is questionable if the contact interactions between the compounds and the protein (Ser149, Val 152, Val328, Ser422 and Gly426) are very important for the methylphenidate's binding mode, as these, on the one hand, are very weak interactions and on the other hand do occur in the rigid docking run in a much higher degree. This may point out the importance of flexibility in

a docking process as this interaction could not be observed within the flexible docking runs. It was published recently in "Current Opinion in Structural Biology" that 50%-70% of the calculated binding poses based on rigid proteins models, only considering one conformation of the receptor, are incorrect. (Totrov and Abagyan 2008)
To circumvent those problems the flexible side chain workflow used in the GOLD docking run was developed.
So even if the results of both docking runs are closely related and one strengthens the other the flexible one may be preferred when it comes to details especially considering weak interactions.

Correlation of activity values of methylphenidate derivatives and insights obtained in QSAR analysis with the proposed docking pose:

The Amine: This model shows the importance of the ionic interaction with Asp79, which is also a proposed key interaction for cocaine binding. Nevertheless, also Oxa- and Carba-cyclic methylphenidate analogues show inhibitory potential. When the nitrogen atom is changed to oxygen the affinity decreases by nearly 3 orders of magnitude. If the aromatic ring is additionally substituted with two chlorine atoms in meta and para position the affinity stays within one order of magnitude of the corresponding nitrogen analogue (Meltzer et al. 2003). These chlorine substituents are likely to have high influence on the interaction strength of the compounds. As it was published by Froimowitz (Froimowitz et al. 2007) even the otherwise inactive erythro diastereomers of alkylated methylphenidate analogues show high interaction potential when the aromatic ring is 3,4- Di-chloro substituted.

The phenyl group and its substitution: Figure 49 on the next page shows the protein surface of the pocket. Green color indicates lipophilic surface area whereas the violet one shows hydrophilic area. As it can be seen the methylphenidate's aromatic system is surrounded by a lipophilic surface area. This is on the one hand plausible by itself - the phenyl group is a lipophilic structure - and on the other hand could be an explanation why halogen substituents in meta and para increase the affinity (Deutsch et al. 1996) as they increase the lipophilicity.

Furthermore, the observed π – stacking interaction with Tyr156 could be an explanation for the increased activity of the halogen substituted compounds, as the aromatic system is enriched in electrons by halogen atoms via mesomeric effects. The 2D QSAR model dealing with the influence of different substituents at the phenyl group (4.1.1 2D-QSAR Model phenyl ring substituted analogues (A7)) approved the positive influence of substituents which increase the lipophilicity in para position as well as in ortho position. However, it also shows that the positive effect of increased lipophilicity of these substituents does not play a role as the hindering effect of the size of these substituents prevails. The docking model cannot explain the nature of this sterical hindrance although the pocket is quite narrow in this region and a slight expansion of the pocket might be necessary to accommodate these ligands.

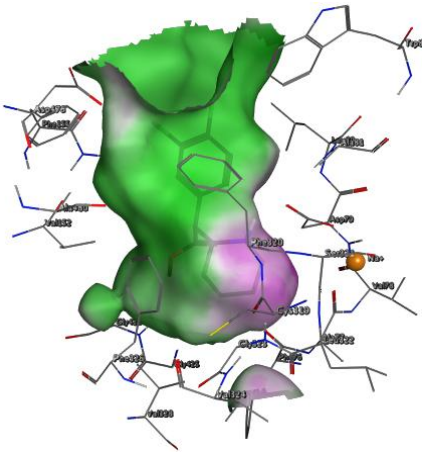


Figure 49 Pocket Surface

Outlook

Docking poses always represent only snapshots of interactions with a receptor or target protein. Even if side chains are kept flexible during the docking run this is a rather small degree of flexibility as the interaction process is a dynamic one.

To meet these requirements better it would make sense to dock into different receptor conformations, which is already known to improve docking results (Totrov and Abagyan 2008). This would be a possibility to consider a broader degree of flexibility as well as different interaction patterns of ligand and protein at different stages of the interactions.

Furthermore molecular dynamic simulations could be a second step to a better understanding of the ligand's behavior in the binding pocket and its environment.

Finally mutational analysis of the proposed interacting amino acid residues would be important for the validation of the model. Asp79, Tyr156, Phe320, Phe326 also Ser422 are the most

interesting candidates for these mutational studies. If methylphenidate's affinity decreases with the exchange of these amino acids this would strengthen the proposed binding mode.

A References:

Benzenhofer, U. and T. Passie (2006). "[The early history of "Ecstasy"]." Nervenarzt **77**(1): 95-96, 98-99.

Beuming, T., J. Kniazeff, et al. (2008). "The binding sites for cocaine and dopamine in the dopamine transporter overlap." Nat Neurosci **11**(7): 780-789.

Bonchev, D. (1983). "Information theoretic indices for characterization of chemical structures " Wiley-Interscience: New York,.

CCDC, C. C. D. C. (2009). GOLD Suite 4.1.1.

CCDC, C. C. D. C. (2010). GOLD Online Documentation, Cambridge Crystallographic Data Centre.

Chemical Computing Group, M., Canada (2009). MOE (Molecular Operating Enviroment) Tutorial, Chemical Computing Group, Montreal, Canada.

Davies, H. M., D. W. Hopper, et al. (2004). "Synthesis of methylphenidate analogues and their binding affinities at dopamine and serotonin transport sites." Bioorg Med Chem Lett **14**(7): 1799-1802.

Deutsch, H. M., Q. Shi, et al. (1996). "Synthesis and pharmacology of potential cocaine antagonists. 2. Structure-activity relationship studies of aromatic ring-substituted methylphenidate analogs." J Med Chem **39**(6): 1201-1209.

Elkashef, A., F. Vocci, et al. (2008). "Pharmacotherapy of methamphetamine addiction: an update." Subst Abus **29**(3): 31-49.

Fleckenstein, A. E., T. J. Volz, et al. (2007). "New insights into the mechanism of action of amphetamines." Annu. Rev. Pharmacol. Toxicol.

A References:

Froimowitz, M., Y. Gu, et al. (2007). "Slow-onset, long-duration, alkyl analogues of methylphenidate with enhanced selectivity for the dopamine transporter." J Med Chem **50**(2): 219-232.

Gatley, S. J., D. Pan, et al. (1996). "Affinities of methylphenidate derivatives for dopamine, norepinephrine and serotonin transporters." Life Sci **58**(12): 231-239.

Gerlach, M., C. Mehler-Wex, et al. (2009). Neuro-Psychopharmaka im Kindes- und Jugendalter Grundlagen und Therapie, Springer-Verlag/Wien

Goldstein, R. A., C. DesLauriers, et al. (2009). "Cocaine: history, social implications, and toxicity--a review." Dis Mon **55**(1): 6-38.

Grabowski, J., J. D. Roache, et al. (1997). "Replacement medication for cocaine dependence: methylphenidate." J Clin Psychopharmacol **17**(6): 485-488.

Greely, H., B. Sahakian, et al. (2008). "Towards responsible use of cognitive-enhancing drugs by the healthy." Nature **456**(7223): 702-705.

Hansch, A. L. C. (1979). Substituent Constants for Correlation Analysis in Chemistry and Biology. NY, Wiley-Interscience.

Huang, X., H. H. Gu, et al. (2009). "Mechanism for cocaine blocking the transport of dopamine: insights from molecular modeling and dynamics simulations." J Phys Chem B **113**(45): 15057-15066.

Kim, D. I., H. M. Deutsch, et al. (2007). "Synthesis and pharmacology of site-specific cocaine abuse treatment agents: restricted rotation analogues of methylphenidate." J Med Chem **50**(11): 2718-2731.

Klebe, G. (2009). Wirkstoffdesign - Entwurf und Wirkung von Arzneistoffen. Heidelberg, Spektrum Akademischer Verlag Heidelberg.

A References:

Kollins, S. H., E. K. MacDonald, et al. (2001). "Assessing the abuse potential of methylphenidate in nonhuman and human subjects: a review." Pharmacol Biochem Behav **68**(3): 611-627.

Konstenius, M., N. Jayaram-Lindstrom, et al. (2009). "Sustained release methylphenidate for the treatment of ADHD in amphetamine abusers: A pilot study." Drug Alcohol Depend.

Lowis, D. R. (1997). "HQSAR: A New, Highly Predictive QSAR Technique." Tripos Technical Notes **1**: 1-17.

Marcel, L. V., C. C. Jason, et al. (2003). "Improved protein-ligand docking using GOLD." Proteins: Structure, Function, and Genetics **52**(4): 609-623.

Meltzer, P. C., P. Wang, et al. (2003). "Synthesis and evaluation of dopamine and serotonin transporter inhibition by oxacyclic and carbacyclic analogues of methylphenidate." J Med Chem **46**(8): 1538-1545.

Mutschler, E., G. Geisslinger, et al. (2008). Mutschler Arzneimittelwirkungen Lehrbuch der Pharmakologie und Toxikologie, Wissenschaftliche Verlagsgesellschaft mbH Stuttgart.

Nutt, D., L. A. King, et al. (2007). "Development of a rational scale to assess the harm of drugs of potential misuse." Lancet **369**(9566): 1047-1053.

Parrott, A. C. (2004). "Is ecstasy MDMA? A review of the proportion of ecstasy tablets containing MDMA, their dosage levels, and the changing perceptions of purity." Psychopharmacology (Berl) **173**(3-4): 234-241.

Pifl, C., H. Drobny, et al. (1995). "Mechanism of the dopamine-releasing actions of amphetamine and cocaine: plasmalemmal dopamine transporter versus vesicular monoamine transporter." Mol Pharmacol **47**(2): 368-373.

A References:

Qureshi, A. I., M. F. Suri, et al. (2001). "Cocaine use and the likelihood of nonfatal myocardial infarction and stroke: data from the Third National Health and Nutrition Examination Survey." Circulation **103**(4): 502-506.

Rothman, R. B. and M. H. Baumann (2002). "Therapeutic and adverse actions of serotonin transporter substrates." Pharmacol Ther **95**(1): 73-88.

Schmutz, S. D. (2004). Die Amphetaminbehandlung verhaltensauffälliger Kinder von 1937 bis in die 70er Jahre in Amerika unter besonderer Berücksichtigung der Substanz Methylphenidat (Ritalin®). University of Freiburg.

Schweri, M. M., H. M. Deutsch, et al. (2002). "Biochemical and behavioral characterization of novel methylphenidate analogs." J Pharmacol Exp Ther **301**(2): 527-535.

Seiden, L. S., K. E. Sabol, et al. (1993). "Amphetamine: effects on catecholamine systems and behavior." Annu Rev Pharmacol Toxicol **33**: 639-677.

Senn, C., A. Bucheli, et al. (2007). "Party Drogen [Club drugs]." Ther Umsch **64**(2): 109-113.

Singh, S. K., C. L. Piscitelli, et al. (2008). "A competitive inhibitor traps LeuT in an open-to-out conformation." Science **322**(5908): 1655-1661.

Steinhilber, Roth, et al. (2005). Medizinische Chemie, Deutscher Apotheker Verlag Stuttgart.

Tauss, M. (2008). "Speed zur Wirkungsgeschichte von Amphetaminen." Wiener Zeitschrift für Suchtforschung **Jg 31 2008**: 51-54.

Todeschini, R. and V. Consonni (2000). Handbook of Molecular Descriptors, WILEY-VCH Verlag GmbH, D-69469 Weinheim.

Torres, G. E., R. R. Gainetdinov, et al. (2003). "Plasma membrane monoamine transporters: structure, regulation and function." Nat Rev Neurosci **4**(1): 13-25.

A References:

Totrov, M. and R. Abagyan (2008). "Flexible ligand docking to multiple receptor conformations: a practical alternative." Curr Opin Struct Biol **18**(2): 178-184.

Tripos™Inc. (2007). Sybyl 8.0, Tripos.

Tripos™Inc. (2007) "Sybyl 8.0 Manual: Tripos Bookshelf > HQSAR > Introduction to HQSAR."

Vaughan, R. A., M. L. Parnas, et al. (2005). "Affinity labeling the dopamine transporter ligand binding site." J Neurosci Methods **143**(1): 33-40.

Volz, T. J. and J. O. Schenk (2005). "A comprehensive atlas of the topography of functional groups of the dopamine transporter." Synapse **58**(2): 72-94.

Weissensteiner, R. (2008). Molecular Modelling Studies for Analysing Interactions of Substrates and Inhibitors of the Serotonin Transporter. Department of Medicinal Chemistry. Vienna, University of Vienna. **Magister Thesis**.

Yang, T., C. M. Breneman, et al. (2007). "Investigation of multi-modal high-salt binding ion-exchange chromatography using quantitative structure-property relationship modeling." J Chromatogr A **1175**(1): 96-105.

B Abstract

The Dopamine Transporter which terminates a dopaminergic neuronal signal by withdrawing the monoamine neurotransmitter Dopamine from the synaptic cleft is the target of many psychoactive compounds. Methylphenidate blocks this transporter very selectively in respect to other monoamine transporters. While Methylphenidate's pharmacological effects are well studied and it is also used widely in therapy the molecular basis of its interaction is still uncovered.

In this study QSAR analyses using different sets of 2D descriptors as well as Hologram-QSAR analyses were performed in the MOE and SYBYL Software packages to point out the influence of different structural modifications of methylphenidate derivatives on their activity. The gained information was also used for interpretational purposes in combination with docking experiments. For the docking experiments a workflow was created, composed of the docking process with flexible side chains in the GOLD Suite, a step of minimization of the complexes, RMSD clustering of the received poses and the evaluation in the MOE Software package. A homology model based on the bacterial Leucine transporter had to be used, since currently no X-Ray structure of the Dopamine transporter is available. Six different ligands, methylphenidate and 5 derivatives, all with relatively high affinity (3.6- 130nM), were docked. A common binding mode for all six docked ligands could be found in Dopamine's binding pocket.

To validate the positioning it was checked if the activity of active and inactive methylphenidate derivatives can be explained with the proposed interactions in the pocket. Furthermore it was checked if observations from the QSAR experiments could be explained by the docking model.

Although there are still open questions concerning the activity of some analogues, which have to be investigated in further experiments, many trends observed in QSAR–Analyses were explained in the docking pose. The proposed binding mode shows interactions with the amino acids Asp79, Tyr156, Phe320, Phe326 and Ser422. These residues are proposed for following mutagenesis studies.

C Zusammenfassung

Der Dopamintransporter, welcher dopaminerge neuronale Signale durch die Wiederaufnahme des Neurotransmitters Dopamin aus dem synaptischen Spalt in das Neuron beendet, ist das Ziel vieler psychoaktiver Verbindungen. Methylphenidat blockiert diesen Transporter sehr selektiv im Bezug auf andere Monoamin-Transporter. Während Methylphenidats pharmakologische Wirkung gut untersucht und es auch in der Therapie weit verbreitet ist, bleibt sein genauer Interaktionsmodus immer noch ungeklärt.

In dieser Arbeit wurden QSAR Analysen unter der Verwendung von 2D Deskriptoren sowie Hologramm-QSAR Analysen mithilfe der Softwarepakete „MOE“ und „Sybyl“ durchgeführt, um die strukturellen Einflüsse verschiedener Methylphenidat Derivatisierungen auf die Aktivität zu untersuchen. Die gewonnen Erkenntnisse wurden anschließend zur Interpretation von Docking Experimente verwendet. Für die Docking Experimente wurde ein Workflow erstellt der das Docken mit flexiblen Seitenketten in der GOLD Suite, einem Komplex-Minimierungsschritt, RMSD Clustering der erhaltenen Posen und die Auswertung im MOE Softwarepaket umfasst. Es wurde ein Homologiemodell auf Basis des bakteriellen Leucin Transporters verwendet, da zurzeit keine Röntgenstruktur des DAT verfügbar ist. Methylphenidat und fünf Derivate mit relativ hoher Aktivität (3.6-130nM) wurden gedockt. Dabei konnte ein gemeinsamer Bindungsmodus für alle sechs Liganden gefunden werden.

Um diese Positionierung zu validieren wurde überprüft ob die Affinitätswerte aktiver sowie inaktiver Derivate durch Interaktionen und Hindernisse im vorgeschlagenen Modell erklärbar sind. Außerdem wurde überprüft ob Erkenntnisse aus den QSAR Analysen im Docking Model wiedergefunden werden können. Obwohl immer noch Fragen die die Aktivität einiger Derivate betreffen offen sind, die in weiteren Experimenten geklärt werden müssen, konnten viele Trends der QSAR-Analysen über die Docking Posen erklärt werden. Der Vorgeschlagene Bindungsmodus zeigt Interaktionen mit Asp79, Tyr156, Phe320, Phe326 und Ser422, welche als Ziel für nachfolgende Mutationsexperimente vorgeschlagen werden.

D Abbreviations:

BB	(Protein) Back Bone
DAT	The Dopamine Transporter
MOE	Molecular Operating Environment
NET	The Noradrenalin Transporter
Nr. C.	Number of Compounds used in a QSAR Model
PLIF	Protein Ligand Interaction Fingerprint
RMSD	Root Mean Square Distance
RMSE	Root Mean Square Error
R2	Correlation Coefficient
SERT	The Serotonin Transporter
SLD	Small Ligand Docking
TMD	Transmembrane domain
XRMSE	Cross-validated Root Mean Square Error
XR2	Cross-validated Correlation Coefficient

E Molecular Descriptors:

Descriptor	Properties	Rescource
diameter	"Largest vertex eccentricity in graph"	(Chemical Computing Group 2009)
Es	Taft Size parameter → The bigger the residue the more negative is his Es Constant	(Hansch 1979),(Todeschini and Consonni 2000)
MR	The molar refractivity is calculated considering the weight, the polarizability and the refractivity index. The descriptor shows the contribution of aromatic substituent's to the compounds MR.	(Hansch 1979),(Todeschini and Consonni 2000)
Pi	Aromatic substituent's hydrophobicity constant: Residues which increase lipophilicity have positive values, Residues which decrease it have negative values	(Hansch 1979)
PEOE_VSA-1	Total VDW surface area of a group of atoms with a certain partial charge (-0.1, -0.05)	(Yang et al. 2007)
vdw_vol	Van der Waals Volume	(Chemical Computing Group 2009)
weinerPath	A topological index descriptor. Calculated on bases of the topological distances within the chemical graph of the structure.	(Chemical Computing Group 2009), (Todeschini and Consonni 2000)

E Molecular Descriptors:

Weight	molecular weight	(Chemical Computing Group 2009)
Zagreb	The Zagreb index belongs to the group of topological indices. It describes a molecules topology and it is calculated based on the sum of squared vertex valences of the molecular graph of the molecule.	(Bonchev 1983; Todeschini and Consonni 2000; Chemical Computing Group 2009)

F Curriculum vitae:

Name:	Martin Wurzer
Date of Birth:	September 9th 1985
Nationality:	Austria
Address:	Waldzeile 6
	1170 Wien
	Österreich

Education:

Sep 1992- Jun 1996	Elementary School: Öffentliche VS Kreindlgasse
	Kreindlgasse 24, 1190 Vienna
Sep 1996- Jun 2004	Secondary School: Bundesgymnasium 19
	Gymnasiumstraße 83, 1190 Vienna
Oct 2005 – Jun 2010:	Diploma study Pharmacy at the University of Vienna
	Althanstr. 14, 1090 Vienna
June 2009	Diploma Thesis: Department of Medicinal Chemistry
	Althanstr. 14, 1090 Vienna

Military Service:

Oct 2004 – May 2005:	Wiener Garde

Working experience:

Aug 2005 -Today:	Internship and following part time work at “Apotheke zur Guten Hoffnung” 1190 Wien, Rathstraße 16
Oct 2008-Nov 2009:	University tutor in the laboratory “PR Allgemeine Mikrobiologie und Hygiene“

G Compounds used for the QSAR Models

Detailed structures: Section (H Methylphenidate derivative Database)

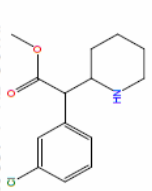
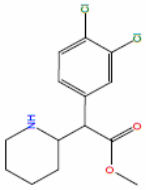
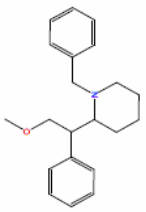
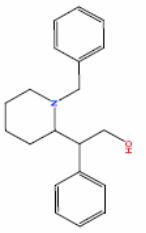
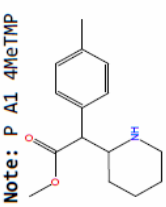
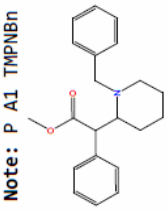
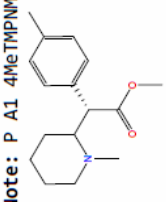
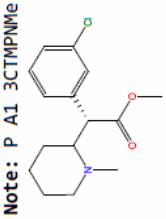
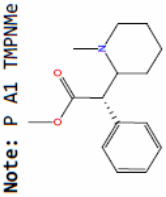
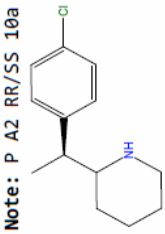
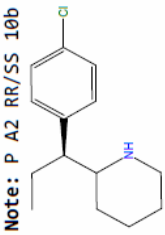
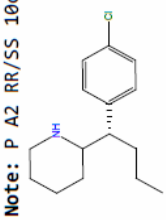
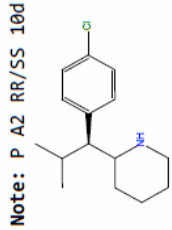
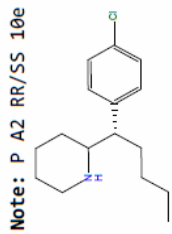
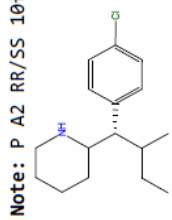
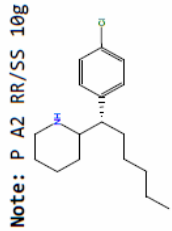
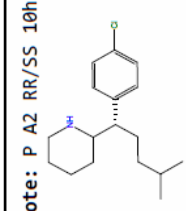
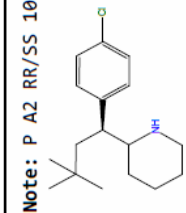
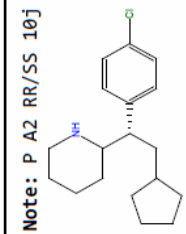
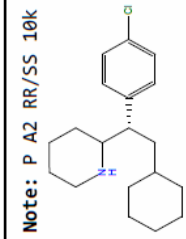
Model 1 (P A7): The compounds mainly differ in the substitution of the phenyl ring.

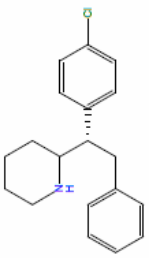
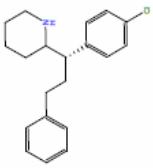
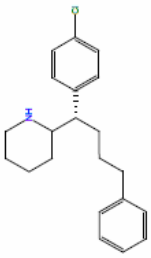
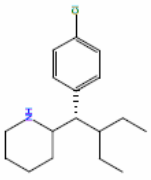
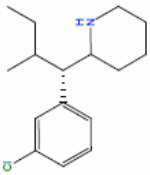
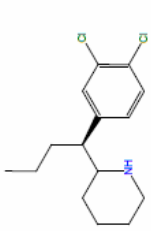
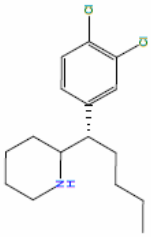
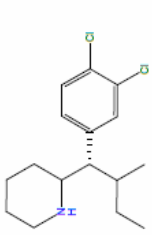
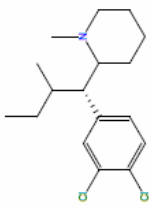
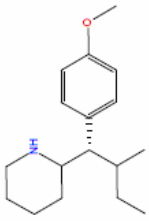
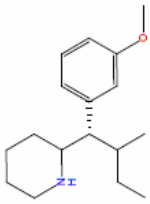
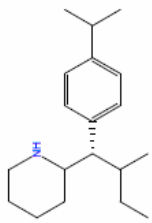
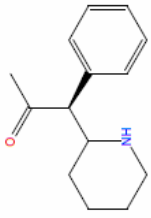
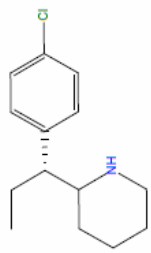
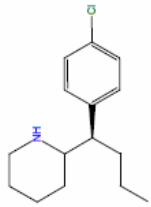
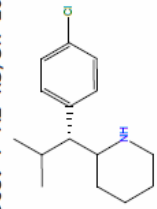
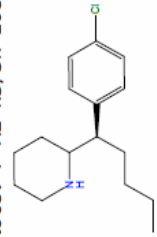
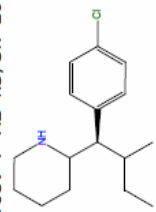
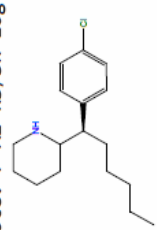
P A7 1aa, P A7 1b, P A7 1d, P A7 1e, P A7 1f, P A7 1g, P A7 1i, P A7 1j, P A7 1k, P A7 1l, P A7 1m, P A7 1n, P A7 1o, P A7 1p, P A7 1q, P A7 1r, P A7 1s RAC SS, P A7 1t P A7 1u, P A7 1v, P A7 1w, P A7 1x, P A7 1y, P A7 1z.

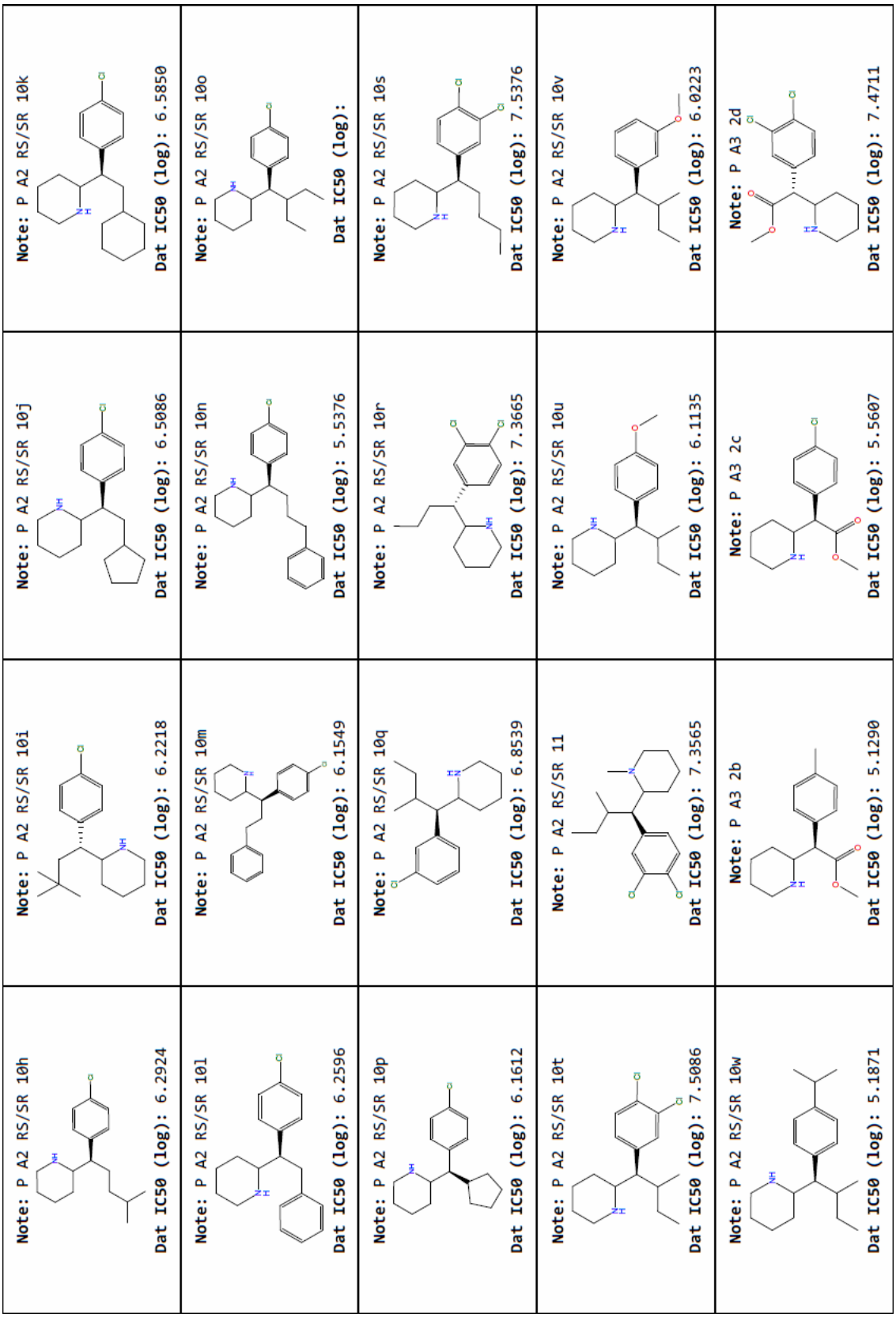
Model 2 (P A2): The compounds mainly differ in the alkyl substitution which replace the carboxy-methyl-ester group of methylphenidate:

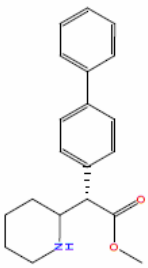
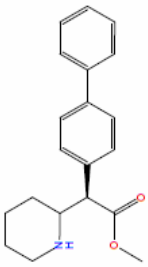
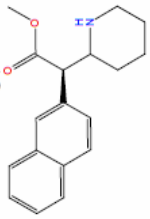
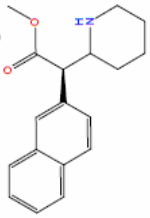
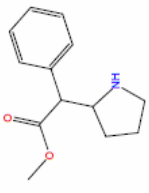
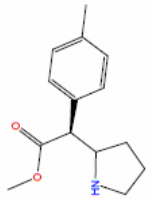
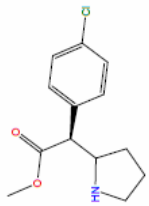
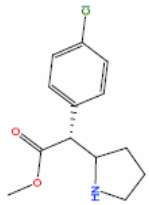
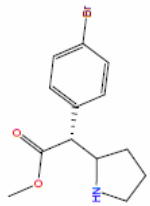
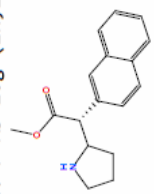
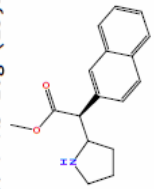
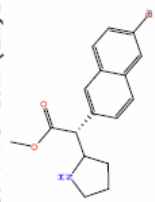
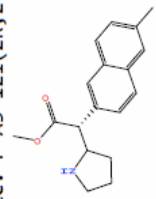
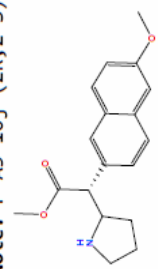
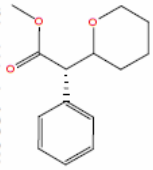
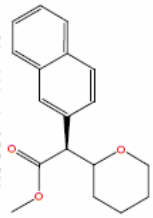
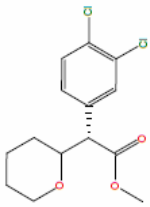
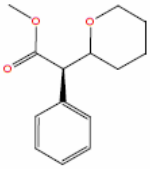
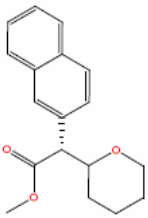
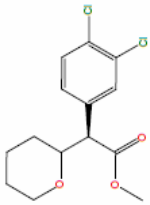
P A2 RR/SS 10a, P A2 RR/SS 10b, P A2 RR/SS 10c, P A2 RR/SS 10d, P A2 RR/SS 10e, P A2 RR/SS10f, P A2 RR/SS 10g, P A2 RR/SS 10h, P A2 RR/SS 10i, P A2 RR/SS 10j, P A2 RR/SS 10k, P A2 RR/SS 10l, P A2 RR/SS 10m, P A2 RR/SS 10n, P A2 RR/SS 10o, P A2 RR/SS 10p, P A2 RR/SS 10q, P A2 RR/SS 10r, P A2 RR/SS 10s, P A2 RR/SS 10t, P A2 RR/SS 11, P A2 RR/SS 11u, P A2 RR/SS 10v, P A2 RR/SS 10w, P A2 RR/SS 15

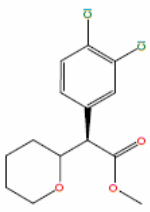
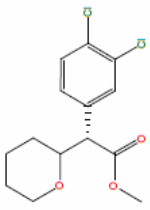
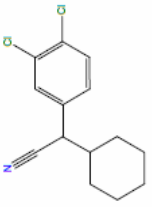
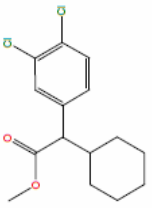
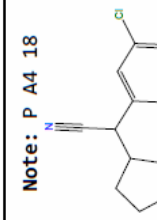
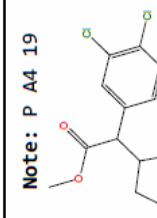
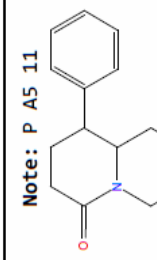
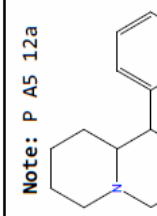
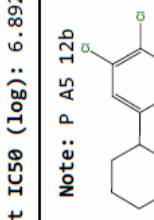
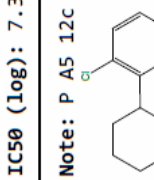
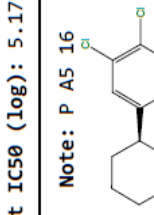
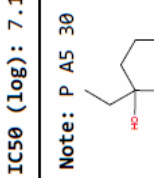
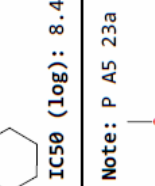
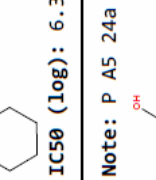
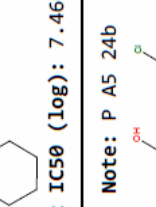
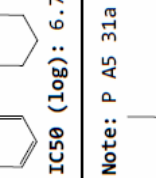
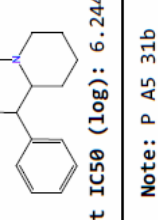
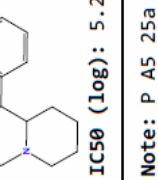
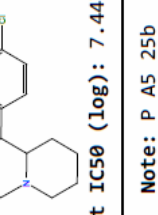
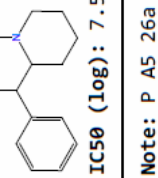
H Methylphenidate derivative Database

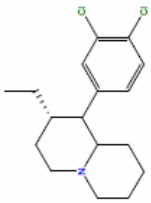
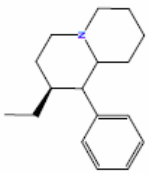
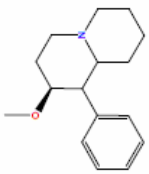
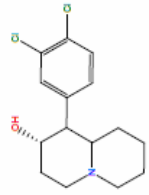
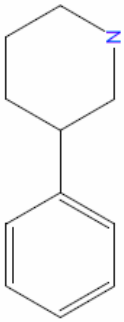
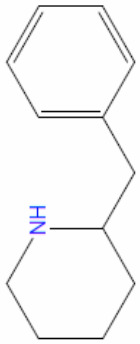
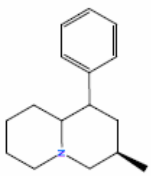
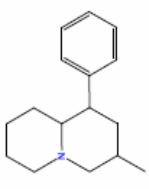
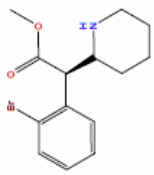
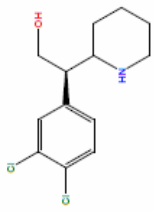
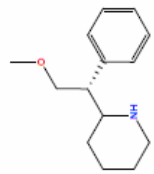
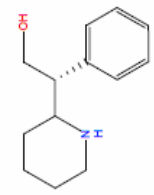
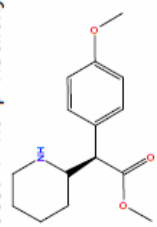
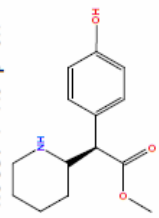
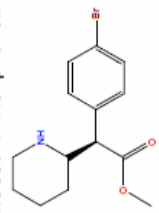
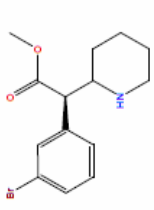
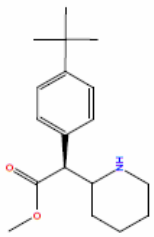
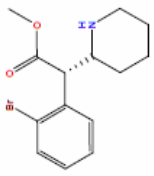
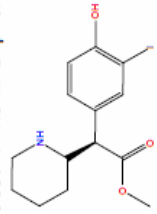
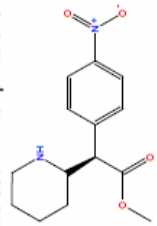
Note: P A1 3CTMP  Dat IC50 (log): 8.2924	Note: P A1 3,4CTMP  Dat IC50 (log): 8.2757	Note: P A1 TROMeNBn  Dat IC50 (log): 7.7496	Note: P A1 TROHNBn  Dat IC50 (log): 7.6253
Note: P A1 4MeTMP  Dat IC50 (log): 7.4815	Note: P A1 TMPNBn  Dat IC50 (log): 7.2765	Note: P A1 4MeTMPNMe  Dat IC50 (log): 6.8539	Note: P A1 3CTMPNMe  Dat IC50 (log): 6.7959
Note: P A1 TMPNMe  Dat IC50 (log): 6.3019	Note: P A2 RR/SS 10a  Dat IC50 (log): 6.7447	Note: P A2 RR/SS 10b  Dat IC50 (log): 7.4318	Note: P A2 RR/SS 10c  Dat IC50 (log): 7.9586
Note: P A2 RR/SS 10d  Dat IC50 (log): 7.3372	Note: P A2 RR/SS 10e  Dat IC50 (log): 8.1079	Note: P A2 RR/SS 10f  Dat IC50 (log): 7.7959	Note: P A2 RR/SS 10g  Dat IC50 (log): 7.6383
Note: P A2 RR/SS 10h  Dat IC50 (log): 8.4437	Note: P A2 RR/SS 10i  Dat IC50 (log): 6.9208	Note: P A2 RR/SS 10j  Dat IC50 (log): 8.0269	Note: P A2 RR/SS 10k  Dat IC50 (log): 6.8861

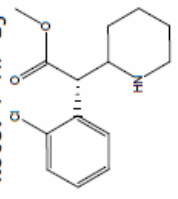
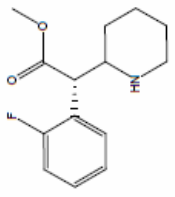
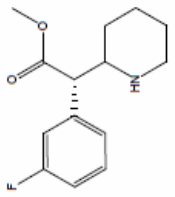
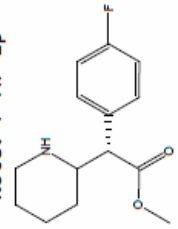
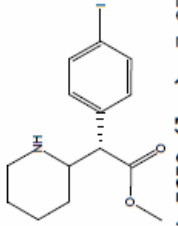
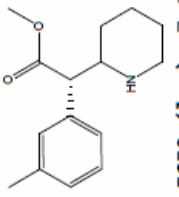
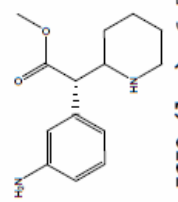
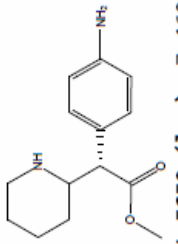
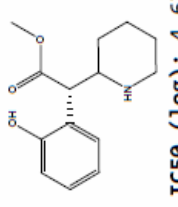
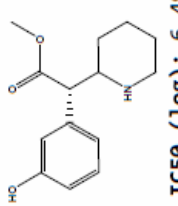
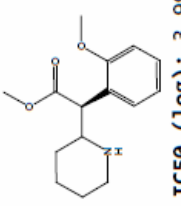
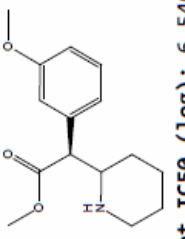
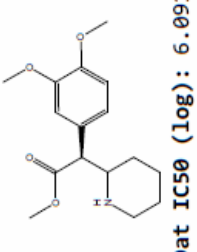
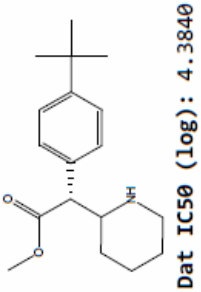
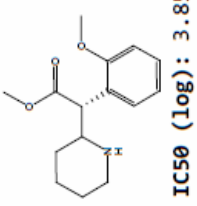
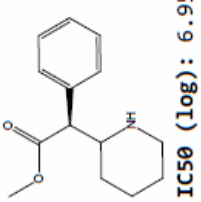
Note: P A2 RR/SS 101  Dat IC50 (log): 6.3565	Note: P A2 RR/SS 10m  Dat IC50 (log): 7.6198	Note: P A2 RR/SS 10n  Dat IC50 (log): 6.3565	Note: P A2 RR/SS 10o  Dat IC50 (log): 6.3979
Note: P A2 RR/SS 10p Dat IC50 (log): 7.4437	Note: P A2 RR/SS 10q  Dat IC50 (log): 8.4318	Note: P A2 RR/SS 10r  Dat IC50 (log): 9.0132	Note: P A2 RR/SS 10s  Dat IC50 (log): 8.6383
Note: P A2 RR/SS 10t  Dat IC50 (log): 9.0000	Note: P A2 RR/SS 11  Dat IC50 (log): 8.1805	Note: P A2 RR/SS 11u  Dat IC50 (log): 7.2840	Note: P A2 RR/SS 10v  Dat IC50 (log): 7.6576
Note: P A2 RR/SS 10w  Dat IC50 (log): 5.4815	Note: P A2 RR/SS 15  Dat IC50 (log): 6.4318	Note: P A2 RS/SR 10b  Dat IC50 (log): 5.7447	Note: P A2 RS/SR 10c  Dat IC50 (log): 6.4202
Note: P A2 RS/SR 10d  Dat IC50 (log): 6.0458	Note: P A2 RS/SR 10e  Dat IC50 (log): 6.5376	Note: P A2 RS/SR 10f  Dat IC50 (log): 6.7696	Note: P A2 RS/SR 10g  Dat IC50 (log): 6.0605


Note: P A2 RS/SR 10k
Dat IC50 (log): 6.5850
Note: P A2 RS/SR 10o
Dat IC50 (log):
Note: P A2 RS/SR 10s
Dat IC50 (log): 7.5376
Note: P A2 RS/SR 10v
Dat IC50 (log): 6.0223
Note: P A3 2d
Dat IC50 (log): 7.4711
Note: P A2 RS/SR 10j
Dat IC50 (log): 6.5086
Note: P A2 RS/SR 10n
Dat IC50 (log): 5.5376
Note: P A2 RS/SR 10r
Dat IC50 (log): 7.3665
Note: P A2 RS/SR 10u
Dat IC50 (log): 6.1135
Note: P A3 2c
Dat IC50 (log): 5.5607
Note: P A2 RS/SR 10i
Dat IC50 (log): 6.2218
Note: P A2 RS/SR 10m
Dat IC50 (log): 6.1549
Note: P A2 RS/SR 10q
Dat IC50 (log): 6.8539
Note: P A2 RS/SR 11
Dat IC50 (log): 7.3565
Note: P A3 2b
Dat IC50 (log): 5.1290
Note: P A2 RS/SR 10h
Dat IC50 (log): 6.2924
Note: P A2 RS/SR 10l
Dat IC50 (log): 6.2596
Note: P A2 RS/SR 10p
Dat IC50 (log): 6.1612
Note: P A2 RS/SR 10t
Dat IC50 (log): 7.5086
Note: P A2 RS/SR 10w
Dat IC50 (log): 5.1871

Note: P A3 1e  Dat IC50 (log): 5.9914	Note: P A3 2e  Dat IC50 (log): 5.1891	Note: P A3 1g(2R,2R')  Dat IC50 (log): 7.9547	Note: P A3 2g  Dat IC50 (log): 5.6819
Note: P A3 9a  Dat IC50 (log): 5.5003	Note: P A3 9b  Dat IC50 (log): 5.7144	Note: P A3 9c  Dat IC50 (log): 6.5654	Note: P A3 10c  Dat IC50 (log):
Note: P A3 10f  Dat IC50 (log): 6.1959	Note: P A3 10g (2R,2'S)  Dat IC50 (log): 6.3325	Note: P A3 10g (2S,2'R)  Dat IC50 (log): 5.7570	Note: P A3 10h (2R,2'S)  Dat IC50 (log): 6.3507
Note: P A3 12i(2R,2'S)  Dat IC50 (log): 6.0706	Note: P A3 10j (2R,2'S)  Dat IC50 (log): 6.0246	Note: P A4 4a  Dat IC50 (log): 5.5229	Note: P A4 4b  Dat IC50 (log): 6.3665
Note: P A4 12  Dat IC50 (log): 7.4685	Note: P A4 5a  Dat IC50 (log): 4.8861	Note: P A4 5b  Dat IC50 (log): 5.3979	Note: P A4 15  Dat IC50 (log): 6.7144

Note: P A4 13  Dat IC50 (log): 7.7696	Note: P A4 14  Dat IC50 (log): 6.1331	Note: P A4 16  Dat IC50 (log): 6.8962	Note: P A4 17  Dat IC50 (log): 6.8356
Note: P A4 18  Dat IC50 (log): 6.8928	Note: P A4 19  Dat IC50 (log): 7.3279	Note: P A5 11  Dat IC50 (log): 5.1798	Note: P A5 12a  Dat IC50 (log): 7.1180
Note: P A5 12b  Dat IC50 (log): 8.4698	Note: P A5 12c  Dat IC50 (log): 6.3188	Note: P A5 16  Dat IC50 (log): 7.4609	Note: P A5 30  Dat IC50 (log): 6.7570
Note: P A5 23a  Dat IC50 (log): 6.2441	Note: P A5 24a  Dat IC50 (log): 5.2041	Note: P A5 24b  Dat IC50 (log): 7.4473	Note: P A5 31a  Dat IC50 (log): 7.5591
Note: P A5 31b  Dat IC50 (log): 8.4634	Note: P A5 25a  Dat IC50 (log): 6.0419	Note: P A5 25b  Dat IC50 (log): 7.8539	Note: P A5 26a  Dat IC50 (log): 6.9666

Note: P 5A 33b  Dat IC50 (log): 8.3851	Note: P A5 33a  Dat IC50 (log): 8.2588	Note: P A5 27a  Dat IC50 (log): 7.9666	Note: P A5 26b  Dat IC50 (log): 8.6091
Note: P 5A 35  Dat IC50 (log): 5.3410	Note: P 5A 34  Dat IC50 (log): 5.1965	Note: P 5A 29b  Dat IC50 (log): 6.9245	Note: P 5A 29a  Dat IC50 (log): 6.7496
Note: P A6 O-Brom  Dat IC50 (log): 6.0555	Note: P 5A 36c  Dat IC50 (log): 8.4078	Note: P 5A 36b  Dat IC50 (log): 7.0462	Note: P 5A 36a  Dat IC50 (log): 6.3915
Note: P A6 p-Methoxy  Dat IC50 (log): 7.3768	Note: P A6 p-OH  Dat IC50 (log): 6.9031	Note: P A6 p-Brom  Dat IC50 (log): 7.6778	Note: P A6 m-Brom  Dat IC50 (log): 8.3979
Note: P A7 1i  Dat IC50 (log): 4.8697	Note: P A6 o-Brom AFF m A7!  Dat IC50 (log): 4.6189	Note: P A6 m-I p-OH  Dat IC50 (log): 7.3768	Note: P A6 p-Nitro  Dat IC50 (log): 6.7447

Note: P A7 1j  Dat IC50 (log): 5.7100	Note: P A7 1n  Dat IC50 (log): 5.8477	Note: P A7 1o  Dat IC50 (log): 7.3925	Note: P A7 1p  Dat IC50 (log): 7.4559
Note: P A7 1q  Dat IC50 (log): 7.8539	Note: P A7 1r  Dat IC50 (log): 7.6696	Note: P A7 1t  Dat IC50 (log): 6.5768	Note: P A7 1u  Dat IC50 (log): 7.4609
Note: P A7 1w  Dat IC50 (log): 4.6364	Note: P A7 1x  Dat IC50 (log): 6.4935	Note: P A7 1y  Dat IC50 (log): 3.9957	Note: P A7 1z  Dat IC50 (log): 6.5406
Note: P A7 1aa  Dat IC50 (log): 6.0915	Note: P A7 11a  Dat IC50 (log): 4.3840	Note: P A7 11e  Dat IC50 (log): 3.8570	Note: P A2 1 TMP  Dat IC50 (log): 6.9586