Experimental and computational investigation of affinity and selectivity factors in CYP2D6 and CYP3A4 mediated metabolism

BRITTA BONN

GÖTEBORGS UNIVERSITET

DOCTORAL THESIS

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry

Experimental and computational investigation of affinity and selectivity factors in CYP2D6 and CYP3A4 mediated metabolism

BRITTA BONN

Department of Chemistry
University of Gothenburg
SE-412 96 Gothenburg

ISBN: 978-91-628-8129-0

Printed by Intellecta Infolog AB
Göteborg, 2010

Abstract

The assessment of ADME properties and metabolic behavior of a drug is central in drug discovery and drug design. The main target for studies of metabolic properties is the Cytochrome P450 (CYP), which is responsible for the metabolism of a majority of drugs on the market and consequently also involved in many drug-drug interactions. Examples of information and tools that could guide drug design towards favorable metabolic properties are structural information of the CYPs, affinity and selectivity towards the CYPs and the site of metabolism (SOM). The present work was initiated to use experimental and computational tools to study the affinity and selectivity for two of the most important isoenzymes, CYP2D6 and CYP3A4, and highlight the benefit of combining different approaches and assays to understand the metabolic properties of a drug.

A novel computational approach resulting in an assessment of enzyme-ligand interaction patterns were successfully used for a comparison of different 3D-structures in order to highlight discriminative amino acid residues. This method could also offer an identification of important interactions when understanding affinity and selectivity for an enzyme. In order to explore affinity and selectivity factors in CYP3A4 and CYP2D6 mediated metabolism two compounds were selected for an experimental evaluation of the catalytic properties of the enzyme. Based on the results from different *in vitro* assays it could be concluded that CYP3A4 was more unselective, producing metabolites as a result of orientations presenting many different parts of the molecules to the heme. CYP2D6 on the other hand showed more restricted binding modes. The combined information from inhibition studies and metabolite identification also gave indications on productive and non-productive binding modes in the two enzymes. With further exploration of CYP2D6, and its pharmacophore, *N*-dealkylation and the effect of blocking the SOM in CYP2D6 substrates were studied. These results were in agreement with the previously stated pharmacophore and also showed that the SOM for these substrates could be successfully assigned with a ligand-based approach, which could also be used to assign selectivity for CYP2D6.

Key information, in designing compounds towards preferable metabolic properties, is obtained from metabolite identification or SOM determinations. I order to enhance metabolite identification a semi-automated software was tested that assigned the metabolite structure from MS raw data with high success rate. This could consequently be beneficial for drug design in that it enables a high throughput of metabolite identification data. During the course of these work important aspects to consider in drug design has been observed, e.g. improving metabolic stability by blocking soft spots tend to result in CYP inhibition. Instead of blocking soft spots the affinity for an enzyme could be diminished and this could be guided by the computational approach used in the first project. In summary, the combined knowledge from *in vitro* and *in silico* tools could be beneficial for the understanding of the metabolic behavior of a drug.

Keywords: Cytochrome P450, *in vitro* metabolism, inhibition, biotransformation, computational modeling, drug design

List of publications

This thesis is based on the following publications, referred to in the text by the corresponding Roman numerals:

I **Exploration of Enzyme-Ligand Interactions in CYP2D6 & 3A4 Homology Models and Crystal Structures Using a Novel Computational Approach**
Kjellander B., Masimirembwa C.M., and Zamora I.
Journal of Chemical Information and Modeling **2007**, 47;1234-1247

II **Exploration of Catalytic Properties of CYP2D6 and CYP3A4 Through Metabolic Studies of Levorphanol and Levallorphan**
Bonn B., Masimirembwa C.M., and Castagnoli N.
Drug Metabolism and Disposition **2010**, 38; 187-199

III **The Molecular Basis of CYP2D6-Mediated *N*-Dealkylation: Balance between Metabolic Clearance Routes and Enzyme Inhibition**
Bonn B., Masimirembwa C.M., Aristei Y., and Zamora I.
Drug Metabolism and Disposition, **2008**, 36; 2199-2210

IV **Enhanced Metabolite Identification with MS^E and a Semi-Automated Software for Structural Elucidation**
Bonn B., Leandersson C., Fontaine F., and Zamora I.
Under revision in Rapid Communications in Mass Spectrometry

The publications are reprinted with permission from the publishers.

Contribution report

Paper I	Minor contribution to the formulation of research problem; performed the experimental and computational work; interpreted the results and wrote the manuscript.
Paper II	Major contribution to the formulation of the research problem; performed all the experimental and computational work; contributed to the interpretation of MS spectra; and wrote the manuscript.
Paper III	Contribution to the formulation of the research problem; performed all the experimental and computational work except for literature search and compilation of CYP2D6 substrates; interpreted the results and wrote the manuscript.
Paper IV	Formulation of the research problem; performed all experimental work and spectral interpretation; wrote the manuscript.

List of Abbreviations

ADME	Absorption, Distribution, Metabolism and Excretion
AMMC	4-aminomethyl-7-methoxycoumarine
BFC	7-benzyloxy-4-trifluoromethylcoumarine
CID	Collision Induced Dissociation
Cl_{int}	Intrinsic clearance
CPCA	Concensus Principal Component Analysis
CPR	Cytochrome P450 Reductase
CYP	Cytochrome P450
DDI	Drug-Drug Interactions
DMPK	Drug Metabolism and Pharmacokinetics
FLAP	Finger Print for Ligand and Proteins
GSH	Glutathione
HLM	Human Liver Microsomes
HPLC	High Pressure Liquid Chromatography
IC_{50}	Inhibitor concentration causing 50% reduced enzyme activity
K_i	Inhibition constant
K_{inact}	Maximum rate of inactivation
K_m	Michaelis-Menten constant
MBI	Mechanism Based Inhibition
MIC	Metabolite Intermediate Complex
MIF	Molecular Interaction Field
MS	Mass Spectrometry
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCE	New Chemical Entity
NR	Normalized Ratio
PCA	Principal Component Analysis
PLS	Partial Least Squares
QM	Quantum Mechanical
QSAR	Quantitative Structure Activity Relationship
RMSD	Root Mean Square Deviation
SMR	Structure Metabolism Relationship
SOM	Site of Metabolism
SRS	Substrate Recognition Sites
TDI	Time Dependent Inhibition
UGT	UDP-glururonosyltransferase
V_{max}	Maximum velocity of enzyme catalyzed reactions

Table of contents

1 Introduction

In the early 1990s the main reason for attrition of drug candidates in drug discovery was due to poor drug metabolism and pharmacokinetic (DMPK) properties. During the following decade this dramatically shifted and in 2000 this number was decreased to less than 10%. Instead there was an increase in failures due to toxicity and safety.[1] The impressive development of *in vitro* assay systems and bioanalytical techniques has greatly improved and a number of **a**bsorption, **d**istribution, **m**etabolism and **e**xcretion (ADME) screening tools have become available, which enable high throughput screening of DMPK properties of new chemical entities (NCEs).[2] This has made it possible to remove compounds with undesired properties early in the drug discovery process. In parallel, the use of *in silico* tools and computational modeling has become more and more common both in medicinal chemistry and within DMPK. Regarding drug metabolism and drug-drug interactions, the availability of homology models, and later, the successful crystallization of human cytochrome P450 (CYP) enzymes, has contributed to a broader knowledge of structure and function of this important super gene family of drug metabolizing enzymes. This has played an important role in the understanding and prediction of DMPK properties in humans.

Even though the ADME screening of NCE has improved the metabolic and kinetic properties of drugs in development, the failures due to efficacy and safety are still an issue. It is therefore essential to consider these different metabolic features together when designing new drugs in order to make the drug discovery cycle more efficient. To achieve this, close collaboration between medicinal chemists and DMPK scientists is of great importance. New trends are also evolving within drug design, i.e. the impact of rational design instead of screening of compound libraries to improve DMPK properties of new compound series. One aim of this strategy is to understand drug-drug interaction (DDIs) and metabolic issues, such as stability or formation of reactive metabolites. Furthermore, the purpose would be to design away from the problems instead of "screening away" from them.

The work in this thesis emphasizes on combining *in vitro* enzyme kinetic parameters, homology models & crystal structures of CYPs and computational modeling to increase our understanding of the metabolic fate of drugs. The use of the combined insights in chemistry, different DMPK properties and enzyme structural information are expected to be a useful approach in molecular design of NCE analogues. When studying the

metabolic properties of a drug, affinity, selectivity and reactivity are considered to be fundamental factors. In this thesis the focus has been put on experimental and computational studies of affinity and selectivity for two metabolizing enzymes, CYP3A4 and CYP2D6, which, together, are responsible for the metabolism of approximately 60% of the most prescribed drugs on the market.[3] In order to introduce the field some background regarding aspects of drug metabolism and computational modeling will be discussed.

2 Metabolism

The pharmacokinetics of an administered drug is determined by its ADME properties, in which metabolism is a key component. The metabolism of a drug, i.e. its biotransformation, is a protection mechanism against chemical insults on the body through the generation of more hydrophilic compounds that can be readily excreted through kidneys and/or bile. Though essentially being a detoxification mechanism, it can also result in bioactivation of drugs with the generation of reactive or toxic species, which can become safety issues in the development of an NCE. Drug metabolism is generally divided into two categories, functionalization (phase I) and conjugation reactions. In phase I polar functionalities are introduced or exposed, resulting in more hydrophilic compounds. This encompasses biotransformations such as oxidation and hydrolysis reactions, but also reductions.[4] These reactions usually create a handle for the conjugation of polar endogenous groups to the metabolite, e.g. glucuronic acid, sulfate, or amino acids, which further facilitate excretion from the body. The conjugation reactions can also have the purpose of detoxification, e.g. conjugation with glutathione, or result in a termination of the effect via e.g. an acetylation or methylation.[4] The phase I reactions are often the rate limiting step for disposition of most drugs and of the phase I enzymes, the cytochrome P450s (CYPs) are essential catalysts for the metabolism of xenobiotics.[5] The drug metabolizing enzymes are found mainly in the liver but are also present in virtually all tissues.[6] An orally administered drug has to pass through the liver after being absorbed from the gastrointestinal tract; accordingly, the metabolism in this organ is one of the key limitations for the bioavailability of a compound.

2.1 Biological and chemical aspects of Cytochrome P450 – A general overview

The superfamily of cytochrome P450 (CYP) enzymes is a family of versatile biological catalysts. The CYPs are promiscuous and their structure is flexible, which consequently enables accommodation of substrates with diverse structures as well as the possibility of multiple orientations and binding modes for a single substrate. This structural/functional adaption enables CYPs to metabolize and dispose most xenobiotics the body is exposed to. Each enzyme's characteristic active site topology and different key interaction partners between enzyme and substrate are important factors for determination of

substrate selectivity, affinity, and reactivity.[7] However, despite characteristic features for the different isoforms of the enzyme system, there is overlapping substrate specificity between them. This may lead to different kinetic spaces, referring to affinity and capacity, when comparing different CYPs. This 'kinetic- space' concept partly explains the body´s concerted efforts to metabolize and dispose xenobiotics, where some CYPs have high affinity and low turnover and others have low affinity and high turnover for the same substrate. The former enzymes are usually responsible for the metabolism and disposition of xenobiotics when at low concentrations but get saturated at high concentrations and then the latter CYPs become key players in the process. This can be exemplified by the metabolism of dextromethorphan by CYP3A4 and CYP2D6 where CYP2D6 acts with high affinity/low turnover and CYP3A4 with low affinity/high turnover.[8] The same relationship, with CYP3A4 acting as a low affinity/high turnover enzyme, is observed in the metabolism of diazepam by CYP3A4 and CYP2C19.[9]

Besides the structural aspects of the CYPs they possess a diversity of catalytic properties with the possibility to catalyze a variety of different reactions. The CYPs predominantly act as monooxygenases but also show enzymatic mechanisms corresponding to reductases, oxidases and hydroperoxidases.[10] This catalytic diversity can be demonstrated by searching the MDL database (Symyx Solutions, Inc.) for CYP mediated reactions, in Figure 2.1 some of the most common reactions and their occurrence in the database are presented.

The CYPs were discovered in the late 1950s[11, 12] and have since then been widely studied due to their central role in metabolism of xeno- and endobiotics. They are heme-thiolate enzymes with the protoporphyrine moiety non-covalently bound to the protein. The heme is common for all CYPs, and is responsible for the catalytic properties of the enzyme family. Omura et al. discovered a characteristic absorption spectrum at 450 nm when carbon monoxide was bound to the reduced heme[13, 14] and this was the origin for the name Cytochrome P450. The genes coding for theses enzymes are divided into families (<40% amino acid identity) and subfamilies (40-55% amino acid identity) which are organized by a universal nomenclature system.[15] Currently this uses Arabic numerals for each family (e.g. CYP3), a capital letter assigning the subfamily (e.g. CYP3A) and number for the individual protein (e.g. CYP3A4). Today there are 58 human CYPs identified[16] and of these, 15 are involved in the metabolism of xenobiotics.[17] These drug-metabolizing CYPs are within the families 1-3 while the other families are involved in the metabolism of endogenous compounds within classes of steroids, fatty acids, eicosanoids and vitamins.[17]

Among the drug metabolizing CYPs the CYP3A family is the most predominant in the liver followed by CYP2C8/9/19, CYP1A2, CYP2E1 and CYP2D6.[18] However, some of these drug metabolizing enzymes show great variability in expression and activity due to genetic polymorphism.[19] This means that for several genes there are allelic variants present and these can express different phenotypes. These are commonly referred to as poor metabolizers (PM), intermediate metabolizers (IM), extensive metabolizers (EM) and ultra rapid metabolizers (UM). Administration of a drug that is a substrate of a

polymorphic enzyme can have a big influence on the exposure depending on the genotype of the patient, leading to interindividual differences. This can consequently result in safety/toxicity and efficacy issues in the use of some drugs in some individuals and some ethnic groups. Major polymorphic enzymes include CYP2D6, 2C9, 2C19, 2C8 and 2A6.[20]

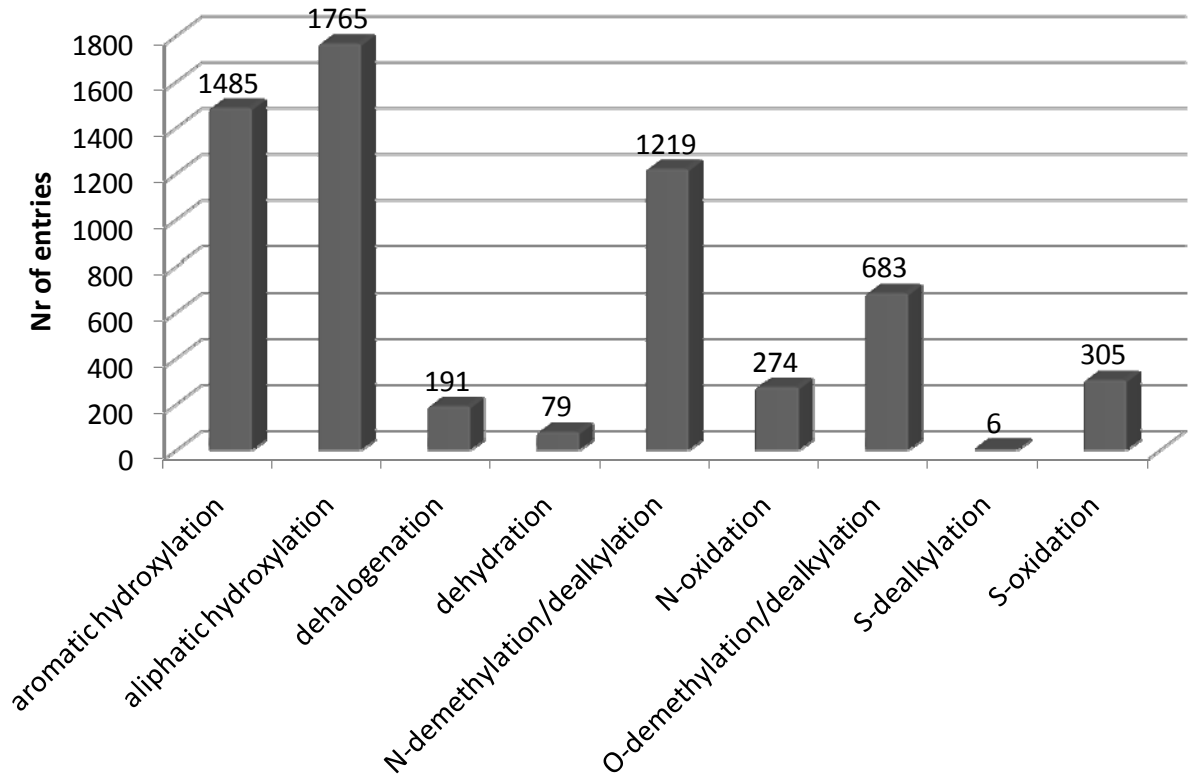


Figure 2.1 Frequency of occurrence of common reactions catalyzed by CYPs found in the MDL database (Symyx Solutions, Inc.)

2.2 Structure of Cytochrome P450 enzymes

As mentioned above the structure of CYPs greatly contribute to the versatile nature of this enzyme family. Structural information and the possibility to assess structure-function data is preferably achieved with the solving of crystal structures. The first CYP crystal structure that was solved was the bacterial P450cam in 1987.[21] Accordingly much of the structural and catalytic understanding of CYPs has been realized from studies of this enzyme. The bacterial CYPs are soluble, as opposed to the membrane bound mammalian counterparts, and therefore easier to crystallize. Additional bacterial structures were solved following this success, e.g. P450BM3,[22] but the crystallization of membrane bound CYPs met with extreme difficulties. There was a long lag period before the first crystal structure of a mammalian CYP was solved in 2000,

the rabbit CYP2C5.[23] This breakthrough was followed by successful crystallization of several human CYPs: 2C9,[24, 25] 2C8,[26] 3A4,[27-29] 2D6[30] 1A2[31] and most recently 2B6.[32]

In the gap between the availability of crystal structures of bacterial CYPs and mammalian CYPs, information about structure and function was achieved with site directed mutagenesis and homology modeling.[33] An important study was done by Gotoh in which six substrate recognition sites (SRS) were defined for the CYP2 families (Table 2.1).[34] This study was performed by a comparative analysis with P450cam, whose SRS were identified from the substrate bound crystal structure.[21] When comparing with the information that later came with the crystallization of mammalian CYPs a lot of this information still seems to hold true.

Table 2.1 Substrate recognition sites (SRS) in CYP2 family[34]

SRS	Structural region
SRS-1	B' helix and flanking areas
SRS-2	C-terminal end of F helix
SRS-3	N-terminal end of G helix
SRS-4	N-terminal half of I helix
SRS-5	β_3 sheet
SRS-6	Central part of β_5 sheet

The overall 3D-structure of CYPs is well conserved. The structure of CYP3A4 is depicted in Figure 2.2. It constitutes of one α-domain and one β-domain. The α-domain is dominated by helical structural elements and contains the heme and its adjacent I and L helices. This domain also comprises the F-G region, and the B' helix, which is involved in substrate recognition. The β-sheet rich region (β-domain) is positioned close to the N-terminal and comprises the A and B helices (on the left side in the structure in Figure 2.2). The different α-helices and β-sheets are connected with loops representing very flexible regions. The secondary structural elements are well conserved across the CYP families, but the relative orientation of these elements is changing between families and even between different structures of one enzyme. The orientation may also depend on the substrate bound. This flexibility can explain the ability of CYPs to accommodate a diversity of substrates, from vitamins and fatty acids to low molecular weight drugs.[35] The highest flexibility is observed within the F and G helices. This was confirmed by comparing enzymes in the CYP2 family[36] but also with different crystal structures of CYP3A4.[29] As could be expected, since all enzymes have the same catalytic moiety, the most conserved regions are those close to the heme, being involved in catalysis and in the binding of the heme.[37] This includes the I and L helices (see Figure 2.2). Prior to the L

helix is the invariant cysteine residue, the axial ligand to the heme, and its environment is important for maintaining the redox potential in the enzyme. On the opposite side of the heme are other well conserved regions, those involved in catalysis and activation of O_2. The region that differs the most between different CYPs is the B' helix, which controls substrate specificity.

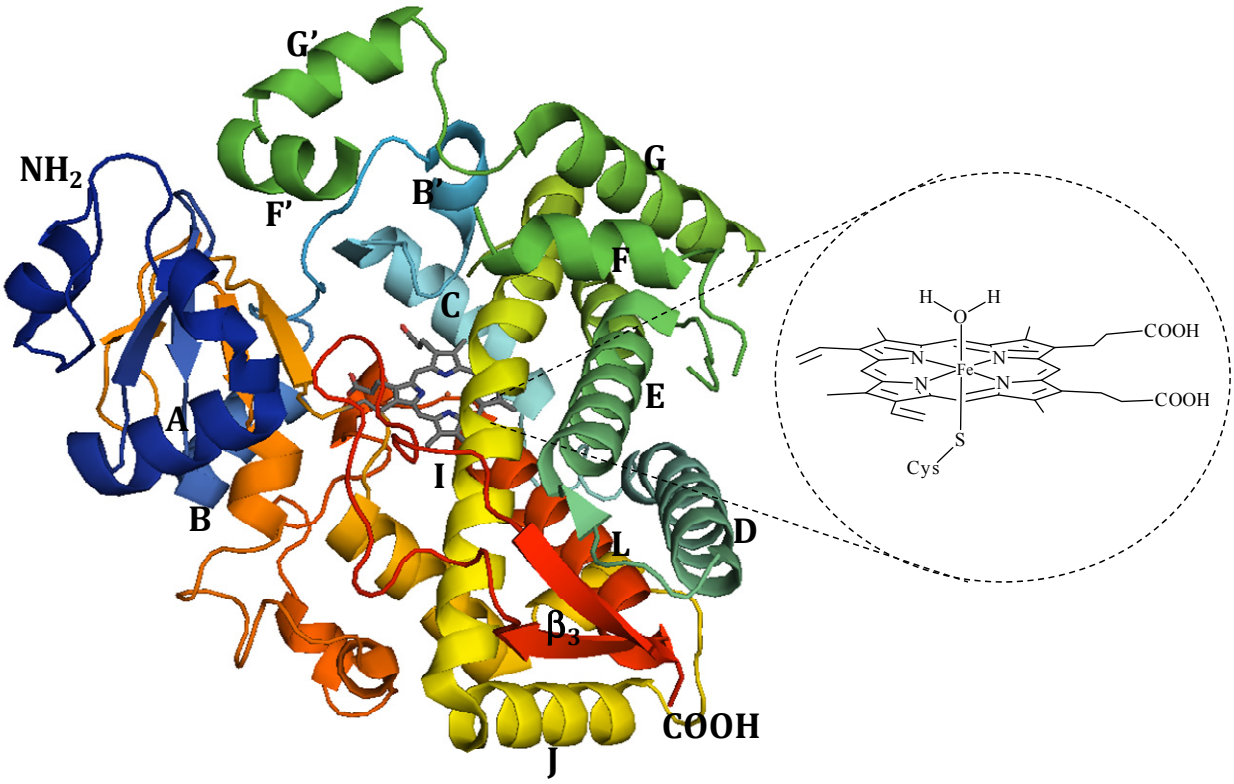


Figure 2.2 Tertiary structure of CYP3A4 (2j0d.pdb) with a close up on the heme.

2.2.1 CYP2D6

CYP2D6 is a polymorphic enzyme, with 5-10% of the Caucasian population being poor metabolizers due to this polymorphism.[38] CYP2D6 has been widely studied and is well characterized regarding structure and function as well as its substrates/inhibitors. The pharmacophore is also well defined and comprises a basic nitrogen 5-7 or 10Å from the site of oxidation, a planar hydrophobic area close to the site of oxidation and a negative electrostatic potential coplanar with the hydrophobic area.[39-41] The active site cavity is relatively small (see Figure 2.4) and the interaction with a ligand is restricted by electrostatic interactions involving two acidic amino acids residues - Glu216, in the F-helix, and Asp301, in the I-helix. With the information from the pharmacophore models substrates/inhibitors are presumed to interact with the basic nitrogen in the ligand. Their importance for substrate recognition and catalytic activity has also been shown with site-directed mutagenesis studies.[42, 43] In addition to these electrostatic interactions, hydrophobic interactions with Phe120 (B'-C loop) and Phe483 (β_{1-4} loop) also play important roles for the regio-specificity and metabolism of CYP2D6

substrates.[44, 45] Some important active site amino acid residues and the CYP2D6 pharmacophore are highlighted in Figure 2.3 where dextromethorphan is docked into the enzyme active site cavity. The contribution of the electrostatic interaction to the binding of the substrate could be an explanation of the relatively high affinity observed for CYP2D6 substrates.[7] This identifies CYP2D6 as a high affinity enzyme with regard to the kinetic space model, at the same time it is relatively easily saturated, i.e. it has a low capacity.

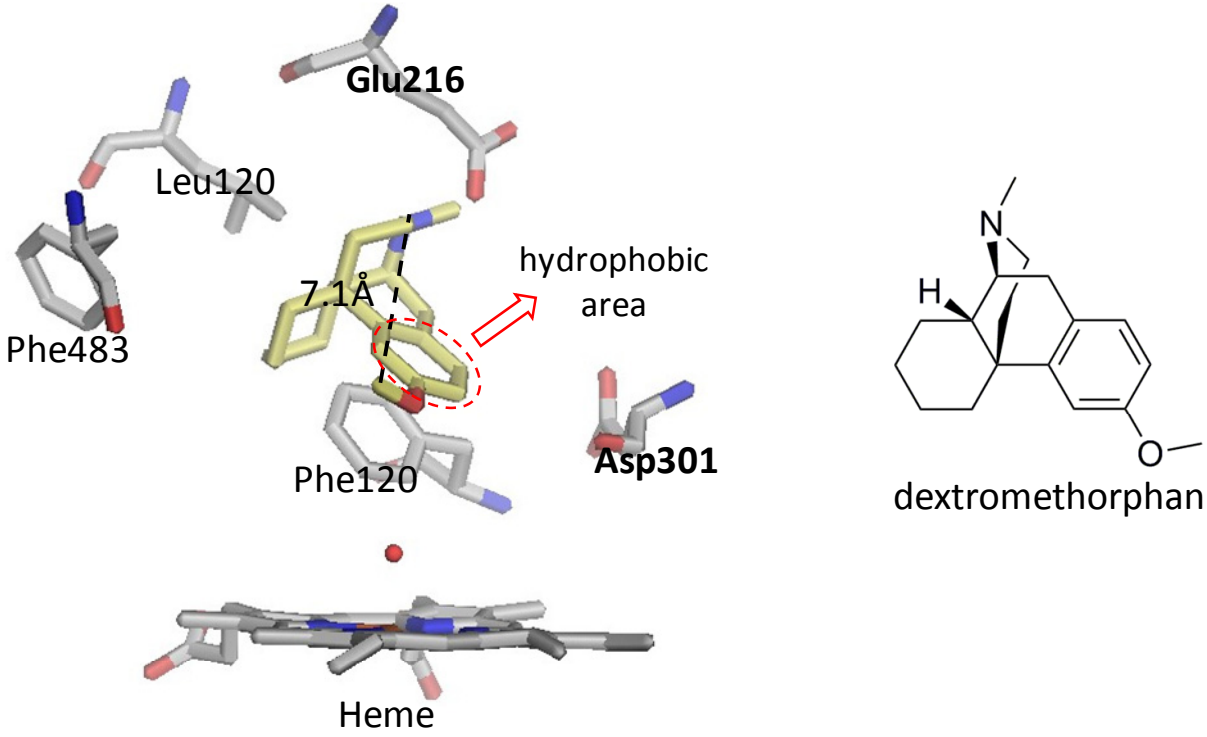
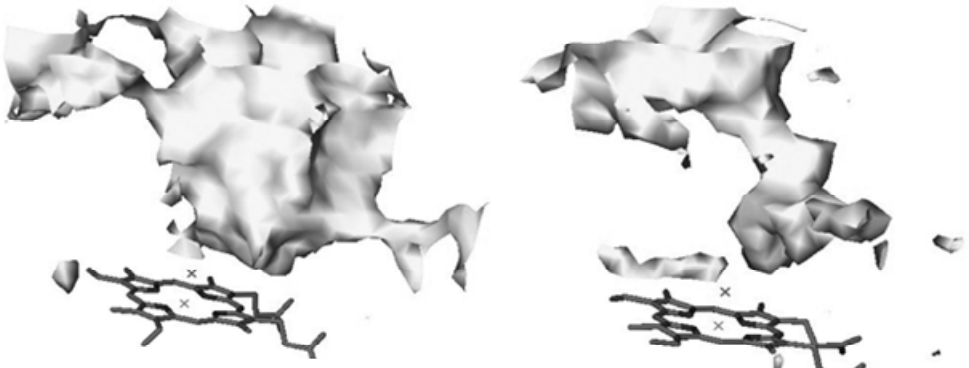


Figure 2.3 Dextromethorphan (yellow) docked into the active site cavity of CYP2D6 with important amino acid residues highlighted. Dextromethorphan exemplifies the CYP2D6 pharmacophore with 7.1Å between the metabolized methoxygroup and the basic nitrogen as well a hydrophobic area close to the site of metabolism.

2.2.2 CYP3A4

CYP3A4 is the most abundant CYP and dominates the biotransformation of xenobiotics.[46] As opposed to CYP2D6 it has low substrate specificity and accommodates a variety of, mostly lipophilic, substrates covering a wide molecular weight range. This is in accordance with the knowledge about the active site cavity, which is relatively big (Figure 2.4) and considerably flexible.[29] Common features of CYP3A4 substrates are a hydrophobic area, two hydrogen bond acceptors 7.7 Å apart and one hydrogen bond donor.[47] The importance of hydrophobic amino acid residues present in the active site has been demonstrated by site-directed mutagenesis studies[48-50] highlighting the presumed dominance of hydrophobic interactions between ligand and enzyme. Apart from the hydrophobic interactions the desolvation component upon binding of a ligand

is probably of importance for substrate binding in an enzyme with a large active site.[7] The nature of the interactions dominating substrate binding in CYP3A4 results in a relatively low affinity enzyme that is not easily saturated. A visualization of CYP2D6 and CYP3A4 active site volumes are presented in Figure 2.4.



2J0D.pdb
(CYP3A4 co-crystallized with erythromycin)

2F9Q.pdb
(CYP2D6 wo ligand)

Figure 2.4 Visualization of the active site cavities of CYP3A4 and CYP2D6 made from calculation of molecular interaction fields with a steric probe in GRID.[51] *The estimated volume of the active site of the CYP3A4 structure is ~2000Å*^{3} [29] *and the CYP2D6 structure ~540Å*^{3} [30]

2.3 Catalytic properties of Cytochrome P450

In addition to the structure of CYPs, their catalytic properties allow for a variety of reactions to be carried out. This is due to the heme and its environment in the active site, which account for the unique catalytic properties of the enzymes. The heme iron is connected to the four pyrrole nitrogens with a cysteine residue as the fifth axial ligand (Figure 2.2). The CYPs are referred to as *mixed function oxidases* or *monooxygenases*[4] and the general reaction involves the transfer of an oxygen atom to the substrate and generation of one molecule of water. The net reaction is described as follows:

$$RH + NADPH + O_2 + H^+ \rightarrow ROH + NADP^+ + H_2O \qquad \textit{Eq. 2.1}^{52}$$

This is a redox-mediated system that contains several components: the CYP, responsible for transfer of oxygen, and the NADPH-dependent cytochrome P450 reductase (CPR) and cytochrome b_5 with their corresponding co-factors that supply electrons. The catalytic cycle of this system is described in more detail below.

2.3.1 Catalytic cycle

A schematic picture of the catalytic cycle is depicted in Figure 2.5. In the resting state, the heme iron is in its ferric state (Fe^{III}) and is hexa-coordinated with a water molecule as a distal ligand. Upon binding of a substrate (RH) the water is displaced (step 1.). In the second step, the ferric iron (Fe^{III}) is reduced by introduction of an electron from CPR. The resulting ferrous state of the iron (Fe^{II}) is capable of binding dioxygen (step 3.). The resulting dioxygen complex is highly unstable and is further reduced by the cytochrome b_5/CPR system (step 4). This step can also be accomplished by direct electron transfer from CPR itself. In step 5 the complex is protonated to afford a hydroperoxy species ($Fe^{III}OOH$), followed by a heterolytic cleavage of the dioxygen bond and loss of water (step 6). The resulting oxyferryl species (FeO^{3+}) is referred to as the activated oxygen species and is considered to be the oxidizing agent in CYP-mediated oxidative reactions, hence the metabolite determining step. The activation of the dioxygen species by protonation of the distal oxygen is supported by a conserved threonine (Thr) residue in the active site. This Thr is suggested to act as a hydrogen bond acceptor interacting with a hydroperoxy intermediate promoting the second protonation step.[53, 54] This is one of the components of a proton shuttle connecting the surface of the enzyme with the ligand. In summary, steps 1-6 in the catalytic cycle of CYPs are responsible for activation of oxygen, they are then followed by an oxidation of the substrate (step 7). Finally the oxidized product is released and the heme returns to its resting state. Different mechanisms accounting for the oxidation of the substrate (step 7) are discussed further below.

2.3.2 Catalytic mechanisms and common reactions

The first six steps in the catalytic cycle describe the activation of oxygen and generation of the oxygenating species FeO^{3+}. The central process of the CYP mediated reactions is however, how this activated species reacts with the substrate (RH) and the formation of the oxidized metabolite. Four general mechanisms have been suggested to account for the different metabolic transformations[55, 56] and these are depicted in Figure 2.6. As seen in Figure 2.1 aliphatic/aromatic hydroxylation and dealkylation reactions are frequently observed within drug metabolism. Aliphatic hydroxylation can be explained by a process involving an initial hydrogen abstraction followed by the so called oxygen rebound (Figure 2.6a). Aromatic hydroxylation is probably initiated by an electrophilic attack by the FeO^{3+} complex (Figure 2.6d). The reaction then proceeds either through a 1,2-migration or via an epoxidation reaction followed by a 1,2-migration. CYPs also catalyze dealkylation reactions (in Figure 2.6 referred to as heteroatom release) that are a consequence of hydroxylation alpha to a heteroatom. The considered heteroatoms are preferably N, O or S. The heteroatom release can be initiated by an aliphatic carbon hydroxylation, described above, but is rather suggested to be initiated by one-electron-transfer.[55] This is followed by hydrogen abstraction and oxygen rebound (Figure 2.6b). Regarding the route for *N* –and *O*-dealkylation the resulting aminol/hemiacetal is

chemically unstable and is cleaved to yield the corresponding amine/alcohol and an aldehyde.

Heteroatom oxygenation has also been rationalized as a one-electron-transfer reaction followed by an additional electron transfer and finally oxygen rebound (Figure 2.6c). Besides oxidative reactions CYPs can also catalyze reductions and hydrolysis, however these reactions are not considered further here.

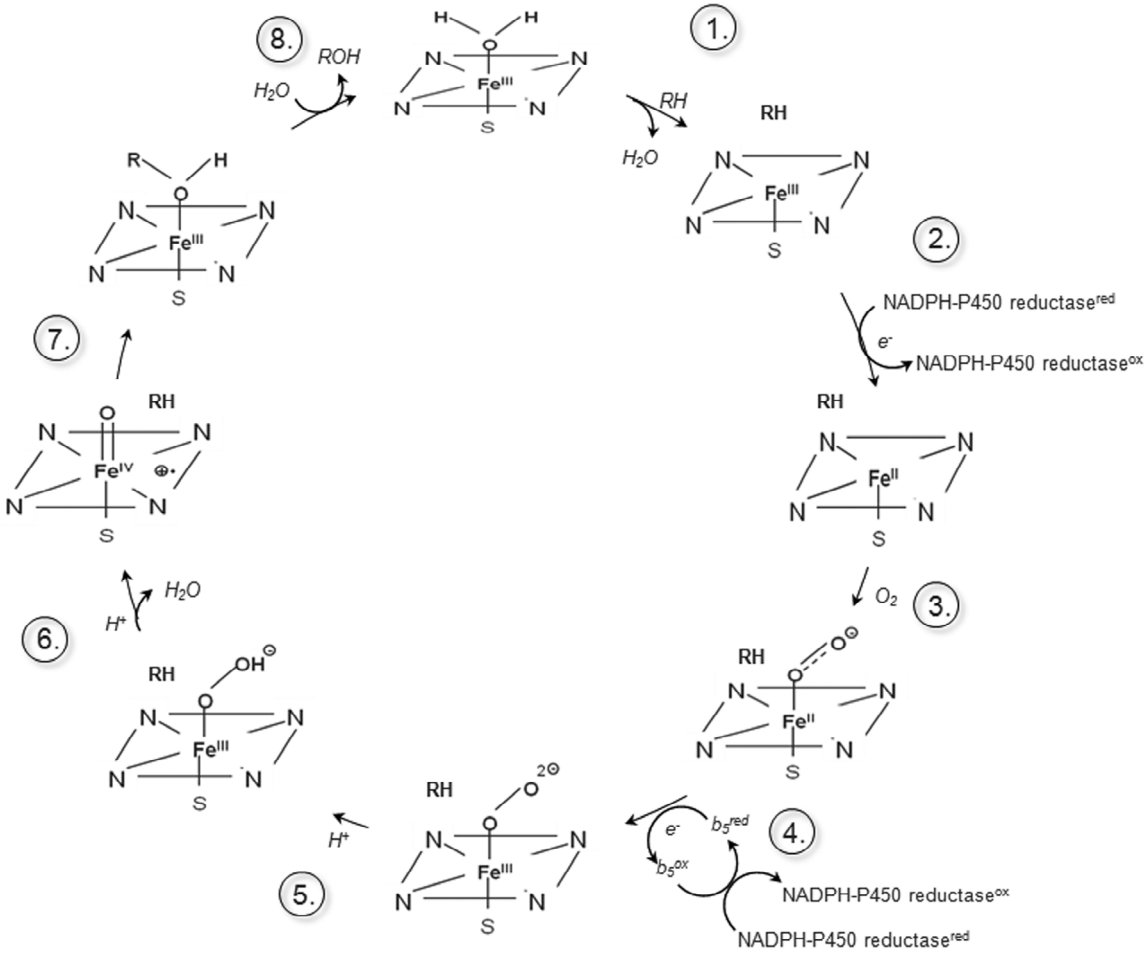


Figure 2.5 The catalytic cycle of cytochrome P450

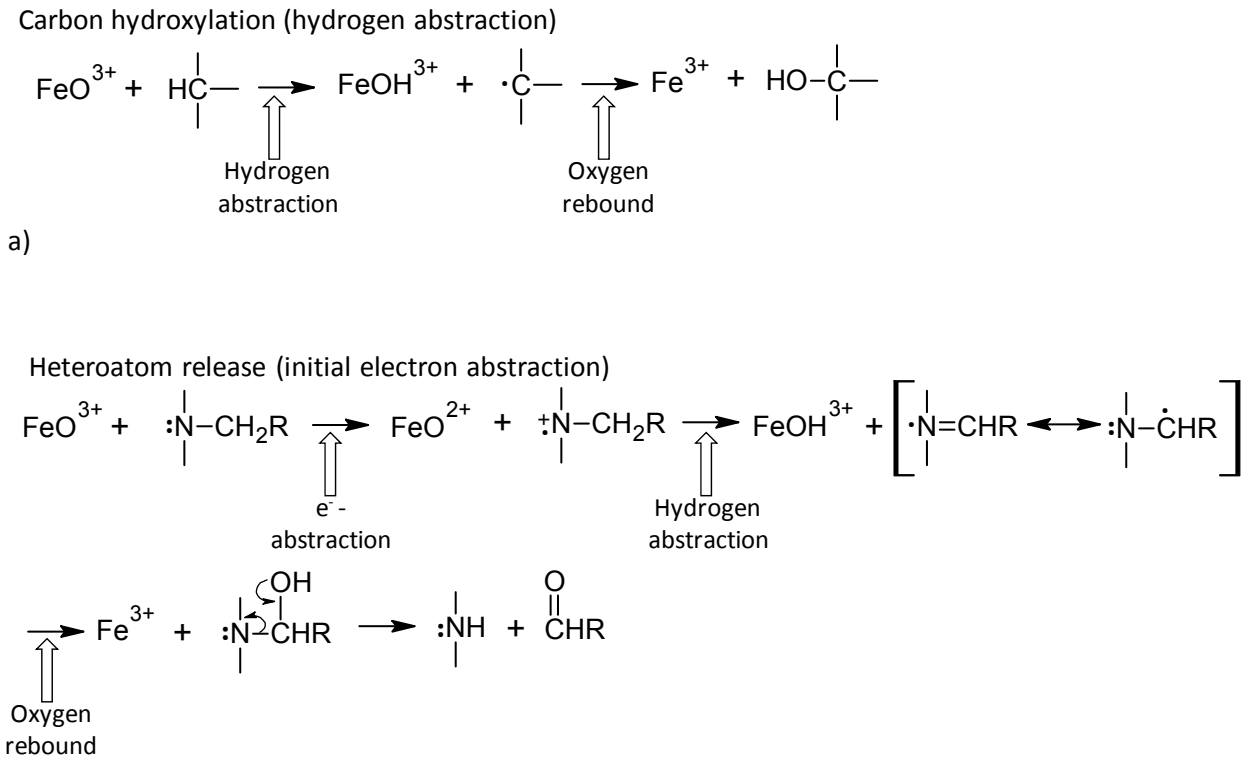


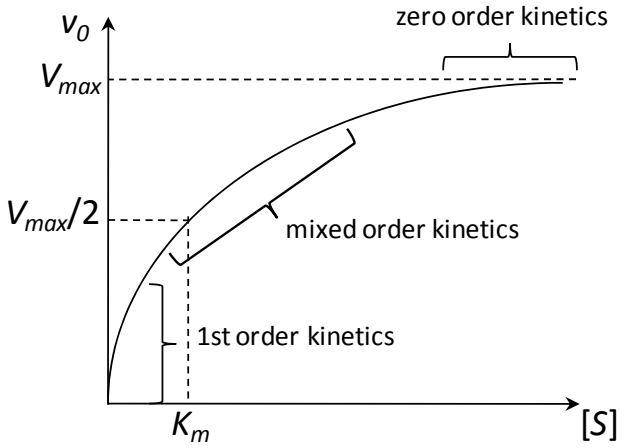
Figure 2.6 General mechanisms for CYP mediated reactions (Adapted from Guengerich 2001[55]) a) carbon hydroxylation, b) heteroatom release, c) heteroatom oxygenation, d) epoxidation and 1,2-group migration.

2.4 Enzyme Kinetics

The enzyme kinetics of CYP-mediated metabolism and inhibition have been observed to follow Michaelis-Menten kinetics, assuming one binding site, or atypical kinetics, assuming two or more binding sites.[57] When including more than one binding site the underlying mechanisms become more complex, however, here only a simplified description assuming Michaelis-Menten kinetics is considered.

2.4.1 Metabolic stability

Metabolic stability and determination of intrinsic clearance (Cl_{int}) is a central assay in drug design and discovery since it determines the rate at which the compound may disappear from the body. The classical concept is based on Michaelis-Menten kinetics describing product formation in an enzyme-catalyzed reaction (Figure 2.7).[57] The hyperbolic curve obtained when plotting the rate of reaction against substrate concentration is shown in Figure 2.7 and can be described by Eq. 2.2. This relationship is only valid using the initial velocity (v_0), i.e. when the substrate concentration is not limiting (less than 10-20% of substrate consumed). As mentioned above it is also assumed that a single substrate and a single substrate-enzyme complex are involved and that no allosteric binding occurs. As can be seen from the plot, a linear relationship between rate of reaction and substrate concentration is observed at low substrate concentrations (1st order kinetics). At high concentrations no change in the rate of reaction is observed with increased substrate concentration, i.e. zero order kinetics is applicable, and the rate at this point is referred to as V_{max}. At this phase, the enzyme's catalytic capacity is said to be saturated. The concentration at 50% of V_{max} equals the K_m-value, which can be used as a measure of the substrate's affinity for the enzyme.



$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Figure 2.7 The concept of an enzyme catalyzed process and a graph describing the change in rate with substrate concentration for a reaction showing Michaelis-Menten kinetics.

$$v = \frac{V_{max} \times S}{K_m + S} \qquad \textit{Eq. 2.2}$$

When using a substrate concentration (*S*) far below K_m the equation describing *v* can be further simplified as a linear relationship dependent on *S* and the V_{max}/K_m ratio.[58] In most cases the V_{max}/K_m ratio can also be referred to as Cl_{int}, a key parameter bridging enzyme kinetics and pharmacokinetics.[57] Cl_{int}, the intrinsic metabolic capacity, is used for the determination of metabolic stability of a drug and defines the rate of metabolism. It can be determined from both product formation and substrate disappearance as described by Eq. 2.3:

$$\frac{V_{max}}{K_m} = Cl_{int} = \frac{ln2 \times V}{t_{1/2}} \qquad \textit{Eq. 2.3}$$

where *V* is the volume of the incubation and $t_{1/2}$ the half-life of the parent compound.

This is a simplified relationship, as some major assumptions have been made: 1) initial velocity is measured i.e. [*S*]>>[*E*], where *E* is the enzyme concentration, and 2) the sampling is done within the linear phase of the Michaelis-Menten curve i.e. [*S*]<<K_m. In this linear phase, the catalytic rate is at its highest and the rate of reaction linearly increases with the increase in substrate concentration. In metabolic stability screening of NCEs a concentration around 1µM is commonly used and applying the above-mentioned calculation of Cl_{int} it is assumed that this is below K_m of the reaction. This is true for most CYP-mediated reactions, but there are exceptions and in these cases the Cl_{int} value will be underestimated.

2.4.2 CYP inhibition

Besides studying how a drug is metabolized by the CYPs it is also of interest to consider its potential to inhibit drug metabolizing enzymes. CYP inhibition is the major cause of drug-drug interactions (DDI) and due to the overlapping substrate specificity this can become an issue when combining several drugs. The consequence of CYP inhibition is increased plasma concentrations for the co-administered drug, which can result in severe effects for drugs with a narrow therapeutic window. Different mechanisms of inhibition have been observed and they are classified as reversible inhibition or irreversible inhibition. The drug whose metabolism is inhibited is referred to as the 'victim' and the inhibitor is referred to as the 'perpetrator'.

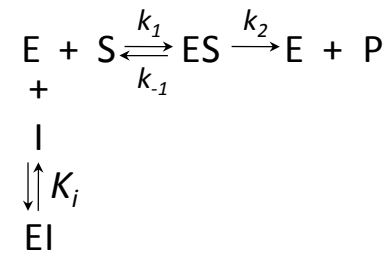
2.4.2.1 Reversible inhibition

Reversible inhibition is divided into competitive, non-competitive/mixed inhibition and uncompetitive inhibition. The type of inhibition can be assigned by analyzing the change in K_m and V_{max} in incubations with and without inhibitor. Competitive inhibition is when the substrate and inhibitor compete for the same site at the enzyme (Scheme 2.1a). This type of inhibition will result in an increase in K_m but V_{max} will remain unchanged. The

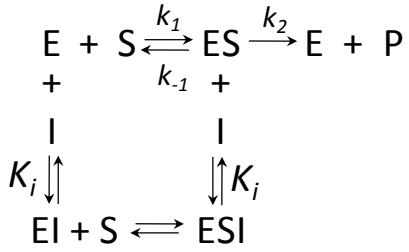
degree of inhibition is determined by the inhibition constant, K_i. However, in screening of NCEs the IC_{50}-value is commonly used as a measure of the inhibition potential. The IC_{50}-value corresponds to the inhibitor concentration resulting in 50% reduced activity for the probe substrate. For competitive inhibition the K_i equals $IC_{50}/2$ when using a substrate concentration equal to its K_m-value.

For non-competitive inhibition the inhibitor has a different binding site compared to the substrate and has the possibility to bind both to the free enzyme and the substrate-enzyme complex (Scheme 2.1b). Consequently a non-competitive inhibitor can either inhibit the substrate from binding or the substrate-enzyme complex from generating a product. In the presence of a non-competitive inhibitor the reaction will not reach the same rate, i.e. V_{max} will decrease but the K_m for a probe substrate will not change. At a substrate concentration at K_m the inhibition constant, K_i equals the IC_{50}-value for this type of inhibition. For mixed inhibition, both competitive and non-competitive inhibition is occurring. In this case, the K_m also increases whereas the V_{max} decreases and is not recovered with increase in substrate concentration.

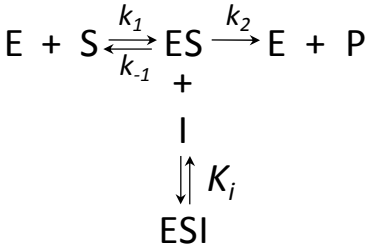
The last type of reversible inhibition is uncompetitive inhibition. An uncompetitive inhibitor binds only to the enzyme-substrate complex and inactivates it (Scheme 2.1c). Since the extent of inhibition is dependent on the amount of the substrate-enzyme complex and consequently the amount substrate present, it increases with increased *S*. This type of inhibition is recognized by a decrease in both K_m and V_{max}. This is a rarely observed type of inhibition



a) Competitive inhibition



b) Non-competitive/mixed inhibition



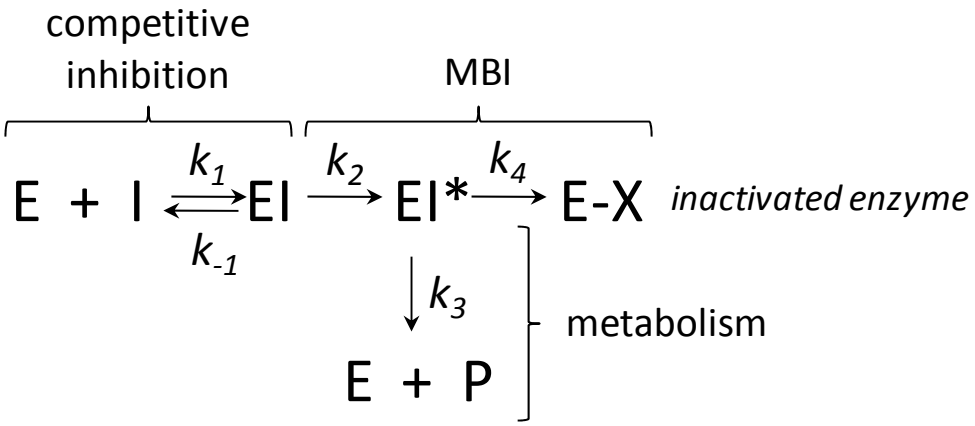
c) Uncompetitive inhibition

Scheme 2.1 Kinetic schemes describing the different types of reversible inhibition a) competitive inhibition b) non-competitive/mixed inhibition and c) uncompetitive inhibition.

2.4.2.2 Irreversible inhibition

Some compounds effect their inhibition through metabolic activation where the metabolites react or strongly bind to the enzyme, effectively inactivating it. Due to the requirement of metabolic activation, the process is time dependent, hence they are commonly being referred to as time-dependent inhibitors (TDI).[59] Irreversible inhibitors bind covalently to the enzyme, either to the heme or to the apoprotein and are referred to as mechanism based inhibitors (MBI). Quasi-irreversible inhibitors show strong coordination to the heme and form metabolite intermediate complexes (MIC). In both cases the enzyme is effectively inactivated and synthesis of new enzyme molecules is required to restore activity. Since TDI involves metabolic activation this process is NADPH– and concentration dependent and unlike reversible inhibitors it is also time-dependent. Accordingly this type of inhibition can be identified by preincubating the inhibitor with the enzyme in the presence of NADPH, for different time periods, before measuring the activity of the enzyme using a probe substrate. One way to assign the degree of inhibition is to use a partition ratio. This is defined as moles of product formed per mole of inactivated enzyme.[59] Translated to Scheme 2.2 the partition ratio is the ratio of k_3/k_4.[60] Additional parameters used in the context of TDI are k_{inact} – maximum rate of inactivation, and K_I – inhibitor concentration at half k_{inact}.

Several structural moieties have been reported to cause mechanism-based inhibition, among these are terminal alkenes, alkynes, furans, thiophenes, epoxides, secondary amines, and benzodioxoles.[61]



Scheme 2.2 Kinetic scheme describing mechanism-based inactivation (adapted from Kalgutkar et. al. 2007[60])

2.5 Metabolism platforms

For determination of metabolic properties of an NCE there are several ADME platforms available[2] and some of the tools and model systems for metabolic studies are described below.

2.5.1 In vitro systems

Initial characterization of metabolic properties like inhibition, metabolite identification etc is done in available *in vitro* systems which vary in complexity and can allow many different questions to be answered. At the lowest level of complexity there are subcellular systems like recombinantly expressed enzymes and microsomal fractions. The progress in biotechnology has made it possible to heterologously express human CYPs in e.g. bacteria, yeast or insect cells.[62, 63] The system gives the opportunity to study the role of one single isoform and is successfully used for inhibition studies, CYP phenotyping and metabolite identification. When using recombinant enzymes to assess the contribution of different CYPs to the metabolism of a drug it is however important to consider their different expression levels in the liver. For the work in this thesis recombinant CYP2D6 and CYP3A4 were used for most studies in order to focus on the properties and interactions involving these two isoforms.

For a more comprehensive view of the metabolism of a drug the microsomal fraction can be used.[64] Microsomes consist of vesicles of the endoplasmic reticulum (ER) and in the preparation process different subcellular fractions are obtained: 1) S9, containing ER and cytosol, 2) only cytosol, and finally 3) microsomes. Most membrane bound drug metabolizing enzymes, e.g. CYPs and UDP-glucuronosyl transferase (UGT), are present in the ER and can be evaluated using microsomes as a test system. However, there are also several soluble enzymes of interest, e.g. dehydrogenases and amidases, which are present in the cytosol. Thus, the choice of system has to be based on which kind of reactions and enzymes that are under study. Microsomes, which are most commonly used, contain both CYPs and the conjugation enzyme UGT. However, sequential metabolism is not possible to follow because of the need to add mutually non-compatible additives for the optimal activity of each of these families of enzymes. As an example, for the activity of UGTs, detergents have to be added which are not compatible with CYP activity. The benefit of using these systems is that the metabolism of a compound is solely dependent on metabolism and not influenced by active/passive transport. From a practical point of view, these microsomal systems are also easy to store and handle.

If there is interest to take active/passive transport into account as well as sequential metabolism of phase I and conjugation reactions, whole cell systems like primary hepatocytes are more appropriate.[65] When using primary hepatocytes the compound will be subjected to both soluble and membrane bound enzymes. The corresponding co-factors will also be present in the system. Most reliable is to use freshly isolated hepatocytes for metabolism studies, but these can be difficult to obtain in which case it is more common to use cryopreserved cell suspensions. The benefit of using cryopreserved cells is that it is possible to pool different batches in order to decrease interindividual variability in the expression of enzymes. The risk, on the other hand, is that some functions might be lost in the cryopreservation process.

All the above mentioned systems can be used for incubations with new compounds in order to address metabolic stability, DDI (inhibition) or for the generation and

identification of metabolites. There is also a possibility to study interspecies differences by using preparations from different species, e.g. human, rat, dog, and monkey.

2.5.2 Metabolite identification

An important assay for the understanding of the metabolic fate of a drug is metabolite identification. Metabolites can be generated *in vivo* or in one of the above mentioned *in vitro* systems. The metabolites are extracted from the biological matrix, plasma, urine or incubation mixture, and analyzed for structural identity. In drug discovery this is almost exclusively done with high performance liquid chromatography coupled to mass spectrometry (HPLC-MS) and the identification of unknown structures is often accomplished with the use of tandem mass spectrometry (MSMS).[66] The first task is to separate drug related material from endogenous or matrix background. A rough filter is to subtract the data achieved from a blank sample, however this is not always enough and it is also important to use an appropriate blank. With the development of highly sensitive mass spectrometers, accurate mass data can now be available and used in metabolite identification.[67, 68] The possibility to have the molecular mass of an unknown peak reported with four decimal points enables the assignment of the elemental composition, which facilitates the discrimination from matrix components as well as allows separation of isobaric ions. There are software available for handling MS spectral data which also assign drug related material and suggest a possible biotransformation route (hydroxylation, dealkylation etc), e.g. MetWorks[69] from Thermo Scientific and Metabolynx[70, 71] from Waters. To be able to assign the site of metabolism (SOM), e.g. site of hydroxylation, interpretation of collision induced dissociation (CID) fragments are done. A comparison of the fragmentation pattern of the parent compound with that of the metabolite can, in most cases, give an idea of the SOM. A more accurate way for structural elucidation is to use NMR spectroscopy. However, this technique is not widely used for metabolite identification. The main reason is that the different components in the sample cannot be separated in an NMR spectrum; consequently there are high demands on sample preparation. Secondly the sensitivity of NMR spectroscopy is far lower than that of MS and more material is needed.[66]

3 Computational modeling of Cytochrome P450

CYPs are versatile enzymes and the activity and selectivity of the enzyme are determined by a combination of studies of reactivity and of the complementarity in the interaction between ligand and enzyme. The complex nature of both the enzyme structure (flexibility and water interactions) and the catalytic cycle (several types of oxidation mechanisms etc) makes it difficult to generate a complete *in silico* prediction tool that covers rate of metabolism, extent/type of inhibition, enzyme-ligand affinity and regio-selectivity of catalytic reactions etc. Several approaches to predict CYP reactivity has been summarized in a review by de Graaf et al.[72] In general these methods can be divided into structure-based or ligand-based approaches depending on whether the methodology involves the structure of the enzyme or not.

3.1 GRID molecular interaction fields (GRID-MIFs)

One approach to describe both protein and/or ligand potential interactions is to use molecular interaction fields (MIFs). Calculating grid molecular interaction fields (GRID-MIFs) is a way to obtain non-covalent binding information of a molecule and can be used to determine potential energetically favorable binding sites.[73, 74] The MIFs for a target (protein or ligand) are calculated using probes representing different chemical interaction properties (hydrogen bond donor/acceptor, electrostatic potential, hydrophobic potential etc) positioned in a 3D grid around the target (Figure 3.1). Examples of probes to use are the amino function (NH_3^+), carbonyl oxygen (O), carboxy oxygen (O^-), hydroxyl (OH), methyl (CH_3) and water (H_2O). A GRID energy value is then calculated with the GRID force field for each probe at each particular position of the grid around the region of interest. With the GRID force field the pair wise energy between a grid probe and a target atom is calculated (Eq 3.1):

$$E_{PAIR} = E_{LJ} + E_Q + E_{HB} \qquad \textit{Eq. 3.1}$$

where E_{LJ} is the Lennard-Jones potential, describing the energy minima for preferred arrangements of two atoms, E_Q is an electrostatic term and E_{HB} the hydrogen bond potential. All force field parameters used are obtained from analysis of crystal structures.

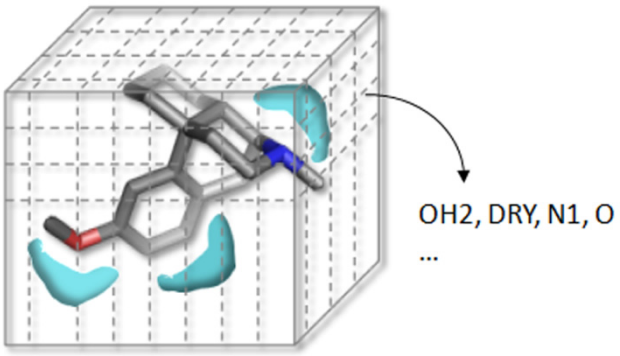


Figure 3.1 Concept of calculation of GRID-MIFs. The interaction with the probes (OH2, DRY, N1, O, …) is calculated for each grid point in the surrounding 3D-grid resulting in an interaction field for each property.

GRID-MIFs can be used to characterize enzyme active sites and ligands and is applicable in both structure-based and ligand-based modeling. MIFs are used to derive descriptors for quantitative structure activity relationships (QSAR), virtual screening and also to define site of metabolism (SOM) in the software MetaSite.[75] In Paper III GRID-MIFs are used in combination with a pharmacophore approach called FLAP.[76] They can also be successfully used for determination of favorable interaction patterns between ligand and enzyme, which can be compared between different 3D-structures. The software used for these calculations is GRID from Molecular Discovery Ltd.[51]

3.2 Structure-based modeling

Structure-based modeling uses the 3D-structure of the protein and includes homology modeling and molecular docking. Homology modeling is a method to predict the 3D-structure of a protein based on its amino acid sequence and the structure of a homologous protein.[77] Homologous proteins have a common genetic origin, and in the case of CYPs the crystal structures of bacterial CYPs have been used for the modeling of the human counterparts.[78, 79] A general process of this comparative modeling is described by Kirton et al.[80] and contains the following steps: 1) identification of a protein template with high sequence similarity 2) sequence alignment of the target protein and the template, 3) building of a 3D-model, and 4) validation and refinement of the constructed model (stereochemical quality, side chain environment and root mean square deviation (RMSD) of the main chain).

The availability of homology models and crystal structures for the human CYPs has made it possible to use molecular docking to study and predict the interaction between ligand and enzyme. There are several docking algorithms available (a comprehensive list of molecular modeling software is available in a review by de Graaf et al.[72]) and for the work in this thesis the docking tool that has been selected is the program GLUE within the GRID software,[51] which is based on the GRID force field. The general principle of

molecular docking is a search for optimal orientations of the ligand in the enzyme active site combined with a scoring function. Most software have the ability to account for flexibility in both the protein and the ligand. Consequently, flexible molecules will result in a wealth of possible docking poses. To handle this, the scoring function has the aim to generate a measure of the complementarity of the interactions between ligand and protein, i.e. it makes a ranking of the potential orientations of the molecule inside the active site cavity. However, an energy-based scoring function will have difficulties to describe certain energetic terms like the entropic contribution or the water interaction, therefore accurate scoring functions are difficult if not impossible to obtain.[81]

From molecular docking energetically favored orientations of the ligand can be visualized and important interactions with the surrounding amino acid residues can be identified. Identifying the part of the substrate that is oriented towards the heme can also give information about the site of metabolism. However, it must be stressed that the CYPs are flexible proteins with relatively large active site cavities which make docking difficult and not infrequently give inconclusive results due to the large number of obtained docking poses. In Paper I a new approach to handle docking results is described in which the energy of interaction between derived docking poses and amino acid residues in the active site is calculated. This can then be used to identify important interaction partners within the active site.

3.3 Ligand-based modeling

Classical ligand-based methods use the information obtained from substrates and/or inhibitors and do not include information from the 3D-structure of the protein. Included here are quantum mechanical (QM) calculations, pharmacophore models and 3D quantitative structure activity relationship (3D-QSAR) models. A pharmacophore is a 3D arrangement of functional groups and describes common characteristics, like hydrogen bond donors/acceptors, hydrophobic areas, positive/negative charges etc, for substrates or inhibitors. A pharmacophore model can be built by superimposing a large set of known ligands. Since all substrates share the characteristic that they should be oriented in the protein cavity with their SOM close to the heme, the atom in this position is a natural anchor point that can be used for successful alignment of the compounds. This is also the underlying hypothesis used in different computational approaches. For inhibitors, on the other hand, it is more difficult since the orientation in the active site and important interaction points are seldom known, although this depends on the type of inhibition. In order to obtain a pharmacophore model, successful alignments of the ligands are needed and therefore flexible compounds could be more difficult to process. It is also important to keep in mind, when using and analyzing these models, that the CYPs allow for some flexibility and multiple pharmacophores can co-exist.[82] A way to come around the uncertainties and to refine/guide the pharmacophore model building is to use structural information from crystal structures and homology models. Pharmacophore models can be used to discriminate substrates/inhibitors from non-

substrates/inhibitors as well as rationalization of SAR and structure metabolism relationships (SMR). In Paper III a pharmacophore approach is used for SOM predictions in CYP2D6.

Another ligand-based approach is 3D-QSAR in which spatial descriptors of the aligned ligands are characterized in 3D and analyzed with multivariate data analysis.[72] Further on these descriptors can be used to predict properties like IC_{50} and K_m by the use of linear regression techniques like partial least squares (PLS).[83]

3.4 Statistical methods

Multivariate data, such as those obtained from MIFs, containing a matrix of variables and objects cannot be visualized and analyzed by ordinary linear regression techniques because of huge intercorrelated X data. In these cases principal component analysis (PCA) is a powerful method that reduces the dimensionality of the data while describing the variation.[84] In this approach linear combinations of the original variables are extracted with the aim to describe as much of the variation in the data as possible. Starting from a multidimensional data space the first linear combination, principal component (PC), is extracted in order to describe as much of the variance as possible. Additional PCs are then extracted orthogonally to the previous one/ones aiming at describing the variance not formerly explained. At a high number of PCs the method will not describe the general trend in the data but particularly the data for each object. This can be referred to as noise. Consequently an optimal number of PCs will describe the data well while including as little noise as possible. The results from a PCA can be visualized in so called score- and loading plots. These are combinations of two PCs with the objects projected in the score plot and the variables in the loading plot. Since this analysis is based on the variance of the data, objects that are projected close to each other in the score plot have equivalent variance and are therefore similar. On the other hand objects far from each other show different variance and could be discriminated. The scores and the loadings are correlated and variables having positive influence on an object are positioned in the corresponding place in the loading plot. The use of scores and loadings is a convenient way to visualize and identify trends and relationships in the data.

In Paper I consensus principal component analysis (CPCA) is used for identification of selectivity in interaction pattern between different proteins. CPCA is a method based on PCA that uses MIFs calculated in the active site of the proteins.[85]

4 Aims and Objectives

Aim

The overall aim of this thesis was to use a combination of *in vitro* metabolism platforms and *in silico* tools to study affinity and selectivity in CYP2D6 and CYP3A4 mediated metabolism, and also to introduce the power of this combined approach within drug design.

Objectives

The specific objectives of the study were:

1. To evaluate and compare crystal structures and homology models for CYP3A4 and CYP2D6

2. To develop a computational tool to determine and identify discriminative enzyme-ligand interactions from dockings and assessment of affinity.

3. To compare catalytic properties of CYP3A4 and CYP2D6 with *in vitro* metabolic studies and *in silico* dockings

4. To explore the CYP2D6 pharmacophore and the affinity for the enzyme using *in silico* and *in vitro* methods, respectively.

5. To evaluate the possibilities for enhanced throughput in site of metabolism (SOM) determinations by linking analytical methods and software for semi-automated assignment of SOM.

5 Methodological considerations

In order to study the affinity and selectivity for CYP2D6 and CYP3A4 different *in vitro* assays and analytical approaches were used. These are summarized in Figure 5.1 and further description of the methods can be found in Papers I-IV.

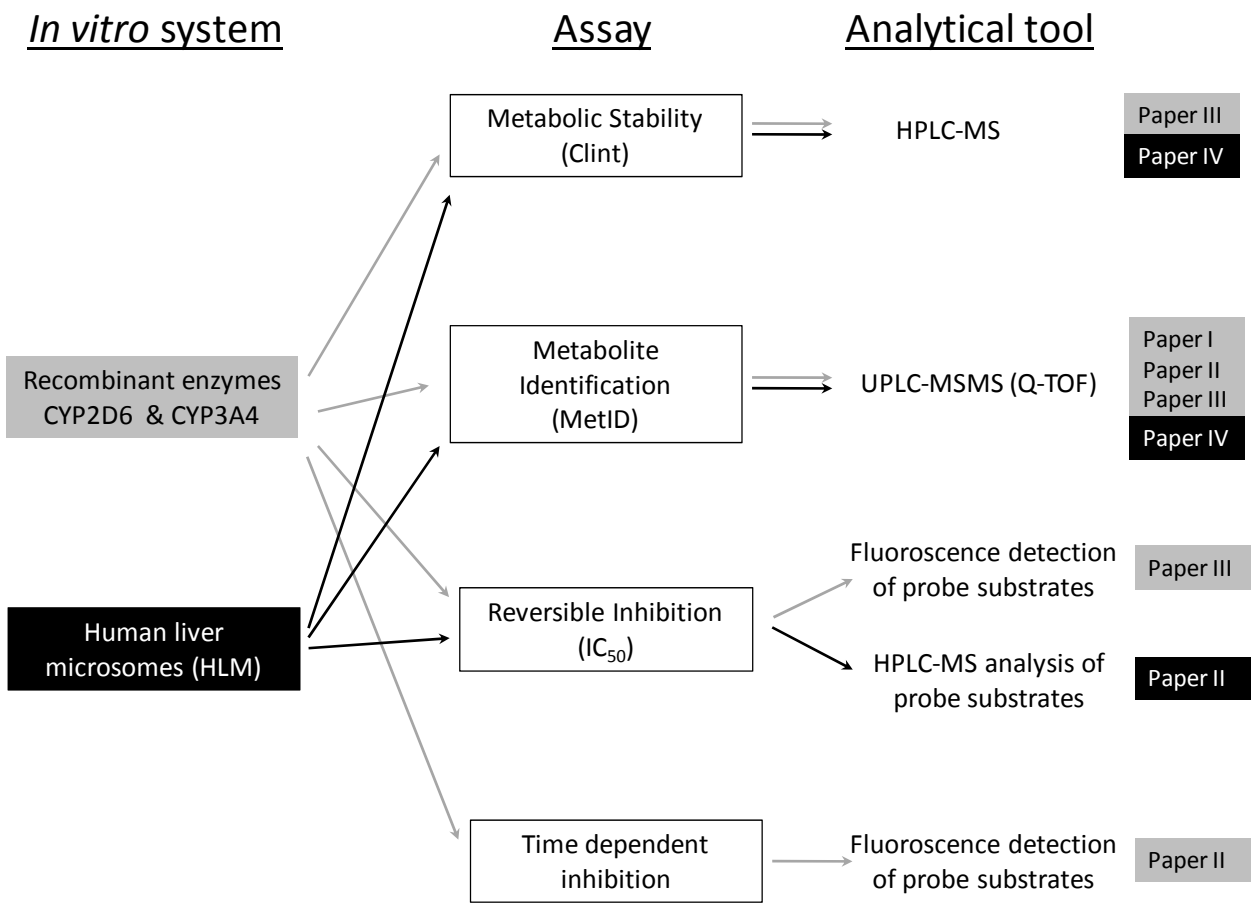


Figure 5.1 Experimental methods used in the different papers. Studies in recombinant enzymes are marked with gray lines and boxes and studies in human liver microsomes are in black.

Besides the experimental methods used, different computational approaches were employed to analyze and visualize the behavior of ligands in the two iso-enzymes. The

computational approaches are reviewed in Table 5.1 and the different techniques are introduced in Section 3.

Table 5.1 Computational approaches used in Paper I-IV

	Technique	**Comment**	**Software**	
Structure based methods	Molecular dockings	The GRID force field and MIFs	Glue	Paper I Paper II
	Energy calculation	The GRID force field and MIFs	GRID	Paper I
	SOM predictions	Enzyme-Ligand complementarity + reactivity	MetaSite	Paper I Paper III
		Spectral data + MetaSite	Mass-MetaSite	Paper IV
Ligand based methods	4-point pharmacophore	The GRID force field and MIFs	FLAP	Paper III
Statistical methods	PCA		GOLPE	Paper I
	CPCA	MIFs	GOLPE	

6 Results and discussion

In this thesis affinity and selectivity factors in CYP mediated metabolism, and how this can impact drug design were investigated. Figure 6.1 summarizes the different parts of this work. The use of structural information of the enzymes together with information about affinity and selectivity as well as SOM is believed to be valuable input in the course of the design of new drugs with desirable DMPK properties.

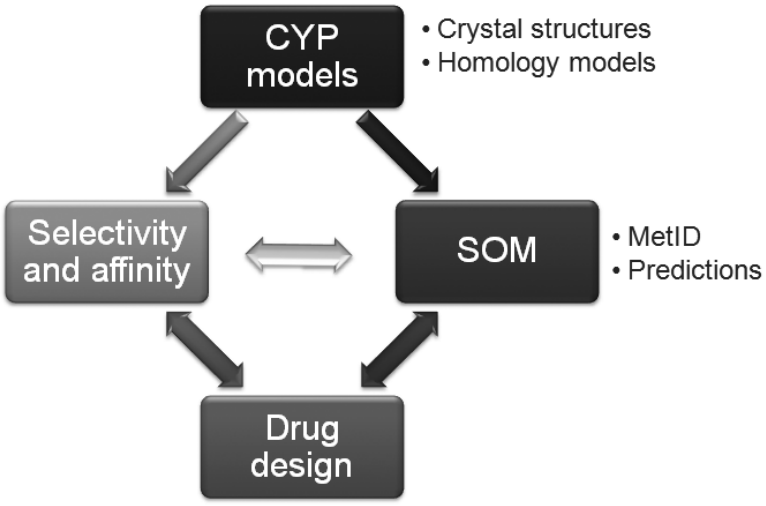


Figure 6.1 Different factors and tools of interest in drug design that are considered in the present work.

In our studies CYP2D6 and CYP3A4 were chosen as model systems for the exploration of affinity and selectivity in CYP mediated metabolism. These enzymes are of major interest since they play an important role in the metabolism of xenobiotics. Together they are responsible for the metabolism and disposition of approximately 60% of the most prescribed drugs on the market.[3] They also differ from each other with the respect to size of active site, ligands and kinetic space. Both enzymes are well studied and in particular CYP2D6, which has a defined active site and pharmacophore. This previous knowledge has been beneficial for the studies performed in this thesis. The structural information of the enzymes can be obtained from studying crystal structures and homology models and since several 3D-structures are available for the two enzymes, the

initial project involved a comparison and evaluation of the protein structures of the two enzymes.

6.1 Evaluation of 3D-structures for CYP2D6 and CYP3A4

The fundamental tools for *in silico* studies and for obtaining structural information of CYPs are crystal structures and homology models. With the current availability of different 3D-structures for each enzyme it was of interest to compare them and identify amino acid residues that differed in orientation. Though this could be obtained from a visual comparison; it would require successful alignments of the structures. The purpose of this study was to compare available crystal structures and homology models for CYP3A4 and CYP2D6 regarding their interaction patterns with ligands and how they perform in SOM predictions. The 3D-structures that were used were a homology model and a ligand-free crystal structure (2f9q.pdb) for CYP2D6 and for CYP3A4 one homology model, two ligand-free crystal structures (1tqn.pdb and 1woe.pdb) and one structure co-crystallized with erythromycin (2j0d.pdb). These are summarized in Table 6.1.

Table 6.1 Homology models and crystal structures of CYP3A4 and CYP2D6 used in Paper I.

		Comment	Resolution (Å)	R-factor
CYP3A4 structures	CYP3A4 homology model[78]			
	1tqn.pdb[28]	Ligand-free	2.05	0.24
	1woe[27]	Ligand-free	2.8	0.24
	2j0d.pdb [29]	Co-crystallized with erythromycin	2.75	0.24
CYP2D6 structures	CYP2D6 homology model [78]			
	2f9q.pdb[30]	Ligand-free	3.0	0.23

6.1.1 Calculation of enzyme-ligand interaction pattern from docking poses (Paper I)

To study the interaction pattern in the different structures, a set of opioid analgesics was used. This family of compounds was reported to be metabolized by both CYP3A4 and CYP2D6,[8, 86-89] they are also relatively rigid in their structure, which simplified the

computational work. Molecular docking of the compounds within the active site of the different structures was performed with the docking module Glue in the program GRID.[51] The resulting docking poses were analyzed with a new computational approach which involved calculation of the accumulated interaction energy between the obtained docking poses and the surrounding amino acid residues in the active site cavity. Since the study aimed to identify the interactions of a substrate within the enzyme, only docking poses within 6 Å from the heme were considered for further analysis. The active site amino acid residues were then defined as those interacting with an ensemble of all filtered docking poses. From the resulting energy values amino acid residues with a high energy of interaction over the set of docking poses were identified. In order to visualize the data and discriminate between amino acid residues among the structures PCA was used. A flowchart of the method is depicted in Figure 6.2.

6.1.1.1 CYP2D6 structures

The resulting scores and loadings, based on the interaction energies between docking poses and active site amino acid residues, for the homology model and crystal structure of CYP2D6 are presented in Figure 6.3. The score plot displays how the observations, in this case the different docking poses, are projected over the first and second principal component (PC). Since the PC describes the variance in the data, docking poses positioned close to each other have equivalent variance and are therefore similar. Hence, the fact that the docking poses from the homology model and the crystal structure are clustered indicates that the differences, in the interaction energies from docking poses, between the two structures are more pronounced than the differences in the interaction energies from poses obtained within one structure. Consequently, the two CYP2D6 structures could be discriminated based on the interaction pattern with the ligands. From the loading plot the discriminative amino acid residues could be identified. Since this data refer to negative energy values a variable, i.e. amino acid residue, that has strong influence on an object, i.e. docking pose, is positioned at the same coordinates but with opposite sign in the loading plot. From an analysis of the score- and the loading plot, hydrophobic interactions seemed to be of importance in the crystal structure where Phe120, Leu213, Val370, and Leu484 could be distinguished as being responsible for the interactions in this structure. From the docking results in the active site of the homology model Ile106, Leu121, Glu216, Glu222, Ser304, Asp301 and Met374 were more important. Studying an overlay of the two active sites (Figure 6.3c) it was apparent that these were also the amino acids that differed in orientation between the structures. From the PCA loading plot Asp301 was identified to be more important for binding in the homology model compared to in the crystal structure. However, from the visual comparison this residue did not differ substantially in the two structures. The reason for the contradicting results could be that the differences in the other amino acid residues were altering the accessibility of Asp301. As mentioned above the discriminative amino acid residues in the crystal structure were mainly of hydrophobic character, which could be an indication of a hydrophobic collapse of the active site due to the absence of a ligand.

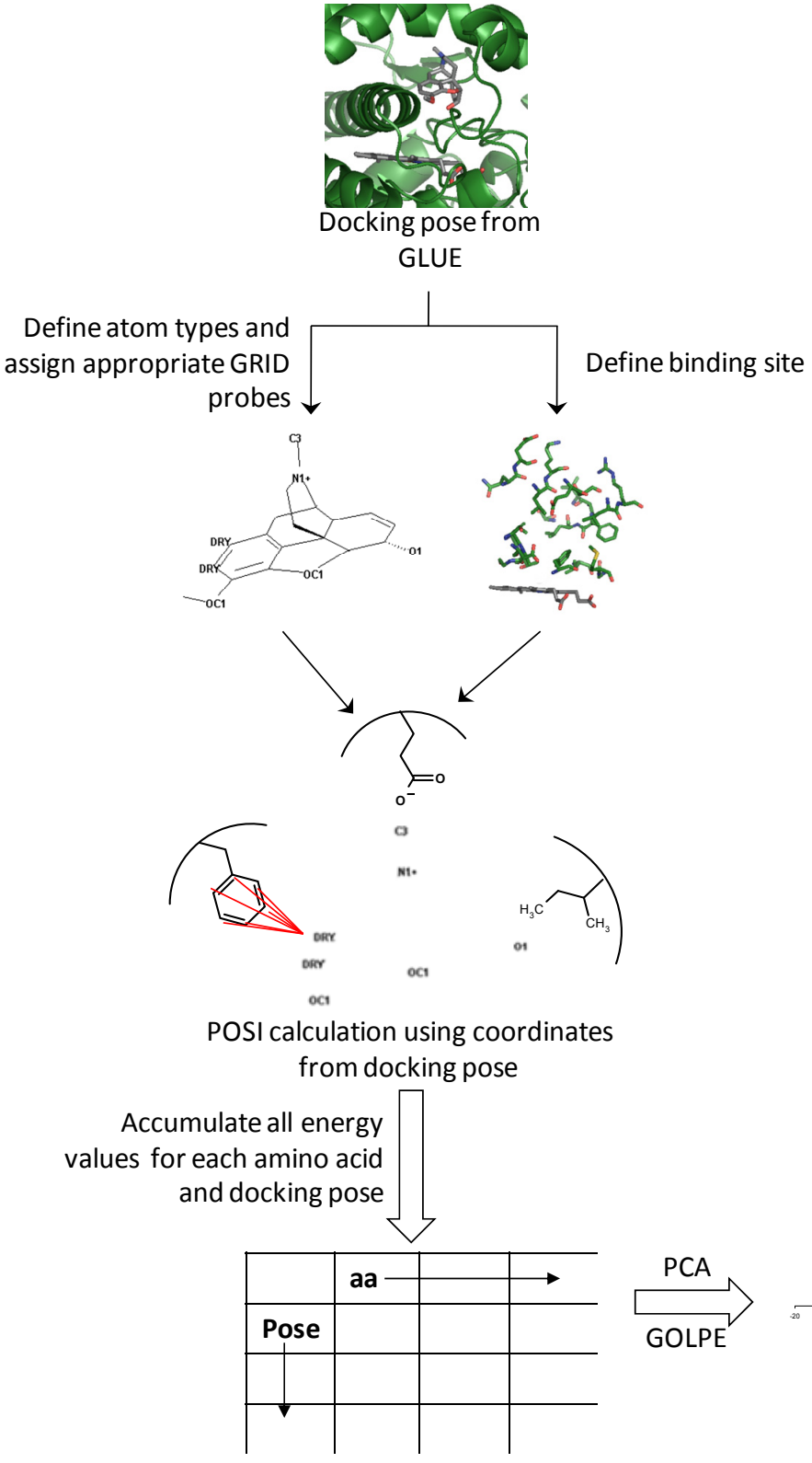


Figure 6.2 Energy calculation method including molecular dockings of compounds in active site cavity followed by assignment of appropriate GRID probes and calculation of accumulated interaction energies between docking pose and active site amino acid residues. The resulting data was analyzed with PCA.

This might also have had consequences for the orientation of Glu216. It should also be stressed that a crystal structure and/or a homology model is a static representation of the enzyme and as mentioned earlier, CYPs are referred to as flexible and promiscuous enzymes. Consequently, it is not surprising if a different conformation is obtained upon binding of a substrate. It is impossible to judge the structure that best describes the enzyme. It could even be that one conformation is better suited for one set of compounds and a different one for another set. From a comparison of the ability to predict SOM using MetaSite, the result from the two structures did not differ. This could be explained by the fact that MetaSite employs flexible GRID in the calculations and the allowed movement of the amino acid residue will cancel out the differences observed from the static representations.

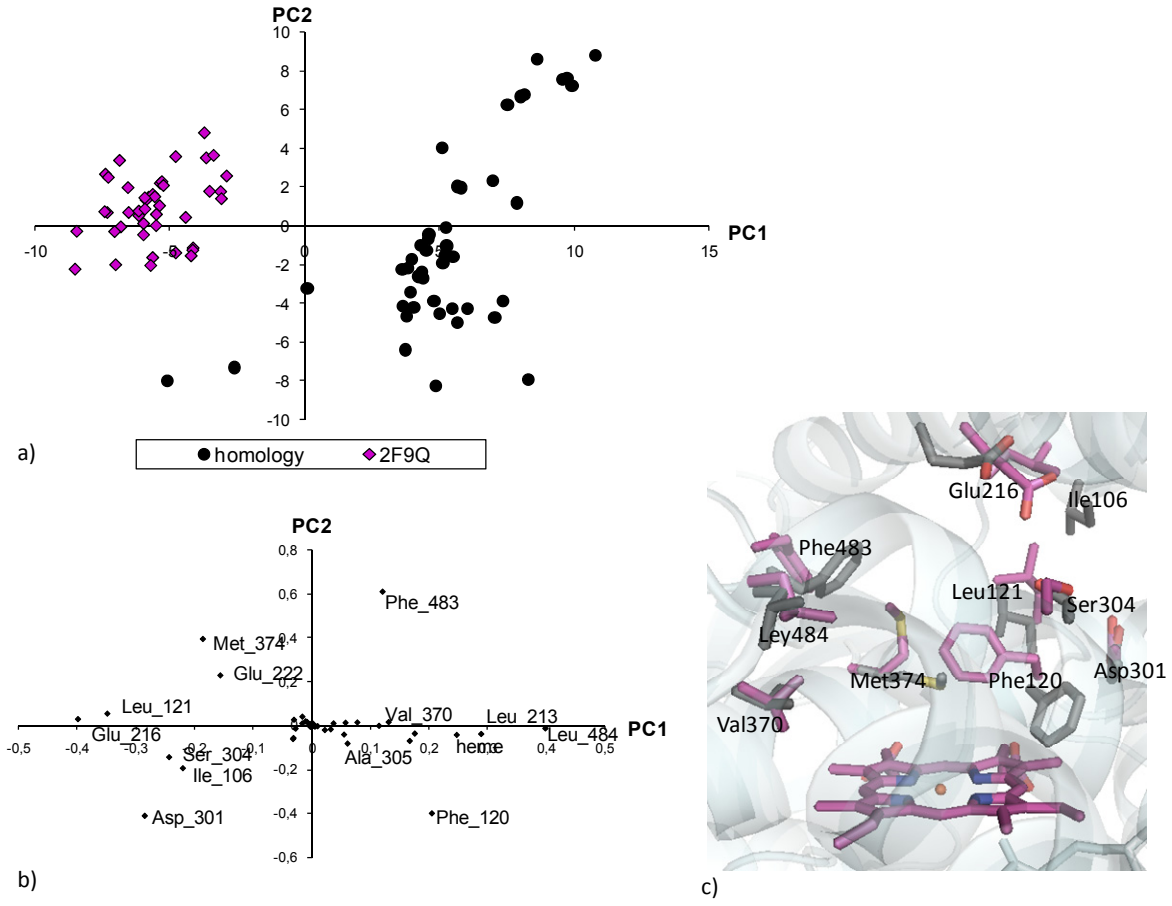


Figure 6.3 Results from comparison of interaction pattern in the two CYP2D6 structures, 2f9q.pdb (magenta) and the CYP2D6 homology model (grey). a) score plot from PCA analysis, projecting the obtained docking poses b) loading plot from PCA analysis, projecting the active site amino acid residues and c) superimposition of the two structures with selected active site amino acid residues highlighted.

6.1.1.2 CYP3A4 structures

Compared to the single crystal structure of CYP2D6 that is available, there were several different structures available for CYP3A4, with or without ligand present in the active site. The four structures studied here were clustered in the PCA (Figure 6.4a) indicating

that they could be differentiated based on their interaction pattern with the docked substrates. The homology model and the ligand-free crystal structure, 1tqn.pdb, were best separated and looking at the overlay of the active sites there were some major differences in the orientation of amino acid residues (Figure 6.4c). These involve Arg105, Arg212 and Phe215 pointing into the active site in the ligand-free crystal structure, giving rise to strong interactions with the docked ligand. Regarding the crystal structure co-crystallized with erythromycin (2j0d.pdb), Ser119 and Phe304 were identified as being important for the ligand-enzyme interactions. These residues could also be picked out from the PCA of the interaction energies. The identification of hydrophobic amino acid residues in the ligand-free crystal structure can be compared with the scenario in CYP2D6 and also here be a consequence of a hydrophobic collapse. In the analysis of the CYP3A4 structures an additional data set of inhibitors with diverse structures were used. The PCA result from the calculated interaction energies with this set of compounds was analogous to the results shown with the opioid analgesics. This was encouraging and put strength to the method since it did not seem to be biased by the choice of compounds for the dockings.

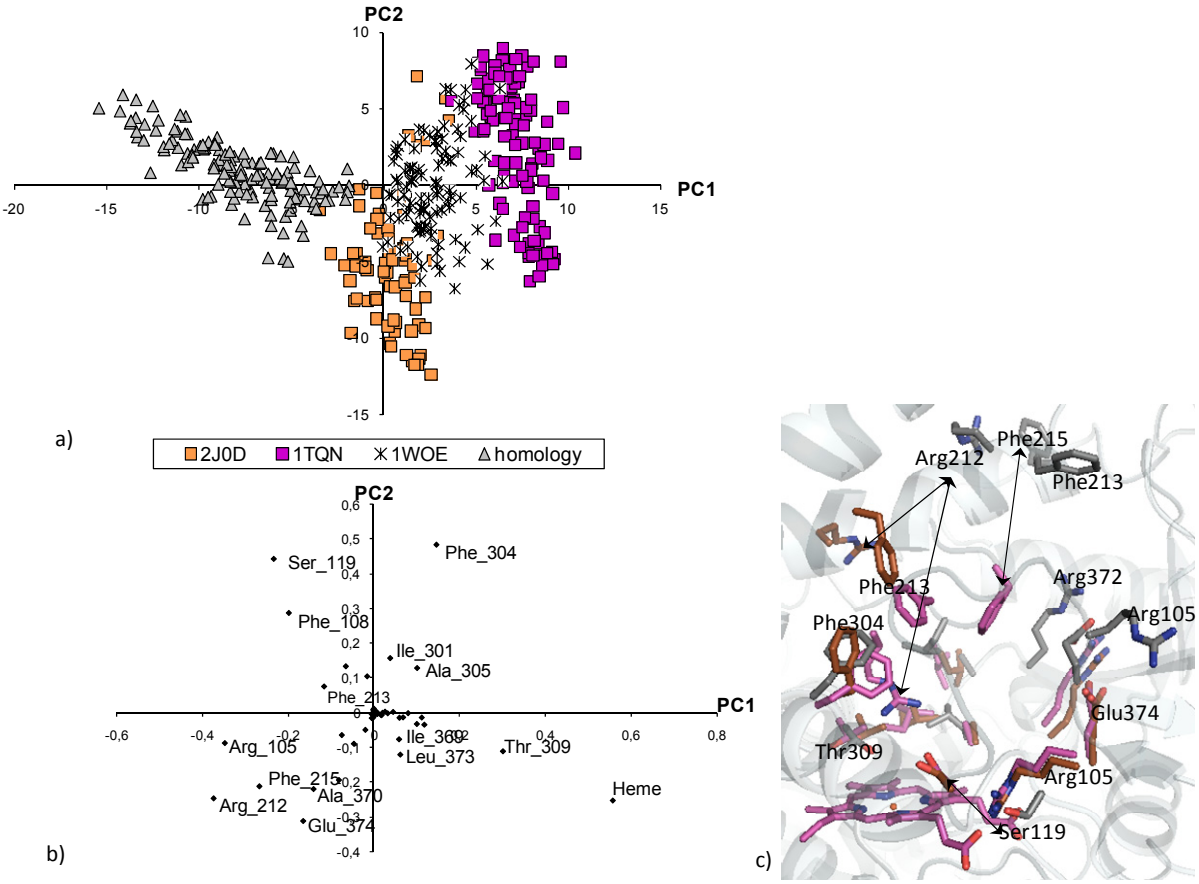


Figure 6.4 Results from comparison of interaction pattern in the CYP3A4 structures a) score plot from PCA analysis, projecting the obtained docking poses, b) loading plot from PCA analysis, projecting the active site amino acid residues and c) superimposition of, 1tqn.pdb (magenta), CYP3A4 homology model (grey) and 2j0d.pdb (orange) with selected active site amino acid residues highlighted. The arrows refer to amino acid residues with pronounced altered orientation in the different structures.

As for CYP2D6 the prediction of SOM using MetaSite was also compared for the different CYP3A4 structures. In this case the homology model performed evidently worse than the crystal structures (50% success rate compared to ~80%). This was not surprising since the homology model had pronounced differences in the orientation of several active site amino acid residues compared to the crystal structures, which could not be accounted for in the allowed flexibility in the MetaSite calculations.

To conclude this study it is obvious that there are differences between the available structures but still it is impossible to judge which one it is that gives the most correct representation of the enzyme. When choosing which structure to use it is however of interest to be aware of the differences and then the energy calculation approach could be a valuable tool. For the CYP3A4 structures this approach was compared with a previously used method for structural comparison, CPCA,[90, 85] which gave similar results regarding discriminative amino acid residues. Nevertheless, in this study it was found that the energy calculation approach is easier to interpret, in addition the CPCA needs successful alignments of the structures. Comparisons of 3D-structures based on their interaction pattern with ligands seems to be a successful approach and the methodology with which the energy of interactions is calculated from docking poses could also be a way to assess the affinity for an enzyme. Accordingly the method could also be valuable for other purposes like identification of wanted/unwanted interactions with ligands, which can be utilized when trying to improve the metabolic properties of a drug. Overall this is an approach that can be employed for the rationalization of affinity as well as selectivity for a CYP.

6.2 Studies of affinity and selectivity for CYP2D6 and CYP3A4

In combination with experimental data the available structures for CYP2D6 and CYP3A4 can be of great help in the understanding of affinity and selectivity for the two isoforms. For example, molecular dockings within the active site cavities can give information of favorable orientations of substrates in one enzyme compared to another. Dockings within the active site cavity of the enzyme together with experimental data on SOM could also be used to visualize possible binding modes and to identify key interactions. For the prediction of affinity, the studies often refer to inhibitors and inhibition constants[81] and in addition the selectivity for an enzyme could be guided by information of SOM, e.g. identification and comparison of selective orientations between enzymes.

From the knowledge of CYP2D6 and CYP3A4 these enzymes accommodate different substrates and probably prefer different binding modes. From the dockings in CYP3A4 a wealth of poses are often obtained from which it can be difficult to draw conclusions while the CYP2D6 results are generally more restricted (see Paper I). This leads us to ask whether this is a disadvantage of the docking protocol or a correct description of what is happening within the active site of the enzyme? To answer this question a study was performed aimed at determining if the differences between CYP2D6 and CYP3A4,

documented from analysis of crystal structures etc, could also be reflected by experimental data. In addition, the CYP2D6 pharmacophore was further explored in order to assess the affinity and selectivity for this particular enzyme.

6.2.1 Comparison of catalytic properties of CYP3A4 and CYP2D6 (Paper II)

From the set of opioid analgesics used in Paper I, two compounds were selected for detailed studies in CYP2D6 and CYP3A4 – levorphanol (**1**) and levallorphan (**2**) (Figure 6.5). With the purpose to explore the differences between CYP2D6 and CYP3A4 regarding binding modes and catalytic properties, metabolite identification in the corresponding recombinant enzymes and the inhibition potential of the compounds were evaluated. The time dependent inhibition (TDI) properties were also explored since the allylic side chain of levallorphan is a structural feature that has been observed in the context of TDI[61] and consequently could lead to reactive metabolites. These results were expected to provide further information about the orientations and interactions of the substrates with CYP3A4 and CYP2D6.

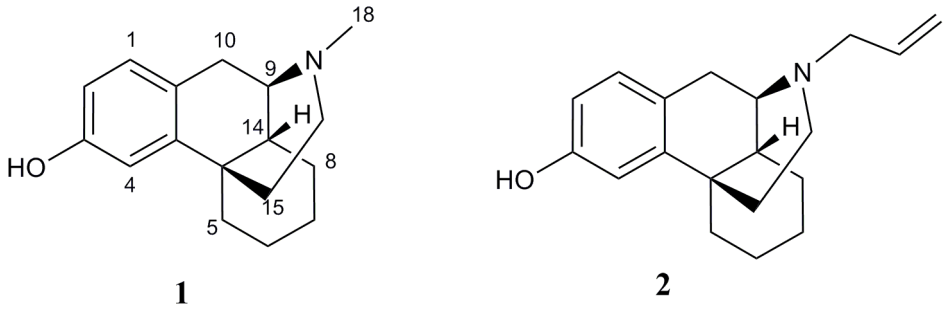
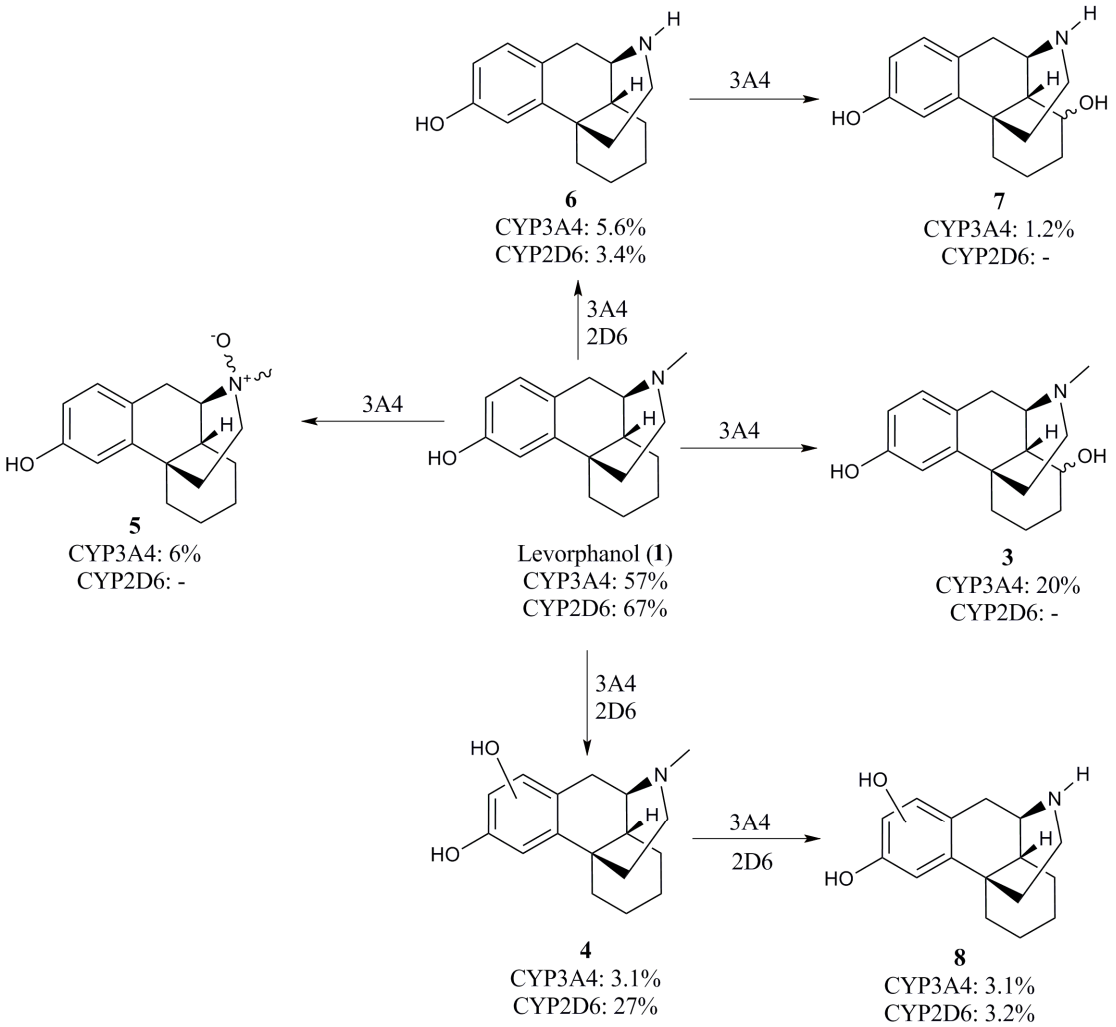


*Figure 6.5 The structure of levorphanol (**1**) and levallorphan (**2**)*

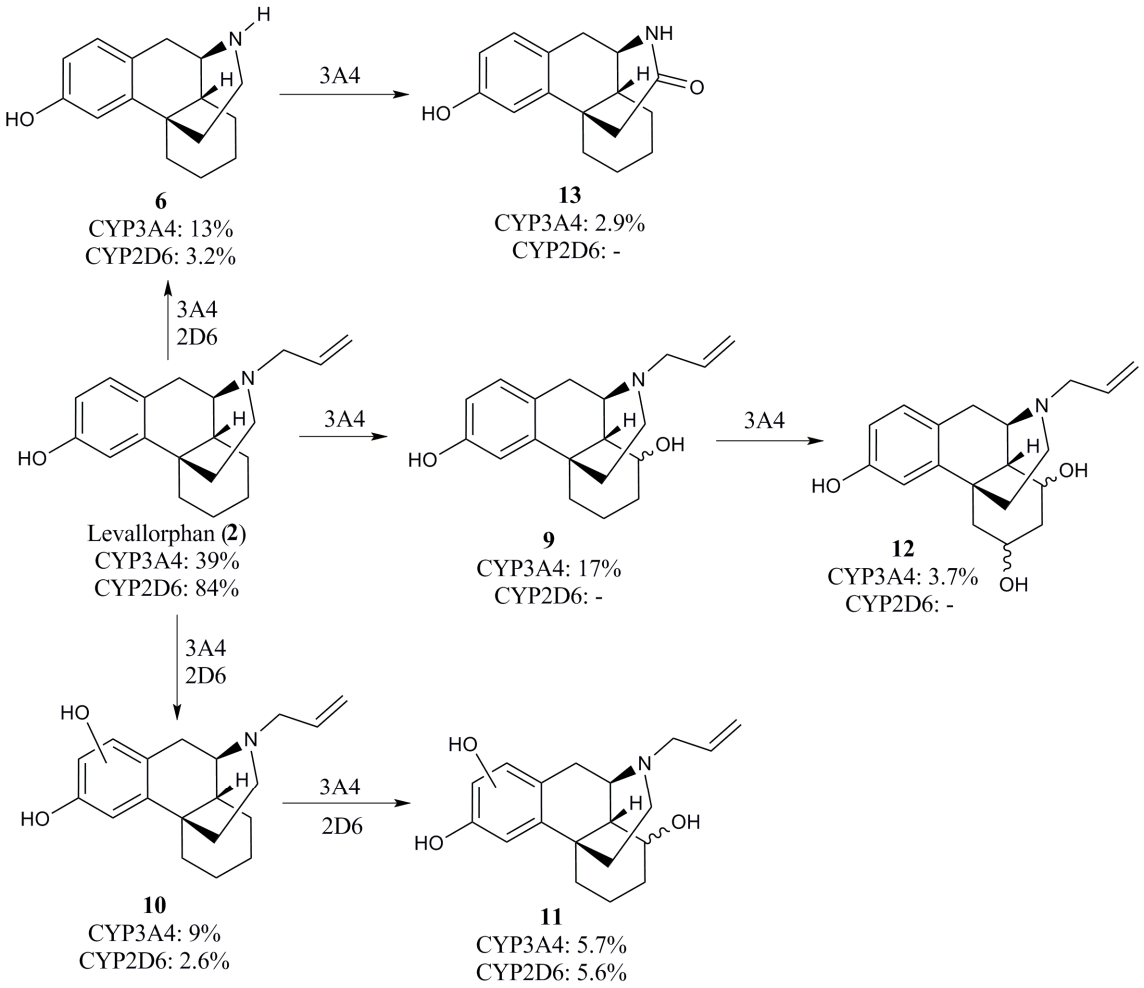
The most straightforward way to analyze the orientation of a substrate in the active site cavity is to determine the SOM. This was done in recombinant CYP2D6 and CYP3A4 for the two compounds and the results are presented in Schemes 6.1 and 6.2. The assignment of metabolites was done by rationalization of collision induced dissociation (CID) spectra. Spectra and the corresponding schemes for the tentative assignment of diagnostic fragment ions are available in paper II and its Supplementary Information.

Studying the percentage remaining after incubation and the half-life of the compounds it was obvious that CYP3A4 was a better catalyst for both levorphanol and levallorphan. In CYP2D6 the half-life was >40 min for both compounds with 67% levorphanol and 84% levallorphan remaining after the 60 min of incubation. Levorphanol did also have a low turnover in CYP3A4 ($t_{1/2}$ >40 min) but still there was 43% converted to metabolites. The half-life of levallorphan in CYP3A4 was 8.5 min, indicating that it is a good substrate for this enzyme. Moreover, since the aim of this study was to explore the orientation of the substrates and their interactions in the active site cavity of the two iso-enzymes the SOM

was of special interest. Reviewing the proposed metabolic schemes it was apparent that CYP3A4 was involved in all transformations including hydroxylation in the aromatic moiety and the cyclohexyl ring as well as *N*-dealkylation and *N*-oxidation. Based on the %MS area the cyclohexyl ring and the amino function seemed to be the most favorable SOMs in CYP3A4. The orientations in CYP2D6 appeared to be more restricted with the aromatic hydroxylation as the dominating pathway. In summary these results indicated that levallorphan and levorphanol could have many different orientations, presenting different parts of the molecules to the heme in CYP3A4, while the orientation of the substrates was more restricted in CYP2D6. *N*-Dealkylation/deallylation was observed in CYP2D6 despite that this orientation would be expected to be unfavored based on the CYP2D6 pharmacophore. However, these metabolites were only minor with respect to the %MS area.



*Scheme 6.1 Suggested metabolic pathways for levorphanol (**1**) in CYP2D6 and CYP3A4 with the %MS area after 60 min incubation indicated.*



*Scheme 6.2 Suggested metabolic pathways for levallorphan (**2**) in CYP2D6 and CYP3A4 with the %MS area after 60 min incubation indicated.*

In addition to the metabolite identification, CYP inhibition studies were performed to further explore the catalytic properties and to determine the affinity for CYP2D6 and CYP3A4. The potential for reversible inhibition was tested in HLM using probes specific for the corresponding isoforms. The probe reactions used were formation of 1'-OH-midazolam in CYP3A4 and the formation of OH-bufuralol in CYP2D6. IC_{50}-curves were obtained by incubating the compounds at eight different concentrations. In order to assess the TDI properties of levallorphan and levorphanol they were incubated with recombinant enzymes at different time points before measuring the enzyme activity with fluorescent probe substrates (4-aminomethyl-7-methoxycoumarine (AMMC) for CYP2D6 and 7-benzyloxy-4-trifluoromethylcoumarine (BFC) for CYP3A4). The preincubations were performed with and without NADPH. As stated earlier a time dependent inhibitor exhibit time-, NADPH- and concentration dependency.[60] Accordingly, the measure of the remaining activity of the enzyme, with respect to the metabolism of a probe substrate, was plotted against time in order to identify any time-dependency (Figure 6.6). In addition a normalized ratio (NR) was calculated for which resulting enzyme activity (*A*) in incubations in the presence of NADPH was normalized to those without NADPH (Eq. 6.1).

A low NR indicated the presence of a NADPH dependency and consequently implies that the compound acted as a time dependent inhibitor. In Table 6.3 are the obtained IC_{50}-values and the NR summarized.

$$NR = \frac{(A_{inh}/A_{ctrl})}{(A_{inh}/A_{ctrl})_{wo\ NADPH}} \qquad \text{Eq. 6.1}$$

Both levorphanol and levallorphan reversibly inhibited CYP2D6 and the IC_{50}-value of 0.65 µM for levallorphan indicated strong inhibition. CYP3A4 on the other hand was not reversibly inhibited by either of the compounds. Conversely CYP3A4 was inhibited in a time dependent manner by levallorphan. The time and concentration dependency can be seen in Figure 6.6a. In addition to the time dependency the low NR for levallorphan in CYP3A4 (Table 6.3) showed evidence of a NADPH dependence, which further indicated the occurrence of TDI. In the results for levallorphan on CYP2D6 a concentration dependent loss of activity was noticeable (Figure 6.6c). This implied a significant reversible inhibition at the concentration used and due to the low IC_{50} of levallorphan (0.65 µM) this reversible component was difficult to eliminate. From the activity versus time curve in Figure 6.6c it could however be understood that this loss of activity was not time-dependent. Considering the NR values around one (Table 6.3), it was also evident that the loss of CYP2D6 activity in the presence of levallorphan was not NADPH dependent. Levallorphan which acts as a strong reversible inhibitor on CYP2D6 and a time dependent inhibitor of CYP3A4, and for which evidence of reversible inhibition was lacking, could be an informative case regarding the catalytic properties and selectivity between the two enzymes. It could be explained by that the strong reversible inhibition of CYP2D6 resulted from an orientation of the ligand that did not produce any metabolites, referred to as a non-productive binding mode. In contrast, CYP3A4 favored orientations that generated metabolites, i.e. productive binding modes, among others an orientation resulting in a reactive metabolite causing the TDI.

Table 6.2 Inhibition results for levallorphan and levorphanol.

	Reversible Inhibition (IC_{50})		Time Dependent Inhibition (NR)[a]	
	CYP3A4	CYP2D6	CYP3A4	CYP2D6
Levorphanol	>50 µM	11.5 µM	0.87[b]/0.79[c]	0.80[b]/0.95[c]
Levallorphan	>50 µM	0.65 µM	0.45[b]/0.27[c]	1.1[b]/0.87[c]

[a] NR calculated from 20 min sample
[b] The NR values refer to incubations at 12.5 µM
[c] The NR value refer to incubations at 25 µM

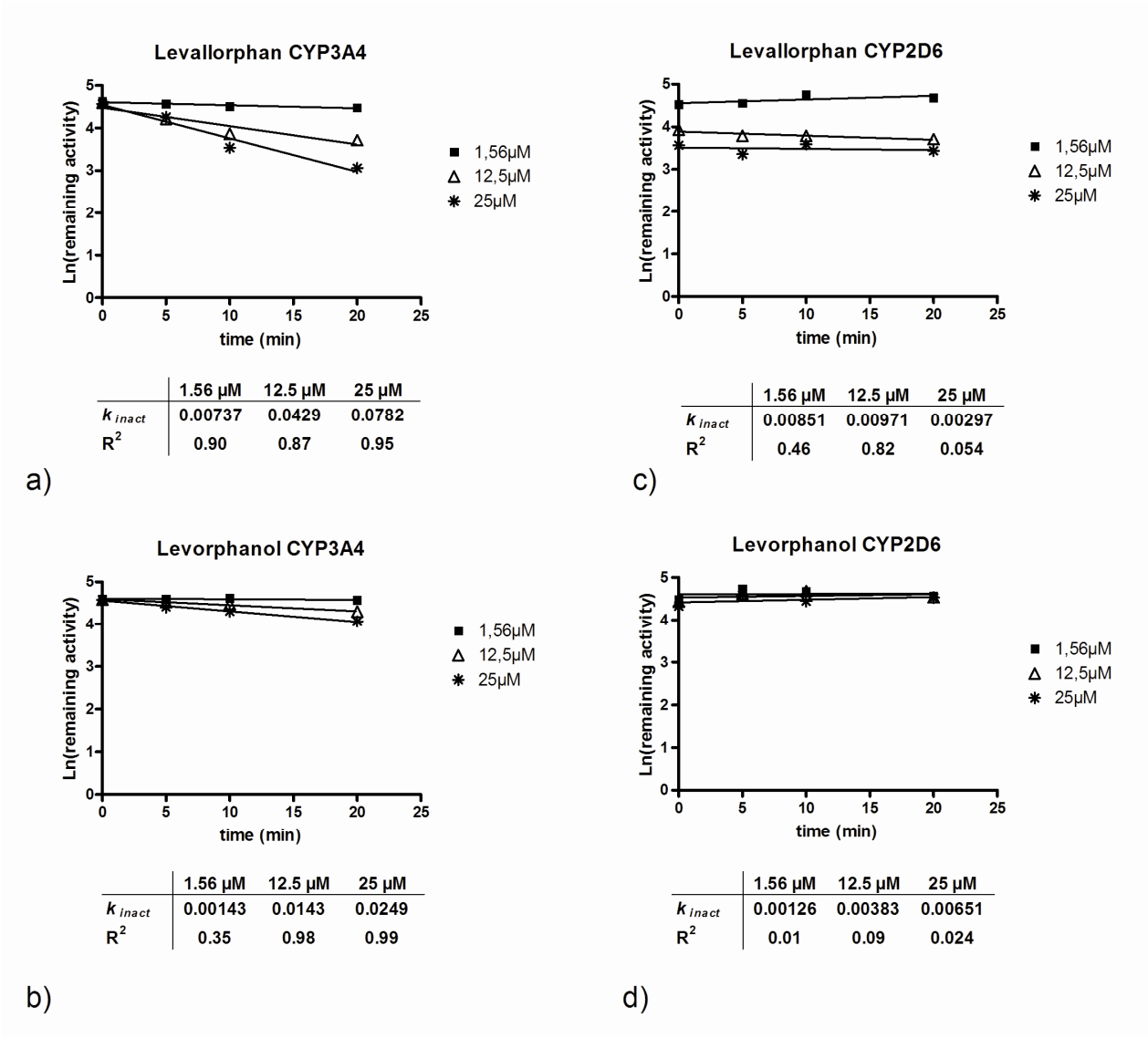


Figure 6.6 Results from TDI experiments a) levallorphan on CYP3A4, b) levorphanol on CYP3A4, c) levallorphan on CYP2D6, and d) levorphanol on CYP2D6.

According to the hypothesis from the TDI results, a productive binding mode was favored in CYP3A4 that was not available in CYP2D6, it was of interest to identify the orientation of levallorphan, i.e. the part of the molecule assumed to cause the TDI. Therefore incubations with glutathione were performed in order to trap possible reactive metabolites and assign the structure of the conjugate. The resulting total ion chromatogram (TIC) is shown in Figure 6.7. There was a large peak present in CYP3A4 at *m/z* 589, corresponding to the glutathionyl conjugate of levallorphan, which was not present in the CYP2D6 incubations. Based on the fragment ions in the MSMS the proposed structure of the conjugate was **16** and a suggested mechanism for its formation is presented in Scheme 6.3. An alternative pathway could go via the epoxide but since the epoxide itself or the diol could not be identified this was not considered likely. Additional argument for the mechanism proposed in Scheme 6.3 was that the allylic position and its attachment to the nitrogen was particularly susceptible to

oxidation. This pathway would also account for the *N*-dealkylation product as a consequence of C-N bond cleavage of the alpha-carbinolamine (**14**).

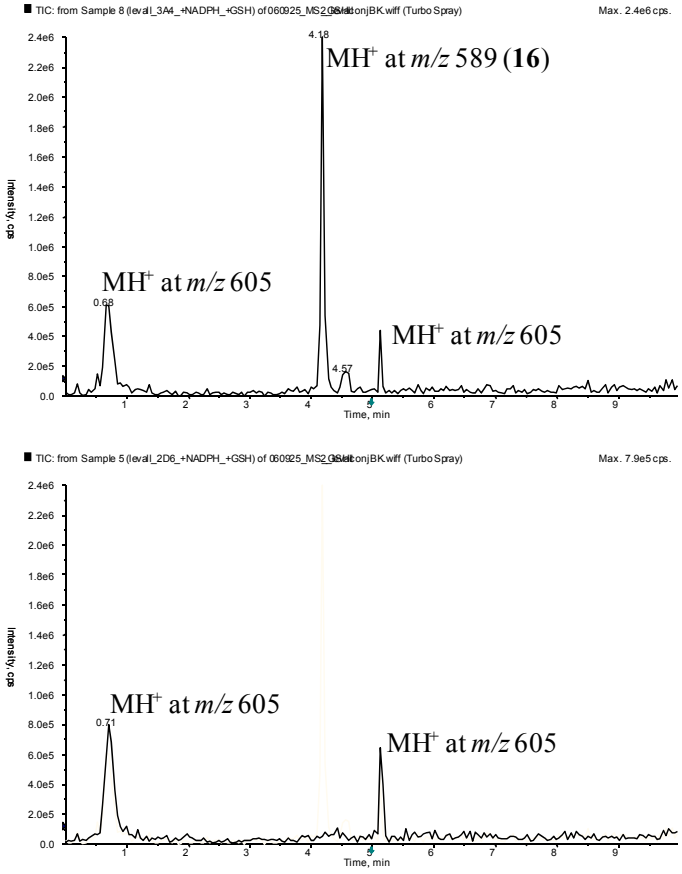
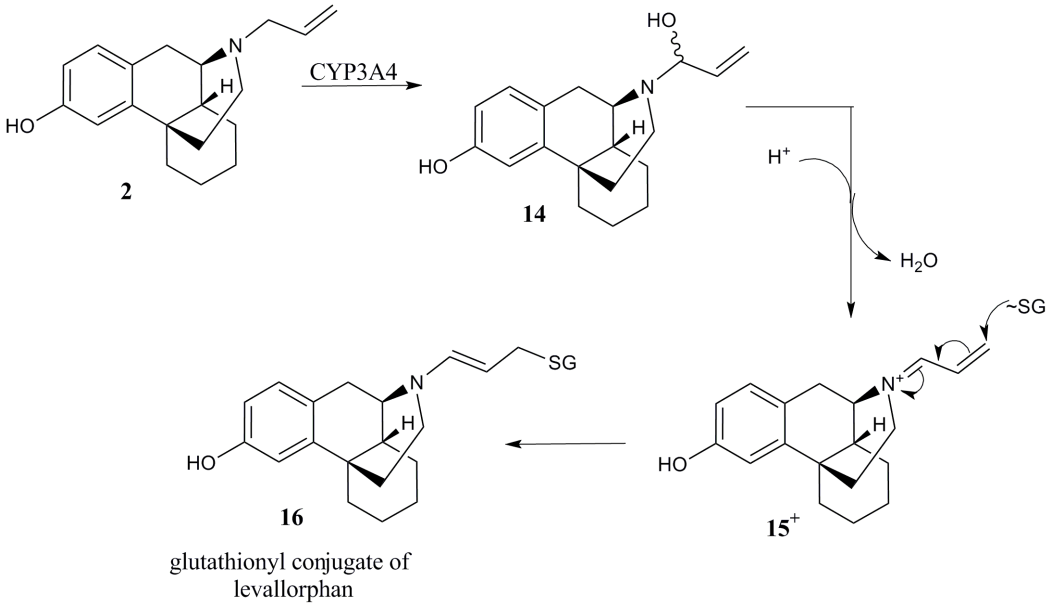


Figure 6.7 Total ion chromatogram for GSH complemented incubations with levallorphan in CYP3A4 (upper pane) and CYP2D6 (lower pane).



*Scheme 6.3 Proposed mechanism for the formation of the levallorphan glutathionyl conjugate (**16**).*

Summarizing the different data on levorphanol and levallorphan in CYP2D6 and CYP3A4 provided a picture of how these two enzymes interacted with the substrates. Results from metabolite identification in the two enzymes showed that in CYP3A4 the substrates had several productive binding modes, which resulted in a variety of metabolites. In CYP2D6 on the other hand the orientation of substrates seemed to be more restricted mainly resulting in hydroxylation of the aromatic moiety. Despite the low metabolic activity of CYP2D6, compared to CYP3A4, the strong reversible inhibition of levallorphan was an indication of a high affinity for the enzyme. This could be explained by the fact that a non-productive binding mode, i.e. a binding mode not resulting in any metabolites, was favored in CYP2D6. On the contrary, the lack of reversible inhibition of CYP3A4 could be explained by that the compounds being readily metabolized and consequently not causing any inhibition. Since TDI is a consequence of a metabolic activation in CYP3A4 this also indicated that levallorphan adopted a productive binding mode in this enzyme but not in CYP2D6. From the results of the GSH trapping experiment this orientation had presumably the nitrogen atom and the allylic side chain pointing towards the heme. The fact that this orientation was not favored in CYP2D6 is in agreement with the CYP2D6 pharmacophore, which states that a basic nitrogen is present 5-7 Å from the site of metabolism that presumably is involved in electrostatic interactions with Glu216 or Asp301.[43] Also the findings in CYP3A4 were in agreement with what has previously been identified from 3D-structures and dockings: a large and flexible active site cavity, lacking specific interactions that consequently resulted in a wealth of docking solutions. The metabolite identification, together with the different inhibition studies, highlighted some of the differences between the two enzymes and could guide the understanding of the affinity and selectivity for CYP3A4 and CYP2D6. To follow up the hypothesis on the orientation of ligands in CYP2D6 the CYP2D6 pharmacophore was further explored.

6.2.2 Exploration of the CYP2D6 pharmacophore (Paper III)

As mentioned earlier the CYP2D6 pharmacophore suggests an orientation of the ligand with the basic nitrogen pointing away from the heme, interacting with acidic amino acids residues in the active site cavity (Figure 2.3). The results from the previous study also indicate that the nitrogen towards the heme is not a preferred orientation in this enzyme. According to this result, *N*-dealkylation reactions should not be commonly seen in CYP2D6 or if observed, they are generally with a low turnover. *N*-dealkylation is a frequently observed reaction and the primary enzyme catalyzing this reaction is CYP3A4[91] but it is also observed in CYP2D6. The hypothesis raised for this study was that *N*-dealkylation occurs in CYP2D6 when the preferred SOM, based on the pharmacophore, was blocked towards oxidative metabolism. In order to explore this hypothesis the literature was searched for CYP2D6 substrates metabolized through *N*-dealkylation as a major pathway and substrates metabolized via other biotransformations. In order to have a comparable environment around the nitrogen only methyl- and ethylamines were included. The resulting data set consisted of 43 compounds of which 13 were assigned as being *N*-dealkylated as primary metabolic

pathway. From the whole data set, 20 compounds were also selected for experimental determination of major metabolites.

In order to do this evaluation the preferred SOM based on the CYP2D6 pharmacophore needed to be assigned. For this purpose a ligand-based computational approach was used – FLAP (Fingerprints for Ligands And Proteins).[76] This software generated 4-point pharmacophores based on GRID-MIFs which could be used as a fingerprint describing the molecule. The considered properties used were hydrophobic center, hydrogen bond donor and hydrogen bond acceptor. For the assessment of preferred SOM in CYP2D6, FLAP fingerprints were generated for two known substrates with characterized SOMs, dextromethorphan, a 7 Å substrate, and tropisterone, a 10 Å substrate (Figure 6.8b). These were then compared with the corresponding fingerprint of a query molecule and a best conformer overlap was obtained (see general procedure of FLAP in Figure 6.8a). From this overlap the preferred SOM was assigned as the part of the query molecule that coincided with the SOM in the template molecule.

The FLAP results for the 13 compounds reported as being metabolized through *N*-dealkylation as the major pathway are highlighted in Figure 6.9. For eight of these compounds (dexfenfluramine, hydromorphone, morphine, chlorpheniramine, Bionet-BB 6N-708, sertraline, citalopram, and fluoxetine) the preferred SOM was a part of the molecule that was blocked towards an oxidative reaction, the remaining five compounds did not have any blocking groups. The second group of compounds contained those that were primarily metabolized by other biotransformations (30 compounds). For the compounds in this group the FLAP alignment did not identify any stable groups at the preferred SOM except for dextrorphan, for which the phenolic OH group was picked. Accordingly, all except one of the compounds that were blocked at the preferred SOM were found in the group with *N*-dealkylation as the major metabolic pathway. However, among the *N*-dealkylated compounds there were some compounds lacking blocking groups, thus not following the hypothesis. Of these compounds lidocaine and Bionet-BB 6N-710 were stable according to their Cl_{int} values and consequently not good substrates for CYP2D6.

In summary, the hypothesis that *N*-dealkylation occurs in CYP2D6 when the preferred SOM is blocked towards metabolism seemed to hold in most cases. However, the number of compounds that were *N*-dealkylated by CYP2D6 was limited and the results only show general trends and could not be assigned statistical relevance. Unfortunately it was not possible to quantify the amount of metabolite formed, in addition the dataset did not contain enough compounds for an appropriate comparison between blocked and unblocked compounds with the same scaffold. Nevertheless, the trend observed from the data was that for the *N*-dealkylated compounds this metabolic pathway was the only one observed, or in combination with other minor metabolites. Hence, one possible postulation could be that *N*-dealkylation is a second choice reaction in CYP2D6 that could be brought forward by inhibition of other more preferred metabolic pathways.

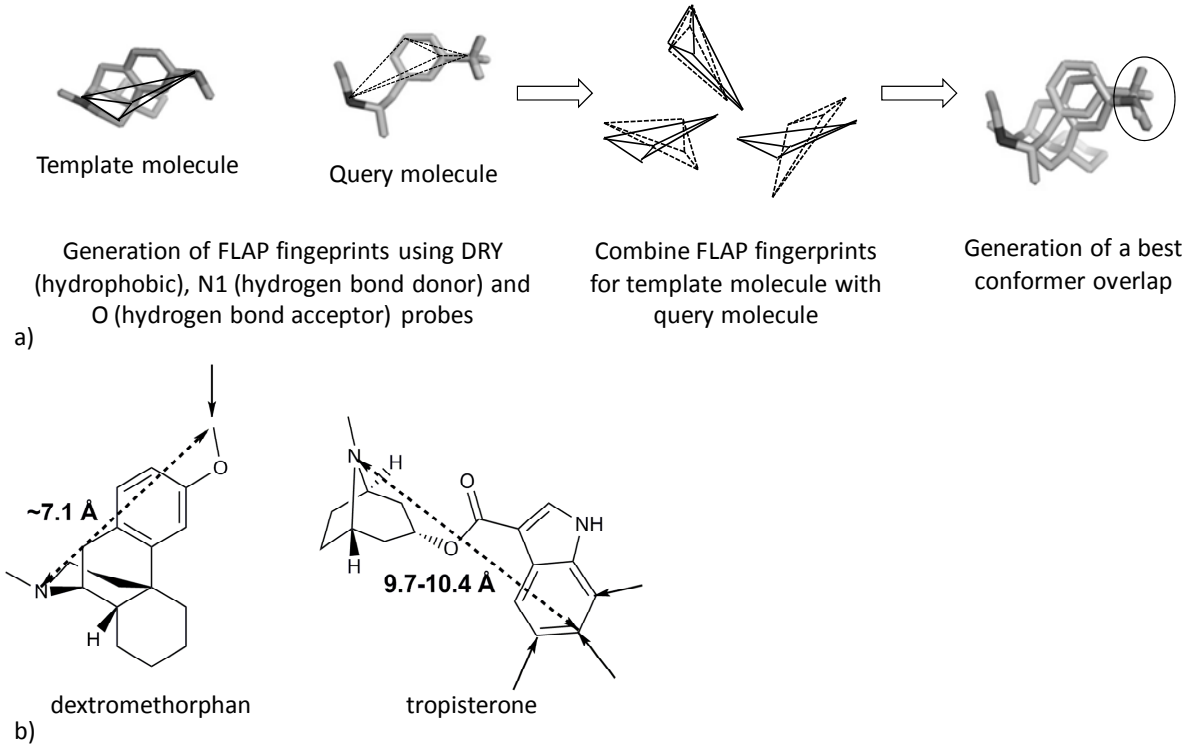


Figure 6.8 a) FLAP procedure and b) the two template molecules used for the FLAP analysis with their SOM indicated with arrows.

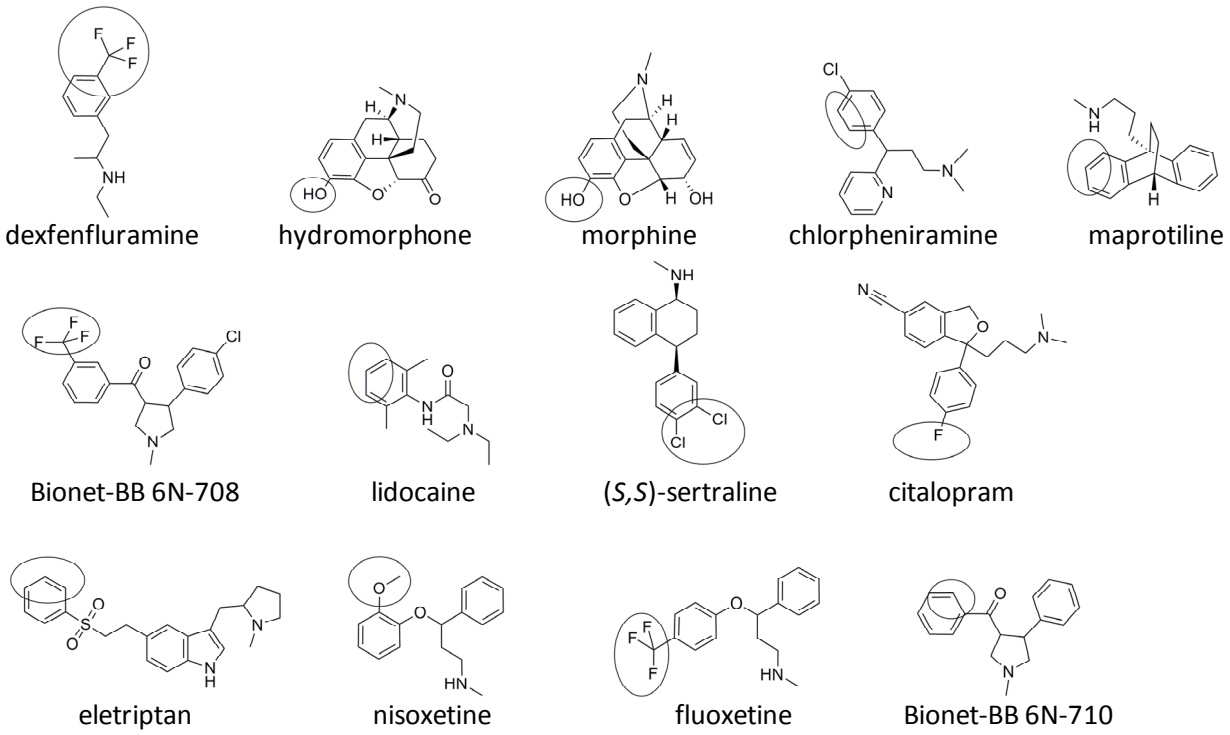


Figure 6.9 Results from FLAP alignment for compounds reported as being metabolized via N-dealkylation as major pathway. Preferred SOM assigned from FLAP is encircled.

If this is the case, it would also be interesting to see if this could be translated to CYP2D6 inhibitors in order to rationalize the inhibitory binding mode. Accordingly, in drug design it is often of importance to consider metabolic stability/metabolite identification together with inhibition of the enzyme for increased knowledge of the interactions and DMPK properties. Some aspects of this will be further discussed in section 6.4.

6.3 Site of metabolism

In the process to understanding and improving the metabolic behavior of a drug, a central assay is determination of SOM. The SOM is also informative when trying to rationalize the selectivity for CYPs. To experimentally assign the structure of a metabolite is time-consuming both regarding time of analysis and the interpretation of data. As a guide to metabolite identification there are also predictive software available. In paper I, MetaSite[75] was used to evaluate how the different 3D-structures performed in predicting SOM. MetaSite is a structure-based approach that considers both the complementarity between enzyme and ligand as well as the reactivity of different parts of the substrate. Despite the differences identified in the 3D-structures the success rate for MetaSite was similar independent on which structure that was used with around ~80-90% correct predictions of SOM within the top three ranked solutions, except for the CYP3A4 homology model that only showed ~50% success rate. MetaSite has also previously been reported to perform well in SOM predictions for a variety of compounds.[92]

In Paper III *N*-dealkylation in CYP2D6 was studied, which is supposed to be a more unfavorable reaction in this enzyme. Hence, it was of interest to see if MetaSite also could predict this reaction. For the group of compounds not *N*-dealkylated the success rate for prediction of SOM was 87% but MetaSite was only able to predict 15% of the *N*-dealkylated compounds. In this study the ligand-based approach, FLAP (see 6.2.2), was also used. For this evaluation it was formulated that *N*-dealkylation would be the principal biotransformation route if the preferred SOM coincides with a part of the query molecule that was blocked towards metabolism. Using this approach, the group with *N*-dealkylated compounds was correctly predicted in 62% of the cases, which was far better than the performance of MetaSite for this group. Regarding the compounds not metabolized by *N*-dealkylation the FLAP method assigned the correct SOM in 77%, which was not as good as MetaSite but still a satisfactory result. To sum up, this ligand-based pharmacophore approach seemed to be a good alternative and complement to MetaSite when studying biotranformation in CYP2D6. Since the CYP3A4 substrates are far more diverse in structure it is doubtful if this approach will be as successful for this enzyme.

6.3.1 Mass-MetaSite – a semi-automated software for metabolite identification (Paper IV)

Predictive software for metabolite identification can be of great help in assigning the SOM, but in most cases this needs to be determined experimentally. As mentioned earlier this is a labor intensive assay hence the interest in the development of software for enhanced handling and interpretation of the data. In Paper IV, Mass-MetaSite, a semi-automated software for metabolite identification was tested. This is an extension of MetaSite, which uses the information from MS raw data in combination with SOM prediction to assign the metabolites. A flowchart of the approach is depicted in Figure 6.10 and further described in Paper IV. The purpose of the software was to mimic the work of a biotransformation scientist in the process of metabolite identification, i.e. rationalization of fragment ions in CID spectra and compare those from the parent compound with those of the metabolite. The software was evaluated with a public data set, containing 20 compounds, and an AstraZeneca in-house data set of 15 compounds. For the public data set metabolite identification was performed in human liver microsomes (HLM) (92 metabolites in total) and for the in-house compounds, previously reported results in HLM were used (49 metabolites). For this evaluation Q-TOF data were used both MSMS and MS^E (fullscan data using a high- and a low-energy scan function).[93] The first task for the program was to identify peaks related to the parent. The recognized peaks were color coded to indicate if they were first or second generation metabolites. There was also a third coloring for peaks originating from reactions not recognized by the software, referred to as missing reactions. In the beta version used here Mass-MetaSite could only suggest a structure for the first generation of metabolites. Consequently, the structural elucidation was done at approximately one half-life of the parent compound with the aim to reduce the generation of many secondary metabolites. For the data set and incubation conditions used, ~70% were identified as first generation metabolites (Figure 6.11) and these were further on assigned a structure. Approximately 10% of the recognized peaks were categorized as missing reactions. This included, for example, some metabolites that were actually second generation metabolites, dealkylations where the corresponding aldehyde was observed, and amide hydrolysis products. Since the software is an extension of MetaSite, it only considers CYP-mediated reactions and consequently amide hydrolysis was not included in the list of reactions. When importing MSMS raw data into Mass-MetaSite the peaks of interest have already been selected by the operator from the full scan experiment. When using MS^E, on the other hand, the software needed to identify the parent related peaks. For the public data set, five metabolites (out of 92) were missed in the peak recognition from MS^E data. This included one co-eluting peak, three with poor mass accuracy and one minor metabolite.

The most important part, which would be the most time-saving, was however the software´s ability to assign structures of the metabolites. For both the public data set and the in-house compounds Mass-MetaSite assigned a structure, in first rank, that agrees to that obtained from manual interpretation in 77% of the cases. The ranking of the results were based on a scoring considering matches/mismatches of fragment ions

for the proposed structures. To this score the MetaSite result could be added for a final rank, and this would discriminate between structures having the same fragment matching score. In most cases structures having the same score coincide with metabolites for which their SOM could not be assigned to a single carbon. The MetaSite prediction factor could also be a guide for metabolites assigned with high uncertainty (low score), for example metabolites with complex fragmentation that was not in accordance with that of the parent. Since the software interprets the data by comparing the spectra of the parent with that of the metabolite it will have difficulties assigning structures of metabolites that fragment differently compared to the parent.

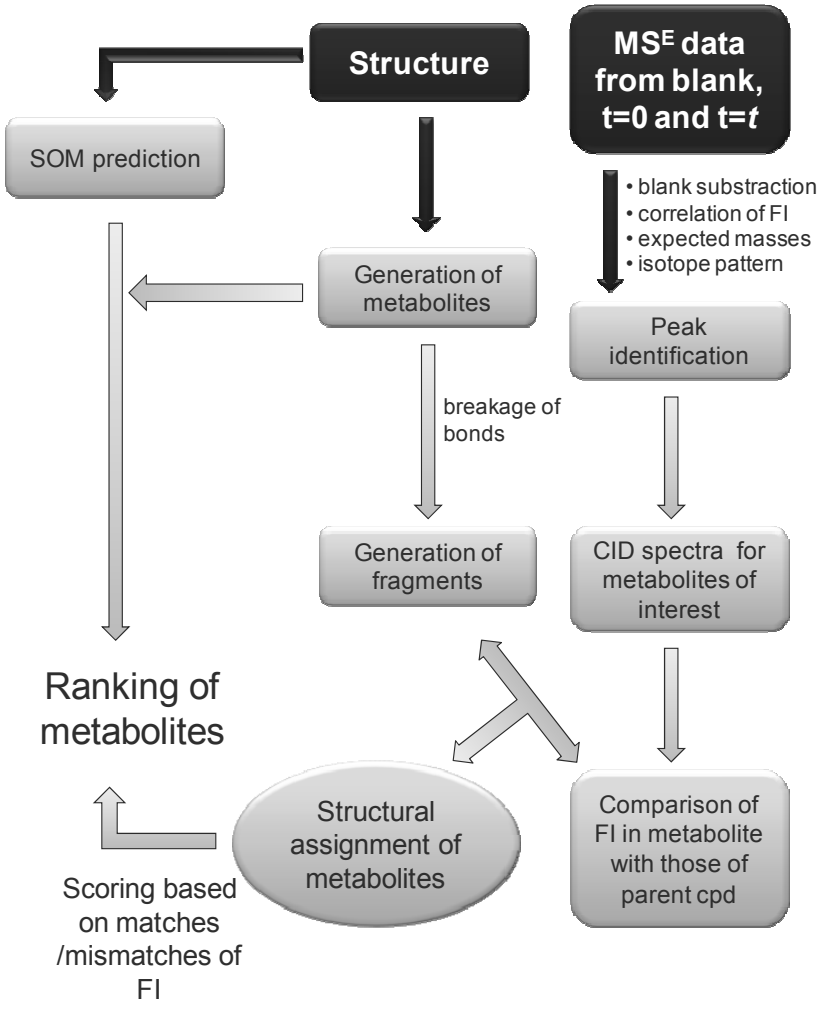


Figure 6.10 Flowchart of the Mass-MetaSite procedure using MS^E data.

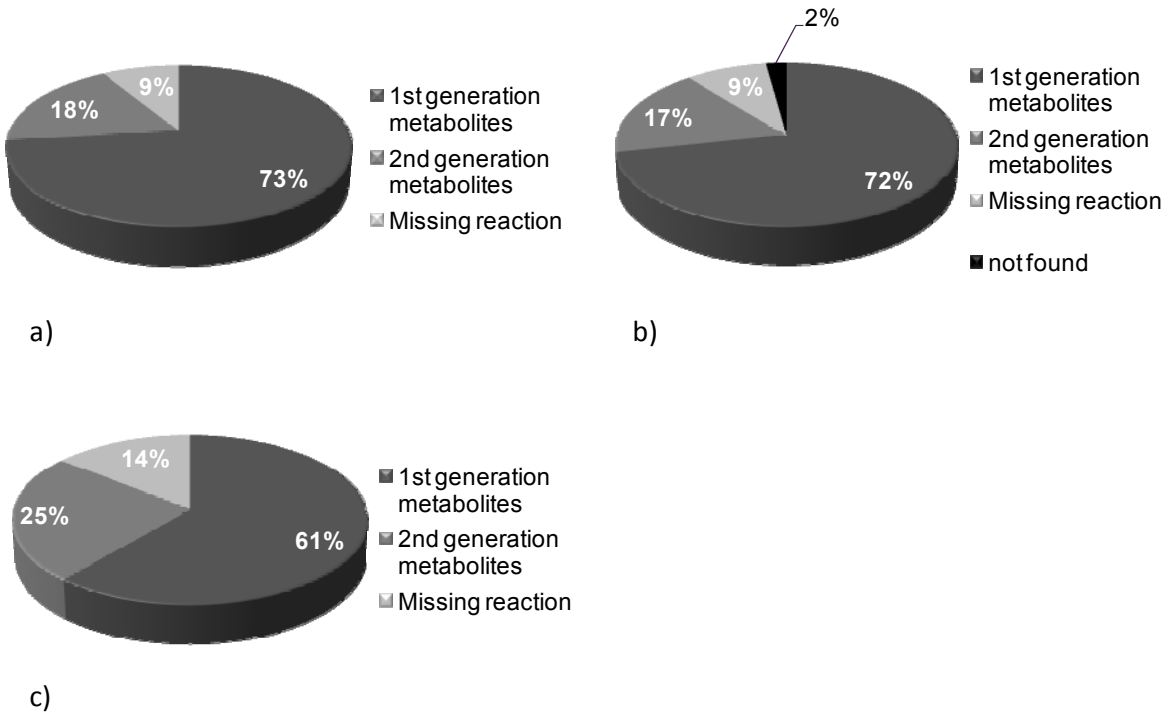


Figure 6.11 Peak recognition from Mass-MetaSite for a) MSMS data on public data set (92 metabolites), b) MS^E data on public data set (92 metabolites) and c) MSMS data on in-house data set (49 metabolites).

The results from the Mass-MetaSite validation indicated that this software could reduce the work for a biotransformation scientist. Even if not all structures could be assigned, time could be saved on interpretation of a majority of metabolites and more effort could be spent on complex fragmentation patterns and those not recognized by the software. Enabling a more high-throughput metabolite identification would clearly contribute to successful drug design and the understanding of ligand-enzyme interactions. Currently there are other software available for handling the of MSMS data, like Metabolynx[70, 71] from Waters and MetWorks[69] from Thermo Scientific. However, these only generate a metabolite profile with elemental compositions. For example, in the case of hydroxylations it is of interest to identify the SOM and especially when searching the orientation of the substrate in the active site. Here Mass-MetaSite could be a positive complement to the existing software and could facilitate metabolite identification. In the context of this thesis, Mass-MetaSite could contribute to the understanding of affinity and selectivity for CYPs by enabling a more high-throughput delivery of SOM data thus providing possibilities to consider larger datasets.

6.4 Some aspects of drug design

The understanding of selectivity and affinity for the CYP enzymes as well as the assessment of SOM all come together as important factors in successful drug design. As shown in this thesis, much can be gained by considering a broader view including inhibition potentials, metabolic stability and SOM to understand the interaction with an enzyme. The use of computational tools will also contribute with the possibility to visualize and identify specific interactions and characteristics within enzymes and ligands. Some important aspects of drug design will be highlighted below.

6.4.1 The effect of blocking SOM on metabolic stability and inhibition potential

One of the primary tasks in drug discovery and drug design is to solve issues with metabolic stability. There are different approaches to improve metabolic properties of a drug for example blocking soft spots, removal or isosteric replacement of metabolically labile spots, or reduce lipophilicity.[94] In Paper III blocked and unblocked CYP2D6 substrates were studied regarding their likelihood to undergo *N*-dealkylation. From this study another aspect of drug design was brought forward. In addition to metabolite identification Cl_{int}- and IC_{50}-values were also determined for the experimental data set, containing 20 compounds. An analysis was done of the distribution of high clearance (>3 µl/min/pmol P450), moderate clearance, and low clearance (<0.7 µl/min/pmolP450) compounds as well as reversible inhibitors (IC_{50} <10 µM) in the different groups (Figure 6.12). From this data two trends could be observed: in the group of compounds that were *N*-dealkylated by CYP2D6, also containing the blocked compounds, there were no high clearance compounds but a majority of low clearance compounds (Figure 6.12a). In the second group of compounds, not *N*-dealkylated as primary metabolic pathway, the three Cl_{int} categories were evenly distributed. Continuing with the inhibition results (Figure 6.12b), it could be seen that a majority of the *N*-dealkylated compounds were inhibitors of CYP2D6 while the others had a fifty-fifty distribution of inhibitors and non-inhibitors. In summary, these results indicated a trend that when changing the structure of the molecule in order to prohibit oxidation at the preferred SOM, e.g. by adding blocking groups, the metabolic stability improved but at the same time it tended to result in increased inhibition of the enzyme. This follows the discussion that an unproductive binding mode gives CYP inhibition. This could be exemplified by Bionet 4N-726 and Bionet 6N-708, two compounds with the same core structure with and without blocking group (Figure 6.13). By introducing an additional blocking group in Bionet 6N-708 the Cl_{int}-value was reduced from high to stable, at the same time the inhibitory potential was increased almost ten times. The consequence of increasing inhibition from addition of blocking groups has also been previously reported for CYP2C9.[95] Another way to come around issues with poor metabolic stability could be to identify and disrupt interactions between ligand and enzyme. For this, the energy calculation method described in Paper I could be valuable. By performing molecular dockings and calculating the interaction energies between the enzyme active site cavity and ligand, important interactions could be identified. From the outcome, as a list of

amino acid residues and interaction energies, the type of a crucial interaction (hydrogen bond, π-π interaction, electrostatic interaction etc.) could be assigned. Subsequently, an evaluation could be done whether or not the molecule could be changed to reducing this interaction. In addition, with existing metabolite identification data the analysis could be refined by only considering docking poses that were in accordance with the productive binding mode.

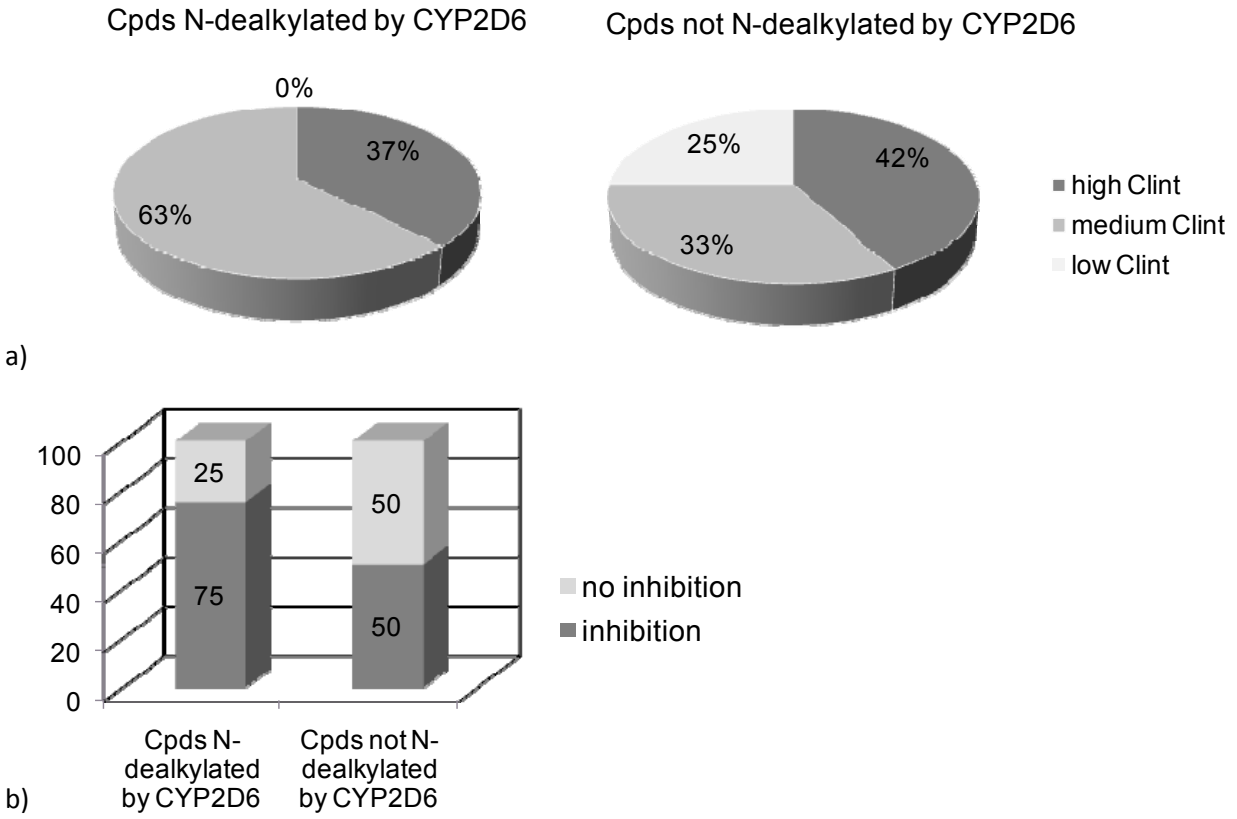
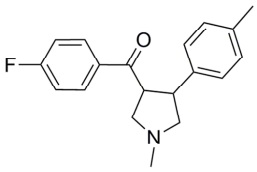
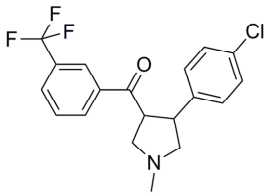


Figure 6.12 Division of CYP2D6 substrates according to their a) Cl_{int}-values and b) inhibition potential.





	4N-726	6N-708
Clint (µl/min/pmolP450)	4.3	stable
IC_{50} (µM)	3.33	0.38

Figure 6.13 Cl_{int}-values and inhibition data for Bionet 4N-726 and Bionet 6N-708

In summary the combined information from different *in vitro* assays, e.g. inhibition, metabolic stability and metabolite identification can gain insight and understanding of the metabolic behavior of a drug. For example in the studies on levallorphan and levorphanol, the reversible inhibition properties on one enzyme and the TDI on the other could be used to further rationalize favorable orientations and differences in catalytic properties in the two enzymes. With the use of computational tools the interactions between enzyme and ligand could also be visualized and allow for identification of crucial interaction points.

7 Concluding remarks and future perspective

This study was initiated in order to study affinity and selectivity factors in CYP2D6 and CYP3A4 mediated metabolism. The procedure includes experimental and computational tools and the aim was also to demonstrate the power of this combined approach within drug discovery.

3D-structures of CYPs are important tools when studying CYP-mediated metabolism and with the improvements in protein engineering and crystallization there are currently several structures available. A computational approach was developed with the purpose to identify discriminative enzyme-ligand interactions using dockings of ligands in the different structures available. It was found that they could be discriminated based on the energy of interaction calculated per amino acid. However, even if the differences can be successfully rationalized using the methodology proposed it is difficult, if not impossible, to judge which 3D-structure is more appropriate to represent the enzymes considering the protein-ligand interactions. In addition to the ability of the method to identify discriminative enzyme-ligand interactions, and amino acid residues, when comparing 3D-structures this technique can also be utilized for identification of specific interactions that can be diminished/strengthen with the aim to change the affinity for an enzyme.

Moreover, the integrative approach to study inhibition properties and metabolite identification in CYP2D6 and CYP3A4 gave the opportunity to highlight some catalytic and selectivity properties for the isoforms. The analysis of levallorphan and levorphanol in these CYP isoforms show that, 1) CYP2D6 demonstrates high affinity but generate preferably non-productive binding modes of the two compounds with a limited number of different potential orientations, 2) CYP3A4, on the contrary, allow for more unrestricted orientations including productive binding modes. These results are in agreement with what is known and observed from 3D-structures and computational studies of the enzymes – CYP3A4 with its large active site results in more unspecific interactions with ligand while CYP2D6 interactions are more constrained. Consequently, using different *in vitro* assays together with computational approaches enhances the understanding of selectivity and affinity for an enzyme.

Additional studies on CYP2D6 showed that, in accordance with the pharmacophore, *N*-dealkylations are in general second choice reactions in this enzyme that might occur as a consequence of inhibtion of other preferred metabolic routes. Based on the well

characterized pharmacophore, SOM predictions in CYP2D6 can be successfully performed with a ligand-based pharmacophore approach (FLAP), which also predicts the unfavorable CYP2D6-mediated *N*-dealkylation.

In addition to the prediction of SOM, metabolite identification is also determined experimentally throughout the thesis. By using the semi-automated software MassMetaSite a majority of the metabolites can be elucidated and the use of this software can alleviate the work load of a biotransformation scientist and enhance the throughput in the metabolite identification process. In turn, this information can help to understand the affinity and selectivity for an enzyme and consequently facilitate drug design since it enables SOM information of a larger number of compounds.

Finally, since the different kinetic and metabolic parameters are correlated it is encouraged to consider these together when optimizing the metabolic properties of a drug. Some further evaluation of the approaches and findings could add more value to their use within drug design:

- Evaluation of whether the energy calculation method for protein-ligand interactions can be used to guide the selection of 3D-structure for a particular study.

- Investigate CYP2D6 inhibitors regarding the presence of blocking groups and evaluate if the pharmacophore approach as implemented in FLAP can also be used for prediction of inhibitors and inhibitory binding modes.

- Evaluation of whether the energy calculation method could be used to identify important interactions to disrupt instead of blocking soft spots with the aim to improve metabolic stability and decrease affinity for an enzyme.

8 Acknowledgements

I would like to express my sincere gratitude to all of you who have contributed and supported me during the course of this work.

Special thanks to my supervisors Prof. Collen Masimirembwa, Dr Ismael Zamora and Prof. Kristina Luthman.

Collen – thank you for pushing and inspiring me! Your never ending lack of crazy ideas and your capability to find new angles of a problem never ends to amaze me. Thank you for all discussions and everything you have taught me about thinking, writing and science.

Ismael – thank you for always being able to turn "less good" data into nice projects and for constantly bringing new ideas and solutions. I have really enjoyed working with you, you are a great teacher and I have learnt a lot from you. Thank you for always being there for me and for insisting on my spring-time visits in Barcelona☺

Kristina – thank you for helping me keep my feet on the ground when the Spanniard and Zimbabwian went crazy with all their ideas☺. Thank you for our long talks about everything and nothing and always being positive and supportive. I always leave your office with a smile on my face!

Prof Neal Castagnoli Jr, I am so happy that I got the opportunity to work with you! Your vast knowledge in biotransformation is a true inspiration, there is no one that "push arrows" and "move electrons" like you!

Prof. Tommy B Andersson and Dr Kurt-Jürgen Hoffman for introducing me to the field of Cytochrome P450 and drug metabolism. With this thesis I hope that I can convince you that I am actually capable of finalizing my projects…☺

Dr Lars Weidolf for you interest and support in my work and for always being available for discussions.

Fabiene Fontaine for your amazing work with Mass-MetaSite!

Members of the biotranformation section, Carina, Anna, Emre, Tove, Martin, Sara, Marianne, Susanne, Bosse and Ulrik for all the support and for offering such a nice working environment.

The “old in vitro group” for introducing me to AZ and drug discovery in the best possible way, and making the time at the HA2 lab always fun and inspiring.

“The Panel”, Christine, Susanne, Sara, Marie, Kajsa, Malin, Maria for all the fruitful discussions and invaluable advices in the coffe breaks, making any thesis-setbacks totally disappear. “The Panel” is always right, that’s for sure!

Lovisa Afzelius for your support and long talks during our walks Korsvägen-AZ and productive discussions over a glass of wine or two…

Dr anders Tunek, Dr Richard Thompson and Dr Hugues Dolgos for supporting this project.

All people I have worked with during my years at AstraZeneca for making it a joy to go to work every day.

Besides the life at AZ there is of course friends and family…

Linda, Elin, Pernilla and Kristin ni är mina klippor! Thank you for always being there to share bad times and good times, set-backs and success!

Anna for always being a friend and for bringing some “non-science” into my life☺!

My parents Hans and Lotta and my brother Johan for being the best family one can wish for and for your persistent efforts to try to understand what I have actually done the past years☺

And finally…

Peter, my personal chemistry-hotline, β-blocker and love of my life, and my lovely daughter Emma, you are the best! Jag Älskar Er!!!

9 References

1. Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates?. *Nat. Rev. Drug Discov.* **2004**, *3,* 711-715.

2. Lin, J.; Sahakian, D. C.; de Morais, S. M.; Xu, J. J.; Polzer, R. J.; Winter, S. M. The role of absorption, distribution, metabolism, excretion and toxicity in drug discovery. *Curr. Top. Med. Chem.* **2003**, *3,* 1125-1154.

3. Wienkers, L. C.; Heath, T. G. Predicting in vivo drug interactions from in vitro drug discovery data. *Nat. Rev. Drug Discov.* **2005**, *4,* 825-833.

4. Cutler, S. J.; Block, J. H. Chapter 4: Metabolic Changes of Drugs and Related Organic Compounds. In *Wilson and Grisvold's Textbook of Organic Medicinal and Pharmaceutical Chemistry* . Lippincott Williams & Wilkins, Philadelphia, **2003**, Vol. 11, pp 65-141

5. Estabrook, R. W. A passion for P450s (rememberances of the early history of research on cytochrome P450). *Drug Metab. Dispos.* **2003**, *31,* 1461-1473.

6. Testa, B.; Krämer, S. D. Part 1: Principles and Overview. In *The Biochemistry of Drug Metabolism: principles, redox reactions, hydrolysis;* Kisakürek, V. M., Kolitzus, T., Ed.; VHCA, Wiley-VCH, Zürich Switzerland, Weinheim Germany, **2008**, pp 1-49.

7. Lewis, D. F.; Eddershaw, P. J.; Dickins, M.; Tarbit, M. H.; Goldfarb, P. S. Structural determinants of cytochrome P450 substrate specificity, binding affinity and catalytic rate. *Chem. Biol. Interact.* **1998**, *115,* 175-199.

8. von Moltke, L. L.; Greenblatt, D. J.; Grassi, J. M.; Granda, B. W.; Venkatakrishnan, K.; Schmider, J.; Harmatz, J. S.; Shader, R. I. Multiple human cytochromes contribute to biotransformation of dextromethorphan in-vitro: role of CYP2C9, CYP2C19, CYP2D6, and CYP3A. *J. Pharm. Pharmacol.* **1998**, *50,* 997-1004.

9. Ono, S.; Hatanaka, T.; Miyazawa, S.; Tsutsui, M.; Aoyama, T.; Gonzalez, F. J.; Satoh, T. Human liver microsomal diazepam metabolism using cDNA-expressed cytochrome P450s: role of CYP2B6, 2C19 and the 3A subfamily. *Xenobiotica* **1996**, *26,* 1155-1166.

10. Testa, B.; Krämer, S. D. The Biochemistry of Cytochromes P450 (CYPs) and Flavin Monooxygenases (FMOs). In *The Biochemistry of Drug Metabolism: Principles, Redox Reactions, Hydrolyses*; Kisakürek, V. M., Kolitzus, T., Ed.; VHCA, Wiley-VCH, Zürich Switzerland, Weinheim Germany, **2008**, pp 55-76.

11. Klingenberg, M. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.* **1958**, *75,* 376-386.

12. Garfinkel, D. Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Arch. Biochem. Biophys.* **1958**, *77,* 493-509.

13. Omura, T.; Sato, R. A new cytochrome in liver microsomes. *J. Biol. Chem.* **1962**, *237,* 1375-1376.

14. Omura, T.; Sato, R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J. Biol. Chem.* **1964**, *239,* 2370-2378.

15. Nebert, D. W.; Adesnik, M.; Coon, M. J.; Estabrook, R. W.; Gonzalez, F. J.; Guengerich, F. P.; Gunsalus, I. C.; Johnson, E. F.; Kemper, B.; Levin, W. The P450 gene superfamily: recommended nomenclature. *DNA* **1987**, *6,* 1-11.

16. Nelson, D. R. Cytochrome P450 Homepage. *http://drnelson.uthsc.edu/CytochromeP450.html*. Accessed 2010.

17. Guengerich, P. F. Human Cytochrome P450 Enzymes. In *Cytochrome P450: Structure, Mechanism, and Biochemistry;* Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers, New York, **2005**, pp 377-530.

18. Rendic, S.; Di Carlo, F. J. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metab. Rev.* **1997**, *29,* 413-580.

19. Meyer, U. A. Genetic polymorphisms of drug metabolism. *Fundam. Clin. Pharmacol.* **1990**, *4,* 595-615.

20. Nagata, K.; Yamazoe, Y. Genetic polymorphism of human cytochrome p450 involved in drug metabolism. *Drug metab. pharmacokinet.* **2002**, *17,* 167-189.

21. Poulos, T. L.; Finzel, B. C.; Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **1987**, *195,* 687-700.

22. Ravichandran, K. G.; Boddupalli, S. S.; Hasermann, C. A.; Peterson, J. A.; Deisenhofer, J. Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. *Science* **1993**, *261,* 731-736.

23. Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. *J. Inorg. Biochem.* **2000**, *81,* 183-190.

24. Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak Vinkovic, D.; Jhoti, H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **2003**, *424,* 464-468.

25. Wester, M. R.; Yano, J. K.; Schoch, G. A.; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-A resolution. *J. Biol. Chem.* **2004**, *279,* 35630-35637.

26. Schoch, G. A.; Yano, J. K.; Wester, M. R.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site. *J. Biol. Chem.* **2004**, *279,* 9497-9503.

27. Williams, P. A.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **2004**, *305,* 683-686.

28. Yano, J. K.; Wester, M. R.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-A resolution. *J. Biol. Chem.* **2004**, *279,* 38091-38094.

29. Ekroos, M.; Sjogren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103,* 13682-13687.

30. Rowland, P.; Blaney, F. E.; Smyth, M. G.; Jones, J. J.; Leydon, V. R.; Oxbrow, A. K.; Lewis, C. J.; Tennant, M. G.; Modi, S.; Eggleston, D. S.; Chenery, R. J.; Bridges, A. M. Crystal structure of human cytochrome P450 2D6. *J. Biol. Chem.* **2006**, *281,* 7614-7622.

31. Sansen, S.; Yano, J. K.; Reynald, R. L.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. *J. Biol. Chem.* **2007**, *282,* 14348-14355.

32. Gay, S. C.; Shah, M. B.; Talakad, J. C.; Maekawa, K.; Roberts, A. G.; Wilderman, P. R.; Sun, L.; Yang, J. Y.; Huelga, S. C.; Hong, W. X.; Zhang, Q.; Stout, C. D.; Halpert, J. R. Crystal structure of a cytochrome P450 2B6 genetic variant in complex with the inhibitor 4-(4-chlorophenyl)imidazole at 2.0-A resolution. *Mol. Pharmacol.* **2010**, *77,* 529-538.

33. Domanski, T. L.; Halpert, J. R. Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis. *Curr. Drug Metab.* **2001**, *2,* 117-137.

34. Gotoh, O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* **1992**, *267,* 83-90.

35. Poulos, T. L. Structural biology of heme monooxygenases. *Biochem. Biophys. Res. Commun.* **2005**, *338,* 337-345.

36. Scott, E. E.; He, Y. A.; Wester, M. R.; White, M. A.; Chin, C. C.; Halpert, J. R.; Johnson, E. F.; Stout, C. D. An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100,* 13196-13201.

37. Poulos, T. L.; Johnson, E. F. Structures of Cytochrome P450 Enzymes. In *Cytochrome P450: Structure, Mechanism, and Biochemistry;* Ortiz de Montellano, P. R., Ed.; Kluwer Academic/Plenum Publishers, New York, **2005**; pp 87-114.

38. Meyer, U. A.; Skoda, R. C.; Zanger, U. M. The genetic polymorphism of debrisoquine/sparteine metabolism - molecular mechanisms. *Pharmacol. Ther.* **1990**, *46,* 297-308.

39. Koymans, L.; Vermeulen, N. P. E.; Van Acker, S. A. B. E.; Te Koppele, J. M.; Heykants, J. J. P.; Lavrijsen, K.; Meuldermans, W.; Donne-Op den Kelder, G. A predictive model for substrates of cytochrome P450-debrisoquine (2D6). *Chem. Res. Toxicol.* **1992**, *5,* 211-219.

40. Lewis, D. F.; Eddershaw, P. J.; Goldfarb, P. S.; Tarbit, M. H. Molecular modelling of cytochrome P4502D6 (CYP2D6) based on an alignment with CYP102: structural studies on specific CYP2D6 substrate metabolism. *Xenobiotica* **1997**, *27,* 319-339.

41. de Groot, M. J.; Bijloo, G. J.; Martens, B. J.; van Acker, F. A.; Vermeulen, N. P. A refined substrate model for human cytochrome P450 2D6. *Chem. Res. Toxicol.* **1997**, *10,* 41-48.

42. Guengerich, F. P.; Hanna, I. H.; Martin, M. V.; Gillam, E. M. Role of glutamic acid 216 in cytochrome P450 2D6 substrate binding and catalysis. *Biochemistry (N. Y.)* **2003**, *42,* 1245-1253.

43. Paine, M. J.; McLaughlin, L. A.; Flanagan, J. U.; Kemp, C. A.; Sutcliffe, M. J.; Roberts, G. C.; Wolf, C. R. Residues glutamate 216 and aspartate 301 are key determinants of substrate specificity and product regioselectivity in cytochrome P450 2D6. *J. Biol. Chem.* **2003**, *278,* 4021-4027.

44. Flanagan, J. U.; Marechal, J. D.; Ward, R.; Kemp, C. A.; McLaughlin, L. A.; Sutcliffe, M. J.; Roberts, G. C.; Paine, M. J.; Wolf, C. R. Phe120 contributes to the regiospecificity of cytochrome P450 2D6: mutation leads to the formation of a novel dextromethorphan metabolite. *Biochem. J.* **2004**, *380,* 353-360.

45. Lussenburg, B. M. A.; Keizers, P. H. J.; de Graaf, C.; Hidestrand, M.; Ingelman-Sundberg, M.; Vermeulen, N. P. E.; Commandeur, J. N. M. The role of phenylalanine 483 in cytochrome P450 2D6 is strongly substrate dependent. *Biochem. Pharmacol.* **2005**, *70,* 1253.

46. Rendic, S. Summary of information on human CYP enzymes: Human P450 metabolism data. *Drug Metab. Rev.* **2002**, *34,* 83-448.

47. Ekins, S.; Bravi, G.; Wikel, J. H.; Wrighton, S. A. Three-dimensional-quantitative structure activity relationship analysis of cytochrome P-450 3A4 substrates. *J. Pharmacol. Exp. Ther.* **1999**, *291,* 424-433.

48. Khan, K. K.; He, Y. Q.; Domanski, T. L.; Halpert, J. R. Midazolam oxidation by cytochrome P450 3A4 and active-site mutants: an evaluation of multiple binding sites and of the metabolic pathway that leads to enzyme inactivation. *Mol. Pharmacol.* **2002**, *61,* 495-506.

49. Harlow, G. R.; Halpert, J. R. Alanine-scanning mutagenesis of a putative substrate recognition site in human cytochrome P450 3A4. Role of residues 210 and 211 in flavonoid activation and substrate specificity. *J. Biol. Chem.* **1997**, *272,* 5396-5402.

50. Stevens, J. C.; Domanski, T. L.; Harlow, G. R.; White, R. B.; Orton, E.; Halpert, J. R. Use of the steroid derivative RPR 106541 in combination with site-directed mutagenesis for enhanced cytochrome P-450 3A4 structure/function analysis. *J. Pharmacol. Exp. Ther.* **1999**, *290,* 594-602.

51. GRID version 22.2.2. Molecular Discovery Ltd, http://moldiscovery.com

52. Powis, G.; Jansson, I. Stoichiometry of the mixed function oxidase. *Pharmacol. Ther.* **1979**, *7,* 297-311.

53. Nagano, S.; Poulos, T. L. Crystallographic study on the dioxygen complex of wild-type and mutant cytochrome P450cam. Implications for the dioxygen activation mechanism. *J. Biol. Chem.* **2005**, *280,* 31659-31663.

54. Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. E.; Sweet, R. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G. The catalytic pathway of cytochrome p450cam at atomic resolution. *Science* **2000**, *287,* 1615-1622.

55. Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* **2001**, *14,* 611-650.

56. Guengerich, F. P.; Macdonald, T. L. Chemical mechanisms of catalysis by cytochromes P-450: a unified view. *Acc. Chem. Res.* **1984**, *17,* 9-16.

57. Zhang, Z-Y.; Wong, Y. N. Enzyme kinetics for clinically relevant CYP inhibition. *Curr. Drug Metab.* **2005**, *6,* 241-257.

58. Houston, J. B. Utility of in vitro drug metabolism data in predicting in vivo metabolic clearance. *Biochem. Pharmacol.* **1994**, *47,* 1469-1479.

59. Riley, R. J.; Grime, K.; Weaver, R. Time-dependent CYP inhibition. *Expert Opin. Drug Metab. Toxicol.* **2007**, *3,* 51-66.

60. Kalgutkar, A. S.; Obach, R. S.; Maurer, T. S. Mechanism-based inactivation of cytochrome P450 enzymes: chemical mechanisms, structure-activity relationships and relationship to clinical drug-drug interactions and idiosyncratic adverse drug reactions. *Curr. Drug Metab.* **2007**, *8,* 407-447.

61. Fontana, E.; Dansette, P. M.; Poli, S. M. Cytochrome p450 enzymes mechanism based inhibitors: common sub-structures and reactivity. *Curr. Drug Metab.* **2005**, *6,* 413-454.

62. Gonzalez, F. J.; Korzekwa, K. R. Cytochromes P450 expression systems. *Annu. Rev. Pharmacol. Toxicol.* **1995**, *35,* 369-390.

63. McGinnity, D. F.; Griffin, S. J.; Moody, G. C.; Voice, M.; Hanlon, S.; Friedberg, T.; Riley, R. J. Rapid characterization of the major drug-metabolizing human hepatic cytochrome P-450 enzymes expressed in Escherichia coli. *Drug Metab. Dispos.* **1999**, *27,* 1017-1023.

64. Powis, G. The use of human liver for foreign compound metabolism and toxicity studies. *Drug Metab. Rev.* **1989**, *20,* 379-394.

65. Li, A. P.; Lu, C.; Brent, J. A.; Pham, C.; Fackett, A.; Ruegg, C. E.; Silber, P. M. Cryopreserved human hepatocytes: characterization of drug-metabolizing enzyme activities and applications in higher throughput screening assays for hepatotoxicity, metabolic stability, and drug-drug interaction potential. *Chem. Biol. Interact.* **1999**, *121,* 17-35.

66. Watt, A. P.; Mortishire-Smith, R. J.; Gerhard, U.; Thomas, S. R. Metabolite identification in drug discovery. *Curr. Opin. Drug Discov. Devel.* **2003**, *6,* 57-65.

67. Zhang, H.; Henion, J.; Yang, Y.; Spooner, N. Application of atmospheric pressure ionization time-of-flight mass spectrometry coupled with liquid chromatography for the characterization of in vitro drug metabolites. *Anal. Chem.* **2000**, *72,* 3342-3348.

68. Sundstrom, I.; Hedeland, M.; Bondesson, U.; Andren, P. E. Identification of glucuronide conjugates of ketobemidone and its phase I metabolites in human urine utilizing accurate mass and tandem time-of-flight mass spectrometry. *J. Mass Spectrom.* **2002**, *37,* 414-420.

69. Lim, H. K.; Chen, J.; Cook, K.; Sensenhauser, C.; Silva, J.; Evans, D. C. A generic method to detect electrophilic intermediates using isotopic pattern triggered data-dependent high-resolution accurate mass spectrometry. *Rapid Commun. Mass Spectrom.* **2008**, *22,* 1295-1311.

70. Mortishire-Smith, R. J.; O'Connor, D.; Castro-Perez, J. M.; Kirby, J. Accelerated throughput metabolic route screening in early drug discovery using high-resolution liquid chromatography/quadrupole time-of-flight mass spectrometry and automated data analysis. *Rapid Commun. Mass Spectrom.* **2005**, *19,* 2659-2670.

71. Hill, A. W.; Mortishire-Smith, R. J. Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Commun. Mass Spectrom.* **2005**, *19,* 3111-3118.

72. de Graaf, C.; Vermeulen, N. P.; Feenstra, K. A. Cytochrome p450 in silico: an integrative modeling approach. *J. Med. Chem.* **2005**, *48,* 2725-2755.

73. Goodford, P. J. The basic Principles of GRID. In *Molecular Interaction Fields;* Cruciani, G., Ed.; Wiley-VCH, Weinheim, **2006**, Vol. 27, pp 3.

74. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28,* 849-857.

75. Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *J. Med. Chem.* **2005**, *48,* 6970-6979.

76. Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J. S. A Common Reference Framework for Analyzing/Comparing Proteins and Ligands. Fingerprints for Ligands And Proteins (FLAP): Theory and Application. *J. Chem. Inf. Model.* **2007**, *47,* 279-294.

77. Blundell, T. L.; Sibanda, B. L.; Sternberg, M. J.; Thornton, J. M. Knowledge-based prediction of protein structures and the design of novel molecules. *Nature* **1987**, *326,* 347-352.

78. De Rienzo, F.; Fanelli, F.; Menziani, M. C.; De Benedetti, P. G. Theoretical investigation of substrate specificity for cytochromes P450 IA2, P450 IID6 and P450 IIIA4. *J. Comput. Aided Mol. Des.* **2000**, *14,* 93-116.

79. de Groot, M. J.; Vermeulen, N. P.; Kramer, J. D.; van Acker, F. A.; Donne-Op den A three-dimensional protein model for human cytochrome P450 2D6 based on the crystal structures of P450 101, P450 102, and P450 108. *Chem. Res. Toxicol.* **1996**, *9,* 1079-1091.

80. Kirton, S. B.; Baxter, C. A.; Sutcliffe, M. J. Comparative modelling of cytochromes P450. *Adv. Drug Deliv. Rev.* **2002**, *54,* 385-406.

81. Stjernschantz, E.; Vermeulen, N. P. E.; Oostenbrink, C. Computational prediction of drug binding and rationalisation of selectivity towards cytochromes P450. *Expert Opin. Drug Metab. Toxicol.* **2008**, *4,* 513-527.

82. Ekins, S.; Waller, C. L.; Swaan, P. W.; Cruciani, G.; Wrighton, S. A.; Wikel, J. H. Progress in predicting human ADME parameters in silico. *J. Pharmacol. Toxicol. Methods* **2000**, *44,* 251-272.

83. Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. Generating Optimal Linear PLS Estimations (GOLPE): An Advanced Chemometric Tool for Handling 3D-QSAR Problems. *Quantitative Structure-Activity Relationships* **1993**, *12,* 9-20.

84. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2,* 37-52.

85. Kastenholz, M. A.; Pastor, M.; Cruciani, G.; Haaksma, E. E.; Fox, T. GRID/CPCA: a new computational tool to design selective ligands. *J. Med. Chem.* **2000**, *43,* 3033-3044.

86. Kirkwood, L. C.; Nation, R. L.; Somogyi, A. A. Characterization of the human cytochrome P450 enzymes involved in the metabolism of dihydrocodeine. *Br. J. Clin. Pharmacol.* **1997**, *44,* 549-555.

87. Yue, Q. Y.; Sawe, J. Different effects of inhibitors on the O- and N-demethylation of codeine in human liver microsomes. *Eur. J. Clin. Pharmacol.* **1997**, *52,* 41-47.

88. Benetton, S. A.; Borges, V. M.; Chang, T. K. H.; McErlane, K. M. Role of individual human cytochrome P450 enzymes in the in vitro metabolism of hydromorphone. *Xenobiotica* **2004**, *34,* 344.

89. Hutchinson, M. R.; Menelaou, A.; Foster, D. J. R.; Coller, J. K.; Somogyi, A. A. CYP2D6 and CYP3A4 involvement in the primary oxidative metabolism of hydrocodone by human liver microsomes. *Br. J. Clin. Pharmacol.* **2004**, *57,* 287-297.

90. Afzelius, L.; Raubacher, F.; Karlen, A.; Jorgensen, F. S.; Andersson, T. B.; Masimirembwa, C. M.; Zamora, I. Structural analysis of CYP2C9 and CYP2C5 and an evaluation of commonly used molecular modeling techniques. *Drug Metab. Disp.* **2004**, *32,* 1218-1229.

91. Coutts, R. T.; Su, P.; Baker, G. B. Involvement of CYP2D6, CYP3A4, and other cytochrome P-450 isozymes in N-dealkylation reactions. *J. Pharmacol. Toxicol. Methods* **1994**, *31,* 177-186.

92. Afzelius, L.; Arnby, C. H.; Broo, A.; Carlsson, L.; Isaksson, C.; Jurva, U.; Kjellander, B.; Kolmodin, K.; Nilsson, K.; Raubacher, F.; Weidolf, L. State-of-the-art tools for computational site of metabolism predictions: Comparative analysis, mechanistical insights, and future applications. *Drug Metab. Rev.* **2007**, *39,* 61-86.

93. Bateman, K. P.; Castro-Perez, J.; Wrona, M.; Shockcor, J. P.; Yu, K.; Oballa, R.; Nicoll-Griffith, D. A. MSE with mass defect filtering for in vitro and in vivo metabolite identification. *Rapid Commun. Mass Spectrom.* **2007**, *21,* 1485-1496.

94. Di, L.; Kerns, E. H.; Carter, G. T. Drug-like property concepts in pharmaceutical design. *Curr. Pharm. Des.* **2009**, *15,* 2184-2194.

95. Ahlstrom, M. M.; Ridderstrom, M.; Zamora, I. CYP2C9 Structure-Metabolism Relationships: Substrates, Inhibitors, and Metabolites. *J. Med. Chem.* **2007**, *50,* 5382-5391.