A $P450_{BM3}$ Toolkit for C–H Activation Synthesis

Xinkun Ren

A thesis submitted to the Board of the Faculty of the Physical Sciences

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy of the University of Oxford

St. Peter's College

Michaelmas Term 2016

A $P450_{BM3}$ Toolkit for C–H Activation Synthesis

Xinkun Ren

St. Peter's College

D.Phil Thesis Abstract

Michaelmas Term 2016

Cytochrome P450s are a superfamily of haem-dependent enzymes that catalyse the oxidation of natural organic compounds, most commonly *via* insertion of an oxygen atom from atmospheric dioxygen into a carbon-hydrogen bond. Since there is widespread interest in the selective functionalisation of C–H bonds, P450s have been extensively studied and evolved, particularly $P450_{BM3}$ from *Bacillus megaterium*, which is catalytically self-sufficient. In this thesis, a $P450_{BM3}$ variant library has been developed based on four 'generic accelerators', which have showed enhanced activity towards a broad range of non-natural substrates such as polyaromatic hydrocarbons and terpenes. The library was refined to 48 variants based on the screening results in order to maintain diversity of substrate pocket topology for oxidising a broad coverage of organic molecules. Successful applications include drug metabolism, drug fragment diversification, synthesis of flavour and fragrance compounds, steroid oxidation, late-stage functionalisation in total synthesis, and C–H amination. Total conversion was observed in most cases with a sufficient turnover number that enabled all the products to be fully characterised by NMR and MS. Two novel P450 activities were discovered as oxidative decarboxylation of an α-hydroxy carboxylic acid and C–H amination. The substrate range and varied product profiles suggest that this enzyme library is a good basis for developing selective C–H activation catalysts while expanding the P450 catalytic repertoire to include reactions important in synthetic chemistry, potentially opening new opportunities for biocatalysis.

Acknowledgements

First and foremost I would like to thank my supervisor Prof. Luet Wong for his tireless support and willingness to help. It has been a great honour to work so closely with one of the brightest minds in the P450 field. His enthusiasm and sense of humour have made the past four years thoroughly enjoyable.

Prof. Jeremy Robertson for teaching me numerous things related to chemistry, for always finding time and being patient to help, and for inspiring me with new ideas. It has been a privilege for having my name on his group website as an “honorary member”. Prof. Kylie Vincent for giving me advices and encouragement on my research career.

Many thanks to Wenyu Chen for her endless efforts on harvesting enzymes, Yushu Li for her enthusiasm and non-stop approach to all the things chemistry related. I am particularly grateful to both of them for finding time to proof read this thesis.

Dr Alize Pennac, for her huge contribution to managing P450 mutant library and for always having a smile on her face and positive attitude around the lab. Kouji Urata, thanks for being a great office buddie. I’ll miss all of the great chat and laughs we had during the past four years. Dr Eachan Johnson, for his guidance and help on taking me through protocols in my first year.

Jack O’Hanlon, for being a fantastic collaborator. His problem solving ability and useful suggestions have been instrumental to the success of lidocaine project. Dr Oleksandr Zhurakovskyi (Alex), for always being ready to help, teaching me things not only chemistry related but also programming.

Sam, Sky, Patri, David and Jessica from Jeremy’s lab for synthesising various compounds as

P450 substrates. Thanks to Dr Nick Rees for running and assigning NMR spectra, Colin Sparrow for running Mass Spectrometry, Dr James Wickens for training me on LC-MS instrument.

Thanks to Peers Lyle, Amy Redston, Elizabeth McFarlane, Yang Cao, Gerogia Lubbock, James Skinner, George Woodward, Jake Jacobs, Leah Barreto, Caitlin Armstrong, Lydia Coyle, Matthew Fisher, and all the members in Jeremy Robertson’s lab for accompany.

Finally, I’m incredibly grateful for the support from my parents and my wife Zifu Wang. Thank you for all your encouragement and patience over the past four years. I couldn’t have done it without you!

List of Publications and Awards in DPhil Program

Publications

1. Drug Oxidation by Cytochrome P450BM3: Metabolite Synthesis and Discovering New P450 Reaction Types, Xinkun Ren, Jake Yorke, Emily Taylor, Ting Zhang, Weihong Zhou, Luet Lok Wong, Chemistry – a European Journal, 2015, 21, 15039-15047

2. Synthesis of Imidazolidin-4-ones via a Cytochrome P450-Catalyzed Intramolecular C–H Amination, Xinkun Ren, Jack A. O'Hanlon, Melloney Morris, Jeremy Robertson, and Luet Lok Wong, ACS Catalysis, 2016, 6 (10), 6833–6837

Awards

1. Tang Award, China Oxford Scholarship Fund, Oxford, UK, 2013

2. Best Poster Award, 49th European Symposium on Biological and Organic Chemistry, Wales, UK, 2015

Table of Contents

List of Figures

List of Tables

Abbreviations

A_x - UV absorbance at wavelength x nm

B. megaterium - *Bacillus megaterium*

COSY - correlation spectroscopy

CYP - cytochrome P450

DEAE - diethylaminoethyl

DMSO - dimethyl sulphoxide

DNA - deoxyribonucleic acid

dNTP - deoxyribonucleoside triphosphate

E. coli - *Escherichia coli*

EDTA - ethylenediaminetetraacetic acid

FAD - flavin adenine dinucleotide

FID - flame ionisation detector

GC - gas chromatography

GDH - glucose dehydrogenase

HAC - heavy atom count

HMBC - heteronuclear multiple bond coherence

HSQC - heteronuclear single quantum correlation spectroscopy

IPTG - isopropyl β-D-1-thiogalactopyranoside

Kan - kanamycin

LB - lysogeny broth

LOS - lead oriented synthesis

Me-β-CD - methyl-β-cyclodextrin

MS - mass spectrometry

MW - molecular weight

min - minutes

NAD(P)$^+$ - nicotine adenine dinucleotide phosphate

NAD(P)H - nicotine adenine dinucleotide phosphate, reduced form

NMR - nuclear magnetic resonance

NOESY - nuclear Overhauser effect spectroscopy

PCR - polymerase chain reaction

rpm - revolutions per minute

SB - substrate-bound

SF - substrate-free

TLC - thin layered chromatography

THQ - 5,6,7,8-tetrahydroquinoxaline

TON - turnover number

TOS - target oriented synthesis

UV-vis - ultra violet-visible

WT - wild type

w/v - weight by volume

v/v - volume by volume

Chapter One

Chapter 1 Introduction

1.1 C–H bond activation

The development of C–H activating catalysts for the hydroxylation (RH → ROH) of complex organic molecules constitutes an area of intense current interest. Synthetic chemists, who are not limited by biological and chemical methods, have developed transition metal catalysts[1-6] and redox enzymes[7-18] for this purpose.

C−H bond oxidation by chemical methods generally requires high-energy reactants such as peroxides and high-valent metal species and precise control of reaction conditions, often leading to low yield and in particular poor selectivity.[19] Therefore, the use of redox enzymes is especially appealing due to the selective, efficient oxidation of substrates requiring low energy input, and the absence of toxic wastes.

1.1.1 Cytochrome P450s

Cytochrome P450s are a superfamily of haem *b*-dependent monooxygenases which are found in virtually all forms of life. The enzymes, which are named after their distinctive 450 nm absorption band of their CO-bound, ferrous form, have been intensively studied during the last 30 years. All P450 enzymes have a highly conserved tertiary structure and share a similar characteristic protein fold with 12 α-helices (slightly variable) and 5 β-sheets, though only one cysteine residue is invariant, to which the prosthetic group is axially ligated (Figure 1.1).[10, 11]

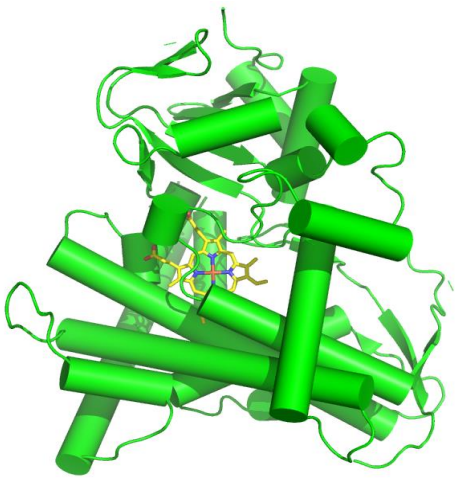


Figure 1.1 Conserved structure of cytochrome P450s (WT $P450_{BM3}$)

The signature P450 reaction is the oxidation of C–H bonds which utilises two electrons commonly derived from NAD(P)H to activate atmospheric dioxygen, one oxygen atom is inserted into the C–H bond while the other oxygen atom is reduced to water (Fig 1.2).[10, 11]

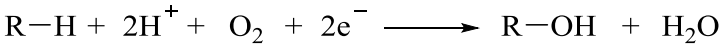


Figure 1.2 Typical reaction types of Cytochrome P450 enzymes

While the insertion of an oxygen atom into a C–H bond is the signature P450 reaction, there are other known reaction types, including desaturation,[20-23] dealkylation,[24-28] olefin epoxidation,[29-35] cyclopropanation,[36-39] heteroatom oxidation,[9, 40-42] ring expansion,[43-46] C–C bond formation,[47] fatty acid decarboxylation,[48, 49] and etc. (Figure 1.3)

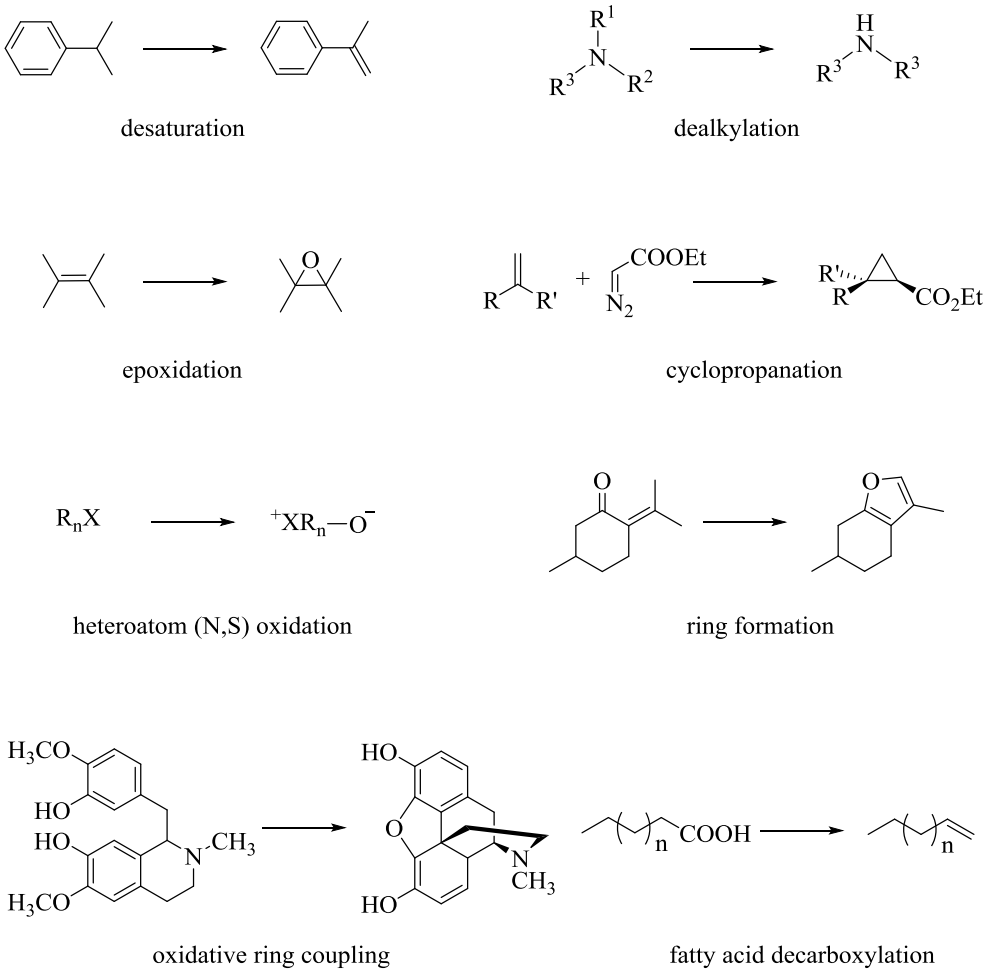


Figure 1.3 Examples of unusual P450 reactions

These diverse activities endow P450 enzymes with a myriad of important physiological roles.[50-53] They enable microorganisms to grow on inert substrates and feature prominently in the metabolic and synthetic pathways in bacteria and fungi, e.g. in the biosynthesis of antibiotics and pharmaceuticals.[54, 55] In higher organisms, they are heavily involved in drug metabolism and steroid synthesis in mammals. Nearly 80% of drugs on the market are oxidised (Phase I metabolism) by P450 enzymes in the liver and small intestine.[56] While in plants, they play vital roles in herbicide resistance and biosynthesis,[57-59] producing many biologically active secondary metabolites such as artemisinin[60], an antimalarial agent isolated from *Artemisia annua*, which is derived from amorpha-4,11-diene oxidation by the P450 enzyme CYP71AV1. Taxol, being another example, is one of the most successful anti-cancer drug currently on the market that is biosynthesised *via* taxadiene oxidation catalysed by a cascade of P450 enzymes.[61-63]

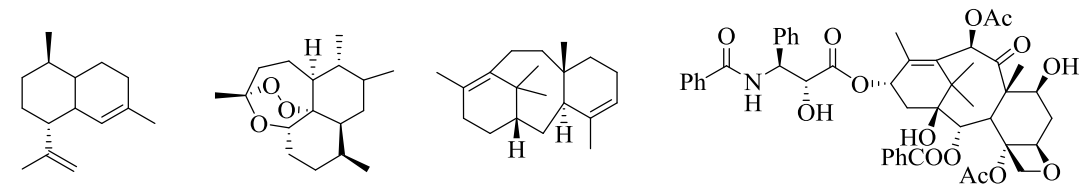


Figure 1.4 The structures of some terpenoid hydrocarbons

So far, over 35,000 CYP genes have been identified in the genome databases (http://dmelson.uthsc.edu/cytochromep450.html) including 63 in human and 287 in the plant *Arabidopsis thaliana*.[64] P450s are relatively rare in prokaryotes – indeed *Escherichia coli*, the most common organism in which prokaryotic P450s are expressed, has none, nor do *Lactobacillus* organisms, which do not synthesize haem.[65] P450s are also found in archaea such as *Sulfolobus solfataricus*, which suggests they are of ancient origin. One possible explanation is that they evolved in response to the toxicity risk posed by rising planetary dioxygen levels, after which the cell membranes of early microbes were formed.[66]

Figure 1.5 shows some different compound types that are naturally oxidised by P450s.[7, 11, 67-71] Some P450 enzymes operate with a high level of substrate specificity such as $P450_{cam}$ which selectively oxidises camphor to 5-*exo*-hydroxycamphor[67] while other members across the CYP superfamily are multifunctional and are active towards a broad range of substrates, notably those involved in the detoxification and metabolism of xenobiotics, are more promiscuous.[7] For example, human CYP3A4 is responsible for the metabolism of approximately 50% of the drug compounds and its versatility makes it one of the most important oxidising enzymes in human body.[68-70]

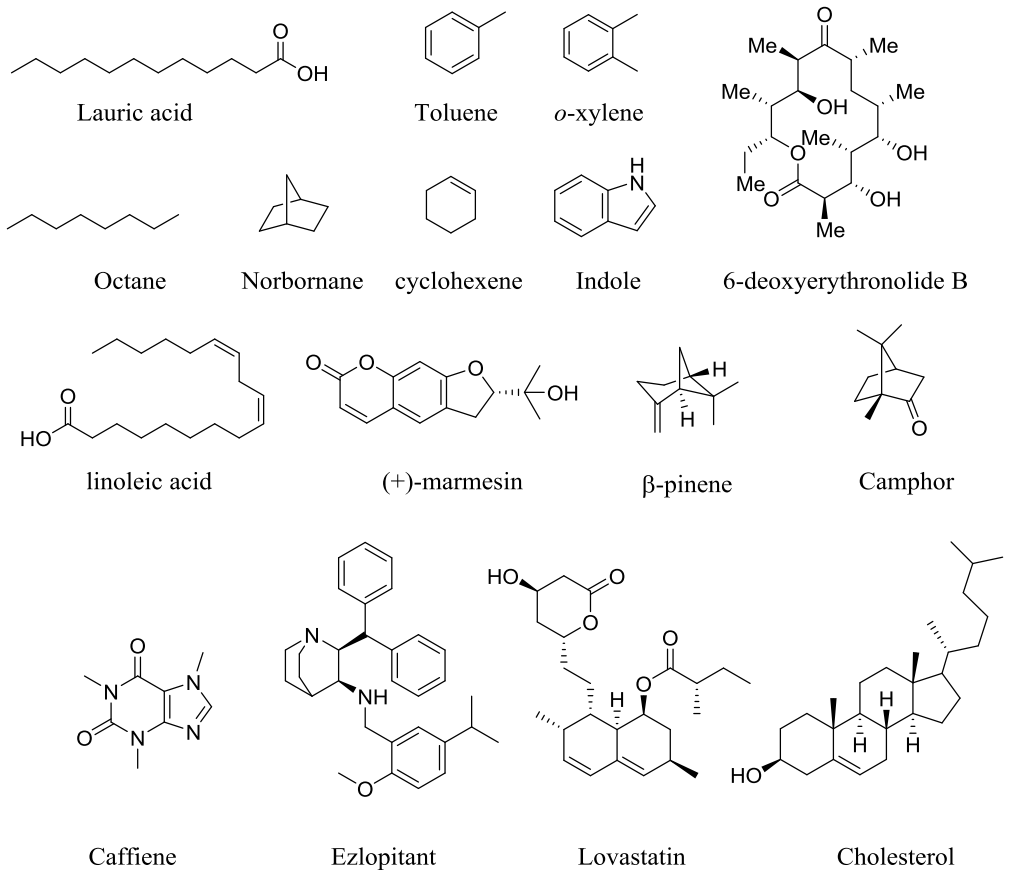


Figure 1.5 Examples of P450 substrates

1.1.2 P450 catalytic cycle

The P450 catalytic cycle, as shown in Figure 1.6, has been widely studied and extensively debated due to the vast variety of reactions catalysed by P450 enzymes and different properties of members across the superfamily.[11] However the general features and steps are now broadly agreed.

The active site of cytochrome P450s contains a haem ring which is tethered to the protein *via* a thiolato ligand from the proximal cysteine. In the resting state (**I**), the ferric centre is low spin (S=1/2) and six coordinate with four equatorial nitrogen ligands to the porphyrin, one proximal thiolate ligand to the cysteine residue, and a water ligand occupies the axial position on the opposite side of the haem. The arrival of a viable substrate induces a conformational change of the active site with the loss of the axial water ligand (**II**) and causes the ferric spin state to change from low-spin to high-spin (S=5/2). This process can be monitored by UV-visible spectroscopy as the Soret band, which arises from a $\pi \rightarrow \pi^*$ electronic transition in the delocalised π system of the porphyrin ring, moves from *ca.* 418 nm to *ca.* 390 nm in a "type I" shift. The displacement of water molecule from the distal axial coordination position also makes the haem reduction potential substantially more oxidising. This facilitates electron transfer from the redox partner to the haem iron to give the ferrous state (**III**) whereupon carbon monoxide binding may interrupt the catalytic cycle and yield the classic CO difference spectrum with a Soret maximum at 450 nm. The binding of molecular oxygen to the ferrous centre gives a dioxygen adduct (**IV**) and a second electron is then transferred from the redox partner, reducing the Fe^{III}-$O_2^{\cdot}$ adduct to generate a short-lived ferric peroxy complex (**V**) which is rapidly protonated to give a hydroperoxide adduct 'Compound 0' (**VI**). A second protonation triggers water loss *via* the heterolytic cleavage of the O–O bond to form a ferryl species 'Compound I' (**VII**), which is thought to be the active entity in most P450 oxidations and was characterised in 2010.[72] A hydrogen atom is then abstracted from the

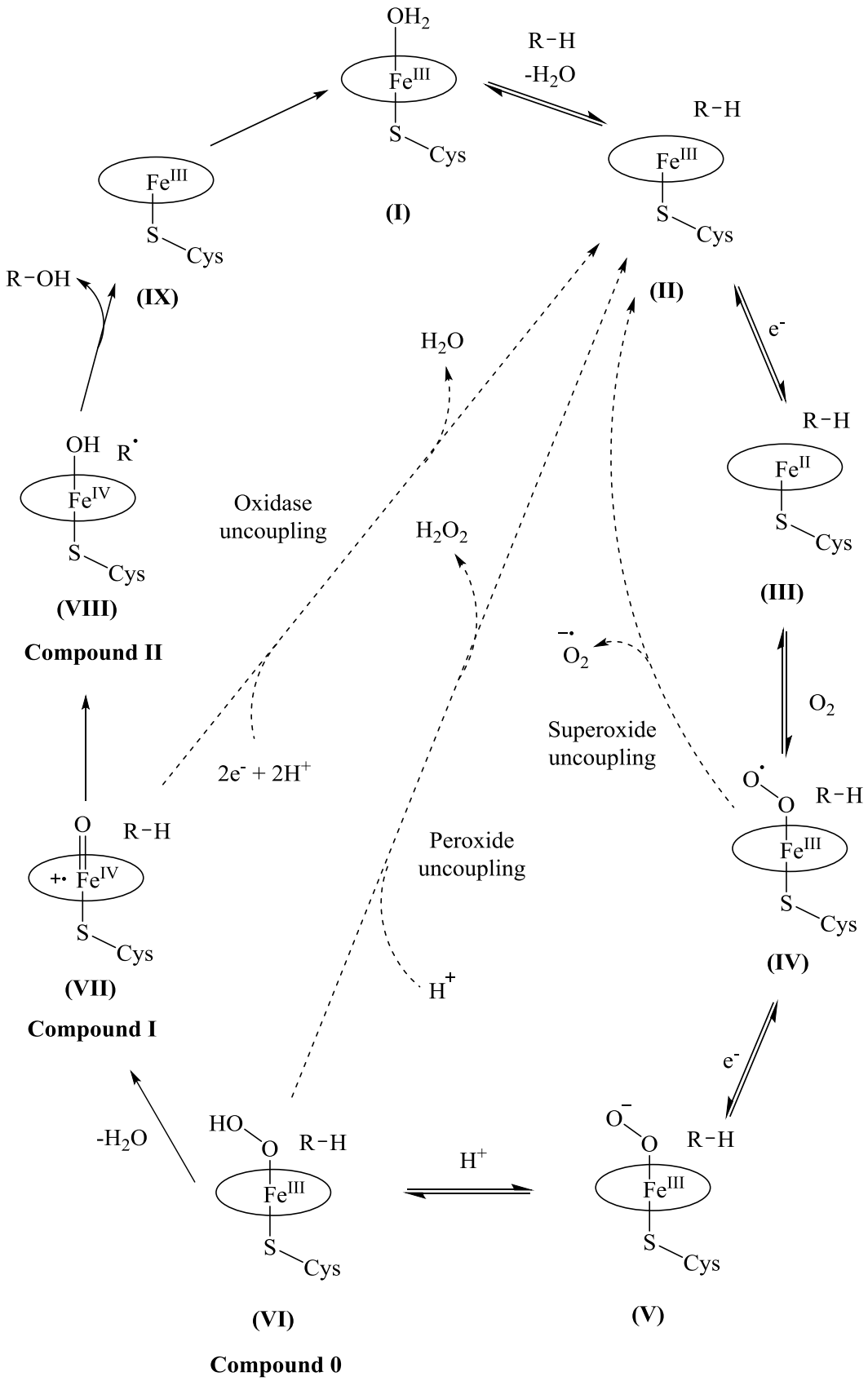


Figure 1.6 The P450 catalytic cycle

substrate (**VIII**) to create a substrate radical which rapidly recombines with •OH to produce the alcohol product, giving the impression that oxygen atom has been "inserted" into the C–H bond. After the alcohol product is released from the enzyme active site, the enzyme returns to its ferric state. It is not known if a water molecule returns to occupy the distal coordination site of the haem iron before another substrate molecule binds within the substrate pocket and the catalytic cycle resumes with the transfer of the first electron.

There are three distinct uncoupling pathways in the P450 catalytic cycle which may hinder substrate oxidation.[73] If the second electron transfer is slow, superoxide may dissociate from **IV**, leading to superoxide uncoupling. If compound 0 is protonated at the iron-bound oxygen atom, hydrogen peroxide is lost, leading to peroxide uncoupling. At high peroxide concentrations, the peroxide uncoupling pathway may be reversed, resulting in the turnover of substrates by peroxide rather than NAD(P)H and oxygen. This process is known as the "peroxide shunt" and is seen with peroxidase enzymes that naturally utilise hydrogen peroxide to function.[74-76] The third uncoupling pathway is oxidase uncoupling, which occurs if a substrate is tightly bound but no hydrogen atom is conveniently positioned for abstraction, compound I (**VII**) may be reduced to give intermediate **II**.

1.2 CYP102A1 ($P450_{BM3}$)

P450s in higher life forms are membrane bound, making handling in a laboratory difficult.[77, 78] For this reason bacterial P450s, which are typically soluble, are preferred. $P450_{BM3}$ (CYP102A1), the subject of this thesis, is among the most well studied bacterial P450[11], with a focus on altering substrate range and product specificity that can be applied in C–H activation synthesis.

1.2.1 General Background

Cytochrome $P450_{BM3}$ is a soluble 119,500 Da enzyme which was the third P450 isolated from *Bacillus megaterium* and is consequently named after this.[79-81] The enzyme consists of a *ca.* 55 kDa P450 or haem domain fused to a *ca.* 65 kDa reductase domain *via* a short linker sequence.[82-84] This built-in electron delivery system confers a high level of catalytic activity unique among P450 cytochromes. The enzyme was first isolated and studied in the Fulco laboratory in the early 1970s and it was identified as a sub-terminal medium- to long-chain fatty acid hydroxylase to give ω–1, ω–2 and ω–3 hydroxy-fatty acids with high activity (>3000 min^{-1}), requiring only NADPH and oxygen to function.[85-88] Fatty amides and alcohols, hydroxylated fatty acids and branched fatty acids were also found to be substrates for $P450_{BM3}$, as well as unsaturated fatty acids,[85, 89, 90] which could be epoxidised.[91-93] It is now known that wild-type $P450_{BM3}$ is not substrate-specific and active towards a broad range of molecules.[11]

1.2.2 Electron transfer

Most P450 enzymes require one or more redox partners to deliver electrons from NAD(P)H to the haem domain and are therefore divided into two broad classes based on the nature of their redox partners. Class I systems includes mitochondrial and bacterial P450s, the enzymes obtain electrons from NADPH *via* adrenodoxin reductase and adrenodoxin.[94, 95] While for class II systems, which includes some bacterial, microsomal and mammalian hepatic P450s, such as human drug metabolising enzymes, the redox partner is a cytochrome P450 reductase (CPR) which contains flavin adnenine dinucleotide (FAD) and flavin mononucleotide (FMN) in separate domains.[96]

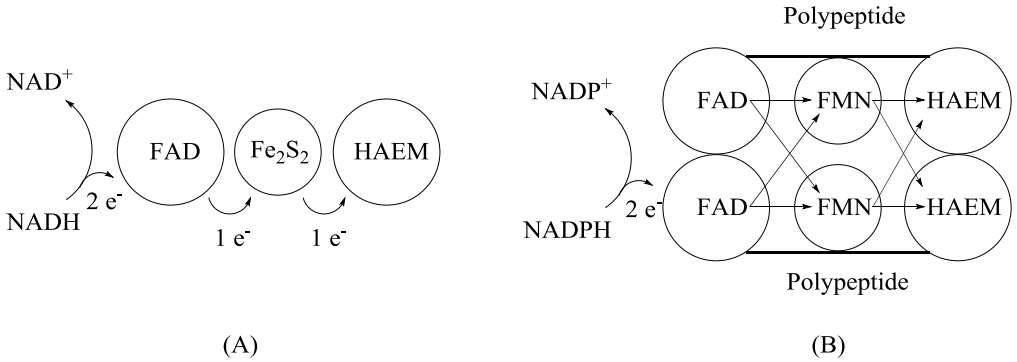


Figure 1.7 Schematic showing electron transfer systems of $P450_{cam}$ and $P450_{BM3}$
(A) PdR/Pdx/$P450_{cam}$ system; **(B)** $P450_{BM3}$ system

$P450_{BM3}$ (CYP102A1), a typical class II system, is in contrast with $P450_{cam}$ (CYP101A1) which is another widely studied enzyme. The schematic electron flow of both systems is shown in Figure 1.7. $P450_{cam}$ utilises PdR/Pdx/$P450_{cam}$ type three-component system and the initial electron acceptor is a FAD-containing ferredoxin reductase which receives two electrons from NADH. The electrons are then passed to the $[Fe_2S_2]$-ferredoxin Pdx which shuttle the electron to the $P450_{cam}$ to perform oxidations.[97]

In $P450_{BM3}$, the haem domain is fused to the reductase domain containing both FAD and FMN binding subdomains. The electrons are passed from NADPH to FAD, then FMN, and finally to the haem iron.[98-100] As both domains are encoded on the same single polypeptide, the electron transfer becomes fast and efficient, which is a major contributing factor to its high activity. Besides its catalytically self-sufficient nature, $P450_{BM3}$ has been proved to function as a dimer, which will be dealt with in detail in section **1.2.3**.

The P450 domain and reductase domain of $P450_{BM3}$ showed little affinity for one another and the holoenzyme may lose its fatty acid hydroxylase activity after treatment with tryptic proteases.[84, 101] These two domains were then expressed and purified separately and it was found that the enzyme activity could not be reconstituted, suggesting the importance of the region linking the two domains.[102-105] Deleting residues from the linker progressively inactivated the enzyme.[106] However, mutagenesis of linker residues may alter the enzyme

activity in certain cases, particularly when glycine and proline were introduced to deter helix formation.[107] It was suggested that the formation of a functional dimer would be impossible when the linker region was shortened, as the haem and the reductase domains might be unable to align correctly, explaining the observed loss of activity.[106]

1.2.3 Dimerisation

In 1977, it was observed by early researchers that the specific fatty acid oxidation activity of $P450_{BM3}$ increases with increasing protein concentration, suggesting that the functional form of the enzyme might be a complex susceptible to dissociation on dilution.[88] The reductive titration with NADPH further supported this hypothesis as it implied the electron transfer could be intermolecular. Sedimentation velocity experiments by Black and co-workers indicated the presence of a mixture of monomers, dimers, trimers and tetramers.[108] On incubation with DTT, the dominant aggregated form was dimeric, with a 1.45-mer also prominent. The 1.45-mer could in principle be a species comprising one full-length enzyme and one P450 domain formed as the result of partial proteolysis.

The two domains (P450 and reductase) were isolated and expressed separately and it was found that they dimerised *via* disulfide bridging.[109] Furthermore, whilst mutating two solvent-exposed cysteine residues, the P450 domain remained monomeric. A mixture of two inactivated $P450_{BM3}$ mutants, one with A264H which inactivated the haem domain while the other included the G570D mutation that inactivated the reductase domain, showed significant levels of hydroxylase activity reconstitution, suggesting that the electron transfer could be intermolecular and the functional form of the protein was dimeric.[110] However, the mechanism for the dimerisation of the full-length protein remains unknown.[109]

1.2.4 Structural studies and key residues

The crystal structure of the haem domain of wild-type (WT) $P450_{BM3}$ in its substrate-free

(SF) form was first published by the Peterson laboratory (PDB code: 2HPD) in 1993[111] while the substrate-bound (SB) form was first solved by by Poulos and co-workers in 1997 (PDB code: 1FAG) with palmitoleate as the substrate.[112] Since then, nearly 40 crystal structures have been reported for $P450_{BM3}$ variants. A variety of computational methods including structure alignment, energy minimisation calculations, molecular dynamics simulations, computer docking, etc., as well as experimental approaches such as Resonance Raman, nuclear magnetic resonance (NMR), magnetic circular dichroism (MCD) and electron paramagnetic resonance (EPR), have been applied to elucidate the mechanism of substrate-enzyme binding behaviour and to identify key amino acid residues that play important roles in determining enzyme activity and selectivity. [103, 113-125]

The $P450_{BM3}$ SF structure showed a long, hydrophobic, largely non-aromatic substrate access channel, which is in contrast to $P450_{cam}$, where no clear substrate channel can be identified from the SF structure.[126] The $P450_{BM3}$ SB structure further supported this hypothesis as the fatty acid substrate extended through the access channel, which is shown in Figure 1.8. Crystal structure analysis also allowed a number of active site residues with potentially important functions to be identified. Most of these have been subjected to site-specific mutagenesis in order to clarify their roles in the catalytic process, most commonly *via* comparison of the fatty acid oxidation rates and product distributions given by the mutants and the WT enzyme.

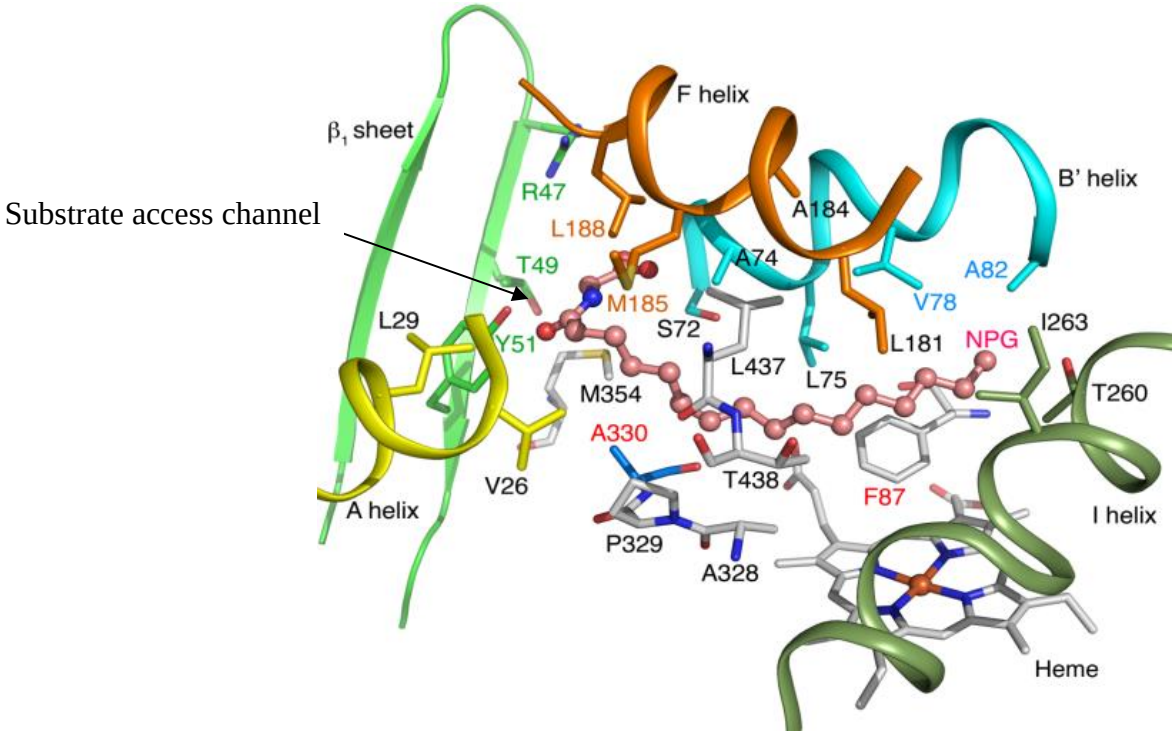


Figure 1.8 The substrate access channel and active site of *N*-palmitoylglycine (NPG)-bound WT $P450_{BM3}$ (key amino acid residues are labelled)

Arg47 and Tyr51, two polar amino acid residues located close to the mouth of substrate channel entrance, were proved to tether carboxylate groups of fatty acid substrates and therefore acting as gatekeepers to the active site by regulating the admission of potential substrates and water.[112, 127]

Phe87, whose side-chain extended into the lumen of the substrate access channel, occupied a prominent position between the haem iron and the fatty acid oxidation sites. Residues with less bulky side-chains such as valine and alanine, are normally induced to help create incremental space in the vicinity of the haem iron and elevate the active-site water population.[112, 127, 128]

Glu267, is in the middle of the I-helix. Its side-chain is typically bonded to at least one water molecule in the SF structure and it was suggested to be a conservative amino acid residue that is responsible for the proton delivery pathways.[129, 130]

Thr268, is one of the most conserved residues among cytochrome P450 superfamily. It has

been suggested by early researchers to play important roles in proton delivery, electron transfer and dioxygen activation.[131-136]

Leu437, has been proposed to act as a "safety catch" for Phe87, and prevents rotation of the Phe87 side-chain, thus keeping the enzyme conformation away from the substrate-bound (SB) form when no substrate is bound.[137, 138]

Other key amino acid residues include Val26, Leu29, Thr49, Ser72, Ala74, Leu75, Val78, Phe81, Ala82, Leu181, Ala184, Leu188, Thr260, Ile263, Ala328, Pro329, etc. and these are also common targets for protein engineering.[110, 124, 139-144]

1.3 Protein engineering

1.3.1 General introduction

Protein engineering is the process of developing useful or valuable proteins. It is a powerful tool for modifying the structure and hence function of an enzyme to allow new catalytic pathways to occur, or to convert non-natural substrates into synthetically useful compounds (such as pharmaceuticals, biofuel, food, and agricultural additives) with a high degree of regio- and enantio-selectivity.

Currently, whenever a certain C−H bond in a given organic molecule needs to be oxidised in a biocatalytic manner, it is best to first test the available wild-type (WT) CYPs.[7-15, 145-148] Unfortunately, this is not routinely successful. Protein engineering can then be invoked in order to develop new P450 variants suitable for biocatalysis, the general approach of which is shown in Figure 1.9.

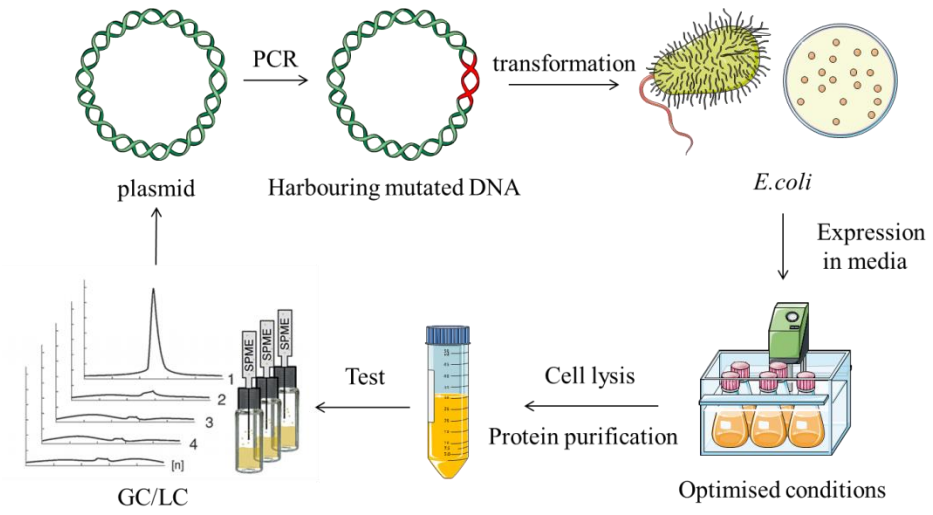


Figure 1.9 General approach of protein engineering

Protein engineering generally starts from the construction of recombinant DNA by molecular cloning or polymerase chain reaction (PCR). The gene encoding target enzyme is introduced onto a cloning vector commonly derived from plasmids or viruses to form the recombinant DNA, which can be genetically modified thus creating new variants. The plasmids harbouring mutated gene can then be transformed into a host organism such as *E. coli* and yeast for protein expression. Proteins with modified amino acid sequence are isolated and purified prior to evaluation. Positive variants are selected and further rounds of protein engineering are often necessary.

1.3.2 Rational design and directed evolution

As creating new variants is the core of protein engineering, two main strategies have emerged to design proteins. Rational design employs site-directed mutagenesis along with protein structural information to rationally design new or improved function. In contrast to this, directed evolution, a non-rational approach, employs recombinant DNA techniques to create numerous variants and is based on high throughput screening to rapidly search for or identify the variant that offers the best solution.[149, 150]

1.3.2.1 Rational design

In rational design, three-dimensional protein structural and computational data are utilised to make reasonable choices for amino acid exchanges around the enzyme's binding pocket. This may require a trial and error strategy involving not just the generation of a single mutant, but also the formation of several likewise 'designed' variants. Because of its root in protein chemistry, rational design was the earliest approach to protein engineering and therefore has been applied to CYPs for a long time, especially by the research groups of Negishi,[151] Poulos,[152, 153] Munro,[154] Urlacher,[155-157] Bernhardt,[158, 159] Vermeulen,[160, 161] Commandeur,[160, 161] and Wong,[162, 163] among others.[7-19]

Successful examples include the L209N mutant of mouse P450-(2*a*-4) designed by Negishi that showed androstenedione hydroxylase activity with high selectivity for the 15α-OH[151] and the F87A/Y96F/L244A/V247L $P450_{cam}$ variant designed within our group that is an excellent catalyst for the regio- and diastereo-selective hydroxylation of (+)-valencene with formation of (+)-*trans*-nootkatol (86% selectivity)[164]. The earliest reports on non-natural substrate oxidation by engineered $P450_{BM3}$ involved mutations at Phe87, Thr268 and Ala328 to oxidise 1,1,2,2-tetrachloroethane and 2-phenylpropanal. The oxidation of polyaromatic hydrocarbons (PAHs) was dramatically enhanced by introducing two space-creating mutations close to the haem iron (F87A and A264G) along with R47L/Y51F to eliminate the polar binding sites at the mouth of the substrate access channel.[162]

This type of research has led to the accumulation of a great deal of information regarding structure-function correlation within the enzyme active site, which can be used as the basis for further protein engineering. However, this method relies on specialised researchers and is always limited by the narrow coverage given by the variants generated by site-directed mutagenesis.

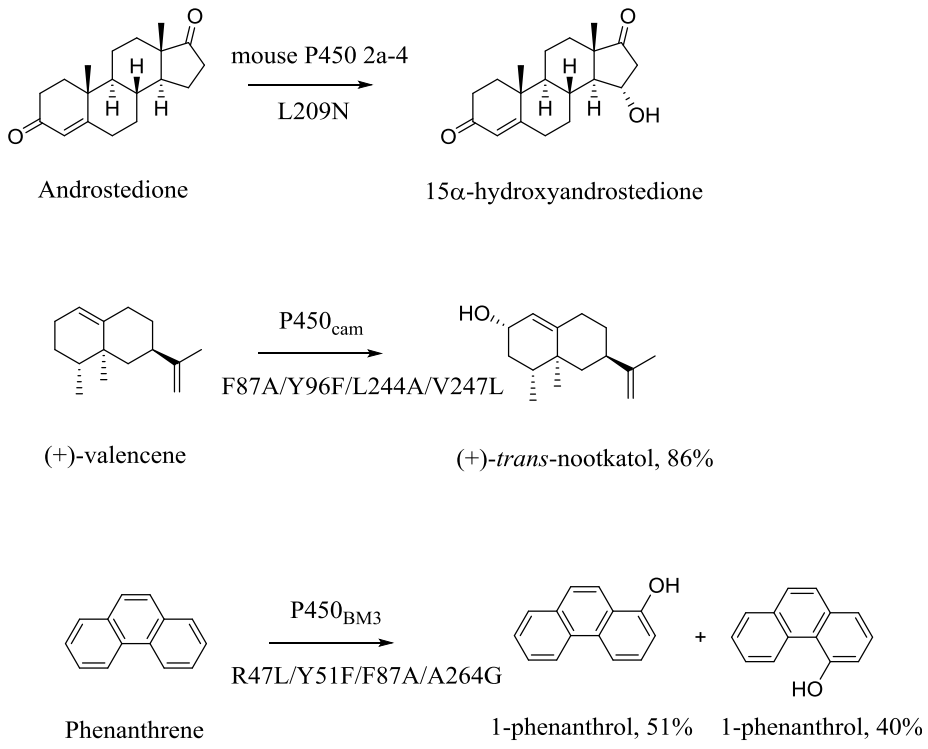


Figure 1.10 Examples of the evolution of CYPs by rational design

1.3.2.2 Directed evolution

Directed evolution, acting as a complementary method to rational design, is a widely applied Darwinian approach based on the generation and high throughput screening of a large number of mutants. The procedure of directed evolution (Figure 1.11) can be divided into three parts: selection of template and sites for mutagenesis, construction of new variants library and high throughput assay or evaluation.

Once a target substrate and a base enzyme which shows some detectable activity towards the target substrate have been selected, a great number of mutants can be developed based on various combinations of pre-determined amino acid residues. A variety of gene mutagenesis methodologies have been developed by molecular biologists in which the most commonly used techniques are error-prone polymerase chain reaction (epPCR) (a shotgun method which address the whole gene/protein and generates random mutants), site-saturation mutagenesis (generation of 20 canonical amino acid residues at predetermined sites), DNA shuffling

(mimic of evolution in nature). DNA degenerate codons are often introduced to decrease the number of mutants and simplify the mutagenesis and screening step. Random mutagenesis was first applied by Sligar to a cytochrome P450 in 1997 which led to the altered regioselectivity in fatty acid hydroxylation, but the enantioselectivity of the reaction was not reported.[165]

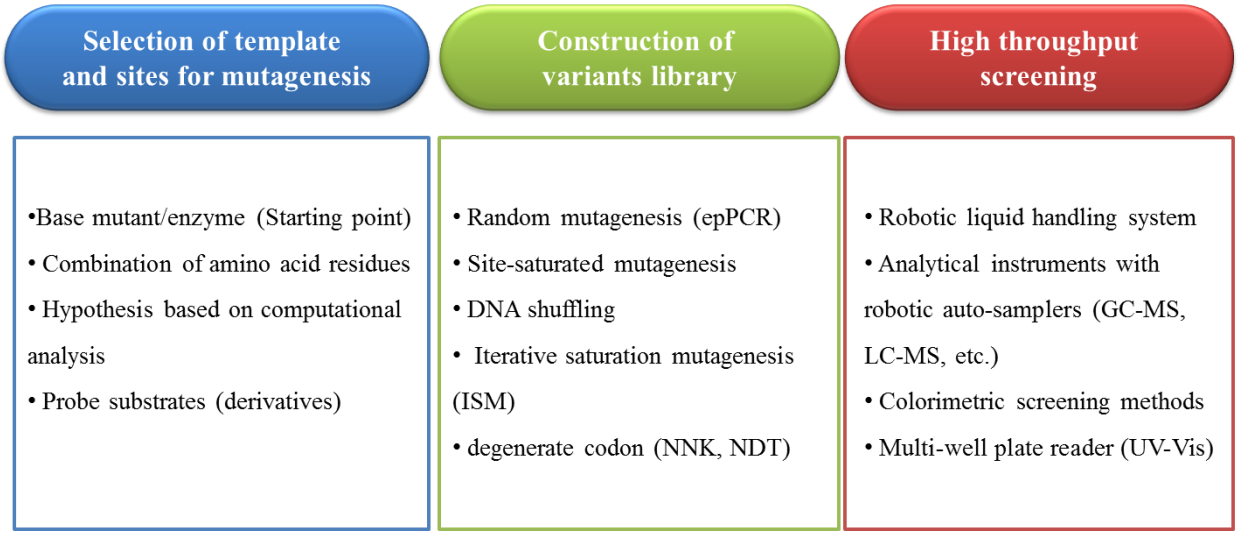


Figure 1.11 General procedure of directed evolution

Since screening is often the bottleneck of directed evolution, much work has been done to establish high-throughput screening assays. Automated chromatography instruments (GC and HPLC) with robotic liquid handling systems taking samples directly from the multi-well microtiter plates have emerged as the preferred technique which may allow 500−1000 reactions to be analysed in one day. Multi-well UV/Vis plate readers present another option as it monitors NADPH consumption rate of reactions in the microtiter plates. However, UV/Vis absorption does not provide information on product profiles. Other screening systems have also been developed, notably the colorimetric screening system which generates coloured compounds to identify active enzymes (Figure 1.12).[21, 166-169]

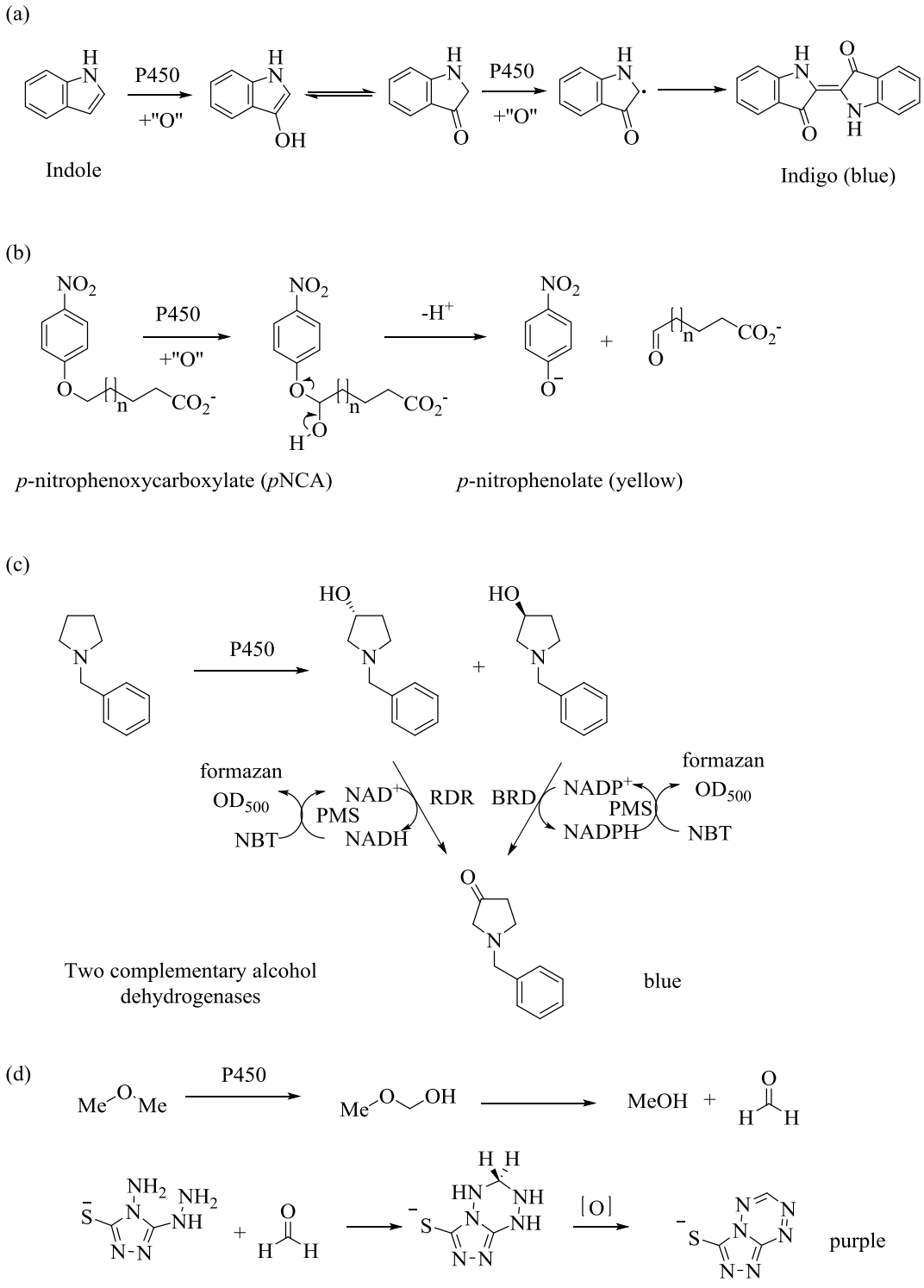


Figure 1.12 Colorimetric assays for high throughput screening

Due to the Darwinian character of this method, directed evolution is currently the most reliable approach to protein engineering and is widely applied by the research groups of Arnold,[170-173] Reetz,[174-180] Schmid,[168] Zhao,[181] Li,[182-185] Fasan[186-189] and others. In a

pioneering study, Schmid and co-workers identified a small number of indigo-producing $P450_{BM3}$ mutants in which three specific mutations (A74G, F87V, L188Q) were found. The A74G/F87V/L188Q mutant was then created and found to be active towards a broad range of organic molecules including fatty acids, polyaromatic hydrocarbons, chlorinated dioxins.[168] In a more systematic approach, Arnold and co-workers applied directed evolution to $P450_{BM3}$ in the oxidation of medium-chain alkanes.[190, 191] Successive rounds of screening yielded an 11-mutation variant, 139-3[191], which showed enhanced activity towards medium- to short-chain alkanes including propane and butane. Further protein engineering resulted in PMO, a variant containing 24 mutations, the most active propane oxidising P450 mutant in the literature, with 99% coupling efficiency and a total turnover number of up to 45,800 from *in vivo* biotransformation.[170] In order to minimise the screening effort, Combinatorial Active-Site Saturation Test (CAST) was put forward by Reetz and co-workers, in which active site residues were classified in subgroups base on the X-ray crystal structure data and homology models. NDT codon degeneracy encoding 12 instead of the normal 20 canonical amino acids was applied to construct variants libraries. These efforts generated over 8,700 transformants of $P450_{BM3}$ that led to the discovery of 2β- and 15β-testosterone hydroxylase activity with over 95% regioselectivity. [192]

CAST/ISM (iterative saturation mutagenesis) strategy was then applied onto the allylic oxidation of methyl cyclohex-1-ene-1-carboxylate. Randomisation was performed with 24 amino acid residues while pre-screening was done by the traditional UV/Vis assay monitoring NADPH consumption rate followed by an automated GC analysis. Some highly (*S*)- and (*R*)-selective variants were identified after 4,600 transformations, which are shown in Figure 1.10.[175] Fasan and co-workers subjected $P450_{BM3}$ to directed evolution to oxidise the anti-malaria therapeutic drug artemisinin. Five semisynthetic chromogenic probes were used as fingerprint probes for first step screening and 522 variants were identified as functional thus

further screened for artemisinin oxidation. One mutant (IV-H4) showed complete selectivity for C7 (*S*)-hydroxylation while another (II-H10) was fully selective towards C7 (*R*)-hydroxylation product. The screening of another 3,000 transformations provided variant X-E12 which showed 94% selectivity for the C6 hydroxylation product.[187]

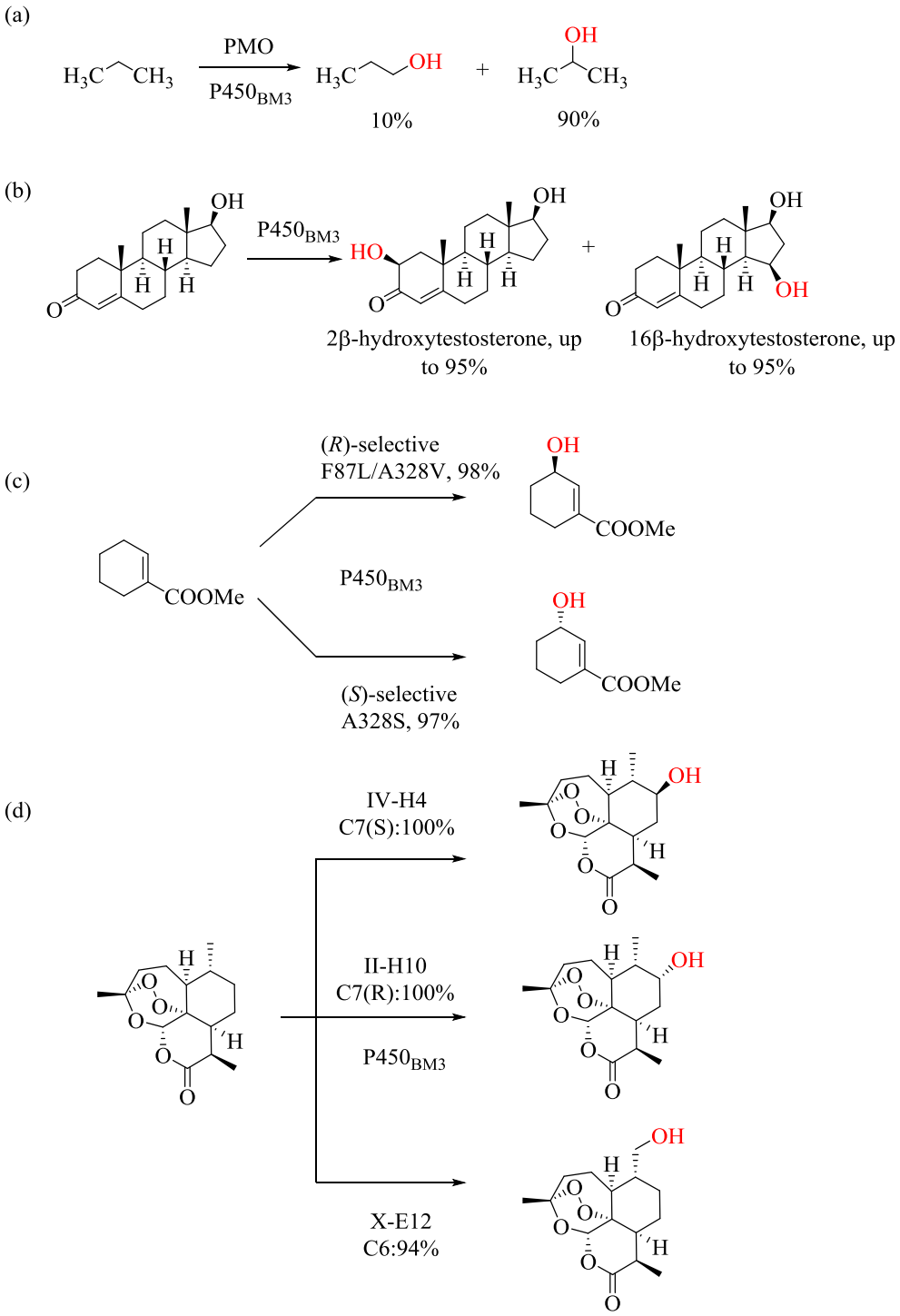


Figure 1.13 Successful examples of CYP engineering by directed evolution

Compared to rational design, directed evolution delivers a far larger number of mutants that covers more sequence space, in which the specific substrates can be oxidised selectively. This method doesn't need specialised researchers and thus can be performed by robotic liquid handling systems. However, the bottleneck of this method is the screening throughput and the selection of starting templates.

1.3.3 Other strategies

(1) Chimeragenesis and protein fusion

Since most protein domains that confer catalytic or other functions can fold independently and are easy to manipulate, creating hybrid proteins that are fusions of pre-existing but unrelated protein domains has now become a standard tool for protein engineering. In an imaginative approach, short sections of amino acids sequence between two key substrate recognition regions of $P450_{BM3}$ (Substrate Recognition Site 1 and Substrate Recognition Site 3) were replaced by those from CYP4C7, a farnesol ω-hydroxylase from *Diploptera punctate.* The recombinant showed an enhancement of enzyme activity by six-fold.[81, 193] Chimeras comprising sections of CYP102A1, A2 and A3 have been reported. It was found that the chimeras oxidised verapamil and astemizole as substrates which were not found in any of their parent proteins.[194-196]

(2) *De novo* design

The ultimate dream in protein engineering is to be able to create any functional protein by simply editing and specifying the amino acid sequence. This is not a completely impossible goal as it is well known that the protein structure is determined by its primary amino acid sequence. However, the mechanism on how amino acids sequences fold into their three dimensional structures is not fully understood although the formation of simple types of secondary structural elements such as α-helix or β-sheet is reasonably well understood. The

ways to associate these simple elements, through hydrophobic interactions, disulphide bonds, and the coordination of metal ions are also reasonably clear. However, due to the complexity of protein sequence-structure-function relationship, examples of *de novo* design of CYP are rare since much of the current effort attempts to combine known secondary structural elements to yield functional proteins.[197]

1.4 The engineering of $P450_{BM3}$ for non-natural substrates oxidation

P450s are attractive enzyme platforms for oxidation and synthesis in biotechnology due to their remarkable C−H activation function and have drawn increasing attention on the commercial applications for specific non-natural substrate categories. $P450_{BM3}$, in particular, has been subjected to protein engineering to oxidise a broad range of non-natural substrates including alkanes, drugs, terpenes, polycyclic aromatic hydrocarbons. New function studies were also performed by researchers to incorporate new activities and therefore expanding the repetoire of $P450_{BM3}$.

1.4.1 Alkane oxidation

Catalytic hydrocarbon oxidation under mild conditions is a highly desirable reaction in petroleum industry, most notably the oxidation of methane. The transformation of this greenhouse gas to methanol will prevent its release and also produce readily transportable liquid feedstock of fuels. The oxidation of aliphatic C–H bonds in alkanes is synthetically difficult due to their high dissociation energy (430 ~ 530 kJ/mol), often requiring heterogeneous catalysts that employ high temperatures and high pressures. Enzymatic systems are therefore attractive alternatives since enzymatic reactions are generally selective, easy to conduct, employ mild conditions and require low energy input.

Enzymes capable of oxidising alkanes are existed in nature, notably the membrane-bound

non-haem terminal alkane hydroxylases from *Rhodococcus rhodochrous* and *Acinetobater calcoaceticus* AlkB and methane monooxygenase (MMO).[198] Certain P450s have also been found to oxidise C_8 to C_{16} alkanes,[199] examples being the CYP52 family[200], however, there are no P450s found to oxidise short-chain gaseous alkanes.

$P450_{BM3}$, a catalytically self-sufficient enzyme, has been engineered as an alkane hydroxylase using a variety of strategies. WT $P450_{BM3}$ has been shown to be active towards hexane and octane while the GVQ (A74G/F87V/L188Q) mutant was found from a colorimetric screen which displays higher activity.[168] Active variants towards propane were achieved through a directed evolution study as mentioned earlier in this thesis. Ethane oxidation has also been reported however the oxidation rate was three orders of magnitude slower than propane oxidation suggesting the difficulty in oxidising small molecules with P450 enzymes.[201]

In order to reduce the enzyme active site volume, an inert co-substrate (perfluorocarboxylic acid) was introduced to fill the enzyme active site by Watanabe and co-workers, which led to higher ethane oxidation rate. Methane oxidation was then reported by Reetz and co-workers using a variety of perfluro fatty acids however the process turned out to be problematic after which the paper was retracted by the authors.[202] Methane oxidation still remains challenging to P450s while the only known valid study of methane oxidation is due to Arnold and co-workers, who reported a turnover number (TON) value of 0.05.[203]

Branched alkanes and cyclohexanes have also been targeted for oxidation by engineered P450s, which may have potential applications in fine chemical synthesis.[156, 190, 204, 205]

1.4.2 Drug oxidation

For the new drugs to be approved to be sold on the market, they have to be tested clinically to be non-toxic and safe. Since human P450s are responsible for over 75% of drug metabolism and clearing in humans and they are difficult to express due to the low stability, highly

variable expression levels and high cost. Microbial P450s were explored as surrogate to produce drug metabolites so that the metabolites can be screened for their efficacy, toxicity and bioavailability. Figure 1.14 shows a panel of drug molecules being successfully oxidised by $P450_{BM3}$.[56, 129, 168, 172, 196, 206-217] However, substrate conversions and turnover numbers were not sufficient for preparative scale metabolites synthesis and full products characterisation.

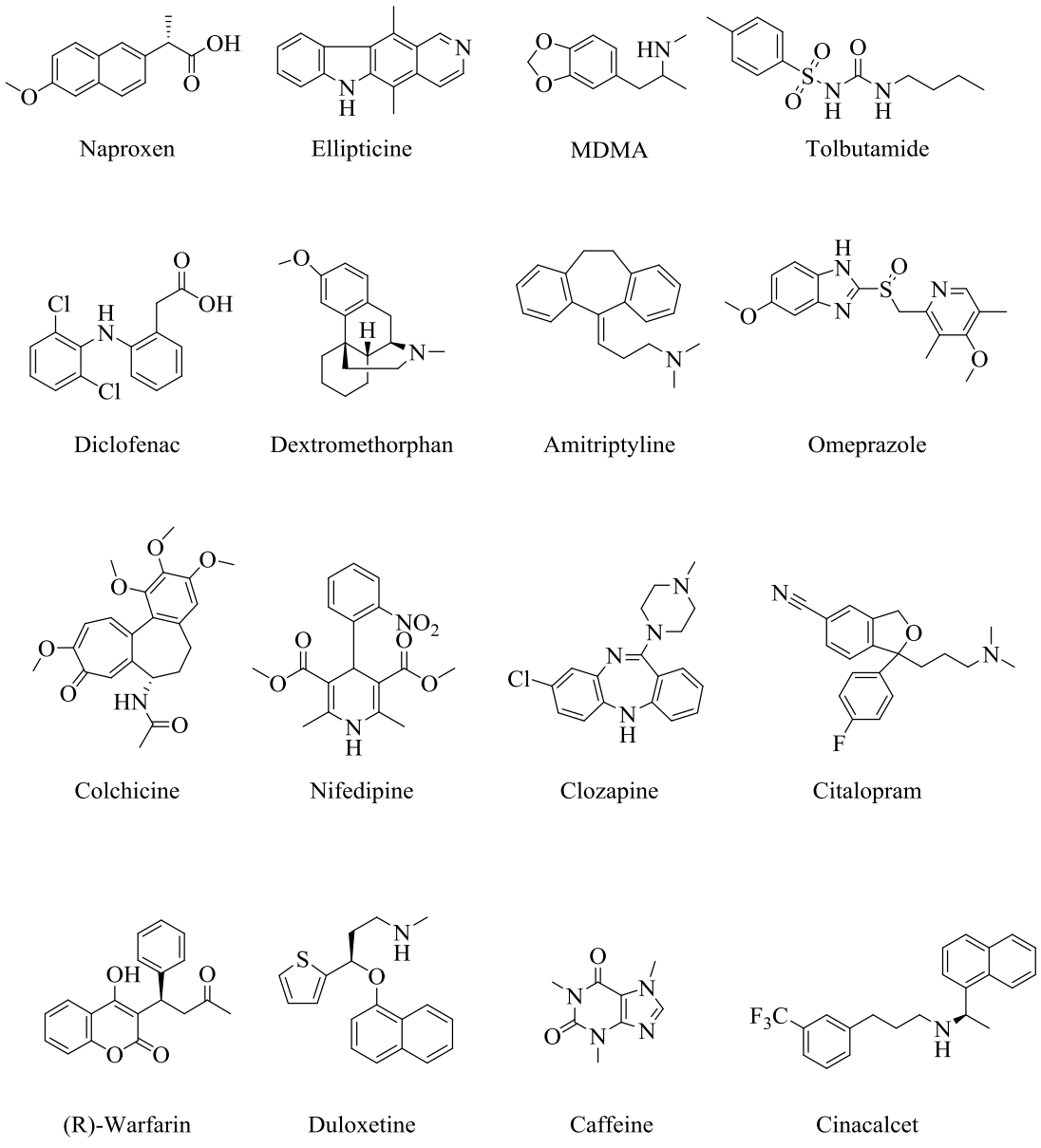


Figure 1.14 Examples of drug molecules oxidised by $P450_{BM3}$ variants

1.4.3 Steroid oxidation

Steroids and their oxygenated derivatives are important compounds in medical and pharmacological applications since the steroid hormones play vital roles in human physiology, with body hormone levels regulating and supporting reproduction, metabolism and immune system, and also promoting muscle and bone synthesis. There has been a widespread interest in synthesising oxygenated steroids compounds, both chemically and biologically.

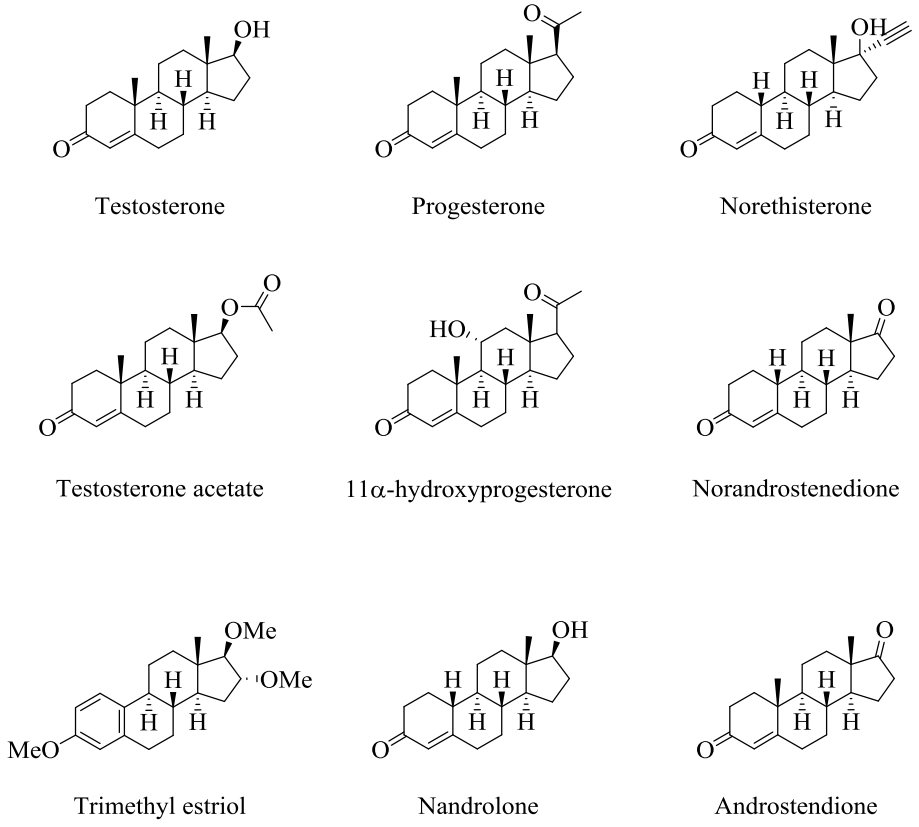


Figure 1.15 Examples of steroids oxidised by $P450_{BM3}$ variants

$P450_{BM3}$ has been engineered towards selective oxidation of a panel of steroids compounds, as shown in Figure 1.15.[161, 172, 192, 210, 218, 219] Since steroids are relatively large molecules for the enzyme active site, Arnold and co-workers introduced combinatorial space-creating alanine substitutions and identified a 17-mutation $P450_{BM3}$ variant F11 which showed moderate activity towards trimethyl estriol, 11α-hydroxyprogesterone and testosterone

acetate.[172] Commandeur and co-workers obtained a triple mutant (R47L/F87V/L188Q) by performing random mutagenesis and this mutant was found to hydroxylate testosterone, norethisterone, nandrolone, progesterone, and androstendione.[210] Vermeulen and co-workers applied site saturation mutagenesis at position 87 and 82 and found that F87 was critical in controlling selectivity of testosterone oxidation while A82W mutation led to a 42-fold increase in oxidation rate for 16β-hydroxylation of both testosterone and norethisterone.[161] Reetz. and co-workers applied a directed evolution study using an ISM/CAST strategy and revealed that A330W/F87A mutant gave highest 16β-regioselectivity at 91% towards progesterone, while A82N/F87A showed 100% 2β-regioselectivity.[192]

1.4.4 Terpene and terpenoid oxidation

Terpenes and terpenoids, being the primary constituents of the essential oils of many types of plants and flowers, are widespread in nature and are widely used as flavours and fragrances in perfumery, cosmetics, foods and wine additives. Synthetic variations and oxygenated derivatives of natural terpenes and terpenoids also greatly expand the variety of flavours and fragrances compounds. For example, (+)-nootkatone and (+)-nootkatol (*trans*- and *cis*-) are derived from valencene oxidation.[164] Santalol, the fragrance compound sandalwood oil, can be synthesised by selectively oxidising santalene.[220] β- and δ-damascone are potential precursors to damascenone, an important rose ketone that finds applications in many fields. Therefore, terpenes and terpenoids have been targeted for selective oxidation by evolved $P450_{BM3}$ enzymes. Some examples are collected in Figure 1.16.[163, 164, 205, 221-237]

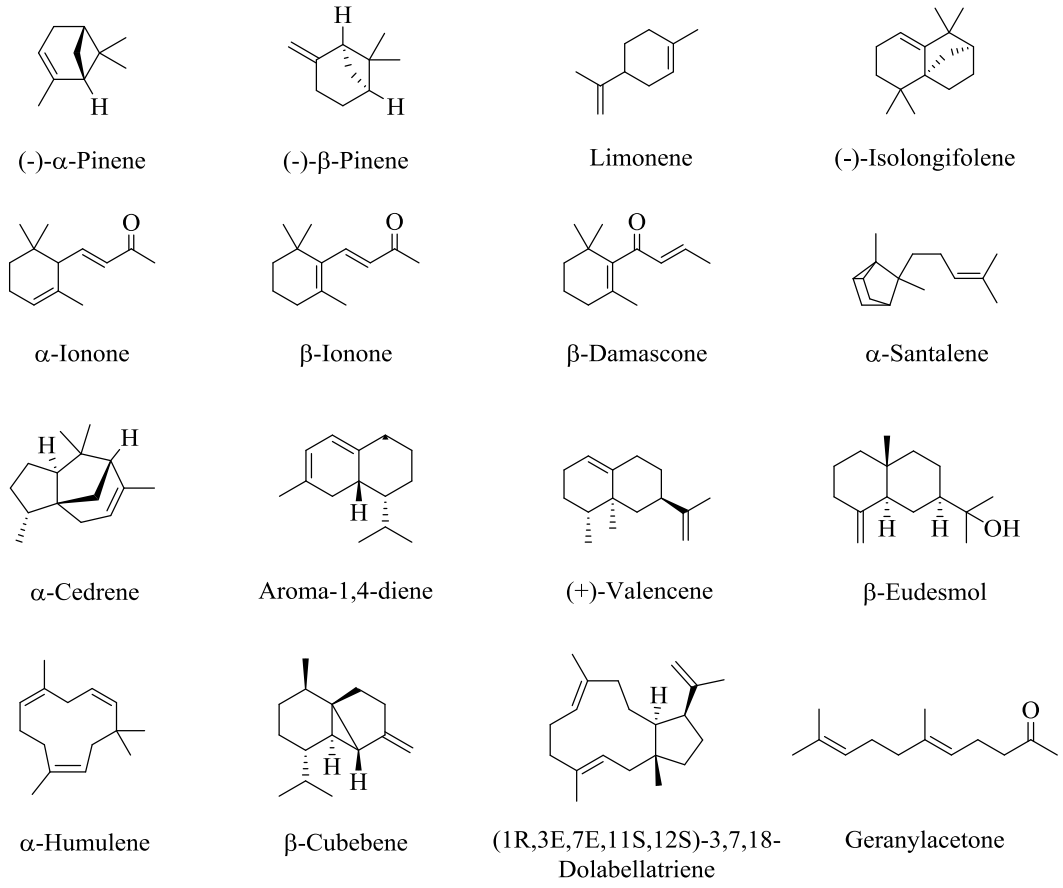


Figure 1.16 Examples of terpenes and terpenoids oxidised by $P450_{BM3}$ variants

1.4.5 Oxidation of aromatics

Aromatic compounds are basic building blocks for a wide range of molecules ranging from pharmaceuticals, detergents, adhensives, petrochemicals, and environmental pollutants. Figure 1.17 shows some aromatic substrates that have been targeted by evolved $P450_{BM3}$ mutants including alkylbenzenes and styrenes,[21, 31, 40, 138, 163, 205, 222, 238-241] phenols and ethers,[194, 207, 211, 212, 238, 242-246] esters and benzoic acids,[108, 171, 211, 238, 242, 247, 248] polyaromatic hydrocarbons (PAHs),[21, 207, 221, 238, 242, 249, 250] and polychlorinated dioxins.[141, 251, 252] The direct hydroxylation of aromatic compounds may provide synthetic handles for structure modification and elaboration and is thus becoming an important methodology in modern synthesis. On the other hand, the monooxygenation of environmental pollutants such as polyaromatic hydrocarbons and polychlorinated dioxins is a potential method of remediation

as P450s can be engineered to include multiple reactions in pollutant degradation such as hydroxylation at an unsubstituted position, hydroxylation with migration of a chloride substituent, hydroxylation with elimination of a chloride substituent, and opening of the dioxin ring, etc.

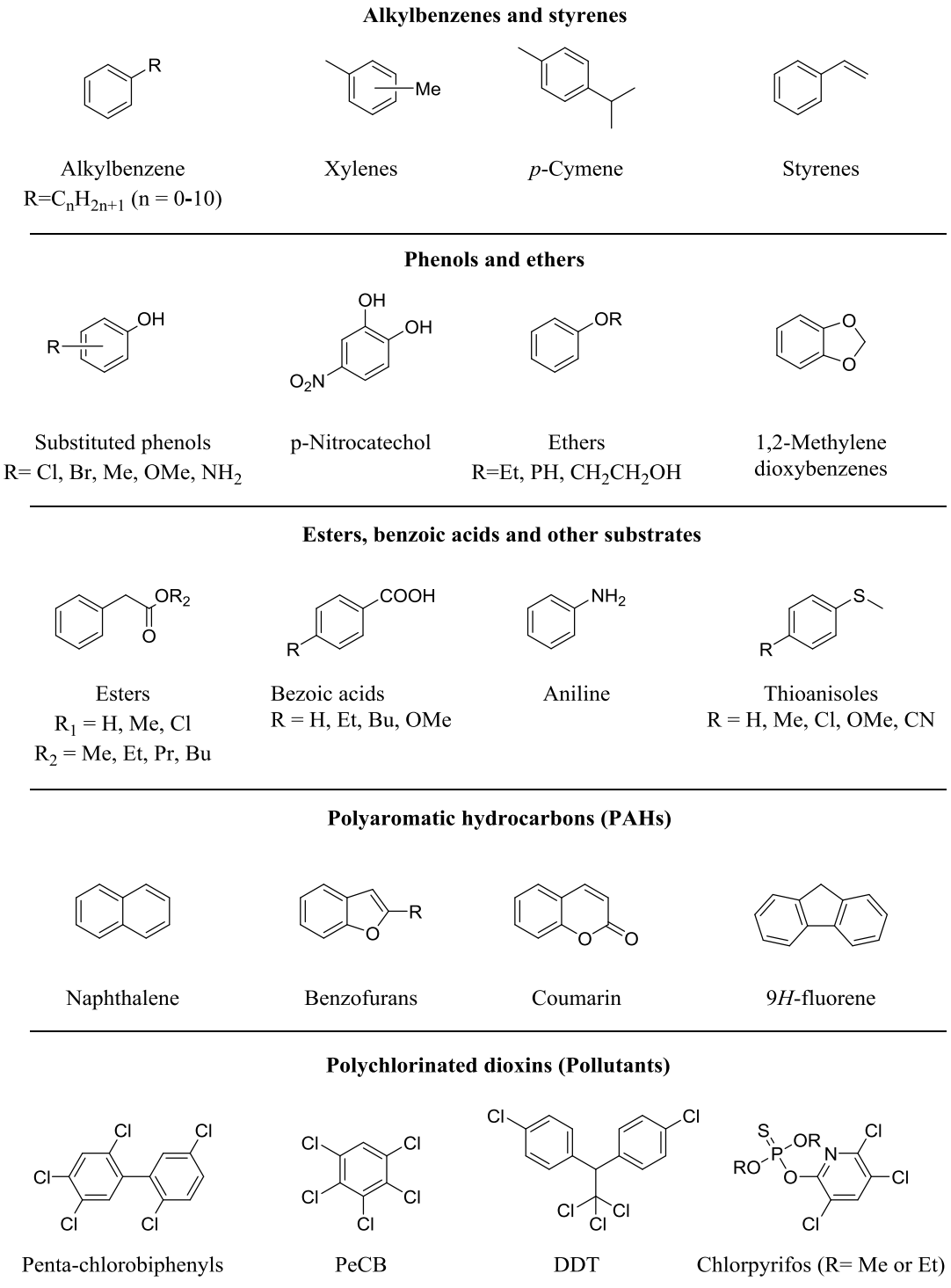


Figure 1.17 Examples of aromatics targeted by $P450_{BM3}$

1.4.6 New function studies

Expanding the enzymes' repertoire to incorporate important synthetic reactions that are not found in nature will open up new opportunities for biocatalysis. It has been demonstrated that $P450_{BM3}$ can be engineered as an iron(II) catalyst to catalyse a variety of carbene and nitrene transfer reactions, in which synthetic reagents are used to generate new reactive intermediates. The cyclopropanation,[37-39, 253-256] N−H insertion,[257] C−H amination,[258] sulfimidation,[259, 260] and aziridination[261] reactions are well known in synthetic chemistry but have no counterparts in nature (as shown in Figure 1.18). Recently, it has been demonstrated that the highly conversed cysteine residue plays a key role in controlling natural and non-natural activities of the enzymes.[262-265] The mutation to either serine or histidine in $P450_{BM3}$ was shown to enhance carbene- and nitrene-transfer activities dramatically and therefore can be used as starting points for the investigations of non-natural P450 activities and evolving highly selective enzymes.

Compared to transition metal catalysts, the enzymes are fully genetically encoded and function under mild conditions. Furthermore, the enzymes can be optimised to exhibit high regio- and stereo- selectivity that may be difficult for the small molecule transition metal catalyst to achieve. By assembling these evolved enzymes, highly selective enzyme cascades can be established to expand the scope of metabolic engineering and offer alternative routes for the synthesis of fine chemicals and natural products.

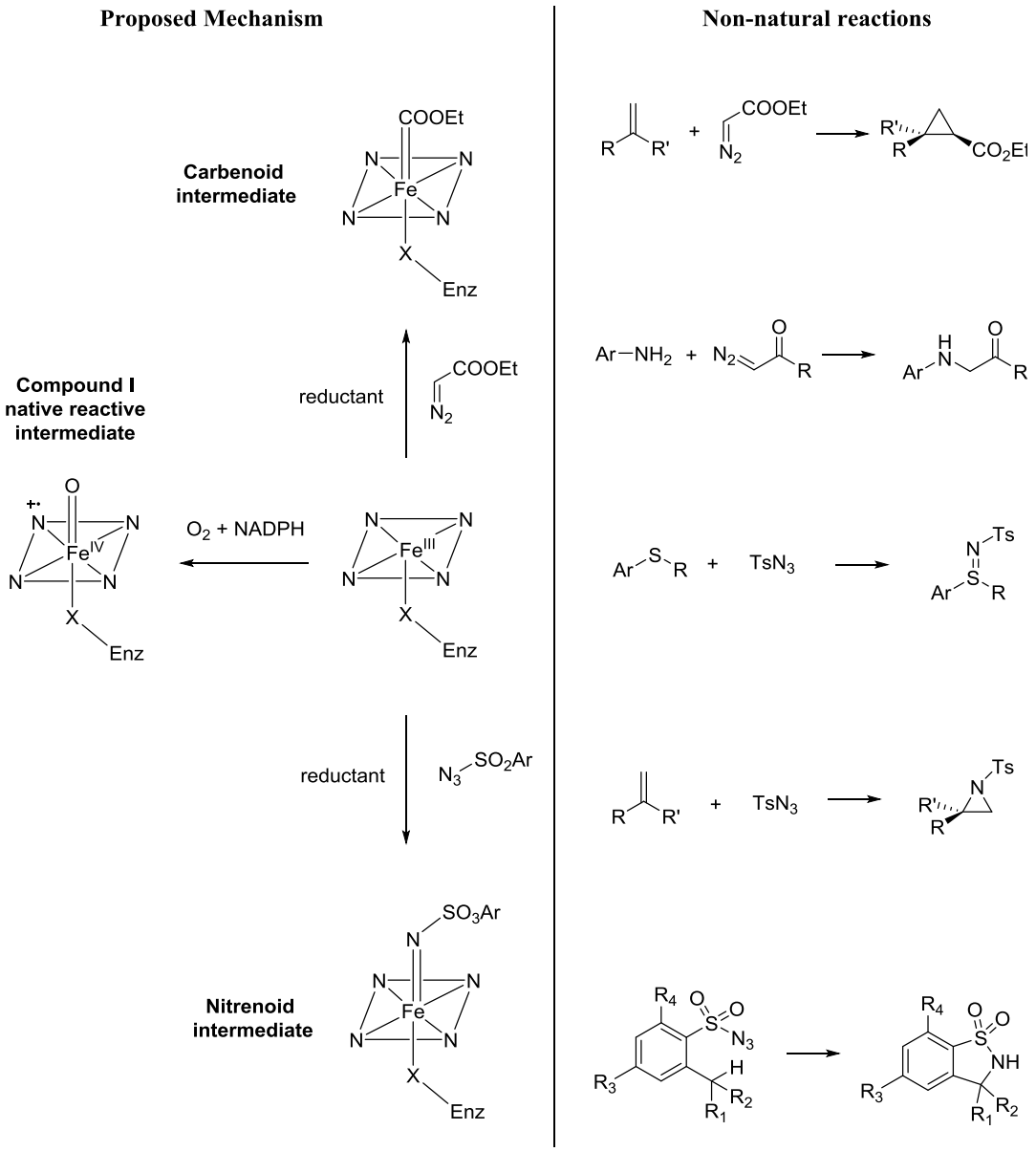


Figure 1.18 Non-natural reactions catalysed by $P450_{BM3}$ and their proposed mechanisms

1.5 Enzymatic functionalisation of C–H bonds

Enzymatic functionalisation of C–H bonds offers significant economic and environmental benefits over traditional synthetic methods. Researchers have discovered a variety of oxidoreductases such as monooxygenases, dioxygenases, oxidases, peroxidases, etc. for such purposes. The reactions catalysed by oxidoreductases are numerous and comprise hydroxylation of aromatic or non-activated carbon atoms, oxidation of primary and secondary amines, heteroatom oxygenation, Baeyer-Villiger oxidation, and double bond epoxidation, etc.[12] The electron acceptors, that is, the compounds that oxidise the substrate of interest, are as diverse as $NAD(P)^+$, cytochromes, molecular dioxygen, hydrogen peroxide, disulfides, quinone or similar compounds, nitrogenous groups, iron-sulfur proteins, and flavin. Since the reactivity of these enzymes is largely dependent on the type of cofactor they employ, this will serve as a convenient way to group enzymes while monoamine oxidases will be discussed in a single section due to its desaturation reaction type and its intrinsic property of activating C–H bonds at α position of an amine.[266-268]

1.5.1 Flavin-dependent enzymes

Flavins are derivatives of riboflavin (vitamin B12) and consist of a tricyclic isoalloxazine moiety. Flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) are two common enzymatic redox-active cofactors derived from riboflavin *via* phosphorylation (FMN) and subsequent adenylation (FAD) at the N10 position (Figure 1.19).

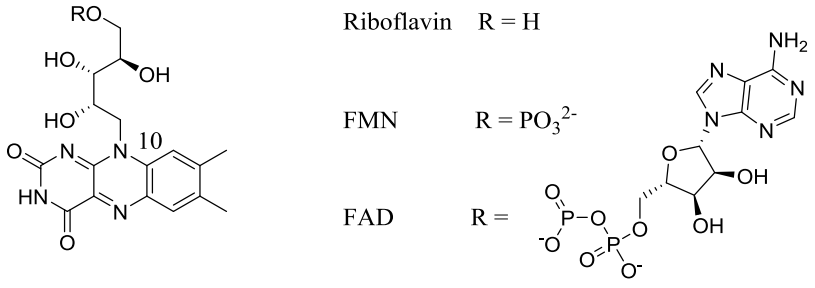


Figure 1.19 Structures of flavin nucleotides

The chemically active portion of flavins is the heterocyclic isoalloxazine ring system, which have three stable oxidation states: fully oxidized, semiquinone (1-electron reduced), and hydroquinone (2-electron reduced) (Figure 1.20). Flavin nucleotides can undergo either one-electron transfer to generate the semiquinone or two-electron reduction into hydroquinone and therefore serving as an important bridge between polar and radical reactions.

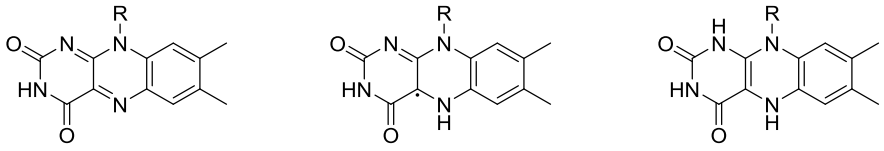


Figure 1.20 Chemical states of flavins

Flavins catalyse a wide variety of chemical transformations, and thus are involved in a host of biological processes: energy production, natural product synthesis, biodegradation, DNA repair, apoptosis, protein folding, xenobiotic detoxification, etc. Thousands of flavoenzymes have been identified and characterized. Current estimates are that 1–3% of bacterial and eukaryotic genomes encode proteins that bind flavin.[269] Moreover, flavoenzymes have been found to possess the ability to hydroxylate C–H bonds in electron-rich (hetero)arenes. The mechanism of this hydroxylation process is shown in Figure 1.21.[270]

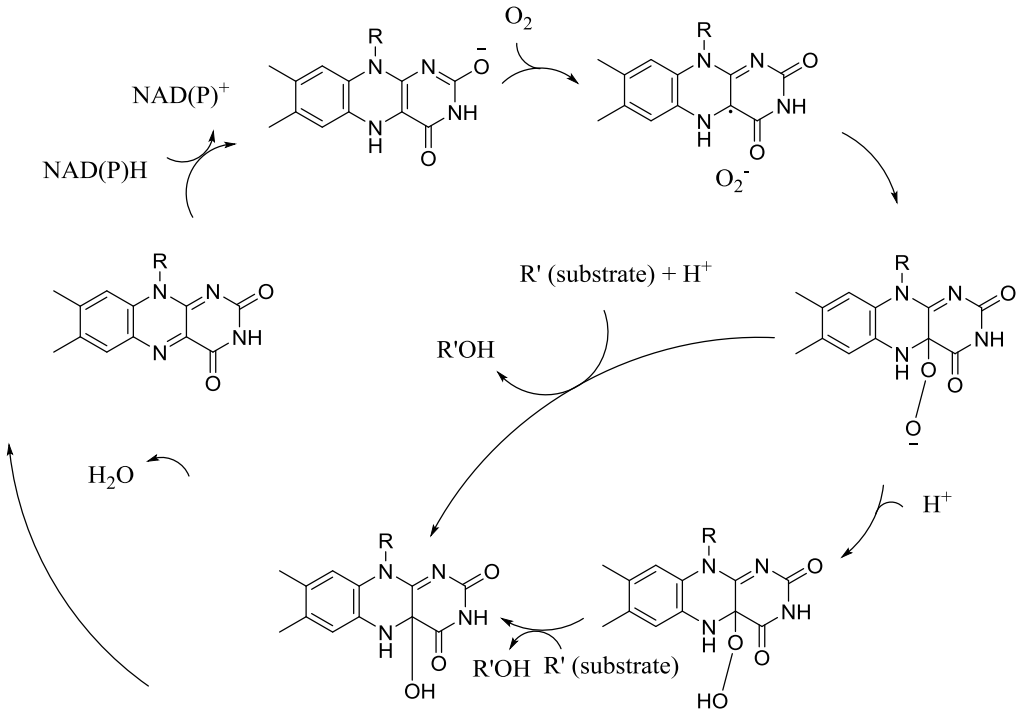


Figure 1.21 Mechannisms of flavin monooxygenases

Flavin monooxgenases have been extensively investigated as biocatalysts due to the importance of aromatic hydroxylation and the difficulties associated with chemical methods for this reaction. Wubbolts and co-workers purified 2-hydroxybiphenyl 3-monooxygenase (HbpA) from *Pseudomonas azelaica* that can be used to produce 3-substituted alcohols from 2-alkyl-, aryl-, or halo-substituted phenols.[271] Regioselective hydroxylation was also achieved with *p*-hydroxybenzoate 3-hydroxylase from *Rhodococcus rhodnii* 135 and *Rhodococcus opacus* 557 by van Berkel and co-workers,[272] as shown in Figure 1.22.

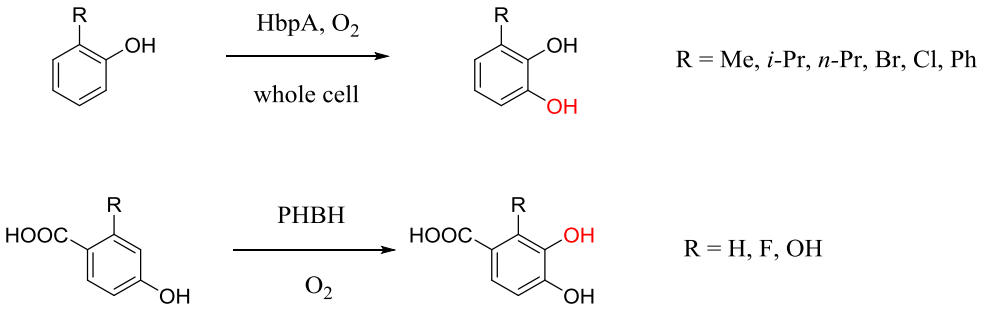


Figure 1.22 Synthesis of substituted phenols using flavin monooxygenases

A variation of the monooxygenase mechanism is used by FAD-dependent halogenases that catalyse the halogenation of electron-rich arenes found in a wide range of antimicrobial natural products. Among the most extensively studied halogenases, several tryptophan halogenases have been characterised. Two tryptophan 7-halogenases RebH and PrnA are shown in Figure 1.23.[273]

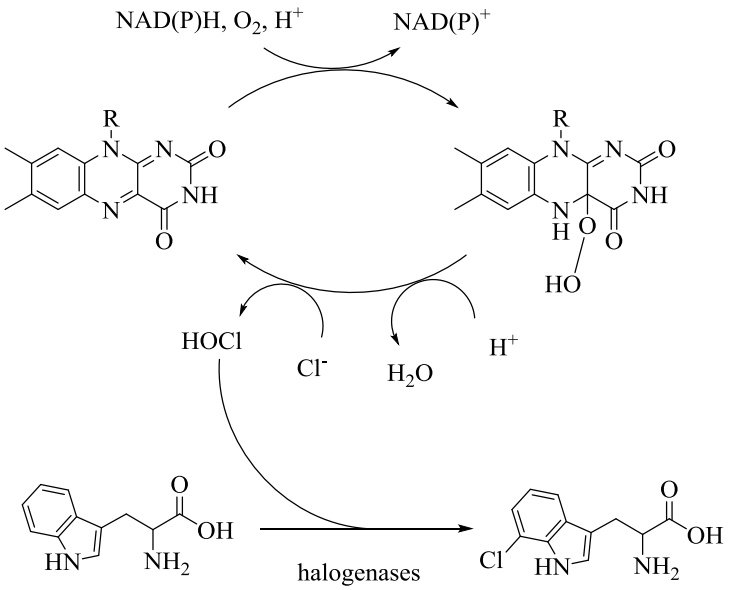


Figure 1.23 FAD-dependent halogenases (RebH/PrnA) catalysing halogenation of tryptophan

1.5.2 Haem and haem enzymes

Haem enzymes are of great research interest due to its capability of functionalising C–H bonds. Cytochrome P450s, the subjects of this thesis, utilize Fe(IV)-oxo radical cation (known as compound I) as reactive species derived from molecular dioxygen activation to effect hydroxylation. A variety of examples have been demonstrated earlier in this chapter. Several other haem enzyme subfamilies also utilize the same reactive intermediate compound I to function. Peroxidases catalyse oxidative coupling reactions *via* sequential one-electron abstractions while haloperoxidases catalyse the halogenation of sp^2 C–H bonds *via* halide oxidation (Figure 1.24).[274, 275]

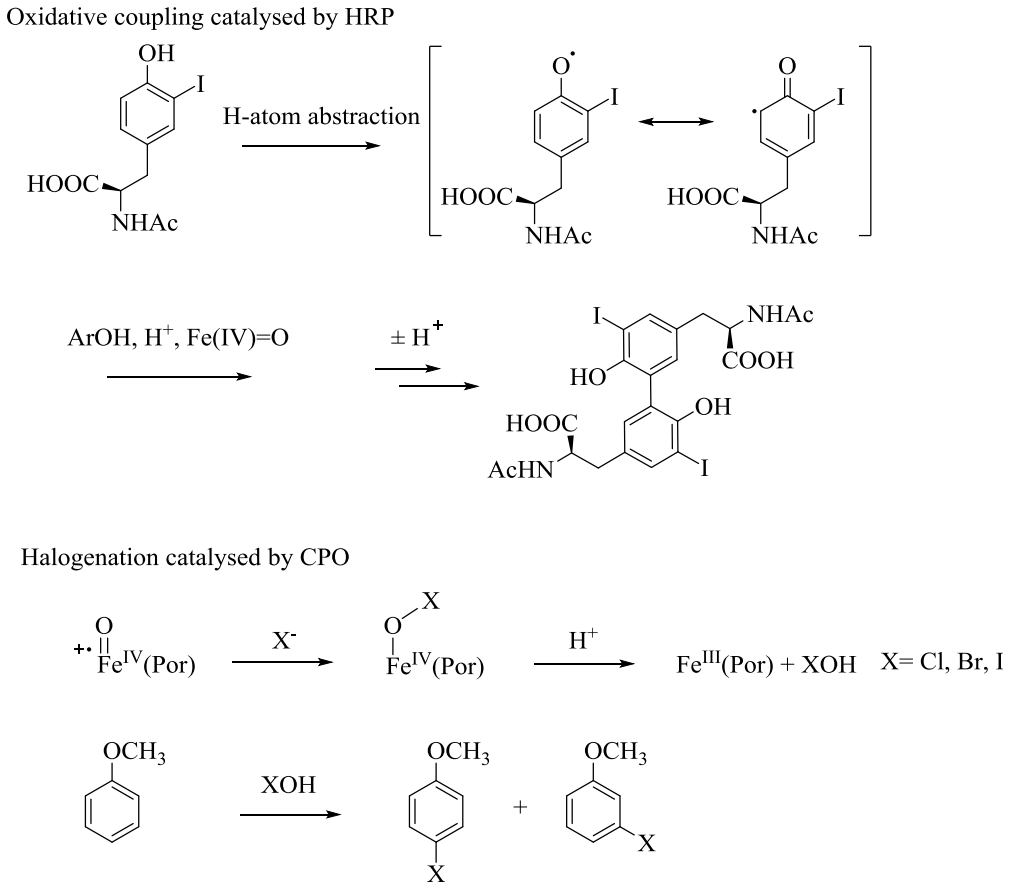


Figure 1.24 Oxidative coupling catalysed by horseradish peroxidase (HRP) and halogenation catalysed by chloroperoxidase (CPO)

1.5.3 Non-haem metalloenzymes

Non-haem metalloenzymes catalyse a wide range of O_2 reactions, substantially paralleling those of haem systems. The active sites of these enzymes range from residue-coordinated metal ions, most notably iron and copper, to enzyme-bound coordination complexes. Based on the number of metal ions coordinated in the active site, these enzymes will be discussed as mononuclear non-haem metalloenzymes and binuclear non-haem metalloenzymes.

a) Mononuclear non-haem metalloenzymes

Mononuclear non-haem iron enzymes are the most prevalent among this subfamily, which utilize as few as two active site residues to coordinate an Fe(II) ion and thus have four sites available to coordinate substrate. In the resting state Fe(II) form of the enzyme, these sites are generally occupied by water or protein ligands. Substrate binding triggers the release of water ligands and therefore offering a vacant coordination site for dioxygen to be bound and generating high-valent iron-oxo intermediates. These intermediates are capable of promoting a range of C–H functionalisation transformations such as hydroxylation, halogenation, desaturation, cyclisation, etc. Some examples are shown in Figure 1.25.[276, 277]

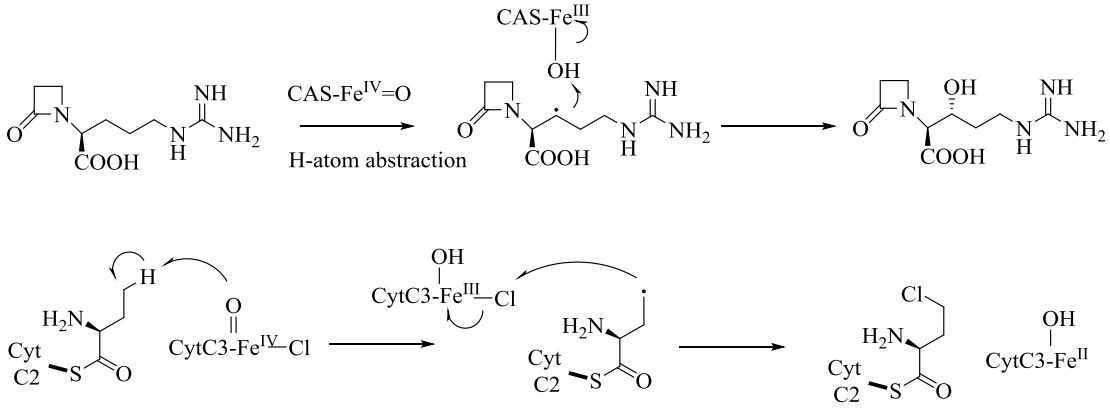


Figure 1.25 Hydroxylation and chlorination reactions catalysed by non-haem iron enzymes

b) Binuclear non-haem metalloenzymes

Bimetallic enzymes, particularly diiron center metalloenzymes and binuclear copper metalloenzymes, have been discovered and extensively studied by early researchers. Bacterial multicomponent monooxygenases (BMMs) constitute a class of the former that utilize residue-ligated iron centers, specifically carboxylate-bridged diiron centers, for hydroxylation of unactivated C–H bonds. Several classes of BMMs have been identified including methane monooxygenase (MMO), alkene monooxygenase (AMO), phenol hydroxylase, and arene monooxygenase. The mechanism of MMO from the methanotroph *Methylococcus capsulatus*

has been studied extensively. The resting state consists of an oxidized dirron (III) species (ox) where the positive charge of the two iron atoms is precisely balanced by four glutamate and two bridging hydroxide ligands. Reduction of the iron (III) centers to the diiron (II) form is required prior to dioxygen activation. A peroxo intermediate is then generated, leading to the fully oxidised diiron (IV) species, which is the proposed active hydroxylating agent in the BMMs. (Figure 1.26)[278]

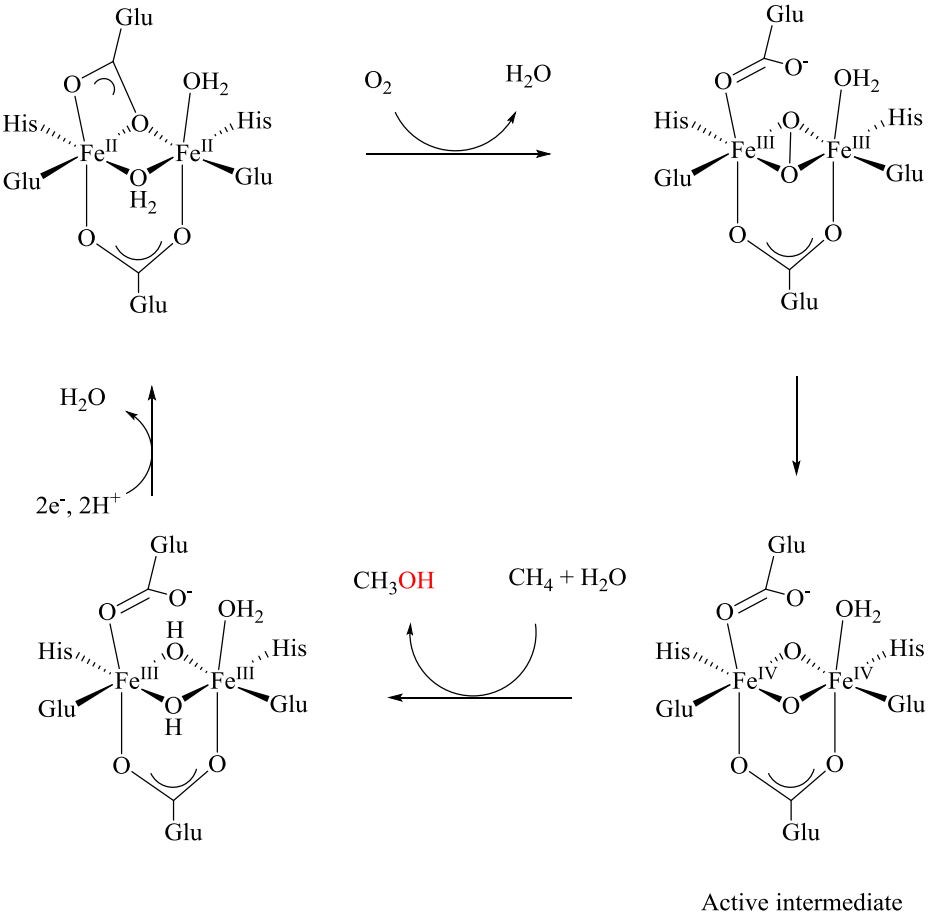


Figure 1.26 The catalytic cycle for the soluble methane monooxygenase

In terms of bimetallic copper enzymes, certain classes of them also possess the ability of hydroxylating C–H bonds, namely tyrosinases, dopamine β-monooxygenase, peptidylglycine α-hydroxylating monooxygenase. Tyrosinase utilizes a coupled binuclear copper center, where two coppers are each ligated by three histidine residues in close proximity to one

another. While for DβM and PHM, they utilize two nonequivalent copper centers that exhibit no observable magnetic interactions. Some of the reactions catalysed by binuclear copper enzymes are shown in Figure 1.27.[279-281]

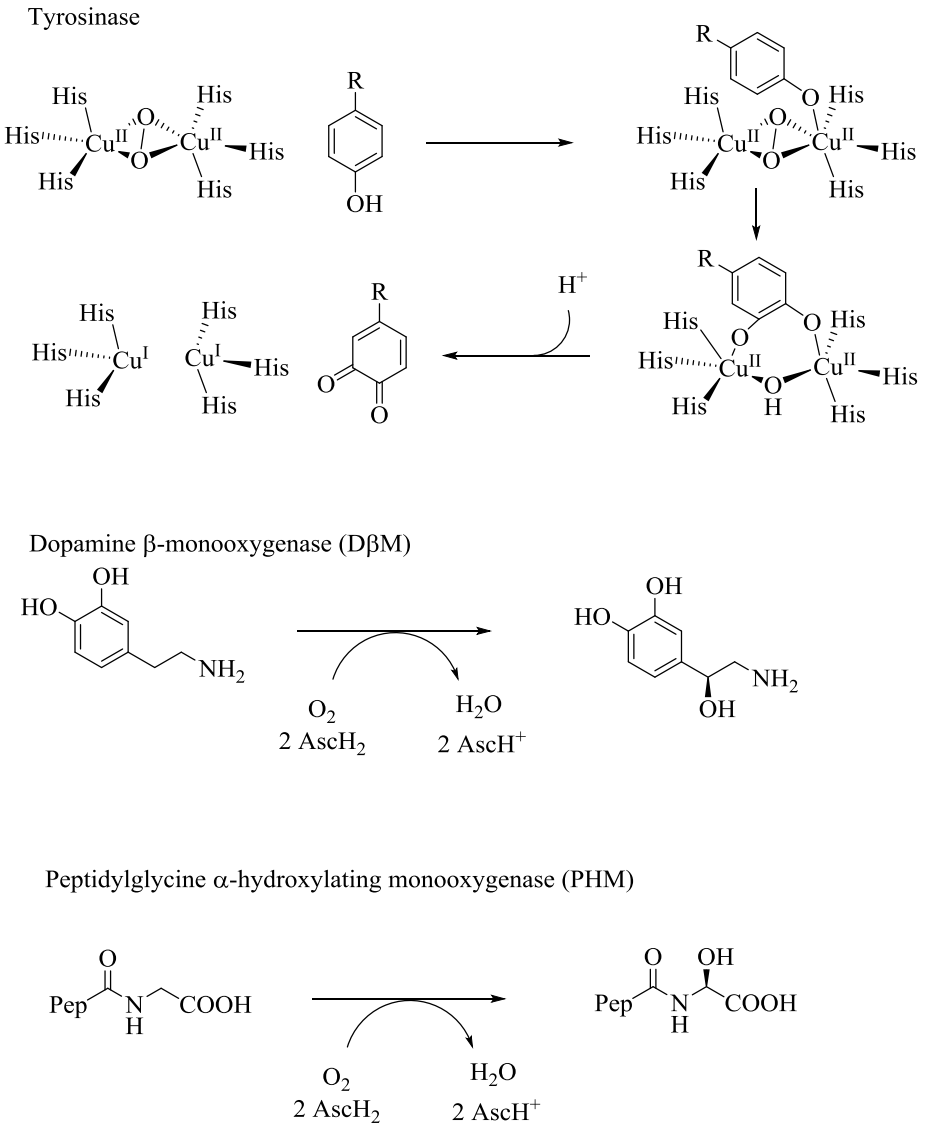


Figure 1.27 Reactions catalysed by tyrosinase, dopamine β-monooxygenase and peptidylglycine α-hydroxylating monooxygenase

1.5.4 Amine oxidases

Amine oxidases are a class of the oxidoreductase enzymes (EC 1.4.3.4) that catalyse redox reactions with amine compounds, in particular monoamine oxidases (MAOs) are able to oxidise amines to imines with simultaneous reduction of oxygen to hydrogen peroxide

(Figure 1.28). [282, 283]

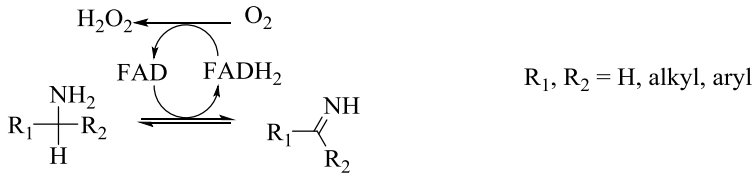


Figure 1.28 General reaction schemes for imines production catalysed by amine oxidases

Amine oxidases are classified into two groups: copper dependent (Type I) and flavin dependent (Type II). Type I amine oxidases form an intermediate imine in their catalytic cycle which remains covalently bound to the protein and therefore are not suitable for biocatalytic applications. While for Type II enzymes, free imines are generated through the reaction process, enabling these enzymes to be applied in synthesis.

Although mechanistic studies have been carried out and a number of mechanisms for Type II MAO-catalysed amine oxidation have been proposed over the years, no agreement concerning the mechanism of these enzymes has been reached. Figure 1.29 shows some possible mechanisms of MAO-catalysed amine oxidation.[282-284] In the aminyl radical cation mechanism, a non-bonded electron of the amine nitrogen is transferred to the flavin to give the flavin semiquinone radical and the amine radical cation. A radical intermediate is then generated by the loss of the α-proton of the amine, which undergoes decomposition either by a second electron transfer to the flavin or radical-radical combination followed by two-electron transfer to the flavin. While in the nucleophilic mechanism, the FAD is non-covalently attached to the enzyme of which the N5 atom is proposed to function as a base and abstract the α-proton of the amine. The amine oxidation is initiated by the nucleophilic attack of the amine nitrogen on the flavin to form a flavin C4a adduct that abstracts the pro-R–H from the substrate to give the imine product.[285]

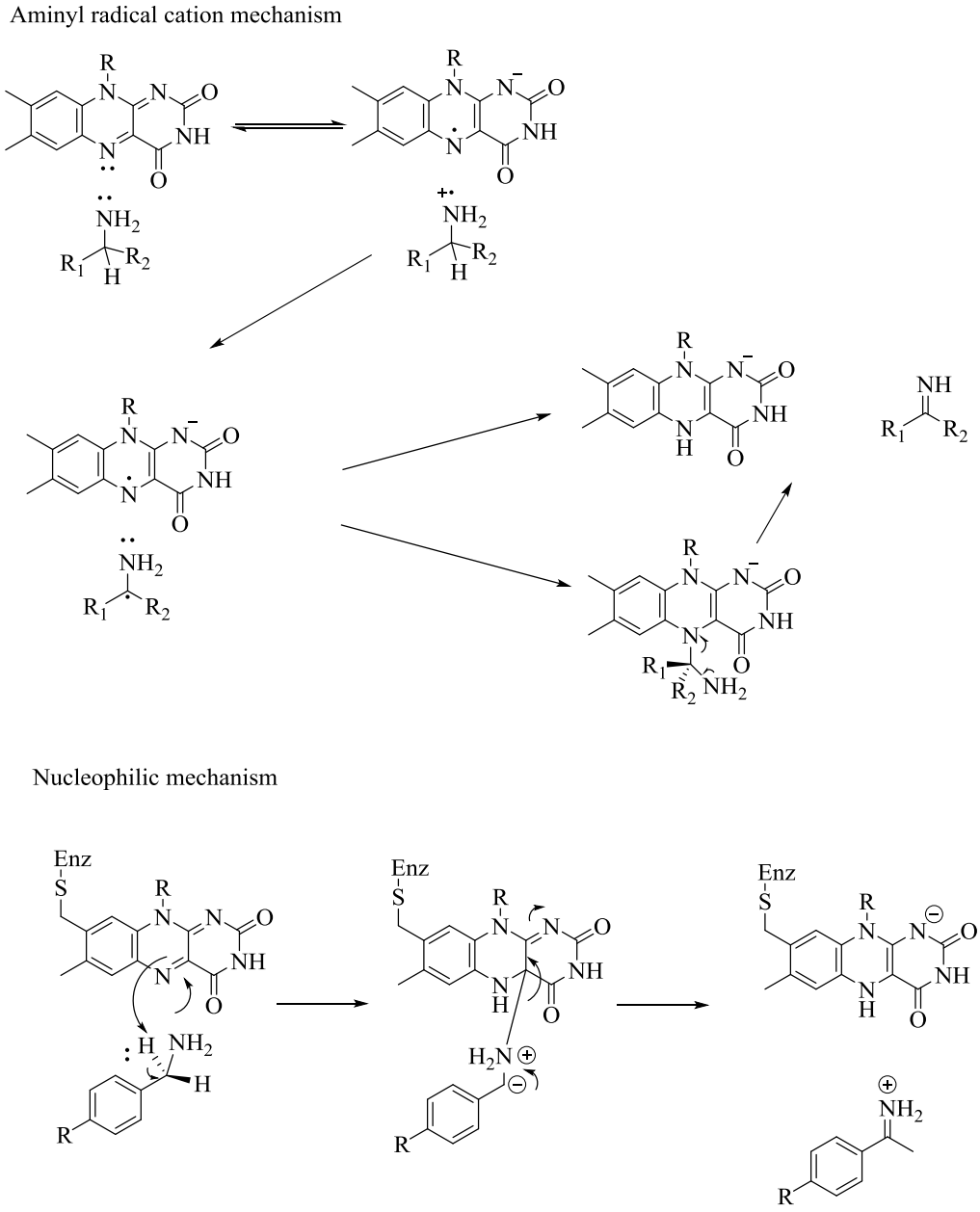


Figure 1.29 Proposed mechanisms for amine oxidation catalysed by MAOs

Since amines and nitrogen heterocycles are valuable building blocks for the synthesis of pharmaceutical drugs and agrochemicals, the MAOs have attracted much attention in the last decade. The MAO from the fungus *Aspergillus niger* (MAO-N), in particular, has been engineered by directed evolution to provide a tool-box of vatiants capable of generating enatiomerically pure amines. These MAO-N variants have been combined with non-selective chemical reducing agents in deracemisation processes,[286] with enantioselective imine

reductases (IREDs) to generate the corresponding chiral amines,[287, 288] with chemical nucleophiles to introduce functional groups at amine α position thus producing key intermediate in the synthesis of the natural product boceprevir[289] (Figure 1.30).

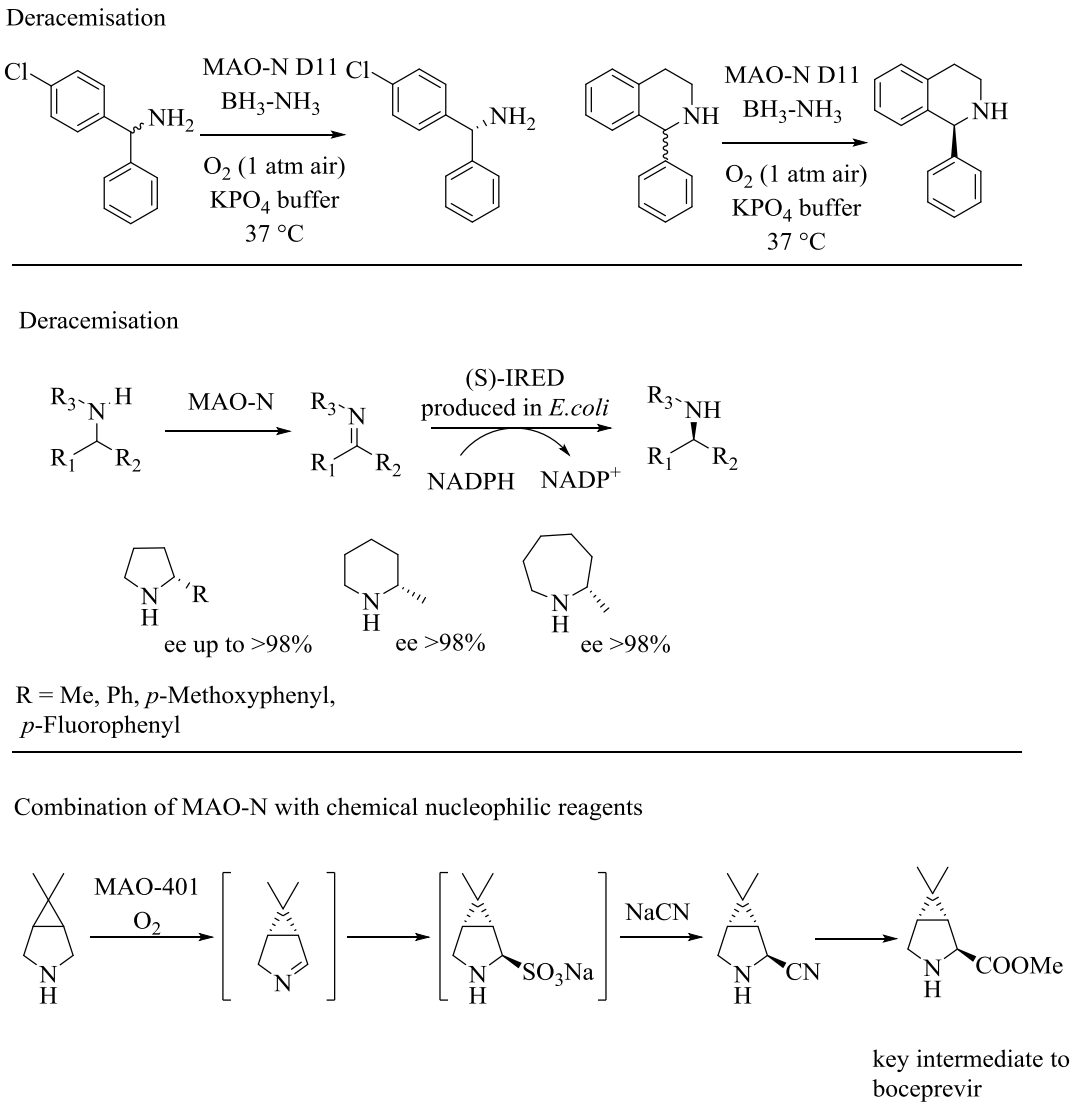


Figure 1.30 Some applications of MAO-N

MAO-N is a highly versatile biocatalyst that can be combined with other biocatalysts and artificial synthetic enzymes and therefore having great potential to become the catalysts of choice in C–H functionalisation synthesis.

1.6 Challenges to be addressed

Biocatalysts are playing an increasingly important role in synthesis. Enzymes, both the isolated forms and those residing inside living cells, have been developed and employed for these tasks. Researchers have put great efforts into enhancing enzyme activity, thermostability, selectivity (chemo-, regio-, enantio-), substrate specificity and robustness in organic solvents by evolving naturally occurring enzymes. Successful examples include ketone reduction,[290-295] Baeyer-Villiger oxidation,[296-305] enantioselective transesterification,[306-310] and epoxide opening.[311] However, only few cytochrome P450 enzymes have been reported in industrial applications.[312-315] The challenges currently facing the development of P450 enzymes as a robust biocatalyst are discussed below.

1.6.1 Protein engineering (laboratory evolution)

As discussed earlier in this chapter (section 1.3), a variety of protein engineering strategies and methods have been developed to evolve $P450_{BM3}$ towards certain targets. However, some hurdles remain to be overcome in the laboratory evolution process.

Rational design and directed evolution are the most widely used strategies in the engineering of $P450_{BM3}$. In terms of rational design strategy, it is generally difficult to predict the effect of even a single mutation due to the highly complex substrate-dependent protein structure-function relationship as well as the substrate-induced protein conformational change. Furthermore, rational design relies on site-directed mutagenesis and therefore delivers only a small number of enzyme variants which give a relatively small coverage of possible sequence space. Screening for mutants with high regio- and particularly enantio- selectivity becomes challenging. High conversion and high total turnover number that enable the reactions to be scalable are also challenging targets since mutations controlling substrate conversion and product inhibition are not clear.

While for directed evolution, although large number of mutants can be produced to cover more sequence space in which it becomes possible to identify highly regio- and enantio-selective variants, the overall process remains time-consuming and robotic high-throughput screening systems are generally required to screen relatively large libraries of enzyme variants, typically 10^3−10^6 transformations. Moreover, within the large number of variants, the number of duplicate and "junk" transformants is large and therefore lowering the probability of hits. For example, Reetz and co-workers evaluated a library of 5,000 clones generated by iterative saturation mutagenesis (ISM) for the enhancement of the enantioselectivity of the hydrolytic kinetic resolution of glycidyl phenyl ether and found the NNK codon library only contain 38 positive mutants.[316] In addition, the variants generated by saturation mutagenesis often lack of diversity since only few amino acid residues can be targeted in order to keep the library to a controllable size. Statistical analysis has pointed out the relationship between percentage coverage of a library and the degree of oversampling, irrespective of codon usage based on the assumption that all sequences occur with equal probability. In order to get 95% coverage of a library generated by NNK codon degeneracy on 3 or 4 amino acid residues, 98,163 and 3,141,251 transformations are required, respectively.[317] The selection of starting templates for enzyme evolution remains another challenging issue while this is also highly relevant to the diversity of variants library. Reetz and co-workers selected F87A as the starting template for the oxidation of testosterone since the F87A mutant produced 2β-hydroxytestosterone and 16β-hydroxytestosterone in 1:1 ratio. Over 8,600 transformations were made but it was found that the variants library could only shift the ratio towards the existing products without "escaping" from the products profiles.[192]

1.6.2 Industrial application (Scale-up optimisation)

Scalability is one of the most important hurdles that need to be overcome in most industrial

applications.[318-320] On the basis that $P450_{BM3}$ is a catalytically self-sufficient enzyme, soluble and can be engineered to a high level of activity towards a wide range of non-natural substrates, it is widely perceived to be the most promising member of the P450 superfamily for industrial applications. The main challenges include overexpressing P450 enzymes at industrial fermentation level, precise design of bioreactors, getting efficient mass transport, increasing substrate conversion and product titre, isolating products from the reaction mixture, waste (biowaste and wastewater) treatment, and among others.

Moreover, $P450_{BM3}$, like other isoforms in its superfamily, is pyridine nucleotide-dependent and requires the expensive co-factor NADPH to function. Enzymatic co-factor regeneration systems have been developed based on robust dehydrogenases and are now commonly used to reduce the co-factor cost such as glucose/glucose dehydrogenase,[227, 242, 321] glucose-6-phosphate/glucose-6-phosphate dehydrogenase,[81, 99, 131, 322, 323] formate/formate dehydrogenase,[204, 324] etc. These components should therefore be taken into consideration with respect to P450 components. An alternative is to perform the biotransformation *in vivo* (whole-cell oxidation) where the onus of co-factor production is undertaken by the host organism. However, this method introduces a few other separate challenges including enhancing membrane permeability that enables hydrophobic substrates to enter the cytoplasm, increasing the life-time of cells thus keeping cultures viable and productive, and isolating organic compounds from cultures. Certain substrates can be toxic to the cells and have to be avoided for whole-cell reactions.

1.7 Thesis aims

The aim of this thesis is to develop and refine a small library of $P450_{BM3}$ variants with sufficient diversity in structural topology in order to oxidise a wide variety of organic molecules including hydrocarbons, oxygen heterocycles, nitrogen heterocycles, carboxylic acids, aromatic compounds, aliphatic amines, long-chain and short-chain compounds. Enhanced activity and high regio-/enantio-selectivity are always primary targets in order to functionalise unreactive C−H bonds at both early and late stages in synthesis. New functions or new reaction types are also important objectives to pursue and they can be developed further as new synthetic methodologies. Fine chemical and natural product synthesis are targeted as well as the scalability of reactions. The library can ultimately be used as a resource for testing non-natural substrates, selecting starting templates, identifying key amino acid residues, collecting information of mutation-selectivity trends and therefore guiding further protein engineering.

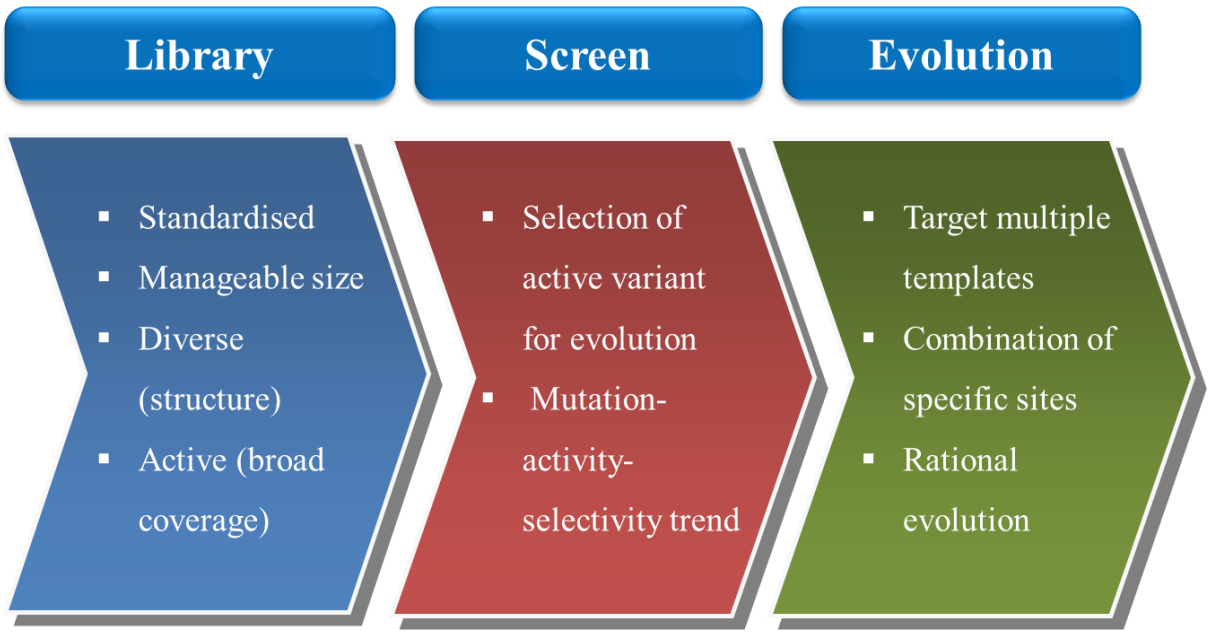


Figure 1.31 Thesis aims and research framework

Chapter Two

Chapter 2 Materials and methods

2.1 Materials, chemicals and equipment

2.1.1 Chemicals, solvents and materials

General reagents and chemicals were from Alfa-Aesar, Fisher Scientific and Sigma-Aldrich. All the drug compounds were from Sigma-Aldrich except chlorzoxazone, which was from Alfa-Aesar. Lidocaine derivatives were kindly synthesised by Mr Jack O'Hanlon at Oxford. Drug fragment compounds were supplied by GSK. Isochroman, 5,6,7,8-tetrahydroquinoline, cyclohexanecarboxamide, ethylpyrazine, acetylpyrazine, ε-caprolactone, ε-carprolactam, 8-methyl-1,2,3,4-tetrahydroquinoline were from Sigma-Aldrich. Cyclopentylethanol was synthesised by reducing cyclopentyl acetic acid with lithium aluminium hydride. Dimethyl anthranilate and methyl anthranilate were obtained from De Monchy. Steroids were purchased from Sigma-Aldrich, Acros and Fluka.

Solvents of HPLC quality were from Sigma-Aldrich and Merck. Media components were from Melford Laboratories, Sigma-Aldrich and Fisher Scientific. Kanamycin was from Melford Laboratories. NADPH (monosodium/disodium salt) were from Apollo Scientific or Melford Laboratories, and NADH (disodium salt) from Roche Diagnostics. Isopropyl-β-D-thiogalactopyranoside (IPTG) was from Melford Laboratories or Fisher Scientific. Lysozyme and catalase were from Sigma-Aldrich. Oligonucleotides were supplied by Eurofins Genetic Services, UK.

2.1.2 Growth media and buffer components

Water used in bacterial growth was purified using Millipore Elix systems. Type 1 water (resistivity ≥8 MΩ) was used for the preparation of media and general enzyme manipulations; type 2, "ultrapure" water (≥18 MΩ) was used for molecular biology experiments. The

components of media (per litre) are shown in Table 2.1. All media were steam sterilised prior to use. LB_{kan}, $2 \times YT_{kan}$, FB_{gly} and TB_{kan} included 30 mg L^{-1} kanamycin, which was added after cooling to room temperature. The pH was adjusted using either concentrated HCl for Tris buffer or orthophosphoric acid for phosphate buffers.

Table 2.1 Growth media components

Medium	**Components (per litre)**
LB	10g tryptone, 5g yeast extract, 10g NaCl.
LBagar	as for LB, with 15g bacto-agar added.
SOB	20g tryptone, 5g yeast extract, 0.5g NaCl, 10 mL 250 mM KCl. Adjusted to pH 7.0 using 10 M NaOH.
SOC	SOB supplemented with 20 mL 20% (w/v) glucose (filter-sterilised).
FBgly	19.2 g $K_2HPO_4 \cdot 3H_2O$, 3.6 g $NaH_2PO_4 \cdot 3H_2O$, 2.5 g $(NH_4)_2SO_4$, 2.0 g Na_2SO_4, 1.2 g $Na_3C_3H_5O(COO)_3 \cdot 2H_2O$, 0.5 g NH_4Cl, 5.0 mL glycerol. (Autoclave and then add 2.0 mL 1 M $MgSO_4 \cdot 7H_2O$, 3.0 ml trace elements solution, 1.0 mL 10% w/v thiamine-HCl).
2×YT	16g tryptone, 10g yeast extract, 5g NaCl.
TB	12g tryptone, 24g yeast extract, 4.0 mL glycerol, 2.31g K_2HPO_4, 12.5 g KH_2PO_4
Trace element	20.1 g Na_2EDTA, 16.7 g $FeCl_3.6H_2O$, 0.74 g $CaCl_2.H_2O$, 0.25 g $CoCl_2.6H_2O$, 0.18 g $ZnSO_4.7H_2O$, 0.132 g $MnSO_4$, 0.1 g $CuSO_4.5H_2O$

LB = Luria-Bertani broth; TB = Terrific Broth;

Buffer components (per litre) are shown in Table 2.2.

Table 2.2 Buffer components

Buffer	**Components (per litre)**
Buffer G	4.9 g K_2HPO_4, 1.6 g KH_2PO_4, 0.15 g 1 mM DTT
50 × TAE, pH 8.5	242.2 g tris, 57.1 mL glacial acetic acid, 100 mL 0.5 M Na2EDTA. Supplied by manufacturer and diluted 50 fold before use.
TFB1, pH 5.8	100 mM RbCl, 50 mM $MnCl_2$, 30 mM KOAc, 10 mM $CaCl_2$. Adjusted to pH 5.8 using 1M acetic acid, filter sterilised prior to use.
TFB2, pH 6.8	10 mM MOPS, 10 mM RbCl, 75 mM CaCl2, 30% v/v glycerol. Adjusted to pH 6.8 using 1 M KOH, filter sterilised prior to use.
200 mM phosphate	34.84 g K_2HPO_4, 27.22 g KH_2PO_4
50 mM Tris-HCl, pH8.5	6.05g $(HOCH_2)_3CNH_2$, adjusted to pH 8.5 using 1M HCl

TFB1, TFB2 = Transformation buffers 1 and 2

2.1.3 Equipment

Sonication was carried out using a Misonix 3000 sonicator and Fisher Scientific FB-505 ultrasonic processor. Mutagenesis and PCR reactions were carried out in a Techne thermocycler from Scientific Laboratory Supplies or a Cetus thermocycler from Perkin-Elmer. UV-visible spectra, enzyme quantitations, activity assays, spin shifts were run on an Agilent Cary 50 spectrophotometer. Gas chromatography analyses were performed on a Thermo Finnigan Trace instrument equipped with a flame ionisation detector (FID) and an AS3000 autosampler using a DB-1 fused silica column (30 m × 0.25 mm, 0.25 μm) with helium as carrier gas. The injectors were maintained at 200 ℃, 250 ℃ or 280 ℃ and the FID at 250 ℃. ^{1}H and ^{13}C NMR spectra were recorded on Bruker Avance AVII 400 (400 MHz) or AVIII

500 (500 MHz) spectrometers in deuterated chloroform or methanol, and are internally referenced to the residual protio solvent. Mass spectra (MS) were obtained on a Micromass GCTOF mass spectrometer operating in electron impact (EI) mode and equipped with a solids probe inlet system. Analytical high performance liquid chromatography was carried out on a Gilson Gradient system using a Pursuit XRS C18 column (4.6 × 150 mm, 5 μm) from Agilent, HPLC grade solvents were from Sigma-Aldrich and Merck.

2.2 DNA manipulations

2.2.1 Site-directed mutagenesis

Site-directed mutagenesis was carried out by a polymerase chain reaction (PCR)-based method using the KOD Hot Start Polymerase kit from Merck using oligonucleotides designed following the manufacturer's guidelines. The reaction set-up is described in Tables 2.3 and the cycling conditions are given in Tables 2.4. Residual parent DNA was digested over 16 hours at 37 °C using 1 μL of the restriction endonuclease, *Dpn* I, which cleaves methylated DNA while leaving unmethylated DNA, such as that produced by PCR, intact.

Table 2.3 50 μL reaction set-up for site-directed mutagenesis using KOD Hot Start

Component	**Volume / μL**
10 × KOD Hot Start buffer	5
10mM dNTPs	5
25mM $MgSO_4$	4
DMSO	2
10 μg μL^{-1} template DNA	1
10 μM forward primer (5'−3')	1.5
10 μM reverse primer (3'−5')	1.5
KOD Hot Start polymerase	1
dd H_2O	29

Table 2.4 Cycling conditions for KOD Hot Start DNA Polymerase

Description	Time /s	Temperature / ℃	Number of cycles
Initial denaturation	180	95	1
Denaturation	20	95	
Annealing	10	55-65	30
Extension	160	70	

2.2.2 Primer design

For each mutagenesis reaction, two complementary oligonucleotides comprising 30 to 50 bases were designed, including at least 10 bases to either side of the nucleotide(s) being altered. Oligonucleotide sequences are given in full in Appendix 2. Melting temperatures were calculated using the equation given in the Stratagene mutagenesis protocol (EQ 2.1):

$$T_m\,(^{o}C) = (675/N) + (0.41 \times GC) - M + 81.5 \qquad \text{(Equation 2.1)}$$

where (T_m) is the melting temperature, N the primer length (in bases), GC the guanine/cytosine percentage, and M the percentage of mismatching nucleotides.

2.2.3 Agarose gel electrophoresis

Diagnostic electrophoresis was employed following mutagenesis to confirm that the new plasmid was of the desired length. 0.5 g agarose was dissolved in 50 mL 1 × TAE buffer by microwave heating. Ethidium bromide (2.5 μL, 10 mg/mL) was added, and the gel was poured and left to solidify. 250 mL of 1 × TAE containing 5 μL ethidium bromide was used as running buffer. Samples were loaded into 25 μL wells created at the cathodic end of the cell, with one being reserved for 10 μL of Promega 1 kbp DNA reference ladder. Preparative

samples (24 µL) included 4 µL of glycerol and were run for 90 min at 60 V, after which the separated DNA fragments were visualised under ultraviolet light. Gel slices of <100 µL were excised using a flame-sterilised scalpel, placed in the nebuliser compartment of a Montage gel extraction device (Millipore) and centrifuged at 5,000 *g* for 10 min. Diagnostic samples (12 µL) included 2 µL Promega 6× loading dye and were run for 30 min at 90 V.

2.2.4 DNA purification

DNA was purified using solutions supplied with the Thermo Scientific GeneJET Plasmid Miniprep kit (#K0503) and manufacturer's protocol. A single DH5α colony harbouring the plasmid of interest was taken from an LB_{agar} plate and grown at 37 ℃ for 16 h in 10 mL LB_{KAN}. The culture was centrifuged at 9,000 *g* for 5 − 10 minutes and the cell pellet was resuspended in 500 µL of the Resuspension Solution (100 µg/mL RNase A) and were transferred into two 1.5 mL microcentrifuge tubes. 250 µL of Lysis Solution was then added into each tube and mixed by inverting the tube 4 − 6 times until the solution becomes viscous and slightly clear prior to incubation for 2 − 3 minutes. 350 µL of the Neutralisation Solution was added and mixed immediately and thoroughly. Cell debris and chromosomal DNA were removed by centrifugation at 16,000 *g* for 5 min. The cleared lysate from each tube was transferred to a GeneJET spin column by decanting or pipetting. Spin columns were centrifuged at 11,000 *g* for 1 min and the flow-through was discarded. 500 µL of Wash Solution was added to the spin column, eluted by centrifugation at 11,000 *g* for 1 min and discarded. A further 500 µL wash solution was added, eluted by centrifugation at 11,000 *g* for 1 min and discarded. The columns were then centrifuged for a final 2 min to ensure thorough solution clearance. 50 µL of nuclease-free water was added to the center of GeneJET spin column membrane, and DNA was eluted by centrifugation for 1 min, after which the eluates were combined. The DNA solution was quantitated by UV spectroscopy.

Each absorbance unit at 260 nm corresponded to a concentration of 5 μg/μL after dilution by a factor of 100.

2.2.5 Confirmation of mutation

Amplified genes were sequenced by Geneservice, UK using two primers, T7F and BM3FP2 (see Table 2.5), to ensure the presence of the desired mutation(s).

Table 2.5 Primers for DNA sequencing

Primers	**Sequence**
T7F	TAATACGACTCACTATAGG
BM3FP2	CTGCAGCGAGCAAATCCAGACGAC

2.3 Enzyme production in *E. coli* and purification

2.3.1 Competent cell preparation

Cells of the desired strain were streaked onto an antibiotic-free LB plate and incubated at 37 ℃ for 16 − 24 h. A single colony was grown overnight in 2.5 mL of antibiotic-free LB in a sterile 50 mL centrifuge tube with the shaking at 250 rpm. The overnight culture was transferred to 250 mL freshly-prepared antibiotic-free LB supplemented with 20 mM $MgSO_4$ and grown for a further 4 − 6 h with the shaking at 120 rpm until OD_{600} reaches 0.4 – 0.6. Cells were pelleted by centrifugation in sterile 50 mL tubes at 5,000 *g* for 5 min at 4 ℃. The supernatant was discarded, and the tubes were centrifuged for a further 1 min, after which the remaining supernatant was removed with a sterile pipette. Each cell pellet was gently resuspended in 20 mL filter-sterilised TFB1 on ice, and left for at least 5 min prior to centrifugation at 4,500 *g* for 5 min at 4 ℃. Each cell pellet was gently resuspended in 2 mL filter-sterilised TFB2 on ice, and left for 1 h. The cells were then transferred in 100 μL

quantities to freshly sterilised microcentrifuge tubes and flash-frozen in liquid nitrogen prior to storage at −80 °C.

2.3.2 Transformation

For protein preparation, 100 ng DNA was inoculated into 100 μL *E. coli* BL21(DE3) competent cells freshly thawed from storage at −80 °C. After 1 h on ice, the cells were heat-shocked at 42 ℃ for 10 seconds and placed back on ice for 2 min. 200 μL filter-sterilised SOC medium preheated to 37 ℃ was added, and the cells were incubated at 37 ℃ for 1 h with shaking at 200 rpm before being spread onto an LB_{agar} plate containing 30 mg L^{-1} kanamycin. For DNA preparation, *E. coli* DH5α was employed, with the same protocol.

2.3.3 Enzyme production and purification

A single *E. coli* BL21(DE3) colony harbouring the appropriate plasmid was inoculated into a 50 mL flask containing 10 mL FB_{gly} media and grown at 37 ℃ with shaking at 100 rpm for 24 h. 30% v/v freshly autoclaved glycerol was added and the mixture was transferred in 1 mL aliquots to sterile microcentrifuge tubes and flash frozen in liquid nitrogen prior to storage at -80 ℃ as a cell bank.

A single tube of the cells was then thawed on ice and inoculated into a 2 L Erlenmeyer flask containing 600 mL of FB_{gly} (30 mgL^{-1} kanamycin) and grown at 37 ℃ with shaking at 100 rpm for 24 h. After sufficient cell growth, the temperature was lowered to 20 ℃, and protein production was induced by adding IPTG to 25 μM from a 400 mM stock. 100 μM of 5-aminolevulinic acid and 10 μM $FeCl_2$ were supplemented to help haem synthesis as well as 3 g/L of tryptone to supply amino acids. Cells were harvested after a further 48 h by centrifugation at 5,250 *g* for 10 min at 4 ℃. The red-brown pellet from each 600 mL growth was resuspended in 25 mL 40 mM potassium phosphate, pH 7.4, 1 mM in DTT (phosphate).

Cells were lysed by sonication, and cell debris was cleared by centrifugation at 37,500 *g* for 20 min at 4 ℃. The first purification step was by ammonium sulfate fractionation. The volume of supernatant was typically *ca.* 120 mL, and 20.37 g of solid $(NH_4)_2SO_4$ was added to take the concentration to 30% saturation (http://www.encorbio.com/protocols/AM-SO4.htm). The solution was stirred at 4 ℃ until all the solids had dissolved and then centrifuged in Oak Ridge tubes at 12,000 *g*, 4 ℃ for 5 mins. 7.00 g of $(NH_4)_2SO_4$ was added to the supernatant and stirred in to take the solution to 40% saturation. The solution was then centrifuged for 5 mins at 12,000 *g*, 4 ℃. The resulting pellet was discarded, and 14.93 g of $(NH_4)_2SO_4$ was added to the supernatant to make it 60% saturation. The mixture was again stirred until all the solids had dissolved, and centrifuged for the final time at 12,000 rcf, 4 ℃ for 5 mins. The supernatant was discarded and the pellet was dissolved in 50 mM phosphate buffer. The protein was concentrated by ultrafiltration (MWCO: 100 kDa) until protein concentration was >10 μM. The protein concentration was then adjusted to 10 μM by diluting with 50 mM phosphate buffer (pH 8.0). Finally the protein was divided into 1 mL aliquots to 1.5-mL microcentrifuge tubes and stored at –20 ℃.

For NADPH consumption assays, the enzyme was further purified with an Amersham-Pharmacia DEAE fast flow Sepharose column (200 × 50 mm) pre-equilibrated with phosphate. The protein was eluted using a linear gradient of 80−400 mM ammonium sulphate in phosphate. The red P450 fractions were collected and concentrated by ultrafiltration using a cellulose membrane with a molecular weight cut-off of 30 kDa. A Sephadex G-25 column pre-equilibrated with phosphate was used to desalt the solution, which was then reconcentrated by ultrafiltration. The solution was centrifuged at 9,250 *g* for 5 min at 4 ℃ and filter sterilised.

2.3.4 Protein qutantification

The P450 protein stock was thawed on ice and diluted by adding 50 μL to 950 μL of buffer G. A few mg of sodium dithionite were added to reduce the protein and the UV spectrum was recorded as baseline. CO was gently bubbled through the reduced protein for about 30 sec until A_{450} was maximised. The protein was quantified using an extinction coefficient of 91000 $M^{-1}cm^{-1}$ for the absorbance at $A_{450} - A_{490}$, as shown in the following:

$$[P450]\ \mu M = \frac{20(A_{450} - A_{490})}{91000} \times 10^6$$

Equation 2.2

2.4 Activity assays

2.4.1 *In vitro* assays

2.4.1.1 NADPH turnover assays

NADPH turnovers were run in 1250 μL of oxygenated 50 mM Tris, pH 7.4 containing 0.1 – 0.5 μM P450. 1 mM substrate was added from a 100 mM stock in methanol/ethanol/DMSO. This solution was equilibrated at 30 ℃ for 1 min, followed by the addition of 1 AU NADPH (*ca.* 160 μM). The rate of decay of A_{340} (Δ) was measured, from which an NADPH consumption rate (*N*) was determined using the equation below, where [P450] is the concentration of enzyme in μM; 6.22 is the millimolar extinction coeffiecient of NADPH at 340 nm.

$$N = \frac{1000\ \Delta}{(6.22 \times [P_{450}])}$$

Equation 2.3

2.4.1.2 Assays using a NADPH regeneration system

Oxidation of specific substrates by $P450_{BM3}$ mutants was screened in 0.5 mL assays in either 7-mL glass vials or 24-well plates using an NADPH regeneration system, 1 – 2 μM P450, 50

mM Tris-HCl (pH 8.5) or 200 mM phosphate buffer (pH 7.5). Substrate (0.5 – 2 mM) was added from a 50 mM – 100 mM methanol/ehanol/DMSO stock. Glucose dehydrogenase (2 U/mL) was added, along with glucose to a final concentration of 100 mM, from a 1 M stock in 50 mM Tris-HCl (pH 8.5) or 200 mM phosphate buffer (pH 7.5). $NADP^+$ monosodium salt was added to a final concentration of 80 μM and the reactions were shaken at 200 rpm overnight at ambient temperature. The aqueous phase was extracted with 300 μL of ethyl acetate or diethyl ether (for cationic compounds) and analysed by GC or HPLC.

2.4.1.3 Total turnover assays

Total turnover numbers (TTNs) were determined in 1000 μL of 50 mM oxygenated Tris, pH 7.4 containing 0.01 μM $P450_{BM3}$, 200 μg bovine liver catalase, 5 mM substrate (added as a 100 mM stock in MeOH/EtOH/DMSO), 80 μM NADPH and an NADPH regenerating system comprising 2 U/mL glucose dehydrogenase (GDH) and 100 mM glucose. Assays were sealed in glass vials and agitated gently at 5 °C for 96 h. The reaction mixture was extracted with 600 μL ethyl acetate and analysed by gas chromatography.

2.4.2 *In vivo* assays

A single *E. coli* BL21(DE3) colony harbouring the appropriate plasmid was inoculated into 4 mL LB_{kan} and grown for 16 h at 37 ℃ with shaking at 200 rpm. The culture was transferred to 100 mL freshly prepared FB_{gly} (with 30 mg/L kanamycin and shaking at 120 rpm) and, after a further 16 h − 20 h growth, the temperature was lowered to 20 ℃ for 1 h and protein expression was induced by adding IPTG to 25 μM. After 24 h at 20 ℃, cells were harvested by centrifugation at 9,250 *g* for 5 min at 4 ℃, returned to the same flasks and resuspended in 100 mL 200 mM phosphate buffer (pH 7.5). 25 - 50 μL substrate (or 25 − 50 mg if a solid) was added as well as 3 g/L of glucose, and again at intervals, typically 3 h and 24 h, at which

times sampling also took place. 600 μL samples were taken in duplicate and centrifuged, after which 500 μL of the supernatant was extracted into 300 μL ethyl acetate for analysis.

2.5 Preparative scale synthesis and product characterisation

2.5.1 General procedure for preparative scale *in vitro* biotransformations

The preparative scale reactions were performed in 50 mL − 200 mL 200 mM phosphate buffer (pH 7.5) or 50 mM Tris-HCl Buffer (pH 8.5) using 1 − 2 μM enzyme, 1 − 10 mM substrate was added from a 100 mM stock in either methanol or ethanol. 2 U/mL of glucose dehydrogenase were added from a 2 U/μL stock in the above buffer, along with glucose to a final concentration of 100 mM. $NADP^+$ monosodium salt was added to a final concentration of 80 μM, to initiate the reaction. The reaction was stirred at 500 rpm at room temperature for 6 h. The aqueous phase was then extracted with ethyl acetate (50 mL × 3), the combined EtOAc was dried over $MgSO_4$, and the solvent was removed after filtration. The residue was purified by silica gel column chromatography to give the products after removal of solvent. The products were characterised by NMR and MS, which were recorded in deuterated solvents and methanol, respectively.

2.5.2 Typical procedure for preparative scale *in vivo* biotransformations

A single *E.coli* BL21(DE3) colony harbouring the appropriate plasmid was inoculated into 250 mL FBgly and grow for 16 − 24 h at 37 °C with shaking at 120 rpm, and protein production was induced by adding IPTG to 40 μM, 5-aminolevulinic acid to 100μM and $FeCl_2$ to 10 μM at 20 °C . After 48 h, 200 mL of culture was taken for *in vivo* reaction while 50 mL of culture was used for enzyme quatitation. Cells were harvested from 200 mL of culture by centrifugation at 9,250 *g* for 5 min at 4 °C, resuspended in 200 mL phosphate buffer (200 mM, pH 7.5). 2 mM to 10 mM lidocaine deriatives were added from a 100 mM

methanol or ethanol stock, along with glucose to a final concentration of 100 mM, from a 1 M stock in phosphate buffer. 600 μL samples were taken in duplicate and centrifuged every 4 hours, after which 500 μL of the supernatant was extracted into 300 μL ethyl acetate or chloroform for GC analysis.

The whole *in vivo* biotransformation mixture was extracted with ethyl acetate or chloroform after 72 h and the combined organic was dried over $MgSO_4$, and the solvent was removed after filtration. The crude residue was purified by silica gel column chromatography to isolate products.

Chapter Three

Chapter 3 Engineering $P450_{BM3}$ into a drug metabolising enzyme

3.1 Introduction

3.1.1 General introduction

For a new drug to be approved by the regulatory bodies, e.g. the US Federal Drug Administration and the European Medicines Evaluation Agency, they have to be tested clinically to be non-toxic and safe. When a xenobiotic enters the human body, Phase I oxidative metabolism often occur as the first reaction, that is responsible for converting pro-drugs into active oxidised species (as with the bioactivation of the pro-drug tegafur to 5-fluorouracil),[325] detoxification,[326] and in some case such as acetonitrile and benzo[*a*]pyrene, toxicification. As cytochrome P450s are responsible for over 75% of drug metabolism and clearing in humans, drug metabolites produced by P450s may have toxic and adverse side effects towards human bodies.[327, 328] There is a need to produce these metabolites on a preparative scale, to screen for their efficacy, toxicity and bioavailability and accordingly assess the feasibility of a candidate pharmaceutical. Chemical synthesis of drug metabolites is notoriously difficult. One way to produce these metabolites is to use human liver microsome P450s directly or express them with the co-factor protein cytochrome P450 reductase (CPR) in yeast,[329] *E. coli*,[330] and insect cells.[331] However, due to the low stability, variable expression levels and high cost, these systems have not yet provided sufficient yields to be practical. Microbial cytochrome P450s are soluble, stable and readily expressed to high levels and have high activities for favoured substrates.[332] These enzymes are therefore attractive for the development of active systems for the synthesis of human phase I metabolites. As the wild type $P450_{BM3}$ enzyme from *Bacillus megaterium* has been shown to oxidise a number of drug molecules to their human metabolites at low but detectable activity,[333] $P450_{BM3}$ has been extensively investigated and engineered as a platform for drug oxidation. Both rational design

and directed evolution have been applied and increased enzymatic activity for the oxidation of numerous drug molecules.[11, 332, 334-336] However, even simple drug molecules such as chlorzoxazone, naproxen and lidocaine remain challenging substrates and there are few reports of product isolation and full characterisation.[11, 56, 206, 209, 213, 337-339]

3.1.2 Review of previous work

Various protein engineering strategies have been applied towards drug oxidation, among which directed evolution has become the main approach for identifying and generating libraries of drug oxidising $P450_{BM3}$ variants.

Arnold and co-workers utilised the 'peroxide shunt' pathway where hydrogen peroxide was used to replace the enzyme reductase domain and NADPH as the only chemical reagent to make the enzyme functional. A panel of 'P450 peroxygenases' was developed using the $P450_{BM3}$ haem domain to oxidise fatty acids[340] and were then evolved through a combination of random mutagenesis and 12-*p*-nitrophenoxycarboxylic acid (12-pNCA) colorimetric assay[168] to identify variants metabolising propranolol (shown in Figure 3.1a). The BM3-H9C1 (F87A/I58V/H100R/I102T/F107L/A135S/M145A/N239H/S274T/L324I/I366V/K434E/E442K/V446I) mutant was selected from the initial screen which contains fourteen mutations and showed 18% conversion of 5 mM propranolol with 5 μM enzyme and 6 mM H_2O_2 to produce 4′-hydroxypropranolol, 5-hydroxypropranolol and desisopropylpropranolol. A library of variants based on BM3-H9C1 was created by directing mutations to seven active-site residues (A74, L75, V78, F81, A82, F87, and T88) which are within 5 Å of the substrate in the crystal structure of $P450_{BM3}$ bound to *N*-palmitoylglycine.[135] The screening led to the discovery of the BM3-H2C1 variant (K24R/R47H/A74V/A82L/F87G/I58V/H100R/I102T/F107L/A135S/M145A/N239H/S274T/L324I/I366V/K434E/E442K/V446I) with 20% conversion under the same conditions as BM3-H9C1 that was able to produce 70

mg products from 1 L of batch culture.[341] The turnover number was improved after a further screening of 1,190 variants with the most active mutant (21B3-14C10) forming 23.2 nmol product with 0.4 nmol enzyme in 2h. (turnover number = 58)[342]

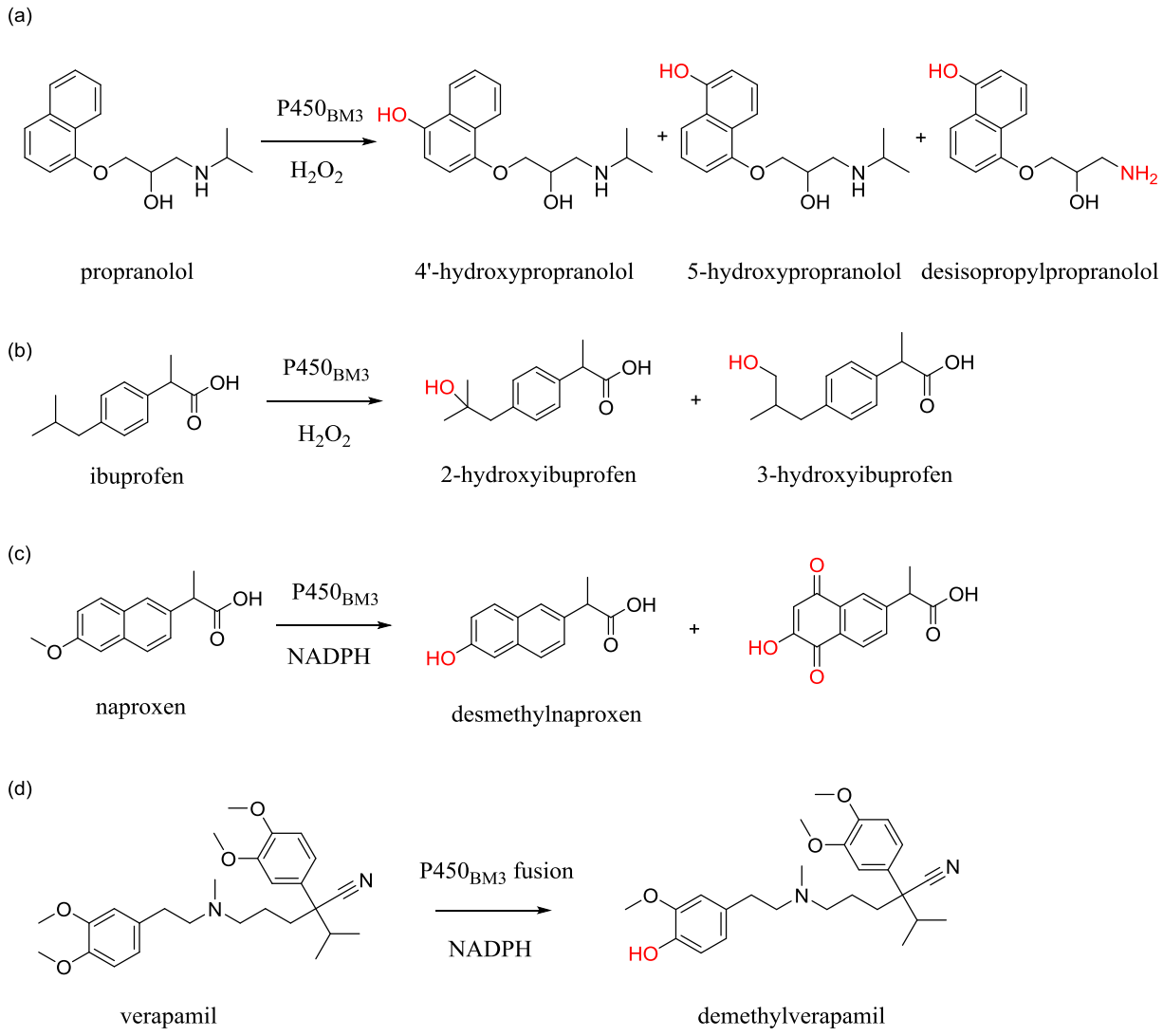


Figure 3.1 Drug metabolism by engineered $P450_{BM3}$

Non-steroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen and naproxen (Figure 3.1b and Figure 3.1c) were also subjected to directed evolution studies. The BM3 haem domain variant 21B3 was identified by screening hydroxylation activity towards 12-pNCA using hydrogen peroxide in place of the NADPH cofactor and oxygen,[340] which was then

used as the template for further protein evolution. 2,210 variants were screened for naproxen oxidation but none of them was active, suggesting difficulties for the negatively charged naproxen molecule to enter the hydrophobic $P450_{BM3}$ substrate-binding pocket. Mutating leucine 75 in the substrate-binding pocket to arginine allowed the identification of the 5H6-13C9 mutant to hydroxylate naproxen by providing a compensating positive charge for the naproxen molecule, though the activity was low, at 12.9 nmol product was formed by 0.4 nmol enzyme in 2 hours.[342] Improved activity towards naproxen was eventually achieved by re-attaching the reductase domain (BMR) to the variant 'peroxygenases' where NADPH regeneration system was used as the electron donor. Conversions of up to 20% for naproxen (3 mM substrate, 0.033 mol% enzyme) were observed with the W7D8 mutant to give desmethylnaproxen and a benzoquinonone compound derived from 5-hydroxylation of desmethylnaproxen.[206] The W7D8 variant was also active towards ibuprofen, with a total turnover number of 650 to produce 2-hydroxyibuprofen and 3-hydroxyibuprofen.[206] Chimeras and fusions which incorporate the haem domain from $P450_{BM3}$ and the reductase domain of CYP102A2 or CYP102A3 were also reported by Arnold and co-workers to be active towards verapamil (Fig. 3.1d). Demethylation was observed with the variant 9-10A F87L(R47C/V78A/F87L/K94I/P142S/T175I/A184V/F205C/S226R/H236Q/E252G/R255S/A290V/L353V) but not with the human microsomes. The metabolite was isolated from a 25 mg reaction with 37% yield.[339]

Gilardi and co-workers carried out similar strategies by combining human drug metabolising enzymes to the $P450_{BM3}$ reductase domain (BMR). The CYP2E1-BMR fusion was successfully expressed and found to be soluble and active towards chlorzoxazone (K_M = 0.65 $\pm$0.08 mM and k_{cat} = 0.95 $\pm$0.10 min^{-1}).[243, 343, 344] Other fusion proteins including CYP2C9-BMR, CYP2C19-BMR and CYP3A4-BMR were also developed by Gilardi group and

reported to have similar drug metabolising activities compared to the parent human cytochrome P450s.[214]

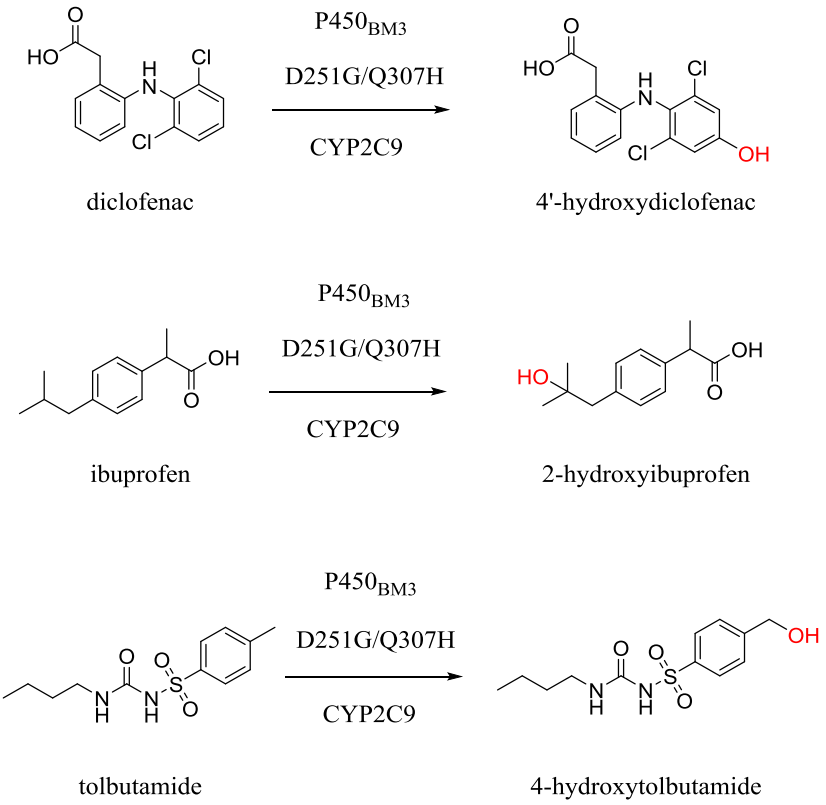
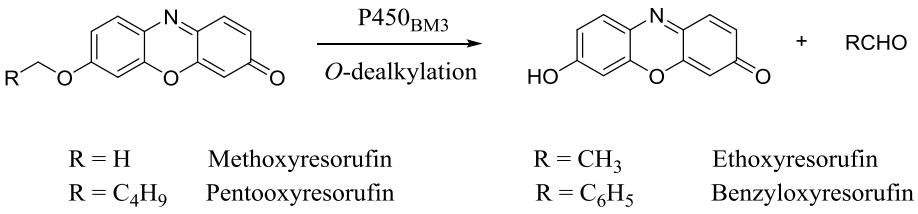


Figure 3.2 Drug metabolism by engineered $P450_{BM3}$ (published by Arnold and co-workers)

Directed evolution was also carried out by Gilardi and co-workers using an alkali assay for screening a library of $P450_{BM3}$ random mutants derived from error-prone polymerase chain reaction.[345] The alkali assay was based on the fluorescent product derived by alkali degradation of the $NAD(P)^+$ generated during the enzymatic turnovers.[346] This approach has led to the identification of mutant Asp251Gly/Gln307His (A2) which converted diclofenac, ibuprofen and tolbutamide to their human CYP2C9 metabolites (Figure 3.2).[345] However, this work was focused on characterising biochemical data such as dissociation constants and K_M values, the mutants did not provide sufficient activity for metabolite synthesis.

The Vermeulen group have reported high throughput screening for drug oxidation. R47L/F87V/L188Q was obtained from a fluorescent assay with 7-benzyloxyresorufin and

was found to be 900 times more active at *O*-demethylation compared to WT.[208] The assay was based on benzyloxyresorufin *O*-dealkylation (BROD) activity that could be applied in whole-cell oxidations (Figure 3.3). A library of 45 drug-like compounds was then tested using the R47L/F87V/L188Q mutant and the BROD assay, of which eight molecules showed high affinity towards this mutant.



R = H	Methoxyresorufin	$R = CH_3$	Ethoxyresorufin
$R = C_4H_9$	Pentooxyresorufin	$R = C_6H_5$	Benzyloxyresorufin

Figure 3.3 Benzyloxyresorufin *O*-dealkylation (BROD) assay

This variant was later shown to be active towards testosterone, dextromethorphan, amodiaquine, acetaminophen, and 3,4-methylenedioxymethylamphetamine (MDMA).[210] By introducing various combinations of mutations at R47, S72, F87, L188, L437, etc., further evolution and screening based on these mutants generating mutants M01, M02, M05, M11, MT35, MT38, and MT43, which showed improved drug oxidation activity of 43 drugs tested, with 33 being more than 20% converted while some substrates were 99% converted, using 0.5 μM variant enzymes and 40 μM drug substrate, corresponding to a TTN of 80 for 100% conversation.[209, 215, 216] These investigations utilised HPLC co-elution and LS-MS/MS for analysis. The drug metabolites were not synthesised or isolated at preparative scale due to the low turnover numbers and therefore the structures of these metabolites remained uncharacterised.

3.1.3 Library development based on generic accelerators (I401P, KSK19, KT2)

The aim of the work reported in this chapter was to develop a panel of $P450_{BM3}$ variants that can mimic the product profiles of human drug metabolising enzymes (CYP3A4, CYP2C9, etc.) and synthesise preparative scale yields of drug metabolites. As described in section 3.1.2, previous reports in the literature have relied on high throughput screening systems where large numbers of variants have to be created and screened case by case based on specific structures of drug molecules. 'Junk' and duplicate transformants are inevitable through this approach since the reaction conditions are difficult to standardise due to the varied expression levels and randomised mutants.

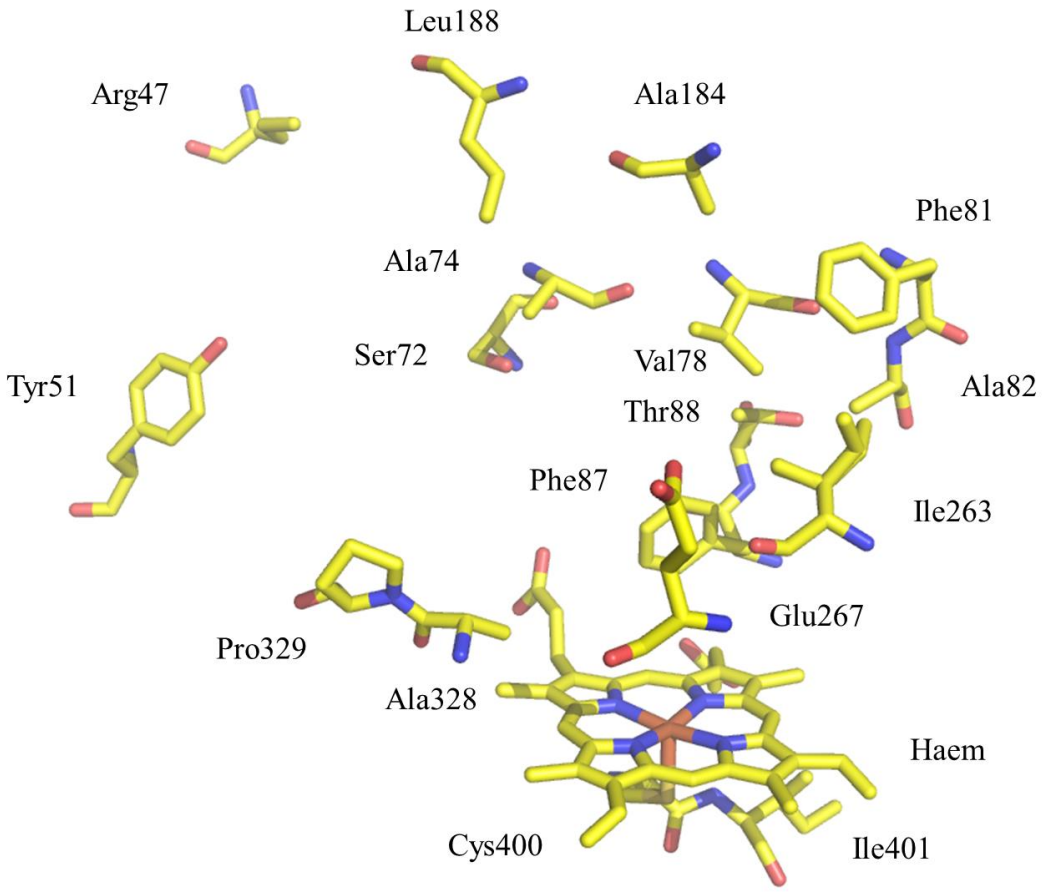


Figure 3.4 Active site of SF WT (PDB code: 1BU7) showing residues commonly mutated in drug metabolising variants

An alternative is to use variants that have shown enhanced activity towards a broad range of molecules as a good basis for protein engineering, by introducing substitutions of different side-chain volumes at specific amino acid residues (Arg47, Tyr51, Ser72, Ala74, Val78, Phe81, Ala82, Phe87, Thr88, Ala184, Leu188, Ile263, Glu267, Ala328, Pro329 , etc.) known to contact or were close to a bound substrate. (as shown in Figure 3.4)[112]

Previous directed evolution studies within our group have identified four 'generic accelerator' $P450_{BM3}$ variants, A330P (AP)[347], I401P (IP)[163], A191T/N239H/I259V/A276T/L353I (KT2)[347] and F87A/H171L/Q307H/N319Y (KSK19)[347] simply derived from Wild-type (WT) that possess increased oxidation activity for a wide range of organic compounds such as alkylbenenes, short- to medium- chain alkanes, fatty acids, polyaromatic hydrocarbons, and etc. The crystal structures of the haem domains of A330P and I401P variants were solved in a previous study within the group. In the A330P structure, the enzyme active site is constricted by the relocation of Proline329 side chain into the substrate entrance channel and therefore providing a basis for the distinctive C–H bond oxidation profiles and the enhanced activity towards small molecules. While in the I401P structure, the introduction of proline at Ile401 position caused the Ala399-Gln402 loop to drop away from the active site, and pull the iron atom to the proximal side towards Cys400 by 0.5Å, as found in the NPG-bound structure. This movement elongates the Fe-OH_2 bond by 1.1Å without a substrate bound, allowing the water molecule to be displaced more easily by a non-natural substrate and lowers the barrier to the rate-limiting first electron transfer step of catalytic cycle, thus increasing the turnover frequency. KT2 and KSK19 variants were discovered through a directed-evolution study within the group using gene-shuffling strategy and indigo screening.

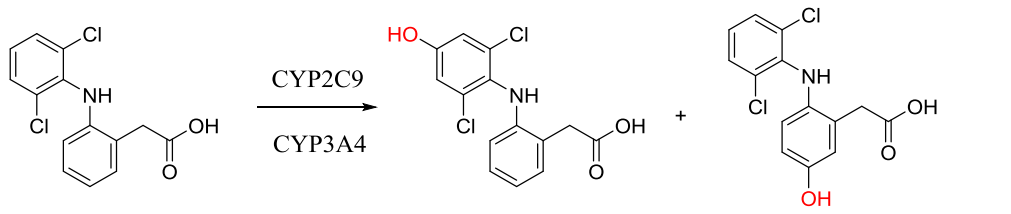
An initial screen of these mutants with common drugs showed detectable but modest activity[348] and they became the basis for the development of drug metabolising variants.

By introducing mutations with different side-chain volume at various sites onto these four base variants, a library of ~50 variant enzymes was constructed and they were screened towards six common drugs that span anionic (diclofenac, naproxen), neutral (testosterone, chlorzoxazone) and cationic (amitriptyline and lidocaine) molecules, which will be discussed in detail in this chapter.

3.2 Anionic drugs

3.2.1 Diclofenac

Diclofenac is a common nonsteroidal anti-inflammatory drug (NSAID) taken to reduce inflammation and as an analgesic reducing pain in certain conditions. It is metabolised primarily by human CYP2C9[349] and CYP3A4[350] into the 4'-hydroxydiclofenac (**3.01**) and 5-hydroxydiclofenac (**3.02**), respectively (Figure 3.5). Despite directed evolution efforts discussed earlier in section 3.1.2, no $P450_{BM3}$ variants have been reported to have high activity and conversion towards this substrate (the highest reported conversion was 50%),[213, 345] although other related NSAIDs such as meclofenamic acid have been found to be oxidised.[338]



Diclofenac | 4'-hydroxydiclofenac, **3.01** | 5-hydroxydiclofenac, **3.02**

Figure 3.5 Diclofenac and the two major human metabolites, 4′- and 5-hydroxydiclofenac

The initial screening was carried out with ~50 mutants based on the four 'generic accelerators' in a total volume of 0.5 mL in both 10.5 mL glass vials (Cat. No. 151064 from Wheaton, UK) and 24-well plates (Cat. No. 83.3922.300 from Sarstedt, UK), with 2 mM diclofenac being

added from a 100 mM stock in MeOH, 5 μM enzyme and 40 μM $NADP^+$ monosodium salt. The regeneration system was composed of 10 U of glucose dehydrogenase (GDH from Codexis) and 100 mM glucose (Figure 3.6).

Table 3.1 Product distribution and substrate conversion of the initial screen with diclofenac

Number	Enzyme	Products [a]				Conversion [b]
		3.01	**3.02**	**X**	**Y**	
1	RP/FV/EV/F81W	31%	60%	9%	-	100%
2	RP/FV/EV	35%	50%	15%	-	91%
3	RP/IA/EV	-	93%	-	7%	35%
4	RP/FV/EV/V78F	23%	47%	15%	15%	22%
5	RT2/F81W	28%	35%	16%	21%	18%
6	RP/EV	-	64%	-	36%	15%
7	F87A/IP	26%	8%	15%	51%	12%
8	KT2/E267F	-	79%	-	21%	10%
9	QP/A82M	29%	-	26%	45%	8%
10	RP/FV	13%	26%	17%	44%	5%
11	KSK19/QP/F87V	20%	63%	17%	-	5%

[a] Diclofenac metabolism was detected using GC and TLC whereas the quantification was based upon peak integration of the ^{1}H NMR spectra of crude extracts of 10 mL reactions. The other products **X** and **Y** were not formed in sufficient quantities to be characterised.
[b] Results represent averages of three independent experiments with errors not higher than 10%. Only mutants with >5% conversion are listed.

Glucose dehydrogenase (GDH) catalyses the dehydrogenation of glucose to form a lactone, with $NADP^+$ being reduced to NADPH.[351] The lactone is spontaniously hydrolysed in aqueous solution to gluconic acid.

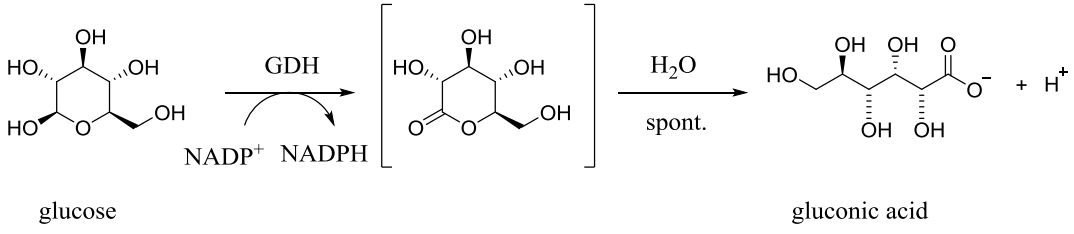


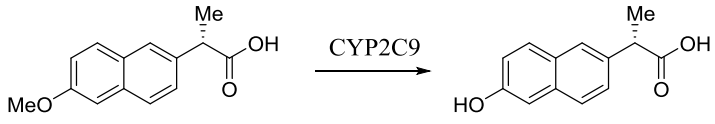
Figure 3.6 NADPH regeneration system using glucose and glucose dehydrogenase (GDH)

The reaction were carried out in 200 mM phosphate buffer, pH 7.5 and left for 12 h at 20 ℃ whilst shaking at 150 rpm. The reaction mixtures were extracted by ethyl acetate and analysed by gas chromatography and crude LC-MS. The product distribution and substrate conversion data are shown in Table 3.1 where inactive variant enzymes are not shown. The screening results revealed that only mutants containing the F81W, F87V, and E267V mutations showed >10% diclofenac conversion. The RP/FV/EV mutant showed 91% conversion while the addition of F81W mutation led to 100% conversion of 2 mM diclofenac. The two major products were then isolated by silica-gel chromatography from a 118 mg preparative scale reaction (0.4 mmol, 200 mL total reaction volume) with 5 μM RP/FV/EV mutant and characterised as the human metabolites 4′-hydroxydiclofenac (**3.01**, 34% isolated yield) and 5-hydroxydiclofenac (**3.02**, 47% isolated yield) by NMR and high-resolution MS. The two minor products were not formed in sufficient quantities for isolation and characterisation.

Given the broad substrate range of the library the low number of hits was unexpected but consistent with the few reports of diclofenac oxidation activity by $P450_{BM3}$. The screening data indicated a crucial role for a mutation at E267, with the E267V being more effective than E267F while mutations to isoleucine and leucine did not increase activity. It is not surprising that the F87V/E267V combination was also present in mutants that oxidised meclofenamic acid.[338]

3.2.2 Naproxen

Naproxen, like diclofenac, is another typical NSAID that is commonly used for relief of a wide variety of pain, fever, inflammations, and stiffness. It is structurally more rigid and presents different steric demands compared to diclofenac. It is metabolised primarily by CYP2C9 to the *O*-demethylation product, 6-desmethylnaproxen (Figure 3.7).[352]



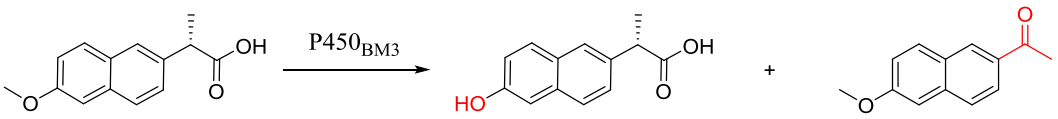
Naproxen 6-desmethylnaproxen

Figure 3.7 Naproxen and its human metabolite, 6-desmethylnaproxen

The initial screening was conducted at 0.5 mL scale in 10.5 mL glass vials, with 2 μM enzyme and 1 mM substrate being added from 100 mM stock in MeOH. NADPH regeneration was used under the same conditions as for diclofenac. The reaction mixtures were extracted with 300 μL ethyl acetate and analysed by gas chromatography. Table 3.2 shows the product distribution and substrate conversion of the screening results.

From a screen of 44 $P450_{BM3}$ variants, 13 variants were found to give >40% conversion to two major products while two minor products were also formed but remained uncharacterised due to the insufficient yield. Most of the active mutants contained the A330P (AP) mutation, with three RT2/AP-based variants converting >95% naproxen. The KSK19 series of mutants with the F87A mutation were also effective. Surprisingly, the WT $P450_{BM3}$ was active and converted 5% naproxen into the two major products, which was not reported in the previous reports.[206, 342]

Table 3.2 Product distribution and substrate conversion of the initial screen with naproxen



Naproxen 6-desmethylnaproxen, **3.03** 2-acetyl-6-methoxynaphthalene, **3.04**

Number	Enzymes	Retention time /min [a]				Conversion [b]
		4.71 (3.04)	4.91 (3.03)	6.67	7.08	
1	RT2AP/F81W	68%	32%	-	-	100%
2	RT2AP/PV	34%	66%	-	-	96%
3	RPAP/A328V	63%	37%	-	-	96%
4	RLYF/AM/IA/QP	70%	30%	-	-	85%
5	RP/FW	66%	34%	-	-	68%
6	KSK19/AI/IP	66%	34%	-	-	67%
7	RT2AP/AI/S72A	62%	38%	-	-	66%
8	KSK19	29%	67%	5%	-	61%
9	RP/FV/EV	60%	40%	-	-	58%
10	RT2/A330W	52%	48%	-	-	43%
11	RPAP/A74W	57%	43%	-	-	42%
12	KSK19/AI/Y51V	50%	45%	5%	-	41%
13	RT2/FW/AI	66%	34%	-	-	41%
14	KSK19/AI/FV	67%	33%	-	-	39%
15	KT2	18%	82%	-	-	28%
16	RT2IP/V78I	56%	44%	-	-	27%
17	GVQ	21%	79%	-	-	26%
18	RP/FV/EV/AI	60%	40%	-	-	16%
19	RT2/AP/AM/AI	73%	27%	-	-	16%
20	RT2/AP	-	21%	40%	39%	15%
21	KSK19/I263A	14%	86%	-	-	11%
22	RLYF/KSK19/AI	56%	44%	-	-	11%
23	F87A	48%	52%	-	-	11%
24	KSK19/AI/AI	15%	75%	-	11%	10%
25	QP/AM	55%	45%	-	-	10%
26	WT	20%	80%	-	-	5%
27	RP/FV	-	-	67%	33%	4%
28	RP/EV	-	-	-	-	0%
29	RPAP	-	-	-	-	0%

[a] Naproxen metabolites were identified using GC. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.
4.71 min product: 2-acetyl-6-methoxynaphthalene **3.04** (characterised by NMR and MS)
4.91 min product: 6-desmethylnaproxen **3.03** (characterised by NMR and MS)
6.67 min product: unidentified
7.08 min product: unidentified
[b] Average of 2 independent experiments with errors not higher than 10%.

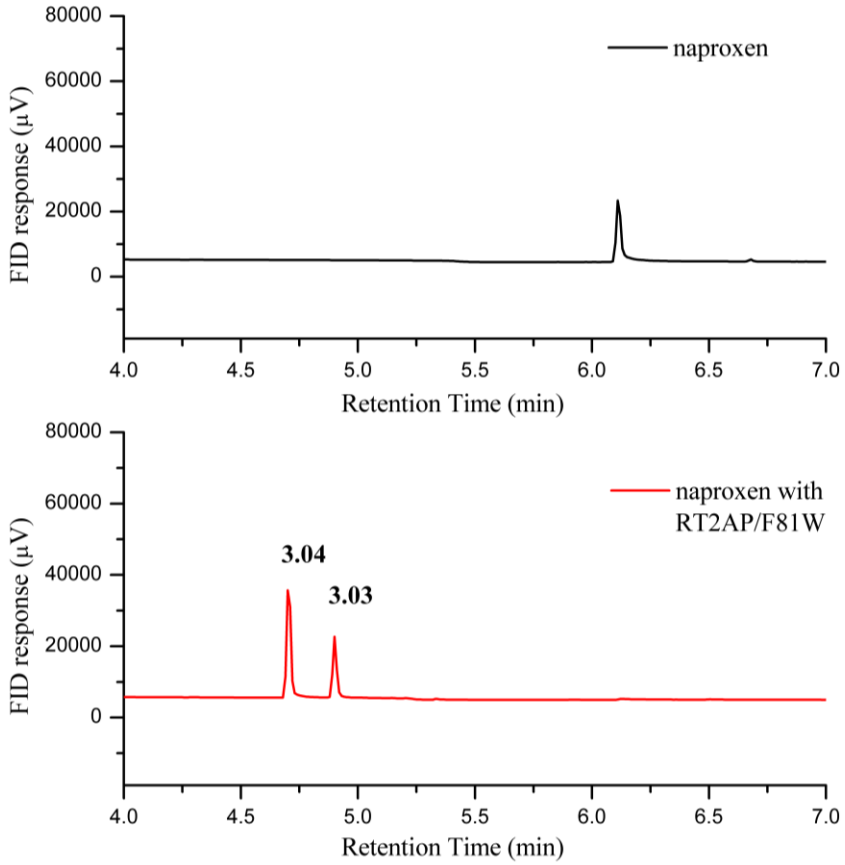


Figure 3.8 Gas chromatography analysis of 1 mM naproxen metabolised by 2 μM RT2/AP/F81W

RT2/AP/F81W was the most active mutant among the panel of 44 enzymes and displayed total conversion towards 1 mM naproxen while RP/AP/A328V gave similar conversion and product profiles. Compared to RT2/AP/F81W, the RT2/AP/PV variant showed 96% conversion but different product distribution, with the *O*-demthylation product formed at 66% selectivity. The parent mutant RT2/AP did not show high activity towards naproxen (15% conversion), suggesting the important roles of F81W and P329V mutations. The high activity of the RP/FV/EV mutant was also of interest since the RP/FV and RP/EV mutants showed no conversion.

The two major products were purified by silica gel chromatography from a 46 mg preparative scale reaction (0.2 mmol substrate, 200 mL total volume) using 2 μM RT2/AP/F81W and

characterised as 6-desmethylnaproxen (**3.03**, 30% isolated yield) and 2-acetyl-6-methoxynaphthalene (**3.04**, 62% isolated yield). Figure 3.8 shows the gas chromatography traces for the conversion of 1 mM naproxen by 2 μM RT2/AP/F81W with the product peaks being labelled.

2-Acetyl-6-methoxynaphthalene (**3.04**) is a novel metabolite (^{1}H NMR and ^{13}C NMR are shown in Figure 3.9) that has not been reported as a product from naproxen oxidation by a P450 enzyme. It is most likely formed through α-hydroxylation followed by the new P450 reaction type of oxidative decarboxylation of the α-hydroxyacid to the ketone. The proposed mechanism is shown in Figure 3.10.

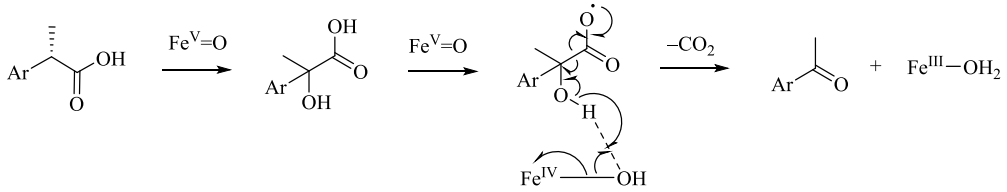


Figure 3.10 Proposed mechanism for the formation of the naproxen oxidative decarboxylation metabolite 2-acetyl-6-methoxynaphthalene (**3.04**)

The carboxyl radical formed by either hydrogen atom abstraction from the acid or electron transfer oxidation of the carboxylate by the ferryl, compound I species loses CO_2 and forms the ketone by abstraction of the alcohol-OH hydrogen by compound II (Fe^{IV}–OH). This mechanism is closely related to that for the decarboxylation of fatty acids with concomitant abstraction of an H-atom on the β carbon to form the terminal alkene catalysed by the enzyme P450 OleTJE (CYP152L1).[353, 354]

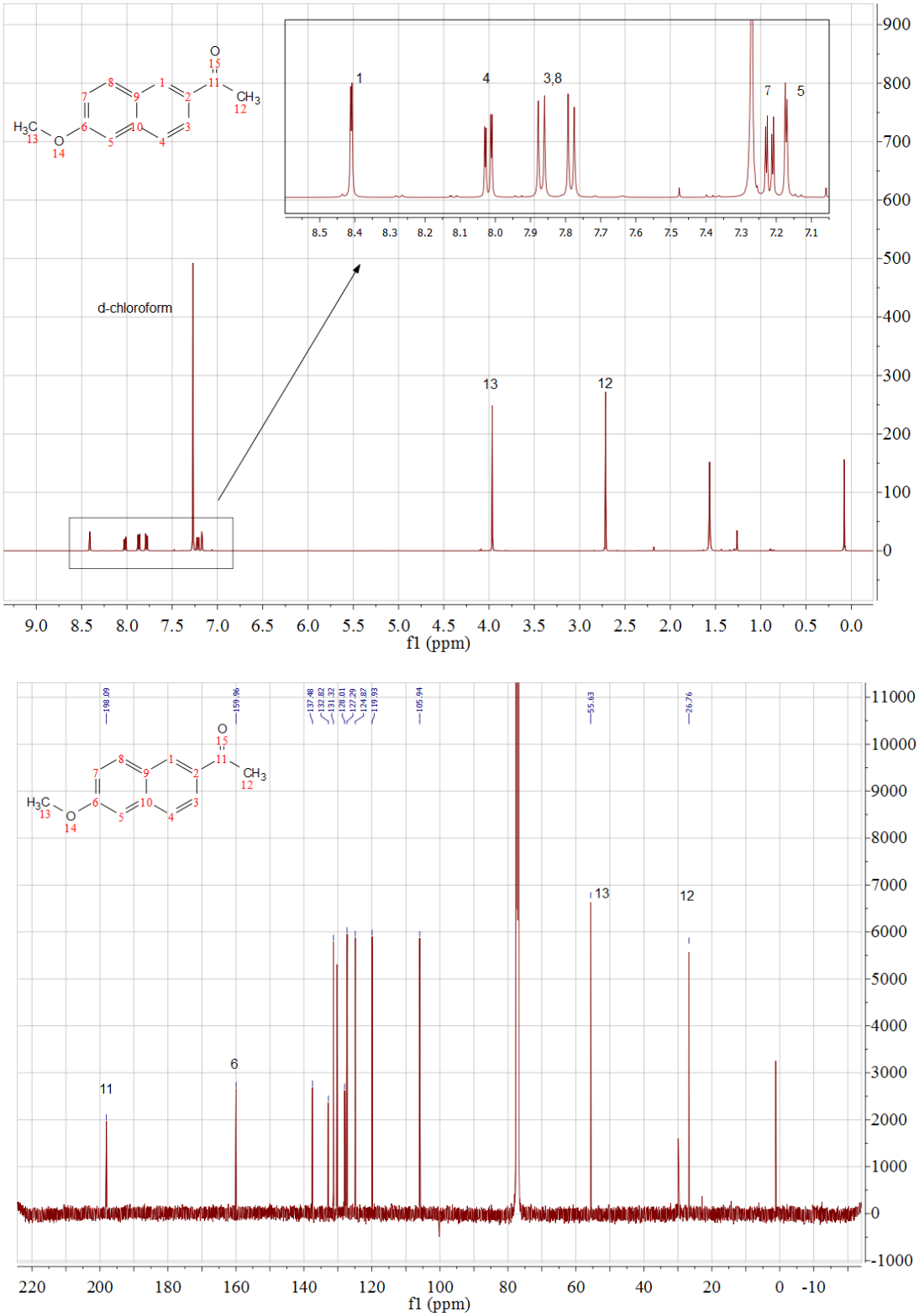


Figure 3.9 ^{1}H and ^{13}C NMR spectra of 2-acetyl-6-methoxynaphthalene (**3.04**)

In Figure 3.10, the β H-atom is from an O–H rather than a C–H bond. α-Hydroxynaproxen, the obligatory intermediate in this mechanism, was not observed in the reactions, indicating that it was rapidly oxidised. Indeed, α-hydroxylation of carboxylic acids is not a common P450 reaction type but is known for fatty acid oxidation by the CYP152A family of enzymes.[355] It is notable that **3.04** was formed by all of the mutants in our library (Table 3.2); even the wild type gave 20% of **3.04** at 5% conversion while the F87A mutant gave 48% (11% conversion). Some broad chemo-selectivity trends are apparent. Bulky substitutions (A328I, A328V, A330P) favoured decarboxylation while smaller side chains (P329V, I263A) promoted demethylation, for example, the RT2/AP/F81W mutant showed 68% selectivity for naproxen decarboxylation to **3.04** while RT2/AP/P329V gave 66% of **3.03**.

Directed evolution has been applied to $P450_{BM3}$ for naproxen oxidation.[206, 342] Five rounds of random mutagenesis and high-throughput screening starting from the 18-mutation variant 13C9 led to the 24-mutation variant X3H1. This mutant (1 μM) gave 35% conversion of naproxen (3 mM) to a mixture of **3.03** and the 5,8-quinone derivative of **3.03** from further oxidation. This latter product was not observed in our reactions. Since the 13C9 and X3H1 mutants also contained the F87A mutation, the other mutations probably altered naproxen binding such that **3.04** was not formed.[206]

3.3 Neutral drugs

3.3.1 Testosterone

Testosterone is a large neutral steroid hormone from the androgen group and is found in mammals, reptiles, birds, and other vertebrates that is primarily metabolised by the human liver CYP3A4, CYP2C9 and CYP2C19 to 1α-, 2α-/2β-, 6β-, 11β-, 15β- and 16β-hydroxy testosterone (Figure 3.11).[356, 357]

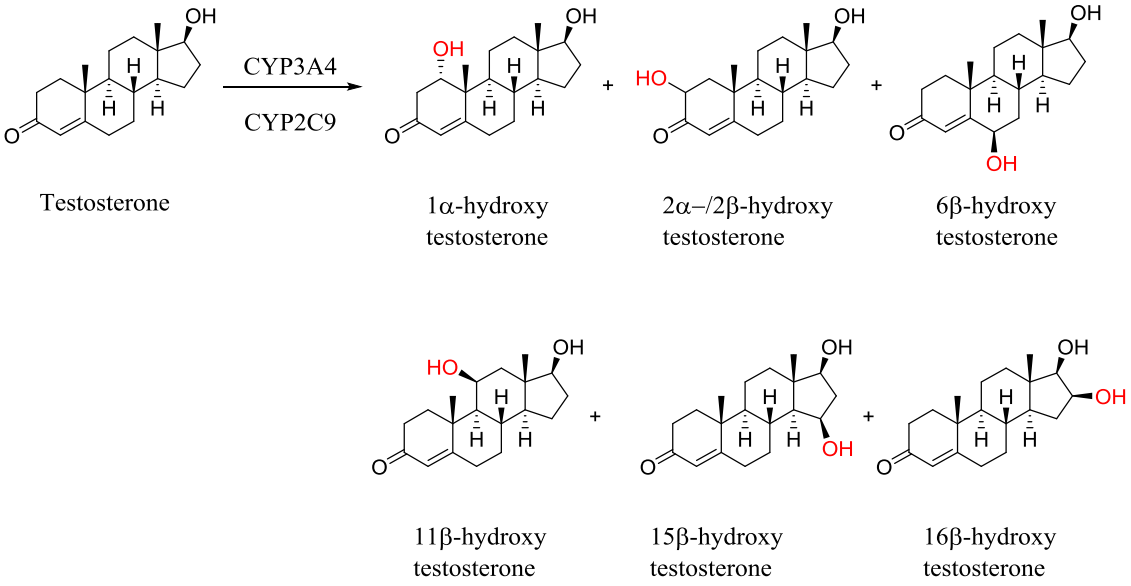
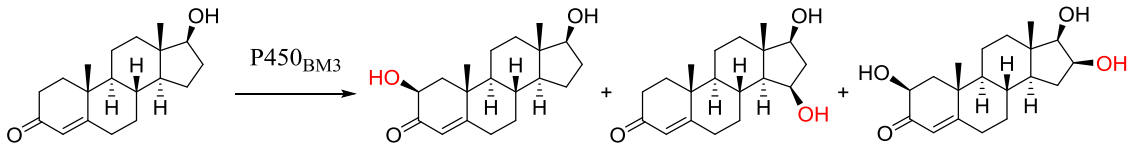


Figure 3.11 Testosterone and its human metabolites

The screening reactions were performed under the same conditions as the naproxen turnovers except for the enzyme concentration, which was 1 μM. The results are shown in Table 3.3. Due to the poor solubility of testosterone in aqueous system, both methanol and ethanol were tested as co-solvent for the addition of substrate and it turned out substrate concentration could only reach 500 μM with ethanol so that methanol was used as the co-solvent to make a 100 mM methanol stock. While previous work carried out by Dr Jake Yorke within our group has demonstrated four RP/FV/EV related variant enzymes (RP/FV/EV, RP/FV/EV/A328I, RP/FV/EV/V78F, RP/FV/EV/F81W) were active for testosterone oxidation the turnover number was not sufficient for product characterisation.[348]

Table 3.3 Product distribution and substrate conversion of the screen with testosterone



Testosterone | 2β-hydroxy testosterone, **3.05** | 15β-hydroxy testosterone, **3.06** | 2β, 16β-dihydroxy testosterone, **3.07**

No	Enzyme	Retention time /min [a]						Conversion [b]
		9.71 2*β*-OH (**3.05**)	**9.84** T-OH	**10.29** T-OH	**10.45** 15*β*-OH (**3.06**)	**10.83** 2*β*,16*β*-$(OH)_2$ (**3.07**)	**10.95** T-$(OH)_2$	
1	KSK19/QP	38%	2%	2%	43%	16%	-	100%
2	RLYF/KSK19	61%	3%	-	33%	3%	-	100%
3	KSK19/AM	10%	2%	-	86%	2%	-	98%
4	RLYF/KSK19/IP	12%	-	-	14%	61%	12%	91%
5	KSK19	-	-	-	60%	40%	-	90%
6	RT2/V78F	5%	-	-	95%	-	-	89%
7	KSK19/QP/FV	4%	-	-	96%	-	-	83%
8	RLYF/FA/L188V	55%	3%	-	42%	-	-	76%
9	F87A	60%	3%	-	37%	-	-	64%
10	KSK19/IA/QP	15%	-	-	82%	4%	-	61%
11	RLYF/FA/AG	38%	2%	-	60%	-	-	55%
12	KSK19/FS	40%	2%	-	58%	-	-	33%
13	RP/FV/EV	28%	5%	11%	56%	-	-	28%
14	KSK19/TA	43%	-	-	57%	-	-	24%
15	KSK19/IA	10%	-	-	90%	-	-	22%
16	RT2/IP/A184I	48%	10%	12%	30%	-	-	20%
17	RP/FV/EV/A184I	80%	-	-	13%	7%	-	20%
18	RT2/AP/A82L	23%	-	-	77%	-	-	13%
19	RP/AM	20%	-	-	80%	-	-	12%
20	RT2	-	-	-	100%	-	-	11%
21	WT	-	-	-	-	-	-	0%

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 µM P450, 1 mM testosterone. After 16 h the reaction mixture was extracted with ethyl acetate (300 µl). The phases were separated by centrifugation at 21,000 *g*. Testosterone metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. The product quantification was based on the integration of peak areas. The GC retention times represent the following products.
9.71 min product: 2β-hydroxytestosterone **3.05** (characterised by NMR and MS)
9.84 min product: monohydroxylation of testosterone, T-OH (MS data)
10.29 min product: monohydroxylation of testosterone, T-OH (MS data)
10.45 min product: 15β-hydroxytestosterone **3.06** (characterised by NMR and MS)
10.83 min product: 2β,16β-hydroxytestosterone **3.07** (characterised by NMR and MS)
10.95 min product: dihydroxylation of testosterone, $T(OH)_2$ (MS data)
[b] Conversion represents the average of 2 independent experiments with errors not higher than 10%.

The WT $P450_{BM3}$ did not show any detectable testosterone oxidation activity. Of the 86 variants screened, 17 showed >20% testosterone conversion and of these, only two (RT2/V78F and RT2/IP/A184I) did not contain a substitution of Phe87 with a smaller side chain (F87A, F87V, F87S). The RLYF/KSK19 mutant (F87A mutation) fully converted 1 mM testosterone to 61% 2β-hydroxytestosterone (**3.05**), 33% 15β-hydroxytestosterone (**3.06**), 3% of the new compound 2β,16β-dihydroxytestosterone (**3.07**) and 3% of a monohydroxylated testosterone (confirmed by MS data). On the other hand, the KSK19 mutant gave 60% **3.06** and 40% **3.07**. Therefore, the R47L/Y51F mutations biased the regioselectivity towards C2 oxidation to give more of **3.05** and disfavored further oxidation of **3.05** at C16 to form **3.07**. Addition of the I401P mutation to RLYF/KSK19 shifted the product mix further towards C2 oxidation (12% **3.05**, 61% **3.07**), with only 14% of **3.06** but a new product (12%) with a mass spectrum consistent with a dihydroxylated testosterone was detected. This is likely to be a further oxidation product of **3.06**, for example, 2β, 15β- or 15β, 16β-dihydroxytestosterone. The increase in the extent of further oxidation is consistent with the previously reported effect of the I401P mutation in promoting the oxidation of a wide range of organic compounds.[163, 358]

Remarkably, the KSK19/QP/FV mutant, in which the F87A mutation in KSK19 (F87A/H171L/Q307H/N319Y) was replaced with F87V, gave 96% selectivity of **3.06**. The RT2/V78F mutant, with V78F being the only active site mutation, gave 95% **3.06**. Other mutants with high selectivity for **3.06** included the KSK19/I263A (90%) and KSK19/A182M (86%). The V78, A82 and I263 side chains are within 6 Å of one another, suggesting that this region of the substrate pocket affects selectivity. Reetz and co-workers performed at V78 and A82 of the F87A mutant (8,700 transformants screened by HPLC) and identified the V78L/A82F/F87A mutant with 91% selectivity for **3.06**. Starting with the mutant

R47Y/T49F/F87A and screening led to the R47Y/T49F/V78L/A82M/F87A mutant which formed 96% **3.06**.[192]

The A82W mutation has been reported to promote testosterone 16β-hydroxylation (82%).[161] Although our mutant library did not give high selectivity for 16β-hydroxytestosterone, C16β oxidation was possible since the 2β,16β-diol **3.07** was formed in the reaction, for example, KSK19 (40%) and RLYF/KSK19/IP (61%). Within the present library the RP/FV/EV/A184I mutant showed the highest selectivity (80%) for C2β oxidation to give **3.05**, compared to the A330W/F87A mutant which gave 97 %.[192] It is notable from the effects of the V78F and A82W mutations that libraries of modest size (30–100 mutants) can produce high regio- and diastereo-selectivity.

The three major hydroxylation products were then isolated from three 57.6 mg reactions (0.2 mmol substrate, 200 mL reaction volume) with three different variants (RLYF/KSK19, KSK19/A82M and RLYF/KSK19/IP) at 1 μM concentration and the GC analysis of the three reactions are shown in Figure 3.12. The stereochemistry was characterised by the comparison of NMR spectra with published data in the literature.[192, 359] The ^{1}H and ^{13}C NMR spectra of compound **3.05**, **3.06**, **3.07** are shown in Figure 3.13, Figure 3.14, Figure 3.15, respectively.

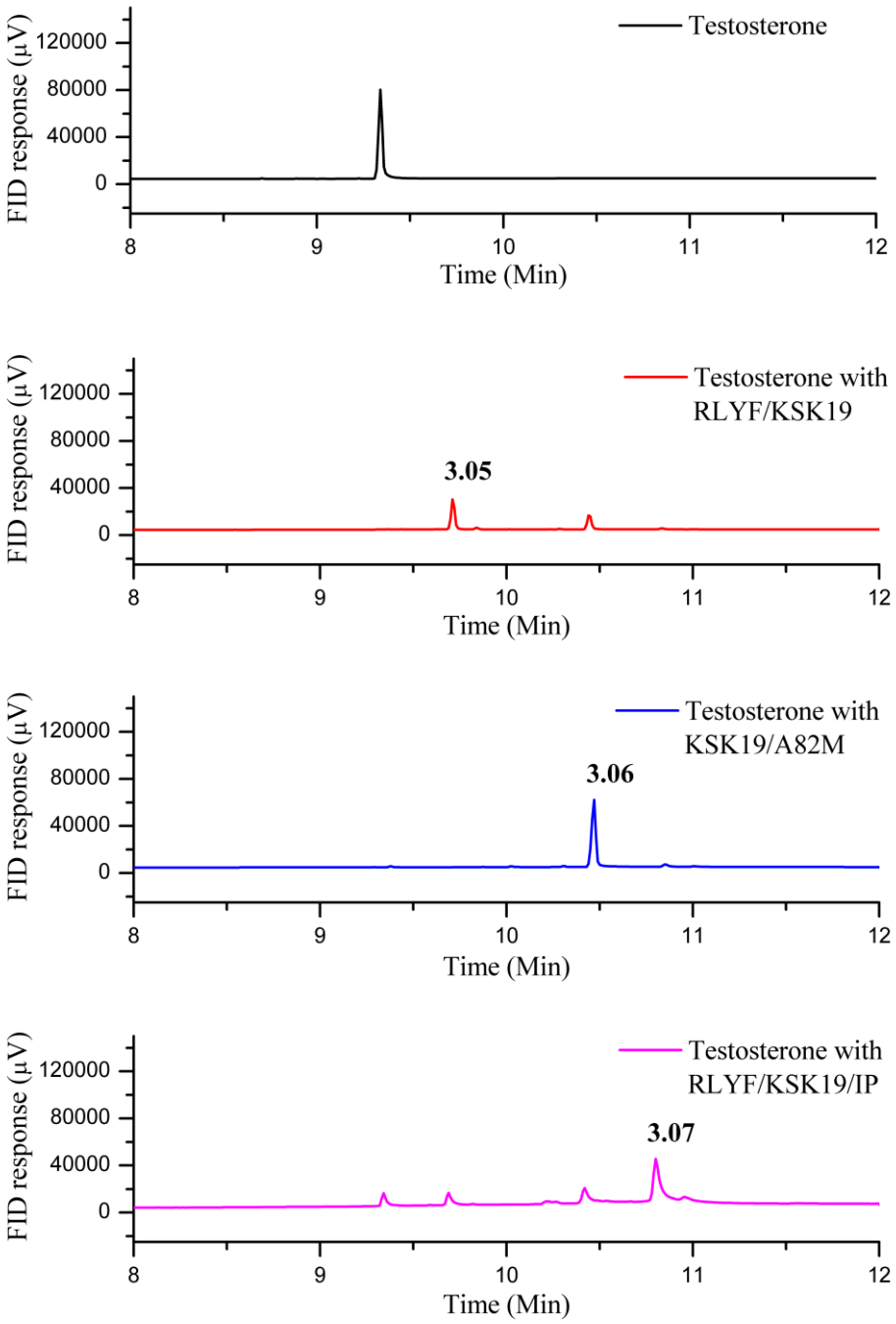


Figure 3.12 Gas chromatography analysis of 1 mM testosterone metabolised by 1 µM $P450_{BM3}$

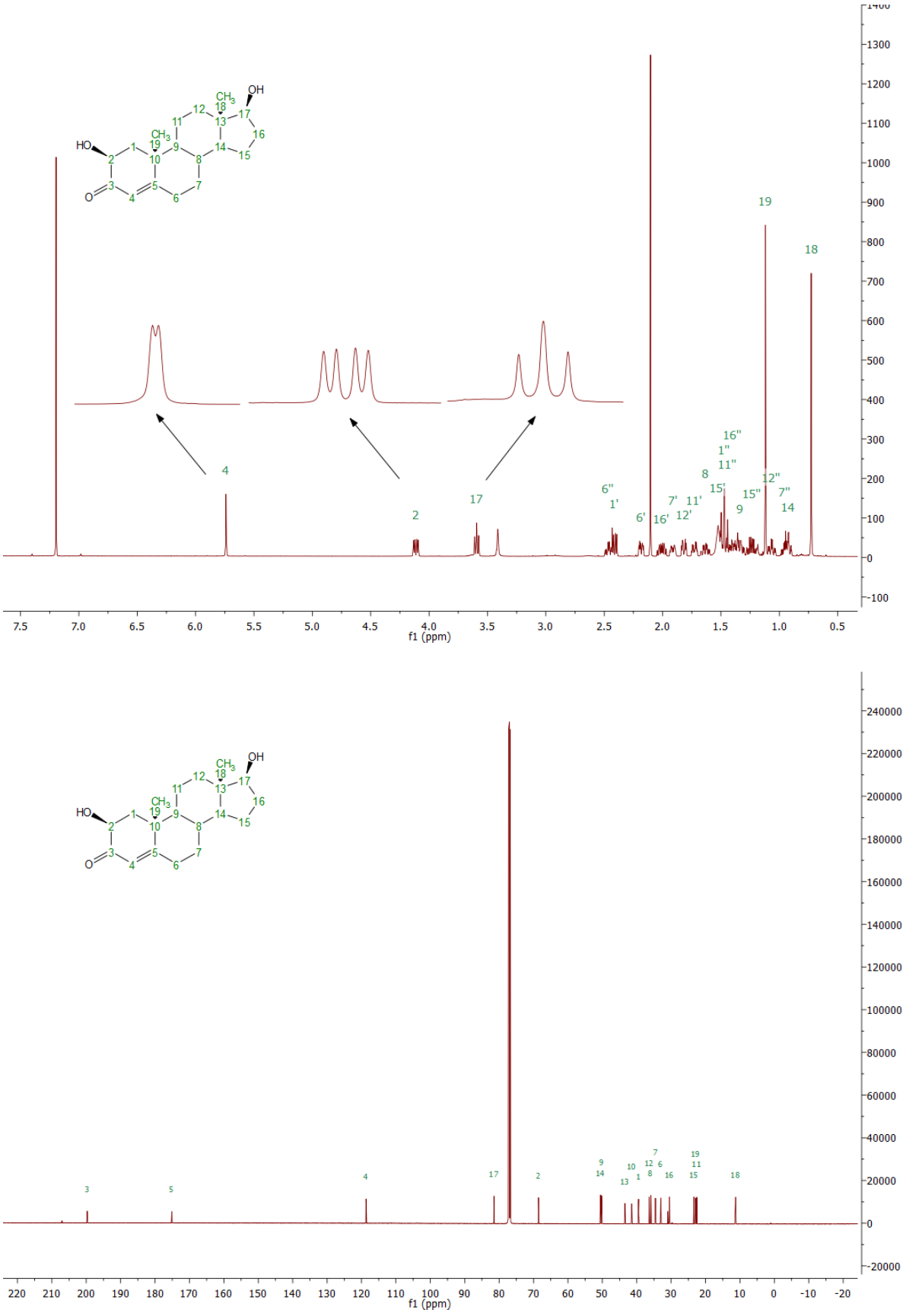


Figure 3.13 ^{1}H and ^{13}C NMR spectra of 2β-hydroxytestosterone, **3.05**

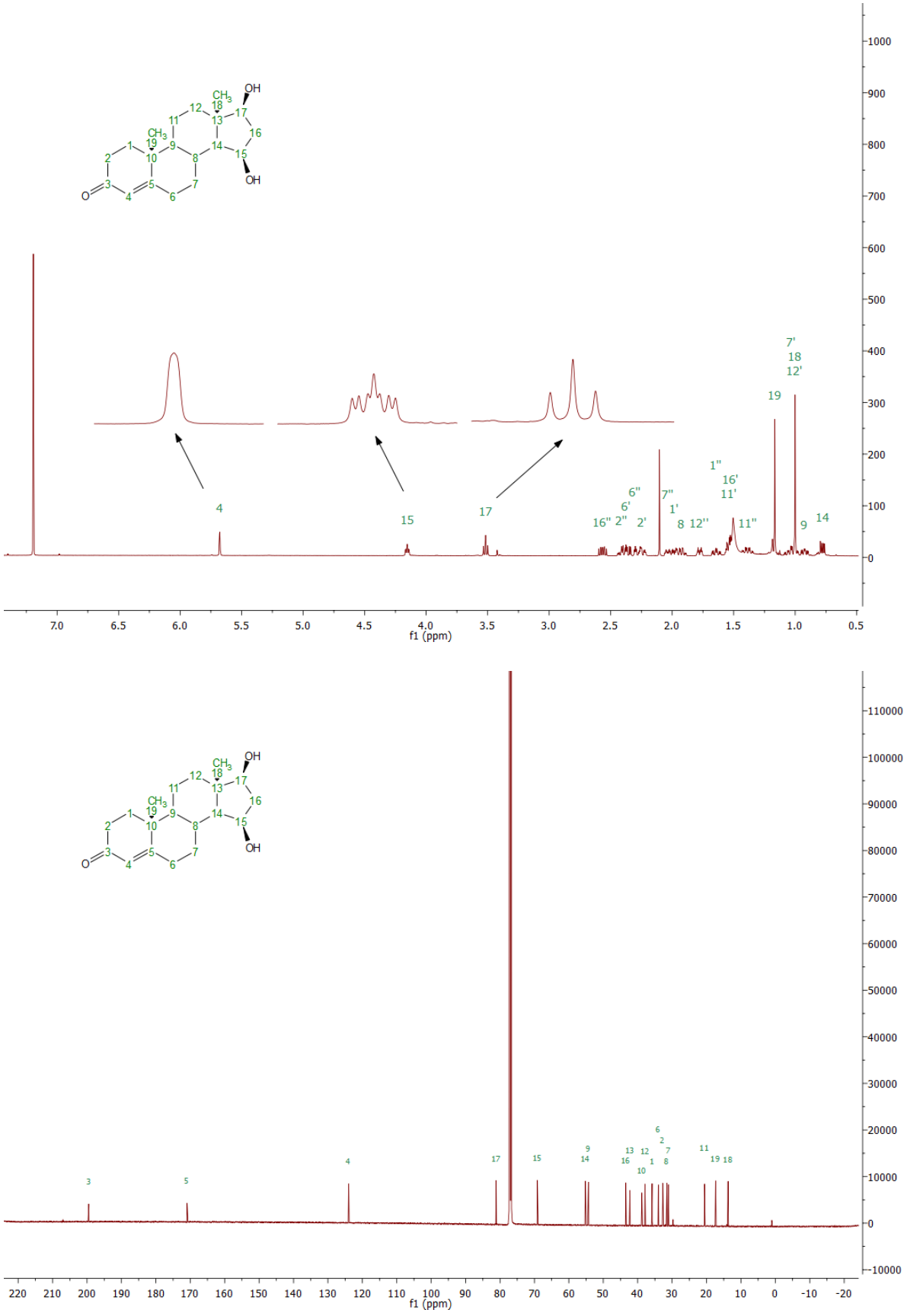


Figure 3.14 ^{1}H and ^{13}C NMR spectra of 15β-hydroxytestosterone, **3.06**

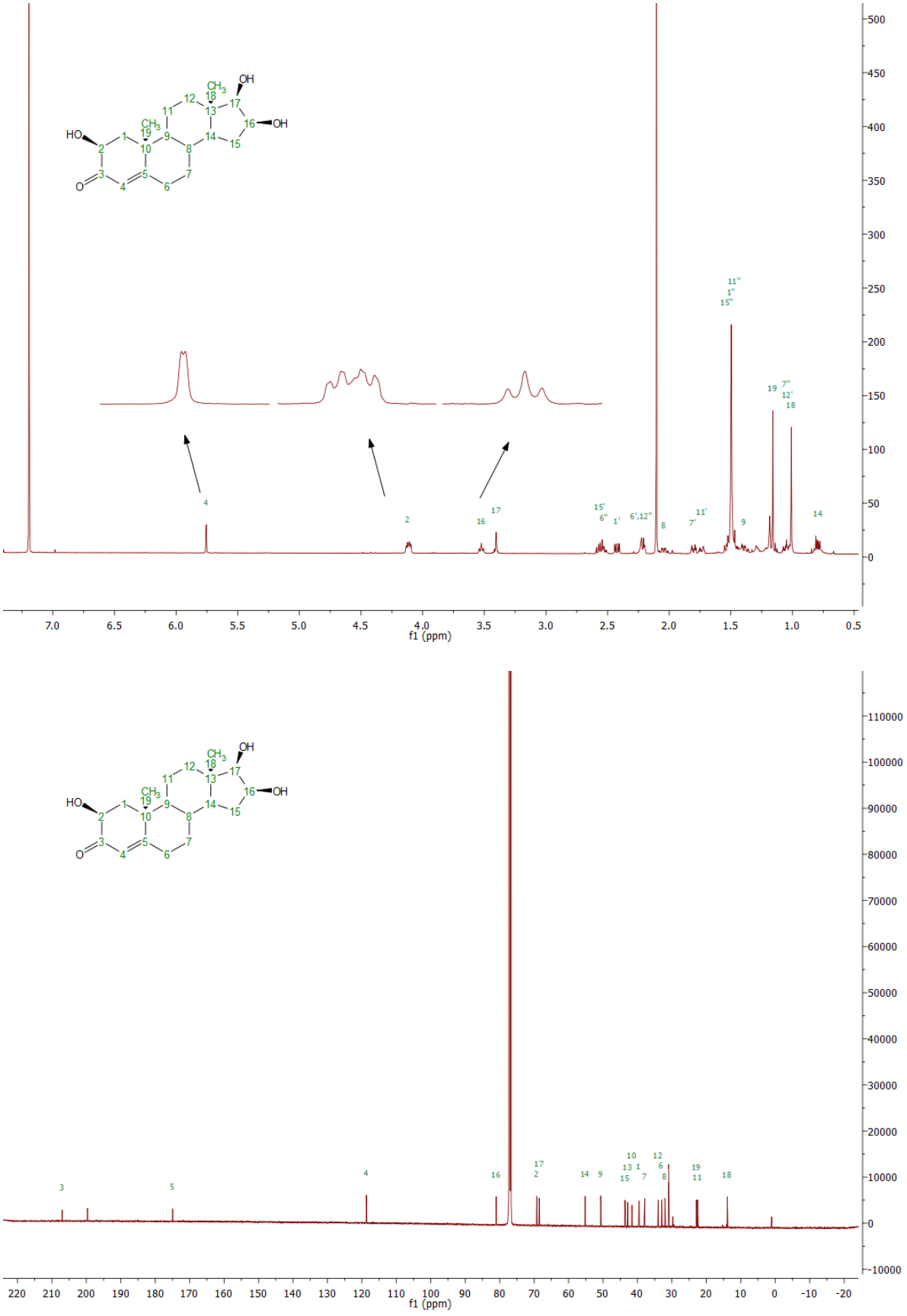
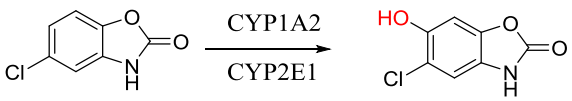


Figure 3.15 ^{1}H and ^{13}C NMR spectra of 2β,16β-dihydroxytestosterone, **3.07**

3.3.2 Chlorzoxazone

Chlorzoxazone is a centrally acting muscle relaxant used to treat muscle spasm and the resulting pain or discomfort. It is primarily metabolised by CYP1A2 and CYP2E1 in human into 6-hydroxychlorzoxazone (Figure 3.16).[360] Although a small 2-ring compound chlorzoxazone proved challenging even compared to the much larger testosterone.



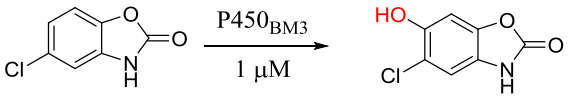
Chlorzoxazone 6-hydroxychlorzoxazone

Figure 3.16 Chlorzoxazone and its human metabolite, 6-hydroxychlorzoxazone

The initial screening was conducted in 200 mM phosphate buffer (pH 7.5), with 1 μM enzyme and 2 mM chlorzoxazone being added from a 100 mM stock in MeOH while the conditions for NADPH regeneration system remained the same as diclofenac, naproxen and testosterone. The substrate conversion tests are shown in Table 3.4.

From a screen of 52 variants 12 was found with >20% chlorzoxazone conversion, of which six showed >50% conversion (Table 3.4). In all cases **3.08** was the only product, which was isolated from a 67.6 mg preparative scale reaction (0.4 mmol substrate, 200 mL reaction volume) and fully characterised by NMR and high resolution MS.

Table 3.4 The initial screening results towards chlorzoxazone



Chlorzoxazone 6-hydroxychlorzoxazone, **3.08**

Number	Enzyme	Conversion [a]
1	RT2/AP/AI	100%
2	RT2/AW/AM	97%
3	RPAP	89%
4	RT2/AP/AL	69%
5	RP/IA/EV	53%
6	RT2	51%
7	RPAP/PV	41%
8	RP/FV	33%
9	RT2/IP/AP	28%
10	KSK19/AI/IP	28%
11	RP	23%
12	RT2/AP/F81W	21%
13	RT2/AP/PV	15%
14	QP	12%
15	WT	0%

[a]. Chlorzoxazone metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. Product quantification was based on the integration of peak areas. Conversion represents the average of 2 independent experiments with errors not higher than 10%.

The most active mutants contained bulky substitutions at A330; the RT2/AP/A184I and RT2/A330W/A82M fully converted 2 mM chlorzoxazone (Figures 3.17). A previous directed evolution study of $P450_{BM3}$ identified the mutant F162L/M185T/L188P/M237I with about 40% conversion of 200 μM chlorzoxazone to **3.08**.[212] None of these four mutations was in the active site. Of the library the most active mutant without a substrate pocket substitution was RT2 (51% conversion). Addition of bulky substitutions to RT2 at active site residues, especially at A330, gave 100% conversion.

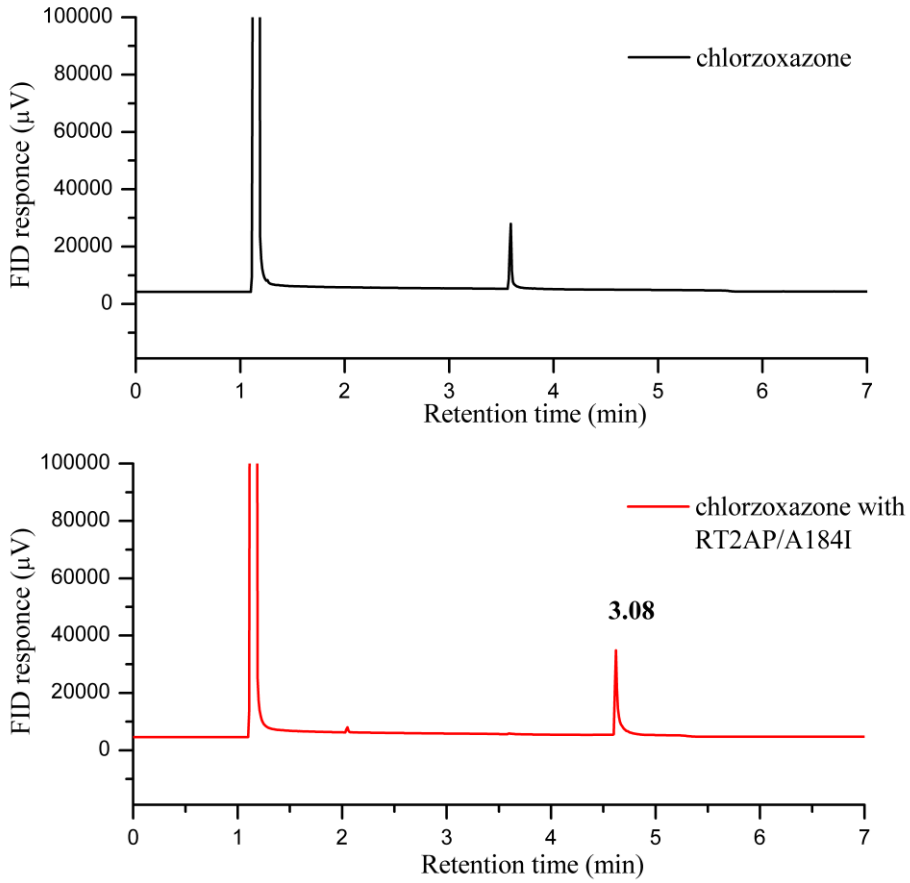


Figure 3.17 GC analysis of 2 mM chlorzoxazone metabolised by 1 μM RT2/AP/A184I

3.4 Cationic drugs

3.4.1 Amitriptyline

Amitriptyline is cationic and the most commonly used tricyclic antidepressant. It is metabolised initially to nortriptyline, *via* demethylation and to hydroxyamitriptyline by hydroxylation, and further to (*E*)- and (*Z*)-10-hydroxynortriptyline (Figure 3.18).[361]

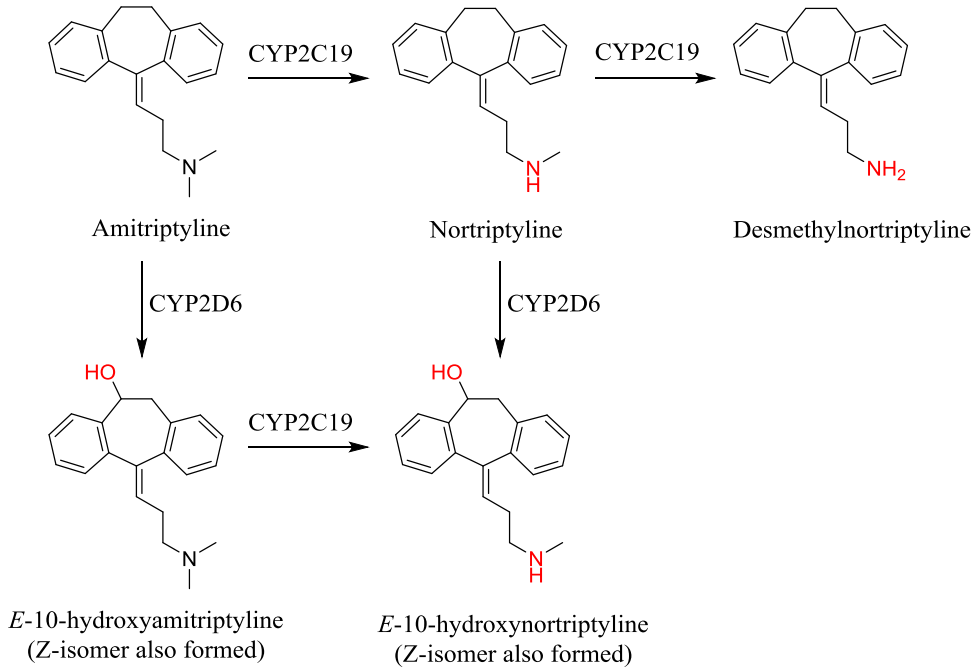
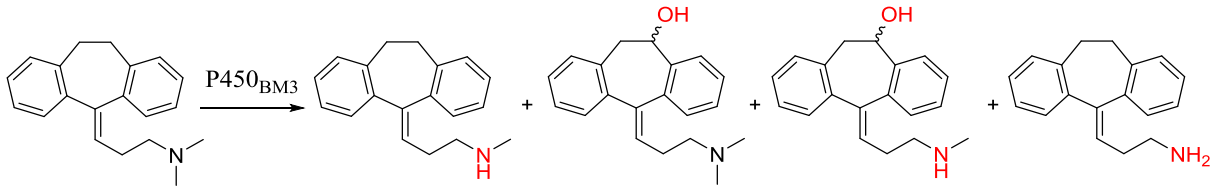


Figure 3.18 Amitriptyline and its human metabolites

The screen was performed with 65 variants in 50 mM Tris-HCl buffer (pH 8.5) while other conditions remained the same as for chlorzoxazone. The reaction mixtures were adjusted to pH>11 prior to the extraction with ethyl acetate. The organic was analysed by gas chromatography (Table 3.5). 23 variants showed >20% amitriptyline conversion and the most active mutant gave 79% conversion. The base mutants RT2, KSK19 and RP had high activity, and the A330P mutation in combination with RT2 also gave high conversion. Even the F87A single-site mutant showed 17% conversion while the RP/FV and RP/FV/EV converted 46% and 21% amitriptyline, respectively.

Table 3.5 Product distribution and substrate conversion of the screen with amitriptyline



Amitriptyline | nortriptyline **3.09** | hydroxyamitriptyline **3.10** | *Z*-10-hydroxynortriptyline **3.11** | desmethylnortriptyline **3.12**

Number	Enzyme	Retention Time /min [a]					Conversion [b]
		6.68 (3.12)	6.91	7.09 (3.09)	7.90 (3.10)	7.96 (3.11)	
1	RT2/AP/VI	1%	3%	44%	50%	5%	79%
2	KSK19	3%	3%	82%	5%	7%	75%
3	RP	-	10%	82%	6%	3%	65%
4	RT2/AP/AI	-	5%	62%	30%	4%	55%
5	RP/FV	7%	4%	83%	6%	-	46%
6	RT2/IP/AP	-	6%	89%	4%	1%	40%
7	RT2	2%	5%	87%	4%	2%	43%
8	RT2/AP/AL	-	-	77%	16%	7%	38%
9	RP/FW	2%	5%	86%	5%	3%	33%
10	KSK19/IA	-	2%	98%	-	-	32%
11	RP/FV/EV	5%	7%	83%	4%	1%	32%
12	RT2/FW/AI	-	-	95%	5%	-	29%
13	RP/FV/EV/A184I	-	3%	97%	-	-	29%
14	RP/EV	-	4%	89%	5%	2%	27%
15	KSK19/AI/YL	1%	4%	89%	5%	1%	27%
16	RT2/FW	-	3%	94%	2%	-	27%
17	RT2/AW/AM	-	-	67%	25%	8%	26%
18	RT2/IP	-	5%	`	3%	-	26%
19	RP/AM	-	-	90%	6%	3%	25%
20	QP	2%	4%	87%	7%	-	22%
21	RT2/AP/AM/AI	-	2%	84%	12%	3%	21%
22	RT2/AP/PV	24%	-	64%	-	12%	21%
23	RP/IA/EV	-	13%	87%	-	-	21%
24	RT2/IP/AI	2%	4%	89%	6%	-	19%
25	KT2	-	-	100%	-	-	18%
26	RLYF/KSK19	-	-	100%	-	-	17%
27	RT2/AW	2%	3%	91%	4%	-	17%
28	RP/FV/EV/VF	-	6%	91%	4%	-	16%
29	FA	-	-	100%	-	-	16%
30	KSK19/QP/FV	-	-	100%	-	-	15%
31	KSK19/AM	6%	-	82%	9%	3%	14%
32	RT2IP/EV	-	7%	89%	3%	-	14%
33	RP/HL	-	3%	91%	6%	-	13%
34	KSK19/QP	-	-	100%	-	-	12%

35	RLYF/KSK19/IP	-	-	100%	-	-	12%
36	KSK19/AI/FV	18%	-	70%	11%	-	12%
37	GVQ	4%	-	96%	-	-	11%
38	KSK19/AI/AI	-	3%	92%	3%	2%	11%
39	KSK19/IA/QP	-	-	100%	-	-	11%
40	RT2/IP/FW	-	-	100%	-	-	11%
41	RT2/AP/VI/AI	4%	-	66%	30%	-	10%
42	RPAP/AV	61%	-	39%	-	-	10%
43	WT	34%	-	66%	-	-	9%
44	KSK19/AI/FL	-	-	90%	8%	2%	9%
45	RQA	7%	-	93%	-	-	9%
46	RP/AM/IA	-	-	100%	-	-	8%
47	RT2/AP/AM	45%	-	55%	-	-	7%
48	QP/AM	-	-	100%	-	-	7%
49	RT2/AP/AI/SA	75%	-	25%	-	-	7%
50	RT2/AP/AI/FW	-	-	87%	13%	-	7%
51	KSK19/AI/YV	10%	-	90%	-	-	6%
52	RPAP	47%	-	53%	-	-	6%
53	RP/FV/EV/A328I	-	-	100%	-	-	6%
54	KSK19/AI/IP	6%	-	94%	-	-	6%
55	RLYF/KSK19/AI	-	-	94%	6%	-	5%
56	RPAP/AM	37%	-	63%	-	-	5%
57	RP/FV/EV/LQ	-	13%	87%	-	-	5%
58	TM/IP	-	-	100%	-	-	5%
59	RPAP/AM/AI	-	-	100%	-	-	5%
60	KSK19/AI	3%	-	86%	8%	3%	4%
61	KSK19/FS	-	-	100%	-	-	4%
62	FA/TA/IP	-	-	100%	-	-	4%
63	AS/PS/QP	-	-	83%	17%	-	4%
64	KSK19/TA	-	-	100%	-	-	3%
65	KSK19/AI/IA	-	-	100%	-	-	3%

[a] The screening reaction mixture (0.5 mL) contained 50 mM Tris–HCL, pH 8.5, 1 μM P450, 2 mM amitriptyline. After 16 h the reaction mixture was adjusted to pH>12 with 5 M KOH and extracted with diethyl ether (300 μl). The phases were separated by centrifugation at 21,000 *g*. Amitriptyline metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. Product quantification was based on the integration of peak areas. The GC retention times represent for the following products:

6.68 min product: desmethylnortriptyline *via* double demethylation of amitriptyline (determined by MS)

6.91 min product: unidentified

7.09 min product: nortriptyline **3.09** (characterised by NMR and MS)

7.90 min product: mono-hydroxylation of amitriptyline **3.10** (characterised by NMR and MS)

7.96 min product: mono-hydroxylation of nortriptyline **3.11** (MS data)

[b] Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

Preparative scale reactions (Figure 3.19) and product isolation (110.8 mg, 0.4 mmol substrate, 200 mL reaction volume, performed with chloroform/methanol system) showed that the human metabolite nortriptyline (**3.09**) was the most common product. The KSK19 mutant gave 82% **3.09**. However, the F87A and F87V mutations strongly biased the product towards **3.09** while mutations added to A330P-containing variants significantly altered product selectivity. The RT2/AP/V78I mutant gave just 44% **3.09** but 50% (Z)-10-hydroxyamitriptyline (**3.10**) and two other products for which the MS data ($[M+H]^+ = 280.17$) were consistent with hydroxynortriptyline (**3.11**, 5%) and the double demethylation product desmethylnortriptyline (**3.12**, 1%), which were identified by comparison with literatures.[362] On the other hand, the RT2/AP/A184I gave only 30% **3.10** while the RT2/AP/P329V formed the highest proportions of **3.11** (12%) and **3.12** (24%) amongst the most active mutants. Interestingly, the RT2/AP/A184I/S72A mutant gave only the demethylation products **3.09** (25%) and **3.12** (75%), albeit at a low conversion of 7%, suggesting a role for the S72 side chain in binding the amine group of amitriptyline.

HPLC-MS analysis in the only report of amitriptyline oxidation by $P450_{BM3}$ showed that a series of ten F87V-based mutants gave **3.09** as the majority product (53–71%) together with 10–15% of **3.12**. However, the compounds were not isolated, and hydroxyl-nortriptyline (**3.11**) was not formed in those reactions.[213]

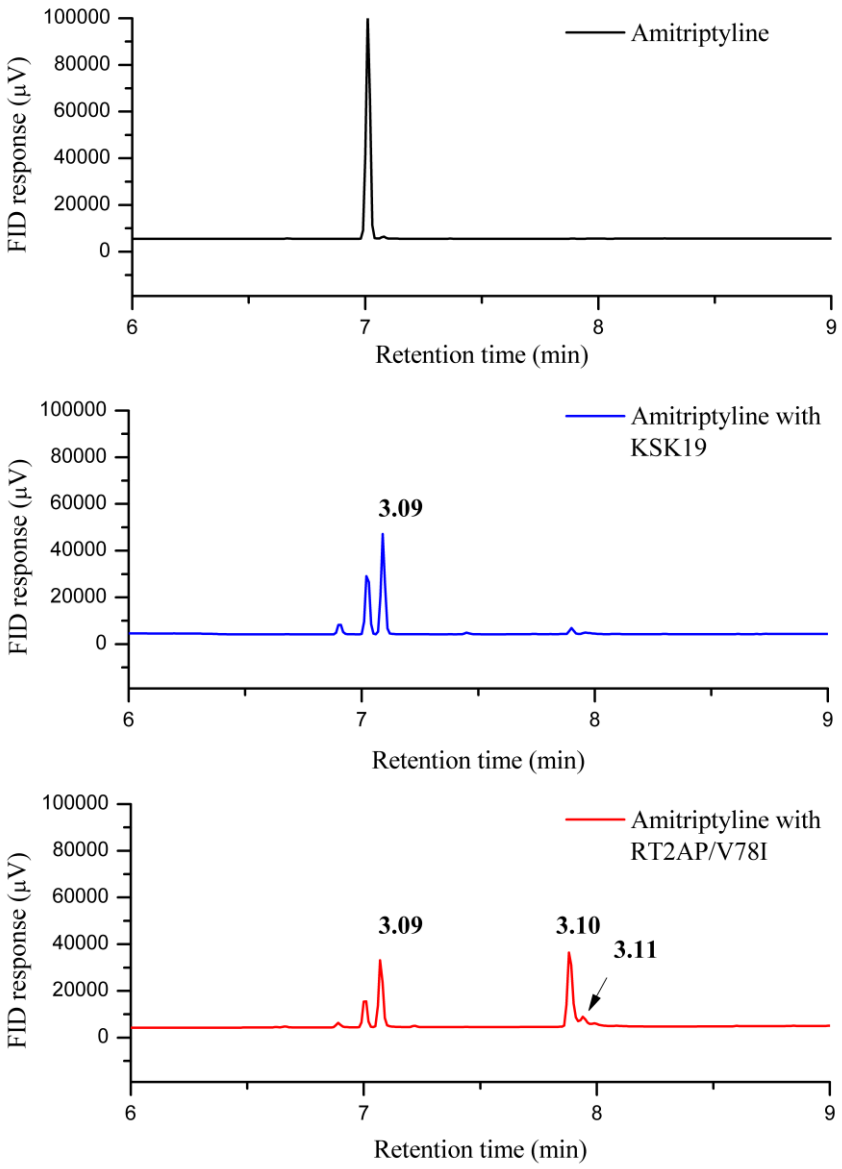
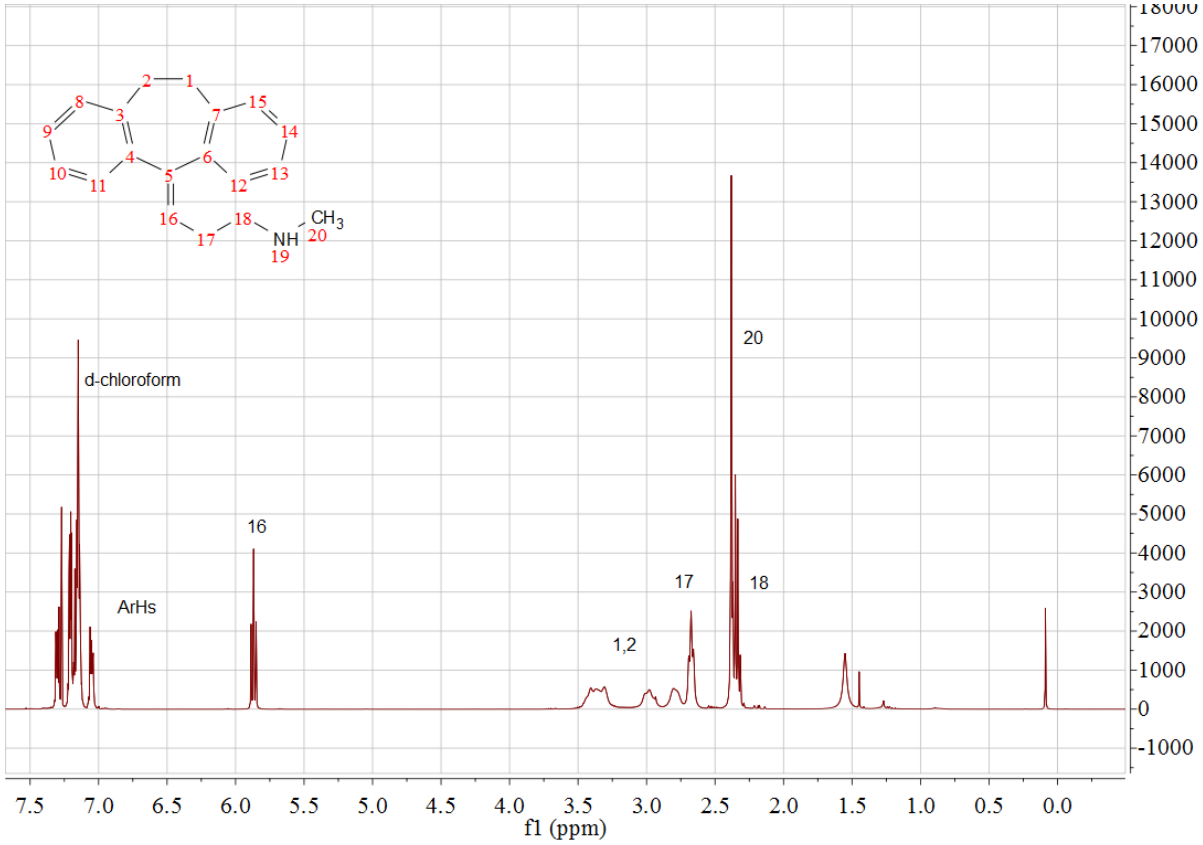


Figure 3.19 GC analysis of 2 mM amitriptyline metabolised by 1 μM $P450_{BM3}$



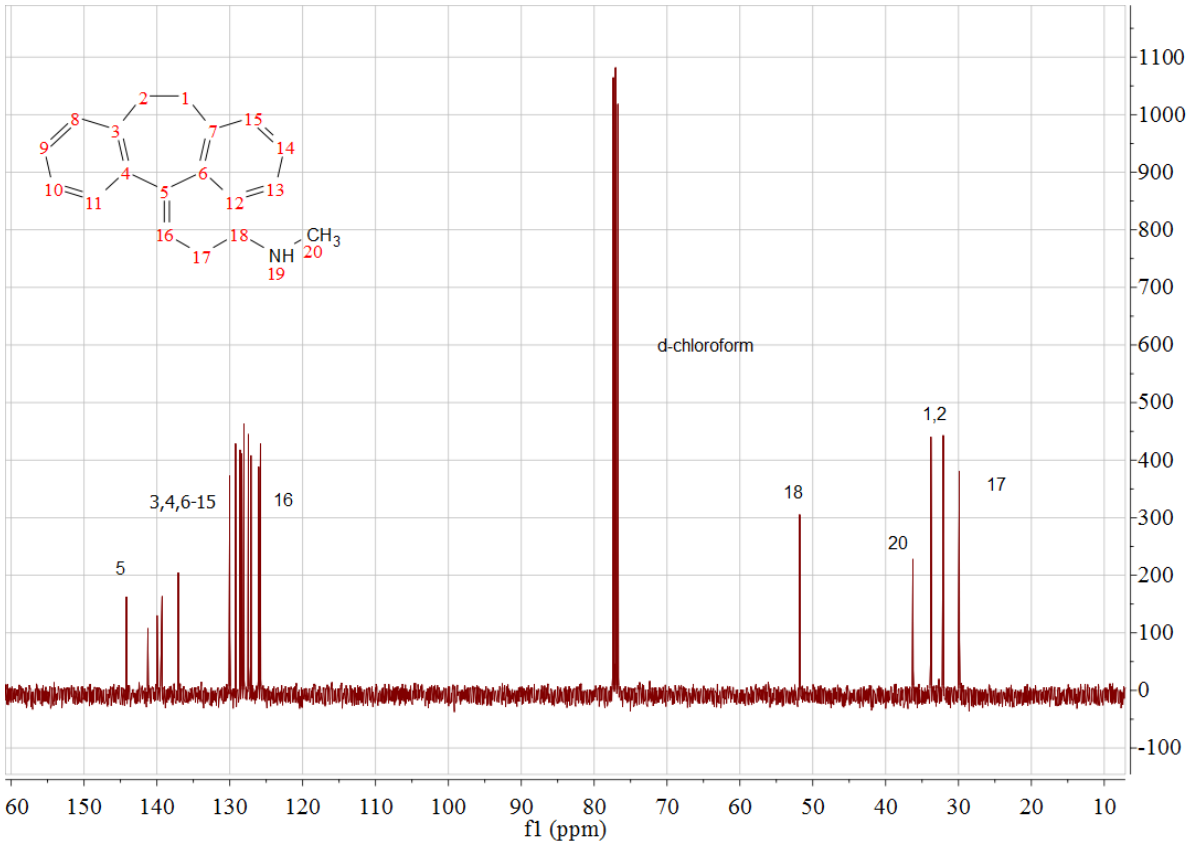


Figure 3.20 ^{1}H and ^{13}C NMR spectra of nortriptyline, **3.09**

3.4.2 Lidocaine

Lidocaine is a common local anesthetic and antiarrhythmic drug, which is primarily metabolised by human CYP3A4 to monoethylglycinexylidide (MEGX) and then glycinexylidide (GX) *via* dealkylation (Figure 3.21),[363] and is of particular interest, as it has not yet been reported as a substrate for $P450_{BM3}$.

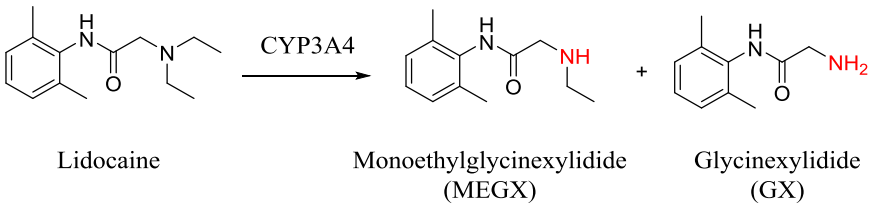
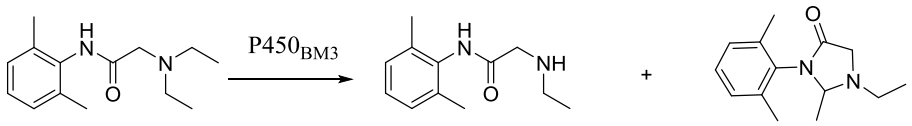


Figure 3.21 Lidocaine and its human metabolites, MEGX and GX

The screening was carried out with 45 variants under the same conditions as amitriptyline. Substrate conversions and product distributions are shown in Table 3.6.

Of 45 variants screened, 24 gave >50% lidocaine conversion. Mutants possessing enlarged (e.g. F87A and F87V) or constrained (e.g. A330P) active sites all showed high activity. The RP/FV/EV/F81W mutant showed 93% conversion and gave 96% of the human metabolite MEGX (**3.13**) which was isolated from a 46.8 mg reaction (0.2 mmol substrate, 100 mL reaction volume) with the RP/FV/EV mutant (1 μM) with 85% isolated yield (39.8 mg). However, the RP/FV/EV mutant formed just 11% of **3.13** but 84% of another compound which was also the majority product for all other variants (GC trace shown in Figure 3.22). This compound was purified from a 46.8 mg reaction (0.2 mmol substrate, 100 mL reaction volume) with the RT2/A330W mutant (1 μM) and characterized as the cyclic *N*,*N*-acetal compound (Z)-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline (**3.14**, 98% isolated yield), the ^{1}H NMR and ^{13}C NMR of which are shown in Figure 3.23.

Table 3.6 Product distribution and substrate conversion of the screen with lidocaine



Lidocaine | monoethylglycinexylidide MEGX, **3.13** | *Z*-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline, **3.14**

Number	Enzyme	Retention time /min [a]			Conversion [b]
		4.29 (3.13)	4.33	4.48 (3.14)	
1	RT2/AW	-	-	100%	100%
2	KSK19/IA	3%	1%	96%	99%
3	RT2/IP	7%	-	93%	97%
4	QP/AM	-	-	100%	95%
5	R19	-	-	100%	95%
6	RP/FV/EV/F81W	96%	4%	-	93%
7	RP/FV/EV	11%	5%	84%	92%
8	RP/AP/AM	-	-	100%	92%
9	RLYF/KSK19	4%	2%	94%	90%
10	RT2/AP	-	-	100%	89%
11	RT2	-	-	100%	87%
12	KSK19/QP	4%	2%	94%	84%
13	RT2AP/AL	-	-	100%	84%
14	RP/EV	-	-	100%	81%
15	RP/FV	-	-	100%	75%
16	RP/FV/EV/A184I	14%	9%	77%	68%
17	KSK19/AM	-	-	100%	68%
18	RLYF/KSK19/AI	-	-	100%	68%
19	FA/IP	-	-	100%	62%
20	KSK19/AI/FV	-	-	100%	61%
21	KSK19/AI/YL	-	-	100%	61%
22	KSK19/AI/YV	-	-	100%	59%
23	KSK19/AI/AI	-	-	100%	55%
24	RT2/AP/AI	31%	-	69%	49%
25	RT2/AP/AM	-	-	100%	46%
26	RT2/AP/AI/SA	-	-	100%	46%
27	KSK19/AI/IP	-	-	100%	46%
28	KT2	-	-	100%	43%
29	FA	-	-	100%	40%
30	KSK19/AI/FL	-	-	100%	38%
31	KSK19/AI	-	-	100%	31%
32	KSK19/AI/YF	-	-	100%	28%
33	KSK19	75%	-	25%	27%
34	RT2/FW/AI	-	-	100%	26%
35	RT2/AP/VI	-	-	100%	25%

36	QP	-	-	100%	24%
37	RP/IA/EV	5%	11%	84%	23%
38	RT2/AP/FW	100%	-	-	23%
39	GVQ	-	-	100%	14%
40	RP	13%	10%	77%	13%
41	A330P	-	-	100%	11%
42	RP/AP/PV	-	-	100%	9%
43	RP/AM/IA	-	-	100%	9%
44	TM/IP	-	-	100%	5%
45	RPAP/AW	-	-	100%	3%

[a] The screening reaction mixture (0.5 mL) contained 50 mM Tris–HCl, pH 8.5, 1 μM P450, 2 mM lidocaine. After 16 h the reaction mixture was adjusted to pH>12 with 5 M KOH and extracted with diethyl ether (300 μl). The phases were separated by centrifugation at 21,000 *g*. Lidocaine metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. Product quantification was based on the integration of peak areas. The GC retention times represent for the following products:
4.29 min product: monoethylglycinexylidide (MEGX) **3.13** from demethylation of lidocaine (characterised by NMR and MS)
4.33 min product: glycinexylidide (GX) from the double demethylation of lidocaine (MS data)
4.48 min product: *N,N*-acetal compound (*Z*)-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline, **3.14** (characterised by NMR and MS)
[b] Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

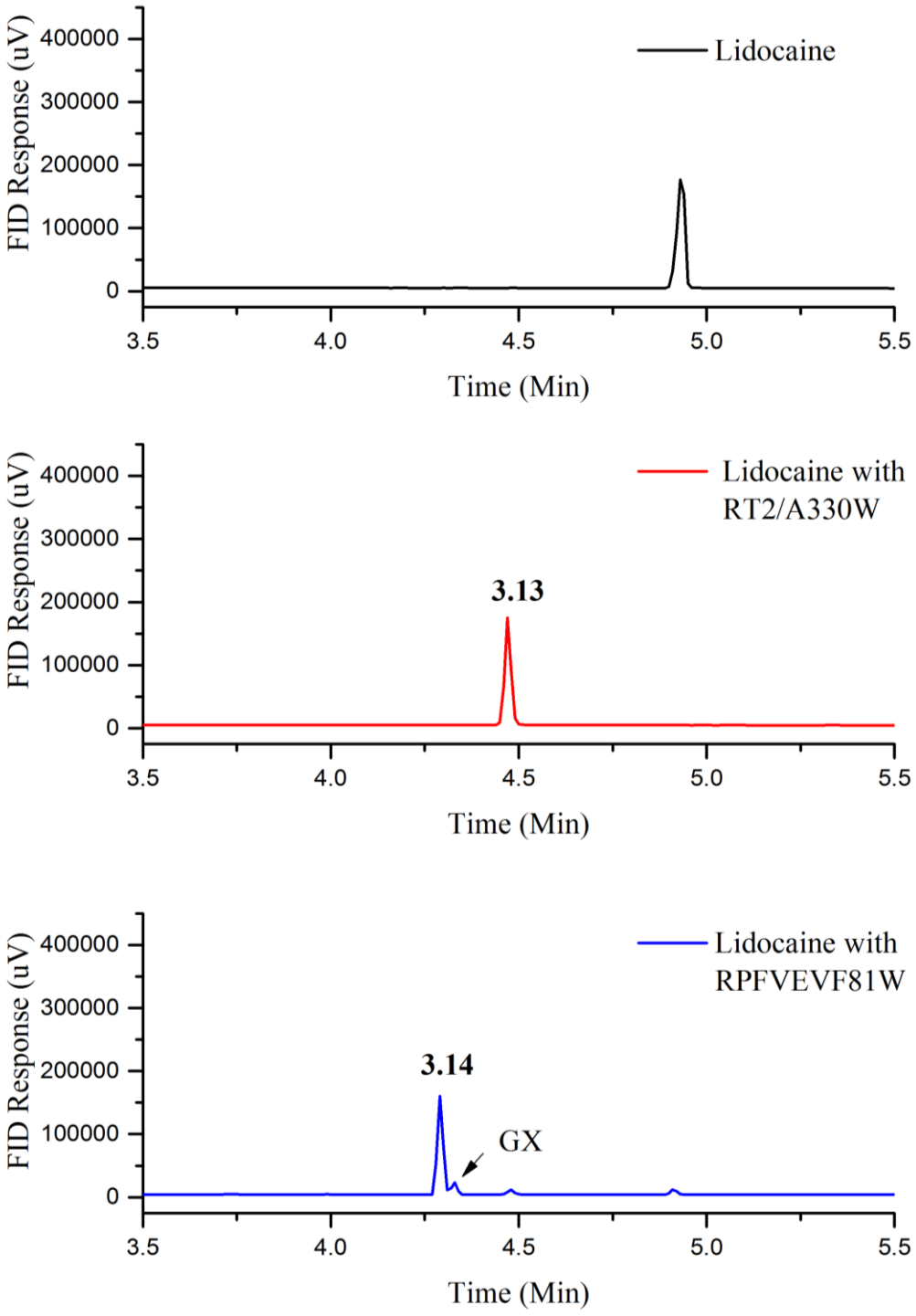
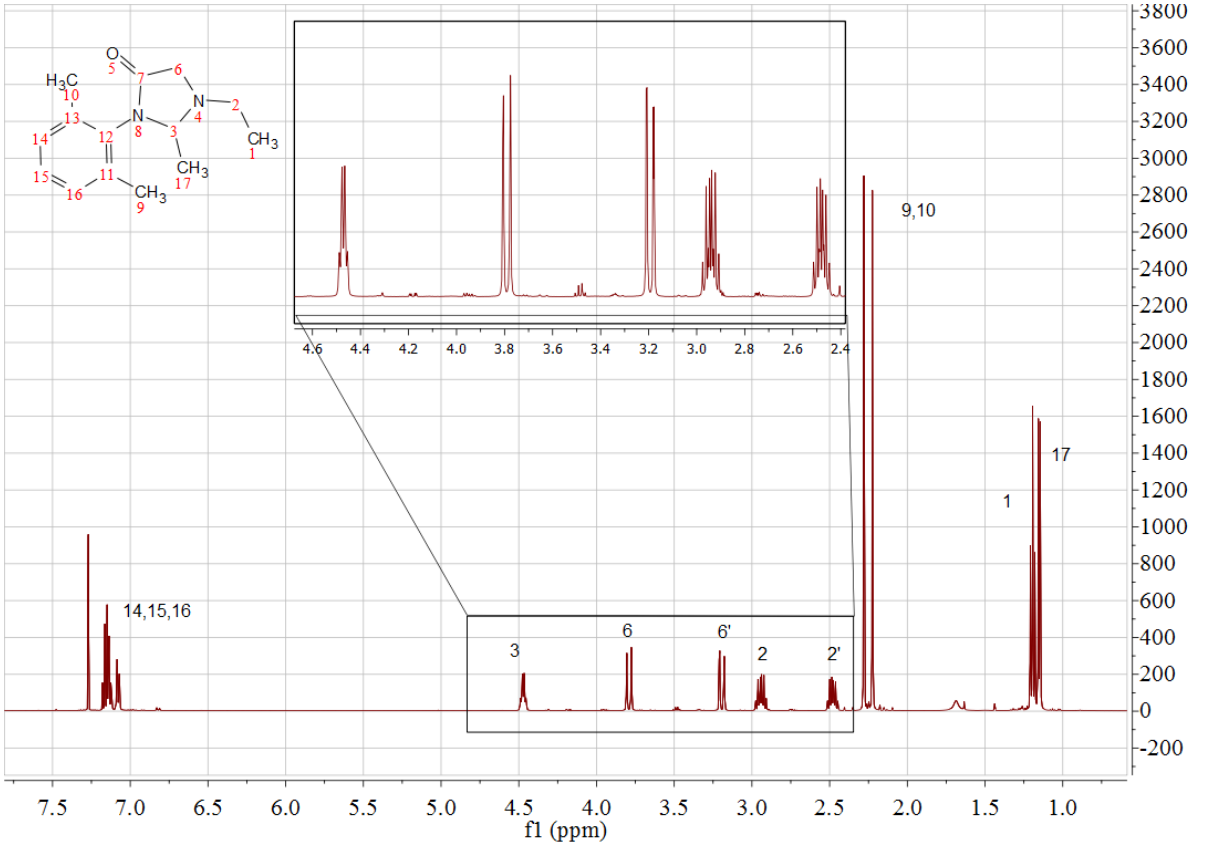


Figure 3.22 Gas chromatography analysis of 2 mM lidocaine metabolised by 1 μM $P450_{BM3}$



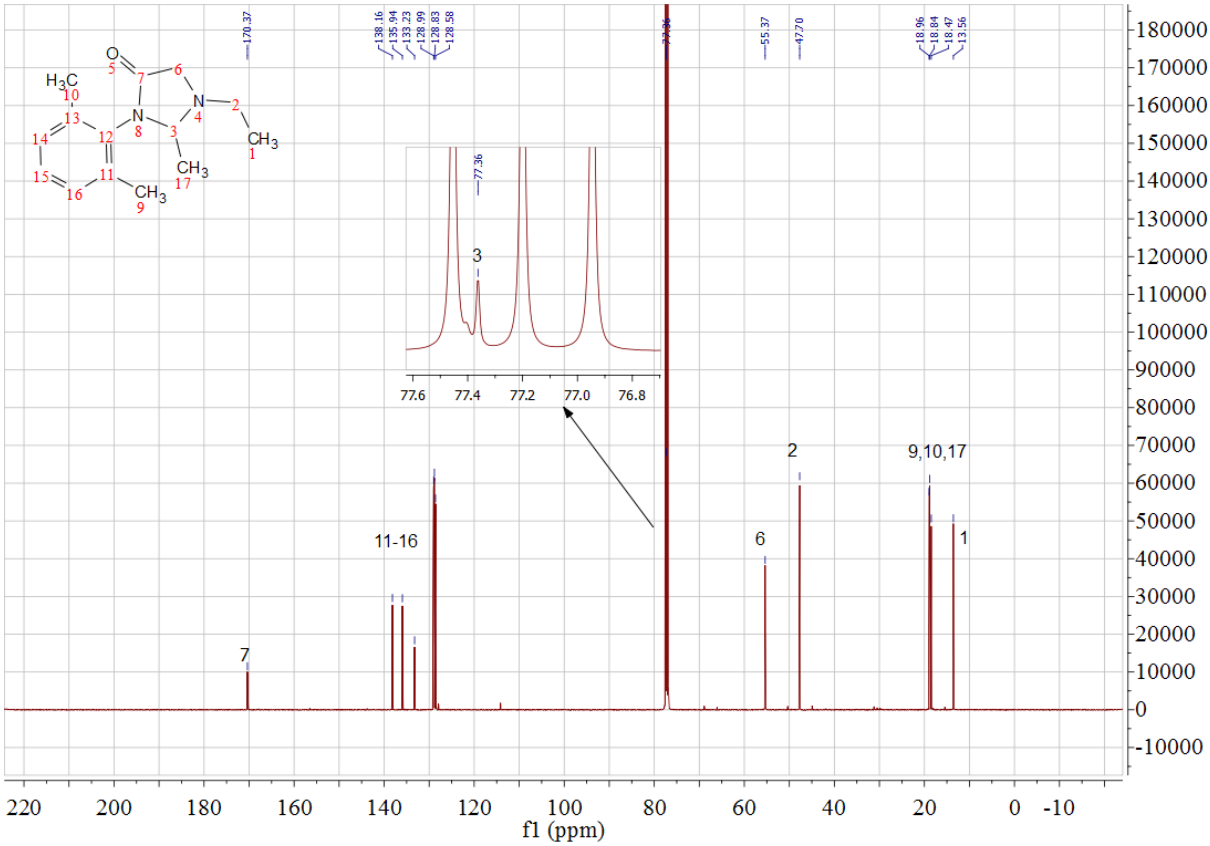


Figure 3.23 ^{1}H and ^{13}C NMR spectra of (Z)-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline, **3.14**

This cyclization product (**3.14**) is not a human metabolite; it was found in lidocaine metabolism in monkeys but the enzyme responsible for its formation is not known.[364] Both **3.13** and **3.14** can be formed by spontaneous breakdown of the carbinolamine (Figure 3.24), with acetaldehyde elimination competing with loss of hydroxide to form an iminium ion that is trapped by intramolecular nucleophilic attack by the amide nitrogen. If these reactions are non-enzymatic, similar product distributions should be observed for the library of $P450_{BM3}$ variants.

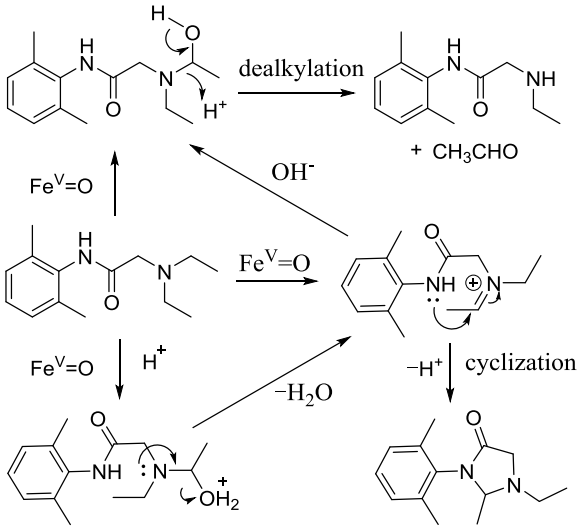


Figure 3.24 Proposed mechanism for the lidocaine intramolecular cyclisation metabolite 2-acetyl-6-methoxynaphthalene **(3.14)**

N-Dealkylation is the common activity, for example, MEGX is the primary human metabolite both *in vitro* and *in vivo*. The different behaviour observed with $P450_{BM3}$ suggests a different chemistry. The carbinolamine may break down within the active site, for example, with an acidic side chain protonating the alcohol group to promote iminium ion formation in preference to acetaldehyde elimination. Alternatively, the reaction may proceed *via* an iminium intermediate formed by α-hydrogen atom abstraction and electron-transfer oxidation. The partition between dealkylation and cyclisation is determined by the competition between iminium trapping by the Fe^{III}–OH versus intramolecular attack by the amide nitrogen. If the

iminium carbon is far away from the Fe^{III}–OH, cyclisation is more likely to occur. The product ratio depends on the effect of mutations on the mobility of the iminium group. Whatever the mechanism, the $P450_{BM3}$ mutants favour cyclisation to form the aminal which is a protected imine and susceptible to nucleophilic attack. Since the amide nitrogen is by far the better leaving group, aminal formation potentially offers a novel route for α-functionalisation of amines if the enzyme can accept lidocaine-like derivatives with more complex amines in place of the diethylamino group.

3.5 Chapter summary and conclusions

The library of $P450_{BM3}$ variants oxidised the panel of cationic (amitriptyline, lidocaine), anionic (diclofenac, naproxen) and neutral (testosterone, chlorzoxazone) drug compounds to their human metabolites, thus mimicking the human CYP functions. No single mutant or family of mutants showed high conversion with all the tested drugs but the mutant library gave full coverage and with varied product profiles. Total conversion was observed with all the test substrates except amitriptyline (highest conversion 79%), which enabled full product characterisation and led to the discovery of the new P450 reaction type of oxidative decarboxylation of an α-hydroxy carboxylic acid (naproxen) and the intramolecular cyclisation of an amine (lidocaine). The formation of a stable, protected imine by cyclisation of lidocaine offers a novel route for α-functionalisation of amines.

There is intense interest in late stage C–H bond functionalisation as a means to streamline synthesis. A prerequisite for this is an enzyme collection with wide substrate range, excellent functional group tolerance, diverse product selectivity and high conversion to maximise the likelihood of initial hits for further selectivity optimisation. The results suggest that the library of $P450_{BM3}$ variants is the basis for one such collection.

Chapter Four

Chapter 4 C–H Functionalisation of Drug Fragments

4.1 Introduction

4.1.1 An overview of drug discovery process and synthetic chemistry

The roles played by synthetic chemistry in the pharmaceutical industry have always been one of the main driving forces in the in the drug discovery process.[365] Figure 4.1 shows the general procedures of drug discovery while the impact of organic synthesis at key phases is evident.[366]

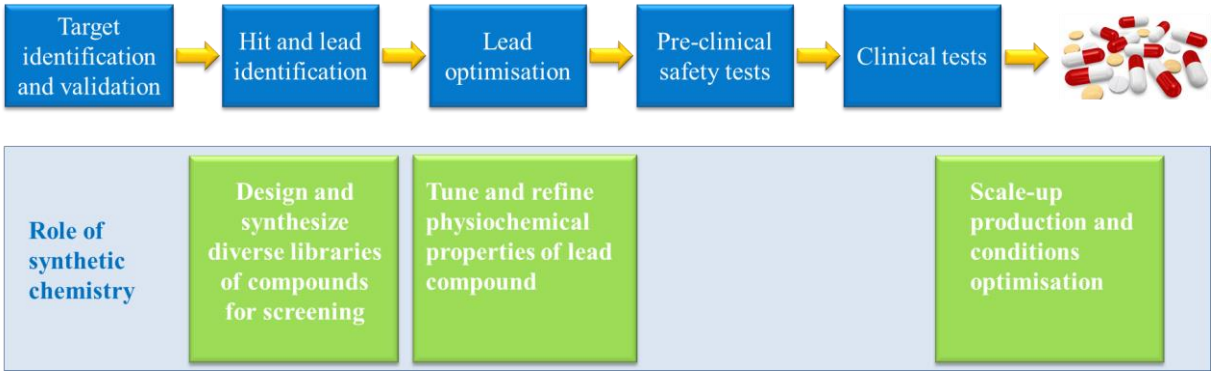


Figure 4.1 The role of synthetic chemistry in the drug discovery process

The drug discovery process generally starts from the identification and validation of proper targets such as proteins, genes, viruses, etc. using a variety of chemical, biological and biophysical techniques.[367] A large number of diverse molecules are then designed and synthesised by a range of robust synthetic methodologies in the search for initial starting points ('hits' or 'leads') that interact with the targets.[368] Once the initial hits and leads have been identified by high throughput screening, the 'lead optimisation' step takes place where leads are optimised through the synthesis of many designed analogues to improve their affinity, selectivity and safety.[367] The resulting 'drug candidate' compounds are subjected to a series of preclinical and clinical tests to deliver a marketable medication.[367] Scale-up and

process chemistry then devises elegant and efficient syntheses to deliver the final compound on multi-kg scales.[366]

Although the drug discovery process has been broadly acknowledged as a well-developed business model, numerous unprecedented challenges have occurred in the pharmaceutical industry, among which the efficiency of making and screening for candidate compounds with more likelihood of success has become one of the most pressing and well-documented challenge. It has been shown by recent cheminformatics data that drug candidates are often prone to failure in clinical trials (66% in phase II and 46% in phase III) due to unforeseen complications, such as poor bioavailability, poor pharmacokinetic properties or unwanted toxicological effects,[369] suggesting the importance of molecular properties of the lead compounds from which the drug candidates are derived in determining their success in clinical trials. By preparing leads with more appropriate physiochemical, drug-like properties, it may be possible to decrease the failure rate, leading to a substantial increase in the productivity and efficiency of the entire drug discovery process.

4.1.2 Characteristics of drug-like molecules

The link between physiochemical properties (size, shape, lipophilicity) of the molecules and their likelihood to become successful drugs has been described by early researchers.[370, 371] For example, the Lipinski's 'rule of five' have set the boundaries and criteria for orally available drugs (Table 4.1)[372], including MW $\leq$ 500, log P $\leq$ 5, hydrogen bonds donors $\leq$ 5, and not more than 10 hydrogen bond acceptors.

Table 4.1 Summary of Lipinski's 'rule of five'

Physiochemical property	Ideal value
Molecular weight	≤ 500
log P	≤ 5
H-bond donors	≤ 5
H-bond acceptors	≤ 10

These rules have previously been used as a guide to describe small-molecule drug-like space. However, the rules are not exclusive. Taxol[373], being one of the examples, is arguably the most effective anti-cancer drug molecule in the market but it lies outside of the chemical space defined by Lipinski's 'rule of five'. In addition, some recent studies also suggested that compounds with significantly lower MW and log P values are much more favoured in the development of successful medicines and this has led to several revisions of these rules, suggesting the physiochemical properties of candidate molecules including log P ≤ 3 and polar surface area >75 $Å^2$ have great safety in pre-clinical studies while molecules with MW ≤ 400 and log P ≤ 4 are more likely to be deliver successful drugs.

Whilst many molecular properties of the candidate compounds have significant effects on the likelihood of success in development of medicines, log P value is generally the most critical factor.[371, 374-376] Increasing hydrophobicity of drug candidates may result in poor aqueous solubility but better binding to their protein targets while excessive lipophilicity can promote promiscuous and uncontrollable off-target binding, leading to undesired biological or toxicological effects. Excessive hydrophilicity (e.g. a negative log P value) should also be forbidden since it may introduce difficulties for the candidate molecules getting across cellular membranes.[377-379] It is now broadly agreed that log P in the range of 1–3 gives the best balance between hydrophilicity and lipophilicity of drug candidates with higher success rate.

Other recently studies also noted the importance of the molecular shapes in the drug discovery process as compounds with higher fractions of sp^3-hybridised carbons (Fsp^3) are favoured[380] while high numbers of aromatic rings (≥ 3)[381] are often associated with undesirable outcomes and ultimately failure. This observation strengthens the impact of lipophilicity of drug substances since the degree of aromaticity is related to their log P values and aqueous solubility.

4.1.3 Characteristics of lead-like molecules

In order to prepare drug candidates with preferred properties which fall into the defined or optimal drug-like chemical space, leads and hits should be targeted with appropriate characteristics to allow for the tendency for increases in molecular size, weight, lipophilicity (log P value) and complexity in the lead optimisation step where numerous analogues are made in the search for desired molecular interactions.

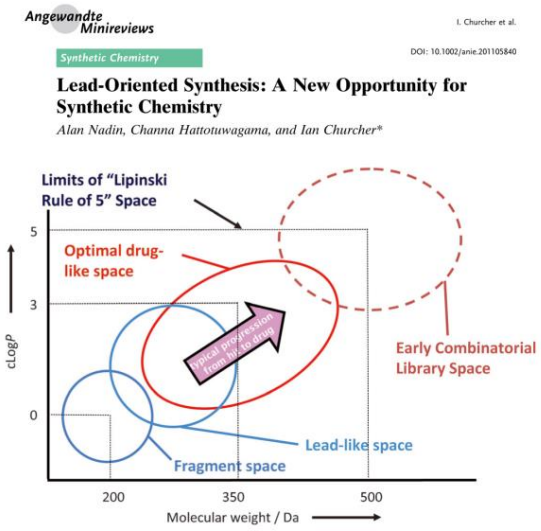


Figure 4.2 Lead-like space defined by Churcher and co-workers at GSK

Therefore, a "lead-like" space (Figure 4.2) has been defined by Churcher and co-workers at Glaxo Smith Kline (GSK) in which the lead molecules usually have well-controlled properties including lipophilicity (log P in the range of –1 to +3), molecular weight (MW in the range of 200 to 350 g mol^{-1}), sub-structure and functional groups (Table 4.2).[366]

Table 4.2 Lead-like rules defined by Churcher and co-workers

Physiochemical property	Ideal value
Molecular size	$14 \leq$ non-hydrogen atoms ≤ 26 $200 \leq MW \leq 350$
Lipophilicity	$-1 \leq \log P \leq +3$
Molecular shape	$1 \leq$ aromatic rings ≤ 2 High degree of Fsp^3
Substructure	Remove chemically reactive, electrophilic or redox active groups

In contrast to many early combinatorial libraries which relied on various synthetic chemistry methodologies to prepare large numbers of heavy molecules with high degree of complexity, the molecules in 'lead-like space' are much smaller, more hydrophilic and incorporate fewer chemically reactive, electrophilic or redox active groups, leading to the requirement of an opposite lead optimisation vector compared to traditional processes in order to reach optimal drug-like space. Lead-like molecules act as a more optimised source of starting points with a higher probability of delivering successful drug candidates compared to non-lead-like molecules, as they allow for maximum flexibility and property inflation through optimisation. However, it should be noted that these rules are not exclusive, with some successful drugs developed from lead molecules outside of the lead-like space.

4.1.4 Approaches to the preparation of libraries and lead-oriented synthesis (LOS)

Although the physiochemical properties of molecules that determine their success in drug development are broadly understood, there is increasing realisation that current methodologies for building-block library synthesis are unintentionally predisposed to producing molecules with poor drug-like properties.

Over the last two decades, medicinal synthetic chemists have developed numerous robust and predictable strategies and methodologies of producing large numbers of novel molecules to screen for drug candidates. Early combinatorial synthesis libraries relied on introducing diverse capping groups and synthesising various analogues to increase their library size and diversity. However, the addition of functional groups at multiple points of such molecules may result in larger molecular size and higher lipophilicity which lie outside of lead-like space.

Another concept, termed diversity-oriented synthesis (DOS) was also put forward by Schreiber and co-workers which utilised cascades to produce large numbers of scaffolds using a small set of transformations thus creating diversity.[382] Many early DOS libraries were found to be large in size in order to incorporate high degrees of diversification but are often accompanied with low lead-likeness.[366, 383]

Churcher and co-workers at GSK analysed 4.9 million compounds from commercially available libraries by their lead-likeness rules and found that only 2.6% exhibited lead-like character, emphasising the current predisposition of synthetic methodologies to produce poor lead-like compounds due to the harsh conditions and side reactions in the combinatorial synthesis methods, the catalysts employed, and the inherent bias towards more lipophilic molecules that are more easily isolated.

Hence, the concept of Lead-Oriented Synthesis (LOS) has been put forward by Churcher and co-workers aiming at synthesising molecules with specific molecular activities. This approach differs from target oriented synthesis (TOS) which targets a single compound, diversity oriented synthesis (DOS) which focuses on the scaffold diversity, and traditional combinatorial chemical synthesis which aims at large numbers of derivatives. Efficient, manageable, and cost effective LOS methodologies should be designed and particular

attention should be paid to the physiochemical properties of target molecules. The ideal LOS methodology and its ideal characters are shown in Figure 4.3.

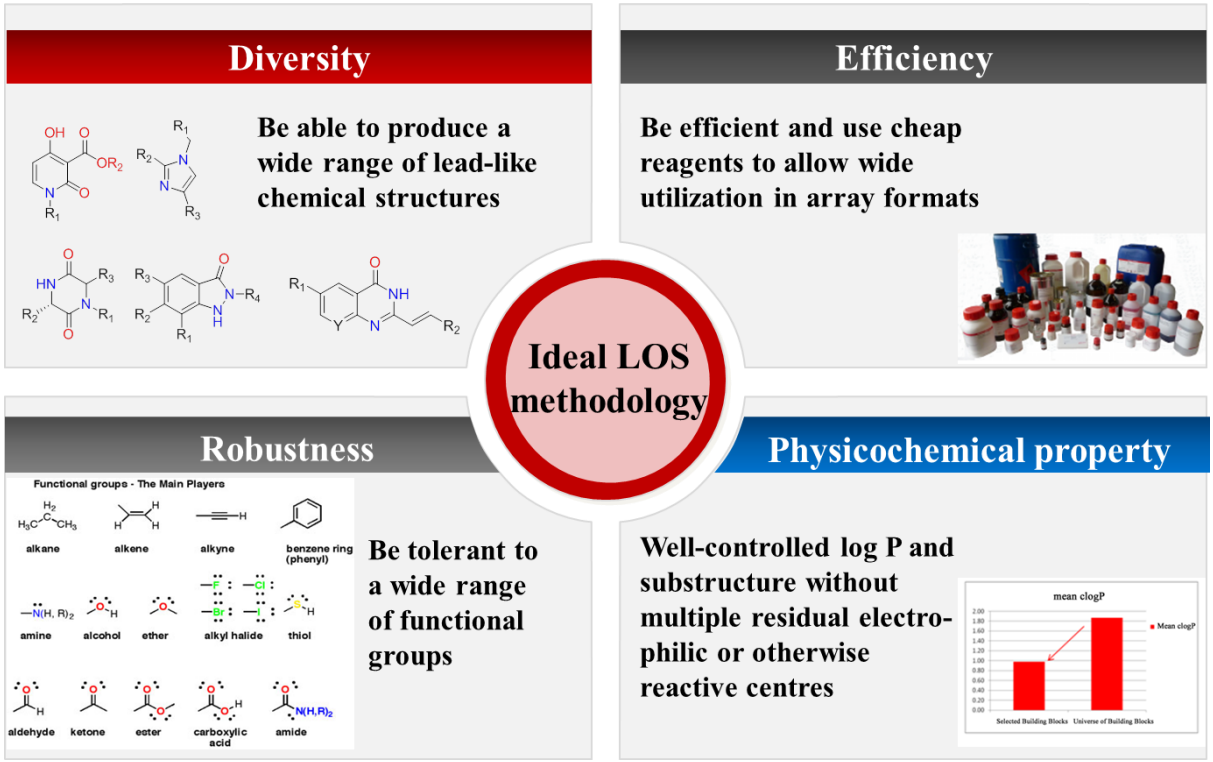


Figure 4.3 Important factors in the development of LOS methodology

4.1.5 LOS and Cytochrome P450s

As mentioned in section **4.1.3** and **4.1.4**, controlling lipophilicity is the main challenge and a critical factor to deliver a successful LOS approach. The simplest approach to overcome the issues of excessive lipophilicity is to increase hydrophilicity by introducing alcohol groups which also add functional diversity and offer sites for further structure elaboration. However, direct functionalisation of C–H bonds is not possible in classical synthesis, and elaborate protection-deprotection strategies are required that are difficult to deploy in combinatorial synthesis, adding to both the time and cost required for drug development.

On the other hand, the activation of C–H bonds is the signature activity of cytochrome P450 enzymes, making P450 enzymes appealing for the development of LOS. Moreover, P450 enzymes are generally promiscuous, able to functionalise more than just the most activated

carbon atoms within in a molecule, which often leading to a diverse set of products with multiple positions without causing polymerisation of electron rich groups.

Considering the intrinsic nature of various synthetic chemistry approaches involved in lead optimisation are often related to a decrease of hydrophilicity, P450 enzymes can act as a complementary tool for controlling molecular lipophilicity as well as promoting molecular diversity for structural modification.

4.1.6 Research objectives and chapter outline

To assess the applicability of using cytochrome P450s for the development of LOS approach to produce potential drug-compounds, it would be remiss not to look at common motifs that have notable interactions with biological systems. Consequently, drug fragments become the targets at the first stage of assessment.

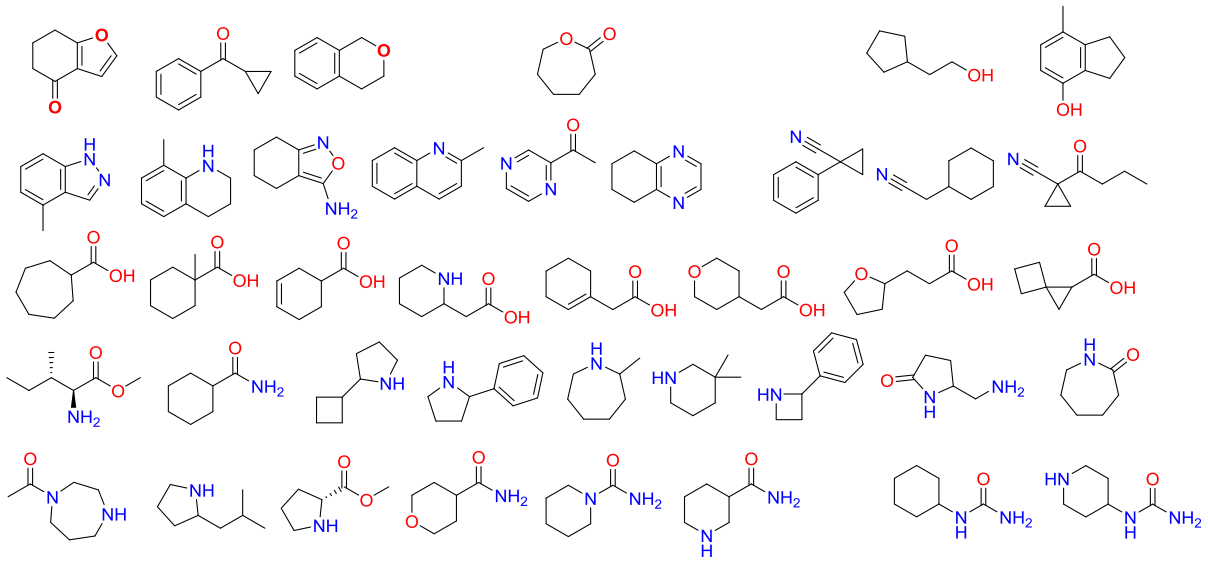


Figure 4.4 40 fragment molecules selected from a GSK library

Working with Nadin and Churcher at GSK, a total of ~40,000 fragment compounds with a heavy atom count (HAC) of 7–11 from commercially available libraries of drug building blocks were assessed for their suitability for LOS. An initial collection of ~1200 molecules

was then examined and 240 were identified as drug-like targets. Of these a library of 40 compounds (Fig. 4.4) was chosen that encompass common drug fragments and contain a variety of functional groups (alcohols, ketones, phenols, ether/ester groups, carboxylic acids, *N*- and *O*-heterocycles, aromatic and unprotected aliphatic amines, and amides). 20 mg samples of each were initially supplied by GSK and tested as substrates of P450 enzymes. Biohydroxylation of this library of 40 small molecules (HAC 7–11) with varied functional groups into a diverse set of lead-like compounds represents a daunting challenge for classical protein engineering and screening approaches but is well suited for the generic accelerator $P450_{BM3}$ mutants. The objective is to generate functional diversity by hydroxylation at as many carbon centres of each substrate as possible, in principle to screen the broad spectrum of products for activity against drug targets and to generate synthetic handles (–OH groups) for building candidate drugs. The research plan is outlined in Figure 4.5.

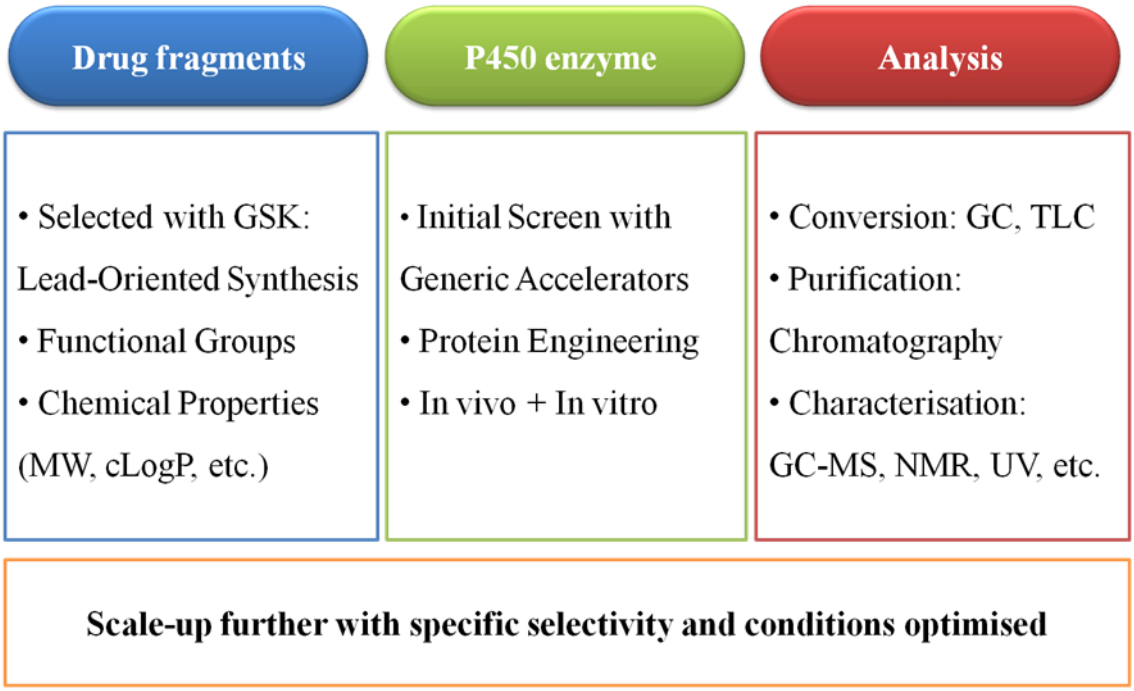


Figure 4.5 Research plan and methods for developing a new LOS approach

The library of molecules was screened for oxidation by a set of 4–5 generic accelerators and new generation variants derived from these, by GC, TLC and MS. Promising hits were scaled up to 10–300 mg scale substrate conversion. The products were isolated and characterised by NMR and MS. Further protein engineering could then be undertaken to alter the product distribution. A panel of ~50 $P450_{BM3}$ variants was created and expressed for screening purposes. Enzymes were adjusted to 10 μM concentration before being divided into 24-well plate (100 μL in each well). Screening reactions were conducted using an NADPH regeneration system as described in chapter 3, with 1−2 μM enzyme and 1−2 mM substrate. The reactions were left shaken at 200 rpm for 8 h prior to extraction and analysis.

4.2 Nitrogen heterocycles

Nitrogen heterocycles (aromatic, aliphatic, etc.) represent a large class of compounds that have the potential to deliver final drugs. The activation and functionalisation of any C–H bond into either an imine or an alcohol would offer possibilities for structural modification and elaboration. Some of the nitrogen heterocycles are presented in Figure 4.6.

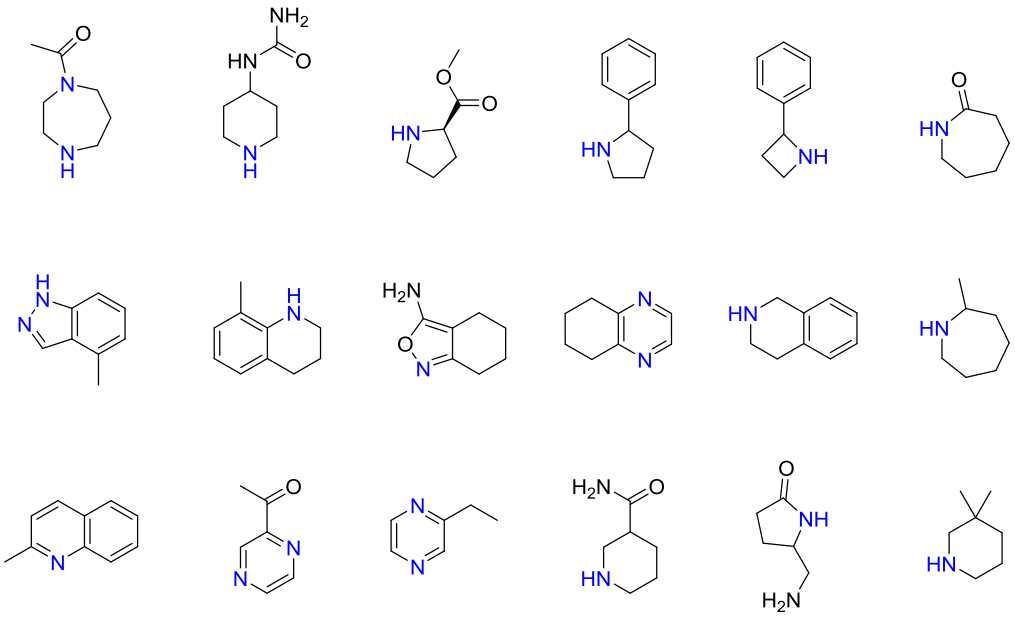


Figure 4.6 Selected nitrogen heterocycles

Monoamine oxidases (MAOs) are a class of oxidoreductase enzymes which are able to oxidise amines to imines with simultaneous reduction of oxygen to hydrogen peroxide.[268] They have been engineered by directed evolution to provide a tool-box of variants which can convert aliphatic amines to imines/iminiums and therefore having great potential to become the catalysts of choice for LOS strategy.[266, 267, 384] However, aromatic nitrogen heterocycles are not suitable substrates for MAOs. Aromatic nitrogen heterocycles were then subjected to oxidation using $P450_{BM3}$ variants while aliphatic nitrogen heterocycles were screened for comparison purposes.

Nitrogen heterocycles presented a significant challenge because aromatic amines (e.g. cyanopyridine) are known to inhibit wild type $P450_{BM3}$ *via* ligation of the *N*-atom to the haem iron.[385] Since not all of the selected nitrogen heterocycle compounds were detectable by gas chromatography, 5,6,7,8-tetrahydroquinoxaline, 8-methyl-1,2,3,4-tetrahydroquinoline, acetylpyrazine were selected for scale-up reactions as they were commercially available while other compounds were only screened for possible conversions. Ethylpyrazine was investigated for the comparison with acetylpyrazine.

4.2.1 5,6,7,8-Tetrahydroquinoxaline

5,6,7,8-Tetrahydroquinoxaline (THQ, **4.01**, Figure 4.7) is a bicyclic amine being protected by conjugated system of pyrazine ring. A library of 30 $P450_{BM3}$ variants were screened for oxidation of 5,6,7,8-tetrahydroquinoxaline (THQ, **4.01**) and analysed by gas chromatography.(Table S4.1 in Appendix, Figure 4.8 and Figure 4.9). The screening results confirmed the inhibitory behaviour of nitrogen heterocycles towards the wild type enzyme (WT), which showed *ca.* 1% conversion of THQ. However, this did not prove to be an obstacle for all the mutants developed on generic accelerators (A330P, I401P, KT2, KSK19); 100% conversion was observed with multiple variants to a range of four products.

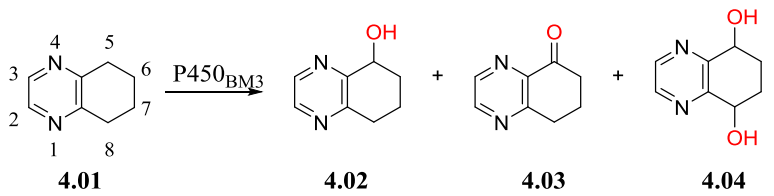


Figure 4.7 $P450_{BM3}$ reaction with 5,6,7,8-tetrahydroquinoxaline **(4.01)**

Three products were characterised from preparative scale reactions. The RP/AP/AM mutant showed total conversion to a single product with a retention time 5.67 min on the GC chromatogram (Figure 4.9a), which was readily purified and found to be 5-hydroxy-THQ (**4.02**). KSK19/A82M gave 62% of **4.02** while 2 minor products also present at 36% (retention time 6.78 min) and 3% (retention time 6.78 min).

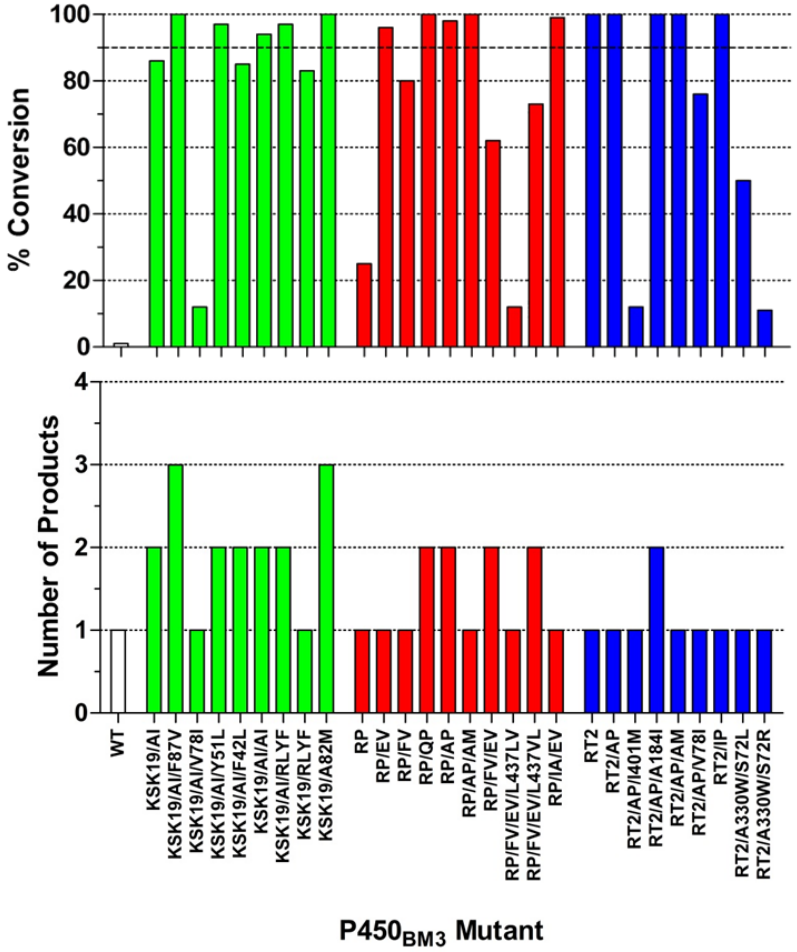
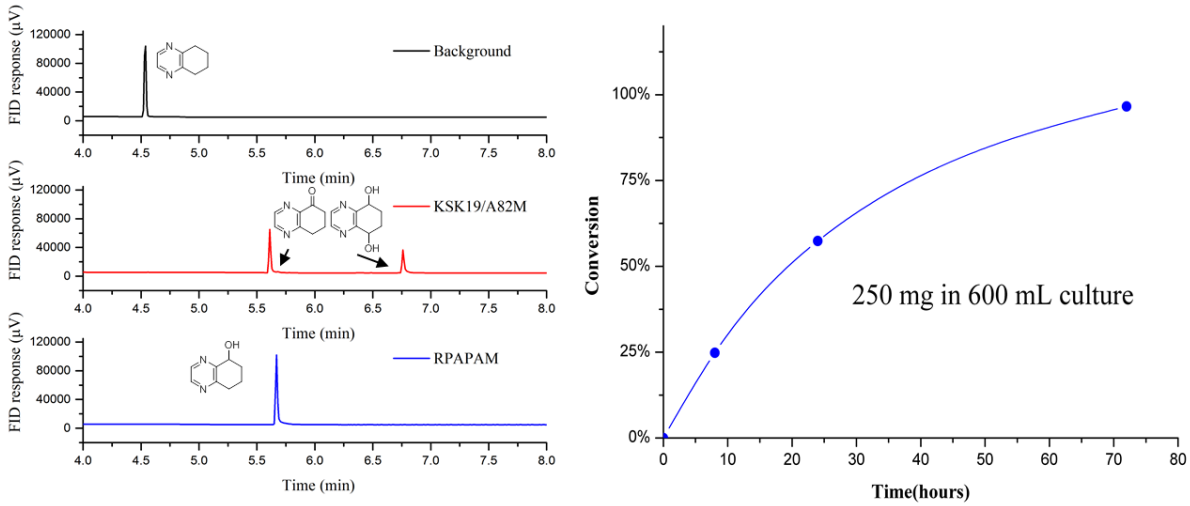


Figure 4.8 Substrate conversion and number of products formed in THQ (**4.01**) oxidation by $P450_{BM3}$ library

A 250 mg reaction was performed *in vivo* using KSK19/A82M expressed in 600 mL FB_{gly} and allowed for the two minor products to be isolated and characterised as THQ-5-one (**4.03**) and THQ-5,8-diol (**4.04**). Figure 4.9b shows the time course study of this *in vivo* reaction, emphasizing that this nitrogen heterocycle substrate is well-tolerated within the $P450_{BM3}$ enzyme system. NMR data allowed for confirmation of the three products with the doublet of doublet present in the 5.0 ppm region corresponding to the C-5 H-atom, while GC-MS analysis showed three products, with MW of **4.01** + 14 (**4.03**), + 16 (**4.02**) and + 32 (**4.04**).



(a) (b)

Figure 4.9 (a) GC analysis of 2 mM THQ (4.01) oxidised by 1 μM RP/AP/AM and KSK19/AM mutants. (b) Time course of 250 mg *in vivo* reaction using the KSK19/AM variant

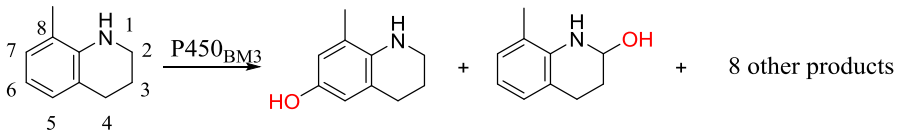
Figure 4.8 shows the substrate conversion and number of products formed among all the 30 screen reactions. Hydroxylation mainly occurred on benzylic position to give the THQ-5-ol (**4.02**) as the major product, indicating all the three generic accelerators (RP, KSK19, KT2) are good basis for THQ oxidation whilst mutations such as V78I, L437LV, I401M suppressed enzyme activity. THQ-5-one (**4.03**), derived from the over-oxidation of THQ-5-ol (**4.02**), was rare as it only appeared with 10 variants among the 30 mutants screened, and only

2 variants gave over 20% selectivity (KSK19/A82M, KSK19/AI/AI). KSK/19AI only displayed 3% selectivity towards **4.03**, suggesting the importance of the A184I mutation for ketone formation. Mutants incorporating A330P or A330W did not give any **4.03**, revealing bulky mutations at A330 disfavoured over oxidation of **4.02**. The RLYF couplet and F87 mutations did not show any dramatic effects on THQ (**4.01**) oxidation. There was another organic-soluble product with some variants (Table S4.1 in Appendix) but then was minimal, if any, oxidation at the ring nitrogens which would be the expected site of attack by chemical oxidising agents in classical synthesis.

4.2.2 8-Methyl-1,2,3,4-tetrahydroquinoline

8-Methyl-1,2,3,4-tetrahydroquinoline (**4.05**) belongs to the family of quinolines and can be readily synthesised by regioselective hydrogenation of 8-methylquinoline catalysed by Pd nanoparticles. Quinolines are important motifs in many drug-like molecules and have been widely used in treatment of illnesses such as malaria, arthritis and lupas.[386]

8-Methyl-1,2,3,4-tetrahydroquinoline (**4.05**) is a slightly protected amine, with the NH group sterically protected by the methyl group and the nitrogen lone pair is delocalised into the π system. Chemical oxidation of its non-methylated analogue, tetrahydroisoquinoline, with H_2O_2/SeO_2 in methanol gives the nitrone but oxidation at other positions is not observed.[387]



4.05 **4.06** **4.07**

Figure 4.10 $P450_{BM3}$ reaction with 8-methyl-1,2,3,4-tetrahydroquinoline (**4.05**)

22 $P450_{BM3}$ variants were screened for the oxidation of **4.05** (1 μM enzyme and 2 mM substrate in 0.5 mL 200 mM pH 7.5 phosphate buffer) using a NADPH regeneration system. The reactions were extracted by 300 μL of ethyl acetate and analysed by gas chromatography.

The substrate conversion and number of products are shown in Figure 4.11 while the full table of product distribution is shown in Table S4.2 in Appendix.

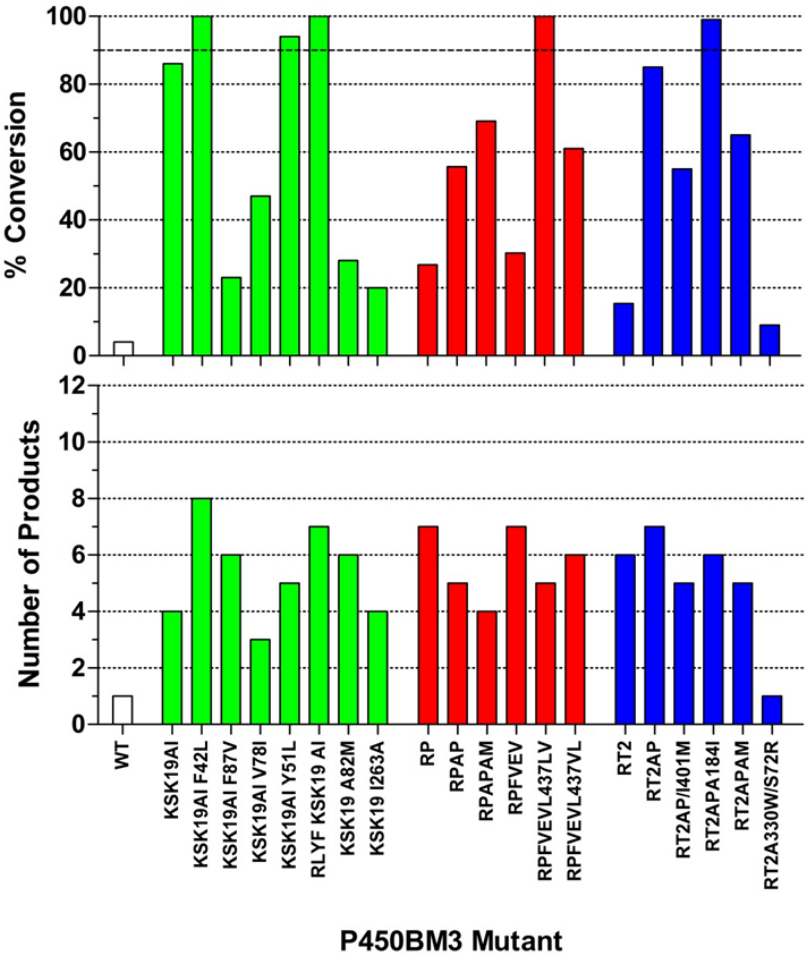


Figure 4.11 Substrate conversion and number of products formed in 8-methyl-1,2,3,4-tetrahydroquinoline (**4.05**) oxidation by a $P450_{BM3}$ library

As shown in Figure 4.11, wild type (WT) $P450_{BM3}$ showed low activity (4 % conversion) to give one product (**4.06**), emphasising the inhibitory behaviour of nitrogen heterocycles towards WT enzyme. While for the variants based on the three families of 'generic accelerators', >90% conversion was observed from multiple variants in each family to a total of 10 products.

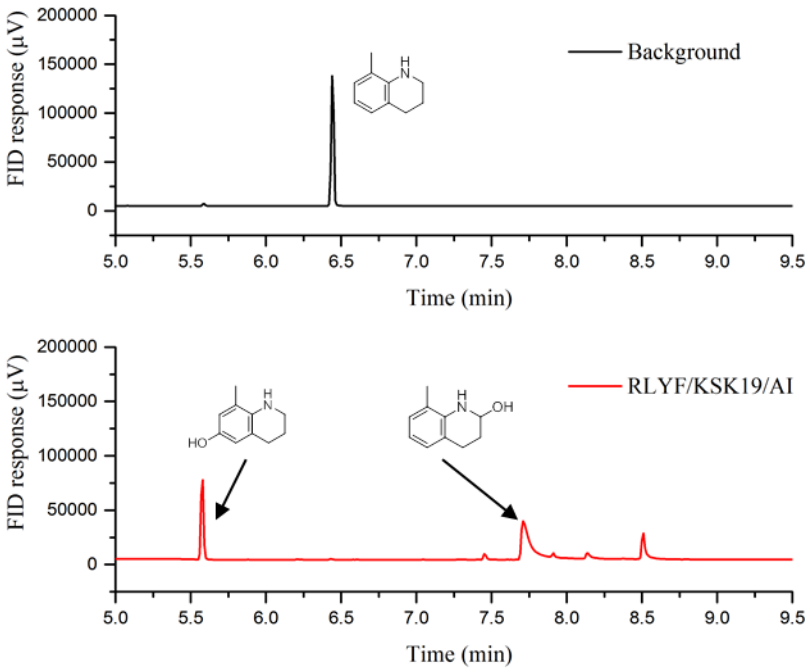


Figure 4.12 GC trace of 12mg prep-scale reaction of 8-Methyl-1,2,3,4-tetrahydroquinoline (**4.05**) using RLYF/KSK19/AI mutant

Two products were characterised from a 12 mg scale reaction using the RLYF/KSK19/AI mutant, which showed >99% conversion to 30% of a product with the retention time 5.59 min and 46% of a product at 7.73 min (Figure 4.12). The NMR data allowed for confirmation of the two productsas 8-methyl-1,2,3,4-tetrahydroquionolin-6-ol (**4.06**) and 8-methyl-1,2,3,4-tetrahydroquinolin-2-ol (**4.07**). Formation of **4.06** was confirmed by the two doublets in the aromatic region at 6.4 ppm, indicative of hydroxylation at the 6 position while the doublet of doublets at 4.75 ppm indicated hydroxylation at the 2 position in **4.07**. Stereochemistry is likely to be racemic due to scrambling of any stereochemistry by loss and re-addition of the OH group. Both compounds arise from oxidation at activated positions (*para* to an amine in an aromatic ring and an amine α-carbon). The site and nature of mutations affected the product profile (Table S4.2). RP/FV/EV/A184I gave the highest selectivity (95%) for **4.06** while the KSK19/AI/F42L was the most selective mutant towards **4.07** (52%). The A330P mutation appeared to disfavour hydroxylation at the amine α position because the highest proportion of **4.07** generated by variants incorporating A330P is 4% by RT2/AP. GC analysis

showed 8 further products. However, they are yet to be characterised (Fig. 4.9). Most of these products had retention times close to **4.07**, indicating the absence of multiple hydroxylation products. In addition, *N*-oxide was not formed since the total GC peak areas are similar among the organic extracts of all the screen reactions. Considering there are eight unique C–H bonds in the starting molecule, these further products could be either hydroxylation products at different sites or the carbonyl products derived from further oxidation from alcohols.

4.2.3 Acetylpyrazine

Acetylpyrazine (**4.08**) is a simple derivative based on the heterocyclic aromatic pyrazine motif, with structural similarities to 8 methyl-1,2,3,4-tetrahydroquinoline (**4.05**) and 5,6,7,8-tetrahydroquinoxaline (**4.01**). Pyrazines are important building blocks in drug discovery due to their antitumor, antibiotic and diuretic activities (e.g. tetramethylpyrazine, also known as ligustrazine).

32 variants were screened for acetylpyrazine (**4.08**) oxidation using the same conditions as for 5,6,7,8-tetrahydroquinoxaline (**4.01**) and 8-methyl-1,2,3,4-tetrahydroquinoline (**4.05**, Figure 4.13). Organics were extracted with ethyl acetate and analysed by gas chromatography (Appendix Table S4.3 in Appendix, Figure 4.15). The wild type enzyme (WT) only gave a negligible conversion of 3%, probably due to inhibition as a result of *N*-coordination, while other variant enzymes did not offer similar activity compared to the oxidation of **4.01** and **4.05**, with the highest conversion of 49% achieved by KSK19/AI to one major product **4.09** (GC retention time 2.01 min) and one minor product (GC retention time 3.12 min). Therefore, a 10 mg reaction with KSK19/AI was performed which enabled the major product **4.09** to be isolated and characterised as 1-(pyrazine-2-yl)ethan-1-ol (Figure 4.14). However, the

preparative scale reaction was insufficient to isolate the minor product and further work is needed to characterise its structure.

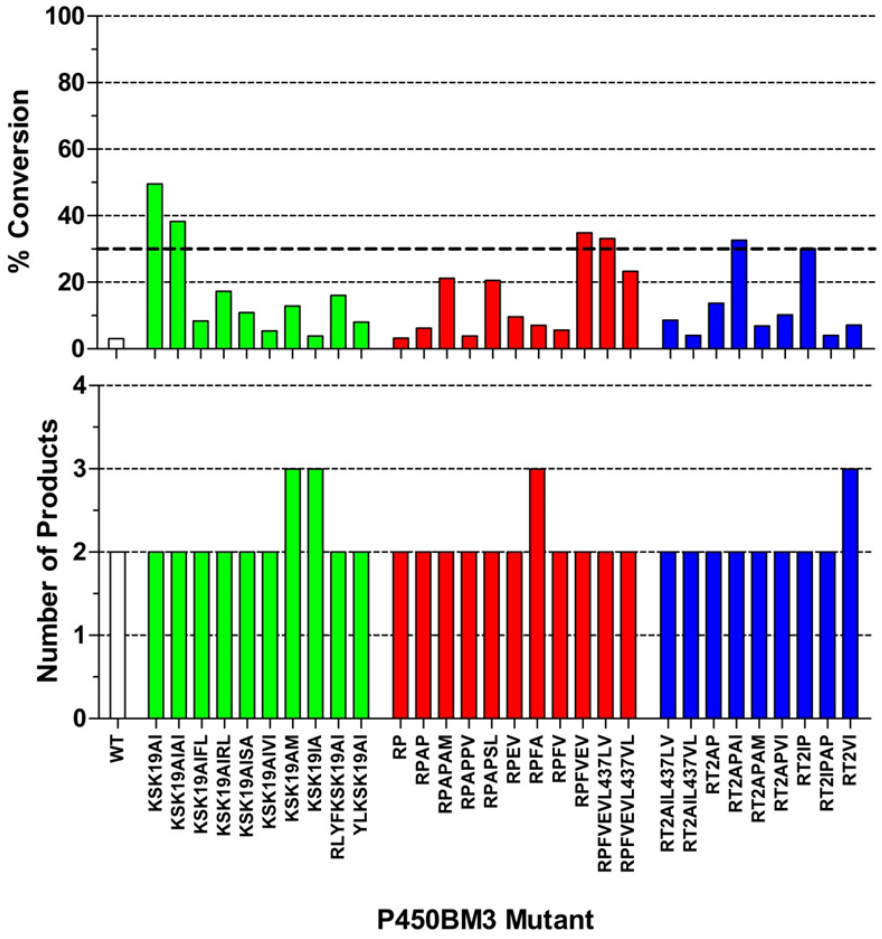
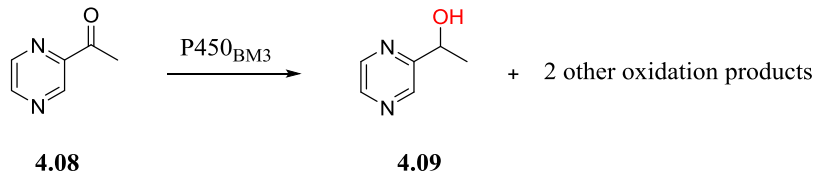


Figure 4.13 Substrate conversion and number of products formed from the oxidation of acetylpyrazine (**4.08**) by $P450_{BM3}$

Among the 32 $P450_{BM3}$ enzymes used for initial screening, the alcohol product (**4.09**) appeared in all the reactions and dominated the product profiles with the majority of the enzymes. The RT2/IP mutant gave the highest selectivity to form 98% alcohol while RLYF/KSK19/AI was 96% selective towards **4.09**. The insertion of a valine before leucine 437 disfavoured alcohol formation, with the RPFVEV/L437VL and RT2AI/L437VL mutants generating only 12% and 8% of **4.09**, respectively. Another product was also observed at retention time of 2.35 min with GVQ, RT2/VI, RP/FA, KSK19/AM and KSK19/IA but not with variants incorporating the A330P mutation, suggesting the importance of smaller side

chain amino acid residue at 330 position. Moreover, the A330P mutation relocated P329 side chain into the active site that may also had an effect on acetylpyrazine oxidation activity.



4.08 **4.09**

Figure 4.14 $P450_{BM3}$ reactions with acetylpyrazine

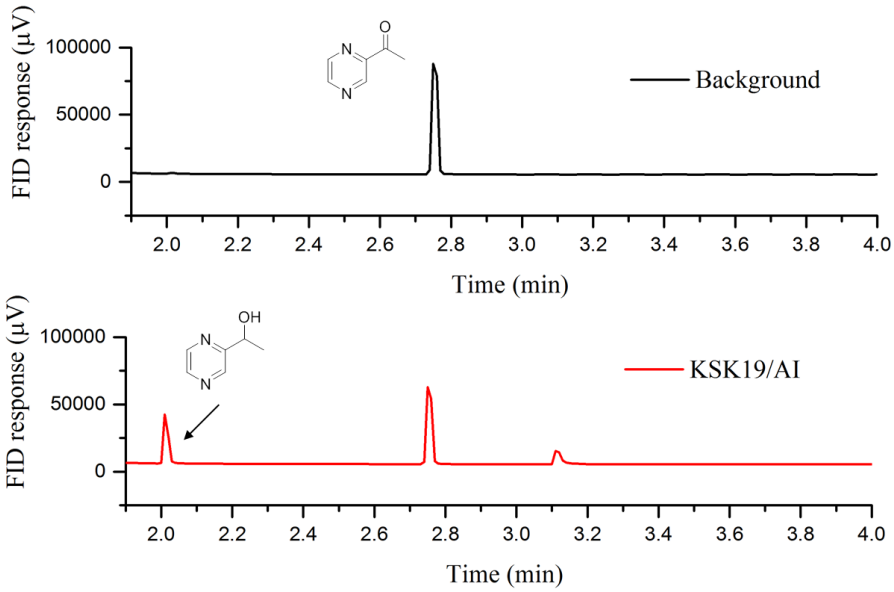


Figure 4.15 GC traces of acetylpyrazine KSK19/AI reaction

The formation of **4.09** by $P450_{BM3}$ was exceptional due to the keto reductase activity displayed by a monooxygenase under aerobic conditions. The proposed mechanism is shown in Figure 4.16, which involves the *N*-ligation of the substrate to the ferrous haem, with significant π back-donation from the metal centre to the π system of the ligand. This increases the negative charge on acetyl oxygen which is protonated to give a nitrene-type intermediate that is reduced by two electrons to give **4.09**. Because **4.09** can be reversely converted into acetylpyrazine by $P450_{BM3}$ variants, this reaction may be in equilibrium and the proportion of **4.09** formed is dependent on substrate binding to the mutant enzymes. Other factors such as dissolved oxygen level and substrate concentration could be key factors in determining the

equilibrium which need to be further investigated, for example, by performing the reaction anaerobically. Moreover, the effect of having multiple ring nitrogen atoms is also interesting, as is the position of the substituent. Further work need to be done with 2-, 3-, 4-acetylpyridine and acetylpyrimidine by the $P450_{BM3}$ library to elucidate mechanisms of keto-reduction activity. It is worth pointing out the possibilities that such keto reductases could be encoded and expressed within the *E.coli* and therefore the control reactions should be performed with highly purified enzymes. An alternative is to express the $P450_{BM3}$ enzymes in different expression systems such as yeast, insect cells, etc.

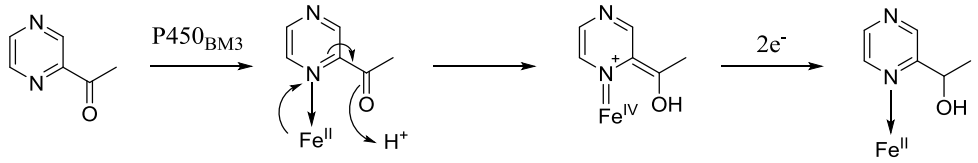


Figure 4.16 A possible mechanism for acetylpyrazine reduction by $P450_{BM3}$

4.2.4 Ethylpyrazine

Ethylpyrazine, another pyrazine compounds, was tested for the comparison with acetylpyrazine. 30 $P450_{BM3}$ variants were screened towards ethylpyrazine (**4.10**) under the same conditions as the acetylpyrazine screen. The organics were extracted with ethyl acetate and analysed by gas chromatography using the same program as acetylpyrazine reactions (Appendix Table S4.4). Figure 4.18 shows the substrate conversion and number of products formed in the screen reactions. Similarly to acetylpyrazine, wild type (WT) $P450_{BM3}$ showed low activity towards ethylpyrazine with 5% conversion while majority of mutants were inactive except three enzymes (RT2/IP, RT2/AP/AI and RP/AP/AM) giving >30% conversion.

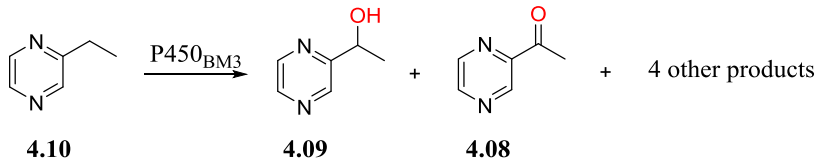


Figure 4.17 $P450_{BM3}$ reactions with ethylpyrazine

RT2/IP was in particular an exceptional, giving 97% conversion into 3 products. Therefore, a 13 mg preparative scale reaction with RT2/IP was carried out *in vitro* and this reaction led to the isolation and characterisation of two compounds: 1-(pyrazine-2-yl)ethan-1-ol (**4.09**, GC retention time 2.01 min) and acetylpyrazine (**4.08**, GC retention time 2.76 min, Figure 4.19).

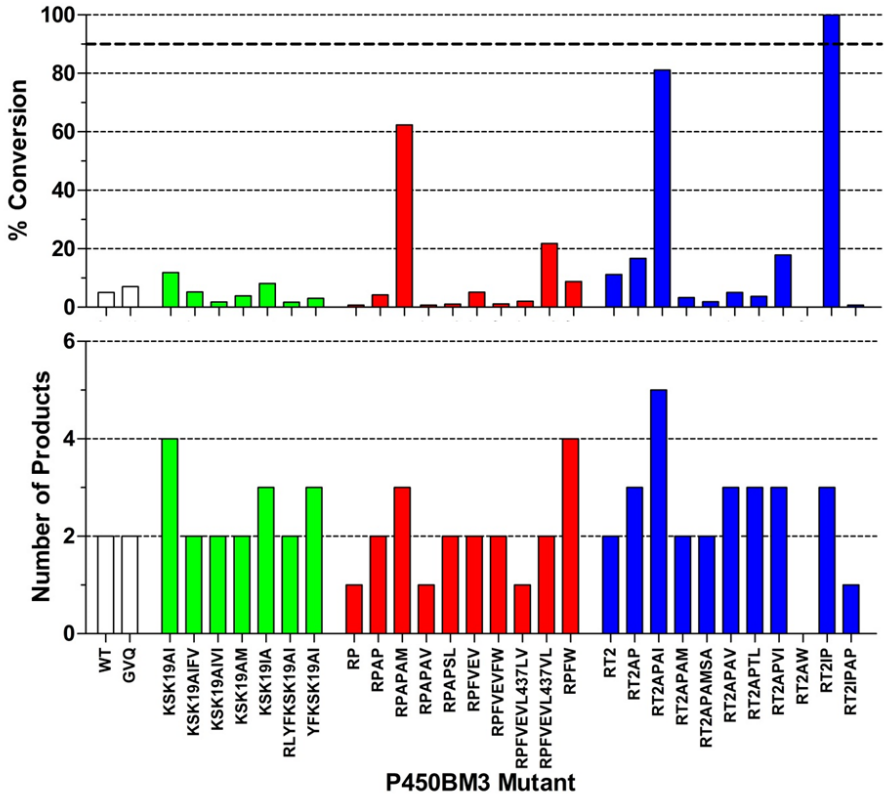


Figure 4.18 Substrate conversion and number of products formed from the oxidation of ethylpyrazine (**4.10**) by $P450_{BM3}$

1-(pyrazine-2-yl)ethan-1-ol (**4.09**) is the benzylic hydroxylation product of ethylpyrazine and appears to be the precursor of acetylpyrazine (**4.08**) *via* further oxidation at the same positon. As mentioned in the above section, the ketone reduction pathway may also take place within

the enzyme system to generate **4.09**. This alcohol product was the most common product among all the screen reactions, with a few mutants being totally selective. However, the poor overall conversions offered by the mutant library are not sufficient for conclusions on mutation-enzymatic activity/selectivity trends being conclusively drawn.

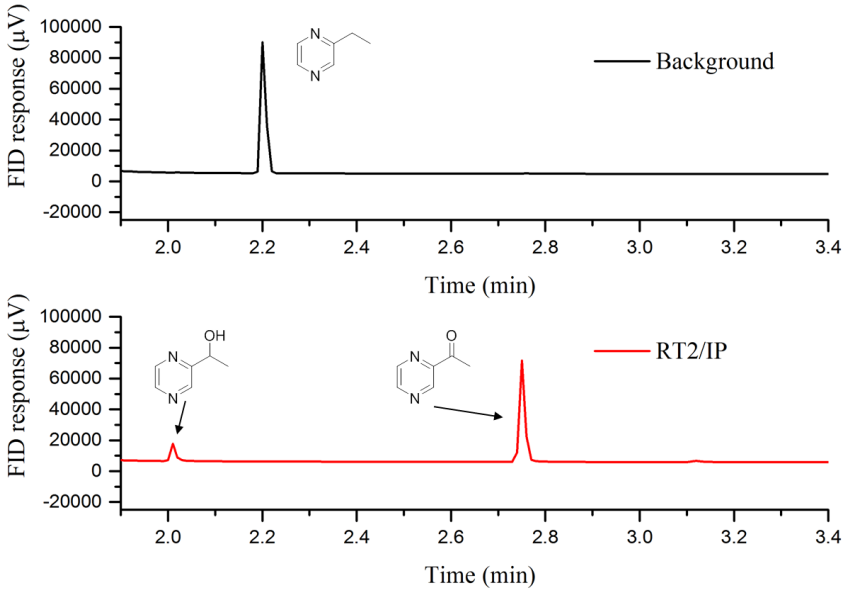


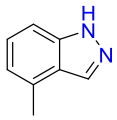
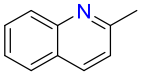
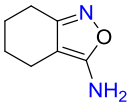
Figure 4.19 GC traces of 13 mg ethylpyrazine prep-scale reaction with RT2/IP

GC analysis also showed the formation of four further products (Appendix Table S4.4). The product at 2.35 min was formed for by GVQ, KSK19/AM and KSK19/IA. This compound was also formed with the same mutants in acetylpyrazine (**4.08**) screen reactions, suggesting these mutants are responsible for the formation of this unknown product. Other products were seen at 2.94 min, 3.12 min, and 3.51 min which are particularly interesting as they could be derived from aromatic or primary C–H oxidation. By expanding the enzyme library, the product selectivity may be shifted towards these uncharacterised products and enable them to be fully characterised.

4.2.5 Other *N*-heterocycles

Other nitrogen heterocycles supplied by GSK were also tested by the enzyme library and were oxidised to multiple products. However, due to the small amount of compounds and time limit, the products have not been characterised. Some examples are shown in Table 3.1

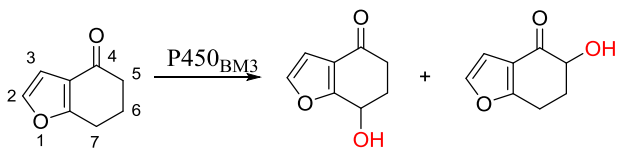
Table 4.3 Nitrogen heterocycles oxidised by P450BM3 variants

Compound			
Number of Products	3-4	5	3
Mutant	RP/FV/EV	KSK19/AI/FV	RP/FV/EV/L437LV

4.3 Oxygen-Heterocycles

4.3.1 6,7-Dihydrobenzofuran-4(*5H*)-one

6,7-Dihydrobenzofuran-4(5*H*)-one (**4.11**) is an important drug building block. Its derivatives have been widely used for the treatment of the hepatitis C virus (HPC) by inhibiting HPC viral replication.



4.11 **4.12** **4.13**

Figure 4.20 $P450_{BM3}$ reactions with 6,7-Dihydrobenzofuran-4(5*H*)-one **(4.11)**

A panel of 28 $P450_{BM3}$ enzymes were tested at 0.5 mL in 200 mM phosphate buffer, with 2 mM substrate and 1 μM enzyme. Each screen reaction was extracted with 300 μL ethyl acetate and analysed by gas chromatography (Appendix Table S4.5, Figure 4.21). Wild type

$P450_{BM3}$ showed poor activity while varied activity was seen with the mutants, suggesting this oxygen heterocycle is well tolerated by the enzyme library.

Total conversion was observed with multiple mutants (RT2/IP, RT2/AP/A184I, RT2/AP/A82L, RP/QP) to generate a range of four products, including 2 major products at 6.08 min and 3.77 min GC retention time and 2 minor products at 5.26 min and 6.12 min. The KSK19/A82M and RLYF/KSK19 mutants gave 97% and 95% conversion, respectively. (Figure 4.22)

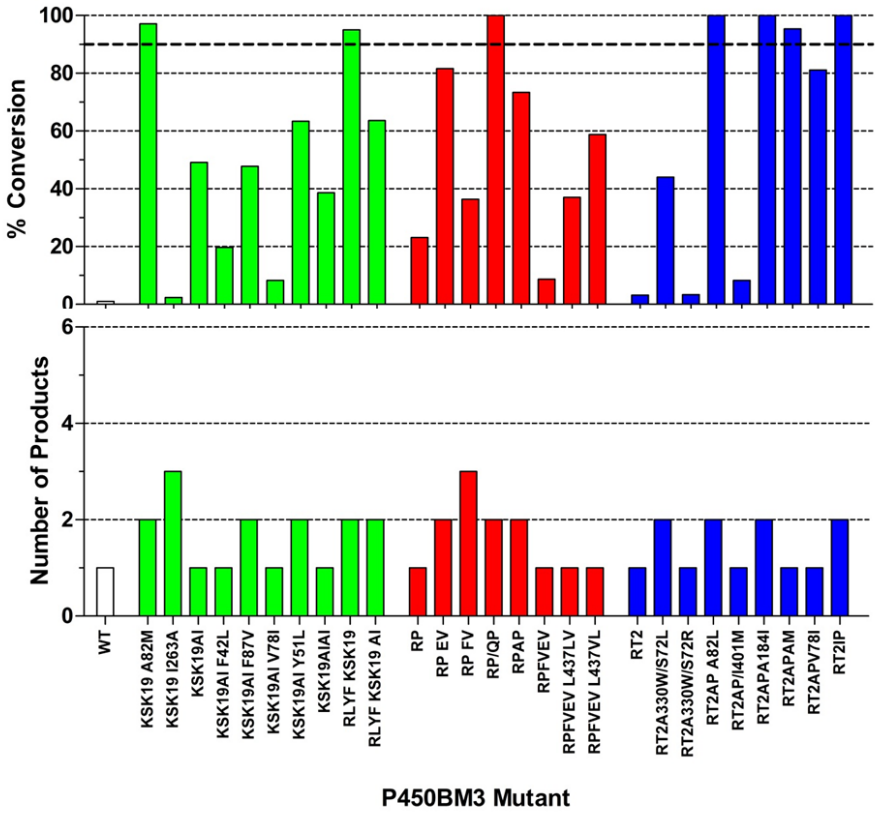


Figure 4.21 Substrate conversion and number of products formed from the oxidation of 6,7-Dihydrobenzofuran-4(5*H*)-one **(4.11)**

Furthermore, high selectivity was also seen within the mutant library. RT2/IP was 98% selective towards 6.08 min product which was then characterised as 7-hydroxy-6,7-dihydrobenzofuran-4(5*H*)-one (**4.12**) from a 6.8 mg preparative scale reaction. The structure

was determined based on the triplet at 5.04 ppm for the H atom at benzylic position on its ^{1}H NMR spectrum. In contrast, another 6.8 mg preparative scale reaction with RT2/AP/A184 mutant displayed 91% selectivity towards 5-hydroxy-6,7-dihydrobenzofuran-4(5*H*)-one (**4.13**, GC retention time 3.77 min) whose indicative peak is the triplet at 3.92 ppm for the H_5 proton in the ^{1}H NMR. Further work need to be undertaken on the determination of enantio- isomers as well as the structure elucidation of the other two minor products. It would be of interest to develop highly enatio-selective $P450_{BM3}$ variant by various protein engineering strategies that mentioned in chapter **1**.

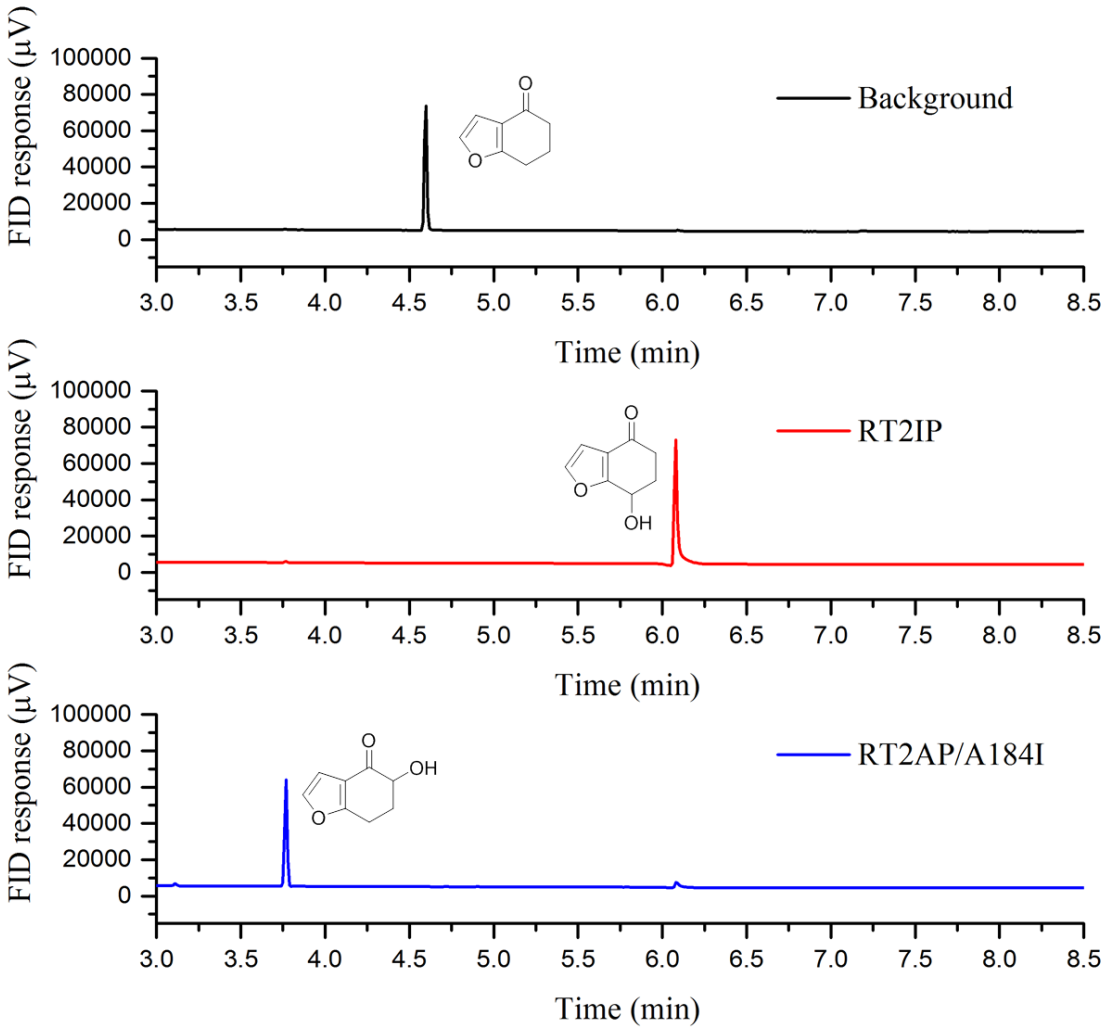
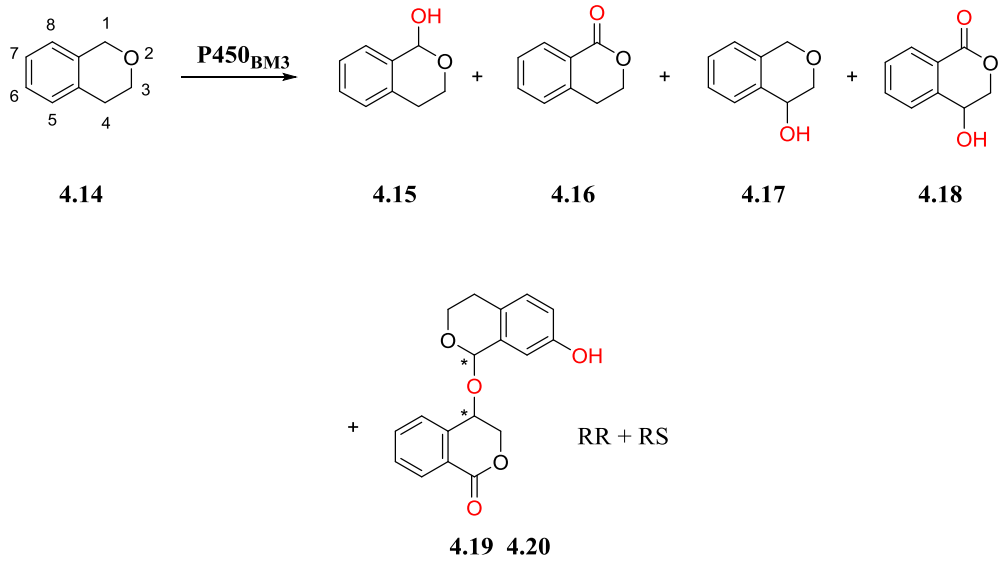


Figure 4.22 GC traces of preparative scale reactions with RT2/IP and RT2/AP/A184I

4.3.2 Isochroman

Isochroman (**4.14**) is a bicyclic oxygen heterocycle consisting of a benzene ring fused to an oxacyclohexane. Isochroman derivatives have been found to exhibit a broad range of potential bioactivities[388] and serve as important building blocks in synthetic chemistry and drug design.[389, 390] Isochroman derivatives have been widely studied and typical preparation methods include reduction of the corresponding esters,[391] intramolecular or intermolecular cyclization reactions[370, 392] and decorations of isochroman,[393] in which the benzylic position is much favoured. The synthesis of hydroxyl substituted isochroman has also been reported *via* a mercury mediated oxidative ring closure reaction of 2-(prop-1-enyl)phenylmethanol derivatives,[394] which doesn't involve starting from pure isochroman and has substitutions in the aromatic region of the product. The use of $P450_{BM3}$ may provide a new method for direct functionalisation of isochroman with controlled activity, either on the aromatic ring or the aliphatic region.



4.14 **4.15** **4.16** **4.17** **4.18**

4.19 **4.20**

Figure 4.23 $P450_{BM3}$ reactions with isochroman **(4.14)**

The screening reactions were conducted under the same conditions as compound **4.11** with a panel of 32 $P450_{BM3}$ variants. (Table S4.6 in Appendix) Figure 4.24 shows the substrate conversion and number of products formed in the screen reactions.

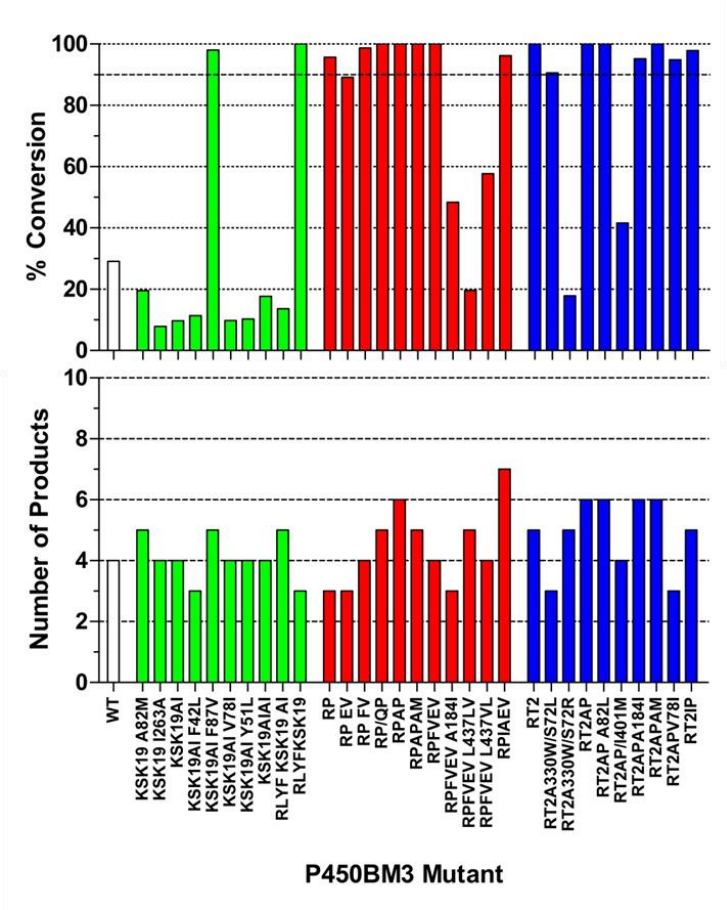


Figure 4.24 Substrate conversion and number of products formed from the oxidation of isochroman by $P450_{BM3}$

Wild type (WT) $P450_{BM3}$ showed reasonable activity and gave 29% conversion to 4 products while multiple variants fully converted isochroman into a range of 7 products in total. The use of three generic accelerators enhanced activity towards isochroman significantly. However, conversions given by KSK19 based mutants were generally lower than 20% except for KSK19/AI/FV and RLYF/KSK19, indicating the importance of the F87V mutation and the RLYF couplet. RP, RP/AP and RP/FV/EV were selected for the preparative scale reactions (13.4 mg each, as shown in Figure 4.25) in order to give a broad coverage towards

most products being formed throughout the screen reactions and led to 4 products being characterised by NMR and MS while the final unknown compound at 7.94 min was thought to be a dimer with MW 278 (2 ×MW of **4.14** + 10). The four characterised products are isochroman-1-ol (**4.15**, GC retention time 5.75 min), isochroman-1-one (**4.16**, GC retention time 5.98 min), isochroman-4-ol (**4.17**, GC retention time 6.68 min) and isochroman-1-one-4-ol (**4.18**, GC retention time 6.91 min).

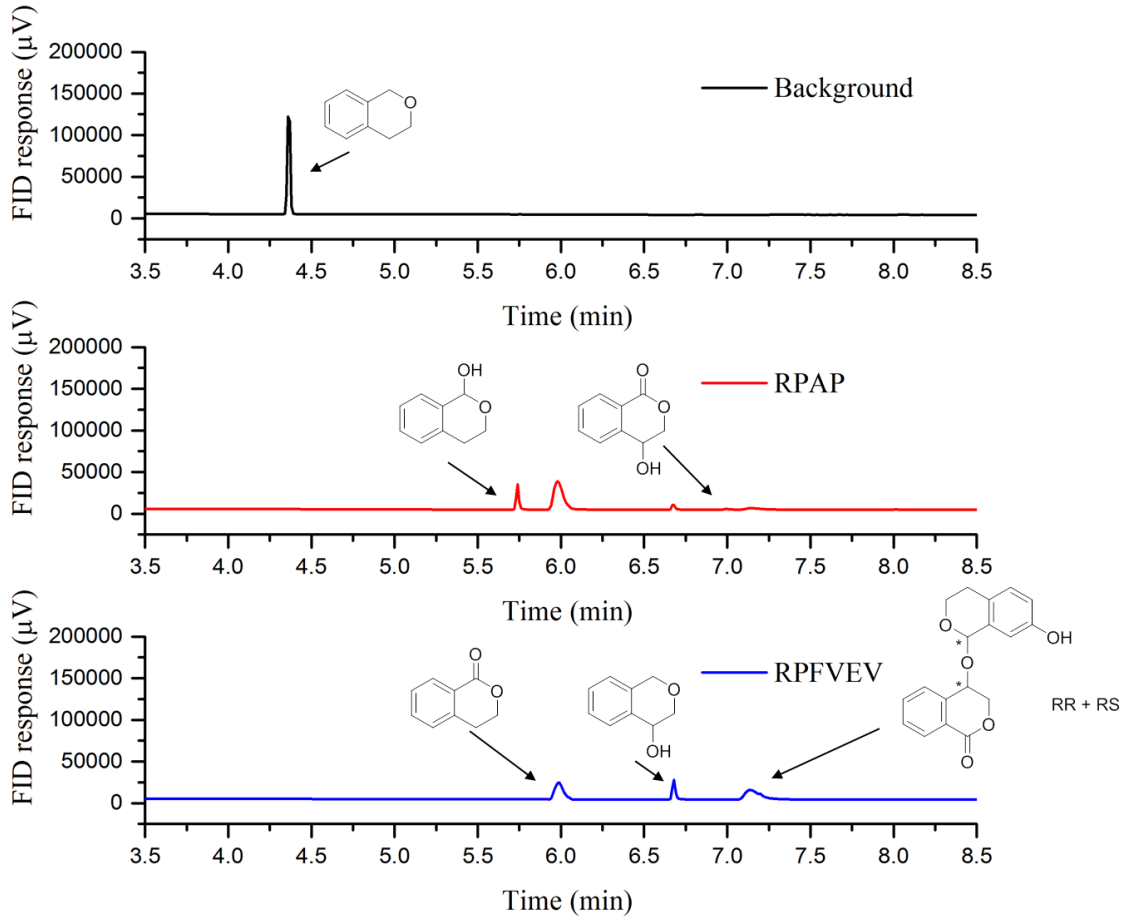


Figure 4.25 GC traces of three 13.4 mg prep-scale reactions with RP, RP/AP and RP/FV/EV

The RP/AP mutant was then used in an *in vivo* reaction (200 mg, 600 mL FB_{gly} culture). The time course study of which is shown in Table 4.4. In the *in vivo* reaction, isochroman-1-ol (**4.15**) was gradually converted into further oxidation products (1-one and 1-ol-4 one) and ultimately being converted into dimers. On the other hand, the proportion of the 4-ol (**4.17**) did not change significantly. The reaction mixture was extracted with ethyl acetate after 24

hours and the products were purified by silica gel chromatography. Three dimer compounds were co-eluted through the column but two of them were characterised by Dr Nick Rees at Oxford as (*R*,*R*)- and (*R*,*S*)- 4-((7-hydroxyisochroman-1-yl)oxy)isochroman-1-one (**4.19**, **4.20**) from the NMR of the crude mixture.

Table 4.4 Isochroman oxidation product profiles by RP/FV/EV and RP/AP mutants

Products	4-ol	1-ol	1-one	1-one-4-ol	Dimers
RP/AP 6h *in vitro*	27.9%	37.1%	24.7%	10.3%	0.0%
RP/AP 24h *in vivo*	26.4%	12.4%	18.9%	33.1%	9.2%

4.3.3 ε-Caprolactone

ε-Caprolactone (**4.21**) is a cyclic ester with a seven-membered ring that belongs to the lactone family. It is generally used for the large-scale production of caprolactam by treating it with ammonia at elevated temperatures. It is also an important monomer used in the manufacture of highly specialised polymers, for example, polyurethanes (as shown in Figure 4.26).

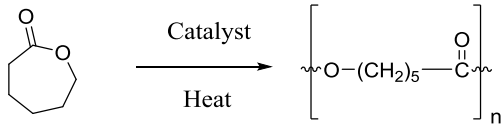


Figure 4.26 Ring opening polymerization of ε-caprolactone

A panel of 32 $P450_{BM3}$ mutants were tested with ε-caprolactone. The screening reaction were conducted at 0.5 mL in 200 mL phosphate buffer (pH 7.5), with 2 mM substrate and 1 μM enzyme. The reactions were quenched after 2 hours by extracting the reaction mixture with 300 μL ethyl acetate and the organics were analysed by gas chromatography (as shown in Figure 4.28 and Table S4.7 in Appendix).

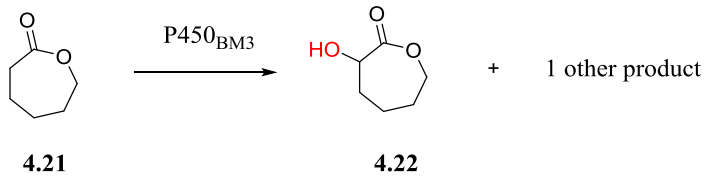


Figure 4.27 $P450_{BM3}$ reactions with ε-caprolactone

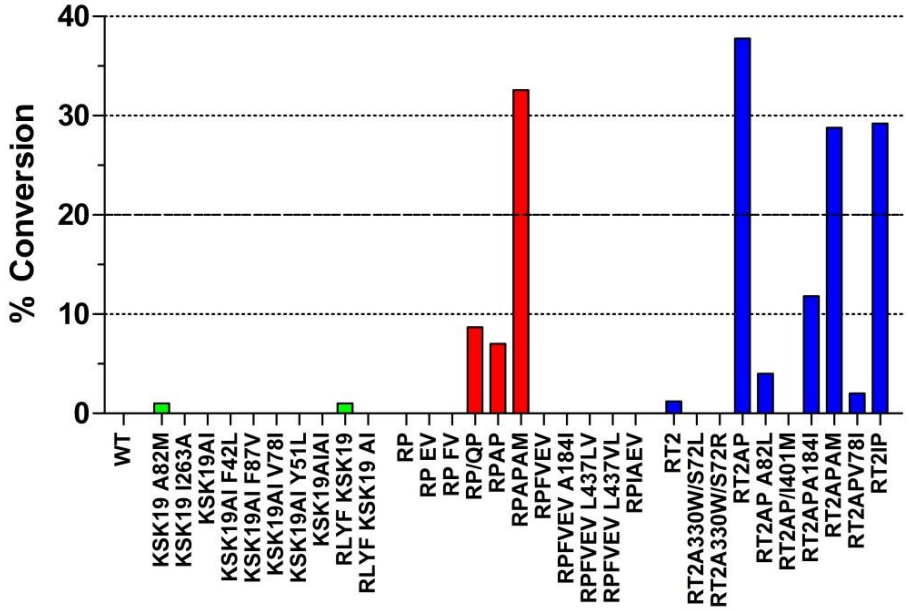


Figure 4.28 Substrate conversion of ε-caprolactone (**4.21**) by $P450_{BM3}$ variants

Wild type (WT) showed no activity towards ε-caprolactone (**4.21**) and this is the same for most of KSK19 based $P450_{BM3}$ variants. Mutants incorporating A330P mutation showed reasonable conversion whereas a tryptophan substitution at the 330 position abolished activity. RT2/AP gave the highest conversion (38%) among the mutants tested but the introduction of the I401M mutation dramatically suppressed enzyme activity. A preparative reaction was performed with the RT2/AP mutant *in vitro* (as shown in Figure 4.29) and the major product was likely to be 3-hydroxy-ε-caprolactone (**4.22**) due to the triplet at 4.33 ppm on ^{1}H NMR of crude reaction mixture however it is not fully characterised. It was problematic to make a large amount of **4.22** as polymerisation can take place in aqueous system which may result in the loss of compound.

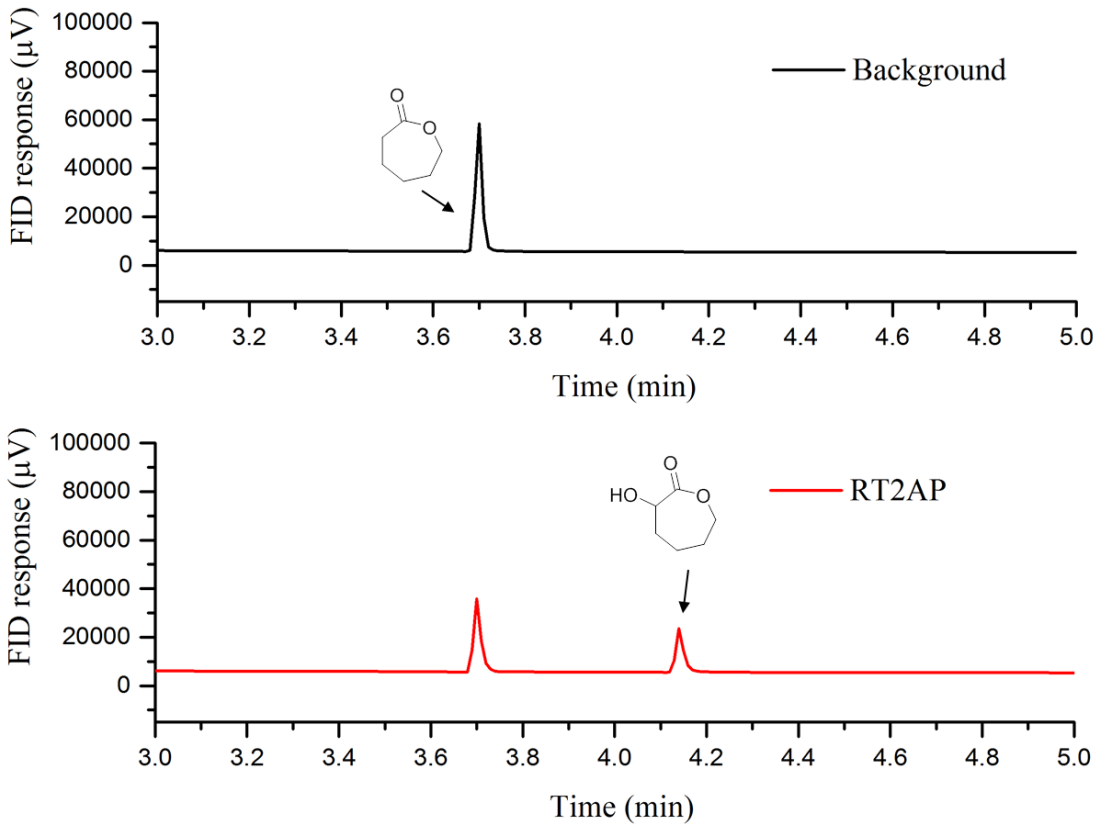
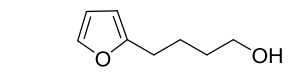


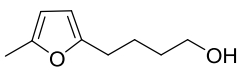
Figure 4.29 GC trace of prep-scale reaction of ε-caprolactone with RT2/AP

4.3.4 Other *O*-heterocycles

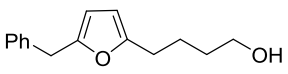
Furan derivatives are versatile synthetic intermediates that can be transformed by a wide range of reactions into structurally diverse molecules found in numerous natural products. Due to the electron-rich nature of furans it readily undergoes oxidation. This transformation is a particular useful method for generating structural complexity.



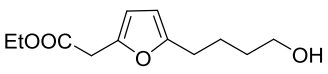
4.23



4.24



4.25



4.26

Figure 4.30 Furan derivatives synthesized by Jessica Thien (Robertson group)

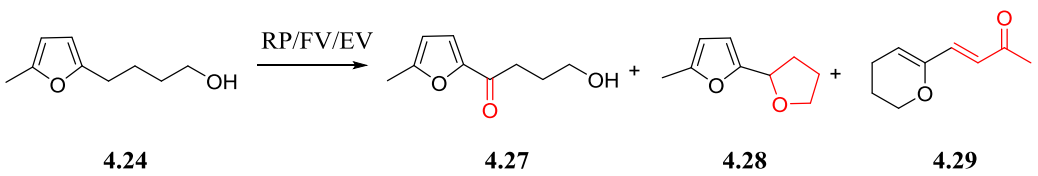


Figure 4.31 The oxidation of **4.24** by $P450_{BM3}$ variants

Four furan derivatives (Figure 4.30, compounds **4.23**, **4.24**, **4.25**, **4.26**) were prepared by Jessica Thien from the Robertson group at Oxford and tested with three mutants RP, RP/AP, and RP/FV/EV. The RP/FV/EV mutant at 1 μM gave 100% conversion of 10 mM of compound **4.24** while other compounds were not transformed at high degree of conversion. (Figure 4.31) This reaction was scaled up to 20 mg and products were purified by prep-TLC. The main products were found to be the ketone (**4.27**), ring formation (**4.28**) and ring rearrangement (**4.29**) products (characterised by Jessica Thien).[395] The suggested mechanisms for their formation are shown in Figure 4.32.

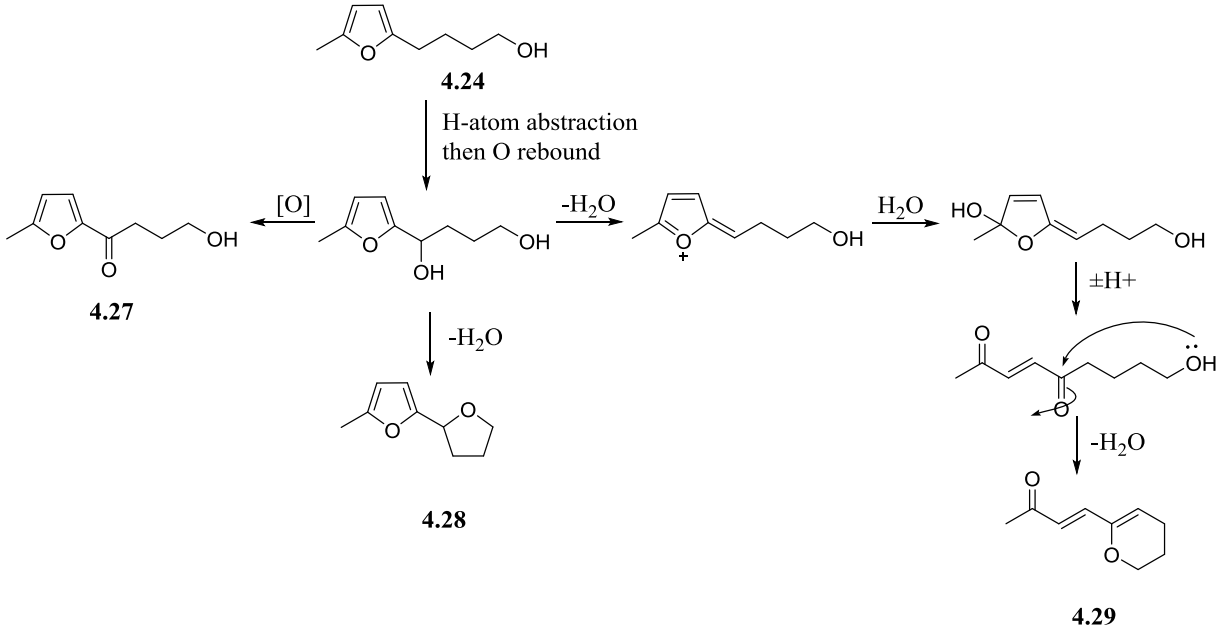


Figure 4.32 Proposed mechanism for the oxidation of **4.24** by $P450_{BM3}$

4.4 Alcohols and phenols

4.4.1 2-cyclopentylethanol

2-Cyclopentylethanol (**4.30**) is a potential precursor to binucleotide-like drugs and prostaglandins. When oxidising or functionalising this primary alcohol with traditional synthetic reagents, the trivial oxidation of the reactive alcohol group is inevitable in most cases. Hence, the use of $P450_{BM3}$ for direct hydroxylation on the cyclopentyl ring without alcohol oxidation will offer novel routes for structural functionalisation as well as new handles for chemical elaboration.

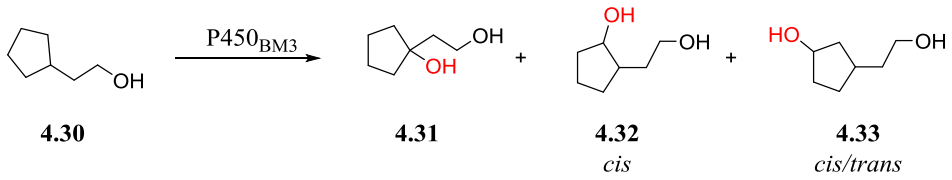


Figure 4.33 $P450_{BM3}$ reactions with 2-cyclopentylethanol

The screening reactions were carried out with a panel of 27 mutants. GC analysis showed a range of 9 products being formed throughout the mutant library while multiple mutants (RP/AP/AM, RP/FV/EV, KSK19/AI/FV, RT2/AP/AM, RT2/IP, RT2/AP, RT2/AP/A184I, RT2/AP/V78I, RP/IA/EV, RP/AP, RP/FV) were found to fully convert 2-cyclopentylethanol with varied product selectivity (Appendix Table S4.8, Figure 4.34). Wild type $P450_{BM3}$ showed 15% substrate conversion and gave four products. The KSK19/IA mutant was inactive, suggesting substrate binding far away from the haem due to space created by the I263A mutation while the A330P mutation promoted substrate conversion since mutants incorporating A330P showed high conversion in most cases (>80% except RT2/AP/I401M).

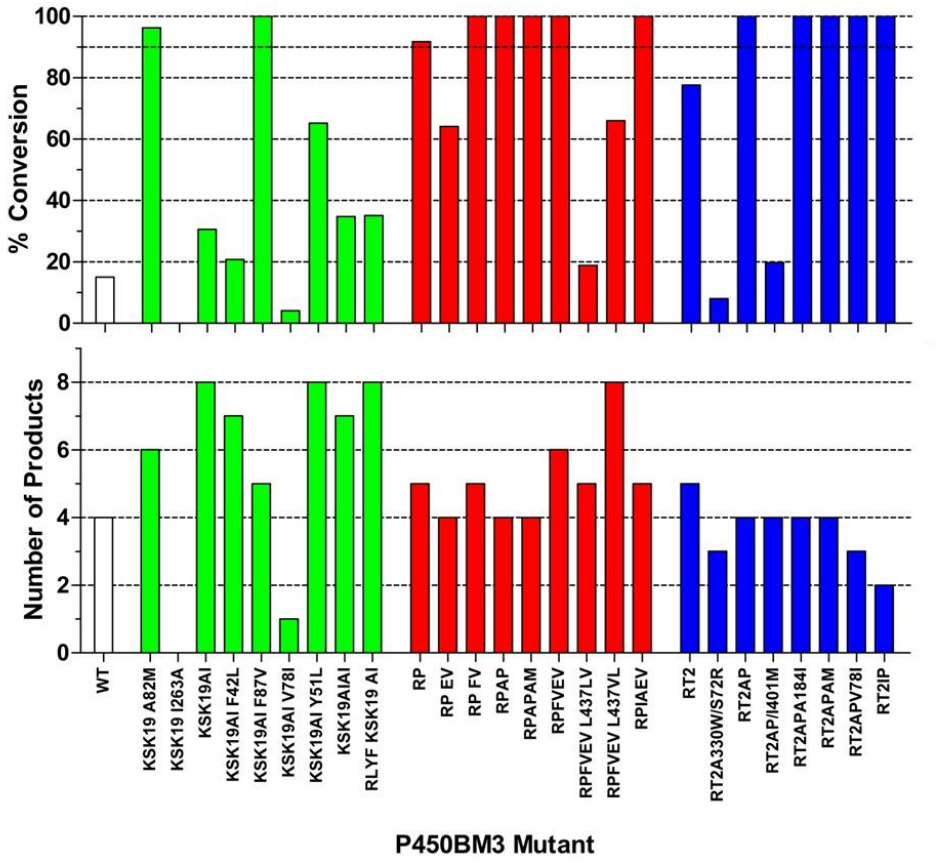


Figure 4.34 Substrate conversion and number of products formed from the oxidation of 2-cyclopentylethanol (**4.30**) by $P450_{BM3}$

Of the mutants that showed total substrate conversion, RT2/IP was the most selective enzyme for the product with the retention time 4.58 min. RT2/AP/A184I gave the highest proportion of the 3.32 min product at 37%. RP/FV/EV displayed highest selectivity (41%) towards the 3.57 min product while KSK19/AI/FV generated largest yield (38%) of the product at 4.55 min. RP/FV/EV was selected to perform a 20 mg preparative scale reaction (GC trace is shown in Figure 4.35) and enabled all the products to be identified with the NMR of crude mixture by comparison with published data. [396-398]

The products were identified as 1-(2-hydroxyethyl)cyclopentan-1-ol (**4.31**, GC retention time 3.32 min), diastereoisomers of *cis*-2-(2-hydroxyethyl)cyclopentan-1-ol (**4.32**, GC retention time 3.57 min) and diastereoisomers of both *trans*- and *cis*- 3-(2-hydroxyethyl)cyclopentan-1-ol (**4.33**, GC retention time 4.55 min, 4.58 min, 4.62 min, 4.64 min).

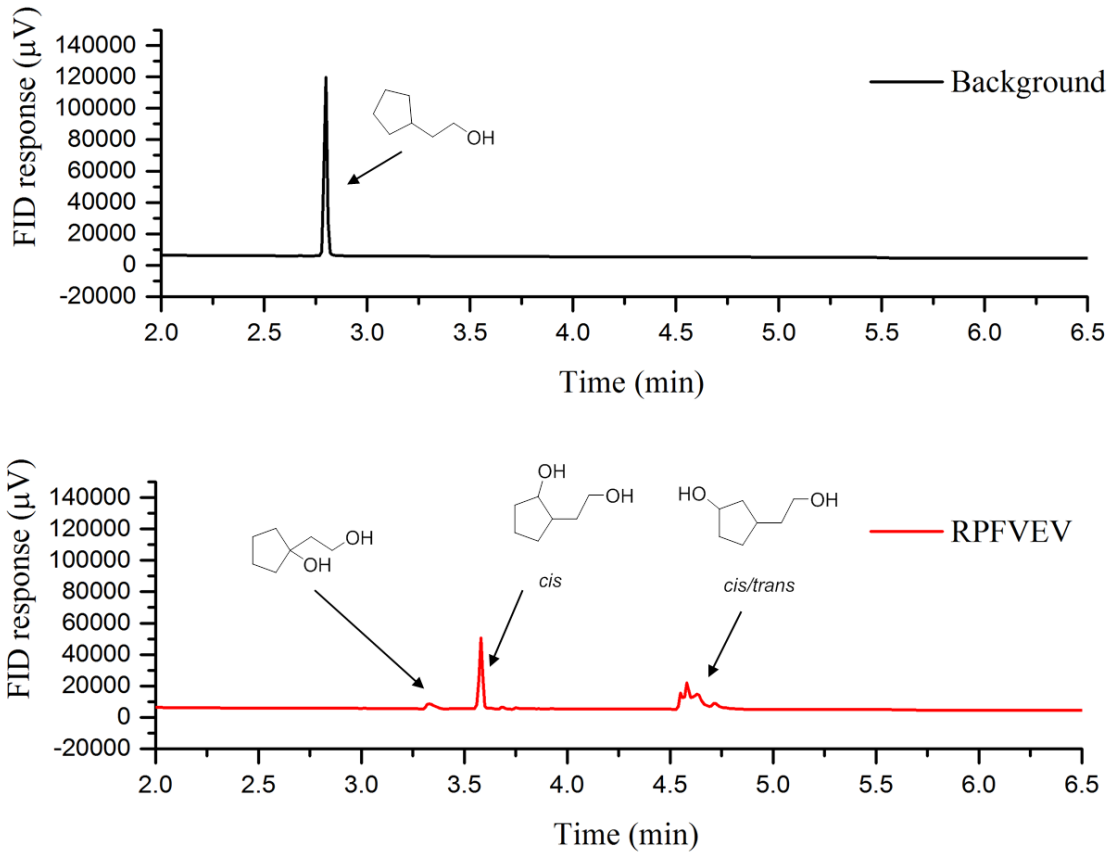


Figure 4.35 GC trace of 20 mg prep-scale reaction of 2-cyclopentylethanol (**4.30**) with the RP/FV/EV mutant

Another 3 products were also seen from the GC analysis including two early retention time products (2.67 min, 2.75 min) and one late product (4.09 min). The two early products may be bicyclic lactones formed from intramolecular cyclisation of hydroxylation products of **4.30** with the loss of water. The observation of hydroxylation at all three positions on the cyclopentyl ring proved the possibility of developing novel methodologies for direct functionalisation of primary alcohols with alcohol groups remaining intact. This collection of mutant enzymes can be used as a good basis for further investigations and protein engineering to identify highly regio- and diastereo- selective variants of $P450_{BM3}$.

4.4.2 7-Methyl-2,3-dihydro-*1H*-inden-4-ol

7-Methyl-2,3-dihydro-1*H*-inden-4-ol (**4.34**) can be seen as a trialkyl substituted phenol and a good comparison with a primary alcohol. Phenols are chemically reactive and susceptible to a wide range of oxidants to generate benzoquinones. As such, the use of $P450_{BM3}$ may offer a new route for direct functionalisation of aliphatic C–H bonds with the phenol alcohol being intact.

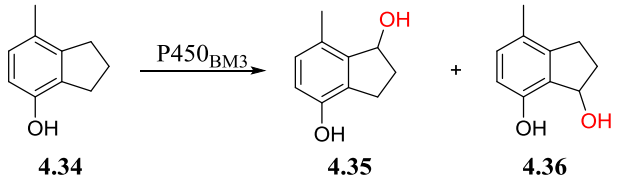


Figure 4.36 $P450_{BM3}$ reactions with 7-methyl-2,3-dihydro-1*H*-inden-4-ol **(4.34)**

A panel of 31 $P450_{BM3}$ variants were screened with **4.34** under the same conditions as 2-cyclopentylethanol (**4.30**). The substrate conversion and number of products formed from screen reactions are shown in Figure 4.38 while the full product distribution data are shown in Appendix Table S4.9.

Wild type (WT) $P450_{BM3}$ was again inactive towards the starting compound, which is consistent with most non-natural substrates that have been tested. High conversions were observed with both the KSK19 and KT2 series of mutants into 2 major products but the RP family mutants showed low activity. RT2/AP/A184I gave the highest conversion throughout the entire mutant library while GVQ converted the substrate 99% with a completely different products profile, producing 96% product at 8.02 min that was significantly higher than RT2/AP/A184I (17%). The most selective mutant was found to be KSK19/AI as it was able to transform 75% of substrate into a single product (GC retention time 8.02 min), which was then isolated from a 10 mg reaction and characterised as 4-methyl-2,3-dihydro-1*H*-indene-1,7-diol (**4.36**, as shown in Figure 4.38). The other major product was isolated from another

10 mg reaction using RT2/AP/A184I which gave 79% selectivity and the structure was elucidated as 7-methyl-2,3-dihydro-1*H*-indene-1,4-diol (**4.35**)

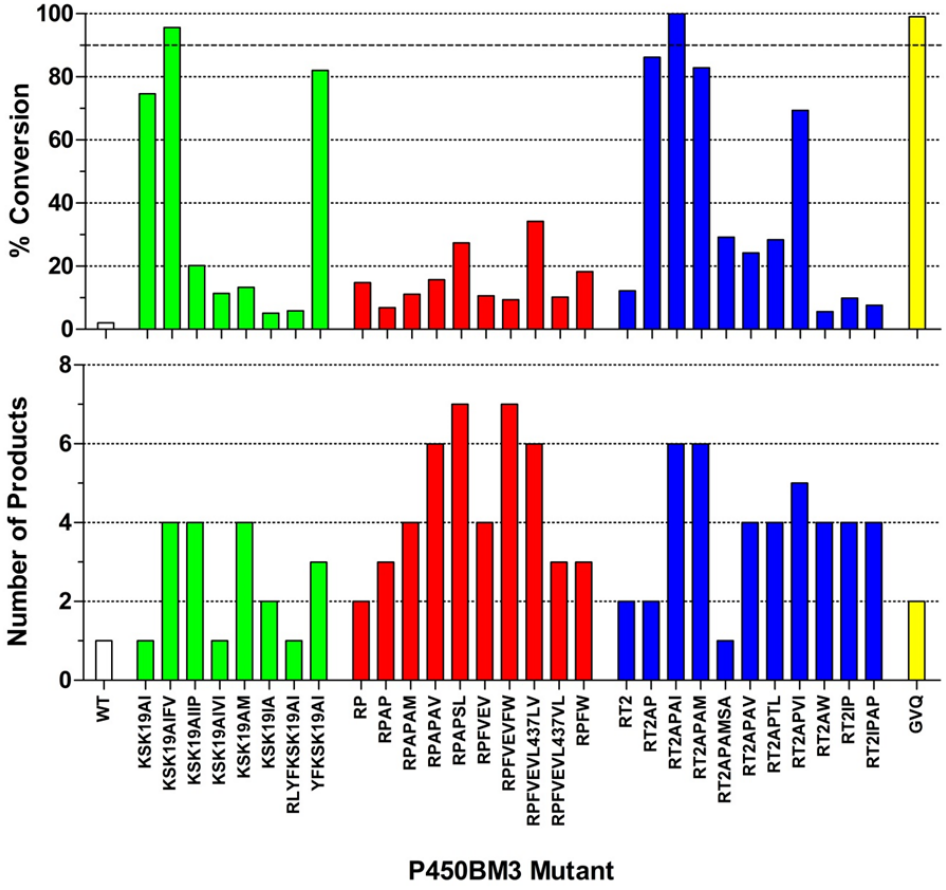


Figure 4.37 Substrate conversion and number of products formed from the oxidation of 7-Methyl-2,3-dihydro-1*H*-inden-4-ol (**4.34**) by $P450_{BM3}$ variants

From the GC analysis, 8 minor products were also formed. However, their structures remained unknown. One product had an earlier retention time (5.72 min) than the substrate (6.11 min) while the other 7 products had retention time between the substrate and **4.36**. Considering the number of C–H bonds within the substrate molecule (**4.34**), the minor products could arise from the mono-hydroxylation at various sites while some of them may be lactones, leading to functional diversity. Further mutagenesis is needed to shift the bias from the oxidation at benzylic positions (**4.35** and **4.36)** towards these minor products especially for the oxidation at non-activated C_3 position which would be of particular interest.

Oxidation at benzylic methyls or phenols were not the major products among the mutant enzymes being tested and these products could be the targets for further engineering of $P450_{BM3}$ enzymes.

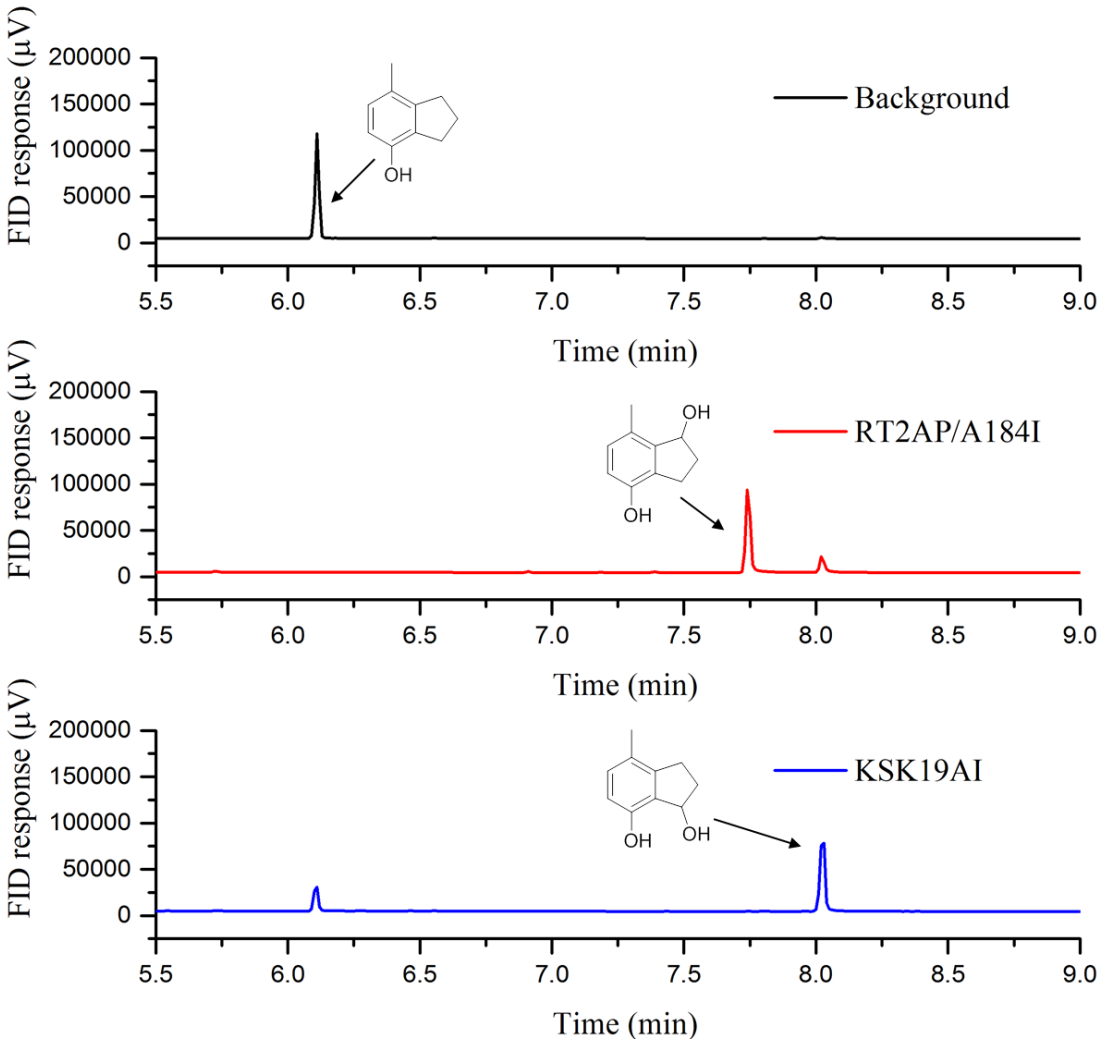
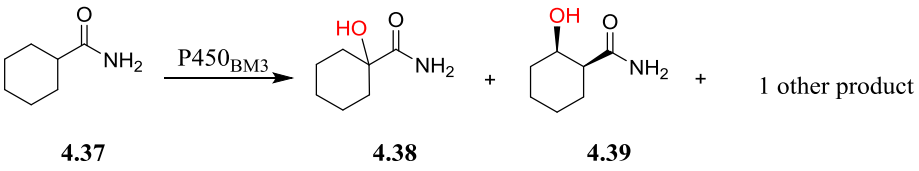


Figure 4.38 GC traces of 10 mg prep-scale reactions of 4.34 with RT2/AP/A184I and KSK19/AI

4.5 Amides

4.5.1 Cyclohexanecarboxamide

Cyclohexanecarboxamide (**4.37**) is a common drug fragment molecule with its derivative *N*-phenyl-*N*-[1-(piperidine-1-carbonyl)cyclohexyl] benzamide (MNRC-5) having inhibitory effects on Schistosoma mansoni cercarial serine protease and cercarial penetration.[399] Cyclohexanecarboxamide (**4.37**) may undergo a wide variety of side reactions, including Hofmann-type rearrangements, dehydration to the nitrile compound and *N*-oxidation when being subjected to chemical oxidants. The use of $P450_{BM3}$ enzymes aimed to avoid these undesirable reactions in favour of C–H bond activation with controlled selectivity.



4.37 **4.38** **4.39**

Figure 4.39 $P450_{BM3}$ reactions with cyclohexane carboxamide **(4.37)**

The initial screening reactions were conducted with a collection of 32 $P450_{BM3}$ mutants (As shown in Appendix Table S4.10 and Figure 4.40). A total of 10 variants including wild type $P450_{BM3}$ (WT) gave no conversion while 24 variants generated conversions lower than 50%. The small substrate size could be the key factor for poor activity since the starting material could be bound far away from the haem group. None of KSK19 based mutants showed good activity, indicating the importance of the F87 side chain. Meanwhile, amino acid substitutions with larger side chain volumes at A330 and A82 significantly enhanced the enzyme activity and led to the identification of 2 mutants (RT2/AP, RT2/AP/A82L) showing 100% conversion. RT2/AP generated two major products in roughly 1:1 ratio whilst the introduction of A82L shifted the selectivity towards the 6.47 min product with 88% selectivity.

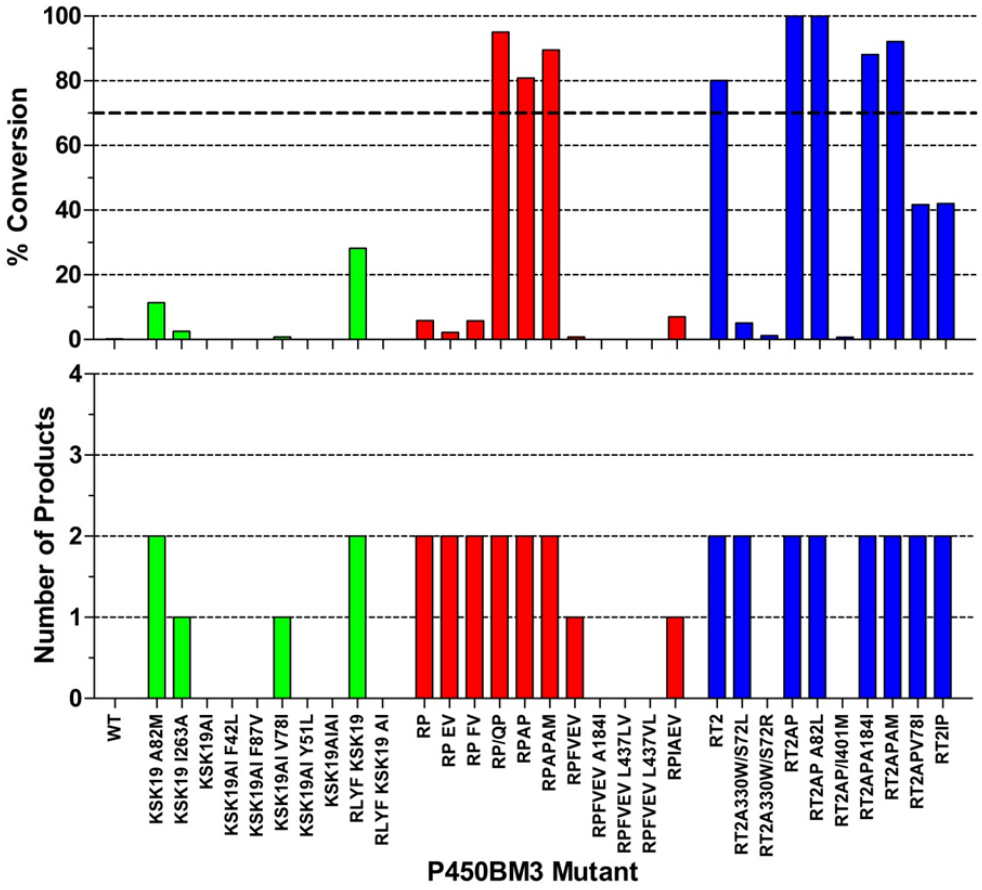


Figure 4.40 Substrate conversion and number of products formed from the oxidation of cyclohexanecarboxamide (**4.37**) by $P450_{BM3}$ variants

A total of three products were observed from the GC analysis including 2 major products and 1 minor product which was only formed in small quantities with 3 mutants (KSK19/A82M, KSK19/I263A, KSK19/AI/V78I). The two major products were then isolated from a 20 mg reaction with RT2/AP (Figure 4.41) and characterised as 1-hydroxycecylohexane-1-carboxamide (**4.38**, GC retention time 5.67 min) and the *cis* isomer of 2-hydroxycyclohexane-1-carboxamide (**4.39**, GC retention time 6.47 min). Formation of **4.38** was confirmed by the loss of the H atom signal at the C1 position on ^{1}H NMR while the *cis* isomer of **4.39** was confirmed by the loss of *trans* coupling between the H atoms at the C1 and C2 positions. However, the enantioselectivity was not determined due to the inaccessibility to a chiral phase GC instrument. Further research into the enantioselectivity

formed by various mutants would be very useful for drug design and discovery. Highly regio-, enantio- and diastereo- selective mutants will offer a novel route for synthesising cyclohexanecarboxamide derivatives or other molecules with more lead-like property.

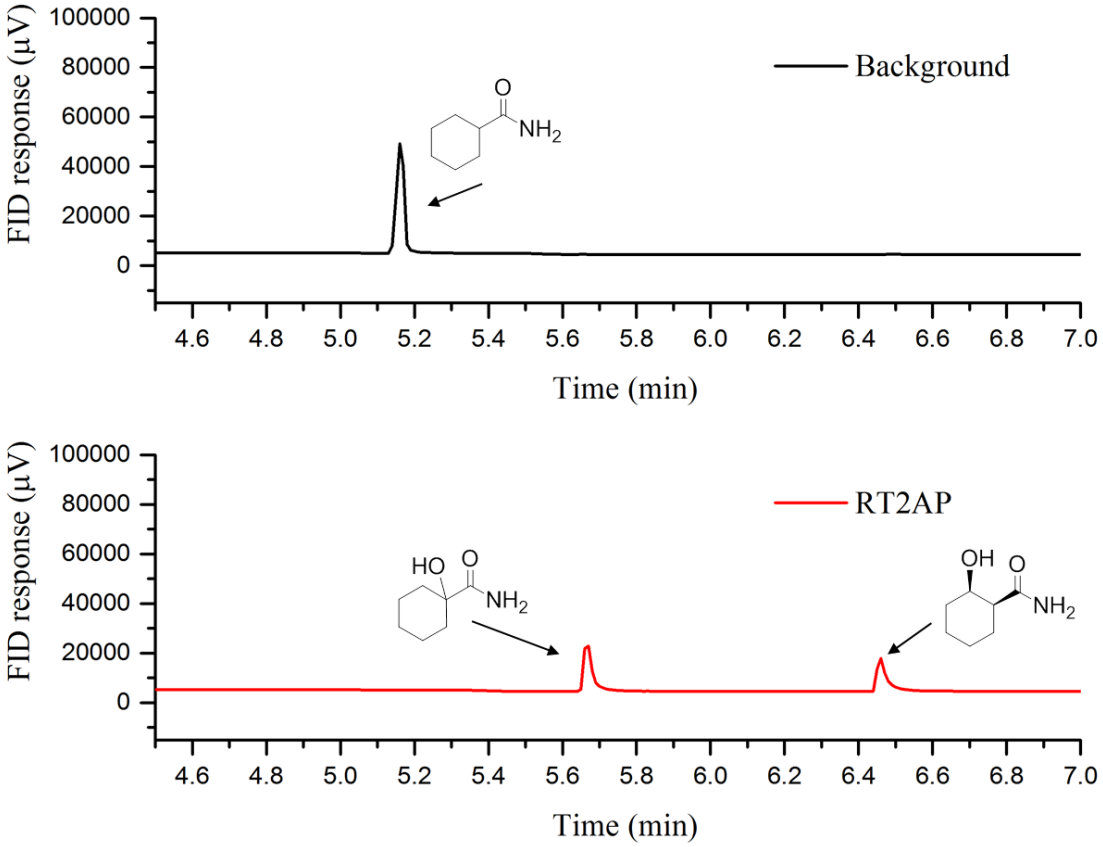


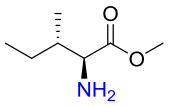
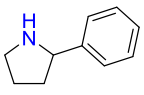
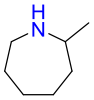
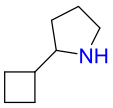
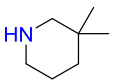
Figure 4.41 GC trace of 20 mg scale reaction with RT2/AP

4.6 Unprotected amines

Unprotected amines play important roles in drug discovery as many drug molecules contain an amine group. However, they are challenging substrates for oxidants including P450 enzymes because they are easily oxidised to amine oxides. Among the 16 aliphatic amines from the GSK panel of 40 drug fragments, 5 were detectable on the GC while the other 11 were detected by TLC as unprotected amines are easy to be oxidised by potassium permanganate.

Therefore, the unprotected amines were screened with a smaller panel of $P450_{BM3}$ mutants (RP, RP/AP, RP/FV/EV, RP/FV/EV/L437LV, KSK19/AI, KSK19/AI/FV, RT2/AP, RT2/AP/V78I, RT2/AP/A184I). It was found that unprotected amines are well-tolerated by the enzyme system, resulting in multiple oxidised products with most amine substrates. Table 4.5 lists a panel of unprotected amine molecules and number of products formed by specific mutant.

Table 4.5 Unprotected amines oxidised by $P450_{BM3}$ variants

Compound					
Number of Products	3-4	5	3-4	3-4	4
Mutant	RT2AP/V78I	KSK19AIFV	RP/FV/EV	RT2AP	RT2AP

L-isoleucine methyl ester was selected for scale-up and transformed to three organic-soluble products with GC retention times consistent with hydroxylation products. However, only 20% conversion was observed with an *in vitro* reaction using RT2/AP/V78I. It is particularly notable that the unprotected primary amine was not oxidised to the amine oxide and the ester was not demethylated, as would be expected in a reaction with any chemical oxidising agent. Unfortunately, due to the strong interaction of amines with silica columns, low conversion (~20%) for most unprotected amines, detection issues and limited amount of materials, their products have not yet been characterised.

4.7 Carboxylic acids

Carboxylic acids were tested for the substrate tolerance of the $P450_{BM3}$ variants but had to undertake a different analysis procedure. The reaction mixtures were acidified by HCl (aq,

1M) before extraction into ethyl acetate and the carboxylic acid was converted to the –$OSiMe_3$ ester, and any alcohol group to the trimethylsilyl ether using a 99:1 mixture of BSTFA/TMCS. These neutral derivatives gave sharp peaks on the GC. Figure 4.42 shows an example of a carboxylic acid reaction with RP and RP/AP mutants.

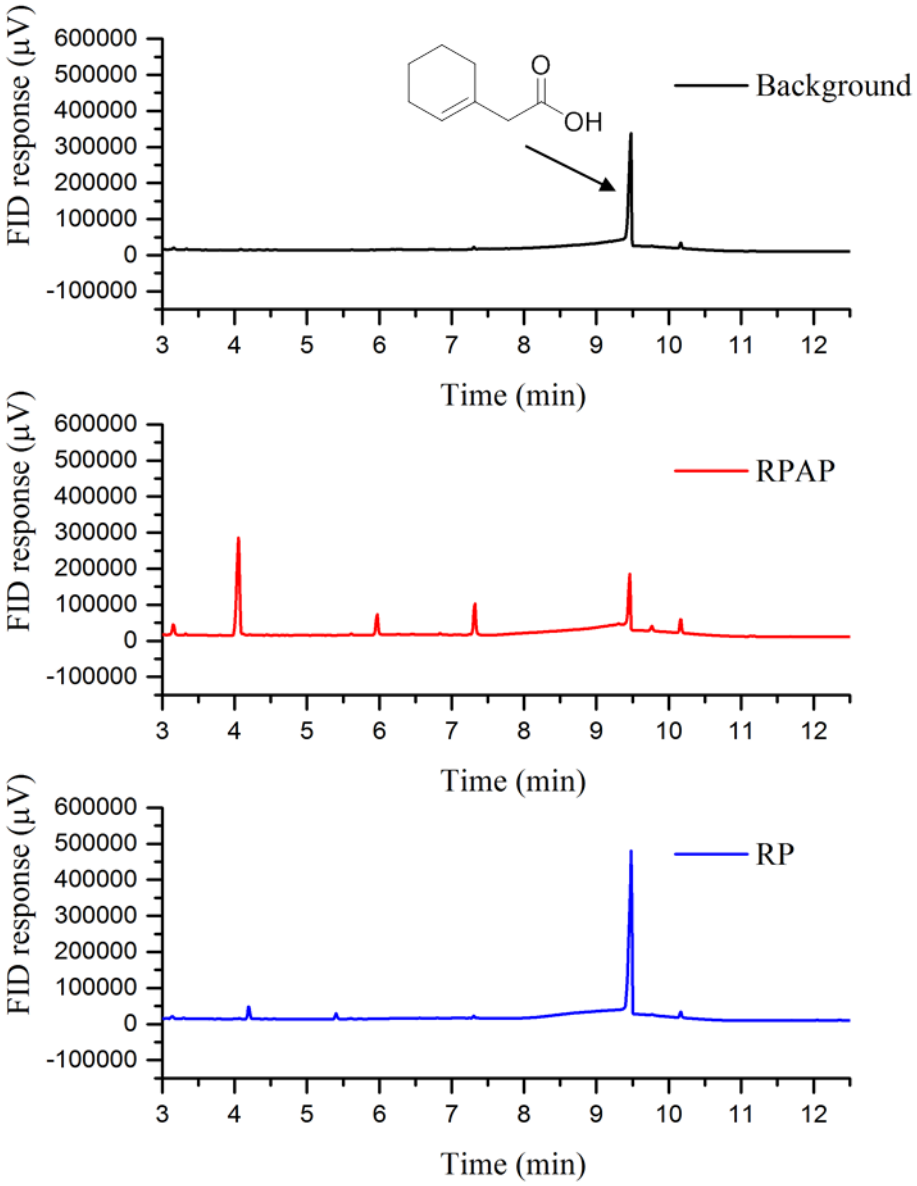
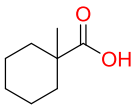
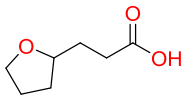
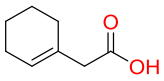
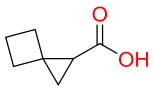


Figure 4.42 GC trace of 2-(cyclohex-1-en-1-yl) acetic acid with RP and RP/AP

The screening was conducted using the same panel of $P450_{BM3}$ mutants due to the complex analysis procedure. Among the eight carboxylic acids in the GSK panel, seven were found to be oxidised to multiple products. Some examples are shown in Table 4.6. No product has

been characterised yet because of the low conversion and lack of compounds. RP/AP and RT2/AP/A184I seem to be the best base mutant for the carboxylic acids since both of them have the A330P mutation, which may help to create a hydrophobic environment and push the carboxylic acids molecules close to the haem iron. More protein engineering needs to be done to improve the conversion level. However, these initial trials indicate that small carboxylic acids molecules are viable substrates for the $P450_{BM3}$ generic accelerators. The use of LC-MS will also speed up screening towards carboxylic acids and product characterisation.

Table 4.6 Carboxylic acids oxidised by $P450_{BM3}$ variants

Compound				
Number of Products	2	5	3-4	4
Mutant	RT2/AP/A184I	RP	RP/FV/EV	RP/AP

4.8 Nitriles

Nitrile compounds are selected to test the functional group tolerance and chemoselectivity of $P450_{BM3}$ mutants. 1-Phenylcyclopropanecarbonitrile contains both a cyanide group and a cyclopropane ring, which will offer multiple sites for the ferryl speicies to attack. A panel of 15 $P450_{BM3}$ variants were screened for the activity towards this compound and the RT2/AP/V78I mutant showed 85% conversion into 7 major products (Figure 4.43). No products have been characterised yet due to the lack of starting material. However, this is an initial hit and a proof-of-concept for nitrile compound being multiply functionalised by $P450_{BM3}$ enzyme library. Once the products are isolated and characterised, further work can be done on protein engineering for the purpose of improving product selectivity.

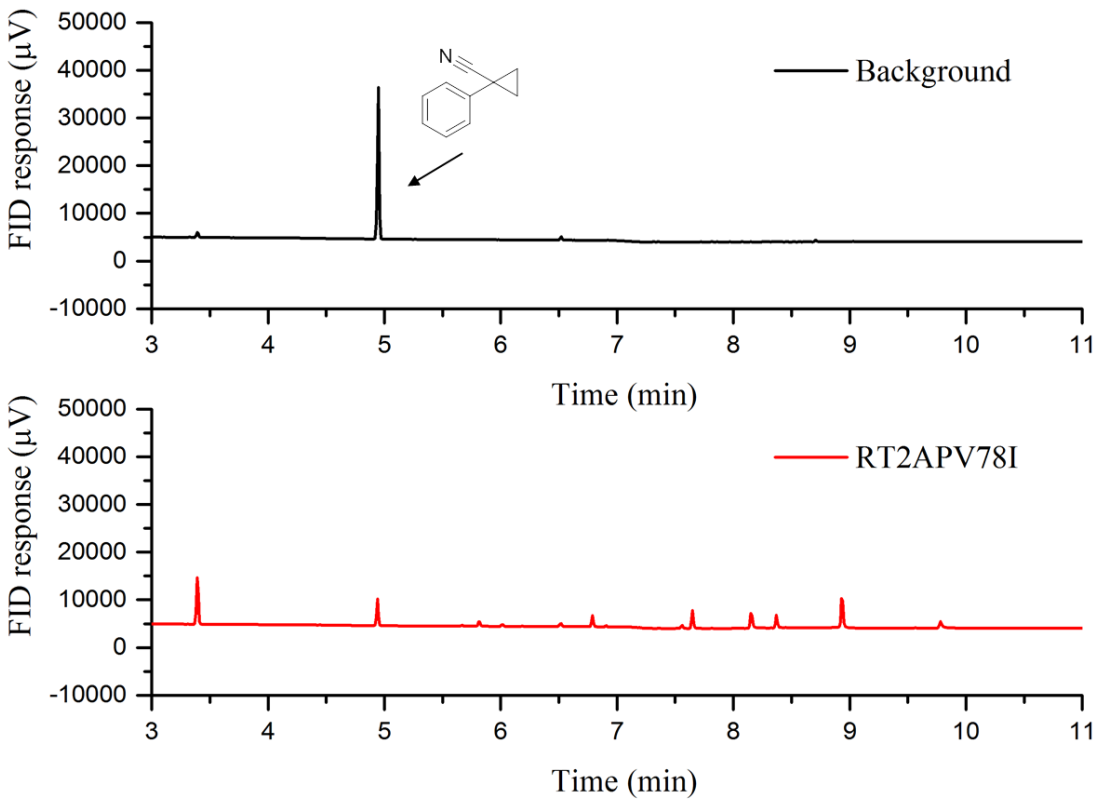


Figure 4.43 GC trace of 1-phenylcyclopropanecarbonitrile being oxidised by RT2/AP/V78I

4.9 Chapter summary and conclusions

Over the last several decades, the drug discovery community has better defined the properties of molecules that are most likely to successfully deliver clinical drugs. However, finding a methodology for synthesising large numbers of small, polar molecules with lead-likeness rather than making heavy, non-polar ones remains a challenge for the medicinal chemists. This chapter is a proof-of-concept research to assess the viability of using engineered $P450_{BM3}$ variants to oxidise drug fragment molecules at various sites thus creating functional diversity.

40 drug fragment molecules representing a variety of subclasses (*N*-heterocycles, *O*-heterocycles, amines, amides, alcohols, carboxylic acids and nitriles) were selected by GSK for the evaluation. Product formation screening were conducted with a library of $P450_{BM3}$ mutants developed on the basis of three generic accelerators (KSK19, RP and KT2)[21, 358] with

the addition of hydrophobic substitutions at key positions such as A330, A184, V78, A82, F87, E267, and etc., which have shown positive results (hits) with 36 compounds thus far, while the 10 molecules were studied in detail with their ease of being analysed by gas chromatography.

Wild type $P450_{BM3}$ is generally inactive towards the 10 fragments that were evaluated while total conversions were observed with generic accelerator mutants towards virtually all of the compounds except acetylpyrazine and ε-caprolactone, suggesting the effectiveness of mutant enzyme in overcoming the issues of substrate inhibition and poor substrate binding. The reactions were scaled up to 10–250 mg both *in vivo* and *in vitro*, which showed similar conversions to the screening reactions, indicating the scalability of using such biocatalysts and enabled products to be isolated and characterised at preparative scale. Multiple products were generated as excepted although the number of products varied widely for each compound depending on the number of available oxidation sites and the ease with which they could be hydroxylated. In contrast to traditional oxidation reactions, chemically vulnerable groups (alcohols, amines, etc.) remained intact through the enzymatic transformations, making the use of $P450_{BM3}$ enzymes particularly attractive for synthesis as protection-deprotection strategies are not necessary.

However, some of the products remained uncharacterised due to the limited supply of compounds. Further work should be done on their structure elucidation and the use of GC/MS and LC/MS would significantly aid product characterisation, particularly for the amines and carboxylic acids which need a complex analytical procedure by GC.

In conclusion, the overall results can be recognised as an exciting beginning and $P450_{BM3}$ enzyme system have shown its potency of being developed into a viable LOS methodology. Unlike traditional organic synthesis methodologies that are often accompanied with harsh conditions and unmanageable generation of side products, $P450_{BM3}$ reaction has its

preponderance of high efficiency, mild conditions, manageable selectivity and environmental friendliness. Large numbers of hydroxylation products can be generated by $P450_{BM3}$ enzyme system that offers a diverse functionality and therefore can be used as a rich source for drug discovery screening. Further investigation into larger lead molecules or previous failed drug candidates with heavy atom counts (HAC) in the range of 11 to 17 would give greater applicability for potential drug-like precursors. The accumulation of such information on selectivity and activity trends for the different mutations would be exceptionally useful for guiding mutant design and targeting residues for mutagenesis.

Chapter Five

Chapter 5 New function studies and other applications of $P450_{BM3}$

5.1 New function studies on intramolecular C–H amination

5.1.1 Introduction

Nitrogen heterocycles play vital roles in the vast majority of drugs which drives chemists to search for efficient and selective methods for the oxidative formation of C–N bonds. Late-stage cyclisation *via* oxidative C–H amination is an attractive strategic choice in this regard. The two main approaches include using a variety of oxidants ($KMnO_4$, $Hg(OAc)_2$, $K_3[Fe(CN)_6]$, etc.) to create electrophilic carbon species that utilise the intrinsic nucleophilicity of nitrogen atoms to form C–N bonds,[400, 401] or generating an electrophilic nitrenoid intermediate by transition metal catalysts (Rh, Ru, Mn, Co, Fe, etc.) that can react with alkenes, nucleophilic atoms, and C–H bonds.[402, 403] However, these chemical routes require the presence of specific functional groups in the substrate and offer imprecise control over the regio- and enantio- selectivity of C–N bond formation.

Enzymes, the natural counterparts of chemical oxidants, act as complementary tools for generating new C–N bonds. For example, transaminases,[404] ammonia lyases,[405] P450TxtE nitrating enzyme[406] and amino acid dehydrogenases[407] target oxidised or chemically activated carbon atoms within the substrates (Figure 5.1).

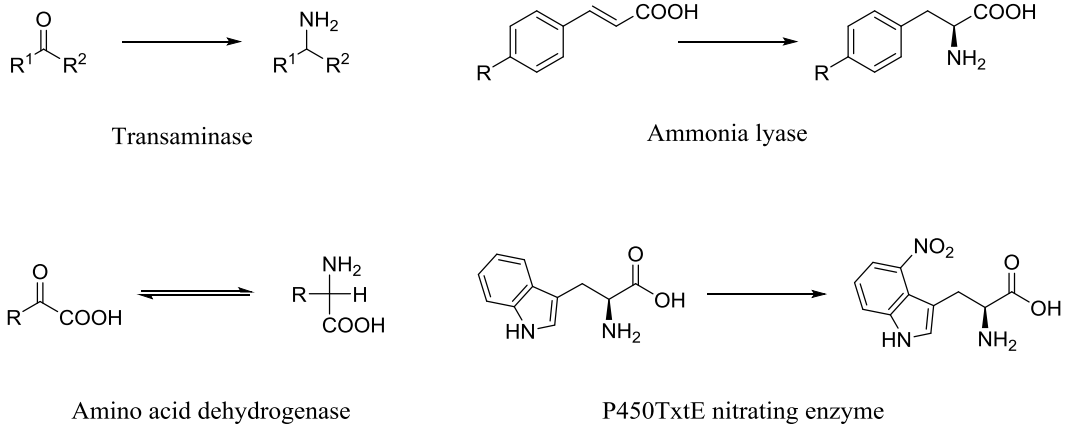


Figure 5.1 Naturally occurring enzymes capable of generating C–N bonds

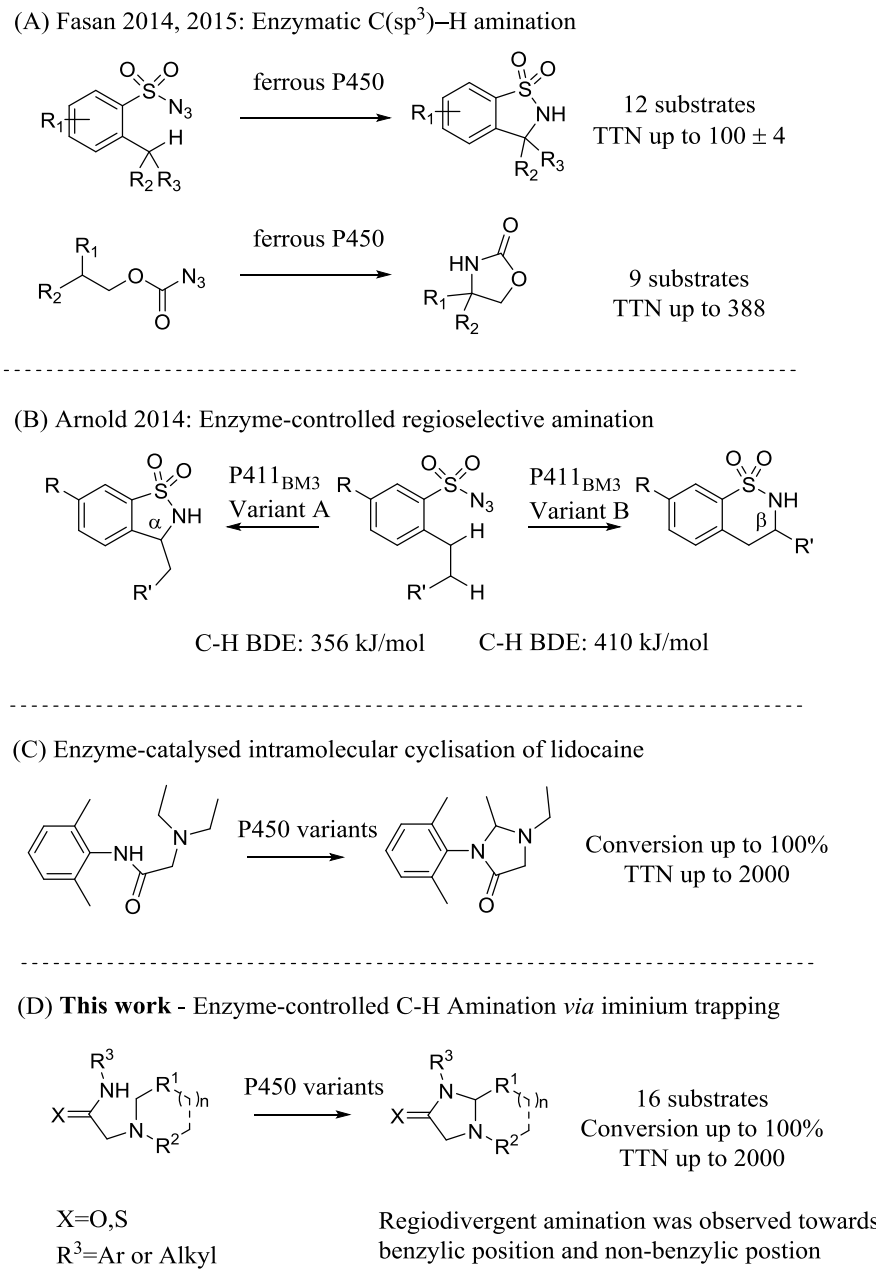


Figure 5.2 Enzyme-catalysed C–H amination

An alternative way for generating C–N bonds is to engineer naturally occurring enzymes to incorporate reactions which are not found in nature's repertoire. As developed from Breslow's pioneering studies,[408, 409] engineered P450 enzymes and myoglobin have been shown to activate azido groups in azidoformates and arylsulfonyl azides to effect nitrenoid formation and intramolecular C–H insertion (as shown in Figure 5.2A).[258, 410] Regioselectivity was achieved by Arnold and co-workers in the C–H animation of 2,5-

dipropylbenzenesulfonylazide at either the benzylic (Bond Dissociation Energy = 356 kJ/mol) or homobenzylic (Bond Dissociation Energy = 405 kJ/mol) position by different P450 mutants, with substrate binding overcoming the preference for insertion into the weaker C–H bond (Figure 5.2B).[411, 412]

Among the intramolecular C–H amination methodologies developed by Fasan and Arnold, they used cytochrome P450s and myoglobins as Fe(II) catalysts, which underwent similar mechanism as synthetic transition metal catalysts to activate nitrogen atoms and generate an nitrenoid intermediate reacting with C–H bonds. Therefore, azido functional groups are generally needed and anaerobic conditions are required in some cases.

Since P450s are well known to possess the ability to oxidise virtually any C–H bond in a variety of complex molecules with well-controlled regio- and enantio- selectivity, we made the hypothesis that P450s can be engineered to generate an electrophilic intermediate by activating carbon atoms.

As part of a larger study on drug metabolism by mutant cytochrome $P450_{BM3}$ (CYP102A1) enzymes, oxidative cyclization of lidocaine to an imidazolidin-4-one was conducted in high yield (Figure 5.2C),[413] which was also mentioned in Chapter 3. The formation of imidazolidinones, by *N*-cyclisation, rather than oxazolidinimines, by *O*-cyclization, was confirmed spectroscopically (see Figure 3.23) and, in the case of the lidocaine product, by comparison with a sample prepared by condensation with acetaldehyde.[414] This outcome accords with the instability of oxazolidinimines in aqueous solution and their conversion into the isomeric imidazolidinones by heating in pyridine.[415-417]

The proposed mechanism was to generate a α-hydroxyamine followed by the loss of hydroxide to form an iminium intermediate, which was then trapped by intramolecular nucleophilic attack by the amide nitrogen. Alternatively, the reaction may proceed *via* an iminium intermediate formed by α-hydrogen atom abstraction and electron-transfer oxidation.

Whatever the mechanism, iminium ion acts as the key activated carbon intermediate in C–N bond formation (Figure 5.3). Therefore, 2-aminoacetamides derived from both cyclic and acyclic amines were selected for the exploration of this intramolecular C–H amination process (Figure 5.2D). The construction of the so-formed imidazolidin-4-ones could be of interest to the pharmaceutical industry as these compounds are 3D templates with polar groups, hydrogen bond acceptors and potentially donors within the core. Further structural and functional group diversity around these cores may give compounds with varied biological activity. Such diversity covers more chemical space (e.g. for drug discovery) but requires the cyclisation catalyst to tolerate varied substituents at both the amine and amide ends of the substrate. Furthermore, the reduced pyrrolo- and pyrido-imidazolone motifs are validated biologically active cores found in the nootropic analgesic dimiracetam and the mannosidase I inhibitor kifunensine, respectively.

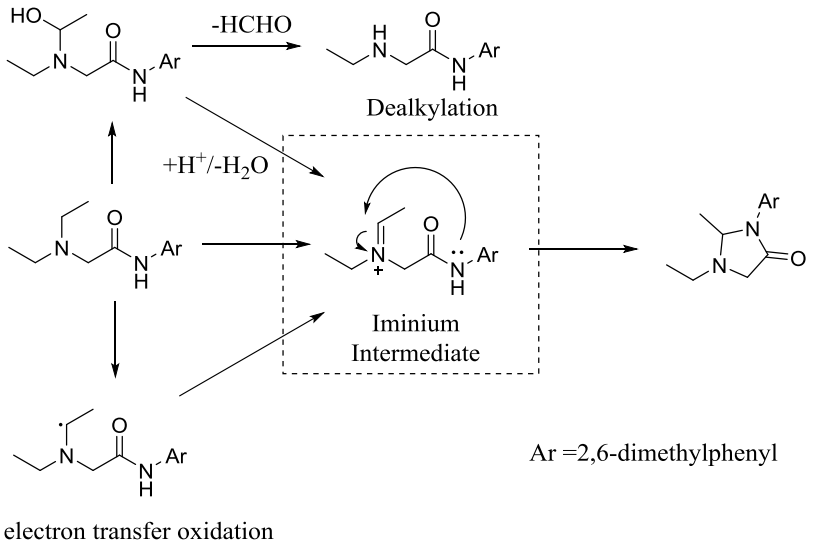


Figure 5.3 Enzyme-catalysed amination of lidocaine (proposed mechanism)

5.1.2 Results and discussion

Working in collaboration with Mr Jack O'Hanlon, lidocaine was selected as the first testing substrate on the intramolecular C–H amination. As shown in Table 5.1, full data see Table S3.6 in Appendix, 45 $P450_{BM3}$ mutants were screened towards lidocaine to give dealkylation

product (MEGX) and intramolecular C–H amination product, both of which can be formed from the breakdown of carbinolamine with either aldehyde elimination or the loss of hydroxide ion.

The product ratio was found to be dependent on the mutants. RP/FV/EV/F81W gave 96% of dealkylation product while RT2/A330W produced cyclisation product with total conversion. This observation indicates that the carbinolamine may breakdown within the enzyme active site otherwise the product ratio would be similar among all the mutant reactions. The partition between dealkylation and cyclisation is determined by the competition between iminium trapping by the Fe^{III}–OH versus intramolecular attack by the amide nitrogen. If the iminium is far away from Fe^{III}–OH, cyclisation is more likely to occur.

Table 5.1 Comparison of activities of $P450_{BM3}$ variants for the reaction of lidocaine to cyclisation and dealkylation products

Entry	Mutant	Dealkylation	Cyclisation	Conversion
1	WT	0%	0%	0%
2	F87A	70%	30%	8%
3	KSK19	75%	25%	27%
4	KSK19/IA	3%	96%	99%
5	RP/FV/EV	11%	84%	92%
6	RP/FV/EV/F81W	96%	0%	93%
7	RT2/A330W	0%	100%	100%
8	RT2AP/A184I	31%	69%	49%

Unlike chemo-selective reactions, this reaction is mutant-dependent and relies on the structural relationship between the substrate and mutant enzyme. Therefore, it is not ideal to apply the same mutant (best for lidocaine) towards a variety of derivatives. Protein engineering strategies are powerful tools to improve the efficiency of certain reactions. However, good starting points are generally necessary and it would be time consuming to

modify the protein active site from the same template as binding sites may vary between even closely related molecules. A library of ~100 $P450_{BM3}$ variants based on four 'generic accelerators' were collected and used for the initial screening towards lidocaine derivatives, which were prepared by Mr. Jack O'Hanlon, as shown in Figure 5.4.

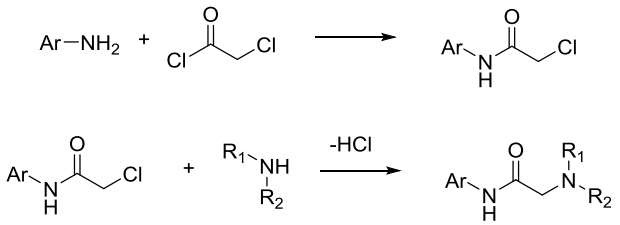
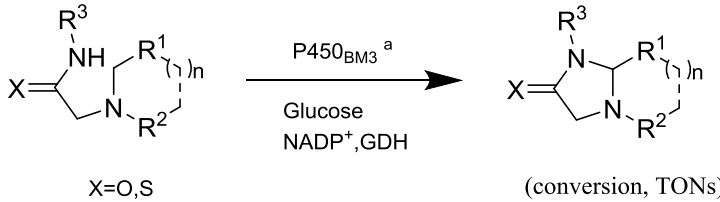


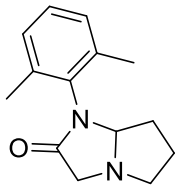
Figure 5.4 Synthesis of lidocaine derivatives (2-aminoacetamides)

5.1.2.1 Cyclic amino derivatives

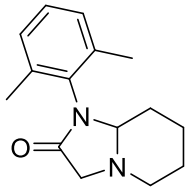
Selected synthetic 2-aminoacetamides bearing cyclic amino moieties were screened for *in vitro* oxidation by the enzyme library (Figure 5.5) using glucose dehydrogenase/glucose to regenerate the NADPH cofactor. At the end of each reaction, the organic-soluble extract was analysed by gas chromatography. Reactions that showed high substrate conversion and product selectivity were scaled up (50 – 100 mg) for product purification and characterisation. 5,5- and 5,6-bicyclic imidazolidin-4-ones were formed from lidocaine-like 2-aminoacetamides with pyrrolidinyl (Figure 5.5, entry **5.01**, Table S5.1 in Appendix) and piperidinyl (entry **5.02**, Table S5.2 in Appendix) groups, with >90% conversions and isolated yields (all characterisation data are in the Appendix 6). The 2-aminoacetamides leading to **5.01** and **5.02** are challenging substrates. Only the RP/FV/EV (for **5.01**) and RT2/AP/A184I (for **5.02**) mutants showed >90% conversion (total turnover number, TTN ~ 2000) and isolated yields for reactions *in vitro*, although more mutants had TTN > 300 (0.33 mol % catalyst loading). The reactions were also readily carried out in whole-cells in shake flasks



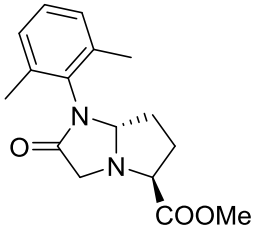
X=O,S (conversion, TONs)



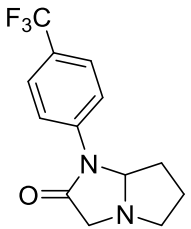
5.01, 99%[b], RP/FV/EV[c] 1976[d], 97%[e] yield



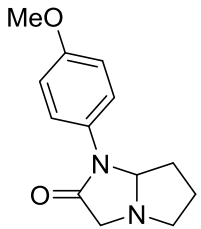
5.02, 94%, RT2AP/A184I, 1888, 92% yield



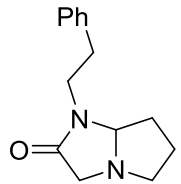
5.03, 100%, RT2/F81W, 1000, 98% yield



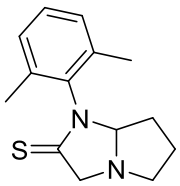
5.04, 84%, KSK19/AM/E267F, 420, 67% yield



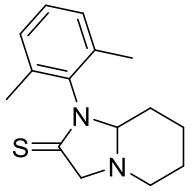
5.05, 95%, KSK19/QP/FV, 473, 88% yield



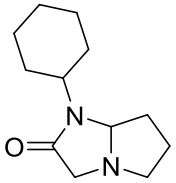
5.06, 79%, KSK19/AM/E267F, 396, 70% yield



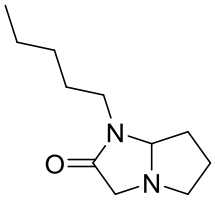
5.07, 95%, RPFVEV 1898, 92% yield



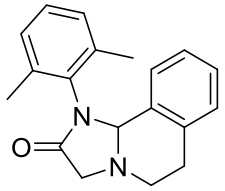
5.08, 84%, KSK19/QP/FV, 837, 66% yield



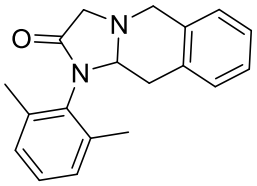
5.09, 93%, RP/FV/EV/V78F, 467, 66% yield



5.10, 72%, RP/FV/EV/L188Q, 358, 55% yield



5.11, 97%, RP/H171L, 483, 81% yield
Benzylic Position



5.12, 94%, RP/IA/EV, 468, 70% yield
Non-benzylic Position

a The data are given as: % substrate conversion, the $P450_{BM3}$ mutant, total turnover number (TTN), isolated yield for reactions *in vitro*. TTNs are given as the ratio of the concentration of substrate converted to the enzyme concentration; where conversion exceeds *ca*. 95%, the enzyme is capable of further substrate conversion and the actual TTN is higher than the value given.

Figure 5.5 $P450_{BM3}$-catalysed Intramolecular C−H amination to Imidazolidin-4-ones

where substrates were added in 2 mM aliquots up to 10 mM total concentration. (Figure 5.6) The results suggest that 2-aminoacetamides and imidazolin-4-ones readily cross the *E.coli* cell wall. Higher product concentrations are likely to be feasible at higher cell densities and more efficient mass transport, e.g. in a bioreactor vessel.

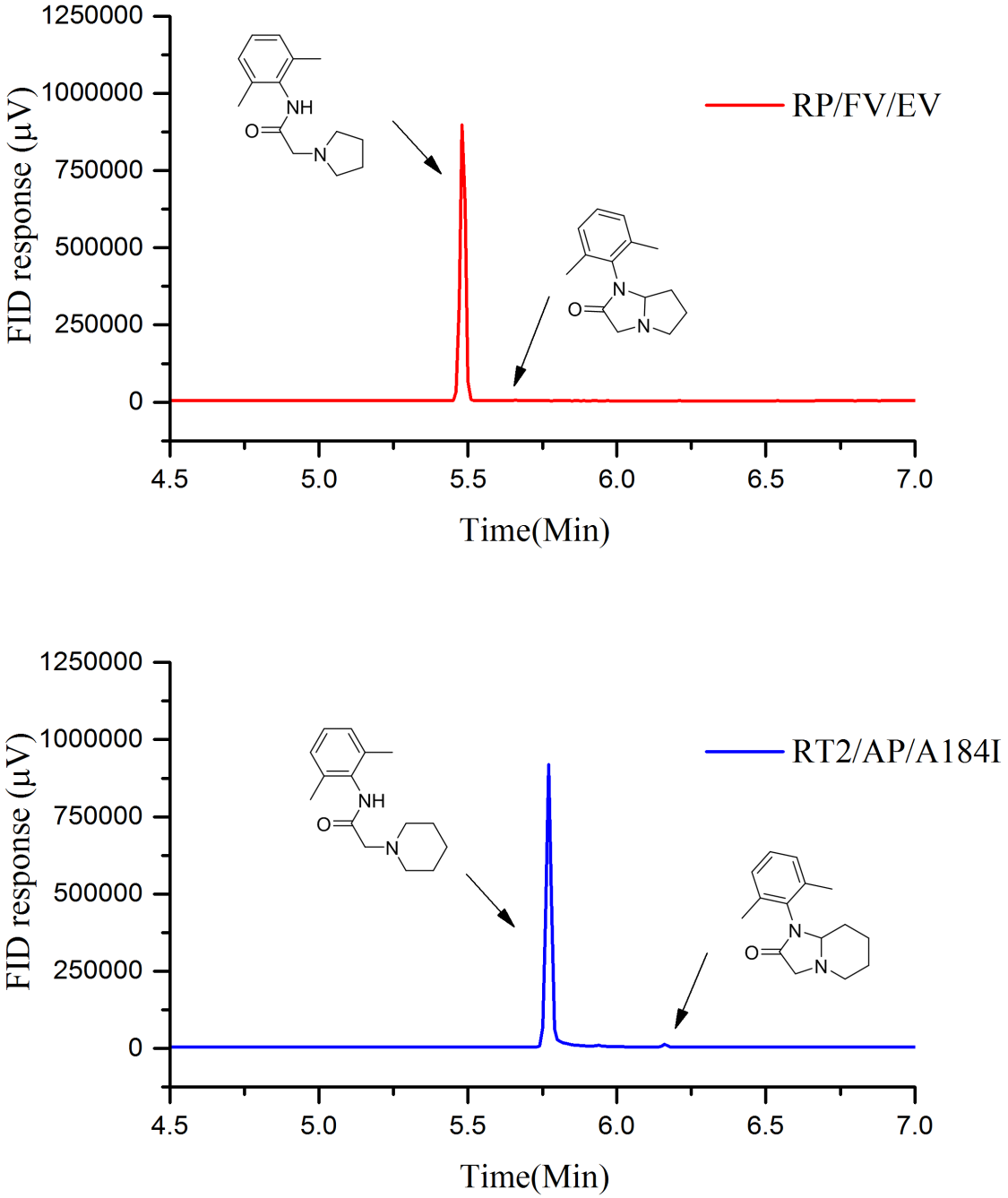


Figure 5.6 *In vivo* biotransformation of 1-(2,6-dimethylphenyl)tetrahydro-*1H*-pyrrolo[1,2-a]imidazol-2(*3H*)-one (**5.01**) and 1-(2,6-dimethylphenyl)hexahydroimidazo[1,2-a]pyridin-2(*3H*)-one (**5.02**) at 10 mM concentration

The functional group tolerance of $P450_{BM3}$ mutant enzymes at both the amine and amide ends of 2-aminoacetamides was then explored. Introduction of substituents to the cyclic amine had no adverse effect on turnover activity or chemoselectivity for C–H amination. The methyl-L-prolinate derivative (entry **5.03**) was converted into >97% of the corresponding cyclisation product by, for example, the RT2/F81W mutant (Table S5.3 in Appendix). Interestingly, the presence of a substituent on the pyrrolidine ring increased the number of mutants showing high conversion and TTN. The 1D NOE spectra (Figure 5.6) showed that compound **5.03** was generated solely as the (5S, 7R)-diastereoisomer by all mutants within the library.

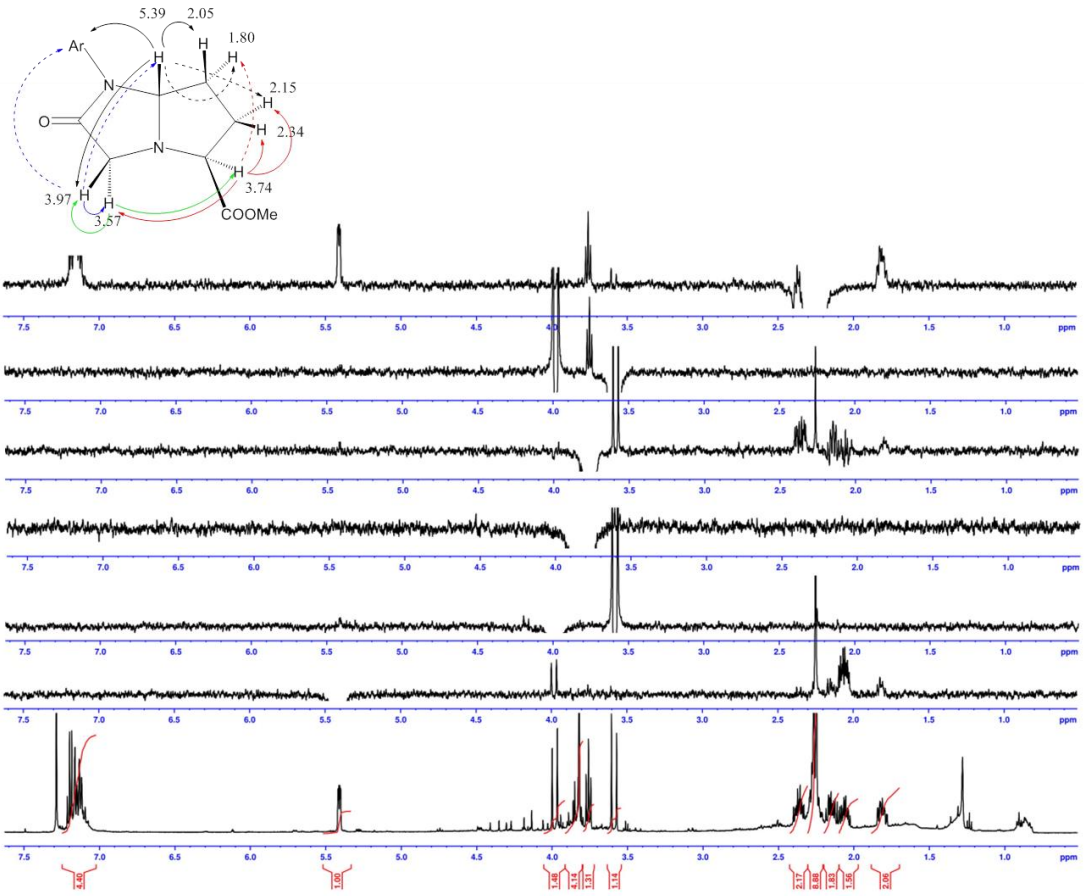


Figure 5.7 1D NOE spectra in $CDCl_3$ of methyl (5S,7R)-1-(2,6-dimethylphenyl)-2-oxohexahydro-1*H*-pyrrolo[1,2-a]imidazole-5-carboxylate (**5.03**)

Theoretical calculations were then performed by Prof. Dr. Jeremy Roberson at Oxford. For each diastereoisomer of compound **5.03**, conformers were obtained in Spartan following a Monte Carlo search (MMFF). Each conformer was then submitted for an equilibrium geometry calculation (B3LYP/6-31G**) with a water solvation model, and a Boltzmann-weighted G^0 was obtained from a thermodynamics calculation.[418] The obtained ΔG^0_{298K} = 2.46 kcal mol^{-1} corresponds to a ~63:1 ratio of *exo*- to *endo*- diasteroisomer at equilibrium, suggesting rapid equilibration of the *N,N*-acetal center likely occurred during the reaction and isolation procedure. This is consistent with the observation that the unsubstituted compounds **5.01** and **5.02**, as well as the products from the other aminoacetamides in Figure 5.4, were obtained in racemic forms.

Electron withdrawing (entry **5.04**) and electron donating substituents (entry **5.05**) were introduced at the 4-position in place of the 2,6-dimethyl substituents on the phenyl group of 2-pyrrolidinylacetamides. The most selective mutants KSK19/A82M/E267F and KSK19/QP/FV provided the cyclised amination products with 84% and 95% conversion and 67% and 88% isolated yield of **5.04** and **5.05**, respectively. Notably, selectivity for the cyclisation product **5.04** was reduced for some variants (as low as 33% for variants giving >60% conversion, Table S5.4) but remained high, often > 85%, for **5.05**. Reduced nucleophilicity of the amide nitrogen in the trifluoromethyl substrate presumably allows other pathways to compete more effectively with cyclisation (Table S5.5). Products from these other pathways were not isolated.

The mutant library also oxidised the 2-phenylethylamide derivative with up to 79% conversion and 70% isolated yield for the cyclised product (entry **5.06**). Selectivity for the cyclisation product was higher than for **5.04** and as high as that for the lidocaine reaction (Fig. S6, Table S5.6), supporting the importance of the nucleophilicity of the amide nitrogen.

These results demonstrate that substitutions on the aromatic ring and different amine groups are tolerated by the enzyme.

Entries **5.07** and **5.08** show that thioamides are also successful substrates for this reaction and the *N*,*N*-acetals were formed with high conversion (Table S5.7 & S5.8).

The conversion of aliphatic and alicyclic amides to cyclisation products **5.09** and **5.10** illustrates that an *N*-aryl substituent at the amide end is not required for 2-aminoacetamide binding or for achieving a binding orientation that facilitates α-amine functionalisation and cyclisation. The RP/FV/EV/V78F mutant converted 93% of the cylcohexyl amine derivative (entry **5.09**, Table S5.09) and gave 66% isolated yield of the cyclisation product, whereas the RP/FV/EV/L188Q mutant showed 72% conversion of the pentyl substituted 2-aminoacetamide (entry **5.10**, Table S5.10) to the corresponding *N*,*N*-acetal in 55% isolated yield.

Cyclisation of the tetrahydroisoquinoline (THIQ) derivative (entries **5.11** and **5.12,** Table S5.11 in Appendix) to a tricyclic compound further highlights the tolerance of the $P450_{BM3}$ mutants for variations in the amine moiety and the ability to control regioselectivity of C–H amination (Figure 5.8).

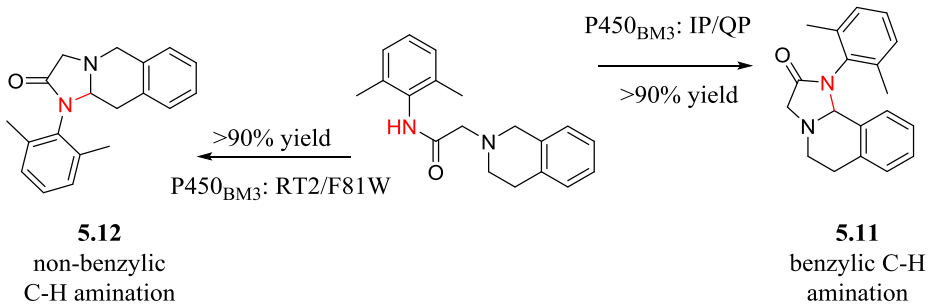


Figure 5.8 Regiodivergent intramolecular C–H amination of THIQ derivative

The RP/H171L mutant showed 91% conversion with 86% regioselectivity for C–N bond formation at the benzylic position (entry **5.11**) while RP/IA/EV showed 94% conversion with 77% selectivity for the non-benzylic position (entry **5.12**). Mutants demonstrating near-

complete shift between the two regioisomers were found from *in vitro* screening of the library. The IP/QP mutant showed 95% selectivity for **5.11** while the RP/F81W gave 94% **5.12** although the conversions under screening conditions were low (~25%). Encouragingly, however, the conversion in both reactions was increased to >95% *in vivo* without loss of chemoselectivity, leading to isolated yields of >90% for each product. If a hydrophobic organic molecule can cross the *E. coli* cell wall, large increases in total turnover for *in vivo* reactions over *in vitro* conditions are possible because the cytoplasmic enzyme is protected from high concentrations of organics and diffusion of products into the external medium reduces product inhibition. The results suggest that 2-aminoacetamides readily cross the *E. coli* cell wall.

This is the first report of selective direct C–H heterofunctionalisation in simple THIQ derivatives. In addition to the many bond-forming processes at the 1-position, a small number of non-selective 1-/3-oxyfunctionalisations are known.[419-422] Selective C–C bond-forming processes at the non-benzylic 3-position have been reported recently *via* two strategies: (1) for carbon substituents capable of supporting a negative charge (malonyl, cyano, nitroalkyl), enrichment of the fraction of 3-derivative can be achieved by thermal equilibration; (2) the 3-C–H bond in *N*-(benzoxazol-2-yl)THIQ is more accessible to bulky Ir(I) catalysts that achieve direct alkylation at that position with terminal alkenes.[423] Both processes require high temperatures and offer limited scope; in contrast, the selective formation of compound **5.12** occurs under benign reaction conditions to provide a stable latent iminium ion for subsequent C–C bond-forming reactions.

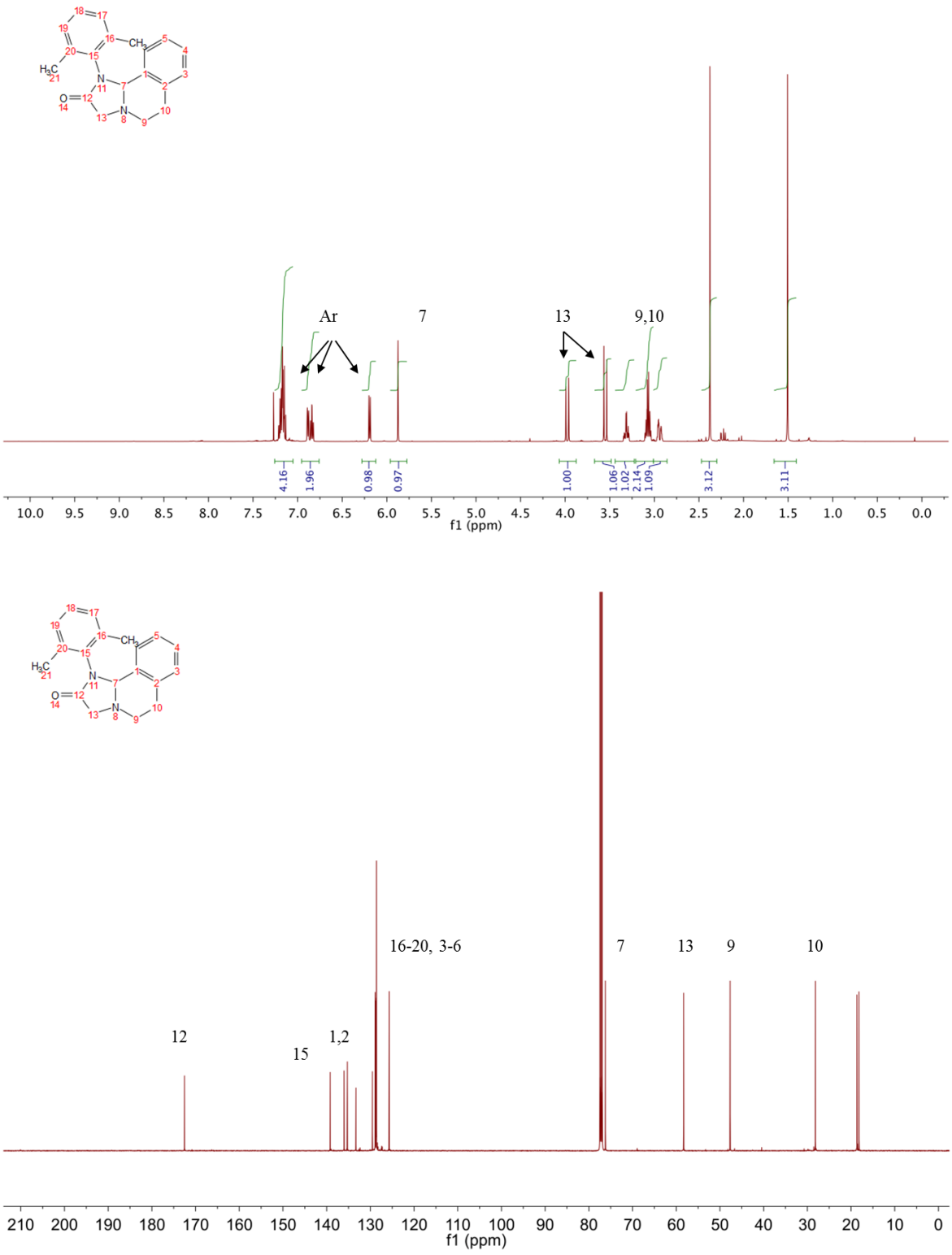


Figure 5.9 The ^{1}H and ^{13}C NMR spectrum in $CDCl_3$ of 1-(2,6-dimethylphenyl)-1,5,6,10b-tetrahydroimidazo[2,1-a]isoquinolin-2(3*H*)-one (**5.11**)

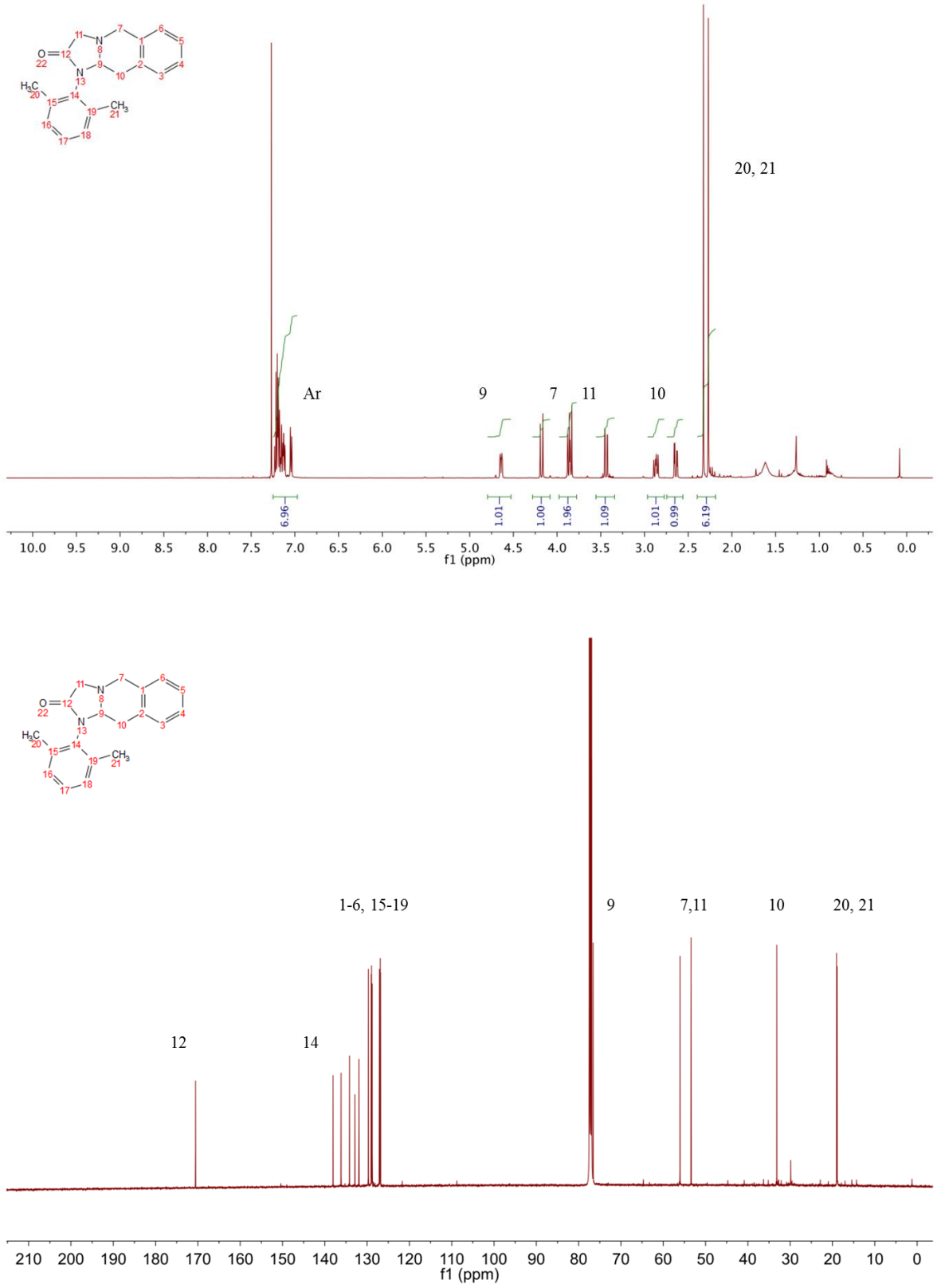


Figure 5.10 The ^{1}H and ^{13}C NMR spectrum in $CDCl_3$ of 1-(2,6-dimethylphenyl)-1,5,10,10a-tetrahydroimidazo[1,2-b]isoquinolin-2(3*H*)-one (**5.12**)

The activity profile of the mutant library (Table S5.1 to S5.11 in Appendix) suggests that the active-site residues F81, A82, F87, I263, E267, and A330 have important effects on 2-aminoacetamide cyclisation *via* C−H amination. However, the base mutants with their enhanced activity for unnatural substrate oxidation are required because the single-site mutants such as F87A, among others, have very low activity. The data showed that the mutants RT2/F81W, RT2/A330P/A184I, RP/F87V/E267V, RP/I263A/E267V, and KSK19/A82M/E267F (Figure 5.5), would have established the possibility of intramolecular C−H amination for all the tested 2-aminoacetamides.

5.1.2.2 Acyclic amino derivatives

Alkyl amine derivatives were also tested with the $P450_{BM3}$ mutant enzyme system for intramolecular C–H amination. 2-(Dimethylamino)-*N*-(2,6-dimethylphenyl)acetamide (compound **5.13**) and *N*-(2,6-dimethylphenyl)-2-(ethyl(methyl)amino)acetamide (compound **5.15**) share identical molecular skeleton with lidocaine, were therefore subjected to an initial screen using 48 $P450_{BM3}$ variants. Substrate conversion and product distribution are shown in Table S5.12 & Table S5.13 in Appendix.

Compound **5.13** was tested for the evaluation of primary alkyl group being activated by P450 enzymes. (Table S5.12-1 in Appendix) It was found that RP/FV/EV converted 65% of **5.13** into ~90% of cyclisation product (Figure 5.11). RP/FV/EV/L437VL showed 100% selectivity to form 3-(2,6-dimethylphenyl)-1-methylimidazolidin-4-one (**5.14**) although substrate conversion was low (~ 15%).

As for compound **5.15**, altered regioselectivity was observed towards primary and secondary alkyl ends at 50% and 74%, forming the cyclisation products **5.16** and **5.17** in 38% and 18% isolated yield, respectively. (Table S5.13-1 in Appendix)

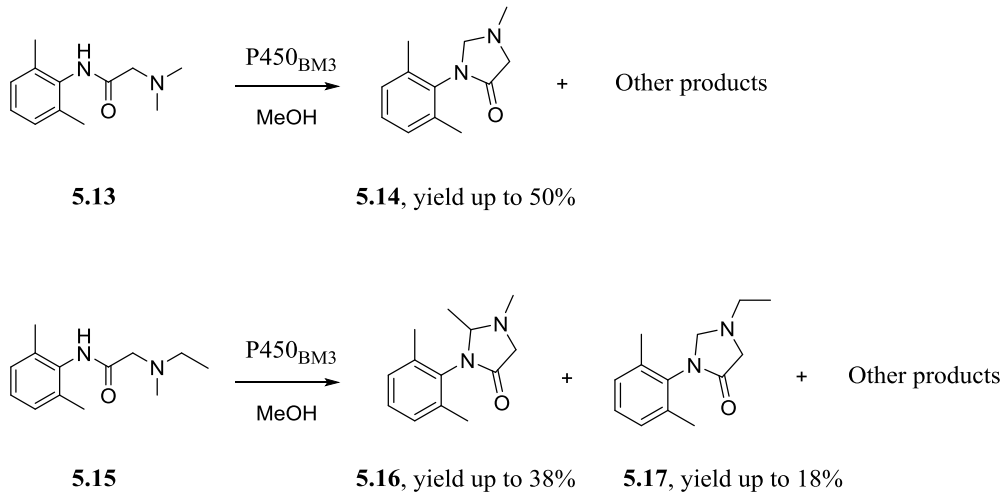


Figure 5.11 Intramolecular cyclisation of lidocaine alkyl derivatives

The monocyclic imidazolidin-4-ones were formed in sufficient quantities that enable the compounds to be fully characterised by NMR and MS. The cyclisation occurred at both primary and secondary alkyl ends with tunable regioselectivity, making the P450-catalysed intramolecular C–H amination a general synthetic methodology for generating monocyclic, bicyclic, tricyclic imidazolidin-4-ones.

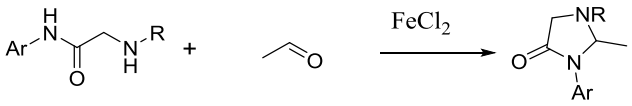


Figure 5.12 Chemical synthesis of monocyclic imidazolidin-4-ones

It is worth pointing out that monocyclic imidazolidin-4-ones are also accessible by chemical methods, through an iron(II) catalysed condensation of aldehyde between amide nitrogen and secondary amines (Figure 5.12). However, this method is not suitable in making bicyclic and tricyclic imidazolidin-4-ones, emphasising the importance of biocatalytic methodology mentioned earlier in this chapter.

Surprisingly, 3-(2,6-dimethylphenyl)-1-methylimidazolidin-4-one (**5.14**) was found in the biotransformation of *N*-(2,6-dimethylphenyl)-2-(ethyl(methyl)amino)acetamide (**5.15**) where

substrate was added in methanol. By changing the co-solvent to DMSO, this compound was not observed, indicating the necessity of methanol.

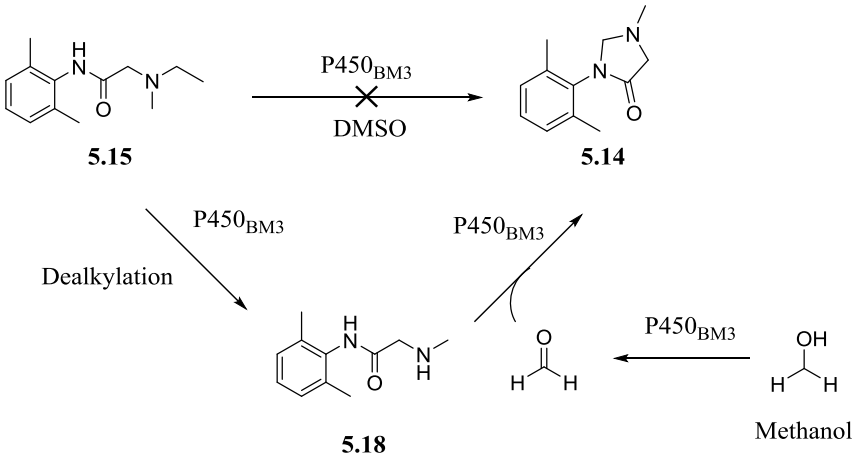


Figure 5.13 Proposed synthetic pathway of unexpected product formation

As shown in Figure 5.13, the dimethyl lidocaine derivative underwent the dealkylation process to give its demethylation product (compound **5.18**) followed by the installation of formaldehyde in which $P450_{BM3}$ enzyme acted as an iron(II) catalyst. Since the usage of DMSO as co-solvent was not able to produce the cyclisation product, it was clear that formaldehyde was generated from methanol oxidation catalysed by $P450_{BM3}$ enzymes. With the amine substrate sitting in the enzyme active site, methanol oxidation became possible within such decoy system otherwise the enzyme active site volume is far too large for methanol being oxidised.

Similar observations were also made from 2-(dimethylamino)-*N*-(2,6-dimethylphenyl) acetamide (**5.13**) oxidation. 3-(2,6-Dimethylphenyl)-1,2-dimethylimidazolidin-4-one (**5.16**) was formed with ethanol as the co-solvent (Table S5.12-2 in Appendix).

To investigate methanol and ethanol oxidation in such decoy systems, two lidocaine derivatives incorporating secondary amines (methyl and isopropyl) were subjected to P450 reactions. The reactions were carried out at 0.5 mL volume in phosphate buffer. Substrate

stocks were made in both methanol and ethanol and the reactions were performed in parallel for comparison.

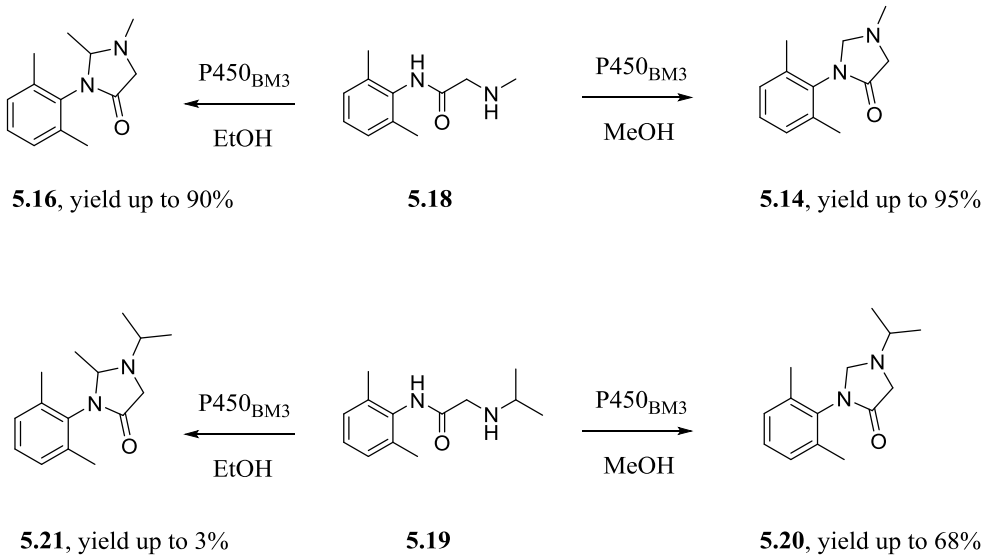


Figure 5.14 P450 catalysed aldehyde insertion

A panel of 48 $P450_{BM3}$ variants was screened towards both compound **5.18** and **5.19** (Figure 5.14, Table S5.14 & Table S5.15 in Appendix). In terms of the oxidation of **5.18**, the most active methanol oxidising mutant was found to be RP/FV/EV/L437VL which gave 95% aldehyde insertion product while ethanol oxidation was more-or-less as efficient as methanol, with RT2 mutant being identified to give 90% yield of acetaldehyde insertion product. Compound **5.19** showed a different result, the proportion of compound **5.20** (68%) was far larger than that of compound **5.21** formation (3%), suggesting methanol oxidation was far more efficient than ethanol oxidation.

5.1.3 Summary

In summary, by screening variants of $P450_{BM3}$ for activity of 2-aminoacetamides, we have developed a scalable enzymatic C–H amination process for the straightforward synthesis of bicyclic imidazolidin-4-ones under mild conditions for both 5,5- and 5,6- fused bicyclic systems. Enzymes offer advantages over transition metal catalysts because they operate under mild conditions in aqueous solvent, are highly active, essentially inexhaustible, and can be evolved to be highly selective for the desired product. The increased conversion for the *in vivo* C–H amination over reactions *in vitro* adds to the versatility of the system. The enzymes accept aromatic and aliphatic amides as well as substituents on the amine ring that also, by thermodynamic stereochemical relay, establish defined stereochemistry at the newly formed ring junction. Cyclisation of the tetrahydroisoquinoline derivative to a tricyclic compound further highlights the tolerance of $P450_{BM3}$ mutants for the variations in the amine moiety and the possibility of controlled site-selective C–H amination. Imidazolidin-4-ones are latent iminium ions, being amendable to Lewis acid-mediated ring-opening with subsequent nucleophilic attack offering new routes of α-functionalisation of tertiary amines. Ongoing engineering studies are directed at identifying $P450_{BM3}$ variants that will hydroxylate the imidazolidin-4-ones products in a second step to introduce additional functional diversity and, potentially, biological activity.

Acyclic amino derivatives have been found to exhibit complex reactivity. For example, compound **5.15** can be converted to its dealkylation products (compound **5.18** and MEGX) where aldehydes are able to be inserted to give cyclisation products (pathway A). Alternatively, the iminium ion can be produced which is then trapped by intramolecular nucleophilic attack to give the desired products (pathway B), as shown in Figure 5.15.

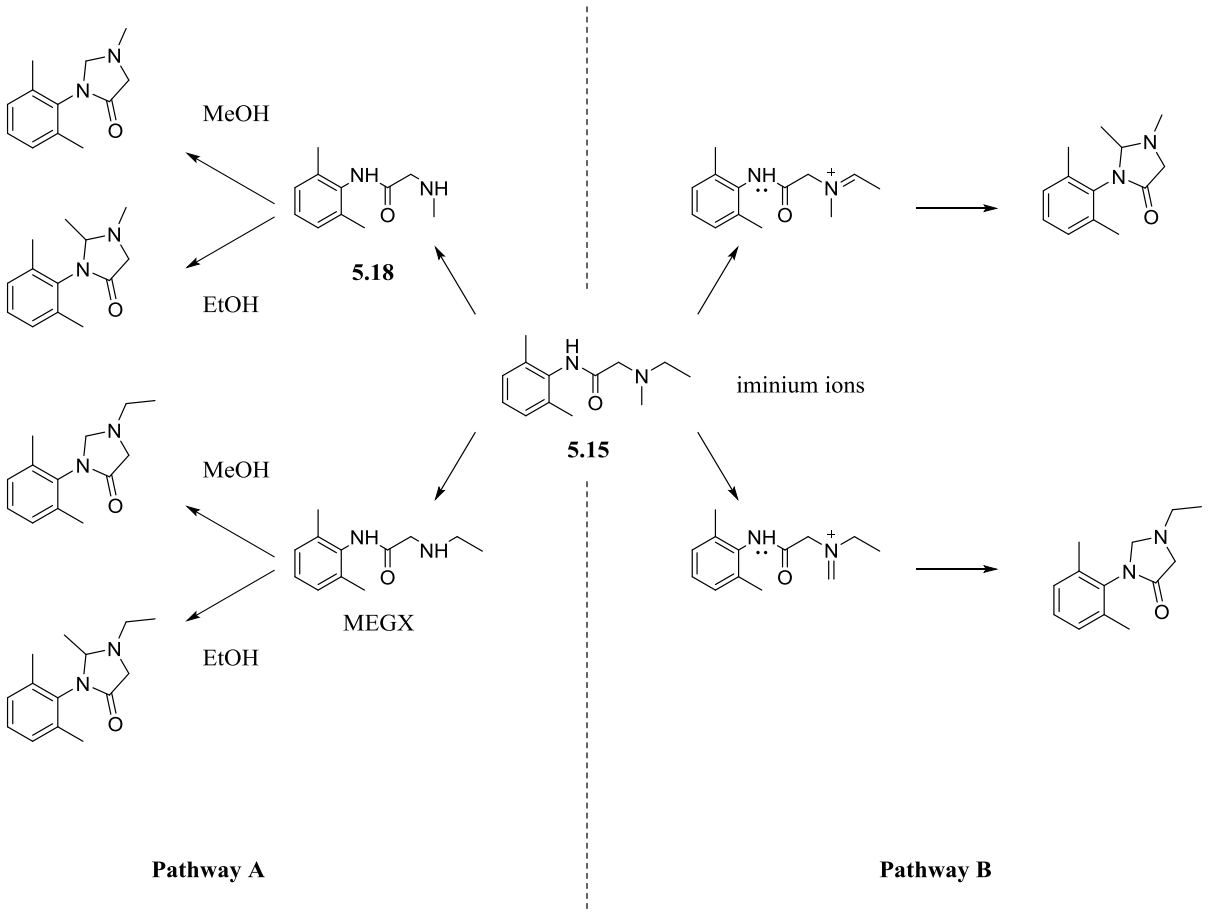


Figure 5.15 Imidazolin-4-ones formation from **5.15** oxidation catalysed by $P450_{BM3}$ mutant enzymes

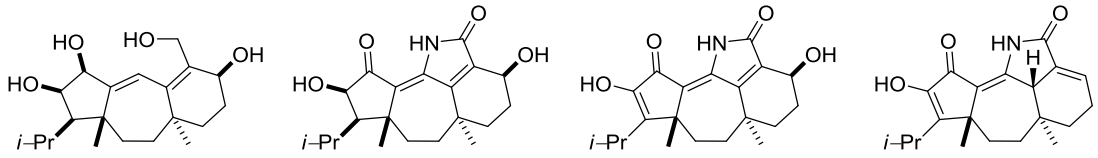
5.2 Late-stage C–H functionalisation by $P450_{BM3}$

5.2.1 Introduction

Late-stage functionalisation of C–H bonds holds the potential to offer more direct routes of chemical synthesis and opportunities to access new molecules that cannot be made easily by conventional synthetic approaches. The introduction of a variety of important functional groups in the last steps of synthesis has become a new strategy in natural product synthesis and drug discovery.

Cytochrome P450 enzymes are capable of activating virtually any C–H bonds in complex organic molecules under mild conditions with a high degree of functional group tolerance and tunable regio- and enantio-selectivity. The application of $P450_{BM3}$ mutant enzyme system

could be a powerful catalytic method to create synthetic handles for structural modification and elaboration. Radianspene system was first subjected to this proof-of-concept research.



Radianspene C | Radianspene J | Radianspene K | Radianspene L

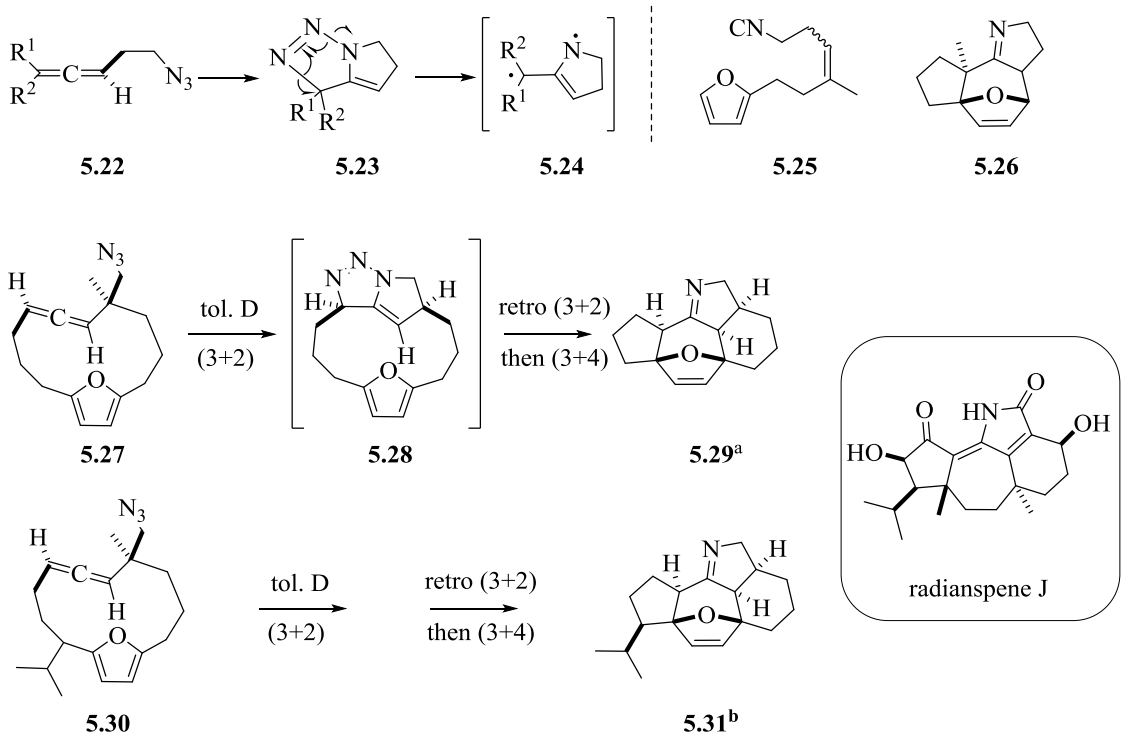
Figure 5.16 Examples of radianspenes isolated by Shen and co-workers

The radianspenes are a class of 13 structurally diverse guanacastepene-type diterpenoids first isolated by Shen and co-workers from the endophytic fungus *Coprinus radians*, which lives within Amanita mushrooms from China. They share the same skeleton consisting of a linear 5-7-6 tricyclic core, with one long face exhibiting a range of oxidation sites, and the opposite face being largely hydrophobic. An evaluation of the antitumour activity of the radianspenes was carried out by Shen and co-workers and radianspene C was found to possess inhibitory activity against the MDA-MB-435 (melanoma) cell line.

No synthetic approaches to the radianspenes have been reported to date, and therefore we wanted to explore the potential of combining cascade cycloaddition reactions and cytochrome P450 reactions in the complex setting of radianspene J total synthesis.

5.2.2 Results and discussion

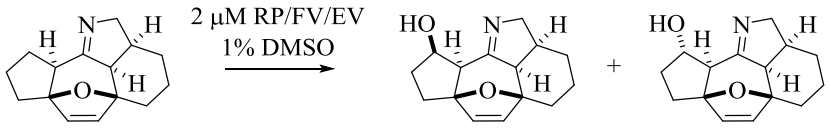
The synthesis of two radianspene cores were carried out by Dr Oleksandr Zhurakovskyi and Mr Samuel Ellis in Prof. Jeremy Robertson's lab at Oxford.[424, 425] The approach involved trapping the azatrimethylenemethane intermediates **5.24** by intramolecular cycloaddition reactions with a tethered furan that produced a mixture of isocyanides **5.25** and cycloadducts **5.26**. The transannular system **5.27** led to the exclusive formation of cycloadducts **5.29** and **5.31**, which were then subjected to oxidation by mutant P450 enzymes.



a synthesised by Dr Oleksandr Zhurakovskyi from Prof. Jeremy Robertson's lab at Oxford
b synthesised by Mr Samuel Ellis from Prof. Jeremy Robertson's lab at Oxford

Figure 5.17 Synthetic route to radianspene core model compounds

In a separate series of experiments, 16 P450 mutant enzymes were screened towards compound **5.29** in small scale reactions (0.1 mg of the substrate) in phosphate buffer (0.5 mL, 200 mM, pH 7.5) using a NADPH regeneration system. The substrate (2 mM) was added from a 100 mM solution in DMSO, followed by glucose dehydrogenase (2 Units/mL), and glucose to final concentration of 100 mM. $NADP^+$ monosodium salt was added to initiate the reactions which were left shaken at 200 rpm at room temperature overnight (16 h). The reaction mixtures were extracted with ethyl acetate (200 μL) and analysed by gas chromatography.



5.29 **5.32**

Figure 5.18 The oxidation of radianspene core compound (5.29) by RP/FV/EV $P450_{BM3}$ variant

The $P450_{BM3}$ RP/FV/EV mutant was then chosen for a preparative scale reaction (14 mg of **5.29**). The monooxygenated product **5.32** was isolated in 15% yield as a 1:0.9 mixture of the "up" and "down" diastereomers (Figure 5.17). The relative configuration of the newly oxygenated stereocentre was established by the analysis of the $^{1}H–^{13}C$ HMBC spectrum and relating it to the predicted dihedral angles in the two isomers. (Figure 5.19) This has enabled C–H activation of a site that would be unreactive to most common chemical methods, and provides a handle for further transformations, such as oxidation of the adjacent C-13 site. This is highly promising as it would enable a large degree of diversification from one intermediate.

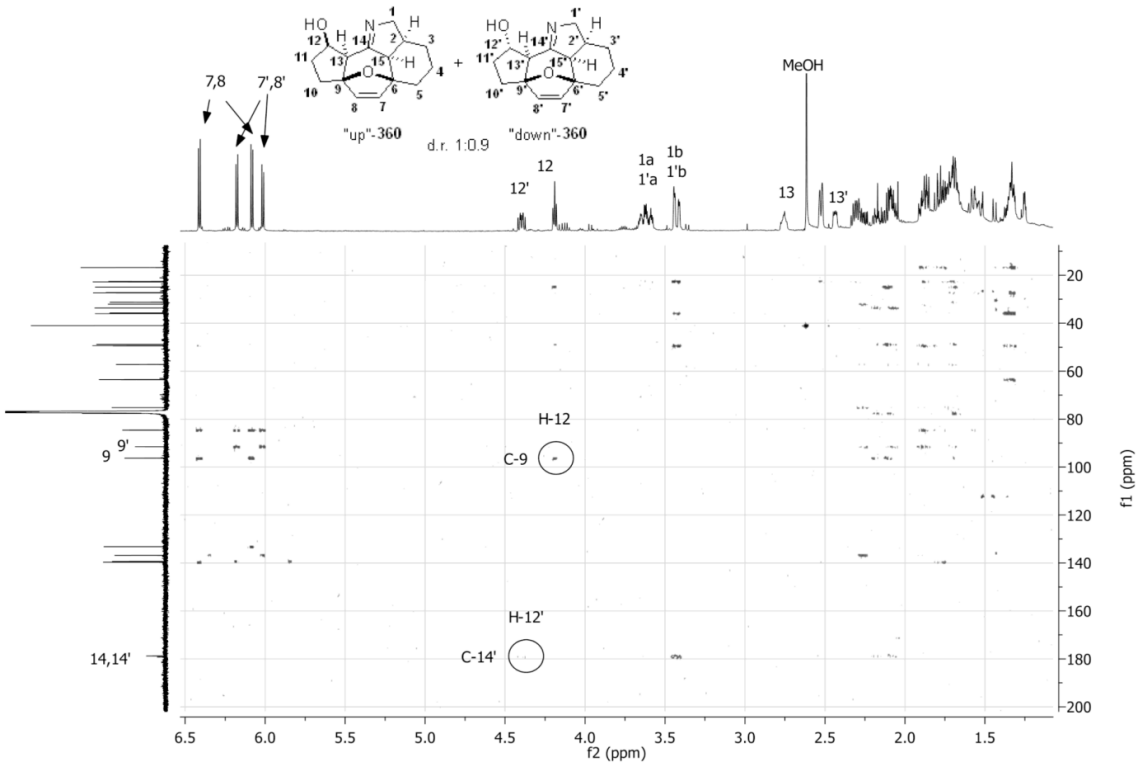


Figure 5.19 $^{1}H–^{13}C$ HMBC spectrum of compound **5.32**

Working in collaboration with Samuel Ellis, the isopropylated radianspene core (compound **5.31**) was also subjected to an initial screening of 16 $P450_{BM3}$ mutants to identify promising mutants. Several interesting observations were made. In some cases, oxidation appears to be selective for one diastereoisomer (for example KSK19/AI), and in all cases where this was observed, the diastereoisomer with the shorter retention time reacted preferentially (Figure 5.20).

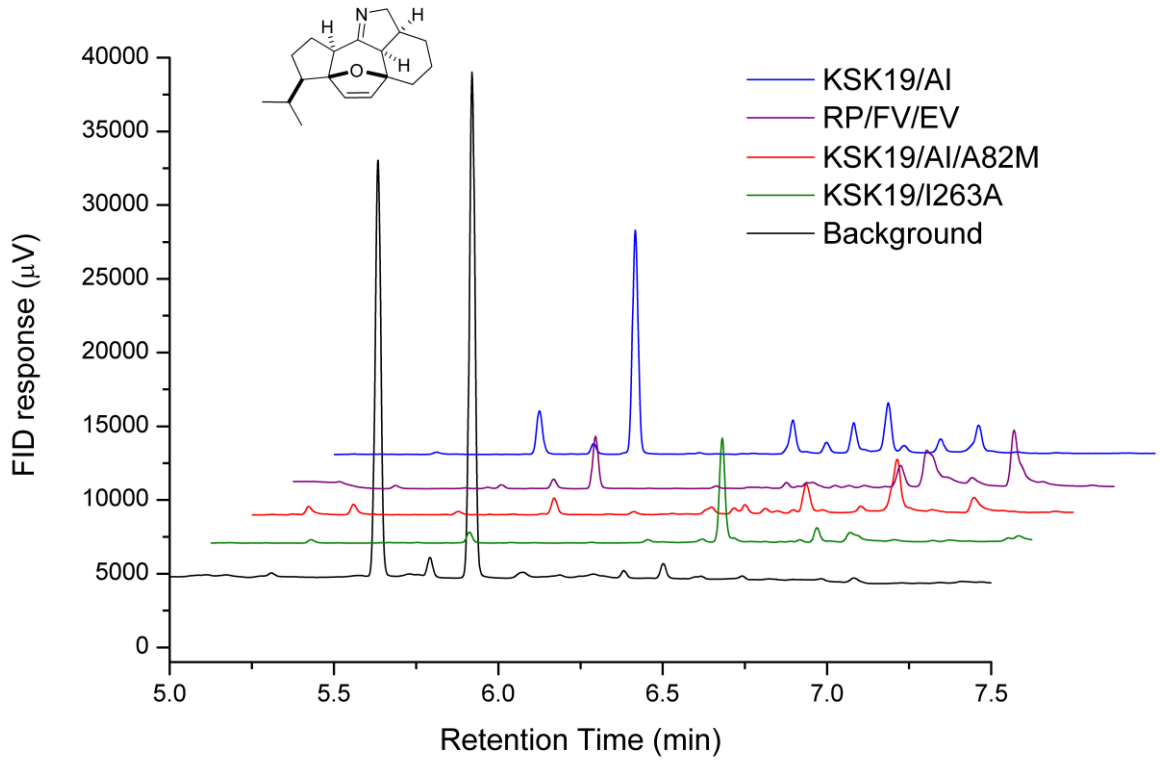


Figure 5.20 GC analysis of the initial screening towards isopropylated radianspene core model **(5.31)**

Most interesting was that mutant KSK19/I263A appears to have converted both diastereoisomers into a single major product based on GC analysis, and TLC indicates two very close spots, potentially two diastereomers. In addition, mass spectrometry (ESI+ 288.2) is consistent with mono-hydroxylation of the isopropylated core model. However, the

products were not fully characterised due to the lack of starting material since the core model was furnished in 2% yield over 17 steps (79% per step). With optimisation of the devised synthetic route, this core model would be an ideal candidate for further exploration, and a preparative scale reaction would allow the identification of the hydroxylated site.

5.2.3 Summary

The preliminary results of the biocatalysed oxygenation of two radianspene core model compounds are encouraging and several promising mutants were identified. However, much more work is to be done before this reaction becomes synthetically useful. The development of $P450_{BM3}$ mutant enzyme systems and the identification of variants capable of fully converting radianspene cores will enable sites of oxidation to be characterised. If synthetically useful oxidation can be achieved, methods to elaborate the structure from these sites will be investigated with the aim of developing the first total synthesis of radianspene J.

5.3 The application of *N*-Dealkylation by $P450_{BM3}$

5.3.1 Introduction

N-Dealkylation has been widely recognised as a major metabolism pathway for many amine drugs. For example, lidocaine, amitriptyline, dextromethorphan, etc. Although there is an argument on the mechanism of these metabolites being formed (either through a hydrogen atom transfer to form a carbon radical which undergoes recombination to form the carbinolamine which collapses non-enzymatically to generate an amine plus a carbonyl compound or through a single electron transfer followed by deprotonation, as shown in Figure 5.21), there is no controversy about what products are formed. The application of P450-catalysed *N*-dealkylation was further investigated.

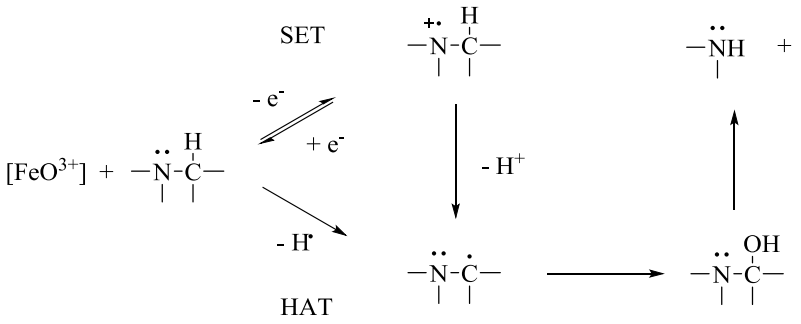
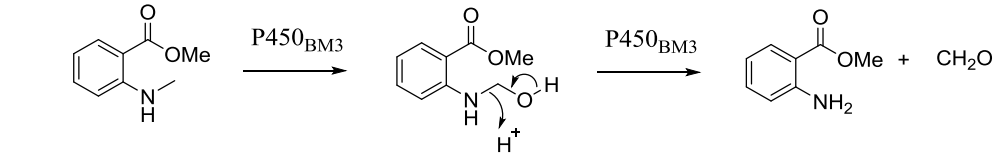


Figure 5.21 Mechanisms hypothesised for oxidative *N*-dealkylation catalysed by P450s

Methyl anthranilate (MA), also known as methyl 2-aminobenzoate, or carbomethoxyaniline, is an ester of anthranilic acid with a fruity grape smell and acts as a bird repellent. It is food-grade and can be used to protect corn, sunflowers, rice, fruit, and golf courses. It is a clear to pale yellow liquid with a light blue fluorescence. At full concentration, it has a fruity grape smell, while it has a sweet, fruity, concord grape-like smell with a musty and berry nuance at 25 ppm.

Methyl anthranilate naturally occurs in the concord grapes and various natural essential oils and has an extensive usage in modern perfumery. It is also an important precursor to a number of Schiff bases which are used in several fragrance formulations, e.g. aurantiol was produced by combining methyl anthranilate and hydroxycitronellol. Methyl anthranilate is synthesised through esterification of anthranilic acid with methanol catalysed by a variety of homogeneous and heterogeneous acid catalysts in the conventional industrial processes. Dimethyl anthranilate (**5.33**, market price $180/kg) was therefore selected as the starting material for the synthesis of methyl anthranilate (**5.34**, market price $1000/kg) through *N*-dealkylation catalysed by $P450_{BM3}$ mutant enzymes, as shown in Figure 5.22.



Dimethyl anthranilate, **5.33**
$180 /kg

Methyl anthranilate, **5.34**
$1000 /kg

Figure 5.22 Proposed synthetic route to Methyl anthranilate

5.3.2 Results and discussion

The initial screen was conducted with 29 $P450_{BM3}$ mutant enzymes in phosphate buffer (200 mM, pH 7.5). The substrate was added from a 200 mM methanol stock to a final concentration of 5 mM while the enzyme concentration was at 1 μM. The reaction mixtures were shaken at 200 rpm for 2 hours prior to extraction by ethyl acetate and analysis by gas chromatography. Table 5.2 shows the substrate conversion and product distribution of dimethyl anthranilate.

Table 5.2 Substrate conversion and product distribution of the initial screen of dimethyl anthranilate (1uM enzyme vs 5mM substrate)

No	Enzyme	Conversion[a]	Methyl anthranilate	FOP[b]	TON[c]
1	KSK19AI Y51L	72%	100%	0%	3607
2	KSK19AI F42L	65%	100%	0%	3262
3	RLYF KSK19 AI	65%	100%	0%	3258
4	KSK19AIAI	65%	100%	0%	3244
5	RP	59%	83%	17%	2959
6	RT2IP	58%	74%	26%	2898
7	KSK19AI	49%	100%	0%	2442
8	KSK19AI F87V	47%	100%	0%	2366
9	RP FV	45%	100%	0%	2275
10	RP/QP	44%	94%	6%	2185
11	RPAPAM	42%	68%	32%	2107
12	KSK19 A82M	41%	100%	0%	2038
13	RPFVEV L437LV	40%	100%	0%	2005
14	RPFVEV	28%	100%	0%	1410

15	KSK19AI V78I	28%	100%	0%	1408
16	RPAP	28%	100%	0%	1394
17	RT2	28%	100%	0%	1382
18	RT2APAM	23%	100%	0%	1147
19	RT2APA184I	23%	100%	0%	1146
20	RPFVEV L437VL	23%	100%	0%	1131
21	RP/IA/EV	22%	100%	0%	1091
22	R19	18%	100%	0%	904
23	RT2IMAP	18%	98%	2%	879
24	RT2AP A82L	17%	100%	0%	861
25	RT2APV78I	14%	100%	0%	685
26	RP EV	14%	100%	0%	677
27	RT2AP	13%	100%	0%	639
28	WT	3%	100%	0%	125
29	KSK19 I263A	1%	100%	0%	44

a The product distributions were determined from the integration of peak areas. Substrate conversion and product distribution were obtained from the average of 2 independent experiments with errors not higher than 10%
b. FOP = Further Oxidation Products
c. TON = Turnover Number

Among the series of 29 $P450_{BM3}$ mutant enzymes, 27 gave over 10% conversion of dimethyl anthranilate (DMA) except WT and KSK19/I263A. Four mutants showed over 65% conversion of DMA to a single product methyl anthranilate (MA) and were all based on KSK19/AI template, suggesting that this reaction was dominated by F87A and A328I mutations. Interestingly KSK19/AI only converted 49% of the substrate, indicating the importance of mutations at the gate of substrate entrance channel (Phe42, Arg47, Tyr51, etc.). In addition, RP, RT2IP, RP/QP, RPAPAM, and RT2AP/IM were the only five $P450_{BM3}$ variants which produced further oxidation products. This observation is consistent with the intrinsic generic accelerating nature of I401P and I401M mutations on the oxidation of non-natural substrates.

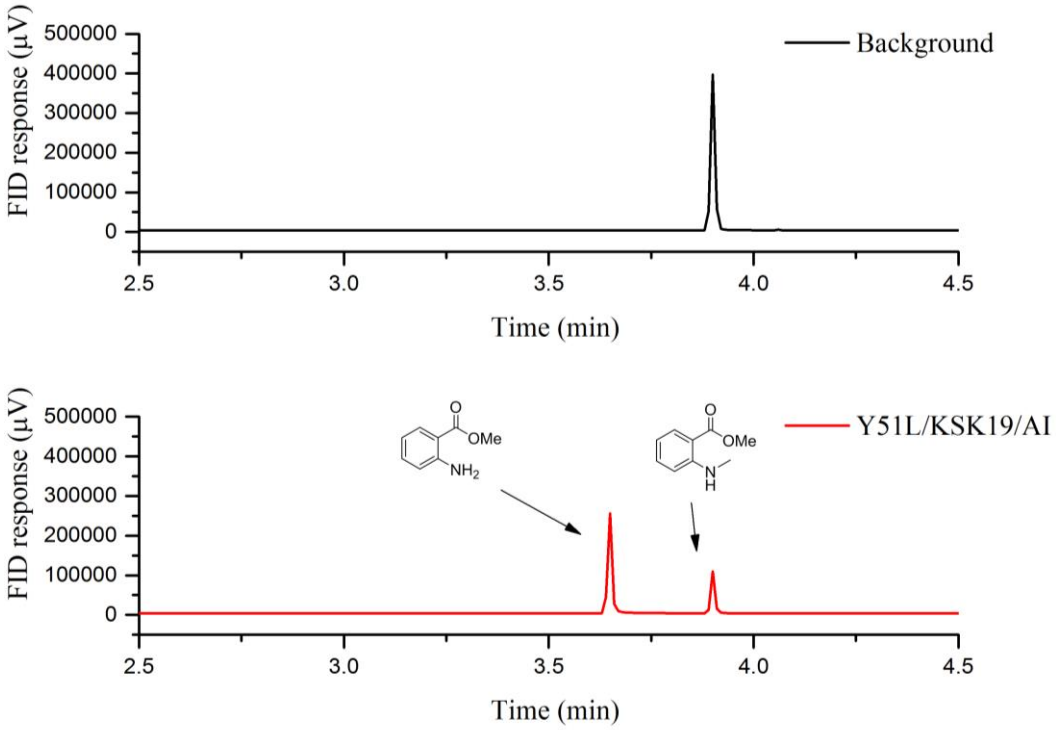


Figure 5.23 GC analysis of dimethyl anthranilate conversion by Y51L/KSK19/AI

The enzymes giving over 40% conversion of DMA substrate were then selected for further investigations in which substrate concentration were increased to 10 mM while enzyme concentration were not changed. The reaction mixtures were shaken at 200 rpm for 8 hours and pH was measured and maintained at ~7.5 by using 1 M KOH solution. Substrate conversion and turnover number data are shown in Table 5.3.

Table 5.3 Small scale reactions of DMA for 8hrs (1uM enzyme vs 10mM DMA)

No	Enzyme	Conversion	TON
1	Y51L KSK19AI	38%	3806
2	RLYF KSK19AI	34%	3411
3	KSK19AIFV	32%	3232
4	KSK19AI F42L	31%	3123
5	KSK19AI	29%	2917
6	KSK19AIAI	28%	2815
7	RT2IP	24%	2423
8	RPQP	23%	2334

9	RPFV	20%	2035
10	KSK19A82M	20%	2029
11	RPFVEVL437LV	19%	1922
12	RP	17%	1749
13	RPAPAM	17%	1747

Similar to the previous reactions (substrate concentration at 5 mM), they were dominated by KSK19/AI based $P450_{BM3}$ vairants. Y51L/KSK19/AI showed 38% conversion of 10 mM DMA to MA, keeping turnover number more or less unchanged. Therefore, Y51L/KSK19/AI was subjected to an *in vivo* biotransformation in phosphate buffer (200 mM pH 7.5). DMA was added in 2 mM aliquots to a final concentration of 20 mM. The reaction was stirred at room temperature for 48 hours after which total conversion was observed and methyl anthranilate was found to be the single product.

These results highlighted the possibility of utilising $P450_{BM3}$ enzymes in the place of homogeneous and heterogeneous acid catalysts in the synthesis of methyl anthranilate. Compared to chemical catalysts, enzymes are efficient, selective, can act under mild conditions and produce natural grade products. Further work needs to be done on the optimisation of such reaction system (buffer, substrate/enzyme ratio, dissolved oxygen level, bioreactor vessel, etc.) that make it a viable industrial approach in methyl anthranilate synthesis.

Chapter Six

Chapter 6 Thesis summary and future work

The aim of this study was to evolve and identify a small panel of $P450_{BM3}$ variants based on 'generic accelerator' mutants (IP, KT2, KSK19, A330P, etc.) that would be useful in the oxidation of natural and non-natural substrates. Mutations at 47 amino acid residues along the substrate entrance channel were introduced to incorporate a decent level of diversity in topology while inactive mutants were discarded in order to keep the library to a manageable size. The purified enzymes were adjusted to 10 μM and stored in 24-well plates (100 μL), making the enzyme system a biological toolkit of C–H activation catalysts that could be applied in drug discovery, fine chemical synthesis, and new function studies, etc. This toolkit may act as a 'fishing net' to obtain hits with new substrates and accumulate preliminary information of the links between key amino acid residues and substrate oxidation activity. Positive mutants would be identified in most cases while oxidation at various sites would be obtained. The toolkit can ultimately be used as a resource for testing non-natural substrates, selecting starting templates, identifying key amino acid residues, collecting information of mutation-selectivity trends and therefore guiding further protein engineering.

Some successful examples were discussed in this thesis. Chapter 3 described the application and refinement of a panel of ~100 $P450_{BM3}$ variants on anionic, cationic and neutral drug molecules for the synthesis of their human metabolites since expression of human drug metabolising cytochrome P450 enzymes remains problematic. Although no single mutant or family of mutants showed high conversion with all the tested drugs, >90% conversion of all the six drug molecules being tested (diclofenac, naproxen, testosterone, chlorzoxazone, lidocaine) was observed with the mutant enzyme library, which enabled full products characterisation and led to the discovery of two new P450 reaction types – decarboxylation of an α-hydroxy carboxylic acid and intramolecular cyclisation of an amine. The formation of a

stable, protected imine by cyclization of lidocaine offers a novel route for α-functionalization of amines.

In chapter 4, the potential of utilising engineered $P450_{BM3}$ mutant enzymes for synthesising large numbers of small, polar fragment molecules with lead-likeness and functional diversity was explored. 40 Drug fragment molecules representing a variety of subclasses (*N*-heterocycles, *O*-heterocycles, amines, amides, alcohols, carboxylic acids and nitriles) were selected by GlaxoSmithKline for evaluation. Positive hits were observed with 36 compounds with the library of $P450_{BM3}$ mutants being tested and 10 of them were studied in more detail. Total conversion was observed towards all the 10 fragments except acetylpyrazine and ε-caprolactone, suggesting the effectiveness of mutant enzymes in overcoming the issues of substrate inhibition and poor substrate binding. Multiple products were generated, thus offering a diverse functionality although the number of products varied widely depending on the number of available oxidation sites and the ease with which they could be hydroxylated. In addition, chemically vulnerable groups (alcohols, amines, etc.) were not attached, making the use of $P450_{BM3}$ mutant enzymes an attractive method for synthesis as protection-deprotection strategies were not necessary. The overall results could be recognised as an exciting beginning for the development of Lead-Oriented Synthesis in drug discovery.

Chapter 5 was focused on the development of a scalable enzymatic C–H amination process for the straightforward synthesis of monocyclic, bicyclic and tricyclic imidazolidin-4-ones while other applications of the $P450_{BM3}$ mutant enzyme library were also reported. By screening the cyclisation activity towards a range of 2-aminoacetimide derivatives, it was found that aromatic and aliphatic amides as well as substituents on the amine ring were accepted. Cyclisation of the tetrahydroisoquinoline derivative to a tricyclic compound further highlights the tolerance of $P450_{BM3}$ mutants for the variations in the amine moiety and the possibility of controlled site-selective C–H amination. The increased conversion for the *in*

vivo C–H amination over reactions *in vitro* adds to the versatility of the system in order to access imidazolidin-4-ones, which are latent iminium ions and amendable to Lewis acid-mediated ring-opening with subsequent nucleophilic attack offering new routes of α-functionalisation of tertiary amines.

Two radianspene core compounds were tested for late-stage functionalisation and found to be hydroxylated at inactivated C–H bonds. Although multiple products remained uncharacterised due to the limited supply of starting materials, these results can be recognised as a proof-of-concept study on introducing synthetic handles at the very last steps with the aim of developing the first total synthesis of radianspene J.

The flavour and fragrance compound methyl anthranilate was another example of successful applications of $P450_{BM3}$ mutant library. 29 Variants were examined towards *N*-dealkylation of dimethyl anthranilate, of which 13 mutants displayed over 40% conversion in 2 hours with turnover number exceeding 2,000. These results highlighted the possibility of utilising $P450_{BM3}$ enzymes in the place of homogeneous and heterogeneous acid catalysts in the synthesis of methyl anthranilate. Further work needs to be done on the optimisation of such reaction system (buffer, substrate/enzyme ratio, dissolved oxygen level, bioreactor vessel, etc.) that make it a viable industrial synthesis of methyl anthranilate.

There is intense interest in late stage C–H bond functionalization as a means to streamline synthesis. Enzymes offer advantages over transition metal catalysts because they operate under mild conditions in aqueous solvent, are highly active, essentially inexhaustible, and can be evolved to be highly selective for the desired product. A prerequisite for this is an enzyme collection with wide substrate range, excellent functional group tolerance, diverse product selectivity and high conversion to maximize the likelihood of initial hits for further selectivity optimization. The results suggest that the library of $P450_{BM3}$ variants is the basis for one such collection.

Throughout this thesis, the $P450_{BM3}$ mutant enzyme library underwent rounds of refinement. The latest compositions are shown in Table 6.1.

Table 6.1 The present $P450_{BM3}$ toolkit in 24-well plates

No	A plate	B plate
1	RLYF KSK19 AI	A330P
2	KSK19 A82M	RPF81W
3	KSK19AI Y51L	RT2APAI S72A
4	KSK19AIAI	RP/H171L/I263G
5	RPFVEV L437VL	RT2APPV
6	KSK19 I263A	GVQ
7	RP	RT2 A330W
8	RP EV	F87A
9	RT2F81W	RLYF KSK19
10	RPFVEV	RP AM IA
11	KSK19AI F87V	RT2 V78F
12	RPFVEV L437LV	RLYF
13	KSK19AI V78I	KSK19/A82M/E267F
14	RT2AP/I401M	RP H171L
15	RT2APAM	RP/E267G
16	RT2IP	KSK48
17	RT2AP	KSK19 QP F87V
18	RT2APA184I	RT2AP VIAI
19	RT2APV78I	RPFVEV L188Q
20	RT2	RPFVEV V78F
21	RPIAEV	RT2F81W A328I
22	WT	RP/E267M
23	RPAP	KSK19 AI IA
24	RP FV	RT2/S72W/A330W

Future work includes the refinement of $P450_{BM3}$ toolkit for C–H activation synthesis, further protein engineering to increase thermal stability and resistance to organic solvents, the development of porous material and immobilisation of enzymes, and industrial fermentation and biotransformation, etc. It is noteworthy to point out some transformations that can be accessed by evolving the present $P450_{BM3}$ mutant enzyme library.

a) fatty acid decarboxylation

As mentioned in chapter 1, P450 OleT has been found to catalyse decarboxylation of linear short-chain fatty acids (C4-C9) to give the corresponding terminal C3 to C8 alkenes. Product titers were increased to *ca.* 0.93 gL^{-1} with TTNs over 2,000 based on the development of enzymatic redox cascade process designed by Faber and co-workers. The system employed reductase and putidaredoxin CamAB in combination with NAD(P)H recycling at the expense of glucose, formate or phosphate. Since $P450_{BM3}$ is a self-sufficient enzyme that shares similar characteristic P450 skeleton and is able to generate highly reactive intermediate (Compound I), there is a potential to engineer $P450_{BM3}$ into a fatty acid decarboxylase. Proposed mechanism is shown in Figure 6.1.

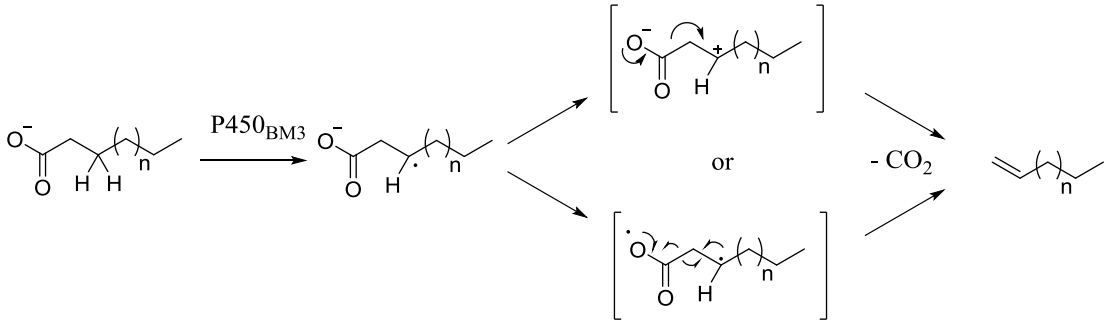


Figure 6.1 Proposed mechanism of fatty acid decarboxylation catalysed by $P450_{BM3}$

b) α-functionalisation of amines

There is intense research interest in developing new methodologies for α-functionalisation of amines in synthetic chemistry. As described in chapter 5, an intramolecular α-amine functionalisation method has been successfully applied in the synthesis of imidazolin-4-ones with well controlled regio-selectivity in which iminium intermediates were generated and acted as electrophiles. If the key iminium ions can be attacked by various nucleophiles intermolecularly, important functional groups would be installed at the α-amine position. (as shown in Figure 6.2) This could be an exciting project as well as a problematic one. Protein engineering could be challenging as the enzyme has to allow at least two molecules stay in the active site while the dimerisation of amine molecules would need to be prevented.

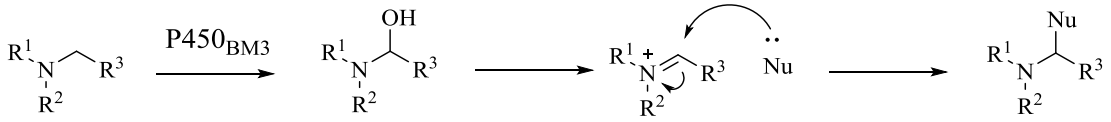


Figure 6.2 Proposed mechanism of α-functionalisation of amines

Appendices

Appendix 1 Full DNA sequence of Wild-type (WT) $P450_{BM3}$

5′ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGT
TATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGGAGAA
ATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTATCAAGTCAGCGTCTAATT
AAAGAAGCATGCGATGAATCACGCTTTGATAAAAACTTAAGTCAAGCGCTTAAATTTGT
ACGTGATTTTGCAGGAGACGGGTTATTTACAAGCTGGACGCATGAAAAAAATTGGAAAA
AAGCGCATAATATCTTACTTCCAAGCTTCAGTCAGCAGGCAATGAAAGGCTATCATGCGA
TGATGGTCGATATCGCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAG
CATATTGAAGTACCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGC
TTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTACAAGTATGG
TCCGTGCACTGGATGAAGCAATGAACAAGCTGCAGCGAGCAAATCCAGACGACCCAGCT
TATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGATGAACGACCTAGTAGA
TAAAATTATTGCAGATCGCAAAGCAAGCGGTGAACAAAGCGATGATTTATTAACGCATA
TGCTAAACGGAAAAGATCCAGAAACGGGTGAGCCGCTTGATGACGAGAACATTCGCTAT
CAAATTATTACATTCTTAATTGCGGGACACGAAACAACAAGTGGTCTTTTATCATTTGCG
CTGTATTTCTTAGTGAAAAATCCACATGTATTACAAAAAGCAGCAGAAGAAGCAGCACG
AGTTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCAT
GGTCTTAAACGAAGCGCTGCGCTTATGGCCAACTGCTCCTGCGTTTTCCCTATATGCAAA
AGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAACTAATGGTTC
TGATTCCTCAGCTTCACCGTGATAAAACAATTTGGGGAGACGATGTGGAAGAGTTCCGTC
CAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGAAACG
GTCAGCGTGCGTGTATCGGTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTA
TGATGCTAAAACACTTTGACTTTGAAGATCATACAAACTACGAGCTGGATATTAAAGAA
ACTTTAACGTTAAAACCTGAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTT
GGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGA
AAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGAACAGCTGAAG
GAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACCGCAGGTCGCA
ACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCTGTATTAATTGTAACGGCG
TCTTATAACGGTCATCCGCCTGATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCG
TCTGCTGATGAAGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAACTGG
GCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCA
GAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGA
AGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAA
ACAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGC
CGCTTGCGAAAATGCACGGTGCGTTTTCAACGAACGTCGTAGCAAGCAAAGAACTTCAA
CAGCCAGGCAGTGCACGAAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTC
TTATCAAGAAGGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCG
TGTAACAGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAAG
AAAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCAATACG
TGGAGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGCAATGGCTGCTAAAACGGTCT
GCCCGCCGCATAAAGTAGAGCTTGAAGCCTTGCTTGAAAAGCAAGCCTACAAAGAACAA
GTGCTGGCAAAACGTTTAACAATGCTTGAACTGCTTGAAAAATACCCGGCGTGTGAAAT
GAAATTCAGCGAATTTATCGCCCTTCTGCCAAGCATACGCCCGCGCTATTACTCGATTTCT
TCATCACCTCGTGTCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGA
AGCGTGGAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAGCTGC

AAGAAGGAGATACGATTACGTGCTTTATTTCCACACCGCAGTCAGAATTTACGCTGCCAA
AAGACCCTGAAACGCCGCTTATCATGGTCGGACCGGGAACAGGCGTCGCGCCGTTTAGA
GGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAACAAGGACAGTCACTTGGAGAAGCACA
TTTATACTTCGGCTGCCGTTCACCTCATGAAGACTATCTGTATCAAGAAGAGCTTGAAAA
CGCCCAAAGCGAAGGCATCATTACGCTTCATACCGCTTTTTCTCGCATGCCAAATCAGCC
GAAAACATACGTTCAGCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTG
ATCAAGGAGCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAA
GCAACGCTTATGAAAAGCTATGCTGACGTTCACCAAGTGAGTGAAGCAGACGCTCGCTT
ATGGCTGCAGCAGCTAGAAGAAAAAGGCCGATACGCAAAAGACGTGTGGGCTGGGTAA3′

Appendix 2 Primers and mutant enzymes

Table S2.1 Oligonucleotide primers used for site-directed mutagenesis in this thesis. Sequence reads from 5′ to 3′ and the codons for the mutated residue are highlighted

Primer name	Sequence	T_M (℃)
L20V	CCGTTAGTAAACACAGATAAACCGGTTCAAG	60.4
L20V rev	GGTTTATCTGTGTTTACTAACGGTAAATTTTTAAGC	58.8
L20A	CCGTTAGAAAACACAGATAAACCGGTTCAAG	60.4
L20A rev	GGTTTATCTGTGTTTTCTAACGGTAAATTTTTAAGC	58.8
L29A	GGTTCAAGCTGCGATGAAAATTGCGGATG	61.5
L29A rev	CAATTTTCATCGCAGCTTGAACCGGTTTATCG	61.8
L29V	GGTTCAAGCTGTGATGAAAATTGCGGATG	60.1
L29V rev	CAATTTTCATCACAGCTTGAACCGGTTTATCG	60.5
F40L	GGAGAAATCTTGAAATTCGAGGCGCCTG	61.4
F40L rev	CGCCTCGAATTTCAAGATTTCTCCTAATTCATCC	62.0
F42V	GAAATCTTTAAAGTCGAGGCGCCTGGTC	61.4
F42V rev	GCGCCTCGACTTTAAAGATTTCTCCTAATTC	60.4
F42L	GAAATCTTTAAATTAGAGGCGCCTGGTCGTGTAAC	63.3
F42L rev	CCAGGCGCCTCTAATTTAAAGATTTCTCCTAATTC	62.1
R47L	GCGCCTGGTCTTGTAACGCGCTACTTATCAAGTCAGCG	69.9
R47L reverse	CGCTGACTTGATAAGTAGCGCGTTACAAGACCAGGCGC	69.9
Y51F	CGTGTAACGCGCTTCTTATCAAGTCAGCGTCTAATTAAAG	65.5
Y51F reverse	CTGACTTGATAAGAAGCGCGTTACACGACCAGGCGCC	70.0
Y51V	CGTGTAACGCGCGTCTTATCAAGTCAGCGTC	65.7
Y51V rev	GACTTGATAAGACGCGCGTTACACGACCAG	64.4
Y51L	CGTGTAACGCGCTTGTTATCAAGTCAGCG	62.9
Y51L rev	GACTTGATAAGAAGCGCGTTACACGACCAG	63.0
S72L	CACGCTTTGATAAAAACTTACTTCAAGCGCTTAAATTTGTACG	63.6
S72L rev	CGTACAAATTTAAGCGCTTGAAGTAAGTTTTTATCAAAGCGTG	63.6
S72A	GATAAAAACTTAGCTCAAGCGCTTAAATTTGTACGTGATTTTG	62.6

S72A rev	CAAATTTAAGCGCTTGAGCTAAGTTTTTATCAAAGCGTGATTC	63.6
S72R	CACGCTTTGATAAAAACTTACGTCAAGCGCTTAAATTTGTACG	64.5
S72R rev	CAAAATCACGTACAAATTTAAGCGCTTGACGTAAGTTTTTATC	62.6
S72V	CGCTTTGATAAAAACTTAGGTCAAGCGCTTAAATTTGTACGTG	64.5
S72V rev	CACGTACAAATTTAAGCGCTTGAGCTAAGTTTTTATCAAAGCG	64.5
S72D	CGCTTTGATAAAAACTTAGATCAAGCGCTTAAATTTGTACGTG	63.6
S72D rev	CACGTACAAATTTAAGCGCTTGATCTAAGTTTTTATCAAAGCG	63.6
S72R/L75Y	CACGCTTTGATAAAAACTTACGTCAAGCGTATAAATTTGTAC	62.6
S72R/L75Y rev	CAAAATCACGTACAAATTTATACGCTTGACGTAAGTTTTTATC	61.7
S72V/V78F	CGCTTTGATAAAAACTTAGGTCAAGCGCTTAAATTTTTCCGTG	64.5
S72V/V78F rev	CACGGAAAAATTTAAGCGCTTGACCTAAGTTTTTATCAAAGCG	64.5
S72D/V78F	CGCTTTGATAAAAACTTAGATCAAGCGCTTAAATTTTTCCGTG	63.6
S72D/V78F rev	CACGGAAAAATTTAAGCGCTTGATCTAAGTTTTTATCAAAGCG	63.6
A74V	GATAAAAACTTAAGTCAAGTGCTTAAATTTGTACG	57.4
A74V rev	CGTACAAATTTAAGCACTTGACTTAAGTTTTTATC	57.4
A74G	GCTTTGATAAAAACTTAAGTCAAGGGCTTAAATTTG	58.8
A74G rev	CGTACAAATTTAAGCCCTTGACTTAAGTTTTTATCA	58.8
A74W	GATAAAAACTTAAGTCAATGGCTTAAATTTGTACGTG	58.9
A74W rev	CACGTACAAATTTAAGCCATTGACTTAAGTTTTTATC	58.9
L75V	CGCTTTGATAAAAACTTAAGCCAAGCGGTTAAATTT	59.9
L75V rev	CCCGTCTCCTGCAAAATCACGTACAAATTTAACCGC	65.5
V78I	AGTCAAGCGCTTAAATTTATACGTGATTTTGCAGGAGACGGG	66.5
V78I reverse	CCCGTCTCCTGCAAAATCACGTATAAATTTAAGCGCTTGACT	66.5
V78F	AGTCAAGCGCTTAAATTTTTCCGTGATTTTGCAGGAGACGGG	67.4
V78F reverse	CCCGTCTCCTGCAAAATCACGGAAAAATTTAAGCGCTTGACT	67.4
V78R	CTTAAGTCAAGCGCTTAAATTTCGACGTGATTTTGCA	62.2
V78R rev	CCGTCTCCTGCAAAATCACGTCGAAATTTAAGCGC	65.6
V78E	CTTAAGTCAAGCGCTTAAATTTGAACGTGATTTTGCA	61.1
V78E rev	CCGTCTCCTGCAAAATCACGTTCAAATTTAAGCGC	64.4
V78I/A82M	GTCAAGCGCTTAAATTTATACGTGATTTTATGGGAGACGGG	65.5
V78I/A82M rev	GTAAATAACCCGTCTCCCATAAAATCACGTATAAATTTAAG	61.5

F81P	GCGCTTAAATTTGTACGTGATCCTGCAGGAGACGGG	67.9
F81P rev	GCTTGTAAATAACCCGTCTCCTGCAGGATCACGTAC	66.7
F81W	GCGCTTAAATTTGTACGTGATTGGGCAGGAGACGGGTTATTTACA	68.2
F81W reverse	TGTAAATAACCCGTCTCCTGCCCAATCACGTACAAATTTAAGCGC	68.2
F81W/A82M	GTACGTGATTGGATGGGAGACGGGTTATTTACAAGC	65.6
F81W/A82M rev	GTAAATAACCCGTCTCCCATCCAATCACGTACAAATTTAAG	64.5
A82M	GCGCTTAAATTTGTACGTGATTTTATGGGAGACGGG	64.4
A82M rev	GCTTGTAAATAACCCGTCTCCCATAAAATCACGTAC	63.3
A82L	GTACGTGATTTTTTAGGAGACGGGTTATTTACAAGCTGGACG	66.5
A82L rev	CCCGTCTCCTAAAAAATCACGTACAAATTTAAGCGCTTG	64.5
A82P	CTTAAATTTGTACGTGATTTTCCAGGAGACGGGTTATTTACAAGC	65.4
A82P rev	GCTTGTAAATAACCCGTCTCCTGGAAAATCACGTACAAATTTAAG	65.4
G83P	GTACGTGATTTTGCACCAGACGGGTTATTTACAAGC	64.4
G83P rev	CCAGCTTGTAAATAACCCGTCTGGTGCAAAATCACG	65.5
L86P	GATTTTGCAGGAGACGGGCCGTTTACAAGCTGGACG	69.0
L86P rev	GCGTCCAGCTTGTAAACGGCCCGTCTCCTGCAAAAT	69.0
F87A	CGTGATTTTGCAGGAGACGGGTTAGGTACAAGCTGGACGCATG	71.2
F87A Reverse	CATGCGTCCAGCTTGTACCTAACCCGTCTCCTGCAAAATCACG	71.2
F87V	CGTGATTTTGCAGGAGACGGGTTAGTAACAAGCTGG	66.7
F87V reverse	CCAGCTTGTTACTAACCCGTCTCCTGCAAAATCACG	66.7
F87P	GCAGGAGACGGGTTACCGACAAGCTGGACGCATG	70.4
F87P rev	CATGCGTCCAGCTTGTCGGTAACCCGTCTCCTGC	70.4
F87S	GCAGGAGACGGGTTATCCACAAGCTGGACGCATG	69.2
F87S rev	CATGCGTCCAGCTTGTGGATAACCCGTCTCCTGC	69.2
F87I	CGTGATTTTGCAGGAGACGGGTTAATTACAAGCTGG	65.6
F87I rev	TTTTTCATGCGTCCAGCTTGTAATTAACCCGTCTCC	64.4
F87L	CGTGATTTTGCAGGAGACGGGTTACTTACAAGCTGG	66.7
F87L rev	TTTTTCATGCGTCCAGCTTGTAAGTAACCCGTCTCC	65.5
F87G	CGTGATTTTGCAGGAGACGGGTTAGGTACAAGCTGG	67.9
F87G rev	CATGCGTCCAGCTTGTACCTAACCCGTCTCCTGCAA	69.0
F87A/A82M	CGCTTAAATTTGTACGTGATTTTATGGGAGACGGGTTAGCGACAA	67.3

F87A/A82M rev	CCAGCTTGTCGCTAACCCGTCTCCCATAAAATCACGTACAAATTT	68.2
F87A/T88A	GCAGGAGACGGGTTAGCGGCAAGCTGGACGCATG	71.7
F87A/T88A rev	CATGCGTCCAGCTTGCCGCTAACCCGTCTCCTGC	71.7
T88P	GCAGGAGACGGGTTATTTCCAAGCTGGACGCATGAAAAAAATTGG	69.1
T88P rev	CCAATTTTTTTCATGCGTCCAGCTTGGAAATAACCC	63.3
T88M	GCAGGAGACGGGTTATTTATGAGCTGGACGCATGAAAAAAATTGG	68.2
T88M rev	CCAATTTTTTTCATGCGTCCAGCTCATAAATAACCC	62.2
H171L	TACCGAGATCAGCCTCTTCCATTTATTACAAGTATGG	63.3
H171L rev	CCATACTTGTAATAAATGGAAGAGGCTGATCTCGGTA	63.3
M177K	CCATTTATTACAAGTAAGGTCCGTGCACTGGATG	63.2
M177K rev	CATCCAGTGCACGGACCTTACTTGTAATAAATGG	63.2
M177V	CCATTTATTACAAGTGTGGTCCGTGCACTGGATG	64.4
M177V rev	CATCCAGTGCACGGACCACACTTGTAATAAATGG	64.4
A184I	CGTGCACTGGATGAAATAATGAACAAGCTGCAGC	64.4
A184I reverse	GCTGCAGCTTGTTCATTATTTCATCCAGTGCACG	64.4
L188A	GCAATGAACAAGGCGCAGCGAGCAAATCC	64.3
L188A rev	GCTCGCTGCGCCTTGTTCATTGCTTCATC	64.3
L188V	GCACTGGATGAAGCAATGAACAAGGTTCAGCGAGCA	66.7
L188V rev	TGCTCGCTGAACCTTGTTCATTGCTTCATCCAGTGC	66.7
L188Q	GAAGCAATGAACAAGCAGCAGCGAGCAAATCCAG	65.6
L188Q reverse	CTGGATTTGCTCGCTGCTGCTTGTTCATTGCTTC	65.6
A191P	GCAATGAACAAGCTGCAGCGACCAAATCCAGACGAC	67.9
A191P rev	CTGGGTCGTCTGGATTTGGTCGCTGCAGCTTGTTCA	69.0
H239N	GATTTATTAACGCATATGCTAAACGGAAAAGATCCA	59.9
H239N rev	CCCGTTTCTGGATCTTTTCCGTTTACGATATGCGTT	64.4
I259V	CATTCGCTATCAAATTGTTACATTCTTAATTGCGGG	61.0
I259V rev	CCCGCAATTAAGAATGTAACAATTTGATAGCGAATG	61.0
I259V/I263W	CAAATTGTTACATTCTTATGGGCGGGACACGAAACAACAAGTGGT	67.3
I259V/I263W rev	ACCACTTGTTGTTTCGTGTCCCGCCCATAAGAATGTAACAATTTG	67.3
I263A	CGCTATCAAATTATTACATTCTTAGCCGCGGGACAC	64.4
I263A reverse	GTGTCCCGCGGCTAAGAATGTAATAATTTGATAGCG	64.4

I263V	CGCTATCAAATTATTACATTCTTAGTTGCGGGACACGAAACAACA	65.4
I263V rev	CTTGTTGTTTCGTGTCCCGCAACTAAGAATGTAATAATTTGATAG	64.5
A264G	CAAATTATTACATTCTTAATTGGGGGACACGAAACA	59.9
A264G rev	CCACTTGTTGTTTCGTGTCCCCCAATTAAGAATGTA	63.3
E267V	CTTAATTGCGGGACACGTAACAACAAGTGGTCTTTT	63.3
E267V reverse	AAAAGACCACTTGTTGTTACGTGTCCCGCAATTAAG	63.3
E267L	CTTAATTGCGGGACACCTAACAACAAGTG	65.3
E267L reverse		63.3
E267V/S270I	CATTCTTAATTGCGGGACACGTAACAACAATTGGTC	63.3
E267V/S270I rev	CAGCGCAAATGATAAAAGACCAATTGTTGTTACGTG	62.2
A328I	CTGCGCTTATGGCCAACTATTCCTGCGTTTTCCC	66.8
A328I reverse	GGGAAAACGCAGGAATAGTTGGCCATAAGCGCAG	66.8
A328G	CGCTTATGGCCAACTGGGCCTGCGTTTTCCCTA	68.1
A328G rev	TAGGGAAAACGCAGGCCCAGTTGGCCGTAAGCG	69.4
A328P	GCGCTGCGCTTATGGCCAACTCCTCCTGCGTTTTCC	71.3
A328P rev	GCATATAGGGAAAACGCAGGAGGAGTTGGCCATAAG	66.7
A328S/P329S	CTGCGCTTATGGCCAACTTCTTCTGCGTTTTCCCTA	66.7
A328S/P329S rev	GCATATAGGGAAAACGCAGAAGAAGTTGGCCATAAGCGCAG	68.5
A328I/A330P	CTGCGCTTATGGCCAACTATTCCTCCTTTTTCCCTATATGC	67.5
A328I/A330P rev	GCATATAGGGAAAAAGGAGGAATAGTTGGCCATAAGCGCAG	67.5
A328V/A330P	CGCTTATGGCCAACTGTTCCTCCGTTTTCCCTATATGC	67.7
A328V/A330P rev	GCATATAGGGAAAACGGAGGAACAGTTGGCCATAAGCG	67.7
P329V/A330P	CTGCGCTTATGGCCAACTGCTGTTCCTTTTTCCCTATATGC	68.5
P329V/A330P rev	GCATATAGGGAAAAAGGAACAGCAGTTGGCCATAAGCGCAG	68.5
A330P	GGCCAACTGCTCCTCCGTTTTCCCTATATGC	65.7
A330P reverse	GCATATAGGGAAAACGGAGGAGCAGTTGGCC	65.7
A330W	CCAACTGCTCCTTGGTTTTCCCTATATGCAAAAGAAGATACG	66.5
A330W rev	GCATATAGGGAAAACCAAGGAGCAGTTGGCCATAAGCGCAGCGCT	71.8
L353I	GAAAAAGGCGACGAAATAATGGTTCTGATTCCTC	62.0
L353I rev	GAGGAATCAGAACCATTATTTCGTCGCCTTTTTC	62.0
I401P	GGAAACGGTCAGCGTGCGTGTCCCGGTCAGCAGTTCGCTCTTC	75.0

I401P reverse	GAAGAGCGAACTGCTGACCGGGACACGCACGCTGACCGTTTCC	75.0
I401M	GAAACGGTCAGCGTGCGTGTATGGGTCAGCAGTTC	69.1
I401M rev	CATGAAGAGCGAACTGCTGACCCATACACGCACG	68.0
I401P/Q403P	GGTCAGCGTGCGTGTCCCGGTCCGCAGTTCGCTCTTC	74.4
I401P/Q403P rev	GAAGAGCGAACTGCGGACCGGGACACGCACGCTGACC	74.4
Q403P	GGTCAGCGTGCGTGTATCGGTCCGCAGTTCGCTCTTCATGAAGC	73.8
Q403P reverse	GCTTCATGAAGAGCGAACTGCGGACCGATACACGCACGCTGACC	73.8
L437S	CTGGATATTAAAGAAACTTCAACGTTAAAACCTGAAGGC	62.4
L437S rev	GCCTTCAGGTTTTAACGTTGAAGTTTCTTTAATATCCAG	62.4
L437F	CTGGATATTAAAGAAACTTTCACGTTAAAACCTGAAGGC	62.4
L437F rev	GCCTTCAGGTTTTAACGTGAAAGTTTCTTTAATATCCAG	62.4
L437VL	CTGGATATTAAAGAAACTGTCTTAACGTTAAAACCTGAAGGC	63.5
L437VL rev	GCCTTCAGGTTTTAACGTTAAGACAGTTTCTTTAATATCCAG	63.5
L437LV	CTGGATATTAAAGAAACTTTAGTCACGTTAAAACCTGAAGGC	63.5
L437LV rev	GCCTTCAGGTTTTAACGTGACTAAAGTTTCTTTAATATCCAG	63.5
T438G	GAGGTCGATATTAAAGAAACTTTAGGGTTAAAACCTGAAGGC	64.5
T438G rev	GCCTTCAGGTTTTAACCCTAAAGTTTCTTTAATATCGACCTC	64.5
T436V/T438V	GATATTAAAGAAGTTTTAGTGTTAAAACCTGAAGG	57.4
T436V/T438V rev	GGTTTTAACACTAAAACTTCTTTAATATCCAGCTC	58.6
T438L	GAGGTCGATATTAAAGAAACTTTACTGTTAAAACCTGAAGGC	63.5
T438L rev	GCCTTCAGGTTTTAACAGTAAAGTTTCTTTAATATCGACCTC	63.5
T438A	GAGGTCGATATTAAAGAAACTTTAGCGTTAAAACCTGAAGGC	64.5
T438A rev	GCCTTCAGGTTTTAACGCTAAAGTTTCTTTAATATCGACCTC	64.5
Q464 STOP	CCTTCACCTAGCACTGAATAGTCTGCTAAAAAAGTAGC	64.5
Q464 STOP rev	GCTACTTTTTTAGCAGACTATTCAGTGCTAGGTGA	64.5

Table S2.2 Purified $P450_{BM3}$ enzymes in this paper and mutations in the variants

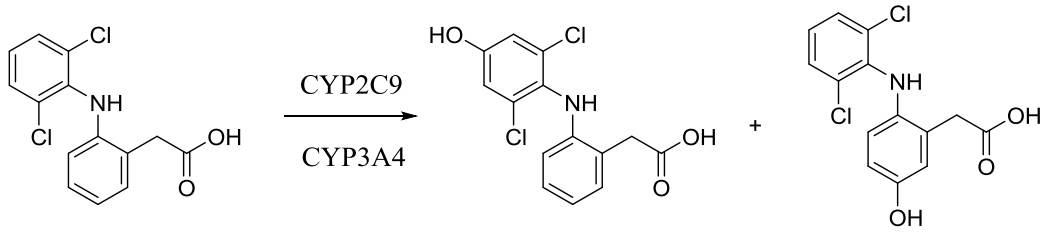
Enzymes	**Mutations**
A191P/Q403P	A191P/Q403P
A328S/P329S/Q403P	A328S/P329S/Q403P
ASPSFAIP	F87A/A328S/P329S/I401P
F87A	F87A
F87A I401P	F87A/I401P
F87A/T88A/I401P	F87A/T88A/I401P
F87S/I401P	F87S/I401P
GVQ	A74G/F87V/L188Q
I401P	I401P
KSK19	F87A/H171L/Q307H/N319Y
KSK19/ IA	F87A/H171L/I263A/Q307H/N319Y
KSK19/ IA/QP	F87A/H171L/I263A/Q307H/N319Y/Q403P
KSK19/A82M	F87A/A82M/H171L/Q307H/N319Y
KSK19/AI	F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/AI	F87A/H171L/A184I/Q307H/N319Y/A328I
KSK19/AI/F42L	F42L/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/FV	F87V/H171L/Q307H/N319Y/A328I
KSK19/AI/IA	F87A/ H171L/I263A/Q307H/N319Y/A328I
KSK19/AI/IP	F87A/H171L/Q307H/N319Y/A328I/I401P
KSK19/AI/L29A	L29A/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/R47L	R47L/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/S72A	S72A/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/VI	V78I/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/Y51F	Y51F/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/Y51L	Y51L/F87A/H171L/Q307H/N319Y/A328I
KSK19/AI/Y51V	Y51V/F87A/H171L/Q307H/N319Y/A328I
KSK19/F87S	F87S/ H171L/Q307H/N319Y
KSK19/QP	F87A/H171L/Q307H/N319Y/Q403P
KSK19/QP/FV	F87V/H171L/Q307H/N319Y/Q403P
KSK19/T88A	F87A/T88A/H171L/Q307H/N319Y
KT2	A191T/N239H/I259V/A276T/L353I

KT2/F87L	F87L/A191T/N239H/I259V/A276T/L353I
QP	Q403P
QP/A82M	A82M/Q403P
RLYF	R47L/Y51F
RLYF/F87A/A82G	R47L/Y51F/A82G/F87A
RLYF/F87A/L75V	R47L/Y51F/L75V/F87A
RLYF/I263A	R47L/Y51F/I263A
RLYF/KSK19	R47L/Y51F/F87A/H171L/Q307H/N319Y
RLYF/KSK19/AI	R47L/Y51F/F87A/H171L/Q307H/N319Y/A328I
RLYF/KSK19/IP	R47L/Y51F/F87A/H171L/Q307H/N319Y/I401P
RLYF/L75V	R47L/Y51F/L75V
RLYF/LA/FA	R47L/Y51F/L75A/F87A
RLYF/LA	R47L/Y51F/L75A
RLYF/QP/A82M/I263A	R47L/Y51F/A82M/I263A/Q403P
RP	R47L/Y51F/I401P
RP/A82M/I263A	R47L/Y51F/A82M/I263A/I401P
RP/EV	R47L/Y51F/E267V/I401P
RP/EV/I263A	R47L/Y51F/I263A/E267V/I401P
RP/F81W	R47L/Y51F/I401P/ F81W
RP/FA	R47L/Y51F//F87A/I401P
RP/FV	R47L/Y51F/F87V/I401P
RP/FV/EV/A328I	R47L/Y51F/F87V/E267V/A328I/I401P
RP/H171L	R47L/Y51F/H171L/I401P
RPAP	R47L/Y51F/A330P/I401P
RPAP/A328V	R47L/Y51F/A328V/A330P/I401P
RPAP/A74W	R47L/Y51F/A74W/A330P/I401P
RPAP/AM	R47L/Y51F/A82M/A330P/I401P
RPAP/AM/A184I	R47L/Y51F/A82M/A184I/A330P/I401P
RPAP/PV	R47L/Y51F/P329V/A330P/I401P
RPAP/S72L	R47L/Y51F/S72L/A330P/I401P
RPFVEV	R47L/Y51F/F87V/E267V/I401P
RPFVEV/A184I	R47L/Y51F/F87V/A184I/E267V/I401P
RPFVEV/F81W	R47L/Y51F/F81W/F87V/E267V/I401P
RPFVEV/L188Q	R47L/Y51F/F87V/L188Q/E267V/I401P

RPFVEV/L437LV	R47L/Y51F/F87V/E267V/I401P/L437LV
RPFVEV/L437VL	R47L/Y51F/F87V/E267V/I401P/L437VL
RPFVEV/V78F	R47L/Y51F/V78F/F87V/E267V/I401P
RT2	R47L/Y51F/A191T/N239H/I259V/A276T/L353I
RT2/A330W	R47L/Y51F/A191T/N239H/I259V/A276T/A330W/L353I
RT2/A330W/A82M	R47L/Y51F/A82M/A191T/N239H/I259V/A276T/A330W/L353I
RT2/AP/A184I/F81W	R47L/Y51F/F81W/A184I/A191T/N239H/I259V/A276T/A330P/L353I
RT2/AP/AI/S72A	R47L/Y51F/S72A/A184I/A191T/N239H/I259V/A276T/A330P/L353I
RT2/AP/AM	R47L/Y51F/A82M/A191T/N239H/I259V/A276T/A330P/L353I
RT2/AP/VI/AI	R47L/Y51F/V78I/A184I/A191T/N239H/I259V/A276T/A330P/L353I
RT2/F81W/A328I	R47L/Y51F/F81W/A191T/N239H/I259V/A276T/A328I/L353I
RT2/IP/A184I	R47L/Y51F/A184I/A191T/N239H/I259V/A276T/L353I/I401P
RT2/IP/AP	R47L/Y51F/A191T/N239H/I259V/A276T/A330P/L353I/I401P
RT2/IP/F81W	R47L/Y51F/F81W/A191T/N239H/I259V/A276T/L353I/I401P
RT2/V78F	R47L/Y51F/V78F/A191T/N239H/I259V/A276T/L353I
RT2AI/L437LV	R47L/Y51F/A191T/N239H/I259V/A276T/A328I/L353I/L437LV
RT2AI/L437VL	R47L/Y51F/A191T/N239H/I259V/A276T/A328I/L353I/L437VL
RT2AP/A184I	R47L/Y51F/A184I/A191T/N239H/I259V/A276T/A330P/L353I
RT2AP/A82L	R47L/Y51F/A82L/A191T/N239H/I259V/A276T/A330P/L353I
RT2AP/AM/AI	R47L/Y51F/A82M/A184I/A191T/N239H/I259V/A276T/A330P/L353I
RT2AP/F81W	R47L/Y51F/F81W/A191T/N239H/I259V/A276T/A330P/L353I
RT2AP/PV	R47L/Y51F/A191T/N239H/I259V/A276T/P329V/A330P/L353I
RT2AP/V78I	R47L/Y51F/V78I/A191T/N239H/I259V/A276T/A330P/L353I
RT2F81W	R47L/Y51F/F81W/A191T/N239H/I259V/A276T/L353I
RT2IP	R47L/Y51F/A191T/N239H/I259V/A276T/L353I/I401P
RT2IP/E267V	R47L/Y51F/A191T/N239H/I259V/E267V/A276T/L353I/I401P
RT2IP/V78I	R47L/Y51F/V78I/A191T/N239H/I259V/A276T/L353I/I401P
T88M/I401P	T88M/I401P
WT	Wild Type

Appendix 3 Screening results of drug metabolism in chapter 3

Table S3.1. Product distribution and substrate conversion of the screen with diclofenac



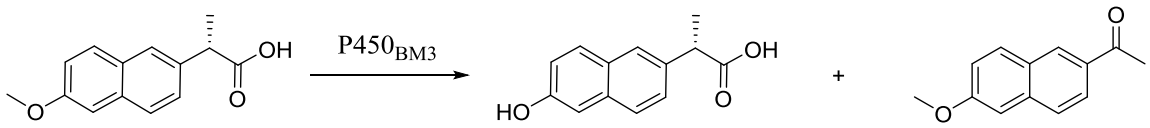
Diclofenac 4'-hydroxydiclofenac, **3.01** 5-hydroxydiclofenac, **3.02**

Number	Enzyme	Products [a]				Conversion [b]
		3.01	**3.02**	**X**	**Y**	
1	RP/FV/EV/F81W	31%	60%	9%	-	100%
2	RP/FV/EV	35%	50%	15%	-	91%
3	RP/IA/EV	-	93%	-	7%	35%
4	RP/FV/EV/V78F	23%	47%	15%	15%	22%
5	RT2/F81W	28%	35%	16%	21%	18%
6	RP/EV	-	64%	-	36%	15%
7	F87A/IP	26%	8%	15%	51%	12%
8	KT2/E267F	-	79%	-	21%	10%
9	QP/A82M	29%	-	26%	45%	8%
10	RP/FV	13%	26%	17%	44%	5%
11	KSK19/QP/F87V	20%	63%	17%	-	5%

[a] Diclofenac metabolism was detected using GC and TLC whereas the quantification was based upon peak integration of the 1H NMR spectra of crude extracts of 10 mL reactions. The other products **X** and **Y** were not formed in sufficient quantities to be characterised.

[b] Results represent averages of three independent experiments with errors not higher than 10%. Only mutants with >5% conversion are listed.

Table S3.2 Product distribution and substrate conversion of the screen with naproxen



Naproxen 6-desmethylnaproxen, **3.03** 2-acetyl-6-methoxynaphthalene, **3.04**

Number	Enzymes	Retention time /min [a]				Conversion [b]
		4.71 (3.04)	4.91 (3.03)	6.67	7.08	
1	RT2AP/F81W	68%	32%	-	-	100%
2	RT2AP/PV	34%	66%	-	-	96%
3	RPAP/A328V	63%	37%	-	-	96%
4	RLYF/AM/IA/QP	70%	30%	-	-	85%
5	RP/FW	66%	34%	-	-	68%
6	KSK19/AI/IP	66%	34%	-	-	67%
7	RT2AP/AI/S72A	62%	38%	-	-	66%
8	KSK19	29%	67%	5%	-	61%
9	RP/FV/EV	60%	40%	-	-	58%
10	RT2/A330W	52%	48%	-	-	43%
11	RPAP/A74W	57%	43%	-	-	42%
12	KSK19/AI/Y51V	50%	45%	5%	-	41%
13	RT2/FW/AI	66%	34%	-	-	41%
14	KSK19/AI/FV	67%	33%	-	-	39%
15	KT2	18%	82%	-	-	28%
16	RT2IP/V78I	56%	44%	-	-	27%
17	GVQ	21%	79%	-	-	26%
18	RP/FV/EV/AI	60%	40%	-	-	16%
19	RT2/AP/AM/AI	73%	27%	-	-	16%
20	RT2/AP	-	21%	40%	39%	15%
21	KSK19/I263A	14%	86%	-	-	11%
22	RLYF/KSK19/AI	56%	44%	-	-	11%
23	F87A	48%	52%	-	-	11%
24	KSK19/AI/AI	15%	75%	-	11%	10%
25	QP/AM	55%	45%	-	-	10%

26	RP/AM/IA	56%	44%	-	-	7%
27	KSK19/QP	51%	49%	-	-	6%
28	WT	20%	80%	-	-	5%
29	RP/FV	-	-	67%	33%	4%
30	KSK19/AI/F42L	-	100%	-	-	4%
31	RLYF/KSK19	-	100%	-	-	4%
32	RT2	16%	46%	38%	-	4%
33	RT2IP	35%	65%	-	-	4%
34	KSK19/AI/VI	43%	57%	-	-	3%
35	RT2AP/AI	55%	45%	-	-	3%
36	RP	42%	58%	-	-	3%
37	RPAP/AM	-	-	100%	-	2%
38	RT2/AP/AM	-	100%	-	-	1%
39	RT2/AP/V78I	-	100%	-	-	1%
40	KSK19/AM	-	-	-	-	0%
41	KSK19/AI	-	-	-	-	0%
42	KSK19/AI/YL	-	-	-	-	0%
43	RP/EV	-	-	-	-	0%
44	RPAP	-	-	-	-	0%

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 8.0, 2 μM P450, 1 mM naproxen. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. Naproxen metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

4.71 min product: 2-acetyl-6-methoxynaphthalene **3.04** (characterised by NMR and MS)

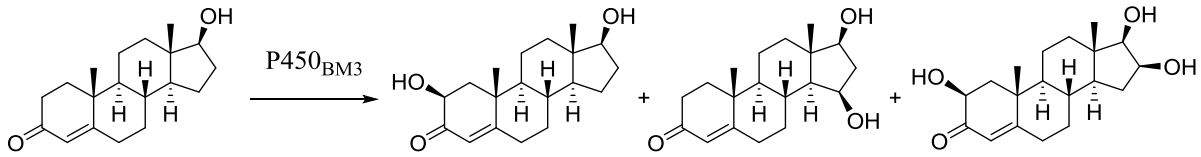
4.91 min product: 6-desmethylnaproxen **3.03** (characterised by NMR and MS)

6.67 min product: unidentified

7.08 min product: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S3.3 Product distribution and substrate conversion of the screen with testosterone



Testosterone | 2β-hydroxy testosterone, **3.05** | 15β-hydroxy testosterone, **3.06** | 2β, 16β-dihydroxy testosterone, **3.07**

No	Enzyme	Retention time /min [a]						Conversion [b]
		9.71	**9.84**	**10.29**	**10.45**	**10.83**	**10.95**	
		2*β*-OH (3.05)	T-OH	T-OH	**15*β*-OH (3.06)**	**2*β*,16*β*-$(OH)_2$ (3.07)**	T-$(OH)_2$	
1	KSK19/QP	38%	2%	2%	43%	16%	-	100%
2	RLYF/KSK19	61%	3%	-	33%	3%	-	100%
3	KSK19/AM	10%	2%	-	86%	2%	-	98%
4	RLYF/KSK19/IP	12%	-	-	14%	61%	12%	91%
5	KSK19	-	-	-	60%	40%	-	90%
6	RT2/V78F	5%	-	-	95%	-	-	89%
7	KSK19/QP/FV	4%	-	-	96%	-	-	83%
8	RLYF/FA/L188V	55%	3%	-	42%	-	-	76%
9	F87A	60%	3%	-	37%	-	-	64%
10	KSK19/IA/QP	15%	-	-	82%	4%	-	61%
11	RLYF/FA/AG	38%	2%	-	60%	-	-	55%
12	KSK19/FS	40%	2%	-	58%	-	-	33%
13	RP/FV/EV	28%	5%	11%	56%	-	-	28%
14	KSK19/TA	43%	-	-	57%	-	-	24%
15	KSK19/IA	10%	-	-	90%	-	-	22%
16	RT2/IP/A184I	48%	10%	12%	30%	-	-	20%
17	RP/FV/EV/A184I	80%	-	-	13%	7%	-	20%
18	RT2/AP/A82L	23%	-	-	77%	-	-	13%
19	RP/AM	20%	-	-	80%	-	-	12%
20	RT2	-	-	-	100%	-	-	11%
21	RLYF/LA/FA	46%	-	-	54%	-	-	9%
22	RP/H171L	45%	6%	8%	41%	-	-	9%

23	RT2/IP/EV	59%	32%	9%	-	-	-	8%
24	RP/FV/EV/V78F	44%	-	-	56%	-	-	7%
25	RT2/AP/FW	-	-	-	100%	-	-	6%
26	RP/FV/EV/L188Q	72%	-	-	28%	-	-	6%
27	RP/FA	47%	-	-	53%	-	-	6%
28	RT2/AP/AV	56%	-	-	44%	-	-	5%
29	QP/AM	43%	-	-	57%	-	-	5%
30	RT2/AW/AM	70%	-	-	30%	-	-	4%
31	KSK19/AI/AI	-	-	-	100%	-	-	4%
32	RPAP/AV	51%	-	-	49%	-	-	3%
33	KT2	100%	-	-	-	-	-	2%
34	RP/IA/EV	-	-	-	100%	-	-	1%
35	KSK19/AI/FV	-	-	-	100%	-	-	1%
36	RP/EV	100%	-	-	-	-	-	1%
37	AP/QP	-	-	-	-	-	-	0%
38	AS/PS/QP	-	-	-	-	-	-	0%
39	AS/PS/FA/IP	-	-	-	-	-	-	0%
40	FA/TA/IP	-	-	-	-	-	-	0%
41	FS/IP	-	-	-	-	-	-	0%
42	IP/QP	-	-	-	-	-	-	0%
43	KSK19/AI/IA	-	-	-	-	-	-	0%
44	KSK19/AI	-	-	-	-	-	-	0%
45	KSK19/AI/FL	-	-	-	-	-	-	0%
46	KSK19/AI/IP	-	-	-	-	-	-	0%
47	KSK19/AI/LA	-	-	-	-	-	-	0%
48	KSK19/AI/RL	-	-	-	-	-	-	0%
49	RLYF/KSK19/AI	-	-	-	-	-	-	0%
50	KSK19/AI/SA	-	-	-	-	-	-	0%
51	KSK19/AI/VI	-	-	-	-	-	-	0%
52	KSK19/AI/YF	-	-	-	-	-	-	0%
53	KSK19/AI/YV	-	-	-	-	-	-	0%
54	KSK43	-	-	-	-	-	-	0%
55	KT2/FL	-	-	-	-	-	-	0%

56	QP	-	-	-	-	-	-	0%
57	RLYF	-	-	-	-	-	-	0%
58	RLYF/IA	-	-	-	-	-	-	0%
59	RLYF/LV	-	-	-	-	-	-	0%
60	RLYF/LA	-	-	-	-	-	-	0%
61	RP	-	-	-	-	-	-	0%
62	RP/FW	-	-	-	-	-	-	0%
63	RP/IA/AM	-	-	-	-	-	-	0%
64	RPAP	-	-	-	-	-	-	0%
65	RPAP/A74W	-	-	-	-	-	-	0%
66	RPAP/SL	-	-	-	-	-	-	0%
67	RPAP/AM	-	-	-	-	-	-	0%
68	RPAP/AM/AI	-	-	-	-	-	-	0%
69	RPAP/PV	-	-	-	-	-	-	0%
70	RP/FV	-	-	-	-	-	-	0%
71	RP/FV/EV/A328I	-	-	-	-	-	-	0%
72	RQA	-	-	-	-	-	-	0%
73	RT2/A330W	-	-	-	-	-	-	0%
74	RT2/AP	-	-	-	-	-	-	0%
75	RT2/AP/VI/AI	-	-	-	-	-	-	0%
76	RT2/AP/AI	-	-	-	-	-	-	0%
77	RT2/AP/AI/FW	-	-	-	-	-	-	0%
78	RT2/AP/AI/SA	-	-	-	-	-	-	0%
79	RT2/AP/AM	-	-	-	-	-	-	0%
80	RT2/AP/AM/AI	-	-	-	-	-	-	0%
81	RT2/AP/VI	-	-	-	-	-	-	0%
82	RT2/FW/AI	-	-	-	-	-	-	0%
83	RT2/IP	-	-	-	-	-	-	0%
84	RT2/IP/FW	-	-	-	-	-	-	0%
85	RT2/IP/AP	-	-	-	-	-	-	0%
86	TM/IP	-	-	-	-	-	-	0%
87	WT	-	-	-	-	-	-	0%

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 1 mM testosterone. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*.

Testosterone metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. The product quantification was based on the integration of peak areas. The GC retention times represent the following products.

9.71 min product: 2β-hydroxytestosterone **3.05** (characterised by NMR and MS)

9.84 min product: monohydroxylation of testosterone, T-OH (MS data)

10.29 min product: monohydroxylation of testosterone, T-OH (MS data)

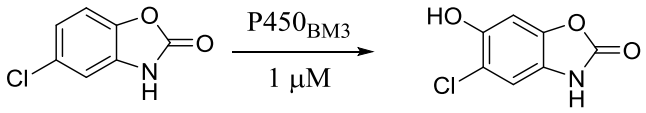
10.45 min product: 15β-hydroxytestosterone **3.06** (characterised by NMR and MS)

10.83 min product: 2β,16β-hydroxytestosterone **3.07** (characterised by NMR and MS)

10.95 min product: dihydroxylation of testosterone, $T(OH)_2$ (MS data)

[b] Conversion represents the average of 2 independent experiments with errors not higher than 10%.

Table S3.4 Substrate conversion of chlorzoxazone



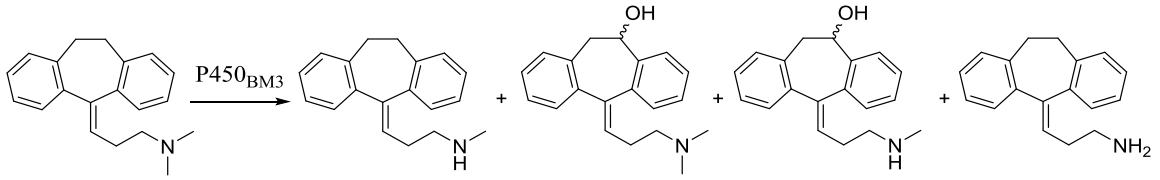
Chlorzoxazone 6-hydroxychlorzoxazone, **3.08**

Number	Enzyme	Conversion [a]
1	RT2/AP/AI	100%
2	RT2/AW/AM	97%
3	RPAP	89%
4	RT2/AP/AL	69%
5	RP/IA/EV	53%
6	RT2	51%
7	RPAP/PV	41%
8	RP/FV	33%
9	RT2/IP/AP	28%
10	KSK19/AI/IP	28%
11	RP	23%
12	RT2/AP/F81W	21%
13	RT2/AP/PV	15%
14	QP	12%
15	RPAP/AM	9%
16	RP/EV	8%
17	KSK19/AM	5%
18	A330P	3%
19	KSK19/AI/FL	0%
20	RLYF/KSK19/AI	0%
21	KSK19/AI/YL	0%
22	KSK19/AI	0%
23	KSK19/AI/AI	0%
24	KSK19/IA	0%
25	RP/FV/EV/A184I	0%
26	RP/FV/EV	0%
27	KSK19/AI/FV	0%

28	KSK19/AI/VI	0%
29	RT2/AP/AM	0%
30	RT2/IP	0%
31	RT2/AP	0%
32	RT2/AP/VI	0%
33	WT	0%
34	R19	0%
35	KSK19/AI/YV	0%
36	RT2/AP/AI/SA	0%
37	FA/IP	0%
38	RT2/FW/AI	0%
39	RPAP/A74W	0%
40	RT2/FW	0%
41	RQA	0%
42	KSK19	0%
43	FP/IP	0%
44	GVQ	0%
45	KSK19/QP	0%
46	FA	0%
47	TM/IP	0%
48	RP/AM/IA	0%
49	KT2	0%
50	RLYF/KSK19	0%
51	RT2/AP/AM/AI	0%
52	KSK19/AI/YF	0%

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 1 mM chlorzoxazone. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. Chlorzoxazone metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. Product quantification was based on the integration of peak areas. All active mutants gave 6-hydroxychlorzoxazone **3.08** as the only product. Conversion represents the average of 2 independent experiments with errors not higher than 10%.

Table S3.5 Product distribution and substrate conversion of the screen with amitriptyline



Amitriptyline | nortriptyline **3.09** | hydroxyamitriptyline **3.10** | *Z*-10-hydroxynortriptyline **3.11** | desmethylnortriptyline **3.12**

Number	Enzyme	Retention Time /min [a]					Conversion [b]
		6.68 (3.12)	6.91	7.09 (3.09)	7.90 (3.10)	7.96 (3.11)	
1	RT2/AP/VI	1%	3%	44%	50%	5%	79%
2	KSK19	3%	3%	82%	5%	7%	75%
3	RP	-	10%	82%	6%	3%	65%
4	RT2/AP/AI	-	5%	62%	30%	4%	55%
5	RP/FV	7%	4%	83%	6%	-	46%
6	RT2/IP/AP	-	6%	89%	4%	1%	40%
7	RT2	2%	5%	87%	4%	2%	43%
8	RT2/AP/AL	-	-	77%	16%	7%	38%
9	RP/FW	2%	5%	86%	5%	3%	33%
10	KSK19/IA	-	2%	98%	-	-	32%
11	RP/FV/EV	5%	7%	83%	4%	1%	32%
12	RT2/FW/AI	-	-	95%	5%	-	29%
13	RP/FV/EV/A184I	-	3%	97%	-	-	29%
14	RP/EV	-	4%	89%	5%	2%	27%
15	KSK19/AI/YL	1%	4%	89%	5%	1%	27%
16	RT2/FW	-	3%	94%	2%	-	27%
17	RT2/AW/AM	-	-	67%	25%	8%	26%
18	RT2/IP	-	5%	`	3%	-	26%
19	RP/AM	-	-	90%	6%	3%	25%
20	QP	2%	4%	87%	7%	-	22%
21	RT2/AP/AM/AI	-	2%	84%	12%	3%	21%
22	RT2/AP/PV	24%	-	64%	-	12%	21%
23	RP/IA/EV	-	13%	87%	-	-	21%
24	RT2/IP/AI	2%	4%	89%	6%	-	19%

25	KT2	-	-	100%	-	-	18%
26	RLYF/KSK19	-	-	100%	-	-	17%
27	RT2/AW	2%	3%	91%	4%	-	17%
28	RP/FV/EV/VF	-	6%	91%	4%	-	16%
29	FA	-	-	100%	-	-	16%
30	KSK19/QP/FV	-	-	100%	-	-	15%
31	KSK19/AM	6%	-	82%	9%	3%	14%
32	RT2IP/EV	-	7%	89%	3%	-	14%
33	RP/HL	-	3%	91%	6%	-	13%
34	KSK19/QP	-	-	100%	-	-	12%
35	RLYF/KSK19/IP	-	-	100%	-	-	12%
36	KSK19/AI/FV	18%	-	70%	11%	-	12%
37	GVQ	4%	-	96%	-	-	11%
38	KSK19/AI/AI	-	3%	92%	3%	2%	11%
39	KSK19/IA/QP	-	-	100%	-	-	11%
40	RT2/IP/FW	-	-	100%	-	-	11%
41	RT2/AP/VI/AI	4%	-	66%	30%	-	10%
42	RPAP/AV	61%	-	39%	-	-	10%
43	WT	34%	-	66%	-	-	9%
44	KSK19/AI/FL	-	-	90%	8%	2%	9%
45	RQA	7%	-	93%	-	-	9%
46	RP/AM/IA	-	-	100%	-	-	8%
47	RT2/AP/AM	45%	-	55%	-	-	7%
48	QP/AM	-	-	100%	-	-	7%
49	RT2/AP/AI/SA	75%	-	25%	-	-	7%
50	RT2/AP/AI/FW	-	-	87%	13%	-	7%
51	KSK19/AI/YV	10%	-	90%	-	-	6%
52	RPAP	47%	-	53%	-	-	6%
53	RP/FV/EV/A328I	-	-	100%	-	-	6%
54	KSK19/AI/IP	6%	-	94%	-	-	6%
55	RLYF/KSK19/AI	-	-	94%	6%	-	5%
56	RPAP/AM	37%	-	63%	-	-	5%
57	RP/FV/EV/LQ	-	13%	87%	-	-	5%
58	TM/IP	-	-	100%	-	-	5%

59	RPAP/AM/AI	-	-	100%	-	-	5%
60	KSK19/AI	3%	-	86%	8%	3%	4%
61	KSK19/FS	-	-	100%	-	-	4%
62	FA/TA/IP	-	-	100%	-	-	4%
63	AS/PS/QP	-	-	83%	17%	-	4%
64	KSK19/TA	-	-	100%	-	-	3%
65	KSK19/AI/IA	-	-	100%	-	-	3%

[a] The screening reaction mixture (0.5 mL) contained 50 mM Tris–HCL, pH 8.5, 1 μM P450, 1 mM amitriptyline. After 16 h the reaction mixture was adjusted to pH>12 with 5 M KOH and extracted with diethyl ether (300 μl). The phases were separated by centrifugation at 21,000 *g*. Amitriptyline metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. Product quantification was based on the integration of peak areas. The GC retention times represent for the following products:

6.68 min product: desmethylnortriptyline *via* double demethylation of amitriptyline (determined by MS)

6.91 min product: unidentified

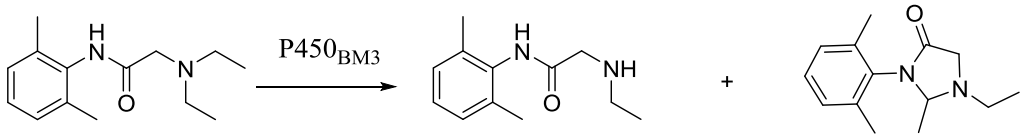
7.09 min product: nortriptyline **3.09** (characterised by NMR and MS)

7.90 min product: mono-hydroxylation of amitriptyline **3.10** (characterised by NMR and MS)

7.96 min product: mono-hydroxylation of nortriptyline **3.11** (MS data)

[b] Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

Table S3.6 Product distribution and substrate conversion of the screen with lidocaine



Lidocaine | monoethylglycinexylidide MEGX, **3.13** | *Z*-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline, **3.14**

Number	Enzyme	Retention time /min [a]			Conversion [b]
		4.29 (3.13)	4.33	4.48 (3.14)	
1	RT2/AW	-	-	100%	100%
2	KSK19/IA	3%	1%	96%	99%
3	RT2/IP	7%	-	93%	97%
4	QP/AM	-	-	100%	95%
5	R19	-	-	100%	95%
6	RP/FV/EV/F81W	96%	4%	-	93%
7	RP/FV/EV	11%	5%	84%	92%
8	RP/AP/AM	-	-	100%	92%
9	RLYF/KSK19	4%	2%	94%	90%
10	RT2/AP	-	-	100%	89%
11	RT2	-	-	100%	87%
12	KSK19/QP	4%	2%	94%	84%
13	RT2AP/AL	-	-	100%	84%
14	RP/EV	-	-	100%	81%
15	RP/FV	-	-	100%	75%
16	RP/FV/EV/A184I	14%	9%	77%	68%
17	KSK19/AM	-	-	100%	68%
18	RLYF/KSK19/AI	-	-	100%	68%
19	FA/IP	-	-	100%	62%
20	KSK19/AI/FV	-	-	100%	61%
21	KSK19/AI/YL	-	-	100%	61%
22	KSK19/AI/YV	-	-	100%	59%
23	KSK19/AI/AI	-	-	100%	55%
24	RT2/AP/AI	31%	-	69%	49%
25	RT2/AP/AM	-	-	100%	46%

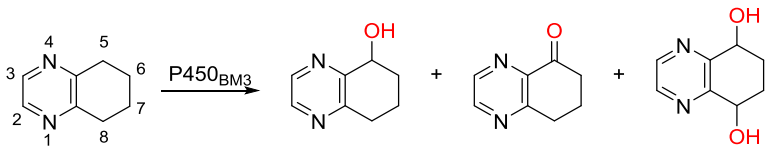
26	RT2/AP/AI/SA	-	-	100%	46%
27	KSK19/AI/IP	-	-	100%	46%
28	KT2	-	-	100%	43%
29	FA	-	-	100%	40%
30	KSK19/AI/FL	-	-	100%	38%
31	KSK19/AI	-	-	100%	31%
32	KSK19/AI/YF	-	-	100%	28%
33	KSK19	75%	-	25%	27%
34	RT2/FW/AI	-	-	100%	26%
35	RT2/AP/VI	-	-	100%	25%
36	QP	-	-	100%	24%
37	RP/IA/EV	5%	11%	84%	23%
38	RT2/AP/FW	100%	-	-	23%
39	GVQ	-	-	100%	14%
40	RP	13%	10%	77%	13%
41	A330P	-	-	100%	11%
42	RP/AP/PV	-	-	100%	9%
43	RP/AM/IA	-	-	100%	9%
44	TM/IP	-	-	100%	5%
45	RPAP/AW	-	-	100%	3%

[a] The screening reaction mixture (0.5 mL) contained 50 mM Tris–HCl, pH 8.5, 1 μM P450, 2 mM lidocaine. After 16 h the reaction mixture was adjusted to pH>12 with 5 M KOH and extracted with diethyl ether (300 μl). The phases were separated by centrifugation at 21,000 *g*. Lidocaine metabolites were identified using GC. The initial oven temperature was 180 ℃ which was held for 2 min, the temperature was then raised at 15 ℃/min to 300 ℃ and kept at 300 ℃ for 5 min. Product quantification was based on the integration of peak areas. The GC retention times represent for the following products:
4.29 min product: monoethylglycinexylidide (MEGX) **3.13** from demethylation of lidocaine (characterised by NMR and MS)
4.33 min product: glycinexylidide (GX) from the double demethylation of lidocaine (MS data)
4.48 min product: *N*,*N*-acetal compound (*Z*)-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline, **3.14** (characterised by NMR and MS)

[b] Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

Appendix 4 Screening results of drug fragment oxidation in chapter 4

Table S4.1 Substrate conversion and product distribution of 5,6,7,8-tetrahydroquinoxaline (THQ)



4.01 **4.02** **4.03** **4.04**

No	Enzyme	Conversion	4.02 (5.67 min)	4.04 (6.78 min)	4.03 (5.73 min)	Product 4 (6.82 min)	No of products
1	KSK19 A82M	100.0%	61.6%	35.5%	2.9%	0.0%	3
2	RPAPAM	100.0%	100.0%	0.0%	0.0%	0.0%	2
3	KSK19AI F87V	100.0%	92.7%	4.4%	2.9%	0.0%	1
4	RT2APAM	100.0%	100.0%	0.0%	0.0%	0.0%	1
5	RT2IP	100.0%	100.0%	0.0%	0.0%	0.0%	1
6	RT2AP	100.0%	100.0%	0.0%	0.0%	0.0%	2
7	RT2APA184I	100.0%	97.6%	2.4%	0.0%	0.0%	2
8	RT2	100.0%	100.0%	0.0%	0.0%	0.0%	1
9	RT2AP A82L	100.0%	100.0%	0.0%	0.0%	0.0%	1
10	RP/QP	100.0%	90.0%	0.0%	0.0%	10.0%	2
11	RPIAEV	98.6%	100.0%	0.0%	0.0%	0.0%	1
12	RPAP	98.2%	95.6%	0.0%	0.0%	4.4%	2
13	KSK19AI Y51L	97.1%	90.9%	9.1%	0.0%	0.0%	2
14	RLYF KSK19 AI	96.8%	90.7%	9.3%	0.0%	0.0%	2
15	RP EV	96.2%	100.0%	0.0%	0.0%	0.0%	1
16	KSK19AIAI	93.9%	79.8%	20.2%	0.0%	0.0%	2

17	KSK19AI	85.6%	96.6%	3.4%	0.0%	0.0%	2
18	KSK19AI F42L	84.5%	96.6%	3.4%	0.0%	0.0%	2
19	RLYF/KSK19	82.5%	100.0%	0.0%	0.0%	0.0%	1
20	RP FV	79.9%	100.0%	0.0%	0.0%	0.0%	1
21	RT2APV78I	75.5%	100.0%	0.0%	0.0%	0.0%	1
22	RPFVEV L437VL	73.3%	98.3%	1.7%	0.0%	0.0%	2
23	RPFVEV	61.9%	95.8%	4.2%	0.0%	0.0%	3
24	RT2A330W/S72L	49.6%	100.0%	0.0%	0.0%	0.0%	1
25	RP	25.5%	100.0%	0.0%	0.0%	0.0%	1
26	RPFVEV L437LV	15.3%	100.0%	0.0%	0.0%	0.0%	1
27	RT2AP/I401M	11.9%	100.0%	0.0%	0.0%	0.0%	1
28	KSK19AI V78I	11.9%	100.0%	0.0%	0.0%	0.0%	1
29	RT2A330W/S72R	11.1%	100.0%	0.0%	0.0%	0.0%	1
30	WT	1.4%	100.0%	0.0%	0.0%	0.0%	1

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 8.0, 1 μM P450, 2 mM 5,6,7,8-tetrahydroquinoxaline. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organic was analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

5.67 min product: THQ-5-ol (**4.02**)

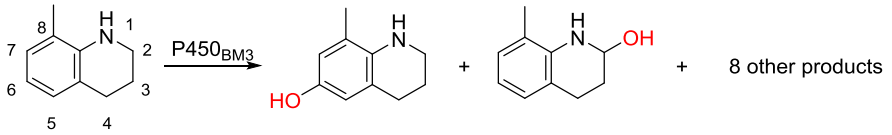
5.73 min product: THQ-5-one (**4.03**)

6.78 min product: THQ-5,8-diol (**4.04**)

6.82 min product: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.2 Substrate conversion and product distribution of 8-methyl-1,2,3,4-tetrahydroquinoline **(4.05)**



4.05 **4.06** **4.07**

No	Enzyme	Conversion	4.06 (5.59 min)	Product 2 (7.46 min)	4.07 (7.73 min)	Product 4 (7.92 min)	Product 5 (8.15 min)	Product 6 (8.52 min)	Product 7 (8.78 min)	Product 8 (6.21 min)	Product 9 (8.35 min)	Product 10 (7.69 min)	No. of products
1	KSK19AI F42L	100.0%	26.0%	2.0%	51.9%	1.5%	2.2%	14.3%	0.6%	1.5%	0.0%	0.0%	8
2	RLYF KSK19 AI	100.0%	30.3%	2.4%	46.3%	2.5%	4.1%	13.7%	0.7%	0.0%	0.0%	0.0%	7
3	RPFVEV L437LV	100.0%	60.0%	0.0%	0.0%	28.9%	1.9%	6.4%	0.0%	0.0%	2.8%	0.0%	5
4	RT2APA184I	99.3%	2.7%	39.2%	1.6%	1.0%	27.3%	28.2%	0.0%	0.0%	0.0%	0.0%	6
5	KSK19AI Y51L	94.7%	71.1%	0.0%	14.8%	2.3%	1.9%	9.8%	0.0%	0.0%	0.0%	0.0%	5
6	KSK19AI	86.5%	72.1%	0.0%	0.0%	1.6%	2.1%	24.2%	0.0%	0.0%	0.0%	0.0%	4
7	RT2AP	85.9%	23.8%	28.1%	4.0%	1.0%	36.9%	5.7%	0.0%	0.0%	0.5%	0.0%	7
8	RPAPAM	69.1%	19.6%	0.0%	0.0%	0.0%	17.7%	57.8%	0.0%	0.0%	0.0%	4.9%	4
9	RT2APAM	65.9%	20.3%	1.5%	0.0%	0.0%	34.3%	39.7%	0.0%	0.0%	0.0%	4.2%	5
10	RPFVEV A184I	63.2%	94.9%	0.0%	0.0%	3.7%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	3
11	RPFVEV L437VL	61.9%	34.8%	17.1%	3.3%	1.3%	28.9%	14.6%	0.0%	0.0%	0.0%	0.0%	6
12	RPAP	55.7%	53.5%	0.0%	0.0%	1.0%	22.7%	21.6%	0.0%	0.0%	0.0%	1.2%	5
13	RT2AP/I401M	55.4%	31.2%	0.0%	0.0%	1.4%	19.2%	46.3%	0.0%	0.0%	0.0%	1.9%	5
14	KSK19AI V78I	47.0%	81.2%	0.0%	0.0%	1.5%	0.0%	17.3%	0.0%	0.0%	0.0%	0.0%	3
15	RPFVEV	30.3%	33.6%	1.4%	2.6%	0.7%	10.2%	50.4%	0.0%	0.0%	1.0%	0.0%	7
16	KSK19 A82M	28.0%	58.8%	0.0%	0.0%	11.2%	3.0%	1.5%	0.0%	0.0%	24.5%	1.1%	6
17	RP	26.8%	18.6%	1.8%	3.2%	1.4%	9.6%	63.3%	0.0%	0.0%	2.3%	0.0%	7
18	KSK19AI F87V	23.2%	61.5%	1.4%	6.3%	0.8%	8.8%	21.1%	0.0%	0.0%	0.0%	0.0%	6
19	KSK19 I263A	20.2%	84.7%	0.0%	0.0%	2.8%	0.0%	9.3%	0.0%	0.0%	3.1%	0.0%	4
20	RT2	15.4%	33.6%	10.1%	6.5%	8.6%	14.4%	26.8%	0.0%	0.0%	0.0%	0.0%	5

21	RT2A330W/S72R	9.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
22	WT	3.7%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 8.0, 1 μM P450, 2 mM 8-methyl-1,2,3,4-tetrahydroquionolin. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

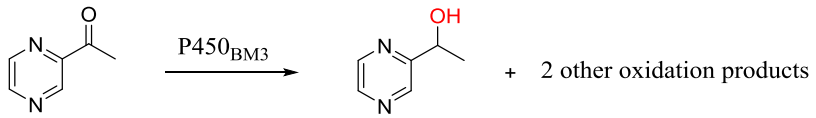
5.59 min product: 8-methyl-1,2,3,4-tetrahydroquionolin-6-ol (**4.06**)

7.73 min product: 8-methyl-1,2,3,4-tetrahydroquinolin-2-ol (**4.07**)

Other products: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.3 Substrate conversion and product distribution of acetylpyrazine **(4.08)**



4.08 **4.09**

No	Enzyme	Conversion	4.09 (2.01 min)	Product 2 (2.35 min)	Product 3 (3.12 min)	No. of product
1	KSK19AI	49.5%	75.4%	0.0%	24.6%	1
2	RPFVEVL437LVAI	44.4%	29.3%	0.0%	70.7%	1
3	KSK19AIAI	38.3%	87.7%	0.0%	12.3%	2
4	RPFVEV	34.9%	88.7%	0.0%	11.3%	2
5	RPFVEVL437LV	33.1%	41.2%	0.0%	58.8%	2
6	RT2APAI	32.6%	89.7%	0.0%	10.3%	2
7	RT2IP	30.1%	97.7%	0.0%	2.3%	2
8	RPFVEVL437VL	23.3%	12.4%	0.0%	87.6%	2
9	RPAPAM	21.2%	96.4%	0.0%	3.6%	2
10	RPAPSL	20.5%	31.7%	0.0%	68.3%	2
11	KSK19AIRL	17.3%	89.0%	0.0%	11.0%	2
12	RLYFKSK19AI	16.0%	95.7%	0.0%	4.3%	2
13	RT2AP	13.7%	88.3%	0.0%	11.7%	2
14	KSK19AM	12.9%	55.8%	40.4%	3.8%	3
15	GVQ	11.7%	77.7%	17.3%	5.1%	3
16	KSK19AISA	10.9%	87.0%	0.0%	13.0%	2
17	RT2APVI	10.2%	85.3%	0.0%	14.7%	2
18	RPEV	9.6%	64.5%	0.0%	35.5%	2
19	RT2AIL437LV	8.6%	80.0%	0.0%	20.0%	2
20	KSK19AIFL	8.4%	73.0%	0.0%	27.0%	2
21	YLKSK19AI	8.3%	77.8%	0.0%	22.2%	2

22	RT2VI	7.1%	32.2%	59.3%	8.4%	3
23	RPFA	7.0%	58.3%	27.8%	13.9%	3
24	RT2APAM	6.9%	78.6%	0.0%	21.4%	2
25	RPAP	6.2%	83.1%	0.0%	16.9%	2
26	RPFV	5.6%	48.5%	0.0%	51.5%	2
27	KSK19AIVI	5.4%	81.3%	0.0%	18.7%	2
28	RT2AIL437VL	4.0%	8.0%	0.0%	92.0%	2
29	RT2IPAP	4.0%	87.5%	0.0%	12.5%	2
30	RPAPPV	3.9%	45.2%	0.0%	54.8%	2
31	KSK19IA	3.9%	28.9%	45.4%	25.6%	3
32	RP	3.2%	45.7%	0.0%	54.3%	2
33	WT	2.7%	32.1%	0.0%	67.9%	2

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM acetylpyrazine. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 40 ℃ which was held for 2 min, the temperature was then raised at 8 ℃/min to 120 ℃. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

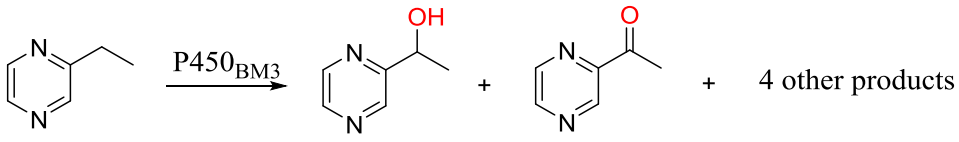
2.01 min product: 1-(pyrazin-2-yl)ethan-1-ol (**4.09**)

2.35 min product: unidentified

3.12 min product: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.4 Substrate conversion and product distribution of ethylpyrazine **(4.10)**



4.10 **4.09** **4.08**

No	Enzyme	Conversion	4.09 (2.01 min)	4.08 (2.35 min)	Product 3 (2.76 min)	Product 4 (2.94 min)	Product 5 (3.12 min)	Product 6 (3.51 min)	No. of products
1	RT2IP	100.0%	15.0%	0.0%	83.2%	0.0%	1.8%	0.0%	3
2	RT2APAI	81.1%	26.5%	0.0%	5.2%	2.0%	58.4%	7.8%	5
3	RPAPAM	62.4%	53.2%	0.0%	5.4%	0.0%	41.4%	0.0%	3
4	RPFVEVL437VL	21.8%	86.0%	0.0%	0.0%	0.0%	14.0%	0.0%	2
5	RT2APVI	17.9%	26.1%	0.0%	0.0%	0.0%	48.8%	25.1%	3
6	RT2AP	16.7%	23.5%	0.0%	0.0%	4.3%	72.2%	0.0%	3
7	KSK19AI	11.9%	24.5%	0.0%	0.0%	12.1%	16.2%	47.2%	4
8	RT2	11.2%	36.1%	0.0%	0.0%	0.0%	63.9%	0.0%	2
9	RPFW	8.8%	5.8%	0.0%	0.0%	11.3%	6.4%	76.5%	4
10	KSK19IA	8.1%	0.0%	40.4%	0.0%	14.0%	0.0%	45.6%	3
11	GVQ	6.8%	55.1%	44.9%	0.0%	0.0%	0.0%	0.0%	2
12	KSK19AIFV	5.2%	67.7%	0.0%	0.0%	0.0%	32.3%	0.0%	2
13	RPFVEV	5.1%	0.0%	0.0%	0.0%	26.2%	0.0%	73.8%	2
14	RT2APAV	5.0%	8.6%	0.0%	0.0%	17.6%	0.0%	73.7%	3
15	WT	4.8%	0.0%	0.0%	0.0%	16.6%	0.0%	83.4%	2
16	RPAP	4.2%	0.0%	0.0%	0.0%	15.7%	0.0%	84.3%	2
17	KSK19AM	3.9%	12.6%	87.4%	0.0%	0.0%	0.0%	0.0%	2
18	RT2APTL	3.7%	41.1%	0.0%	0.0%	23.6%	35.4%	0.0%	3
19	YFKSK19AI	3.6%	46.2%	0.0%	0.0%	0.0%	30.2%	23.6%	3
20	RT2APAM	3.3%	40.8%	0.0%	0.0%	0.0%	59.2%	0.0%	2

21	RPFVEVL437LV	2.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
22	RT2APAMSA	1.9%	32.8%	0.0%	0.0%	0.0%	0.0%	67.2%	2
23	KSK19AIVI	1.8%	45.8%	0.0%	0.0%	54.2%	0.0%	0.0%	2
24	RLYFKSK19AI	1.7%	0.0%	0.0%	0.0%	66.6%	0.0%	33.4%	2
25	RPFVEVFW	1.1%	0.0%	0.0%	0.0%	57.8%	0.0%	42.2%	2
26	RPAPSL	1.0%	51.0%	0.0%	0.0%	0.0%	49.0%	0.0%	2
27	RP	0.7%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
28	RT2IPAP	0.7%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
29	RPAPAV	0.7%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
30	RQA	0.5%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
31	RT2AW	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM ethylpyrazine. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography The initial oven temperature was 40 ℃ which was held for 2 min, the temperature was then raised at 8 ℃/min to 120 ℃. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

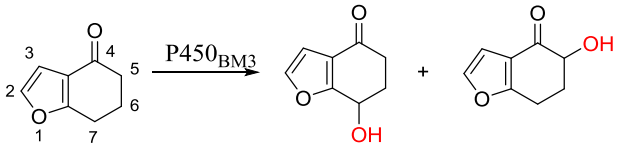
2.01 min product: 1-(pyrazin-2-yl)ethan-1-ol (**4.09**)

2.35 min product: acetylpyrazine (**4.08**)

Other products: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.5 Substrate conversion and product distribution of 6,7-dihydrobenzofuran-4(5*H*)-one **(4.11)**



4.11 **4.12** **4.13**

No	Enzyme	Conversion	Product 1 (5.26 min)	4.12 (6.08 min)	Product 3 (6.12 min)	4.13 (3.77 min)	No. of products
1	RT2IP	100.0%	0.0%	97.7%	2.3%	0.0%	2
2	RT2APA184I	100.0%	0.0%	9.3%	0.0%	90.7%	2
3	RT2AP A82L	100.0%	0.0%	34.9%	0.0%	65.1%	2
4	RP/QP	100.0%	0.0%	97.2%	0.0%	2.8%	2
5	KSK19 A82M	97.1%	2.6%	97.4%	0.0%	0.0%	2
6	RT2APAM	95.4%	0.0%	100.0%	0.0%	0.0%	1
7	R19	95.0%	0.0%	43.3%	0.0%	56.7%	2
8	RP EV	81.6%	2.4%	97.6%	0.0%	0.0%	2
9	RT2APV78I	81.1%	0.0%	100.0%	0.0%	0.0%	1
10	RPAP	73.4%	3.6%	96.4%	0.0%	0.0%	2
11	RLYF KSK19 AI	63.6%	3.5%	96.5%	0.0%	0.0%	2
12	KSK19AI Y51L	63.4%	1.6%	98.4%	0.0%	0.0%	2
13	RPFVEV L437VL	58.8%	0.0%	100.0%	0.0%	0.0%	1
14	KSK19AI	49.1%	0.0%	100.0%	0.0%	0.0%	1
15	KSK19AI F87V	47.8%	0.0%	79.9%	0.0%	20.1%	2
16	RT2A330W/S72L	44.0%	0.0%	85.9%	0.0%	14.1%	2
17	KSK19AIAI	38.6%	0.0%	100.0%	0.0%	0.0%	1
18	RPFVEV L437LV	37.0%	0.0%	100.0%	0.0%	0.0%	1
19	RP FV	36.4%	3.2%	90.7%	0.0%	6.2%	3
20	RP	23.1%	0.0%	100.0%	0.0%	0.0%	1

21	KSK19AI F42L	19.7%	0.0%	100.0%	0.0%	0.0%	1
22	RPFVEV	8.7%	0.0%	100.0%	0.0%	0.0%	1
23	RT2AP/I401M	8.3%	0.0%	100.0%	0.0%	0.0%	1
24	KSK19AI V78I	8.3%	0.0%	100.0%	0.0%	0.0%	1
25	RT2A330W/S72R	3.4%	0.0%	100.0%	0.0%	0.0%	1
26	RT2	3.2%	0.0%	100.0%	0.0%	0.0%	1
27	KSK19 I263A	2.4%	24.9%	48.4%	26.7%	0.0%	3
28	WT	1.1%	0.0%	100.0%	0.0%	0.0%	1

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM naproxen. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

5.26 min product: unidentified

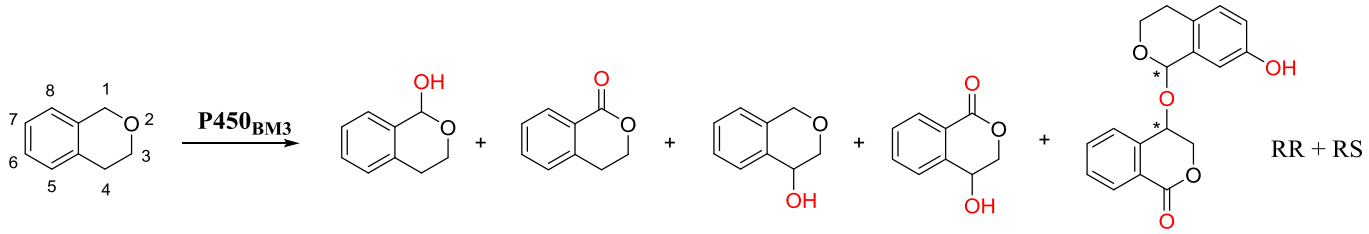
6.08 min product: 7-hydroxy-6,7-dihydrobenzofuran-4(5H)-one (**4.12**)

6.12 min product: unidentified

3.77 min product: 5-hydroxy-6,7-dihydrobenzofuran-4(5H)-one (**4.13**)

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.6 Substrate conversion and product distribution of isochroman **(4.14)**



4.14 **4.15** **4.16** **4.17** **4.18** **4.19** **4.20**

No	Enzyme	Conversion	4.15 (5.75 min)	4.16 (5.98 min)	4.17 (6.68 min)	4.18 (6.91 min)	4.19 (7.14 min)	4.20 (7.18 min)	Product 7 (7.94 min)	No. of products
1	RPAPAM	100.0%	37.2%	47.2%	1.1%	0.0%	12.0%	0.0%	2.6%	5
2	RPFVEV	100.0%	0.0%	42.6%	16.9%	0.0%	29.8%	10.7%	0.0%	4
3	RT2APAM	100.0%	36.8%	41.9%	1.2%	10.7%	7.4%	0.0%	1.9%	6
4	RT2AP	100.0%	7.1%	16.0%	2.7%	0.0%	41.7%	11.4%	21.0%	5
5	RT2	100.0%	6.0%	61.7%	1.2%	0.0%	25.4%	0.0%	5.8%	5
6	RPAP	100.0%	18.6%	70.6%	4.6%	0.6%	5.1%	0.0%	0.6%	6
7	RT2AP A82L	100.0%	20.8%	60.1%	1.7%	1.6%	12.1%	0.0%	3.7%	6
8	R19	100.0%	8.6%	89.0%	2.4%	0.0%	0.0%	0.0%	0.0%	3
9	RP/QP	100.0%	11.1%	45.0%	3.8%	0.0%	29.8%	0.0%	10.3%	5
10	RP FV	98.7%	15.0%	81.0%	2.4%	1.5%	0.0%	0.0%	0.0%	4
11	KSK19AI F87V	98.0%	4.6%	87.2%	4.9%	0.0%	2.4%	1.0%	0.0%	5
12	RT2IP	97.9%	9.3%	68.6%	0.3%	7.1%	14.7%	0.0%	0.0%	5
13	RPIAEV	96.2%	0.7%	6.1%	1.0%	2.5%	72.5%	4.4%	12.8%	7
14	RP	95.7%	14.9%	82.5%	2.6%	0.0%	0.0%	0.0%	0.0%	3
15	RT2APA184I	95.2%	10.4%	31.8%	1.4%	1.4%	46.6%	0.0%	8.3%	6
16	RT2APV78I	94.9%	26.7%	69.5%	3.8%	0.0%	0.0%	0.0%	0.0%	3
17	RT2A330W/S72L	90.6%	14.4%	82.4%	3.2%	0.0%	0.0%	0.0%	0.0%	3
18	RP EV	89.1%	12.3%	84.8%	2.8%	0.0%	0.0%	0.0%	0.0%	3

19	RPFVEV L437VL	57.7%	3.3%	89.2%	6.3%	0.0%	0.0%	1.2%	0.0%	4
20	RPFVEV A184I	48.4%	2.8%	89.5%	7.7%	0.0%	0.0%	0.0%	0.0%	3
21	RT2AP/I401M	41.6%	12.8%	78.3%	6.1%	0.0%	0.0%	2.8%	0.0%	4
22	WT	29.1%	9.8%	78.1%	8.0%	0.0%	0.0%	4.1%	0.0%	4
23	RPFVEV L437LV	19.6%	4.7%	55.5%	13.1%	19.5%	0.0%	7.2%	0.0%	5
24	KSK19 A82M	19.5%	3.2%	72.8%	13.1%	0.0%	0.0%	8.9%	2.0%	5
25	RT2A330W/S72R	17.9%	18.2%	33.5%	12.6%	27.9%	0.0%	7.7%	0.0%	5
26	KSK19AIAI	17.7%	0.0%	73.2%	15.7%	0.0%	0.0%	9.3%	1.8%	4
27	RLYF KSK19 AI	13.6%	2.1%	61.3%	20.4%	0.0%	0.0%	13.8%	2.4%	5
28	KSK19AI F42L	11.4%	0.0%	64.4%	20.8%	0.0%	0.0%	14.9%	0.0%	3
29	KSK19AI Y51L	10.3%	0.0%	57.1%	22.6%	0.0%	0.0%	17.4%	2.9%	4
30	KSK19AI V78I	9.8%	0.0%	42.8%	31.8%	0.0%	0.0%	21.1%	4.4%	4
31	KSK19AI	9.7%	0.0%	56.6%	22.5%	0.0%	0.0%	17.0%	3.8%	4
32	KSK19 I263A	7.9%	0.0%	43.4%	31.2%	0.0%	0.0%	21.6%	3.7%	4

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM isochroman. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

5.75 min product: isochroman-1-ol (**4.15**)

5.98 min product: isochroman-1-one (**4.16**)

6.68 min product: isochroman-4-ol (**4.17**)

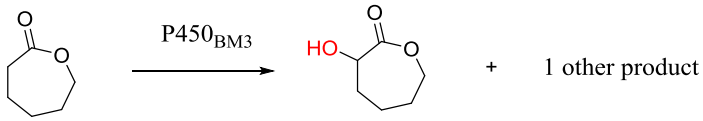
6.91 min product: isochroman-1-one-4-ol (**4.18**)

7.14 min, 7.18 min products: (*R*,*R*)- and (*R*,*S*)- 4-((7-hydroxyisochroman-1-yl)oxy)isochroman-1-one (**4.19**, **4.20**)

7.94 min product: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.7 Substrate conversion and product distribution of ε-Caprolactone **(4.21)**



4.21 **4.22**

No	Enzyme	Conversion	4.22 (4.15 min)	Product 2 (4.09 min)	No. of Products
1	RT2AP	37.8%	100.0%	0.0%	1
2	RPAPAM	32.6%	80.8%	19.2%	2
3	RT2IP	29.2%	100.0%	0.0%	1
4	RT2APAM	28.8%	100.0%	0.0%	1
5	RT2APA184I	11.8%	100.0%	0.0%	1
6	RP/QP	8.7%	100.0%	0.0%	1
7	RPAP	7.0%	100.0%	0.0%	1
8	RT2AP A82L	3.7%	100.0%	0.0%	1
9	RT2APV78I	1.7%	100.0%	0.0%	1
10	RT2	1.2%	100.0%	0.0%	1
11	KSK19 A82M	1.0%	100.0%	0.0%	1
12	R19	0.9%	100.0%	0.0%	1
13	KSK19AI F42L	0.0%	0.0%	0.0%	0
14	RLYF KSK19 AI	0.0%	0.0%	0.0%	0
15	KSK19AI Y51L	0.0%	0.0%	0.0%	0
16	KSK19AI	0.0%	0.0%	0.0%	0
17	KSK19AIAI	0.0%	0.0%	0.0%	0
18	RPFVEV L437VL	0.0%	0.0%	0.0%	0
19	KSK19 I263A	0.0%	0.0%	0.0%	0
20	RP	0.0%	0.0%	0.0%	0

21	RPFVEV A184I	0.0%	0.0%	0.0%	0
22	RP EV	0.0%	0.0%	0.0%	0
23	RPFVEV	0.0%	0.0%	0.0%	0
24	KSK19AI F87V	0.0%	0.0%	0.0%	0
25	RPFVEV L437LV	0.0%	0.0%	0.0%	0
26	KSK19AI V78I	0.0%	0.0%	0.0%	0
27	RT2AP/I401M	0.0%	0.0%	0.0%	0
28	RPIAEV	0.0%	0.0%	0.0%	0
29	WT	0.0%	0.0%	0.0%	0
30	RP FV	0.0%	0.0%	0.0%	0
31	RT2A330W/S72R	0.0%	0.0%	0.0%	0
32	RT2A330W/S72L	0.0%	0.0%	0.0%	0

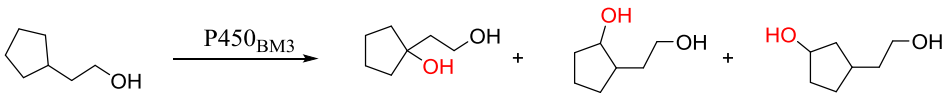
[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM naproxen. After 2 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 40 ℃ which was held for 2 min, the temperature was then raised at 8 ℃/min to 220 ℃. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

4.15 min product: 3-hydroxy-ε-caprolactone

4.09 min product: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.8 Substrate conversion and product distribution of 2-cyclopentylethanol **(4.30)**



4.30 **4.31** **4.32** *cis* **4.33** *cis/trans*

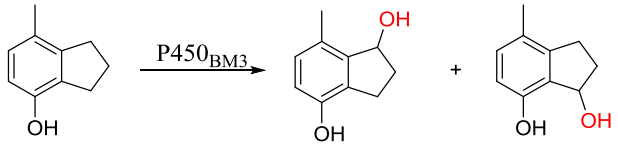
No	Enzyme	Conversion	Product 1 (2.67 min)	Product 2 (2.75 min)	4.31 (3.32 min)	4.32 (3.57 min)	Product 5 (4.09 min)	Product 6 (4.55 min)	Product 7 (4.58 min)	Product 8 (4.62 min)	Product 9 (4.64 min)	No. Of products
1	RPAPAM	100.0%	0.0%	0.0%	22.1%	15.4%	0.0%	0.9%	61.5%	0.0%	0.0%	4
2	RPFVEV	100.0%	0.0%	0.0%	7.4%	41.0%	0.0%	9.1%	20.0%	6.6%	15.9%	6
3	KSK19AI F87V	100.0%	0.0%	0.0%	0.0%	10.5%	0.0%	37.7%	23.1%	9.7%	19.0%	5
4	RT2APA M	100.0%	0.0%	0.0%	27.5%	12.8%	0.0%	3.2%	56.5%	0.0%	0.0%	4
5	RT2IP	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	5.3%	94.7%	0.0%	0.0%	2
6	RT2AP	100.0%	0.0%	0.0%	31.2%	4.5%	0.0%	1.8%	62.6%	0.0%	0.0%	4
7	RT2APA184I	100.0%	0.0%	0.0%	36.9%	11.6%	0.0%	2.0%	49.5%	0.0%	0.0%	4
8	RT2APV78I	100.0%	0.0%	0.0%	7.5%	0.0%	0.0%	1.3%	91.2%	0.0%	0.0%	3
9	RPIAEV	100.0%	0.0%	0.0%	0.0%	32.5%	0.0%	3.6%	45.9%	4.6%	13.4%	5
10	RPAP	100.0%	0.0%	0.0%	3.1%	5.8%	0.0%	2.5%	88.6%	0.0%	0.0%	4
11	RP FV	100.0%	0.0%	0.0%	15.2%	1.8%	6.0%	3.2%	73.9%	0.0%	0.0%	5
12	KSK19 A82M	96.3%	0.0%	0.0%	0.0%	3.3%	45.1%	12.0%	25.5%	5.1%	9.0%	6
13	RP	91.8%	2.6%	0.7%	0.0%	4.4%	0.0%	3.9%	88.4%	0.0%	0.0%	5
14	RT2	77.6%	3.3%	1.4%	0.0%	4.5%	0.0%	4.1%	86.8%	0.0%	0.0%	5
15	RPFVEV L437VL	66.0%	1.5%	1.7%	0.0%	3.3%	1.4%	12.2%	55.5%	7.6%	16.9%	8
16	KSK19AI Y51L	65.2%	12.3%	16.8%	0.0%	7.1%	10.0%	23.4%	8.3%	6.1%	16.0%	8
17	RP EV	64.1%	5.1%	3.2%	0.0%	0.0%	0.0%	5.1%	86.7%	0.0%	0.0%	4
18	RLYF KSK19 AI	35.1%	17.0%	20.6%	0.0%	11.4%	5.8%	17.3%	6.4%	6.1%	15.3%	8

19	KSK19AI AI	34.8%	15.7%	12.7%	0.0%	24.8%	0.0%	17.5%	7.3%	5.2%	16.8%	7
20	KSK19AI	30.6%	16.7%	20.8%	0.0%	3.5%	7.5%	22.4%	7.9%	6.3%	14.9%	8
21	KSK19AI F42L	20.8%	19.2%	24.9%	0.0%	0.0%	6.2%	20.1%	7.6%	6.6%	15.3%	7
22	RT2AP/I401M	19.8%	7.3%	5.4%	0.0%	0.0%	0.0%	2.6%	84.7%	0.0%	0.0%	4
23	RPFVEV L437LV	18.9%	4.5%	0.0%	0.0%	0.0%	0.0%	19.2%	39.8%	8.8%	27.7%	5
24	WT	15.2%	8.9%	7.0%	0.0%	0.0%	0.0%	2.0%	82.2%	0.0%	0.0%	4
25	RT2A330 W/S72R	8.0%	9.7%	8.4%	0.0%	0.0%	0.0%	0.0%	81.9%	0.0%	0.0%	3
26	KSK19AI V78I	4.1%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1
27	KSK19 I263A	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM 2-cyclopentylethanol. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas.

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.9 Substrate conversion and product distribution of 7-methyl-2,3-dihydro-1H-inden-4-ol **(4.34)**



4.34 **4.35** **4.36**

No	Enzyme	Conversion	Product 1 (5.72 min)	Product 2 (6.28 min)	Product 3 (6.47 min)	Product 4 (6.56 min)	Product 5 (6.85 min)	Product 6 (7.18 min)	Product 7 (7.39 min)	4.35 (7.74 min)	Product 9 (7.80 min)	4.36 (8.02 min)	NO. of products
1	RT2APAI	100.0%	1.7%	0.0%	0.0%	0.0%	1.6%	0.5%	0.8%	78.8%	0.0%	16.6%	2
2	GVQ	98.8%	0.0%	4.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	95.6%	2
3	KSK19AIFV	95.6%	0.0%	8.7%	0.9%	0.0%	2.5%	0.0%	0.0%	0.0%	0.0%	87.9%	3
4	RT2AP	86.2%	0.0%	6.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	93.1%	2
5	RT2APAM	82.9%	2.3%	0.0%	3.1%	0.0%	0.0%	2.6%	0.9%	55.4%	0.0%	35.7%	5
6	YFKSK19AI	81.8%	0.0%	9.6%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	89.1%	2
7	KSK19AI	74.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	1
8	RT2APVI	69.4%	0.0%	0.0%	1.2%	0.0%	0.0%	1.4%	4.8%	69.8%	0.0%	22.9%	3
9	RPFVEVL437LV	34.2%	0.0%	1.5%	4.7%	0.0%	0.0%	2.8%	2.1%	5.3%	0.0%	83.6%	5
10	RT2APAMSA	29.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	1
11	RT2APTL	28.4%	0.0%	10.5%	3.1%	0.0%	0.0%	3.0%	0.0%	0.0%	0.0%	83.4%	4
12	RPAPSL	27.4%	0.0%	24.4%	2.3%	1.3%	3.4%	3.7%	0.0%	0.0%	1.2%	63.7%	5
13	RT2APAV	24.2%	0.0%	23.8%	0.0%	0.0%	2.5%	2.6%	0.0%	0.0%	0.0%	71.0%	4
14	KSK19AIIP	20.2%	0.0%	18.7%	3.1%	0.0%	0.0%	3.0%	0.0%	0.0%	0.0%	75.2%	4
15	RPFW	18.3%	0.0%	13.8%	0.0%	0.0%	0.0%	2.8%	0.0%	0.0%	0.0%	83.4%	3
16	RPAPAV	15.7%	0.0%	6.3%	5.6%	0.0%	0.0%	3.5%	0.0%	5.2%	2.7%	76.7%	6
17	RP	14.8%	0.0%	20.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	79.2%	2
18	KSK19AM	13.3%	2.7%	10.8%	0.0%	0.0%	0.0%	10.8%	0.0%	0.0%	0.0%	75.7%	4
19	RT2	12.2%	0.0%	22.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	77.2%	2
20	KSK19AIVI	11.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	1

21	RPAPAM	11.1%	0.0%	44.0%	24.9%	0.0%	0.0%	17.3%	0.0%	0.0%	0.0%	13.7%	4
22	RPFVEV	10.6%	0.0%	6.9%	6.6%	0.0%	0.0%	0.0%	0.0%	6.8%	0.0%	79.8%	4
23	RPFVEVL437VL	10.2%	0.0%	25.4%	0.0%	3.0%	0.0%	0.0%	0.0%	0.0%	0.0%	71.6%	3
24	RT2IP	9.9%	0.0%	7.9%	0.0%	0.0%	0.0%	0.0%	0.0%	22.7%	8.9%	60.5%	4
25	RPFVEVFW	9.4%	0.0%	5.9%	5.9%	3.4%	0.0%	3.9%	0.0%	6.0%	4.5%	70.4%	7
26	RT2IPAP	7.6%	0.0%	8.9%	0.0%	0.0%	0.0%	0.0%	0.0%	13.6%	7.1%	70.3%	4
27	RPAP	6.9%	0.0%	17.7%	10.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	72.0%	3
28	RLYFKSK19AI	5.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	1
29	RQA	5.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	1
30	RT2AW	5.6%	0.0%	8.9%	4.6%	5.1%	0.0%	0.0%	0.0%	0.0%	0.0%	81.5%	4
31	KSK19IA	5.1%	0.0%	21.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	78.4%	2
32	WT	1.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	1

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM 7-methyl-2,3-dihydro-1H-inden-4-ol. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

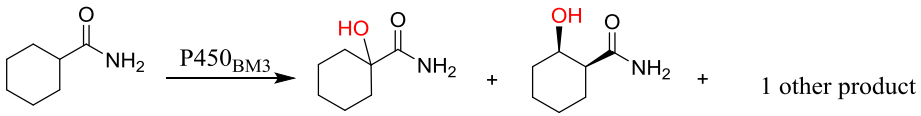
7.74 min product: 7-methyl-2,3-dihydro-1H-indene-1,4-diol (**4.35**)

8.02 min product: 4-methyl-2,3-dihydro-1H-indene-1,7-diol (**4.36**)

Other products: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Table S4.10 Substrate conversion and product distribution of cyclohexanecarboxamide **(4.37)**



4.37 **4.38** **4.39**

No.	Enzyme	Conversion	4.38 (5.67 min)	4.39 (6.47 min)	Product 3 (7.18 min)	No. of product
1	RT2AP	100.0%	54.9%	45.1%	0.0%	1
2	RT2AP A82L	100.0%	11.7%	88.3%	0.0%	0
3	RP/QP	95.0%	58.7%	41.3%	0.0%	0
4	RT2APAM	92.1%	73.4%	26.6%	0.0%	2
5	RPAPAM	89.5%	65.6%	34.4%	0.0%	2
6	RT2APA184I	88.1%	62.6%	37.4%	0.0%	1
7	RPAP	80.9%	57.8%	42.2%	0.0%	0
8	RT2	80.1%	12.2%	87.8%	0.0%	0
9	RT2IP	42.0%	17.6%	82.4%	0.0%	1
10	RT2APV78I	41.7%	65.0%	35.0%	0.0%	1
11	R19	28.2%	11.6%	88.4%	0.0%	0
12	KSK19 A82M	11.4%	78.7%	0.0%	21.3%	2
13	RPIAEV	7.0%	0.0%	100.0%	0.0%	0
14	RP	5.9%	29.8%	70.2%	0.0%	2
15	RP FV	5.8%	29.2%	70.8%	0.0%	0
16	RT2A330W/S72L	5.1%	41.8%	58.2%	0.0%	0
17	KSK19 I263A	2.5%	0.0%	0.0%	100.0%	2
18	RP EV	2.2%	32.5%	67.5%	0.0%	2
19	RT2A330W/S72R	1.2%	0.0%	0.0%	0.0%	0
20	RPFVEV	0.8%	100.0%	0.0%	0.0%	1

21	KSK19AI V78I	0.8%	0.0%	0.0%	100.0%	2
22	RT2AP/I401M	0.7%	0.0%	100.0%	0.0%	1
23	KSK19AI F42L	0.0%	0.0%	0.0%	0.0%	2
24	RLYF KSK19 AI	0.0%	0.0%	0.0%	0.0%	2
25	KSK19AI Y51L	0.0%	0.0%	0.0%	0.0%	2
26	KSK19AI	0.0%	0.0%	0.0%	0.0%	2
27	KSK19AIAI	0.0%	0.0%	0.0%	0.0%	2
28	RPFVEV L437VL	0.0%	0.0%	0.0%	0.0%	2
29	RPFVEV A184I	0.0%	0.0%	0.0%	0.0%	2
30	KSK19AI F87V	0.0%	0.0%	0.0%	0.0%	2
31	RPFVEV L437LV	0.0%	0.0%	0.0%	0.0%	2
32	WT	0.0%	0.0%	0.0%	0.0%	0

[a] The screening reaction mixture (0.5 mL) contained 200 mM phosphate, pH 7.5, 1 μM P450, 2 mM cyclohexanecarboxamide. After 16 h the reaction mixture was extracted with ethyl acetate (300 μl). The phases were separated by centrifugation at 21,000 *g*. The organics were analysed by gas chromatography. The initial oven temperature was 100 ℃ which was held for 2 min, the temperature was then raised at 16 ℃/min to 220 ℃ and kept at 220 ℃ for 2 min. The product distributions were determined from the integration of peak areas. The GC retention times represent the following products.

5.67 min product: 1-hydroxycecylohexane-1-carboxamide (**4.38**)

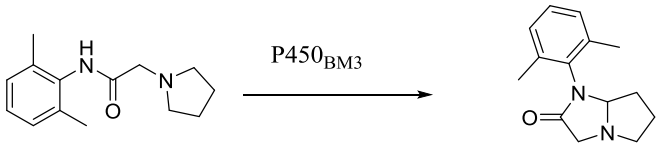
6.47 min product: *cis*-2-hydroxycyclohexane-1-carboxamide (**4.39**)

7.18 min product: unidentified

[b] Average of 2 independent experiments with errors not higher than 10%.

Appendix 5 Screening results of lidocaine derivatives in chapter 5

Table S5.1 Substrate conversion of the initial screen towards lidocaine pyrrolidine derivative



N-(2,6-dimethylphenyl)-2-(pyrrolidin-1-yl)acetamide

1-(2,6-dimethylphenyl)tetrahydro-1*H*-pyrrolo[1,2-*a*]imidazol-2(3*H*)-one, **5.01**

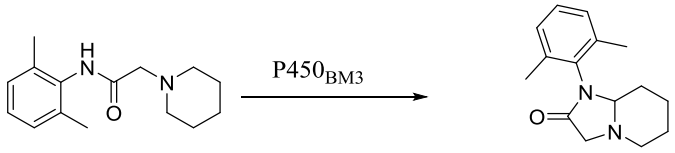
No	enzyme	conversion	TON
1	RP/FV/EV	98.8%	1976
2	KSK19/I263A	34.9%	698
3	RP/FV/EV/A184I	24.9%	499
4	KSK19/QP	23.7%	474
5	RT2/IP	19.3%	386
6	KSK19	15.6%	313
7	RP/IA/EV	11.9%	238
8	RT2/AP/A184I	3.9%	79
9	RP/E267V	3.1%	62
10	KSK19/AI/A184I	2.8%	55
11	WT	2.7%	54
12	RT2/AP/A82M	2.3%	46
13	RP/FV	1.6%	32
14	RT2/AP/V78I	1.5%	30
15	KSK19/AI/Y51L	1.2%	24
16	RT2/AP	1.2%	23
17	RT2/I401P/A330P	1.1%	22
18	RP/AP	1.1%	21
19	RT2/A330W	1.1%	21
20	KSK19/AI/F42L	0.9%	19
21	RP/A330P/A82M	0.9%	18
22	KSK19/AI/V78I	0.9%	18
23	RLYF/KSK19/AI	0.9%	18
24	KSK19/AI	0.9%	18
25	RT2	0.7%	14
26	KSK19/A82M	0.7%	13
27	RP/F81W	0.5%	11
28	RP	0.4%	9

(a) The initial screen was conducted using 1 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.2 Substrate conversion of the initial screen towards lidocaine piperidine derivative



N-(2,6-dimethylphenyl)-2-(piperidin-1-yl)acetamide

1-(2,6-dimethylphenyl)hexahydroimidazo[1,2-*a*]pyridin-2(3*H*)-one, **5.02**

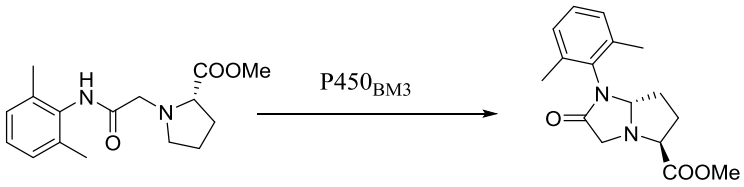
No	enzyme	conversion	TON
1	RT2AP/A184I	94.4%	1888
2	RP/FV/EV	77.4%	1549
3	RT2IP	32.3%	647
4	RP/FV/EV/A184I	26.8%	536
5	RP/IA/EV	24.0%	480
6	RPAP	21.8%	436
7	RP/EV	8.6%	173
8	RP/FV/EV/L437LV	8.6%	171
9	KSK19/AI/AI	8.4%	167
10	RP	7.9%	159
11	KSK19/IA	7.5%	149
12	RT2AP/V78I	7.4%	147
13	RP/FV	6.1%	122
14	KSK19/AI/FV	5.7%	115
15	RP/FV/EV/L437VL	4.1%	82
16	KSK19/AI/Y51L	3.6%	73
17	RT2AP/AM	3.2%	63
18	KSK19/A82M	3.0%	61
19	RT2	2.7%	55
20	KSK19/AI	2.4%	48
21	KSK19/AI/F42L	2.4%	48
22	RLYF/KSK19/AI	2.3%	45
23	RT2AP/IP	2.1%	41
24	RT2AP	1.9%	39
25	RPAPAM	1.3%	26
26	KSK19/AI/VI	1.3%	26
27	WT	0.8%	15

(a) The initial screen was conducted using 1 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.3 Substrate conversion of the initial screen towards lidocaine proline methyl ester derivative



methyl (2-((2,6-dimethylphenyl)amino)-2-oxoethyl)-*L*-prolinate

methyl (5*S*,7a*R*)-1-(2,6-dimethylphenyl)-2-oxohexahydro-1*H*-pyrrolo[1,2-*a*]imidazole-5-carboxylate, **5.03**

No	Enzyme	Conversion	Cyclised Product	Others	TON
1	RT2F81W	100.0%	100.0%	0.0%	1000
2	RP/E267I	100.0%	97.6%	2.4%	1000
3	RP EV	86.2%	97.9%	2.1%	862
4	RPFVEV	85.0%	98.9%	1.1%	850
5	RPFVEV L188Q	76.0%	98.9%	1.1%	760
6	EVVILI	70.3%	98.4%	1.6%	703
7	RPIAEV	68.8%	97.1%	2.9%	688
8	RT2IP A184I	66.3%	98.5%	1.5%	663
9	KSK19 A82M	52.5%	97.5%	2.5%	525
10	KSK19/A82M/I401M	48.1%	96.6%	3.4%	481
11	RT2IP/F87V	43.0%	95.0%	5.0%	430
12	KSK19 QP F87V	39.1%	95.5%	4.5%	391
13	RPFVEV V78F	38.3%	98.4%	1.6%	383
14	KT2 E267L	38.0%	98.6%	1.4%	380
15	RP/E267M	35.2%	94.6%	5.4%	352
16	RT2AP/A82F	28.4%	61.9%	38.1%	284
17	KSK19 QP	27.1%	97.9%	2.1%	271
18	RT2AP/E267L	23.5%	97.1%	2.9%	235
19	RP FV	22.7%	100.0%	0.0%	227
20	KSK19/A82M/E267F	21.2%	94.6%	5.4%	212
21	RPFVEVF81W	19.6%	95.6%	4.4%	196
22	RP/FA	17.4%	95.2%	4.8%	174
23	RPFVEV L437LV	14.5%	91.0%	9.0%	145
24	GVQ	13.9%	91.6%	8.4%	139
25	KSK19AIAI	12.9%	84.7%	15.3%	129
26	RPFVEV L437VL	12.2%	90.9%	9.1%	122
27	F87A I401P	11.2%	100.0%	0.0%	112
28	RP H171L	11.1%	87.4%	12.6%	111
29	KSK19	11.0%	96.0%	4.0%	110
30	RT2APAMAI	10.9%	72.7%	27.3%	109
31	RT2IP	10.6%	94.1%	5.9%	106
32	KT2E267F	10.3%	93.0%	7.0%	103
33	RPF81W	8.1%	92.4%	7.6%	81
34	RLYF KSK19	7.3%	92.6%	7.4%	73
35	RT2AP A82L	6.9%	47.9%	52.1%	69
36	QP A82M	6.9%	85.1%	14.9%	69

37	RT2 A330W A82M	6.5%	55.9%	44.1%	65
38	RP	5.9%	65.1%	34.9%	59
39	RP AM IA	5.5%	64.1%	35.9%	55
40	RT2APA184I	5.3%	86.6%	13.4%	53
41	RT2APAM	5.2%	87.3%	12.7%	52
42	RT2AP/E267F	4.8%	52.9%	47.1%	48
43	KSK19AI Y51L	4.7%	39.6%	60.4%	47
44	RT2F81W A328I	4.1%	45.7%	54.3%	41
45	KSK19AI F87V	3.4%	55.1%	44.9%	34
46	RPAPAM	3.4%	47.2%	52.8%	34
47	RP/IA	3.3%	59.1%	40.9%	33
48	RPAP	3.3%	65.9%	34.1%	33
49	KSK19 I263A	2.9%	55.1%	44.9%	29
50	RT2APA184I F81W	2.9%	81.5%	18.5%	29
51	RP/F87A/A328I	2.8%	75.7%	24.3%	28
52	KSK44 a	2.6%	32.4%	67.6%	26
53	KSK19 T88A	2.6%	41.9%	58.1%	26
54	RLYF	2.5%	0.0%	100.0%	25
55	KSK19 F87S	2.3%	46.3%	53.7%	23
56	KSK48	2.2%	46.3%	53.7%	22
57	A328V	2.2%	0.0%	100.0%	22
58	RT2	2.1%	46.3%	53.7%	21
59	RLYF KSK19 AI	2.0%	58.0%	42.0%	20
60	A328S P329S Q403P	1.9%	0.0%	100.0%	19
61	KSK19 AI IA	1.8%	0.0%	100.0%	18
62	KSK19AI V78I	1.7%	0.0%	100.0%	17
63	RQA	1.7%	39.5%	60.5%	17
64	N70S	1.7%	59.8%	40.2%	17
65	RT2AP	1.6%	52.1%	47.9%	16
66	RT2 V78F	1.6%	0.0%	100.0%	16
67	KSK43	1.6%	0.0%	100.0%	16
68	RT2AI L437LV	1.6%	0.0%	100.0%	16
69	A328P	1.5%	50.6%	49.4%	15
70	RT2AP/I401M	1.5%	0.0%	100.0%	15
71	KT2	1.5%	0.0%	100.0%	15
72	RT2AP VIAI	1.4%	0.0%	100.0%	14
73	RPAPPV	1.4%	0.0%	100.0%	14
74	A330P	1.4%	0.0%	100.0%	14
75	RT2 A330W	1.3%	0.0%	100.0%	13
76	IP QP	1.3%	53.3%	46.7%	13
77	RT2AI L437VL	1.3%	0.0%	100.0%	13
78	I263F	1.3%	40.9%	59.1%	13
79	F87A	1.3%	0.0%	100.0%	13
80	KT2 F87I	1.2%	0.0%	100.0%	12
81	F87P I401P	1.2%	0.0%	100.0%	12
82	QP	1.1%	0.0%	100.0%	11
83	RT2APPV	1.1%	0.0%	100.0%	11
84	R19	0.8%	0.0%	100.0%	8
85	KSK19AI	0.8%	0.0%	100.0%	8

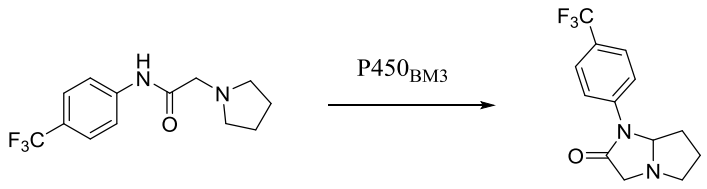
86	RT2APAI S72A	0.8%	0.0%	100.0%	8
87	A328G	0.8%	0.0%	100.0%	8
88	RT2APV78I	0.8%	0.0%	100.0%	8
89	A330W	0.8%	0.0%	100.0%	8
90	WT	0.7%	0.0%	100.0%	7
91	RT2 I263W	0.7%	0.0%	100.0%	7
92	F87A T88A I401P	0.7%	0.0%	100.0%	7
93	T88M IP	0.7%	0.0%	100.0%	7
94	KT2A82G	0.7%	0.0%	100.0%	7
95	RT2A330W I263A	0.7%	0.0%	100.0%	7
96	I263V	0.7%	0.0%	100.0%	7

(a) The initial screen was conducted using 1 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.4 Substrate conversion of the initial screen towards Lidocaine CF_3 derivative



2-(pyrrolidin-1-yl)-*N*-(4-(trifluoromethyl)phenyl)acetamide

1-(4-(trifluoromethyl)phenyl)tetrahydro-1*H*-pyrrolo[1,2-*a*]imidazol-2(3*H*)-one, **5.04**

No	Enzymes	Conversion	Cyclisation product	Others	TON
1	RPFVEV V78F	98.9%	44.4%	55.6%	495
2	RPFVEV L188Q	97.1%	61.5%	38.5%	485
3	RT2F81W	90.0%	76.6%	23.4%	450
4	RPFVEV	86.6%	69.2%	30.8%	433
5	KSK19/A82M/E267F	83.9%	83.8%	16.2%	420
6	KSK19 A82M	71.8%	76.7%	23.3%	359
7	RT2IP/F87V	71.1%	33.0%	67.0%	356
8	RP H171L	71.0%	53.1%	46.9%	355
9	RPIAEV	66.2%	48.1%	51.9%	331
10	RPFVEVF81W	65.7%	80.3%	19.7%	329
11	RP	64.9%	64.0%	36.0%	324
12	KSK19/A82M/I401M	63.9%	83.5%	16.5%	320
13	RP FV	62.9%	51.3%	48.7%	315
14	RP/E267M	58.5%	47.4%	52.6%	293
15	RPFVEV L437VL	58.0%	64.6%	35.4%	290
16	RPF81W	57.0%	52.0%	48.0%	285
17	RP EV	46.8%	59.9%	40.1%	234
18	RP/E267I	46.2%	65.9%	34.1%	231
19	RPFVEV L437LV	44.2%	73.3%	26.7%	221
20	KSK19 QP F87V	43.6%	47.8%	52.2%	218
21	RT2IP	42.8%	53.0%	47.0%	214
22	RT2APPV	41.0%	87.0%	13.0%	205
23	RT2 I263W	40.6%	85.3%	14.7%	203
24	RT2	39.4%	80.3%	19.7%	197
25	RT2IP A184I	39.1%	16.0%	84.0%	196
26	R19	39.1%	92.3%	7.7%	195
27	QP A82M	38.7%	79.1%	20.9%	193
28	EVVILI	37.9%	58.6%	41.4%	189
29	RP/FA	37.5%	77.4%	22.6%	187
30	RP AM IA	37.2%	88.1%	11.9%	186
31	RLYF KSK19	34.9%	78.9%	21.1%	174
32	KT2E267F	34.1%	77.3%	22.7%	170
33	KT2 E267L	33.0%	57.9%	42.1%	165
34	RP/IA	32.4%	57.5%	42.5%	162

35	KT2A82G	27.5%	69.7%	30.3%	137
36	RT2AP/A82F	26.3%	16.3%	83.7%	131
37	KT2	26.3%	88.8%	11.2%	131
38	KSK19AI F87V	25.2%	100.0%	0.0%	126
39	KSK19 QP	20.9%	66.2%	33.8%	104
40	F87A I401P	20.5%	78.6%	21.4%	102
41	RP/F87A/A328I	20.4%	100.0%	0.0%	102
42	KSK19	20.4%	80.7%	19.3%	102
43	RT2F81W A328I	20.3%	82.3%	17.7%	102
44	RQA	20.2%	100.0%	0.0%	101
45	GVQ	17.8%	60.4%	39.6%	89
46	RT2AP A82L	15.9%	100.0%	0.0%	79
47	RT2 V78F	15.2%	100.0%	0.0%	76
48	KT2 F87I	14.4%	100.0%	0.0%	72
49	RT2APAMAI	12.9%	44.4%	55.6%	64
50	KSK19AIAI	12.4%	100.0%	0.0%	62
51	RPAPAM	12.1%	100.0%	0.0%	61
52	RT2APA184I	11.3%	34.0%	66.0%	56
53	RT2AP	10.7%	65.7%	34.3%	53
54	RT2AP VIAI	10.5%	52.8%	47.2%	52
55	RT2APA184I F81W	10.4%	63.7%	36.3%	52
56	RPAP	10.2%	60.3%	39.7%	51
57	RT2APAM	10.0%	71.3%	28.7%	50
58	RT2AP/E267L	9.1%	31.1%	68.9%	46
59	RT2 A330W A82M	8.9%	68.7%	31.3%	45
60	QP	8.1%	100.0%	0.0%	41
61	RT2AP/E267F	8.0%	43.8%	56.2%	40
62	KSK19 F87S	7.6%	100.0%	0.0%	38
63	KSK19AI Y51L	7.3%	100.0%	0.0%	36
64	RT2 A330W	6.9%	81.1%	18.9%	34
65	RT2AP/I401M	5.7%	100.0%	0.0%	28
66	A328P	5.3%	100.0%	0.0%	27
67	N70S	5.3%	100.0%	0.0%	26
68	RLYF KSK19 AI	5.0%	100.0%	0.0%	25
69	IP QP	4.8%	100.0%	0.0%	24
70	RT2AI L437LV	4.2%	100.0%	0.0%	21
71	A328S P329S Q403P	4.2%	100.0%	0.0%	21
72	A328V	4.1%	100.0%	0.0%	21
73	KSK48	4.1%	100.0%	0.0%	20
74	KSK19 T88A	3.9%	100.0%	0.0%	19
75	I263F	3.7%	100.0%	0.0%	19
76	KSK43	3.2%	100.0%	0.0%	16
77	RT2APAI S72A	3.1%	100.0%	0.0%	15
78	KSK19 I263A	2.9%	100.0%	0.0%	15
79	RT2AI L437VL	2.8%	100.0%	0.0%	14
80	RPAPPV	2.7%	100.0%	0.0%	14

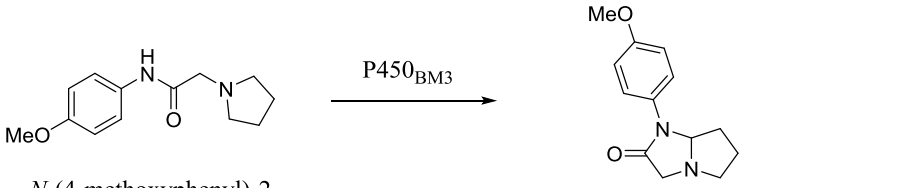
81	RT2APV78I	2.6%	100.0%	0.0%	13
82	KSK44 a	2.4%	100.0%	0.0%	12
83	KSK19AI V78I	2.3%	100.0%	0.0%	12
84	KSK19AI	1.9%	100.0%	0.0%	9
85	RLYF	1.8%	100.0%	0.0%	9
86	F87P I401P	1.8%	100.0%	0.0%	9
87	F87A	1.6%	100.0%	0.0%	8
88	KSK19 AI IA	1.6%	100.0%	0.0%	8
89	F87A T88A I401P	1.5%	100.0%	0.0%	8
90	WT	1.2%	100.0%	0.0%	6
91	A330P	1.1%	100.0%	0.0%	5
92	I263V	0.8%	100.0%	0.0%	4
93	T88M IP	0.8%	100.0%	0.0%	4
94	RT2A330W I263A	0.7%	100.0%	0.0%	4
95	A328G	0.6%	100.0%	0.0%	3
96	A330W	0.5%	100.0%	0.0%	3

(a) The initial screen was conducted using 2 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.5 Substrate conversion of the initial screen towards OMe derivative



N-(4-methoxyphenyl)-2-(pyrrolidin-1-yl)acetamide

1-(4-methoxyphenyl)tetrahydro-1*H*-pyrrolo[1,2-*a*]imidazol-2(3*H*)-one, **5.05**

No	Enzymes	Conversion	Cyclisation product	Others	TON
1	RPFVEV	100.0%	71.7%	28.3%	500
2	KSK19 QP F87V	94.5%	96.7%	3.3%	473
3	KSK19/A82M/E267F	90.7%	97.8%	2.2%	453
4	KSK19 A82M	89.7%	100.0%	0.0%	449
5	RPFVEV V78F	89.0%	92.5%	7.5%	445
6	RPFVEVF81W	77.7%	90.1%	9.9%	389
7	KSK19/A82M/I401M	75.7%	98.6%	1.4%	379
8	RP H171L	69.6%	83.7%	16.3%	348
9	RP/E267M	67.7%	93.5%	6.5%	338
10	RPIAEV	63.6%	92.0%	8.0%	318
11	RP/E267I	61.3%	97.5%	2.5%	306
12	RPFVEV L437VL	59.8%	97.5%	2.5%	299
13	RT2IP/F87V	49.5%	96.0%	4.0%	248
14	RP	49.4%	100.0%	0.0%	247
15	KT2 E267L	48.7%	100.0%	0.0%	244
16	RP EV	48.1%	90.6%	9.4%	240
17	EVVILI	48.1%	96.3%	3.7%	240
18	RP FV	47.6%	94.7%	5.3%	238
19	RPFVEV L437LV	43.9%	91.9%	8.1%	219
20	GVQ	42.5%	100.0%	0.0%	213
21	QP A82M	40.6%	97.8%	2.2%	203
22	RP/FA	34.3%	100.0%	0.0%	172
23	KT2E267F	32.8%	100.0%	0.0%	164
24	RT2AP A82L	27.6%	89.8%	10.2%	138
25	RT2IP	25.7%	88.5%	11.5%	128
26	KSK19 QP	23.1%	100.0%	0.0%	116
27	KT2A82G	22.5%	100.0%	0.0%	113
28	RP AM IA	22.0%	100.0%	0.0%	110
29	RT2 I263W	21.8%	100.0%	0.0%	109
30	R19	20.7%	96.2%	3.8%	103
31	RT2	19.0%	95.8%	4.2%	95
32	KT2 F87I	18.2%	100.0%	0.0%	91
33	KSK19AI F87V	16.0%	93.9%	6.1%	80
34	RPF81W	15.2%	100.0%	0.0%	76

35	RT2IP A184I	15.1%	88.1%	11.9%	75
36	KSK19AIAI	14.3%	100.0%	0.0%	71
37	RLYF KSK19	14.3%	100.0%	0.0%	71
38	RT2APAMAI	13.5%	100.0%	0.0%	68
39	RP/IA	5.8%	92.8%	7.2%	29
40	RT2AP/A82F	5.6%	80.4%	19.6%	28
41	RT2APA184I	5.1%	75.5%	24.5%	25
42	RT2APAM	4.9%	84.9%	15.1%	25
43	KSK19AI Y51L	4.9%	67.1%	32.9%	24
44	RT2APA184I F81W	4.7%	88.4%	11.6%	23
45	RPAP	4.4%	53.1%	46.9%	22
46	RP/F87A/A328I	4.3%	84.3%	15.7%	21
47	RPAPAM	4.0%	84.8%	15.2%	20
48	KSK48	4.0%	100.0%	0.0%	20
49	RT2AP	3.9%	80.7%	19.3%	19
50	RT2AP/I401M	3.8%	77.0%	23.0%	19
51	RT2F81W A328I	3.7%	100.0%	0.0%	19
52	RLYF KSK19 AI	3.5%	100.0%	0.0%	17
53	RT2 A330W A82M	3.5%	83.2%	16.8%	17
54	RT2APAI S72A	3.1%	100.0%	0.0%	15
55	KSK19	3.1%	100.0%	0.0%	15
56	KSK19 F87S	2.6%	100.0%	0.0%	13
57	A328P	2.4%	100.0%	0.0%	12
58	F87A I401P	2.4%	100.0%	0.0%	12
59	RT2AP/E267L	2.2%	100.0%	0.0%	11
60	RT2AI L437LV	2.1%	100.0%	0.0%	11
61	KSK19AI V78I	2.1%	42.8%	57.2%	10
62	RT2AP/E267F	2.1%	100.0%	0.0%	10
63	T88M IP	1.8%	30.0%	70.0%	9
64	KSK19 I263A	1.7%	100.0%	0.0%	9
65	RT2APV78I	1.7%	52.1%	47.9%	9
66	RT2APPV	1.7%	100.0%	0.0%	9
67	KT2	1.6%	100.0%	0.0%	8
68	KSK19AI	1.5%	51.5%	48.5%	8
69	RQA	1.4%	100.0%	0.0%	7
70	RT2AP VIAI	1.4%	100.0%	0.0%	7
71	A328V	1.4%	100.0%	0.0%	7
72	KSK44 a	1.2%	100.0%	0.0%	6
73	KSK43	1.2%	100.0%	0.0%	6
74	WT	1.2%	0.0%	100.0%	6
75	RT2 V78F	1.1%	100.0%	0.0%	6
76	RT2AI L437VL	1.0%	100.0%	0.0%	5
77	IP QP	0.9%	100.0%	0.0%	4
78	RPAPPV	0.8%	100.0%	0.0%	4
79	QP	0.8%	100.0%	0.0%	4
80	KSK19 T88A	0.8%	100.0%	0.0%	4

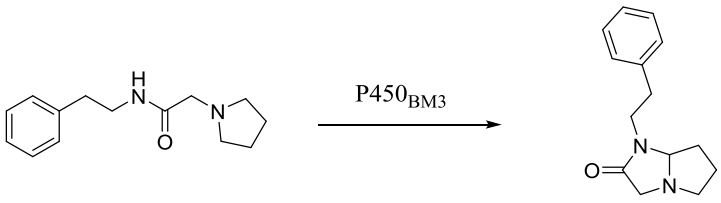
81	RT2 A330W	0.6%	100.0%	0.0%	3
82	F87A T88A I401P	0.0%	0.0%	0.0%	0
83	KSK19 AI IA	0.0%	0.0%	0.0%	0
84	A330P	0.0%	0.0%	0.0%	0
85	RT2F81W	0.0%	0.0%	0.0%	0
86	F87P I401P	0.0%	0.0%	0.0%	0
87	F87A	0.0%	0.0%	0.0%	0
88	RLYF	0.0%	0.0%	0.0%	0
89	A328S P329S Q403P	0.0%	0.0%	0.0%	0
90	RT2A330W I263A	0.0%	0.0%	0.0%	0
91	A328G	0.0%	0.0%	0.0%	0
92	A330W	0.0%	0.0%	0.0%	0
93	I263V	0.0%	0.0%	0.0%	0
94	I263F	0.0%	0.0%	0.0%	0
95	N70S	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.6 Substrate conversion of the initial screen towards 2 extra carbons derivative



N-phenethyl-2-(pyrrolidin-1-yl)acetamide

1-phenethyltetrahydro-1*H*-pyrrolo[1,2-*a*]imidazol-2(3*H*)-one, **5.06**

No	Enzyme	Conversion	Cyclisation Product	others	TON
1	KSK19/A82M/E267F	79.2%	93.5%	6.5%	396
2	RLYF KSK19	71.3%	78.4%	21.6%	356
3	KSK19	69.8%	83.6%	16.4%	349
4	KSK19 QP	58.5%	68.9%	31.1%	292
5	RT2 F87A A330W	42.3%	100.0%	0.0%	212
6	RP H171L	41.3%	97.3%	2.7%	206
7	RPFVEV V78F	39.0%	96.1%	3.9%	195
8	RP/E267I	35.9%	58.6%	41.4%	179
9	RPFVEVF81W	35.6%	92.5%	7.5%	178
10	RPIAEV	35.1%	78.4%	21.6%	176
11	KSK19 A82M	33.4%	94.5%	5.5%	167
12	RP/FA	33.4%	72.6%	27.4%	167
13	RP/E267M	32.8%	100.0%	0.0%	164
14	RPFVEV L437VL	29.9%	100.0%	0.0%	149
15	RPFVEV L188Q	29.6%	91.9%	8.1%	148
16	KT2 E267L	29.0%	91.4%	8.6%	145
17	RPFVEV	28.8%	93.9%	6.1%	144
18	KSK19AIAI	28.4%	100.0%	0.0%	142
19	KT2E267F	25.8%	97.2%	2.8%	129
20	RP	25.0%	100.0%	0.0%	125
21	RP EV	24.6%	100.0%	0.0%	123
22	EVVILI	24.3%	100.0%	0.0%	121
23	RPFVEV L437LV	23.6%	97.1%	2.9%	118
24	KSK19/A82M/I401M	23.6%	94.5%	5.5%	118
25	KSK19 QP F87V	22.2%	96.1%	3.9%	111
26	RT2F81W	22.1%	91.0%	9.0%	111
27	RP/IA	20.3%	97.7%	2.3%	102
28	RP FV	20.0%	89.1%	10.9%	100
29	RT2IP/F87V	20.0%	74.5%	25.5%	100
30	RP IA EV H171L	19.0%	100.0%	0.0%	95
31	RT2APAMAI	18.7%	92.0%	8.0%	93
32	RPF81W	18.4%	100.0%	0.0%	92
33	RT2APPV	18.1%	100.0%	0.0%	90
34	GVQ	17.5%	100.0%	0.0%	87

35	RP AM IA	17.2%	96.2%	3.8%	86
36	QP A82M	16.0%	96.0%	4.0%	80
37	RT2AP A82L	15.5%	90.5%	9.5%	77
38	RT2IP	15.3%	100.0%	0.0%	77
39	R19	15.3%	95.1%	4.9%	77
40	RT2IP A184I	14.7%	88.8%	11.2%	73
41	F87A I401P	13.7%	100.0%	0.0%	68
42	RT2APA184I	13.5%	96.0%	4.0%	67
43	RT2AP/A82F	11.8%	90.5%	9.5%	59
44	RT2	11.2%	100.0%	0.0%	56
45	RT2APAM	11.1%	100.0%	0.0%	55
46	RLYF KSK19 AI	11.0%	100.0%	0.0%	55
47	KSK19 I263A	10.8%	100.0%	0.0%	54
48	RT2 A330W A82M	9.5%	86.1%	13.9%	47
49	RT2 I263W	9.1%	91.5%	8.5%	45
50	KSK19 F87S	8.5%	92.6%	7.4%	43
51	RT2AP/E267L	8.4%	84.6%	15.4%	42
52	RT2AP/E267F	8.4%	89.3%	10.7%	42
53	KT2	8.1%	92.0%	8.0%	40
54	KSK19AI Y51L	7.6%	77.3%	22.7%	38
55	KSK19 AI IA	6.7%	66.7%	33.3%	34
56	RT2 A330W	6.7%	100.0%	0.0%	34
57	F87A	6.6%	90.5%	9.5%	33
58	KSK19AI F87V	6.3%	100.0%	0.0%	32
59	RT2APA184I F81W	6.3%	92.2%	7.8%	31
60	RPAPAM	5.5%	100.0%	0.0%	28
61	RP/F87A/A328I	5.4%	100.0%	0.0%	27
62	RT2AP VIAI	4.8%	87.9%	12.1%	24
63	RT2F81W A328I	4.2%	81.1%	18.9%	21
64	RT2 V78F	4.2%	81.1%	18.9%	21
65	KT2A82G	4.1%	86.8%	13.2%	21
66	RT2AP/I401M	3.9%	64.8%	35.2%	20
67	RPAP	3.9%	77.0%	23.0%	20
68	KSK19 T88A	3.8%	80.4%	19.6%	19
69	RQA	3.5%	100.0%	0.0%	18
70	KT2 F87I	3.5%	81.8%	18.2%	17
71	QP	3.4%	81.6%	18.4%	17
72	RT2AI L437LV	3.0%	72.6%	27.4%	15
73	KSK43	2.5%	77.5%	22.5%	13
74	KSK44 a	2.5%	78.3%	21.7%	12
75	A328P	2.4%	100.0%	0.0%	12
76	A328V	2.4%	79.2%	20.8%	12
77	RT2AP	2.4%	100.0%	0.0%	12
78	A328S P329S Q403P	2.3%	70.4%	29.6%	12
79	WT	2.3%	74.4%	25.6%	11
80	RT2APAI S72A	2.2%	100.0%	0.0%	11

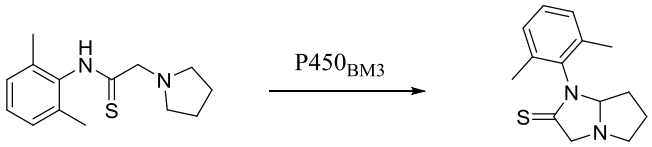
81	KSK19AI	2.0%	61.9%	38.1%	10
82	RT2AI L437VL	1.9%	66.3%	33.7%	10
83	KSK19AI V78I	1.8%	100.0%	0.0%	9
84	F87P I401P	1.7%	52.9%	47.1%	9
85	KSK48	1.6%	100.0%	0.0%	8
86	IP QP	1.6%	100.0%	0.0%	8
87	F87A T88A I401P	1.5%	53.0%	47.0%	8
88	RPAPPV	1.4%	55.7%	44.3%	7
89	RLYF	1.4%	59.6%	40.4%	7
90	RT2A330W I263A	1.3%	49.5%	50.5%	6
91	I263V	1.2%	42.7%	57.3%	6
92	T88M IP	1.2%	100.0%	0.0%	6
93	I263F	1.2%	63.3%	36.7%	6
94	A328G	1.1%	50.7%	49.3%	5
95	A330W	1.0%	46.9%	53.1%	5
96	N70S	1.0%	57.3%	42.7%	5
97	RT2APV78I	0.9%	100.0%	0.0%	4
98	A330P	0.8%	100.0%	0.0%	4

(a) The initial screen was conducted using 2 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.7 Substrate conversion of the initial screen towards sulphur pyrrolidine derivative



N-(2,6-dimethylphenyl)-2-(pyrrolidin-1-yl)ethanethioamide

1-(2,6-dimethylphenyl)tetrahydro-1*H*-pyrrolo[1,2-*a*]imidazole-2(3*H*)-thione, **5.07**

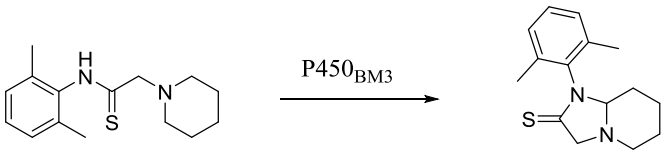
No	enzyme	conversion	TON
1	RPFVEV	95%	1898
2	RPFVEV L437VL	90%	1806
3	RT2AP	42%	831
4	KSK19AI F87V	38%	766
5	RPIAEV	29%	585
6	RT2IP	26%	514
7	RT2APA184I	25%	493
8	RP/QP	24%	481
9	RT2	19%	377
10	RT2AP A82L	16%	326
11	RPAP	16%	311
12	RPAPAM	15%	308
13	RPFVEV A184I	14%	288
14	RP EV	14%	280
15	RT2APAM	14%	274
16	RT2AP/I401M	13%	268
17	RP FV	13%	267
18	RP	13%	260
19	R19	12%	237
20	RT2APV78I	10%	208
21	KSK19 I263A	9%	171
22	KSK19AIAI	8%	163
23	RLYF KSK19 AI	8%	158
24	KSK19AI V78I	8%	158
25	KSK19AI	7%	143
26	KSK19AI F42L	6%	112
27	KSK19AI Y51L	4%	72
28	KSK19 A82M	3%	70

(a) The initial screen was conducted using 1 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.8 Substrate conversion of the initial screen towards sulphur piperidine derivative



N-(2,6-dimethylphenyl)-2-(piperidin-1-yl)ethanethioamide

1-(2,6-dimethylphenyl)hexahydroimidazo[1,2-*a*]pyridine-2(3*H*)-thione, **5.08**

No	Enzymes	Conversion	Cyclisation Product	others	TON
1	KSK19 QP F87V	83.7%	81.9%	18.1%	837
2	GVQ	55.8%	86.6%	13.4%	558
3	RT2AP/E267L	54.9%	90.6%	9.4%	549
4	RPFVEV L188Q	53.5%	75.0%	25.0%	535
5	RT2F81W	53.0%	61.3%	38.7%	530
6	KSK19AI F87V	50.7%	100.0%	0.0%	507
7	RT2AP/A82F	48.8%	90.8%	9.2%	488
8	RP/E267M	43.8%	90.1%	9.9%	438
9	RPFVEV V78F	41.5%	78.2%	21.8%	415
10	RT2IP A184I	39.8%	54.2%	45.8%	398
11	RPFVEV	33.1%	79.8%	20.2%	331
12	RT2AP/E267F	32.1%	94.7%	5.3%	321
13	EVVILI	32.0%	87.5%	12.5%	320
14	F87A I401P	31.4%	44.2%	55.8%	314
15	RPF81W	31.3%	56.8%	43.2%	313
16	RLYF KSK19	28.2%	29.5%	70.5%	282
17	KSK19 A82M	27.6%	74.4%	25.6%	276
18	KSK19	27.4%	37.5%	62.5%	274
19	RP FV	27.1%	81.5%	18.5%	271
20	RP EV	25.0%	91.1%	8.9%	250
21	KSK19 QP	24.9%	39.2%	60.8%	249
22	RPFVEVF81W	23.9%	77.1%	22.9%	239
23	RP/F87A/A328I	23.5%	62.6%	37.4%	235
24	RT2 A330W A82M	23.3%	100.0%	0.0%	233
25	KSK19/A82M/I401M	22.8%	57.2%	42.8%	228
26	RT2F81W A328I	22.8%	100.0%	0.0%	228
27	RT2APAM	22.7%	100.0%	0.0%	227
28	RT2AP A82L	22.1%	100.0%	0.0%	221
29	RT2APA184I F81W	20.8%	100.0%	0.0%	208
30	RT2APAMAI	17.5%	100.0%	0.0%	175
31	RT2APA184I	15.7%	100.0%	0.0%	157
32	RPFVEV L437VL	14.1%	100.0%	0.0%	141
33	RT2 A330W	13.8%	56.3%	43.7%	138
34	QP A82M	13.6%	100.0%	0.0%	136
35	KSK19 F87S	13.4%	55.5%	44.5%	134

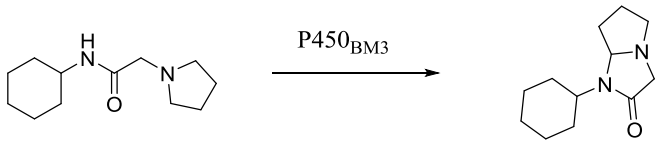
36	RP/IA	12.9%	55.2%	44.8%	129
37	KSK19 T88A	12.7%	64.8%	35.2%	127
38	RPAPAM	12.2%	100.0%	0.0%	122
39	RT2AP	12.0%	92.2%	7.8%	120
40	RP AM IA	11.2%	100.0%	0.0%	112
41	RPFVEV L437LV	11.0%	100.0%	0.0%	110
42	RT2AP/I401M	10.9%	100.0%	0.0%	109
43	RT2AP VIAI	10.8%	100.0%	0.0%	108
44	RT2AI L437LV	9.8%	100.0%	0.0%	98
45	RT2IP	9.7%	80.5%	19.5%	97
46	RPIAEV	9.5%	54.9%	45.1%	95
47	KSK19 AI IA	8.8%	17.1%	82.9%	88
48	KSK19AI Y51L	8.3%	76.4%	23.6%	83
49	RT2AI L437VL	7.7%	100.0%	0.0%	77
50	RT2APV78I	7.7%	100.0%	0.0%	77
51	KSK19AI F42L	7.4%	100.0%	0.0%	74
52	KT2	6.9%	100.0%	0.0%	69
53	RLYF KSK19 AI	6.7%	100.0%	0.0%	67
54	RP	6.6%	100.0%	0.0%	66
55	KSK19AIAI	6.6%	100.0%	0.0%	66
56	KSK19AI	6.6%	100.0%	0.0%	66
57	RT2APPV	6.4%	100.0%	0.0%	64
58	RPAP	6.2%	48.9%	51.1%	62
59	RT2A330W/S72R	5.6%	100.0%	0.0%	56
60	QP	5.4%	100.0%	0.0%	54
61	RT2A330W/S72L	5.2%	100.0%	0.0%	52
62	KSK19 I263A	5.2%	69.2%	30.8%	52
63	KSK43	5.1%	100.0%	0.0%	51
64	RT2	5.1%	100.0%	0.0%	51
65	R19	5.0%	100.0%	0.0%	50
66	F87A	4.7%	100.0%	0.0%	47
67	RP/QP	4.7%	100.0%	0.0%	47
68	F87P I401P	4.6%	100.0%	0.0%	46
69	RPFVEV A184I	4.6%	74.5%	25.5%	46
70	KSK48	4.5%	100.0%	0.0%	45
71	RQA	4.4%	100.0%	0.0%	44
72	KSK19AI V78I	4.1%	100.0%	0.0%	41
73	KSK44 a	4.0%	100.0%	0.0%	40
74	F87A T88A I401P	3.6%	100.0%	0.0%	36
75	RPAPPV	3.4%	100.0%	0.0%	34
76	WT	3.4%	75.4%	24.6%	34
77	A330P	2.6%	100.0%	0.0%	26

(a) The initial screen was conducted using 1 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.9 Substrate conversion of the initial screen towards cyclohexyl derivative



N-cyclohexyl-2-(pyrrolidin-1-yl)acetamide

1-cyclohexyltetrahydro-1*H*-pyrrolo[1,2-*a*]imidazol-2(3*H*)-one, **5.09**

No	Enzymes	Conversion	Cyclisation product	Others	TON
1	RPFVEV V78F	93.4%	73.7%	19.4%	467
2	RT2F81W	90.2%	61.0%	37.0%	451
3	RPIAEV	80.7%	22.7%	54.8%	403
4	RPFVEV L188Q	79.5%	78.9%	21.1%	397
5	RP EV	72.5%	26.7%	64.3%	363
6	RPFVEV	69.8%	74.7%	23.1%	349
7	RP/E267I	66.8%	39.5%	54.2%	334
8	RPFVEV L437LV	66.1%	70.8%	29.2%	331
9	RP H171L	51.8%	80.8%	19.2%	259
10	RPFVEV L437VL	51.8%	89.7%	10.3%	259
11	RPFVEVF81W	51.0%	86.8%	13.2%	255
12	KT2 E267L	45.0%	69.4%	30.6%	225
13	RP/E267M	43.2%	63.5%	30.4%	216
14	EVVILI	42.6%	65.1%	34.9%	213
15	RT2IP/F87V	35.6%	79.5%	20.5%	178
16	RP FV	30.1%	81.4%	18.6%	151
17	RP FV EV A330W	27.5%	82.7%	17.3%	137
18	KSK19 QP F87V	24.8%	88.5%	11.5%	124
19	RP FV EV A330W repeat	23.6%	86.8%	13.2%	118
20	RT2IP A184I	16.9%	64.3%	35.7%	85
21	RT2 FV EV IA A330W	15.7%	100.0%	0.0%	78
22	KSK19/A82M/E267F	15.3%	100.0%	0.0%	76
23	RP IA EV H171L	14.3%	72.2%	27.8%	72
24	RT2IP	13.6%	100.0%	0.0%	68
25	RP	11.5%	100.0%	0.0%	58
26	GVQ	11.0%	100.0%	0.0%	55
27	KT2E267F	10.5%	100.0%	0.0%	52
28	RPF81W	10.4%	100.0%	0.0%	52
29	RT2APAMAI	6.6%	100.0%	0.0%	33
30	RT2AP/A82F	5.8%	100.0%	0.0%	29
31	QP A82M	5.7%	100.0%	0.0%	28
32	RT2AP/E267L	5.1%	68.3%	31.7%	25
33	RP/FA	3.6%	100.0%	0.0%	18
34	KSK19/A82M/I401M	3.4%	100.0%	0.0%	17
35	RT2	3.2%	100.0%	0.0%	16

36	KSK19 A82M	3.2%	100.0%	0.0%	16
37	RT2 I263W	3.1%	100.0%	0.0%	16
38	RT2APAM	2.9%	100.0%	0.0%	15
39	RP/IA	2.8%	100.0%	0.0%	14
40	RP AM IA	2.6%	100.0%	0.0%	13
41	KSK19 QP	2.5%	100.0%	0.0%	13
42	RT2AP A82L	2.4%	100.0%	0.0%	12
43	RT2 A330W A82M	2.3%	100.0%	0.0%	12
44	RPAP	2.1%	100.0%	0.0%	10
45	R19	2.0%	100.0%	0.0%	10
46	KSK19AI F87V	2.0%	100.0%	0.0%	10
47	RT2 V78F	1.8%	100.0%	0.0%	9
48	KSK19AIAI	1.8%	100.0%	0.0%	9
49	RT2APA184I	1.7%	100.0%	0.0%	9
50	F87A I401P	1.7%	100.0%	0.0%	8
51	RT2APPV	1.6%	100.0%	0.0%	8
52	RT2AP/E267F	1.6%	100.0%	0.0%	8
53	RLYF KSK19	1.6%	100.0%	0.0%	8
54	KSK19	1.6%	100.0%	0.0%	8
55	KSK19AI Y51L	1.4%	100.0%	0.0%	7
56	RP/F87A/A328I	1.3%	100.0%	0.0%	7
57	KT2A82G	1.3%	100.0%	0.0%	7
58	KSK19 I263A	1.3%	100.0%	0.0%	6
59	RPAPAM	1.3%	100.0%	0.0%	6
60	RT2APA184I F81W	1.2%	100.0%	0.0%	6
61	KT2	1.2%	100.0%	0.0%	6
62	RLYF KSK19 AI	1.1%	100.0%	0.0%	6
63	RT2AP	1.0%	100.0%	0.0%	5
64	KSK19 F87S	1.0%	100.0%	0.0%	5
65	KSK19AI V78I	0.5%	100.0%	0.0%	3
66	RT2APV78I	0.5%	100.0%	0.0%	2
67	KSK19AI	0.0%	0.0%	0.0%	0
68	RT2AP/I401M	0.0%	0.0%	0.0%	0
69	WT	0.0%	0.0%	0.0%	0
70	KSK48	0.0%	0.0%	0.0%	0
71	RT2AP VIAI	0.0%	0.0%	0.0%	0
72	RT2F81W A328I	0.0%	0.0%	0.0%	0
73	KSK19 T88A	0.0%	0.0%	0.0%	0
74	F87A T88A I401P	0.0%	0.0%	0.0%	0
75	KSK19 AI IA	0.0%	0.0%	0.0%	0
76	RT2AI L437VL	0.0%	0.0%	0.0%	0
77	RT2AI L437LV	0.0%	0.0%	0.0%	0
78	KSK44 a	0.0%	0.0%	0.0%	0
79	A330P	0.0%	0.0%	0.0%	0
80	RQA	0.0%	0.0%	0.0%	0
81	F87P I401P	0.0%	0.0%	0.0%	0

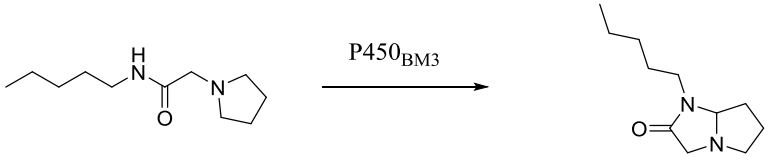
82	RT2 A330W	0.0%	0.0%	0.0%	0
83	QP	0.0%	0.0%	0.0%	0
84	F87A	0.0%	0.0%	0.0%	0
85	RPAPPV	0.0%	0.0%	0.0%	0
86	KSK43	0.0%	0.0%	0.0%	0
87	A328V	0.0%	0.0%	0.0%	0
88	RLYF	0.0%	0.0%	0.0%	0
89	KT2 F87I	0.0%	0.0%	0.0%	0
90	A328S P329S Q403P	0.0%	0.0%	0.0%	0
91	RT2A330W I263A	0.0%	0.0%	0.0%	0
92	A328G	0.0%	0.0%	0.0%	0
93	A330W	0.0%	0.0%	0.0%	0
94	I263V	0.0%	0.0%	0.0%	0
95	I263F	0.0%	0.0%	0.0%	0
96	N70S	0.0%	0.0%	0.0%	0
97	A328P	0.0%	0.0%	0.0%	0
98	RT2APAI S72A	0.0%	0.0%	0.0%	0
99	T88M IP	0.0%	0.0%	0.0%	0
100	IP QP	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.10 Substrate conversion of the initial screen towards n-pentyl derivative



N-pentyl-2-(pyrrolidin-1-yl)acetamide

1-pentyltetrahydro-1*H*-pyrrolo[1,2-*a*]imidazol-2(3*H*)-one, **5.10**

No	Enzymes	Conversion	Cyclisation product	Others	TON
1	RT2F81W	85.5%	65.4%	34.6%	427
2	RPFVEV L188Q	71.5%	81.3%	18.7%	358
3	RPFVEV	68.4%	79.4%	20.6%	342
4	RPFVEV V78F	68.2%	84.5%	15.5%	341
5	RP H171L	60.7%	86.1%	13.9%	304
6	RP/E267I	55.5%	70.8%	29.2%	278
7	RP EV	44.7%	51.4%	48.6%	224
8	RPFVEV L437VL	40.1%	100.0%	0.0%	200
9	RPFVEVF81W	39.6%	100.0%	0.0%	198
10	RPFVEV L437LV	39.4%	100.0%	0.0%	197
11	RT2IP/F87V	27.0%	100.0%	0.0%	135
12	EVVILI	26.6%	100.0%	0.0%	133
13	KSK19/A82M/E267F	25.9%	100.0%	0.0%	130
14	KT2 E267L	25.2%	100.0%	0.0%	126
15	RP/E267M	22.9%	100.0%	0.0%	115
16	RP IA EV H171L	22.7%	100.0%	0.0%	113
17	RT2IP A184I	22.6%	100.0%	0.0%	113
18	RT2APAMAI	21.4%	100.0%	0.0%	107
19	RP FV EV A330W	21.2%	100.0%	0.0%	106
20	RT2 FV EV IA A330W	21.0%	100.0%	0.0%	105
21	RP FV	20.9%	100.0%	0.0%	105
22	RP	18.8%	100.0%	0.0%	94
23	KT2 F87I E267F	18.4%	100.0%	0.0%	92
24	RT2IP	17.7%	100.0%	0.0%	89
25	RT2 F87A A330W	15.4%	100.0%	0.0%	77
26	RT2AP/A82F	15.3%	100.0%	0.0%	77
27	KSK19 QP F87V	14.3%	100.0%	0.0%	72
28	RPF81W	13.9%	100.0%	0.0%	69
29	KT2E267F	13.5%	100.0%	0.0%	68
30	QP A82M	13.1%	100.0%	0.0%	65
31	RT2AP A82L	12.9%	100.0%	0.0%	65
32	GVQ	12.7%	100.0%	0.0%	63
33	RT2APA184I	11.6%	100.0%	0.0%	58
34	RT2 A330W A82M	11.0%	100.0%	0.0%	55
35	RP/FA	10.7%	100.0%	0.0%	53

36	RT2APAM	10.0%	100.0%	0.0%	50
37	KSK19/A82M/I401M	10.0%	100.0%	0.0%	50
38	KSK19 QP	9.7%	100.0%	0.0%	48
39	RLYF KSK19	9.6%	100.0%	0.0%	48
40	RT2AP/E267L	7.3%	100.0%	0.0%	36
41	RT2F81W A328I	7.2%	100.0%	0.0%	36
42	KSK19 A82M	6.6%	100.0%	0.0%	33
43	KSK19	6.5%	100.0%	0.0%	32
44	RP AM IA	5.7%	100.0%	0.0%	28
45	RT2AP/E267F	5.3%	100.0%	0.0%	26
46	RT2 I263W	5.2%	100.0%	0.0%	26
47	R19	4.7%	100.0%	0.0%	24
48	RT2	4.5%	100.0%	0.0%	23
49	RP/IA	4.5%	100.0%	0.0%	22
50	KSK19AIAI	4.4%	100.0%	0.0%	22
51	F87A I401P	4.0%	100.0%	0.0%	20
52	RPAP	3.9%	100.0%	0.0%	19
53	RPAPAM	3.8%	100.0%	0.0%	19
54	RT2AP VIAI	3.7%	100.0%	0.0%	18
55	RT2APA184I F81W	3.4%	100.0%	0.0%	17
56	KSK19AI F87V	3.1%	100.0%	0.0%	15
57	RT2APPV	2.5%	100.0%	0.0%	13
58	RT2 V78F	2.4%	100.0%	0.0%	12
59	A328P	2.2%	100.0%	0.0%	11
60	KSK19 I263A	2.1%	100.0%	0.0%	11
61	RP/F87A/A328I	2.1%	100.0%	0.0%	11
62	KT2A82G	2.1%	100.0%	0.0%	10
63	KSK19AI Y51L	1.9%	100.0%	0.0%	10
64	A328V	1.9%	100.0%	0.0%	10
65	KT2	1.9%	100.0%	0.0%	9
66	RT2AP/I401M	1.6%	100.0%	0.0%	8
67	KSK19 F87S	1.6%	100.0%	0.0%	8
68	RT2AP	1.6%	100.0%	0.0%	8
69	RT2 A330W	1.5%	100.0%	0.0%	7
70	RT2APAI S72A	1.4%	100.0%	0.0%	7
71	KSK44 a	1.4%	100.0%	0.0%	7
72	RT2AI L437LV	1.3%	100.0%	0.0%	6
73	RLYF KSK19 AI	1.2%	100.0%	0.0%	6
74	RQA	1.2%	100.0%	0.0%	6
75	T88M IP	1.1%	100.0%	0.0%	5
76	KSK19 T88A	1.0%	100.0%	0.0%	5
77	IP QP	1.0%	100.0%	0.0%	5
78	KT2 F87I	1.0%	100.0%	0.0%	5
79	QP	0.8%	100.0%	0.0%	4
80	F87P I401P	0.6%	100.0%	0.0%	3
81	F87A T88A I401P	0.5%	100.0%	0.0%	2

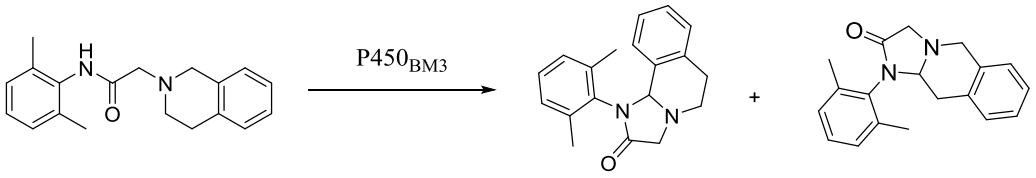
82	RT2AI L437VL	0.5%	100.0%	0.0%	2
83	KSK19 AI IA	0.5%	100.0%	0.0%	2
84	KSK19AI	0.4%	100.0%	0.0%	2
85	F87A	0.4%	100.0%	0.0%	2
86	WT	0.4%	100.0%	0.0%	2
87	RT2APV78I	0.4%	100.0%	0.0%	2
88	KSK48	0.4%	100.0%	0.0%	2
89	KSK19AI V78I	0.4%	100.0%	0.0%	2
90	A328S P329S Q403P	0.4%	100.0%	0.0%	2
91	RPIAEV	0.0%	0.0%	0.0%	0
92	A330P	0.0%	0.0%	0.0%	0
93	RPAPPV	0.0%	0.0%	0.0%	0
94	KSK43	0.0%	0.0%	0.0%	0
95	RLYF	0.0%	0.0%	0.0%	0
96	RT2A330W I263A	0.0%	0.0%	0.0%	0
97	A328G	0.0%	0.0%	0.0%	0
98	A330W	0.0%	0.0%	0.0%	0
99	I263V	0.0%	0.0%	0.0%	0
100	I263F	0.0%	0.0%	0.0%	0
101	N70S	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.11 Substrate conversion of the initial screen towards isotetrahydroquinoline derivative



isotetrahydroquinoline deriative

1-(2,6-dimethylphenyl)-1,5,6,10b-tetrahydroimidazo[2,1-*a*]isoquinolin-2(3*H*)-one, **5.11**
benzylic position

1-(2,6-dimethylphenyl)-1,5,10,10a-tetrahydroimidazo[1,2-*b*]isoquinolin-2(3*H*)-one, **5.12**
non-benzylic position

No	Enzymes	Conversion	Benzylic	Non-benzylic	Others	TON
1	RP H171L	97%	86%	10%	4%	483
2	RPIAEV	94%	14%	77%	10%	468
3	RPFVEV	92%	22%	68%	11%	459
4	RT2IP/F87V	84%	43%	42%	15%	420
5	RPFVEV L188Q	69%	15%	75%	9%	345
6	RT2F81W	58%	40%	41%	19%	292
7	EVVILI	58%	34%	42%	23%	291
8	RT2IP	45%	4%	54%	42%	224
9	RPFVEV V78F	42%	4%	36%	60%	212
10	RPFVEVF81W	41%	9%	72%	20%	206
11	RT2AP/E267L	36%	25%	66%	9%	182
12	RP/E267M	32%	0%	85%	15%	161
13	KSK19AI	31%	0%	0%	100%	155
14	RPF81W	25%	0%	94%	6%	125
15	RPFVEV L437VL	24%	39%	34%	27%	120
16	RT2AP/A82F	23%	3%	19%	78%	113
17	RPFVEV L437LV	22%	9%	19%	72%	111
18	IP QP	22%	95%	2%	3%	109
19	KT2E267F	20%	11%	89%	0%	101
20	KSK19AIAI	18%	33%	67%	0%	91
21	RP	16%	0%	47%	53%	80
22	RT2AP A82L	16%	0%	11%	89%	78
23	RP FV	15%	17%	83%	0%	77
24	KSK19/A82M/E267F	15%	6%	76%	18%	76
25	GVQ	15%	5%	79%	16%	76
26	RT2APA184I	12%	12%	24%	63%	59
27	KSK19 QP F87V	11%	2%	44%	54%	56
28	RT2AP/E267F	8%	9%	81%	10%	40
29	RT2APA184I F81W	8%	7%	83%	9%	39
30	RT2APAM	7%	25%	17%	59%	34
31	RP/FA	6%	0%	100%	0%	29
32	RT2 A330W	5%	0%	100%	0%	27

33	RT2 I263W	5%	5%	78%	17%	26
34	RT2APV78I	5%	38%	13%	49%	25
35	QP A82M	5%	0%	100%	0%	24
36	KSK19AI V78I	5%	70%	30%	0%	23
37	F87A I401P	4%	0%	100%	0%	19
38	RT2	3%	0%	70%	30%	17
39	RT2AP VIAI	3%	0%	80%	20%	16
40	RT2AP	3%	74%	26%	0%	16
41	KSK19 A82M	3%	48%	52%	0%	15
42	KSK19 QP	3%	0%	100%	0%	15
43	KT2	3%	0%	73%	27%	15
44	RT2 A330W A82M	2%	0%	67%	33%	12
45	KT2A82G	2%	0%	75%	25%	11
46	KSK19	2%	0%	100%	0%	11
47	RP AM IA	2%	0%	100%	0%	11
48	KSK19AI F87V	2%	60%	40%	0%	10
49	R19	2%	21%	51%	28%	9
50	RPAP	2%	21%	79%	0%	9
51	RLYF	1%	100%	0%	0%	7
52	KSK19/A82M/I401M	1%	31%	69%	0%	6
53	RT2APAMAI	1%	0%	43%	57%	6
54	RT2APPV	1%	0%	100%	0%	6
55	A330P	1%	56%	44%	0%	5
56	N70S	1%	0%	100%	0%	5
57	KSK19 AI IA	1%	0%	100%	0%	4
58	KSK19 I263A	1%	0%	100%	0%	4
59	KSK19 F87S	1%	0%	100%	0%	4
60	KSK44 a	1%	0%	100%	0%	3
61	KSK43	1%	0%	100%	0%	3
62	RP/IA	1%	0%	100%	0%	3
63	RT2A330W I263A	1%	0%	100%	0%	3
64	KSK48	1%	0%	100%	0%	3
65	I263F	1%	0%	100%	0%	3
66	KSK19 T88A	1%	0%	100%	0%	3
67	RP/F87A/A328I	1%	0%	100%	0%	3
68	RT2F81W A328I	0%	0%	100%	0%	2
69	KT2 F87I	0%	0%	100%	0%	2
70	RLYF KSK19 AI	0%	0%	0%	0%	0
71	KSK19AI Y51L	0%	0%	0%	0%	0
72	RPAPAM	0%	0%	0%	0%	0
73	RT2AP/I401M	0%	0%	0%	0%	0
74	WT	0%	0%	0%	0%	0
75	F87A T88A I401P	0%	0%	0%	0%	0
76	RT2AI L437VL	0%	0%	0%	0%	0
77	RT2AI L437LV	0%	0%	0%	0%	0
78	RQA	0%	0%	0%	0%	0

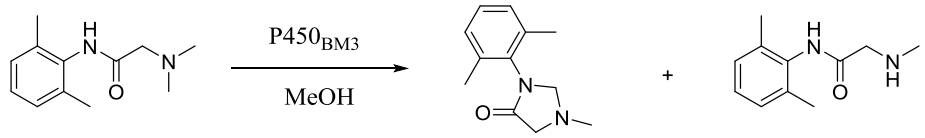
79	F87P I401P	0%	0%	0%	0%	0
80	QP	0%	0%	0%	0%	0
81	F87A	0%	0%	0%	0%	0
82	RLYF KSK19	0%	0%	0%	0%	0
83	RPAPPV	0%	0%	0%	0%	0
84	RT2 V78F	0%	0%	0%	0%	0
85	A328V	0%	0%	0%	0%	0
86	A328S P329S Q403P	0%	0%	0%	0%	0
87	A328G	0%	0%	0%	0%	0
88	A330W	0%	0%	0%	0%	0
89	I263V	0%	0%	0%	0%	0
90	A328P	0%	0%	0%	0%	0

(a) The initial screen was conducted using 2 μM enzyme and 1 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.12-1 Substrate conversion of the initial screen towards lidocaine dimethyl derivative (substrate added in MeOH)



2-(dimethylamino)-*N*-(2,6-dimethylphenyl)acetamide, **5.13**

5.14 Cyclisation

5.18 Dealkylation

No	Enzyme	Conversion	Dealkylation	Cyclisation	others	TON
1	RP EV	68.0%	3.0%	76.1%	20.9%	680
2	RPFVEV	64.9%	2.5%	89.8%	7.7%	649
3	RP/E267M	47.8%	5.3%	88.2%	6.4%	478
4	RPFVEV V78F	43.9%	6.0%	90.9%	3.1%	439
5	RP/H171L/I263G	40.5%	2.9%	50.0%	47.1%	405
6	RPFVEV L188Q	40.1%	9.3%	87.0%	3.6%	401
7	RT2IP	37.0%	3.8%	89.2%	7.0%	370
8	RPFVEV L437LV	35.1%	2.5%	94.6%	2.9%	351
9	RP H171L	30.5%	2.1%	60.3%	37.6%	305
10	RPFVEV L437VL	18.1%	0.0%	100.0%	0.0%	181
11	RT2F81W	13.3%	0.0%	52.4%	47.6%	133
12	KSK19/A82M/E267F	9.6%	13.5%	86.5%	0.0%	96
13	RT2APA184I	9.6%	0.0%	90.4%	9.6%	96
14	RT2APAM	9.1%	0.0%	100.0%	0.0%	91
15	RT2APPV	9.0%	0.0%	28.4%	71.6%	90
16	RPAP	8.0%	0.0%	100.0%	0.0%	80
17	RP FV	7.5%	17.2%	82.8%	0.0%	75
18	RP AM IA	7.5%	0.0%	79.2%	20.8%	75
19	RT2 V78F	7.3%	0.0%	26.3%	73.7%	73
20	RP	7.1%	0.0%	49.8%	50.2%	71
21	RLYF KSK19	6.3%	12.3%	87.7%	0.0%	63
22	RT2AP	6.0%	5.0%	64.5%	30.4%	60
23	GVQ	5.5%	0.0%	40.1%	59.9%	55
24	RPIAEV	5.4%	0.0%	100.0%	0.0%	54
25	KSK19AIAI	4.5%	13.2%	64.3%	22.4%	45
26	RT2 A330W	4.4%	0.0%	24.0%	76.0%	44
27	KSK19 QP F87V	4.3%	0.0%	81.9%	18.1%	43
28	RT2	4.2%	0.0%	23.5%	76.5%	42
29	RT2APAI S72A	3.8%	0.0%	100.0%	0.0%	38
30	RT2AP/I401M	2.7%	0.0%	100.0%	0.0%	27
31	RT2/S72W/A330W	2.6%	0.0%	65.4%	34.6%	26
32	RT2AP VIAI	2.0%	0.0%	100.0%	0.0%	20
33	RPF81W	1.3%	0.0%	0.0%	100.0%	13
34	RLYF KSK19 AI	1.2%	0.0%	100.0%	0.0%	12

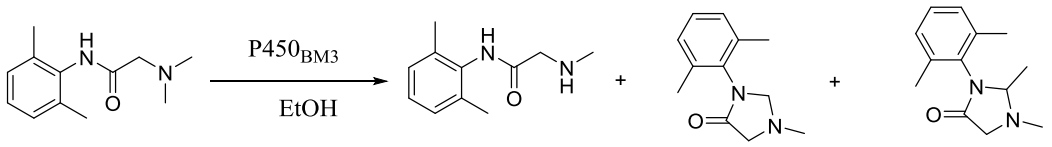
35	KSK19AI F87V	1.1%	0.0%	100.0%	0.0%	11
36	KSK19 A82M	0.5%	0.0%	0.0%	100.0%	5
37	KSK19AI Y51L	0.0%	0.0%	0.0%	0.0%	0
38	KSK19 I263A	0.0%	0.0%	0.0%	0.0%	0
39	KSK19AI V78I	0.0%	0.0%	0.0%	0.0%	0
40	RT2APV78I	0.0%	0.0%	0.0%	0.0%	0
41	WT	0.0%	0.0%	0.0%	0.0%	0
42	A330P	0.0%	0.0%	0.0%	0.0%	0
43	F87A	0.0%	0.0%	0.0%	0.0%	0
44	RLYF	0.0%	0.0%	0.0%	0.0%	0
45	RP/E267G	0.0%	0.0%	0.0%	0.0%	0
46	KSK48	0.0%	0.0%	0.0%	0.0%	0
47	RT2F81W A328I	0.0%	0.0%	0.0%	0.0%	0
48	KSK19 AI IA	0.0%	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.12-2 Substrate conversion of the initial screen towards lidocaine dimethyl derivative (substrate added in EtOH)



2-(dimethylamino)-*N*-(2,6-dimethylphenyl)acetamide, **5.13**

5.18 Dealkylation

5.14 MeOH cyclisation

5.16 EtOH cyclisation

No	Enzyme	Conversion	Dealkylation 5.18	MeOH cyclisation 5.14	EtOH cyclisation 5.16	Others	TON
1	RP EV	64.8%	6.5%	53.8%	23.1%	16.5%	648
2	RPFVEV	64.6%	3.5%	63.3%	27.4%	5.9%	646
3	RP/E267M	38.4%	9.0%	75.6%	12.7%	2.6%	384
4	RP/H171L/I263G	31.9%	6.4%	18.5%	69.9%	5.3%	319
5	RPFVEV V78F	29.0%	10.3%	76.4%	8.9%	4.4%	290
6	RPFVEV L188Q	27.8%	18.4%	25.6%	42.1%	13.8%	278
7	RT2IP	25.6%	9.8%	68.5%	21.7%	0.0%	256
8	RP H171L	22.2%	5.5%	25.3%	69.2%	0.0%	222
9	RT2F81W	16.9%	5.9%	30.9%	63.3%	0.0%	169
10	RP FV	11.4%	14.0%	49.5%	36.5%	0.0%	114
11	RPFVEV L437LV	11.2%	11.6%	56.5%	31.9%	0.0%	112
12	RPIAEV	8.3%	16.8%	46.4%	36.8%	0.0%	83
13	RT2 V78F	7.9%	0.0%	0.0%	77.1%	22.9%	79
14	RT2APPV	7.6%	0.0%	0.0%	100.0%	0.0%	76
15	KSK19/A82M/E267F	7.5%	20.1%	54.5%	25.4%	0.0%	75
16	RP	6.4%	0.0%	23.4%	76.7%	0.0%	64
17	RP AM IA	5.6%	8.9%	18.7%	72.4%	0.0%	56
18	RT2APA184I	5.0%	0.0%	45.2%	54.8%	0.0%	50
19	GVQ	4.5%	0.0%	0.0%	100.0%	0.0%	45
20	RT2	4.5%	0.0%	0.0%	100.0%	0.0%	45
21	RT2 A330W	4.5%	0.0%	0.0%	100.0%	0.0%	45
22	RT2AP	4.4%	0.0%	18.6%	81.2%	0.2%	44
23	RPAP	3.0%	0.0%	51.0%	49.0%	0.0%	30
24	RT2/S72W/A330W	2.2%	0.0%	0.0%	100.0%	0.0%	22
25	KSK19 QP F87V	1.7%	0.0%	0.0%	100.0%	0.0%	17
26	KSK19AIAI	1.2%	0.0%	0.0%	100.0%	0.0%	12
27	RT2APAI S72A	1.0%	0.0%	0.0%	100.0%	0.0%	10
28	RPF81W	1.0%	0.0%	0.0%	0.0%	0.0%	10
29	RLYF KSK19	0.9%	0.0%	0.0%	100.0%	0.0%	9
30	KSK19 A82M	0.6%	0.0%	0.0%	100.0%	0.0%	6
31	RLYF KSK19 AI	0.3%	0.0%	0.0%	100.0%	0.0%	3
32	KSK19AI Y51L	0.0%	0.0%	0.0%	0.0%	0.0%	0
33	RPFVEV L437VL	0.0%	0.0%	0.0%	0.0%	0.0%	0
34	KSK19 I263A	0.0%	0.0%	0.0%	0.0%	0.0%	0

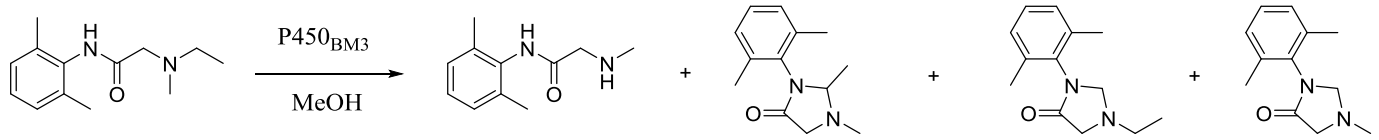
35	KSK19AI F87V	0.0%	0.0%	0.0%	0.0%	0.0%	0
36	KSK19AI V78I	0.0%	0.0%	0.0%	0.0%	0.0%	0
37	RT2AP/I401M	0.0%	0.0%	0.0%	0.0%	0.0%	0
38	RT2APAM	0.0%	0.0%	0.0%	0.0%	0.0%	0
39	RT2APV78I	0.0%	0.0%	0.0%	0.0%	0.0%	0
40	WT	0.0%	0.0%	0.0%	0.0%	0.0%	0
41	A330P	0.0%	0.0%	0.0%	0.0%	0.0%	0
42	F87A	0.0%	0.0%	0.0%	0.0%	0.0%	0
43	RLYF	0.0%	0.0%	0.0%	0.0%	0.0%	0
44	RP/E267G	0.0%	0.0%	0.0%	0.0%	0.0%	0
45	KSK48	0.0%	0.0%	0.0%	0.0%	0.0%	0
46	RT2AP VIAI	0.0%	0.0%	0.0%	0.0%	0.0%	0
47	RT2F81W A328I	0.0%	0.0%	0.0%	0.0%	0.0%	0
48	KSK19 AI IA	0.0%	0.0%	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.13-1 Substrate conversion of the initial screen towards lidocaine methyl ethyl derivative (substrate added in MeOH)



N-(2,6-dimethylphenyl)-2-(ethyl(methyl)amino)acetamide, **5.15**

5.18 Dealkylation

5.16 Ethyl cyclisation

5.17 Methyl Cyclisation

5.14 Methyl MeOH cyclisation

No	Enzyme	Conversion	Dealkylation	Ethyl cyclisation	Methyl cyclisation	Methyl MeOH cyclisation	Others	TON
1	RPFVEV	56.3%	1.1%	74.0%	0.0%	19.0%	5.9%	563
2	RP EV	46.2%	3.9%	44.8%	7.3%	30.6%	13.4%	462
3	RPIAEV	32.8%	5.9%	32.1%	0.0%	49.2%	12.8%	328
4	RP/H171L/I263G	30.6%	1.9%	7.4%	50.7%	18.7%	21.3%	306
5	RP H171L	20.9%	0.0%	14.8%	42.4%	17.1%	25.8%	209
6	RT2F81W	18.0%	2.3%	12.3%	45.6%	21.3%	18.6%	180
7	RP/E267M	16.7%	16.0%	13.5%	5.3%	61.6%	3.7%	167
8	RT2IP	13.5%	8.5%	13.2%	11.7%	63.3%	3.2%	135
9	RT2APPV	12.4%	0.0%	0.0%	33.5%	6.1%	60.4%	124
10	RPFVEV V78F	10.4%	7.4%	54.2%	0.0%	38.3%	0.1%	104
11	RPFVEV L437VL	9.1%	0.0%	56.3%	0.0%	43.7%	0.0%	91
12	RT2	9.0%	0.0%	0.0%	59.4%	8.7%	31.9%	90
13	RPFVEV L437LV	7.7%	0.0%	32.6%	0.0%	67.4%	0.0%	77
14	RPFVEV L188Q	7.5%	0.0%	27.8%	0.0%	59.8%	12.4%	75
15	RP AM IA	6.1%	0.0%	0.0%	44.2%	36.4%	19.4%	61

16	RT2 V78F	5.6%	0.0%	0.0%	59.0%	11.3%	29.7%	56
17	RT2APAM	5.5%	0.0%	0.0%	11.2%	88.8%	0.0%	55
18	RT2 A330W	5.3%	0.0%	0.0%	50.5%	21.7%	27.8%	53
19	RP	5.2%	0.0%	0.0%	56.3%	22.3%	21.3%	52
20	RPAP	3.9%	0.0%	0.0%	0.0%	100.0%	0.0%	39
21	RT2APA184I	3.6%	0.0%	0.0%	20.7%	79.3%	0.0%	36
22	RP FV	3.3%	0.0%	55.1%	0.0%	44.6%	0.3%	33
23	GVQ	3.2%	5.6%	0.0%	47.4%	23.1%	24.0%	32
24	KSK19/A82M/E267F	3.2%	33.1%	0.0%	0.0%	67.2%	0.0%	32
25	RT2/S72W/A330W	2.9%	0.0%	0.0%	44.0%	42.3%	13.7%	29
26	KSK19AIAI	2.9%	16.0%	0.0%	30.6%	41.7%	11.8%	29
27	RT2APAI S72A	2.3%	13.7%	0.0%	13.7%	72.1%	0.4%	23
28	RT2AP/I401M	1.8%	12.6%	0.0%	22.4%	64.5%	0.5%	18
29	RLYF KSK19	1.6%	29.7%	0.0%	0.0%	70.9%	0.0%	16
30	RPF81W	1.5%	0.0%	0.0%	0.0%	44.4%	55.6%	15
31	KSK19 A82M	1.3%	0.0%	0.0%	0.0%	100.0%	0.0%	13
32	KSK48	1.3%	38.3%	0.0%	0.0%	61.7%	0.0%	13
33	RLYF KSK19 AI	1.2%	39.0%	0.0%	0.0%	61.0%	0.0%	12
34	KSK19AI F87V	1.2%	29.9%	0.0%	0.0%	70.9%	0.0%	12
35	RP/E267G	1.2%	0.0%	0.0%	0.0%	100.0%	0.0%	12
36	RT2AP	1.1%	0.0%	0.0%	0.0%	100.0%	0.0%	11
37	WT	1.0%	31.7%	0.0%	0.0%	67.3%	1.0%	10
39	A330P	1.0%	37.0%	0.0%	0.0%	63.0%	0.0%	10
38	RLYF	1.0%	0.0%	0.0%	45.0%	55.0%	0.0%	10
40	KSK19AI Y51L	0.9%	35.2%	0.0%	0.0%	64.8%	0.0%	9
41	F87A	0.8%	29.1%	0.0%	0.0%	70.9%	0.0%	8
42	KSK19 QP F87V	0.8%	0.0%	0.0%	0.0%	100.0%	0.0%	8

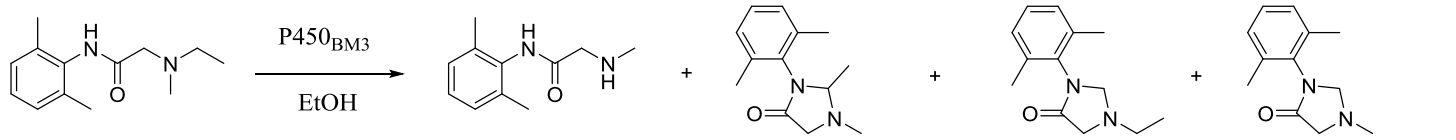
43	RT2F81W A328I	0.7%	0.0%	0.0%	0.0%	100.0%	0.0%	7
44	KSK19 I263A	0.7%	0.0%	0.0%	0.0%	100.0%	0.0%	7
45	RT2AP VIAI	0.6%	0.0%	0.0%	0.0%	100.0%	0.0%	6
46	RT2APV78I	0.5%	0.0%	0.0%	0.0%	100.0%	0.0%	5
47	KSK19AI V78I	0.4%	100.0%	0.0%	0.0%	0.0%	0.0%	4
48	KSK19 AI IA	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.13-2 Substrate conversion of the initial screen towards lidocaine methyl ethyl derivative (substrate added in EtOH)



N-(2,6-dimethylphenyl)-2-(ethyl(methyl)amino)acetamide, **5.15**

5.18 Dealkylation

5.16 Ethyl cyclisation

5.17 Methyl Cyclisation

3.14 Ethyl EtOH cyclisation

No	Enzyme	Conversion	Dealkylation	Ethyl cyclisation	Methyl Cyclisation	Others	TON
1	RPFVEV	47.8%	0.0%	10.5%	66.5%	23.0%	478
2	RP EV	36.4%	6.6%	31.5%	36.2%	25.7%	364
3	RP/H171L/I263G	29.2%	2.1%	66.1%	0.0%	31.9%	292
4	RP H171L	19.3%	0.0%	56.0%	6.1%	38.0%	193
5	RT2F81W	15.6%	1.6%	67.6%	0.0%	30.8%	156
6	RPIAEV	14.5%	24.7%	29.3%	19.4%	26.6%	145
7	RP/E267M	14.0%	18.6%	47.2%	9.3%	24.9%	140
8	RPFVEV L188Q	12.5%	11.5%	13.4%	34.8%	40.3%	125
9	RT2IP	11.1%	10.4%	39.7%	29.4%	20.5%	111
10	RT2APPV	10.5%	0.0%	49.4%	0.0%	50.6%	105
11	RT2	9.7%	0.0%	62.5%	0.0%	37.5%	97
12	RPFVEV V78F	8.0%	10.4%	10.5%	70.5%	8.6%	80
13	RT2 V78F	6.9%	0.0%	67.1%	0.0%	32.9%	69
14	RT2 A330W	6.6%	0.0%	62.0%	0.0%	38.0%	66
15	RP AM IA	6.2%	0.0%	70.8%	0.0%	29.2%	62

16	RP	5.6%	0.0%	69.5%	0.0%	30.5%	56
17	GVQ	4.7%	0.0%	59.1%	0.0%	40.9%	47
18	RT2/S72W/A330W	4.6%	0.0%	70.5%	0.0%	29.5%	46
19	KSK19/A82M/E267F	4.1%	31.3%	50.8%	0.0%	18.0%	41
20	RT2APA184I	2.8%	16.4%	72.6%	0.0%	11.0%	28
21	RPFVEV L437LV	2.7%	0.0%	45.3%	0.0%	54.7%	27
22	RP FV	2.3%	28.1%	51.3%	0.0%	20.6%	23
23	KSK19AIAI	2.2%	10.2%	56.9%	0.0%	32.9%	22
24	KSK19 QP F87V	2.0%	0.0%	67.2%	0.0%	32.8%	20
25	RLYF KSK19	1.8%	21.4%	59.9%	0.0%	18.7%	18
26	KSK19AI F87V	1.7%	9.1%	64.8%	0.0%	26.1%	17
27	WT	1.6%	27.8%	72.8%	0.0%	0.0%	16
28	RT2APAI S72A	1.6%	0.0%	85.8%	0.0%	14.2%	16
29	RLYF	1.5%	0.0%	76.7%	0.0%	23.3%	15
30	KSK19 I263A	1.4%	32.1%	46.7%	0.0%	21.2%	14
31	RPAP	1.4%	0.0%	100.0%	0.0%	0.0%	14
32	RT2AP/I401M	1.3%	0.0%	83.0%	0.0%	17.0%	13
33	A330P	1.2%	31.5%	54.8%	0.0%	13.7%	12
34	KSK48	1.2%	26.2%	57.4%	0.0%	16.4%	12
35	RPF81W	1.2%	0.0%	40.7%	0.0%	59.3%	12
36	RP/E267G	1.0%	0.0%	100.0%	0.0%	0.0%	10
37	KSK19 A82M	0.9%	38.4%	62.8%	0.0%	0.0%	9
38	RT2APAM	0.9%	0.0%	100.0%	0.0%	0.0%	9
39	RLYF KSK19 AI	0.8%	57.8%	42.2%	0.0%	0.0%	8
40	F87A	0.8%	25.0%	75.0%	0.0%	0.0%	8
41	KSK19AI Y51L	0.7%	47.3%	52.7%	0.0%	0.0%	7
42	RT2AP VIAI	0.5%	0.0%	100.0%	0.0%	0.0%	5

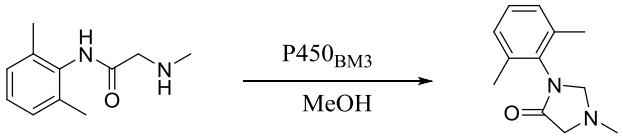
43	KSK19AI V78I	0.4%	100.0%	0.0%	0.0%	0.0%	4
44	RT2AP	0.2%	100.0%	0.0%	0.0%	0.0%	2
45	RPFVEV L437VL	0.0%	0.0%	0.0%	0.0%	0.0%	0
46	RT2APV78I	0.0%	0.0%	0.0%	0.0%	0.0%	0
47	RT2F81W A328I	0.0%	0.0%	0.0%	0.0%	0.0%	0
48	KSK19 AI IA	0.0%	0.0%	0.0%	0.0%	0.0%	0

(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.14-1 Substrate conversion of the initial screen towards lidocaine single methyl derivative (substrate added in MeOH)



N-(2,6-dimethylphenyl)-2-(methylamino)acetamide, **5.18**

5.14
MeOH cyclisation product

No	Enzyme	Conversion	MeOH Cyclisation	Others	TON
1	RPFVEV L437VL	97.4%	97.1%	2.9%	974
2	RPAP	95.2%	98.1%	1.9%	952
3	RT2APAM	93.7%	93.9%	6.1%	937
4	RT2AP/I401M	93.0%	99.6%	0.4%	930
5	RT2 A330W	91.4%	100.0%	0.0%	914
6	RP EV	90.6%	83.4%	16.6%	906
7	RPFVEV L437LV	90.0%	97.1%	2.9%	900
8	RP AM IA	88.5%	100.0%	0.0%	885
9	RP/H171L/I263G	87.9%	84.8%	15.2%	879
10	RPFVEV L188Q	87.1%	75.3%	24.7%	871
11	RPFVEV	87.0%	87.6%	12.4%	870
12	RP H171L	85.1%	96.9%	3.1%	851
13	RT2APA184I	83.9%	99.3%	0.7%	839
14	RPFVEV V78F	81.2%	95.2%	4.8%	812
15	RT2APAI S72A	79.7%	100.0%	0.0%	797
16	RT2IP	77.8%	97.6%	2.4%	778
17	RT2APPV	75.8%	100.0%	0.0%	758
18	KSK19 QP F87V	75.0%	98.6%	1.4%	750
19	RT2F81W A328I	73.2%	100.0%	0.0%	732
20	RP/E267M	71.8%	95.9%	4.1%	718
21	RT2/S72W/A330W	71.4%	100.0%	0.0%	714
22	RT2F81W	71.3%	98.3%	1.7%	713
23	RP/E267G	70.4%	100.0%	0.0%	704
24	RT2 V78F	69.4%	96.4%	3.6%	694
25	RPF81W	68.1%	96.2%	3.8%	681
26	RT2AP VIAI	64.6%	100.0%	0.0%	646
27	RLYF KSK19	57.6%	100.0%	0.0%	576
28	RLYF	53.1%	100.0%	0.0%	531
29	RP FV	52.5%	100.0%	0.0%	525
30	RPIAEV	52.4%	97.8%	2.2%	524
31	KSK19/A82M/E267F	50.9%	91.8%	8.2%	509
32	RT2AP	49.4%	100.0%	0.0%	494
33	RP	48.6%	87.0%	13.0%	486
34	KSK19 AI IA	45.9%	100.0%	0.0%	459

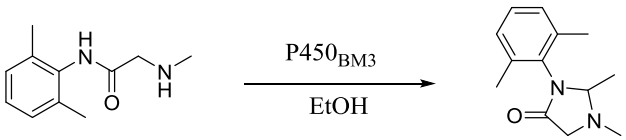
35	RT2	43.3%	100.0%	0.0%	433
36	KSK19AI F87V	42.7%	100.0%	0.0%	427
37	KSK19 I263A	42.3%	91.8%	8.2%	423
38	KSK48	40.5%	100.0%	0.0%	405
39	F87A	40.2%	100.0%	0.0%	402
40	KSK19AIAI	38.8%	100.0%	0.0%	388
41	GVQ	38.3%	95.3%	4.7%	383
42	WT	35.5%	100.0%	0.0%	355
43	RLYF KSK19 AI	35.4%	100.0%	0.0%	354
44	RT2APV78I	34.5%	100.0%	0.0%	345
45	KSK19AI V78I	34.2%	100.0%	0.0%	342
46	A330P	33.5%	100.0%	0.0%	335
47	KSK19AI Y51L	33.5%	100.0%	0.0%	335
48	KSK19 A82M	27.5%	100.0%	0.0%	275

(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.14-2 Substrate conversion of the initial screen towards lidocaine single methyl derivative (substrate added in EtOH)



N-(2,6-dimethylphenyl)-2-(methylamino)acetamide, **5.18**

5.16
EtOH cyclisation product

No	Enzymes	Conversion	EtOH Cyclisation Product	Others	TON
1	RT2	99.3%	93.9%	6.1%	993
2	RT2APPV	98.4%	91.9%	8.1%	984
3	RPFVEV	93.8%	56.7%	43.3%	938
4	RPFVEV L437LV	92.8%	72.3%	27.7%	928
5	RT2F81W	92.1%	86.5%	13.5%	921
6	RT2 V78F	89.9%	87.0%	13.0%	899
7	RP H171L	88.9%	85.1%	14.9%	889
8	RPAP	88.8%	71.5%	28.5%	888
9	RP EV	88.7%	57.5%	42.5%	887
10	RT2APAM	88.4%	60.0%	40.0%	884
11	RPFVEV L437VL	88.1%	58.4%	41.6%	881
12	RPFVEV L188Q	86.4%	43.2%	56.8%	864
13	RT2IP	84.9%	74.5%	25.5%	849
14	KSK19 QP F87V	84.5%	80.9%	19.1%	845
15	KSK19AIAI	83.3%	80.4%	19.6%	833
16	RPIAEV	81.1%	48.3%	51.7%	811
17	RT2APA184I	80.8%	77.1%	22.9%	808
18	KSK19AI F87V	77.8%	86.7%	13.3%	778
19	RP FV	77.6%	69.2%	30.8%	776
20	RT2AP	77.4%	53.5%	46.5%	774
21	RPFVEV V78F	76.6%	52.0%	48.0%	766
22	RP AM IA	76.4%	85.4%	14.6%	764
23	RLYF	75.7%	78.1%	21.9%	757
24	RT2AP VIAI	75.5%	70.8%	29.2%	755
25	RP/E267M	74.0%	63.0%	37.0%	740
26	KSK19 A82M	73.9%	69.9%	30.1%	739
27	KSK19 I263A	72.8%	85.8%	14.2%	728
28	RT2AP/I401M	70.8%	70.9%	29.1%	708
29	RT2APV78I	70.3%	49.8%	50.2%	703
30	RT2F81W A328I	68.6%	55.6%	44.4%	686
31	RPF81W	67.0%	54.7%	45.3%	670
32	RP	66.9%	78.6%	21.4%	669
33	RT2APAI S72A	65.3%	73.7%	26.3%	653

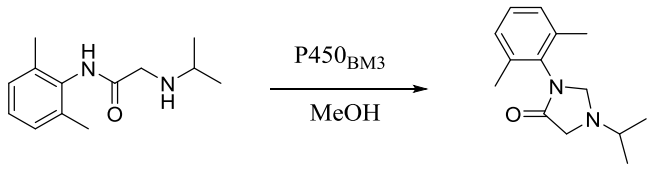
34	RT2/S72W/A330W	64.8%	80.2%	19.8%	648
35	KSK19AI Y51L	63.0%	80.0%	20.0%	630
36	GVQ	62.2%	77.9%	22.1%	622
37	KSK19 AI IA	62.1%	62.9%	37.1%	621
38	KSK48	61.1%	73.3%	26.7%	611
39	RP/H171L/I263G	61.1%	69.6%	30.4%	611
40	WT	61.0%	81.6%	18.4%	610
41	RT2 A330W	59.5%	79.1%	20.9%	595
42	RP/E267G	58.7%	74.3%	25.7%	587
43	RLYF KSK19	54.3%	73.3%	26.7%	543
44	KSK19/A82M/E267F	54.2%	58.1%	41.9%	542
45	F87A	51.6%	71.0%	29.0%	516
46	RLYF KSK19 AI	50.8%	70.0%	30.0%	508
47	A330P	42.6%	63.2%	36.8%	426
48	KSK19AI V78I	39.0%	66.3%	33.7%	390

(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.15-1 Substrate conversion of the initial screen towards lidocaine isopropyl derivative (substrate added in MeOH)



N-(2,6-dimethylphenyl)-2-(isopropylamino)acetamide, **5.19**

5.20
MeOH cyclisation product

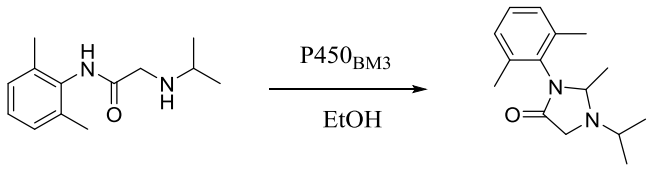
No	enzyme	conversion	MeOH cyclisation product	Others	TON
1	RT2AP/A184I	82.5%	83.8%	16.2%	1651
2	RP/FV/EV	82.5%	95.9%	4.1%	1650
3	RPAP	81.5%	100.0%	0.0%	1629
4	RP/EV	67.5%	95.9%	4.1%	1350
5	RT2AP/V78I	57.7%	89.3%	10.7%	1154
6	KSK19/AI/AI	55.7%	97.7%	2.3%	1113
7	RT2AP/AM	54.6%	78.5%	21.5%	1092
8	RP/FV/EV/A184I	52.4%	93.9%	6.1%	1049
9	RT2AP/IP	50.5%	96.3%	3.7%	1009
10	KSK19/A82M	47.9%	95.9%	4.1%	958
11	RP/FV	47.2%	95.7%	4.3%	945
12	RT2IP	45.3%	93.6%	6.4%	906
13	KSK19/AI/FV	45.0%	94.0%	6.0%	901
14	KSK19/AI/VI	44.1%	94.2%	5.8%	882
15	KSK19/IA	43.8%	97.7%	2.3%	877
16	RP	43.7%	96.9%	3.1%	874
17	RPAPAM	43.3%	91.2%	8.8%	867
18	KSK19/AI	43.1%	97.5%	2.5%	863
19	RLYF/KSK19/AI	42.0%	100.0%	0.0%	841
20	RP/IA/EV	41.5%	90.8%	9.2%	830
21	RT2	37.1%	98.1%	1.9%	741
22	RT2AP	36.6%	92.6%	7.4%	733
23	WT	32.4%	100.0%	0.0%	648

(a) The initial screen was conducted using 1 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

Table S5.15-2 Substrate conversion of the initial screen towards lidocaine isopropyl derivative (substrate added in EtOH)



N-(2,6-dimethylphenyl)-2-(isopropylamino)acetamide, **5.19**

5.21
EtOH cyclisation product

No	Enzyme	Conversion	TON
1	RPFVEV	3.7%	37
2	RT2F81W	2.0%	20
3	RT2	1.8%	18
4	RT2AP	1.7%	17
5	RP	1.3%	13
6	KSK19AI F87V	1.2%	12
7	KSK19AIAI	1.2%	12
8	RT2APV78I	1.0%	10
9	RT2APA184I	0.9%	9
10	RT2IP	0.9%	9
11	KSK19 A82M	0.8%	8
12	WT	0.8%	8
13	RPAP	0.7%	7
14	KSK19AI Y51L	0.7%	7
15	KSK19AI V78I	0.7%	7
16	KSK19 I263A	0.6%	6
17	RLYF KSK19 AI	0.6%	6
18	RP FV	0.5%	5
19	RT2APAM	0.0%	0
20	RPFVEV L437VL	0.0%	0
21	RPFVEV L437LV	0.0%	0
22	RT2AP/I401M	0.0%	0
23	RPIAEV	0.0%	0
24	RP EV	0.0%	0

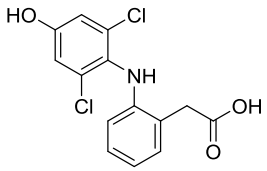
(a) The initial screen was conducted using 2 μM enzyme and 2 mM substrate

(b) Conversion results represent averages of 2 independent experiments with errors not higher than 10%.

(c) TON= Turnover Number. The values represent averages calculated from two independent experiments with 10 % error and is reported as mol product converted per mol enzyme.

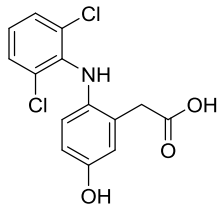
Appendix 6 Compound characterisation data

4′-Hydroxydiclofenac (3.01)



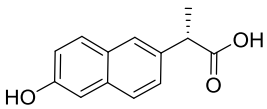
1H NMR (500 MHz, CD3OD) δ 7.91 (s, 1H), 7.18 (dd, J = 7.5, 1.4 Hz, 1H), 7.04 (td, J = 7.8, 1.5 Hz, 1H), 6.89 (s, 2H), 6.82 (td, J = 7.4, 1.1 Hz, 1H), 6.27 (dd, J = 8.1, 1.1 Hz, 1H), 3.72 (s, 2H). 13C NMR (126 MHz, CD3OD) δ 176.32, 156.59, 145.77, 133.91, 132.11, 130.78, 129.12, 124.62, 121.57, 117.17, 116.68, 39.65. HRMS (APCI) [M+Na+] Calcd for C14H11Cl2NO3: 334.0008, Found: 334.0013.

5-Hydroxydiclofenac (3.02)



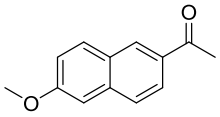
1H NMR (500 MHz, CD3OD) δ 7.34 (d, J = 8.1 Hz, 1H), 6.96 (t, J = 8.1 Hz, 1H), 6.70 (d, J = 2.8 Hz, 1H), 6.54 (dd, J = 8.6, 2.8 Hz, 1H), 6.36 (d, J = 8.6 Hz, 1H), 3.70 (s, 1H). 13C NMR (126 MHz, CD3OD) δ 176.20, 154.39, 140.84, 136.63, 130.32, 129.69, 129.44, 122.12, 118.59, 115.49, 39.58. HRMS (APCI) [M+Na+] Calcd for C14H11Cl2NO3: 334.0008, Found: 334.0013

6-desmethylnaproxen (3.03)



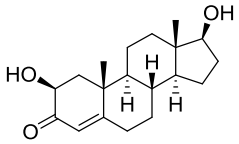
1H NMR (500 MHz, CD3OD) δ = 7.69 – 7.57 (m, 3H), 7.35 (dd, J=8.5, 1.9, 1H), 7.09 – 7.01 (m, 2H), 3.81 (q, J=7.1, 1H), 1.51 (d, J=7.3, 3H). 13C NMR (126 MHz, CD3OD) δ = 177.18, 155.01, 135.40, 134.09, 128.89, 128.45, 126.13, 125.67, 125.53, 118.10, 108.34, 45.16, 17.63. HRMS (APCI) [M+Na+] Calcd for C12H18N2O: 239.0683, Found: 239.0679

2-acetyl-6-methoxynaphthalene (3.04)



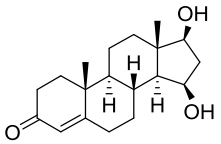
1H NMR (500 MHz, CDCl3) δ = 8.41 (d, J=2.0, 1H), 8.02 (dd, J=8.7, 1.9, 1H), 7.87 (d, J=8.8, 1H), 7.78 (d, J=8.7, 1H), 7.22 (dd, J=9.0, 2.5, 1H), 7.17 (d, J=2.5, 1H), 3.96 (s, 3H), 2.71 (s, 3H). 13C NMR (126 MHz, CDCl3) δ = 198.09, 159.96, 137.48, 132.82, 131.32, 128.01, 127.29, 124.87, 119.93, 105.94, 55.63, 26.76. HRMS (APCI) [M+Na+] Calcd for C13H12O3: 233.0730, Found: 233.0731

2β-Hydroxytestosterone (3.05)



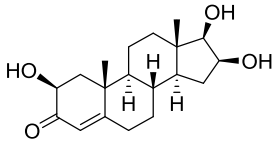
1H NMR (500 MHz, CDCl3) δ = 5.81 (d, J=1.2, 1H), 4.19 (dd, J=13.9, 5.6, 1H), 3.67 (t, J=8.6, 1H), 3.51 – 3.46 (m, 1H), 2.57 – 2.46 (m, 2H), 2.29 – 2.23 (m, 1H), 2.18 (s, 3H), 2.08 (tdd, J=9.3, 7.6, 4.2, 1H), 2.02 – 1.95 (m, 1H), 1.92 – 1.86 (m, 1H), 1.80 (ddd, J=13.4, 7.0, 3.8, 1H), 1.71 (ddd, J=22.6, 10.7, 4.1, 1H), 1.64 – 1.37 (m, 9H), 1.36 – 1.25 (m, 2H), 1.19 (s, 3H), 1.14 (td, J=12.8, 4.1, 1H), 1.06 – 0.96 (m, 2H), 0.80 (s, 3H). 13C NMR (126 MHz, CDCl3) δ = 199.90, 175.28, 118.86, 81.66, 68.72, 50.67, 50.39, 43.58, 41.64, 39.64, 36.56, 36.05, 34.74, 33.18, 30.65, 23.53, 23.02, 22.73, 11.42. HRMS (APCI) [M+H+] Calcd for C19H28O3: 305.2111, Found: 305.2111

15β-Hydroxytestosterone (3.06)



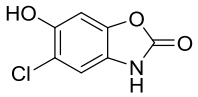
1H NMR (500 MHz, CDCl3) δ = 5.75 (s, 1H), 4.23 (ddd, J=7.8, 5.7, 2.4, 1H), 3.59 (t, J=8.7, 1H), 2.64 (ddd, J=14.6, 8.6, 7.7, 1H), 2.53 – 2.28 (m, 4H), 2.18 (s, 1H), 2.14 – 1.95 (m, 3H), 1.85 (dt, J=12.3, 3.2, 1H), 1.72 (td, J=14.0, 4.7, 1H), 1.66 – 1.51 (m, 9H), 1.46 (ddd, J=26.0, 13.1, 3.9, 2H), 1.24 (d, J=5.8, 3H), 1.18 – 0.96 (m, 7H). 13C NMR (126 MHz, CDCl3) δ = 199.70, 171.12, 124.16, 81.31, 69.33, 55.36, 54.47, 43.63, 42.43, 38.97, 38.06, 35.98, 34.16, 32.86, 31.67, 31.24, 20.76, 17.52, 13.90. HRMS (APCI) [M+H+] Calcd for C19H28O3: 305.2111, Found: 305.2111.

2β,16β-Dihydroxytestosterone (3.07)



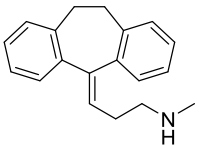
1H NMR (500 MHz, CDCl3) δ = 5.83 (d, J=1.0, 1H), 4.19 (dd, J=13.1, 5.0, 1H), 3.60 (t, J=8.7, 1H), 3.52 – 3.44 (m, 1H), 2.69 – 2.56 (m, 2H), 2.50 (dd, J=13.8, 5.6, 1H), 2.30 (dt, J=7.1, 6.0, 2H), 2.18 (s, 3H), 1.95 – 1.77 (m, 2H), 1.68 – 1.31 (m, 6H), 1.32 – 1.19 (m, 2H), 1.19 – 1.04 (m, 4H). 13C NMR (126 MHz, CDCl3) δ = 199.69, 174.97, 118.71, 81.02, 69.23, 68.54, 55.16, 50.69, 43.62, 42.83, 41.62, 39.53, 37.89, 33.99, 32.95, 32.01, 22.84, 22.51, 13.92. HRMS (APCI) [M+H+] Calcd for C19H29O4: 321.2060, Found: 321.2062

6-Hydroxychlorzoxazone (3.08)



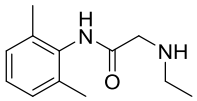
1H NMR (500 MHz, CD3OD) δ = 6.94 (s, 1H), 6.75 (s, 1H). 13C NMR (126 MHz, CD3OD) δ = 155.92, 148.86, 143.29, 123.25, 115.57, 110.21, 98.82. HRMS (APCI) [M+Na+] Calcd for C7H4ClNO3: 207.9772, Found: 207.9767

Nortriptyline (3.09)



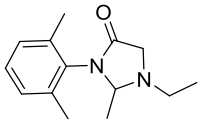
1H NMR (400 MHz, CDCl3) δ 7.34 – 7.28 (m, 1H), 7.24 – 7.11 (m, 6H), 7.08 – 7.03 (m, 1H), 5.87 (t, J = 7.4 Hz, 1H), 3.48 – 3.25 (m, 2H), 3.08 – 2.90 (m, 1H), 2.79 (d, J = 8.4 Hz, 1H), 2.67 (t, J = 6.9 Hz, 2H), 2.40 (s, 3H), 2.34 (dt, J = 14.3, 7.2 Hz, 4H). 13C NMR (101 MHz, CDCl3) δ = 144.32, 141.38, 140.13, 139.42, 137.22, 130.15, 129.36, 128.73, 128.49, 128.20, 127.59, 127.25, 126.18, 125.94, 51.94, 36.40, 33.94, 32.24, 30.03. HRMS (APCI) [M+Na+] Calcd for C19H21N: 286.1566, Found: 286.1568

MEGX (3.13)



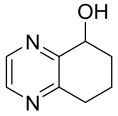
1H NMR (500 MHz, CDCl3) δ = 8.84 (br s, 1H), 7.16 – 7.05 (m, 3H), 3.47 (s, 2H), 2.80 (q, J=7.1, 2H), 2.24 (s, 2H), 1.61 (br s, 1H), 1.19 (t, J=7.1, 1H). 13C NMR (126 MHz, CDCl3) δ = 170.36 (C(O)NH), 135.25, 134.02, 128.39, 127.25 (Ph), 52.75 (CH2), 45.03 (CH2CH3), 18.73 (2CH3), 15.69 (CH3). HRMS (APCI) [M+Na+] Calcd for C12H18N2O: 229.1311, Found: 229.1314

(*Z*)-*N*-(3-ethyl-2-methyloxazolidin-5-ylidene)-2,6-dimethylaniline (3.14)



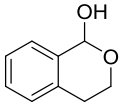
1H NMR (500 MHz, CDCl3) δ = 7.20 – 7.05 (m, 3H), 4.50 – 4.45 (m, 1H), 3.79 (dd, J=14.7, 1.2, 1H), 3.19 (dd, J=14.7, 1.7, 1H), 3.00 – 2.89 (m, 1H), 2.54 – 2.43 (m, 1H), 2.28 (s, 3H), 2.22 (s, 3H), 1.19 (t, J=7.2, 3H), 1.15 (d, J=5.7, 3H). 13C NMR (126 MHz, CDCl3) δ = 170.37, 138.16, 135.94, 133.23, 128.99, 128.83, 128.58, 77.36, 55.37, 47.70, 18.96, 18.84, 18.47, 13.56. HRMS (APCI) [M+Na+] Calcd for C12H18N2O: 255.1468, Found: 255.1477

5,6,7,8-tetrahydroquinoxalin-5-ol (4.02)



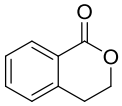
1H NMR (200 MHz, CDCl3) δ = 8.38 (m, 2H), 4.80 (br t, 1H, CH), 2.99 (m, 2H), 1.64 - 2.55 (m, 4H) ms: (m/e) 150 (80), 131 (7) 122(55), 107 (6), 95 (12), 94 (100), 93 (23), 68 (11), 67 (9), 66 (7), 65 (7), 52 (16)

Isochroman-1-ol (4.15)



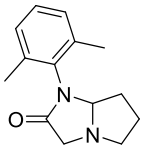
1H NMR (CDCl3, 600 MHz) δ =7.37−7.31 (m,1H), 7.30−7.23 (m, 2H), 7.15 (d, J= 7.2 Hz, 1H), 5.99 (d, J= 5.4 Hz,1H), 4.23 (td, J= 3.6, 12.6 Hz, 1H), 3.97 (ddd, J= 2.4, 5.4, 11.4 Hz,1H), 3.02 (d, J= 5.4 Hz, H), 2.70 (d, J= 16.2 Hz, 1H);13C NMR(CDCl3, 125 MHz) δ =134.9, 134.1, 128.5, 128.3, 127.3, 126.5, 91.5,58.3, 28.0; HRMS (APCI) [M + H]+ calcd for C9H10O2: 151.0754, found: 151.0748.

Isochroman-1-one (**4.16**)



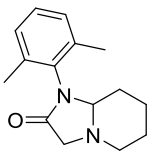
1H NMR (400 MHz, CDCl3): δ = 8.10 (d, J =7.6 Hz, 1 H), 7.55 (t, J = 7.4 Hz, 1 H), 7.40 (t, J = 7.6 Hz, 1 H), 7.27 (d, J = 7.6 Hz, 1 H), 4.55 (t, J = 6.0 Hz, 2 H), 3.07 (t, J = 6.0 Hz, 2 H).13C NMR (100 MHz, CDCl3): δ = 164.78, 139.36, 133.38, 129.90, 127.30, 127.03, 124.95, 67.05, 27.46; LRMS(APCI) [M+H]+ calcd for C9H8O2:149.1, found 149.1

1-(2,6-Dimethylphenyl)tetrahydro-1H-pyrrolo[1,2-a]imidazol-2(3H)-one (5.01)



^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.19–7.06 (m, 3H), 5.34–5.14 (m, 1H), 3.83 (d, *J* = 16.5 Hz, 1H), 3.45 (d, *J* = 16.5 Hz, 1H), 3.32–3.24 (m, 1H), 2.89–2.79 (m, 1H), 2.23 (br s, 6H), 2.04–1.81 (m, 3H), 1.77–1.64 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ = 172.1, 137.6, 135.7, 133.4, 129.1, 129.0, 128.5, 82.1, 57.0, 56.1, 28.9, 24.1, 18.7, 18.7; HRMS (APCI) calc. for $C_{14}H_{18}N_2NaO$ [MNa^+] 253.1312 found 253.1311.

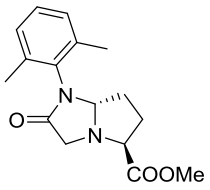
1-(2,6-Dimethylphenyl)hexahydroimidazo[1,2-a]pyridin-2(3H)-one (5.02)



^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.17–7.04 (m, 3H), 4.11–4.06 (m, 1H), 3.63 (d, *J* = 13.5 Hz, 1H), 3.20 (dd, *J* = 13.5, 2.0 Hz, 1H), 3.07 (dt, *J* = 11.0, 3.0 Hz, 1H), 2.52–2.41 (m, 1H), 2.29 (s, 3H), 2.23 (s, 3H), 1.84–1.76 (m, 1H), 1.74–1.62 (m, 2H), 1.52–1.41 (m, 2H), 1.41–1.29 (m, 1H); ^{13}C NMR (126

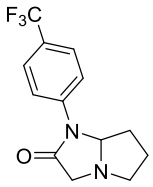
MHz, $CDCl_3$) δ = 170.9, 137.9, 135.9, 132.6, 128.7, 128.6, 128.4, 79.4, 55.9, 49.8, 27.6, 24.5, 22.2, 18.9, 18.7; HRMS (APCI) calc. for $C_{15}H_{21}N_2O$ [MH^+] 245.1648, found 245.1642.

Methyl (5S,7aR)-1-(2,6-dimethylphenyl)-2-oxohexahydro-1H-pyrrolo[1,2-a]imidazole-5-carboxylate (5.03)



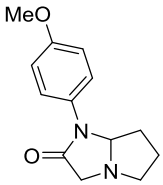
^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.21–7.08 (m, 3H), 5.39 (dd, *J* = 6.0, 2.5 Hz, 1H), 3.97 (d, *J* = 17.0 Hz, 1H), 3.80 (s, 3H), 3.74 (t, *J* = 8.0 Hz, 1H), 3.57 (d, *J* = 17.0 Hz, 1H), 2.40–2.30 (m, 1H), 2.25 (s, 3H), 2.23 (s, 3H), 2.18–1.99 (m, 2H), 1.85–1.74 (m, 1H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 173.2, 171.7, 137.5, 135.8, 133.1, 129.4, 129.1, 128.8, 82.3, 68.8, 56.8, 52.7, 28.5, 28.2, 18.9, 18.8; HRMS (ES^+) calc. for $C_{16}H_{20}N_2NaO_3$ [MNa^+] 311.1366, found 311.1358.

1-[4-(Trifluoromethyl)phenyl]tetrahydro-1H-pyrrolo[1,2-a]imidazol-2(3H)-one (5.04)



^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.72–7.50 (m, 4H), 5.45 (dd, *J* = 6.0, 3.5 Hz, 1H), 3.78 (d, *J* = 17.0 Hz, 1H), 3.51 (d, *J* = 17.0 Hz, 1H), 3.32–3.19 (m, 1H), 2.72 (dd, *J* = 17.0, 8.0 Hz, 1H), 2.42–2.25 (m, 1H), 1.98–1.72 (m, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ = 172.3, 140.2, 126.8 (q, *J* = 33.0 Hz), 126.3 (q, *J* = 4.0 Hz), 124.1 (q, *J* = 272 Hz), 120.8, 81.2, 57.4, 55.8, 31.2, 24.5; HRMS (ES^+) calc. for $C_{13}H_{13}N_2NaOF_3$ [MNa^+] 293.0872, found 293.0868.

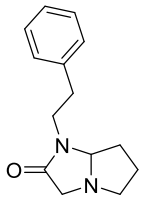
1-(4-Methoxyphenyl)tetrahydro-1*H*-pyrrolo[1,2-a]imidazol-2(3*H*)-one (5.05)



^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.32–7.28 (m, 2H), 6.96–6.91 (m, 2H), 5.35 (dd, *J* = 6.0, 3.0 Hz, 1H), 3.81 (s, 3H), 3.78 (d, *J* = 17.0 Hz, 1H), 3.46 (d, *J* = 17.0 Hz, 1H), 3.30–3.23 (m, 1H), 2.76–2.69 (m, 1H), 2.21–2.11 (m, 1H), 1.94–1.76 (m, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 171.9, 157.6, 129.8,

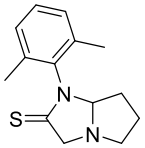
124.0, 114.6, 82.0, 57.7, 56.2, 55.7, 30.8, 24.3; HRMS (ES^+) calc. for $C_{13}H_{16}N_2NaO$ [MNa^+] 255.1104, found 255.1010.

1-Phenethyltetrahydro-1*H*-pyrrolo[1,2-a]imidazol-2(3*H*)-one (5.06)



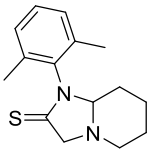
^1H NMR (500 MHz, $CDCl_3$) δ = 7.34–7.28 (m, 2H), 7.26–7.20 (m, 3H), 4.67–4.63 (m, 1H), 3.75 (ddd, *J* = 14.0, 8.5, 7.0 Hz, 1H), 3.58 (d, *J* = 16.5, 1H), 3.25 (d, *J* = 16.5 Hz, 1H), 3.25–3.20 (m, 1H), 3.19–3.13 (m, 1H), 2.97–2.82 (m, 2H), 2.55 (td, *J* = 9.0, 6.5 Hz, 1H), 2.03–1.92 (m, 1H), 1.86–1.70 (m, 2H), 1.69–1.62 (m, 1H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 172.7, 138.8, 128.9, 128.8, 126.8, 80.7, 57.3, 56.3, 42.4, 33.9, 30.3, 24.2; HRMS (ES^+) calc. for $C_{14}H_{18}N_2NaO$ [MNa^+] 253.1311, found 253.1307.

1-(2,6-Dimethylphenyl)tetrahydro-1*H*-pyrrolo[1,2-a]imidazole-2(3*H*)-thione (5.07)



^1H NMR (500 MHz, $CDCl_3$) δ = 7.26–7.12 (m, 3H), 5.45 (d, *J* = 5.0 Hz, 1H), 4.28 (d, *J* = 17.5 Hz, 1H), 4.08 (d, *J* = 17.5 Hz, 1H), 3.34 (s, 1H), 2.91–2.79 (m, 1H), 2.22 (d, *J* = 9.5 Hz, 3H), 2.20 (s, 3H), 2.06–1.93 (m, 3H), 1.92–1.80 (m, 1H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 198.3, 136.9, 134.8, 129.5, 129.2, 89.7, 68.9, 55.5, 28.2, 24.3, 18.6, 18.5; HRMS (ES^+) calc. for $C_{14}H_{19}N_2S$ [MH^+] 247.1264, found 247.1258.

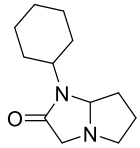
1-(2,6-Dimethylphenyl)hexahydroimidazo[1,2-a]pyridine-2(3*H*)-thione (5.08)



^1H NMR (500 MHz, $CDCl_3$) δ = 7.23–7.08 (m, 3H), 4.34–4.26 (m, 1H), 4.13 (dd, *J* = 15.0, 0.5 Hz, 1H), 3.61 (dd, *J* = 15.0, 3.0 Hz, 1H), 3.10–3.01 (m, 1H), 2.55–2.44 (m, 1H), 2.29 (s, 3H), 2.20 (s, 3H), 1.83–1.72 (m, 1H), 1.72–1.64 (m, 2H), 1.58–1.48 (m, 2H), 1.40–1.28 (m, 1H); ^{13}C NMR (126 MHz,

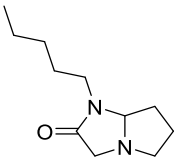
$CDCl_3$) δ = 198.5, 137.1, 135.1, 134.8, 129.0, 129.0, 128.9, 84.3, 68.0, 49.6, 27.5, 24.5, 22.2, 18.8, 18.7; HRMS (ES^+) calc. for $C_{15}H_{20}N_2NaS$ [MNa^+] 282.1239, found 282.1233.

1-Cyclohexyltetrahydro-1*H*-pyrrolo[1,2-a]imidazol-2(3*H*)-one (5.09)



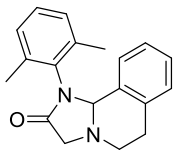
^1H NMR (500 MHz, $CDCl_3$) δ = 4.83 (t, *J* = 5.5 Hz, 1H), 3.79–3.69 (m, 1H), 3.61 (d, *J* = 16.0 Hz, 1H), 3.26 (dd, *J* = 16.0, 1.5 Hz, 1H), 3.23–3.17 (m, 1H), 2.64–2.53 (m, 1H), 2.21–2.11 (m, 1H), 1.98–1.89 (m, 1H), 1.86–1.71 (m, 4H), 1.69–1.65 (m, 2H), 1.64–1.55 (m, 1H), 1.54–1.42 (m, 1H), 1.41–1.28 (m, 2H), 1.21–1.06 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 172.1, 79.6, 57.9, 56.2, 52.4, 33.5, 32.0, 30.3, 26.05, 26.00, 25.6, 24.9; HRMS (ES^+) calc. for $C_{12}H_{21}N_2O$ [MH^+] 209.1648, found 209.1649.

1-Pentyltetrahydro-1*H*-pyrrolo[1,2-a]imidazol-2(3*H*)-one (5.10)



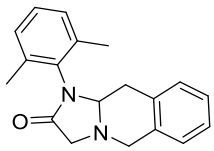
^1H NMR (500 MHz, $CDCl_3$) δ = 4.82 (t, *J* = 4.5 Hz, 1H), 3.64 (d, *J* = 16.5 Hz, 1H), 3.52 (ddd, *J* = 13.5, 9.0, 7.0 Hz, 1H), 3.27 (d, *J* = 16.5 Hz, 1H), 3.24–3.18 (m, 1H), 3.02–2.94 (m, 1H), 2.64–2.54 (m, 1H), 2.13–2.04 (m, 1H), 1.95–1.74 (m, 3H), 1.56–1.46 (m, 2H), 1.40–1.23 (m, 4H), 0.95–0.87 (m, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 172.4, 80.4, 57.5, 56.4, 40.8, 30.5, 29.3, 27.3, 24.3, 22.5, 14.1; HRMS (ES^+) calc. for $C_{11}H_{21}N_2O$ [MH^+] 197.1648, found 197.1649.

1-(2,6-Dimethylphenyl)-1,5,6,10b-tetrahydroimidazo[2,1-a]isoquinolin-2(3*H*)-one (5.11)



^1H NMR (500 MHz, $CDCl_3$) δ = 7.23–7.12 (m, 4H), 6.88 (d, *J* = 7.0 Hz, 1H), 6.86–6.81 (m, 1H), 6.19 (d, *J* = 8.0 Hz, 1H), 5.87 (s, 1H), 3.98 (d, *J* = 16.0 Hz, 1H), 3.55 (d, *J* = 16.0 Hz, 1H), 3.37–3.26 (m, 1H), 3.13–3.02 (m, 2H), 2.99–2.89 (m, 1H), 2.38 (s, 3H), 1.51 (s, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 172.5, 139.2, 136.0, 135.3, 133.3, 129.5, 128.9, 128.8, 128.7, 128.6, 128.6, 128.6, 125.7, 76.2, 58.3, 47.7, 28.2, 18.6, 18.2; HRMS (ES^+) calc. for $C_{19}H_{20}N_2NaO$ [MNa^+] 315.1468, found 315.1461.

1-(2,6-Dimethylphenyl)-1,5,10,10a-tetrahydroimidazo[1,2-b]isoquinolin-2(3*H*)-one (5.12)



^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.25–7.09 (m, 6H), 7.04 (d, *J* = 7.5 Hz, 1H), 4.67–4.61 (m, 1H), 4.18 (d, *J* = 14.5 Hz, 1H), 3.90–3.81 (m, 2H), 3.44 (dd, *J* = 14.0, 2.0 Hz, 1H), 2.91–2.83 (m, 1H), 2.64 (dd, *J* = 15.0, 3.5 Hz, 1H), 2.33 (s, 3H), 2.27 (s, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 170.6, 138.0, 136.2, 134.2, 132.9, 131.9, 129.7, 129.0, 129.0, 128.8, 127.1, 126.9, 126.8, 76.6, 56.0, 53.4, 33.2, 19.0, 18.9. HRMS (ES^+) calc. for $C_{19}H_{20}N_2NaO$ [MNa^+] 315.1468, found 315.1461.

Appendix 7: 1H NMR figure of drug fragments in Chapter 4

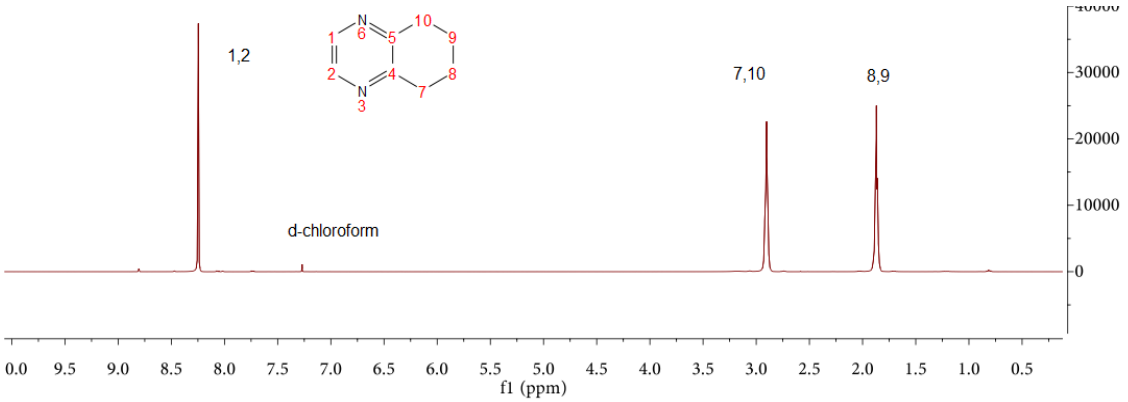


Figure S7.1 ^{1}H NMR of 5,6,7,8-tertrahydroquinoxaline **(4.01)**

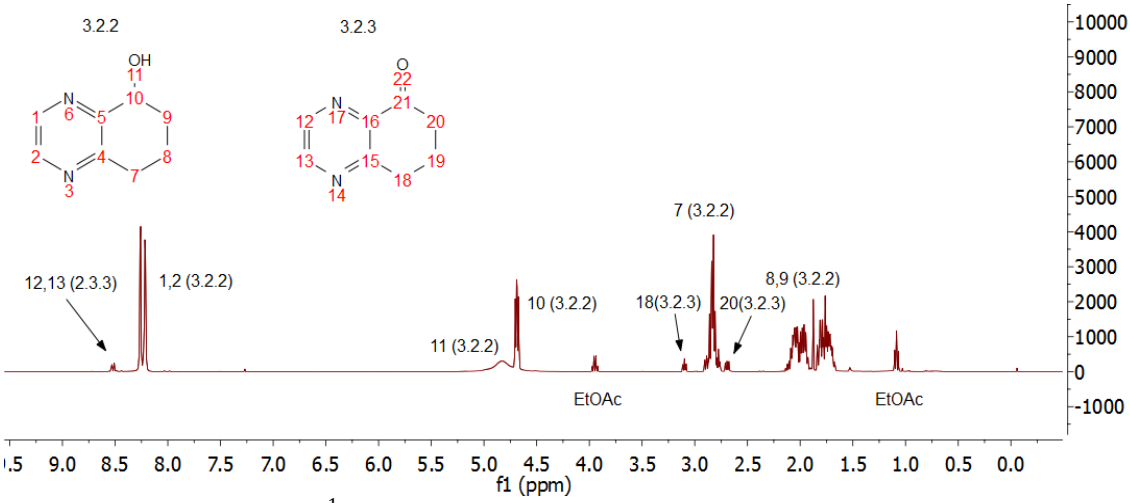


Figure S7.2 ^{1}H NMR of 5-hydroxy-tetrahydroquinoxaline **(4.02)** and tetrahydroquinoxaline-5-one **(4.03)**

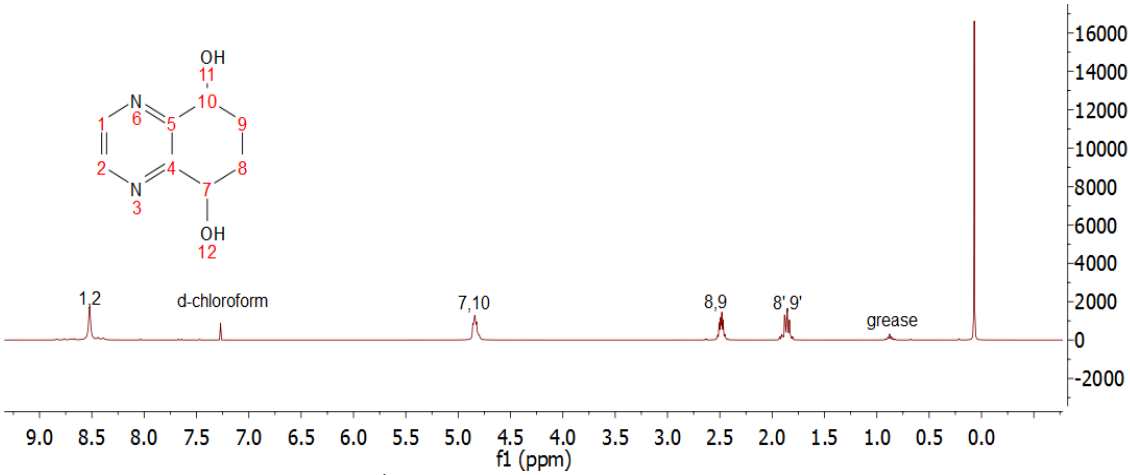


Figure S7.3 ^{1}H NMR of tetrahydroquinoaline-5,8-diol **(4.04)**

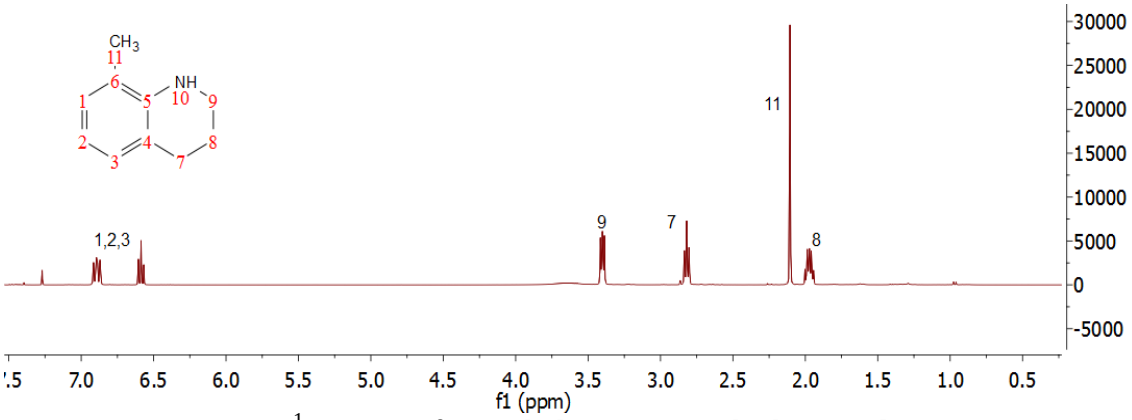


Figure S7.4 ^{1}H NMR of 8-methyl-1,2,3,4-tetrahydroquinoline **(4.05)**

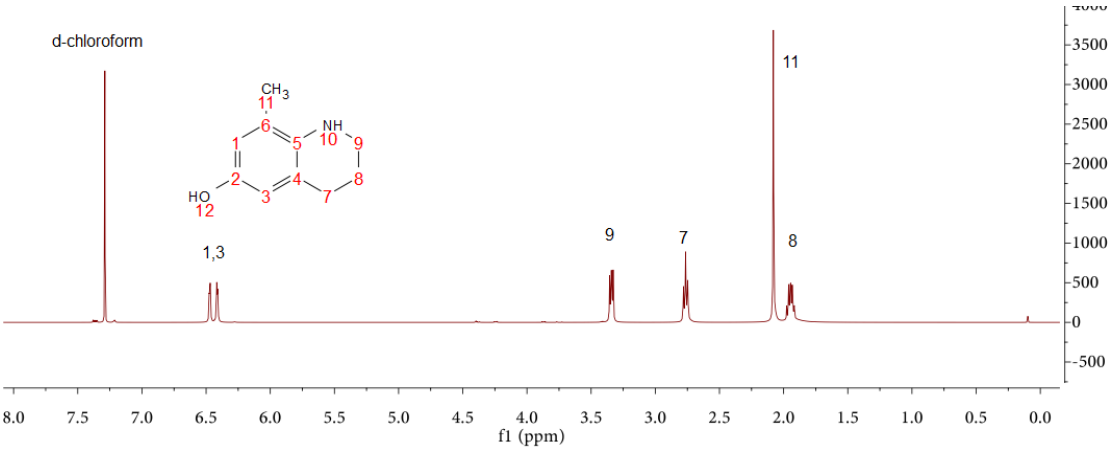


Figure S7.5 ^{1}H NMR of 8-methyl-1,2,3,4-tetrahydroquinolin-6-ol **(4.06)**

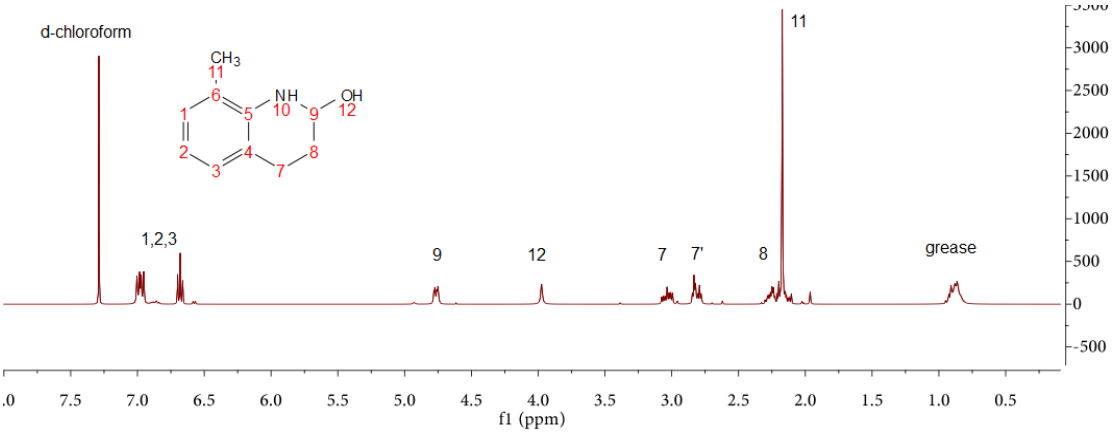


Figure S7.6 ^{1}H NMR of 8-methyl-1,2,3,4-tetrahydroquinolin-2-ol **(4.07)**

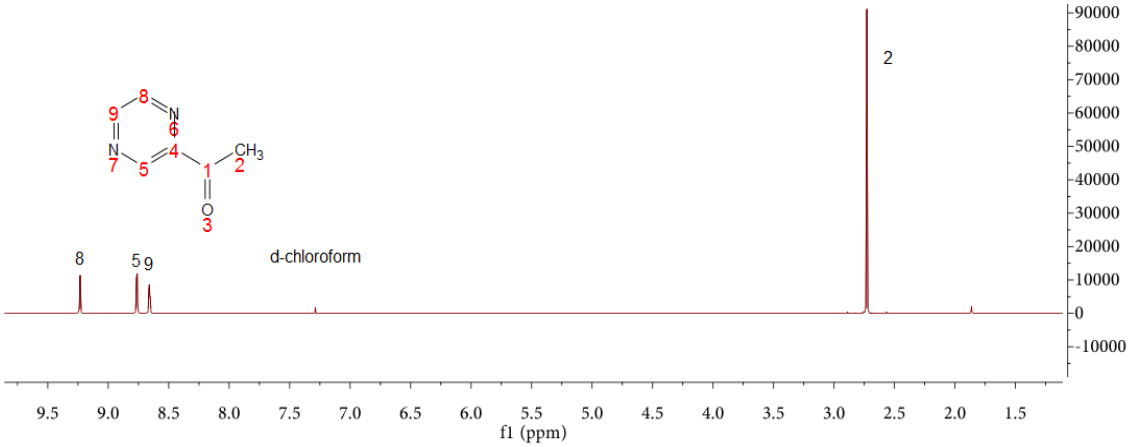


Figure S7.7 ^{1}H NMR of acetylpyrazine **(4.08)**

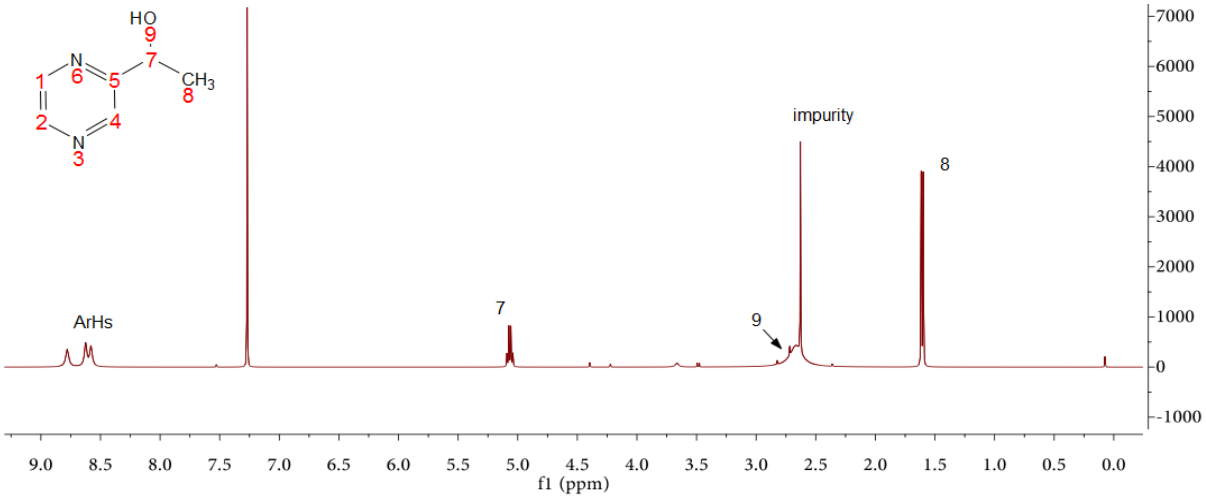


Figure S7.8 ^{1}H NMR of 1-(pyrazine-2-yl)ethan-1-ol **(4.09)**

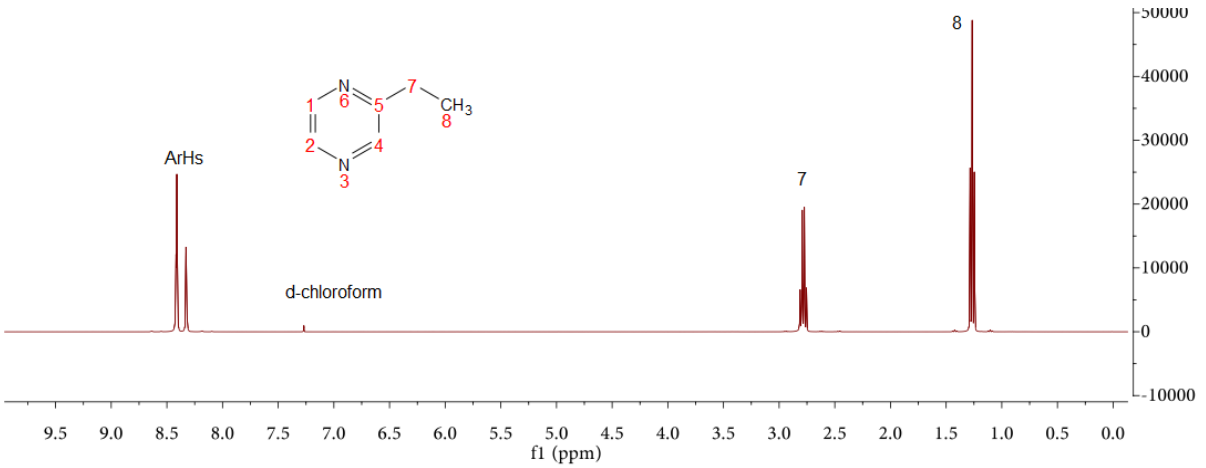


Figure S7.9 ^{1}H NMR of ethylpyrazine **(4.10)**

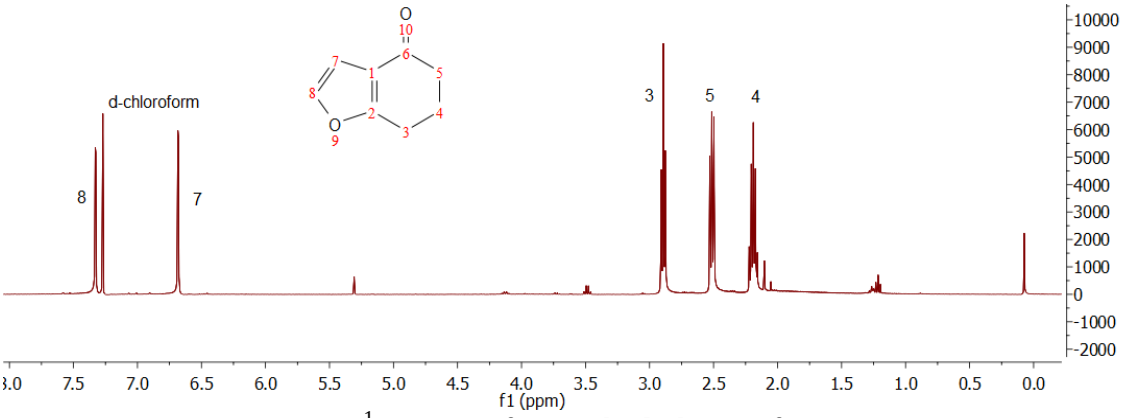


Figure S7.10 ^{1}H NMR of 6,7-Dihydrobenzonfu-4(5*H*)-one **(4.11)**

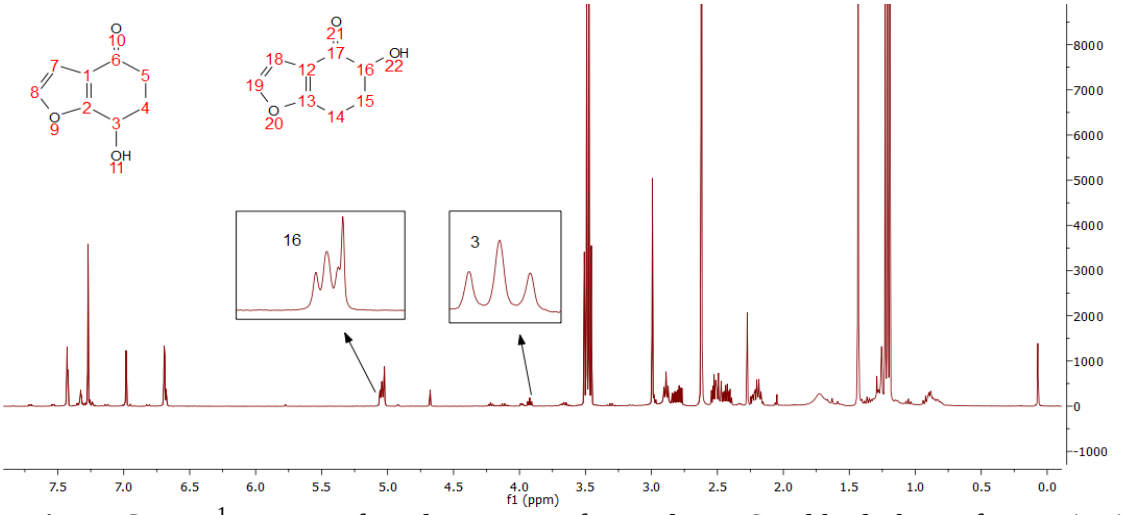


Figure S7.11: ^{1}H NMR of crude mixture of 7-Hydroxy-6,7-dihydrobenzofuran-4(5*H*)-one (**4.12**) and (5-hydroxy-6,7-dihydrobenzofuran-4(5*H*)-one **(4.13)**

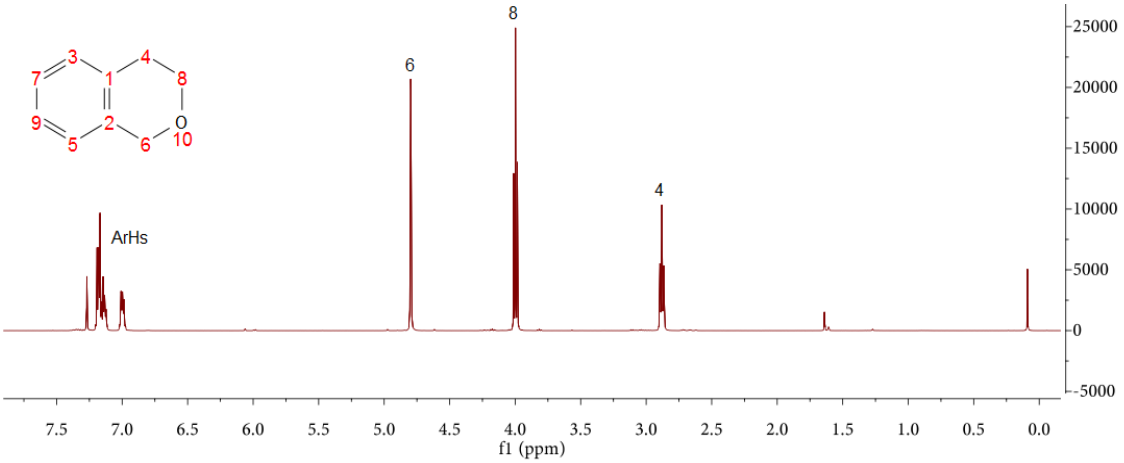


Figure S7.12 ^{1}H NMR of isochroman **(4.14)**

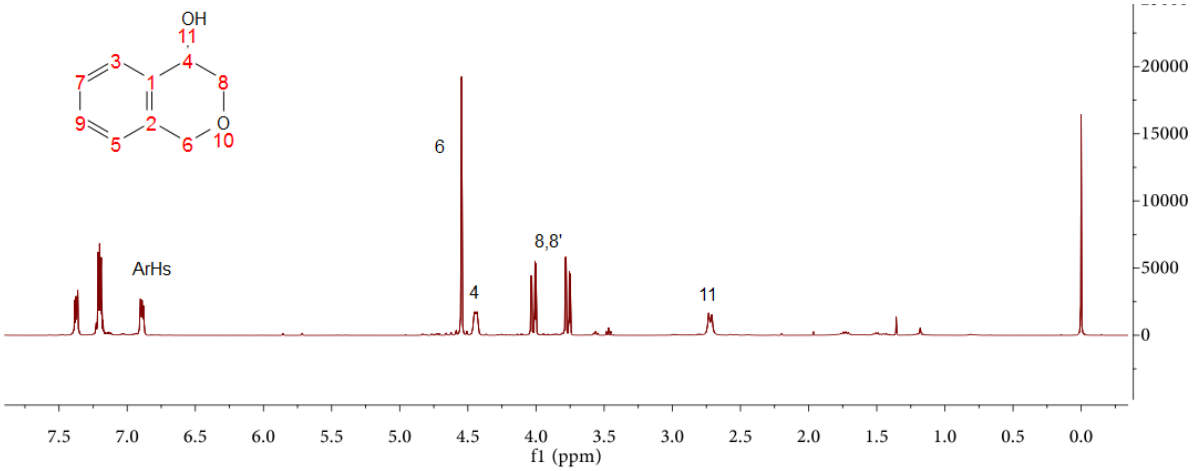


Figure S7.13 ^{1}H NMR of isochroman-4-ol **(4.17)**

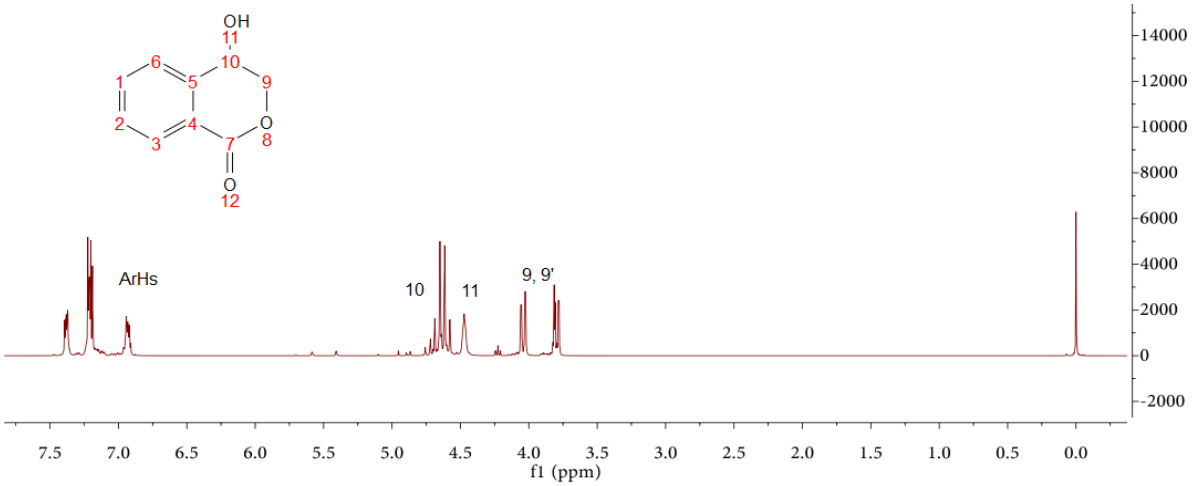


Figure S7.14 ^{1}H NMR of 4-hydroxyisochroan-1-one **(4.18)**

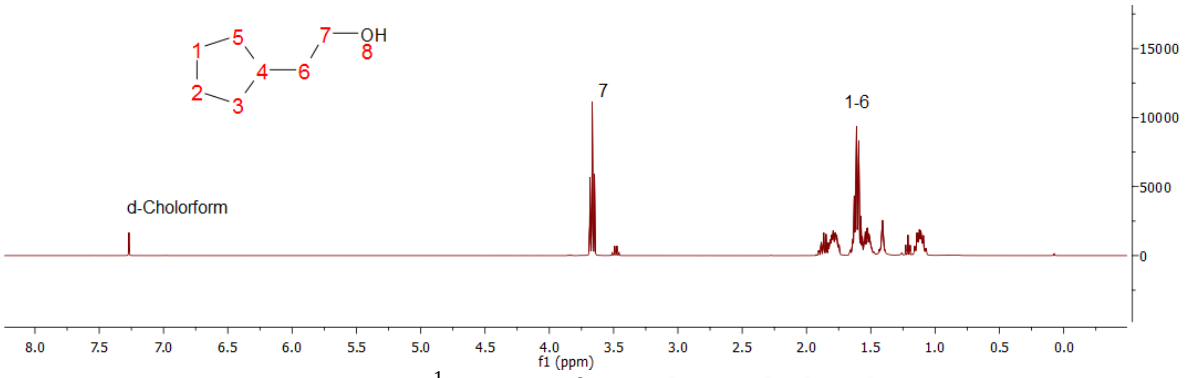


Figure S7.15 ^{1}H NMR of 2-Cyclopentyl ethanol **(4.30)**

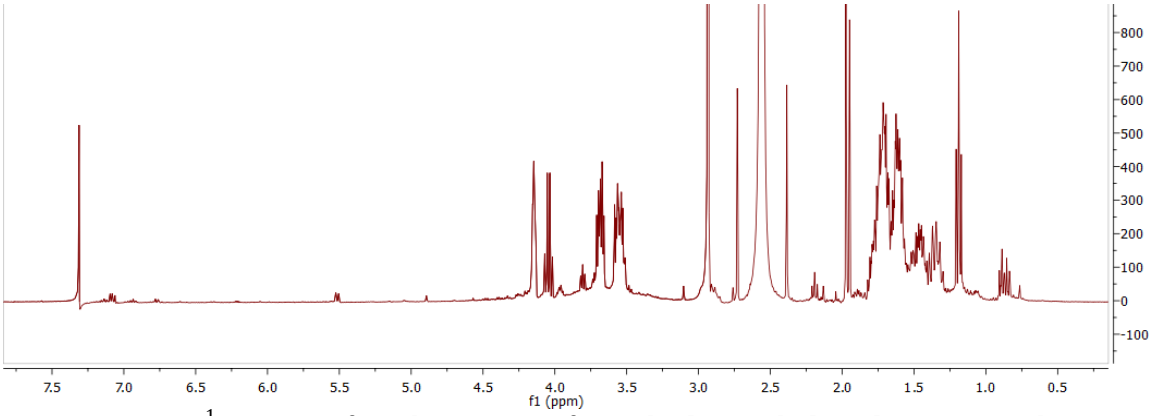


Figure S7.16 ^{1}H NMR of crude mixture of 2-(2-hydroxyethyl)cyclopentan-1-ol (**4.32**) and 1-(2-hydroxyehtyl)cyclopentan-1-ol (**4.33**)

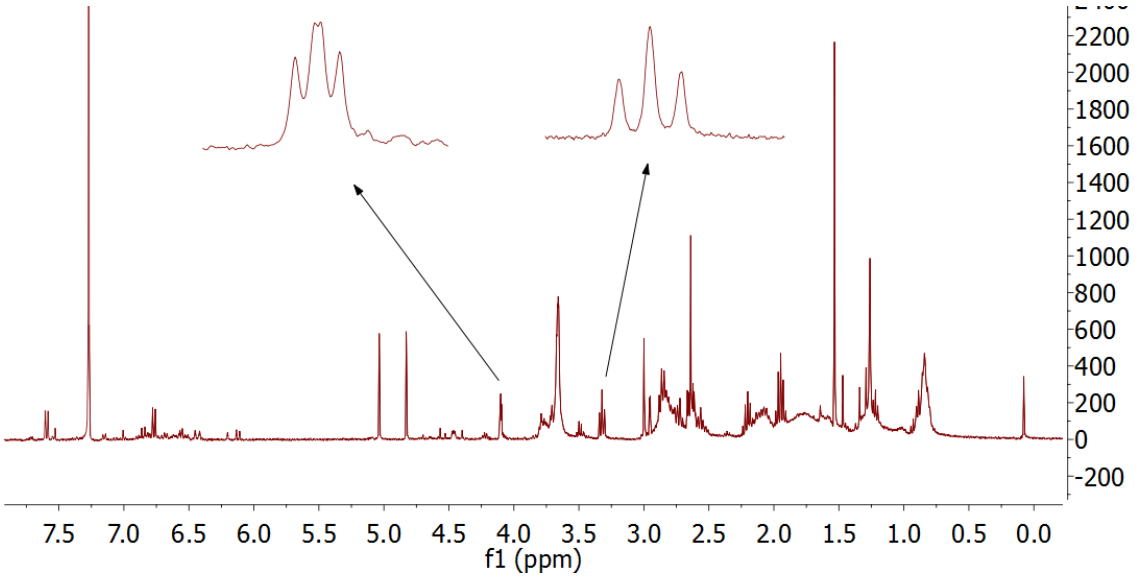


Figure S7.17: ^{1}H NMR of crude mixture of 7-methyl-2,3-dihydro-1*H*-indene-1,4-diol, (**4.35**) and 4-methyl-2,3-dihydro-1H-indene-1,7-diol **(4.36)**

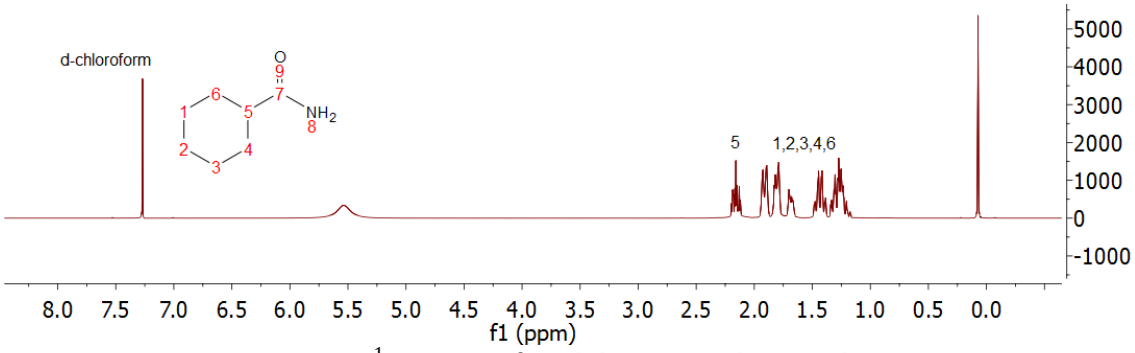


Figure S7.18 ^{1}H NMR of cyclohexane carboxamide **(4.37)**

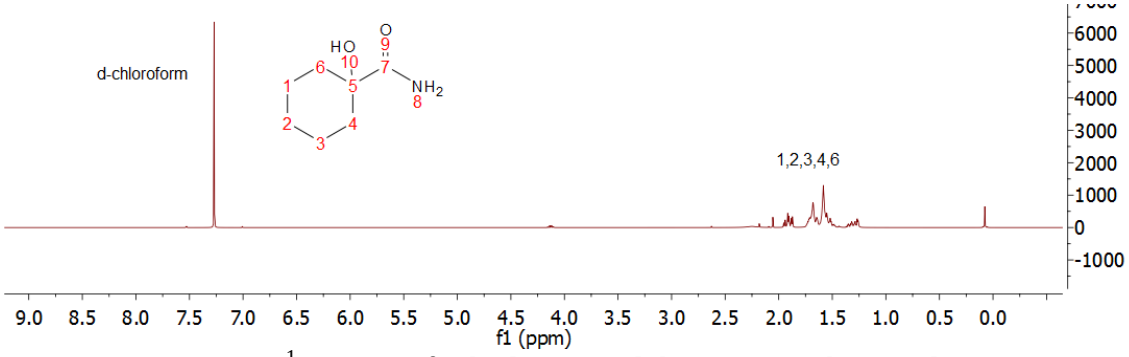


Figure S7.19 ^{1}H NMR of 1-hydroxycecylohexane-1-carboxamide **(4.38)**

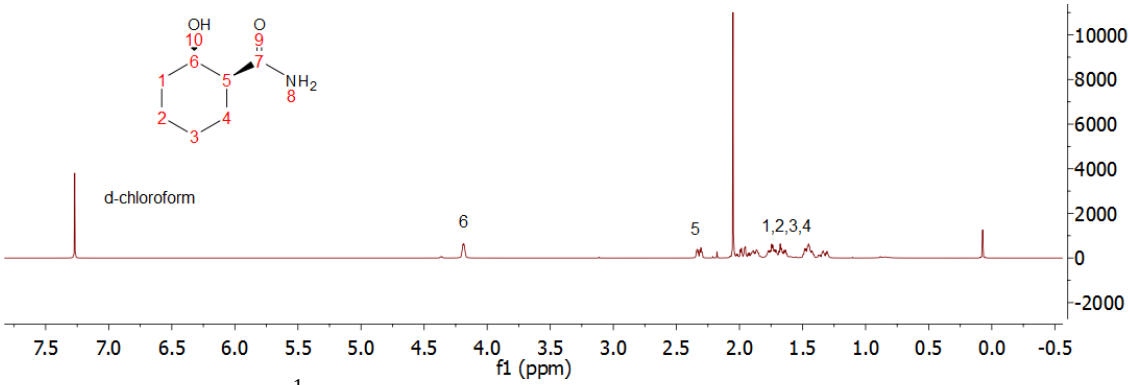


Figure S7.20 ^{1}H NMR of 2-hydroxycyclohexane-1-carboxamide **(4.39)**

References

References

1. T. Newhouse and P. S. Baran, *Angew. Chem., Int. Ed.*, 2011, **50**, 3362-3374.
2. M. C. White, *Science*, 2012, **335**, 807-809.
3. S. R. Neufeldt and M. S. Sanford, *Acc. Chem. Res.*, 2012, **45**, 936-946.
4. M. Sun, J. Zhang, P. Putaj, V. Caps, F. Lefebvre, J. Pelletier and J.-M. Basset, *Chem. Rev.* , 2014, **114**, 981-1019.
5. E. Roduner, W. Kaim, B. Sarkar, V. B. Urlacher, J. Pleiss, R. Glaeser, W.-D. Einicke, G. A. Sprenger, U. Beifuss, E. Klemm, C. Liebner, H. Hieronymus, S.-F. Hsu, B. Plietker and S. Laschat, *ChemCatChem*, 2013, **5**, 82-112.
6. M. Canta, D. Font, L. Gomez, X. Ribas and M. Costas, *Adv. Synth. Catal.*, 2014, **356**, 818-830.
7. D. R. Nelson, *J. Am. Chem. Soc.*, 2005, **127**, 12147-12148.
8. A. W. Munro, H. M. Girvan and K. J. McLean, *Nat. Prod. Rep.*, 2007, **24**, 585-609.
9. P. R. Ortiz de Montellano, *Chem. Rev.*, 2010, **110**, 932-948.
10. S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar and W. Thiel, *Chem. Rev.* , 2010, **110**, 949-1017.
11. C. J. C. Whitehouse, S. G. Bell and L.-L. Wong, *Chem. Soc. Rev.*, 2012, **41**, 1218-1260.
12. J. C. Lewis, P. S. Coelho and F. H. Arnold, *Chem. Soc. Rev.*, 2011, **40**, 2003-2021.
13. K. Drauz, H. Groeger, O. May and Editors, *Enzyme Catalysis in Organic Synthesis, Volume 1: Third Edition*, Wiley-VCH Verlag GmbH & Co. KGaA, 2011.
14. E. O'Reilly, V. Kohler, S. L. Flitsch and N. J. Turner, *Chem. Commun.*, 2011, **47**, 2490-2501.
15. R. Fasan, *ACS Catal.*, 2012, **2**, 647-666.
16. Y. Khatri, F. Hannemann, M. Girhard, R. Kappl, A. Meme, M. Ringle, S. Janocha, E. Leize-Wagner, V. B. Urlacher and R. Bernhardt, *Biotechnol. Appl. Biochem.*, 2013, **60**, 18-29.
17. D. Holtmann, M. W. Fraaije, I. W. C. E. Arends, D. J. Opperman and F. Hollmann, *Chem. Commun.*, 2014, **50**, 13180-13200.
18. G.-D. Roiban and M. T. Reetz, *Chem. Commun.* , 2015, **51**, 2208-2224.
19. M. T. Reetz, *J. Am. Chem. Soc.*, 2013, **135**, 12480-12496.
20. L. Ji, J. Zhang, W. Liu, A. S. Faponle, M. G. Quesne, M. A. Sainna, V. S. P. de, A. Franke, E. R. van, D. Kumar and E. R. van, *Chemistry*, 2015, **21**, 9083-9092.
21. C. J. C. Whitehouse, S. G. Bell, H. G. Tufton, R. J. P. Kenny, L. C. I. Ogilvie and L.-L. Wong, *Chem. Commun.* , 2008, 966-968.
22. A. E. Rettie, P. R. Sheffels, K. R. Korzekwa, F. J. Gonzalez, R. M. Philpot and T. A. Baillie, *Biochemistry*, 1995, **34**, 7889-7895.
23. X. Guan, M. B. Fisher, D. H. Lang, Y. M. Zheng, D. R. Koop and A. E. Rettie, *Chemico-Biological Interactions*, 1998, **110**, 103-121.
24. T. S. Dowers and J. P. Jones, *Drug Metab. Dispos.*, 2006, **34**, 1288-1290.
25. M. N. Bhakta and K. Wimalasena, *Eur. J. Org. Chem.*, 2005, 4801-4805.
26. M. N. Bhakta, P. F. Hollenberg and K. Wimalasena, *J. Am. Chem. Soc.*, 2005, **127**, 1376-1377.
27. H. Chen, M. J. de Groot, N. P. E. Vermeulen and R. P. Hanzlik, *J. Org. Chem.*, 1997, **62**, 8227-8230.
28. K. E. Anderson, G. J. Hammons, F. F. Kadlubar, J. D. Potter, K. R. Kaderlik, K. F. Ilett, R. F. Minchin, C. H. Teitel, H. C. Chou, M. V. Martin, F. P. Guengerich, G. W. Barone, N. P. Lang and L. A. Peterson, *Carcinogenesis*, 1997, **18**, 1085-1092.
29. T. Kubo, M. W. Peters, P. Meinhold and F. H. Arnold, *Chem. - Eur. J.*, 2006, **12**, 1216-1220.

30. M. Alcalde, E. T. Farinas and F. H. Arnold, *J. Biomol. Screening*, 2004, **9**, 141-146.
31. E. T. Farinas, M. Alcalde and F. Arnold, *Tetrahedron*, 2004, **60**, 525-528.
32. S. Jin, T. M. Makris, T. A. Bryson, S. G. Sligar and J. H. Dawson, *J. Am. Chem. Soc.*, 2003, **125**, 3406-3407.
33. A. D. N. Vaz, D. F. McGinnity and M. J. Coon, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 3555-3560.
34. J. A. S. J. Razenberg, R. J. M. Nolte and W. Drenth, *J. Chem. Soc., Chem. Commun.*, 1986, 277-279.
35. J. P. Collman, J. I. Brauman, B. Meunier, T. Hayashi, T. Kodadek and S. A. Raybuck, *J. Am. Chem. Soc.*, 1985, **107**, 2000-2005.
36. P. S. Coelho, E. M. Brustad, A. Kannan and F. H. Arnold, *Science*, 2013, **339**, 307-310.
37. T. Heel, J. A. McIntosh, S. C. Dodani, J. T. Meyerowitz and F. H. Arnold, *ChemBioChem*, 2014, **15**, 2556-2562.
38. H. Renata, Z. J. Wang, R. Z. Kitto and F. H. Arnold, *Catal. Sci. Technol.*, 2014, **4**, 3640-3643.
39. Z. J. Wang, H. Renata, N. E. Peck, C. C. Farwell, P. S. Coelho and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2014, **53**, 6810-6813.
40. C. J. C. Whitehouse, N. H. Rees, S. G. Bell and L.-L. Wong, *Chem. - Eur. J.*, 2011, **17**, 6862-6868.
41. C. J. C. Whitehouse, S. G. Bell and L.-L. Wong, *Chemistry*, 2008, **14**, 10905-10908.
42. A. Jousserandot, J.-L. Boucher, Y. Henry, B. Niklaus, B. Clement and D. Mansuy, *Biochemistry*, 1998, **37**, 17179-17191.
43. M. Frey, P. Chomet, E. Glawischnig, C. Stettner, S. Grun, A. Winklmair, W. Eisenreich, A. Bacher, R. B. Meeley, S. P. Briggs, K. Simcox and A. Gierl, *Science*, 1997, **277**, 696-699.
44. L. M. Halo, M. N. Heneghan, A. A. Yakasai, Z. Song, K. Williams, A. M. Bailey, R. J. Cox, C. M. Lazarus and T. J. Simpson, *J. Am. Chem. Soc.*, 2008, **130**, 17988-17996.
45. Y. Matsuda, T. Iwabuchi, T. Wakimoto, T. Awakawa and I. Abe, *J. Am. Chem. Soc.*, 2015, **137**, 3393-3401.
46. d. M. P. R. Ortiz and S. D. Nelson, *Arch Biochem Biophys*, 2011, **507**, 95-110.
47. M. Makino, H. Sugimoto, Y. Shiro, S. Asamizu, H. Onaka and S. Nagano, *Proc Natl Acad Sci U S A*, 2007, **104**, 11591-11596.
48. J. L. Grant, C. H. Hsieh and T. M. Makris, *J. Am. Chem. Soc.*, 2015, **137**, 4940-4943.
49. A. Dennig, M. Kuhn, S. Tassoti, A. Thiessenhusen, S. Gilch, T. Buelter, T. Haas, M. Hall and K. Faber, *Angew. Chem., Int. Ed.*, 2015, **54**, 8819-8822.
50. D. S. Riddick, X. Ding, C. R. Wolf, T. D. Porter, A. V. Pandey, Q.-Y. Zhang, J. Gu, R. D. Finn, S. Ronseaux, L. A. McLaughlin, C. J. Henderson, L. Zou and C. E. Fluck, *Drug Metab. Dispos.*, 2013, **41**, 12-23.
51. O. O. Pentyuk, S. O. Kachula and O. K. Gerich, *Ukr. Biokhim. Zh.*, 2004, **76**, 16-28.
52. S. E. Bulun, S. Sebastian, K. Takayama, T. Suzuki, H. Sasano and M. Shozu, *J. Steroid Biochem. Mol. Biol.*, 2003, **86**, 219-224.
53. V. Rago, F. Romeo, S. Aquila, D. Montanaro, S. Ando and A. Carpino, *Reprod. Biol. Endocrinol.*, 2005, **3**, 72.
54. C. Brieke, M. Peschke, K. Haslinger and M. J. Cryle, *Angew. Chem., Int. Ed.*, 2015, **54**, 15715-15719.
55. C. Brieke, V. Kratzig, K. Haslinger, A. Winkler and M. J. Cryle, *Org. Biomol. Chem.*, 2015, **13**, 2012-2021.
56. M. C. Damsten, B. M. A. van Vugt-Lussenburg, T. Zeldenthuis, J. S. B. de Vlieger, J. N. M. Commandeur and N. P. E. Vermeulen, *Chem.-Biol. Interact.*, 2008, **171**, 96-107.

57. R. A. Kahn, T. Fahrendorf, B. A. Halkier and B. L. Moller, *Arch. Biochem. Biophys.*, 1999, **363**, 9-18.
58. E. Hayashi, K. Fuzimoto and H. Imaishi, *Plant Biotechnol.* , 2007, **24**, 393-396.
59. A. Gesell, M. Blaukopf, L. Madilao, M. M. S. Yuen, S. G. Withers, J. Mattsson, J. H. Russell and J. Bohlmann, *Plant Physiol.*, 2015, **168**, 94-106.
60. K. H. Teoh, D. R. Polichuk, D. W. Reed, G. Nowak and P. S. Covello, *FEBS Lett.*, 2006, **580**, 1411-1416.
61. R. Kaspera and R. Croteau, *Phytochem. Rev.*, 2006, **5**, 433-444.
62. S. Jennewein, C. D. Rithner, R. M. Williams and R. B. Croteau, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 13595-13600.
63. M. Chau, S. Jennewein, K. Walker and R. Croteau, *Chem. Biol.*, 2004, **11**, 663-672.
64. D. Leys, C. G. Mowat, K. J. McLean, A. Richmond, S. K. Chapman, M. D. Walkinshaw and A. W. Munro, *J. Biol. Chem.*, 2003, **278**, 5141-5147.
65. A. Souter, K. J. McLean, W. E. Smith and A. W. Munro, *J. Chem. Technol. Biotechnol.*, 2000, **75**, 933-941.
66. R. L. Wright, K. Harris, B. Solow, R. H. White and P. J. Kennelly, *FEBS Lett.*, 1996, **384**, 235-239.
67. I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko and S. G. Sligar, *Science* 2000, **287**, 1615-1622.
68. D. F. V. Lewis and Y. Ito, *Expert Opin. Drug Metab. Toxicol.*, 2008, **4**, 1181-1186.
69. J. S. Warrington, R. I. Shader, L. L. Von Moltke and D. J. Greenblatt, *Drug Metab. Dispos.*, 2000, **28**, 392-397.
70. H. Z. Bu, *Curr. Drug Metab.*, 2006, **7**, 231-249.
71. J. F. Andersen, K. Tatsuta, H. Gunji, T. Ishiyama and C. R. Hutchinson, *Biochemistry*, 1993, **32**, 1905-1913.
72. J. Rittle and M. T. Green, *Science* 2010, **330**, 933-937.
73. P. J. Loida and S. G. Sligar, *Biochemistry*, 1993, **32**, 11530-11538.
74. W. D. Woggon, H. A. Wagenknecht and C. Claude, *J. Inorg. Biochem.*, 2001, **83**, 289-300.
75. P. F. Hollenberg, *FASEB J.*, 1992, **6**, 686-694.
76. C. C. Winterbourn, M. C. Vissers and A. J. Kettle, *Curr Opin Hematol*, 2000, **7**, 53-58.
77. T. Omura, *J Biochem*, 2010, **147**, 297-306.
78. P. Sandhu, Z. Guo, T. Baba, M. V. Martin, R. H. Tukey and F. P. Guengerich, *Arch Biochem Biophys*, 1994, **309**, 168-177.
79. A. J. Fulco, *Annu. Rev. Pharmacol. Toxicol.*, 1991, **31**, 177-203.
80. C. F. Oliver, S. Modi, W. U. Primrose, L.-Y. Lian and G. C. K. Roberts, *Biochem. J.*, 1997, **327**, 537-544.
81. C.-K. J. Chen, T. K. Shokhireva, R. E. Berry, H. Zhang and F. A. Walker, *JBIC, J. Biol. Inorg. Chem.*, 2008, **13**, 813-824.
82. L. O. Narhi and A. J. Fulco, *J. Biol. Chem.*, 1986, **261**, 7160-7169.
83. L. O. Narhi, B. H. Kim, P. M. Stevenson and A. J. Fulco, *Biochem. Biophys. Res. Commun.*, 1983, **116**, 851-858.
84. L. O. Narhi and A. J. Fulco, *J. Biol. Chem.*, 1987, **262**, 6683-6690.
85. Y. Miura and A. J. Fulco, *J. Biol. Chem.*, 1974, **249**, 1880-1888.
86. R. S. Hare and A. J. Fulco, *Biochem. Biophys. Res. Commun.*, 1975, **65**, 665-672.
87. P. P. Ho and A. J. Fulco, *Biochim. Biophys. Acta, Lipids Lipid Metab.*, 1976, **431**, 249-256.
88. R. S. Matson, R. S. Hare and A. J. Fulco, *Biochim. Biophys. Acta, Lipids Lipid Metab.*, 1977, **487**, 487-494.

89. Y. Miura and A. J. Fulco, *Biochim. Biophys. Acta, Lipids Lipid Metab.*, 1975, **388**, 305-317.
90. N. Shirane, Z. Sui, J. A. Peterson and P. R. Ortiz de Montellano, *Biochemistry*, 1993, **32**, 13732-13741.
91. J. F. Buchanan and A. J. Fulco, *Biochem. Biophys. Res. Commun.*, 1978, **85**, 1254-1260.
92. R. T. Ruettinger and A. J. Fulco, *J. Biol. Chem.*, 1981, **256**, 5728-5734.
93. J. H. Capdevila, S. Wei, C. Helvig, J. R. Falck, Y. Belosludtsev, G. Truan, S. E. Graham-Lorence and J. A. Peterson, *J. Biol. Chem.*, 1996, **271**, 22663-22671.
94. A. Y. Lu and M. J. Coon, *J Biol Chem*, 1968, **243**, 1331-1332.
95. A. Y. H. Lu, K. W. Junk and M. J. Coon, *J. Biol. Chem.*, 1969, **244**, 3714-3721.
96. T. Omura, E. Sanders, R. W. Estabrook, D. Y. Cooper and O. Rosenthal, *Arch. Biochem. Biophys.*, 1966, **117**, 660-673.
97. I. F. Sevrioukova, C. Garcia, H. Li, B. Bhaskar and T. L. Poulos, *J Mol Biol*, 2003, **333**, 377-392.
98. J. T. Hazzard, S. Govindaraj, T. L. Poulos and G. Tollin, *J. Biol. Chem.*, 1997, **272**, 7922-7926.
99. M. B. Murataliev, M. Klein, A. Fulco and R. Feyereisen, *Biochemistry*, 1997, **36**, 8401-8412.
100. I. F. Sevrioukova, J. T. Hazzard, G. Tollin and T. L. Poulos, *J. Biol. Chem.*, 1999, **274**, 36097-36106.
101. S. S. Boddupalli, R. W. Estabrook and J. A. Peterson, *J. Biol. Chem.*, 1990, **265**, 4233-4239.
102. H. Li, K. Darwish and T. L. Poulos, *J. Biol. Chem.*, 1991, **266**, 11909-11914.
103. J. S. Miles, A. W. Munro, B. N. Rospendowski, W. E. Smith, J. McKnight and A. J. Thomson, *Biochem. J.*, 1992, **288**, 503-509.
104. S. Modi, W. U. Primrose, L.-Y. Lian and G. C. K. Roberts, *Biochem. J.*, 1995, **310**, 939-943.
105. A. W. Munro, J. G. Lindsay, J. R. Coggins, S. M. Kelly and N. C. Price, *FEBS Lett.*, 1994, **343**, 70-74.
106. S. Govindaraj and T. L. Poulos, *Biochemistry*, 1995, **34**, 11221-11226.
107. S. Govindaraj and T. L. Poulos, *Protein Sci.*, 1996, **5**, 1389-1393.
108. S. D. Black and S. T. Martin, *Biochemistry*, 1994, **33**, 12056-12062.
109. S. Govindaraj and T. L. Poulos, *J. Biol. Chem.*, 1997, **272**, 7915-7921.
110. R. Neeli, H. M. Girvan, A. Lawrence, M. J. Warren, D. Leys, N. S. Scrutton and A. W. Munro, *FEBS Lett.*, 2005, **579**, 5582-5588.
111. K. G. Ravichandran, S. S. Boddupalli, C. A. Hasemann, J. A. Peterson and J. Deisenhofer, *Science* 1993, **261**, 731-736.
112. H. Li and T. L. Poulos, *Nat. Struct. Biol.*, 1997, **4**, 140-146.
113. J. McKnight, M. R. Cheesman, A. J. Thomson, J. S. Miles and A. W. Munro, *Eur. J. Biochem.*, 1993, **213**, 683-687.
114. A. W. Munro, J. G. Lindsay, J. R. Coggins, I. MacDonald, W. E. Smith and B. N. Rospendowski, *Biochem. Soc. Trans.*, 1994, **22**, 54S.
115. I. D. G. Macdonald, W. E. Smith and A. W. Munro, *FEBS Lett.*, 1996, **396**, 196-200.
116. K. K. Khan, S. Mazumdar, S. Modi, M. Sutcliffe, G. C. K. Roberts and S. Mitra, *Eur. J. Biochem.*, 1997, **244**, 361-370.
117. J. Hudeeek, V. Baumruk, P. Anzenbacher and A. W. Munro, *Biochem. Biophys. Res. Commun.*, 1998, **243**, 811-815.
118. I. D. G. Macdonald, A. W. Munro and W. E. Smith, *Biophys. J.*, 1998, **74**, 3241-3249.
119. T.-j. Deng, L. M. Proniewicz, J. R. Kincaid, H. Yeom, I. D. G. Macdonald and S. G. Sligar, *Biochemistry*, 1999, **38**, 13699-13706.
120. I. D. G. Macdonald, W. E. Smith and A. W. Munro, *Eur. Biophys. J.*, 1999, **28**, 437-445.

121. J. Hudecek, E. Anzenbacherova, P. Anzenbacher, A. W. Munro and P. Hildebrandt, *Arch. Biochem. Biophys.*, 2000, **383**, 70-78.
122. P. Anzenbacher and J. Hudecek, *J. Inorg. Biochem.*, 2001, **87**, 209-213.
123. S. J. Smith, A. W. Munro and W. E. Smith, *Biopolymers*, 2003, **70**, 620-627.
124. Z. Chen, T. W. B. Ost and J. P. M. Schelvis, *Biochemistry*, 2004, **43**, 1798-1808.
125. T. Jovanovic and A. E. McDermott, *J. Am. Chem. Soc.*, 2005, **127**, 13816-13821.
126. T. L. Poulos, B. C. Finzel and A. J. Howard, *J. Mol. Biol.*, 1987, **195**, 687-700.
127. L. A. Cowart, J. R. Falck and J. H. Capdevila, *Arch. Biochem. Biophys.*, 2001, **387**, 117-124.
128. S. Graham-Lorence, G. Truan, J. A. Peterson, J. R. Falck, S. Wei, C. Helvig and J. H. Capdevila, *J. Biol. Chem.*, 1997, **272**, 1127-1135.
129. E. Stjernschantz, B. M. A. van Vugt-Lussenburg, A. Bonifacio, S. B. A. de Beer, G. van der Zwan, C. Gooijer, J. N. M. Commandeur, N. P. E. Vermeulen and C. Oostenbrink, *Proteins: Struct., Funct., Bioinf.*, 2008, **71**, 336-352.
130. H. Yeom and S. G. Sligar, *Arch. Biochem. Biophys.*, 1997, **337**, 209-216.
131. T. J. Volz, D. A. Rock and J. P. Jones, *J. Am. Chem. Soc.*, 2002, **124**, 9724-9725.
132. M. J. Cryle and J. J. De Voss, *Angew. Chem., Int. Ed.*, 2006, **45**, 8221-8223.
133. J. P. Clark, C. S. Miles, C. G. Mowat, M. D. Walkinshaw, G. A. Reid, S. N. Daff and S. K. Chapman, *J. Inorg. Biochem.*, 2006, **100**, 1075-1090.
134. G. Truan and J. A. Peterson, *Arch. Biochem. Biophys.*, 1998, **349**, 53-64.
135. D. C. Haines, D. R. Tomchick, M. Machius and J. A. Peterson, *Biochemistry*, 2001, **40**, 13456-13465.
136. H. Yeom, S. G. Sligar, H. Li, T. L. Poulos and A. J. Fulco, *Biochemistry*, 1995, **34**, 14733-14740.
137. W.-C. Huang, A. C. G. Westlake, J.-D. Marechal, M. G. Joyce, P. C. E. Moody and G. C. K. Roberts, *J. Mol. Biol.*, 2007, **373**, 633-651.
138. C. J. C. Whitehouse, W. Yang, J. A. Yorke, H. G. Tufton, L. C. I. Ogilvie, S. G. Bell, W. Zhou, M. Bartlam, Z. Rao and L.-L. Wong, *Dalton Trans.*, 2011, **40**, 10383-10396.
139. K. G. Ravichandran, S. S. Boddupalli, C. A. Hasemann, J. A. Peterson and J. Deisenhofer, *Science*, 1993, **261**, 731-736.
140. T. W. B. Ost, C. S. Miles, A. W. Munro, J. Murdoch, G. A. Reid and S. K. Chapman, *Biochemistry*, 2001, **40**, 13421-13429.
141. T. W. B. Ost, J. Clark, C. G. Mowat, C. S. Miles, M. D. Walkinshaw, G. A. Reid, S. K. Chapman and S. Daff, *J. Am. Chem. Soc.*, 2003, **125**, 15010-15020.
142. T. W. B. Ost, A. W. Munro, C. G. Mowat, P. R. Taylor, A. Pesseguiero, A. J. Fulco, A. K. Cho, M. A. Cheesman, M. D. Walkinshaw and S. K. Chapman, *Biochemistry*, 2001, **40**, 13430-13438.
143. M. L. Klein and A. J. Fulco, *J. Biol. Chem.*, 1993, **268**, 7553-7561.
144. A. W. Munro, K. Malarkey, J. McKnight, A. J. Thomson, S. M. Kelly, N. C. Price, J. G. Lindsay, J. R. Coggins and J. S. Miles, *Biochem. J.*, 1994, **303**, 423-428.
145. E. O'Reilly, M. Corbett, S. Hussain, P. P. Kelly, D. Richardson, S. L. Flitsch and N. J. Turner, *Catal. Sci. Technol.*, 2013, **3**, 1490-1492.
146. L. M. Podust and D. H. Sherman, *Nat. Prod. Rep.*, 2012, **29**, 1251-1266.
147. C. J. von Buehler and V. B. Urlacher, *Chem. Commun.*, 2014, **50**, 4089-4091.
148. Y.-C. Yin, H.-L. Yu, Z.-J. Luan, R.-J. Li, P.-F. Ouyang, J. Liu and J.-H. Xu, *ChemBioChem*, 2014, **15**, 2443-2449.
149. M. J. Dougherty and F. H. Arnold, *Curr. Opin. Biotechnol.*, 2009, **20**, 486-491.
150. P. A. Romero and F. H. Arnold, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 866-876.

151. M. Iwasaki, T. A. Darden, L. G. Pedersen and M. Negishi, *Biochemistry*, 1995, **34**, 5054-5059.
152. H. Furuya, T. Shimizu, K. Hirano, M. Hatano, Y. Fujii-Kuriyama, R. Raag and T. L. Poulos, *Biochemistry*, 1989, **28**, 6848-6857.
153. T. L. Poulos, *Nature*, 1989, **339**, 580-581.
154. M. Hodgkin, S. Gardner and M. Wakelam, *Biochem. Soc. Trans.*, 1993, **21**, 490S.
155. C. von Buehler, P. Le-Huu and V. B. Urlacher, *ChemBioChem*, 2013, **14**, 2189-2198.
156. E. Weber, A. Seifert, M. Antonovici, C. Geinitz, J. Pleiss and V. B. Urlacher, *Chem. Commun.*, 2011, **47**, 944-946.
157. M. Dietrich, T. A. Do, R. D. Schmid, J. Pleiss and V. B. Urlacher, *J. Biotechnol.*, 2009, **139**, 115-117.
158. M. Lisurek, B. Simgen, I. Antes and R. Bernhardt, *ChemBioChem*, 2008, **9**, 1439-1449.
159. K. T. Nguyen, C. Virus, N. Guennewich, F. Hannemann and R. Bernhardt, *ChemBioChem*, 2012, **13**, 1161-1166.
160. H. Venkataraman, S. B. A. de Beer, D. P. Geerke, N. P. E. Vermeulen and J. N. M. Commandeur, *Adv. Synth. Catal.*, 2012, **354**, 2172-2184.
161. V. Rea, A. J. Kolkman, E. Vottero, E. J. Stronks, K. A. M. Ampt, M. Honing, N. P. E. Vermeulen, S. S. Wijmenga and J. N. M. Commandeur, *Biochemistry*, 2012, **51**, 750-760.
162. A. B. Carmichael and L.-L. Wong, *Eur. J. Biochem.*, 2001, **268**, 3117-3125.
163. C. J. C. Whitehouse, S. G. Bell, W. Yang, J. A. Yorke, C. F. Blanford, A. J. F. Strong, E. J. Morse, M. Bartlam, Z. Rao and L.-L. Wong, *ChemBioChem*, 2009, **10**, 1654-1656.
164. R. J. Sowden, S. Yasmin, N. H. Rees, S. G. Bell and L.-L. Wong, *Org. Biomol. Chem.*, 2005, **3**, 57-64.
165. S. A. Maves, H. Yeom, M. A. McLean and S. G. Sligar, *FEBS Lett.*, 1997, **414**, 213-218.
166. W. L. Tang, Z. Li and H. Zhao, *Chem. Commun.* , 2010, **46**, 5461-5463.
167. Q.-S. Li, U. Schwaneberg, P. Fischer and R. D. Schmid, *Chem. - Eur. J.*, 2000, **6**, 1531-1536.
168. U. Schwaneberg, C. Schmidt-Dannert, J. Schmitt and R. D. Schmid, *Anal. Biochem.*, 1999, **269**, 359-366.
169. U. Schwaneberg, C. Otey, P. C. Cirino, E. Farinas and F. H. Arnold, *J. Biomol. Screening*, 2001, **6**, 111-117.
170. R. Fasan, M. M. Chen, N. C. Crook and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2007, **46**, 8414-8415.
171. M. Landwehr, L. Hochrein, C. R. Otey, A. Kasrayan, J.-E. Baeckvall and F. H. Arnold, *J. Am. Chem. Soc.*, 2006, **128**, 6058-6059.
172. J. C. Lewis, S. M. Mantovani, Y. Fu, C. D. Snow, R. S. Komor, C.-H. Wong and F. H. Arnold, *ChemBioChem*, 2010, **11**, 2502-2505.
173. T. S. Wong, F. H. Arnold and U. Schwaneberg, *Biotechnol. Bioeng.*, 2004, **85**, 351-358.
174. R. Agudo, G.-D. Roiban, R. Lonsdale, A. Ilie and M. T. Reetz, *J. Org. Chem.*, 2015, **80**, 950-956.
175. R. Agudo, G.-D. Roiban and M. T. Reetz, *ChemBioChem*, 2012, **13**, 1465-1473.
176. A. Ilie, R. Agudo, G.-D. Roiban and M. T. Reetz, *Tetrahedron*, 2015, **71**, 470-475.
177. A. Ilie, R. Lonsdale, R. Agudo and M. T. Reetz, *Tetrahedron Lett.*, 2015, **56**, 3435-3437.
178. G.-D. Roiban, R. Agudo, A. Ilie, R. Lonsdale and M. T. Reetz, *Chem. Commun.*, 2014, **50**, 14310-14313.
179. G.-D. Roiban, R. Agudo and M. T. Reetz, *Tetrahedron*, 2013, **69**, 5306-5311.
180. G.-D. Roiban, R. Agudo and M. T. Reetz, *Angew. Chem., Int. Ed.*, 2014, **53**, 8659-8663.

181. C. A. Denard, H. Huang, M. J. Bartlett, L. Lu, Y. Tan, H. Zhao and J. F. Hartwig, *Angew. Chem., Int. Ed.*, 2014, **53**, 465-469.
182. Y. Yang, J. Liu and Z. Li, *Angew. Chem., Int. Ed.*, 2014, **53**, 3120-3124.
183. Y. Yang and Z. Li, *Chimia*, 2015, **69**, 136-141.
184. Y. Yang, Y. T. Chi, H. H. Toh and Z. Li, *Chem. Commun.*, 2015, **51**, 914-917.
185. S. Q. Pham, G. Pompidor, J. Liu, X.-D. Li and Z. Li, *Chem. Commun.* , 2012, **48**, 4618-4620.
186. J. N. Kolev, K. M. O'Dwyer, C. T. Jordan and R. Fasan, *ACS Chem. Biol.*, 2014, **9**, 164-173.
187. K. Zhang, B. M. Shafer, M. D. Demars, H. A. Stern and R. Fasan, *J. Am. Chem. Soc.*, 2012, **134**, 18695-18704.
188. K. Zhang, S. El Damaty and R. Fasan, *J. Am. Chem. Soc.*, 2011, **133**, 3242-3245.
189. R. Fasan, N. C. Crook, M. W. Peters, P. Meinhold, T. Buelter, M. Landwehr, P. C. Cirino and F. H. Arnold, *Biotechnol. Bioeng.*, 2011, **108**, 500-510.
190. A. Glieder, E. T. Farinas and F. H. Arnold, *Nat. Biotechnol.*, 2002, **20**, 1135-1139.
191. M. W. Peters, P. Meinhold, A. Glieder and F. H. Arnold, *J. Am. Chem. Soc.*, 2003, **125**, 13442-13450.
192. S. Kille, F. E. Zilly, J. P. Acevedo and M. T. Reetz, *Nat. Chem.*, 2011, **3**, 738-743.
193. M. B. Murataliev, L. N. Trinh, L. V. Moser, R. B. Bates, R. Feyereisen and F. A. Walker, *Biochemistry*, 2004, **43**, 1771-1780.
194. C. R. Otey, M. Landwehr, J. B. Endelman, K. Hiraga, J. D. Bloom and F. H. Arnold, *PLoS Biol*, 2006, **4**, e112.
195. C. R. Otey, J. J. Silberg, C. A. Voigt, J. B. Endelman, G. Bandara and F. H. Arnold, *Chem. Biol.*, 2004, **11**, 309-318.
196. M. Landwehr, M. Carbone, C. R. Otey, Y. Li and F. H. Arnold, *Chem. Biol.*, 2007, **14**, 269-278.
197. H. Ji, W. Zhang, M. Zhang, M. Kudo, Y. Aoyama, Y. Yoshida, C. Sheng, Y. Song, S. Yang, Y. Zhou, J. Lu and J. Zhu, *J Med Chem*, 2003, **46**, 474-485.
198. J. B. van Beilen and E. G. Funhoff, *Curr. Opin. Biotechnol.*, 2005, **16**, 308-314.
199. J. G. Leahy and R. R. Colwell, *Microbiol. Rev.*, 1990, **54**, 305-315.
200. T. Sumita, T. Iida, A. Hirata, H. Horiuchi, M. Takagi and A. Ohta, *FEMS Microbiol. Lett.*, 2002, **214**, 31-38.
201. P. Meinhold, M. W. Peters, M. M. Y. Chen, K. Takahashi and F. H. Arnold, *ChemBioChem*, 2005, **6**, 1765-1768.
202. F. E. Zilly, J. P. Acevedo, W. Augustyniak, A. Deege, U. W. Haeusig and M. T. Reetz, *Angew. Chem., Int. Ed.*, 2011, **50**, 2720-2724.
203. M. M. Chen, P. S. Coelho and F. H. Arnold, *Adv. Synth. Catal.*, 2012, **354**, 964-968.
204. S. C. Maurer, K. Kuehnel, L. A. Kaysser, S. Eiben, R. D. Schmid and V. B. Urlacher, *Adv. Synth. Catal.*, 2005, **347**, 1090-1098.
205. A. Seifert, M. Antonovici, B. Hauer and J. Pleiss, *ChemBioChem*, 2011, **12**, 1346-1351.
206. A. Rentmeister, T. R. Brown, C. D. Snow, M. N. Carbone and F. H. Arnold, *ChemCatChem*, 2011, **3**, 1065-1071.
207. J. D. Bloom, S. T. Labthavikul, C. R. Otey and F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 5869-5874.
208. B. M. A. Lussenburg, L. C. Babel, N. P. E. Vermeulen and J. N. M. Commandeur, *Anal. Biochem.*, 2005, **341**, 148-155.
209. B. M. A. van Vugt-Lussenburg, E. Stjernschantz, J. Lastdrager, C. Oostenbrink, N. P. E. Vermeulen and J. N. M. Commandeur, *J. Med. Chem.*, 2007, **50**, 455-461.

210. B. M. A. van Vugt-Lussenburg, M. C. Damsten, D. M. Maasdijk, N. P. E. Vermeulen and J. N. M. Commandeur, *Biochem. Biophys. Res. Commun.*, 2006, **346**, 810-818.
211. G. Nardo, A. Fantuzzi, A. Sideri, P. Panicco, C. Sassone, C. Giunta and G. Gilardi, *JBIC, J. Biol. Inorg. Chem.*, 2007, **12**, 313-323.
212. S.-H. Park, D.-H. Kim, D. Kim, D.-H. Kim, H.-C. Jung, J.-G. Pan, T. Ahn, D. Kim and C.-H. Yun, *Drug Metab. Dispos.*, 2010, **38**, 732-739.
213. J. Reinen, J. S. van Leeuwen, Y. Li, L. Sun, P. D. J. Grootenhuis, C. J. Decker, J. Saunders, N. P. E. Vermeulen and J. N. M. Commandeur, *Drug Metab. Dispos.*, 2011, **39**, 1568-1576.
214. V. R. Dodhia, A. Fantuzzi and G. Gilardi, *JBIC, J. Biol. Inorg. Chem.*, 2006, **11**, 903-916.
215. J. S. van Leeuwen, G. Vredenburg, S. Dragovic, T. F. J. Tjong, J. C. Vos and N. P. E. Vermeulen, *Toxicol. Lett.*, 2011, **200**, 162-168.
216. J. Reinen, S. Ferman, E. Vottero, N. P. E. Vermeulen and J. N. M. Commandeur, *J. Biomol. Screening*, 2011, **16**, 239-250.
217. K. K. Dubey, A. Jawed and S. Haque, *Eng. Life Sci.*, 2011, **11**, 598-606.
218. C. G. Acevedo-Rocha, S. Hoebenreich and M. T. Reetz, *Methods Mol Biol*, 2014, **1179**, 103-128.
219. H. Venkataraman, S. B. A. de Beer, L. A. H. van Bergen, N. van Essen, D. P. Geerke, N. P. E. Vermeulen and J. N. M. Commandeur, *ChemBioChem*, 2012, **13**, 520-523.
220. L. Daviet, L. Rocci and M. Schalk, *Patent Application*, 2015, WO2015040197A2015040191.
221. D. Appel, S. Lutz-Wahl, P. Fischer, U. Schwaneberg and R. D. Schmid, *J. Biotechnol.*, 2001, **88**, 167-171.
222. S. Eiben, L. Kaysser, S. Maurer, K. Kuehnel, V. B. Urlacher and R. D. Schmid, *J. Biotechnol.*, 2006, **124**, 662-669.
223. V. B. Urlacher, A. Makhsumkhanov and R. D. Schmid, *Appl. Microbiol. Biotechnol.*, 2006, **70**, 53-59.
224. S. C. Maurer, H. Schulze, R. D. Schmid and V. Urlacher, *Adv. Synth. Catal.*, 2003, **345**, 802-810.
225. R. J. F. Branco, A. Seifert, M. Budde, V. B. Urlacher, M. J. Ramos and J. Pleiss, *Proteins: Struct., Funct., Bioinf.*, 2008, **73**, 597-607.
226. H. Schewe, B.-A. Kaup and J. Schrader, *Appl Microbiol Biotechnol*, 2008, **78**, 55-65.
227. H. Schewe, D. Holtmann and J. Schrader, *Appl. Microbiol. Biotechnol.*, 2009, **83**, 849-857.
228. A. Seifert, S. Vomund, K. Grohmann, S. Kriening, V. B. Urlacher, S. Laschat and J. Pleiss, *ChemBioChem*, 2009, **10**, 853-861.
229. Y. Watanabe, S. Laschat, M. Budde, O. Affolter, Y. Shimada and V. B. Urlacher, *Tetrahedron*, 2007, **63**, 9413-9422.
230. R. Fasan, Y. T. Meharenna, C. D. Snow, T. L. Poulos and F. H. Arnold, *J. Mol. Biol.*, 2008, **383**, 1069-1080.
231. J. A. Dietrich, Y. Yoshikuni, K. J. Fisher, F. X. Woolard, D. Ockey, D. J. McPhee, N. S. Renninger, M. C. Y. Chang, D. Baker and J. D. Keasling, *ACS Chem. Biol.*, 2009, **4**, 261-267.
232. N. Misawa, M. Nodate, T. Otomatsu, K. Shimizu, C. Kaido, M. Kikuta, A. Ideno, H. Ikenaga, J. Ogawa, S. Shimizu and K. Shindo, *Appl. Microbiol. Biotechnol.*, 2011, **90**, 147-157.
233. L. L. Wong, S. G. Bell and C. Whitehouse, *Patent Application*, 2009, WO2009047498A2009047492.
234. F. Takahashi, K. Fukuda, Y. Takimura and T. Asada, 2010, WO2010082684A2010082681.
235. G. Lubbock, Part II Thesis, University of Oxford, 2014.
236. G. Woodward, Part II Thesis, University of Oxford, 2015.
237. S. Lee, PhD Thesis, University of Kentucky, 2008.

238. T. S. Wong, N. Wu, D. Roccatano, M. Zacharias and U. Schwaneberg, *J. Biomol. Screening*, 2005, **10**, 246-252.
239. C. J. C. Whitehouse, S. G. Bell and L.-L. Wong, *Chem. - Eur. J.*, 2008, **14**, 10905-10908.
240. D. A. Rock, A. E. Boitano, J. L. Wahlstrom, D. A. Rock and J. P. Jones, *Bioorg. Chem.*, 2002, **30**, 107-118.
241. W.-C. Huang, P. M. Cullis, E. L. Raven and G. C. K. Roberts, *Metallomics*, 2011, **3**, 410-416.
242. W. T. Sulistyaningdyah, J. Ogawa, Q.-S. Li, C. Maeda, Y. Yano, R. D. Schmid and S. Shimizu, *Appl. Microbiol. Biotechnol.*, 2005, **67**, 556-562.
243. M. Fairhead, S. Giannini, E. M. J. Gillam and G. Gilardi, *JBIC, J. Biol. Inorg. Chem.*, 2005, **10**, 842-853.
244. M. Landwehr, M. Carbone, C. R. Otey, Y. Li and F. H. Arnold, *Chem. Biol.* , 2007, **14**, 269-278.
245. A. Fantuzzi, Y. T. Meharenna, P. B. Briscoe, C. Sassone, B. Borgia and G. Gilardi, *Chem. Commun.*, 2006, 1289-1291.
246. G. M. Raner, J. A. Hatchell, M. U. Dixon, T. L. Joy, A. E. Haddy and E. R. Johnston, *Biochemistry*, 2002, **41**, 9601-9610.
247. A. Rentmeister, F. H. Arnold and R. Fasan, *Nat. Chem. Biol.*, 2009, **5**, 26-28.
248. J. A. Fruetel, R. L. Mackman, J. A. Peterson and P. R. Ortiz de Montellano, *J. Biol. Chem.*, 1994, **269**, 28815-28821.
249. B. S. Kim, S. Y. Kim, J. Park, W. Park, K. Y. Hwang, Y. J. Yoon, W. K. Oh, B. Y. Kim and J. S. Ahn, *J. Appl. Microbiol.*, 2007, **102**, 1392-1400.
250. Q.-S. Li, J. Ogawa, R. D. Schmid and S. Shimizu, *Appl. Environ. Microbiol.*, 2001, **67**, 5735-5739.
251. D. C. Haines, *Protein Pept. Lett.*, 2006, **13**, 977-980.
252. D. R. Davydov, A. A. Kariakin, N. A. Petushkova and J. A. Peterson, *Biochemistry*, 2000, **39**, 6489-6497.
253. P. S. Coelho, E. M. Brustad, A. Kannan and F. H. Arnold, *Science* 2013, **339**, 307-310.
254. P. S. Coelho, Z. J. Wang, M. E. Ener, S. A. Baril, A. Kannan, F. H. Arnold and E. M. Brustad, *Nat. Chem. Biol.*, 2013, **9**, 485-487.
255. Z. J. Wang, P. Coelho and F. Arnold, 2013.
256. P. S. Coelho, E. M. Brustad, F. H. Arnold, Z. Wang and J. C. Lewis, *Patent Application*, 2014, WO2014058744A2014058742.
257. Z. J. Wang, N. E. Peck, H. Renata and F. H. Arnold, *Chem. Sci.*, 2014, **5**, 598-601.
258. R. Singh, J. N. Kolev, P. A. Sutera and R. Fasan, *ACS Catal.*, 2015, **5**, 1685-1691.
259. C. C. Farwell, J. A. McIntosh and F. H. Arnold, *Patent Application*, 2015, US20150232814A20150232811.
260. C. C. Farwell, J. A. McIntosh, T. K. Hyster, Z. J. Wang and F. H. Arnold, *J. Am. Chem. Soc.*, 2014, **136**, 8766-8771.
261. C. C. Farwell, R. K. Zhang, J. A. McIntosh, T. K. Hyster and F. H. Arnold, *ACS Cent. Sci.*, 2015, **1**, 89-93.
262. F. H. Arnold, *Quart. Rev. Biophys.*, 2015, **48**, 404-410.
263. T. K. Hyster and F. H. Arnold, *Isr. J. Chem.*, 2015, **55**, 14-20.
264. J. A. McIntosh, T. Heel, A. R. Buller, L. Chio and F. H. Arnold, *J. Am. Chem. Soc.*, 2015, **137**, 13861-13865.
265. H. Renata, Z. J. Wang and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2015, **54**, 3351-3367.
266. K. E. Atkin, R. Reiss, N. J. Turner, A. M. Brzozowski and G. Grogan, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2008, **64**, 182-185.

267. I. Rowles, K. J. Malone, L. L. Etchells, S. C. Willies and N. J. Turner, *ChemCatChem*, 2012, **4**, 1259-1261.
268. D. Ghislieri and N. J. Turner, *Top. Catal.*, 2014, **57**, 284-300.
269. C. A. Abbas and A. A. Sibirny, *Microbiol. Mol. Biol. Rev.*, 2011, **75**, 321-360.
270. V. Massey, *J. Biol. Chem.*, 1994, **269**, 22459-22462.
271. M. Held, W. Suske, A. Schmid, K.-H. Engesser, H.-P. E. Kohler, B. Witholt and M. G. Wubbolts, *J. Mol. Catal. B: Enzym.*, 1998, **5**, 87-93.
272. A. P. Jadan, M. J. H. Moonen, S. Boeren, L. A. Golovleva, I. M. C. M. Rietjens and W. J. H. Van Berkel, *Adv. Synth. Catal.*, 2004, **346**, 367-375.
273. P. K.-H. van and E. P. Patallo, *Appl Microbiol Biotechnol*, 2006, **70**, 631-641.
274. M. Hofrichter and R. Ullrich, *Appl. Microbiol. Biotechnol.*, 2006, **71**, 276-288.
275. H. Takahashi, K. Matsumoto, M. Ueda, Y. Miyake and Y. Fukuyama, *Heterocycles*, 2002, **56**, 245-256.
276. K. H. Baggaley, A. G. Brown and C. J. Schofield, *Nat. Prod. Rep.*, 1997, **14**, 309-333.
277. M. Ueki, D. P. Galonic, F. H. Vaillancourt, S. Garneau-Tsodikova, E. Yeh, D. A. Vosburg, F. C. Schroeder, H. Osada and C. T. Walsh, *Chem. Biol.* , 2006, **13**, 1183-1191.
278. M. Merkx, D. A. Kopp, M. H. Sazinsky, J. L. Blazyk, J. Muller and S. J. Lippard, *Angew. Chem., Int. Ed.*, 2001, **40**, 2782-2807.
279. S. Itoh and S. Fukuzumi, *Acc. Chem. Res.*, 2007, **40**, 592-600.
280. J. P. Klinman, *Chem. Rev.*, 1996, **96**, 2541-2561.
281. P. Chen and E. I. Solomon, *J. Am. Chem. Soc.*, 2004, **126**, 4991-5000.
282. D. E. Edmondson, C. Binda and A. Mattevi, *Arch. Biochem. Biophys.*, 2007, **464**, 269-276.
283. P. F. Fitzpatrick, *Arch. Biochem. Biophys.*, 2010, **493**, 13-25.
284. N. S. Scrutton, *Nat. Prod. Rep.*, 2004, **21**, 722-730.
285. J. R. Miller and D. E. Edmondson, *Biochemistry*, 1999, **38**, 13670-13683.
286. D. Ghislieri, A. P. Green, M. Pontini, S. C. Willies, I. Rowles, A. Frank, G. Grogan and N. J. Turner, *J. Am. Chem. Soc.*, 2013, **135**, 10863-10869.
287. F. Leipold, S. Hussain, D. Ghislieri and N. J. Turner, *ChemCatChem*, 2013, **5**, 3505-3508.
288. M. Rodriguez-Mata, A. Frank, E. Wells, F. Leipold, N. J. Turner, S. Hart, J. P. Turkenburg and G. Grogan, *ChemBioChem*, 2013, **14**, 1372-1379.
289. T. Li, J. Liang, A. Ambrogelly, T. Brennan, G. Gloor, G. Huisman, J. Lalonde, A. Lekhal, B. Mijts, S. Muley, L. Newman, M. Tobin, G. Wong, A. Zaks and X. Zhang, *J. Am. Chem. Soc.*, 2012, **134**, 6467-6472.
290. S.-W. Han, E.-S. Park, J.-Y. Dong and J.-S. Shin, *Adv. Synth. Catal.*, 2015, **357**, 1732-1740.
291. K. Hoelsch and D. Weuster-Botz, *Biotechnol. Appl. Biochem.*, 2010, **56**, 131-140.
292. M. Katzberg, N. Skorupa-Parachin, M.-F. Gorwa-Grauslund and M. Bertau, *Int. J. Mol. Sci.*, 2010, **11**, 1735-1758.
293. C. Krump, M. Vogl, L. Brecker, B. Nidetzky and R. Kratzer, *Protein Eng., Des. Sel.*, 2014, **27**, 245-248.
294. H. Li, D. Zhu, L. Hua and E. R. Biehl, *Adv. Synth. Catal.*, 2009, **351**, 583-588.
295. L. J. Ye, H. H. Toh, Y. Yang, J. P. Adams, R. Snajdrova and Z. Li, *ACS Catal.*, 2015, **5**, 1119-1122.
296. H. M. Dudek, P. Popken, E. van Bloois, W. A. Duetz and M. W. Fraaije, *J. Biomol. Screening*, 2013, **18**, 678-687.
297. S. Wu, J. P. Acevedo and M. T. Reetz, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 2775-2780.
298. K. Balke, M. Kadow, H. Mallin, S. Sass and U. T. Bornscheuer, *Org. Biomol. Chem.*, 2012, **10**, 6249-6265.

299. H. M. Dudek, M. J. Fink, A. V. Shivange, A. Dennig, M. D. Mihovilovic, U. Schwaneberg and M. W. Fraaije, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 4009-4020.
300. L. P. Parra, R. Agudo and M. T. Reetz, *ChemBioChem*, 2013, **14**, 2301-2309.
301. M. T. Reetz and S. Wu, *J. Am. Chem. Soc.*, 2009, **131**, 15424-15432.
302. D. E. Torres Pazmino, A. Riebel, J. de Lange, F. Rudroff, M. D. Mihovilovic and M. W. Fraaije, *ChemBioChem*, 2009, **10**, 2595-2598.
303. Z.-G. Zhang, R. Lonsdale, J. Sanchis and M. T. Reetz, *J. Am. Chem. Soc.*, 2014, **136**, 17262-17272.
304. Z.-G. Zhang, L. P. Parra and M. T. Reetz, *Chem. - Eur. J.*, 2012, **18**, 10160-10172.
305. Z.-G. Zhang, G.-D. Roiban, J. P. Acevedo, I. Polyak and M. T. Reetz, *Adv. Synth. Catal.*, 2013, **355**, 99-106.
306. J. G. Oakeshott, A. Devonshire, C. W. Coppin, R. Heidari, S. J. Dorrian and R. J. Russell, *Patent Application*, 2003, WO2003066873A2003066871.
307. S. Lutz and Z. Qian, *Patent Application*, 2006, WO2006086607A2006086602.
308. L. Baltzer, K. S. Broo, H. Nilsson and J. Nilsson, *Bioorg. Med. Chem.*, 1999, **7**, 83-91.
309. I. Roy, A. Sharma and M. N. Gupta, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 887-889.
310. J. Steyaert, *Eur. J. Biochem.*, 1997, **247**, 1-11.
311. J. F. Kennedy and L. Xu, *Carbohydr. Polym.*, 2006, **63**, 567.
312. T. Sakaki, *Biol. Pharm. Bull.*, 2012, **35**, 844-849.
313. C. A. Martinez and S. G. Rupashinghe, *Curr. Top. Med. Chem.*, 2013, **13**, 1470-1490.
314. S. Kumar, *Expert Opin. Drug Metab. Toxicol.*, 2010, **6**, 115-131.
315. K.-R. Kim and D.-K. Oh, *Biotechnol. Adv.*, 2013, **31**, 1473-1485.
316. D. Wahler and J.-L. Reymond, *Angew. Chem., Int. Ed.*, 2002, **41**, 1229-1232.
317. M. T. Reetz, D. Kahakeaw and R. Lohmer, *ChemBioChem*, 2008, **9**, 1797-1804.
318. V. B. Urlacher, S. Lutz-Wahl and R. D. Schmid, *Appl. Microbiol. Biotechnol.*, 2004, **64**, 317-325.
319. R. Wichmann and D. Vasic-Racki, *Adv. Biochem. Eng./Biotechnol.*, 2005, **92**, 225-260.
320. F. Hollmann, K. Hofstetter and A. Schmid, *Trends Biotechnol.*, 2006, **24**, 163-171.
321. H. Schewe, B.-A. Kaup and J. Schrader, *Appl. Microbiol. Biotechnol.*, 2008, **78**, 55-65.
322. A. Celik, D. Sperandio, R. E. Speight and N. J. Turner, *Org. Biomol. Chem.*, 2005, **3**, 2688-2690.
323. F. Ahmed, E. H. Al-Mutairi, K. L. Avery, P. M. Cullis, W. U. Primrose, G. C. K. Roberts and C. L. Willis, *Chem. Commun.*, 1999, 2049-2050.
324. K. Kuehnel, S. C. Maurer, Y. Galeyeva, W. Frey, S. Laschat and V. B. Urlacher, *Adv. Synth. Catal.*, 2007, **349**, 1451-1461.
325. T. Komatsu, H. Yamazaki, N. Shimada, M. Nakajima and T. Yokoi, *Drug Metab. Dispos.*, 2000, **28**, 1457-1463.
326. M. B. Cohen, M. R. Berenbaum and M. A. Schuler, *J. Chem. Ecol.*, 1989, **15**, 2347-2355.
327. A. Baldwin, Z. Huang, Y. Jounaidi and D. J. Waxman, *Arch. Biochem. Biophys.*, 2002, **409**, 197-206.
328. B. Testa, *Chem. Biodiversity*, 2009, **6**, 2055-2070.
329. I. Neunzig, A. Goehring, C.-A. Dragan, J. Zapp, F. T. Peters, H. H. Maurer and M. Bureik, *J. Biotechnol.*, 2012, **157**, 417-420.
330. R. B. Vail, M. J. Homann, I. Hanna and A. Zaks, *J. Ind. Microbiol. Biotechnol.*, 2005, **32**, 67-74.
331. T. H. Rushmore, P. J. Reider, D. Slaughter, C. Assang and M. Shou, *Metab Eng*, 2000, **2**, 115-125.

332. R. Bernhardt, *J. Biotechnol.*, 2006, **124**, 128-145.
333. D.-H. Kim, T. Ahn, H.-C. Jung, J.-G. Pan and C.-H. Yun, *Drug Metab. Dispos.*, 2009, **37**, 932-936.
334. V. B. Urlacher and M. Girhard, *Trends Biotechnol.*, 2012, **30**, 26-36.
335. S. T. Jung, R. Lauchli and F. H. Arnold, *Curr. Opin. Biotechnol.*, 2011, **22**, 809-817.
336. G.-D. Roiban and M. T. Reetz, *Chem. Commun. (Cambridge, U. K.)*, 2015, **51**, 2208-2224.
337. G. Di Nardo and G. Gilardi, *Int. J. Mol. Sci.*, 2012, **13**, 15901-15924.
338. H. Venkataraman, M. C. A. Verkade-Vreeker, L. Capoferri, D. P. Geerke, N. P. E. Vermeulen and J. N. M. Commandeur, *Bioorg. Med. Chem.*, 2014, **22**, 5613-5620.
339. A. M. Sawayama, M. M. Y. Chen, P. Kulanthaivel, M.-S. Kuo, H. Hemmerle and F. H. Arnold, *Chem. - Eur. J.*, 2009, **15**, 11723-11729.
340. P. C. Cirino and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2003, **42**, 3299-3301.
341. C. R. Otey, G. Bandara, J. Lalonde, K. Takahashi and F. H. Arnold, *Biotechnol. Bioeng.*, 2006, **93**, 494-499.
342. J. D. Bloom, S. T. Labthavikul, C. R. Otey and F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 5869-5874.
343. W. Chen, L. L. Koenigs, S. J. Thompson, R. M. Peter, A. E. Rettie, W. F. Trager and S. D. Nelson, *Chem. Res. Toxicol.*, 1998, **11**, 295-301.
344. R. Peter, R. Boecker, P. H. Beaune, M. Iwasaki, F. P. Guengerich and C. S. Yang, *Chem. Res. Toxicol.*, 1990, **3**, 566-573.
345. G. E. Tsotsou, A. Sideri, A. Goyal, G. Di Nardo and G. Gilardi, *Chem. - Eur. J.*, 2012, **18**, 3582-3588.
346. G. E. Tsotsou, A. E. G. Cass and G. Gilardi, *Biosens. Bioelectron.*, 2002, **17**, 119-131.
347. C. J. C. Whitehouse, S. G. Bell, H. G. Tufton, R. J. P. Kenny, L. C. I. Ogilvie and L.-L. Wong, *Chem. Commun. (Cambridge, U. K.)*, 2008, 966-968.
348. J. Yorke, PhD Thesis, University of Oxford, 2012.
349. C. Transon, S. Lecoeur, T. Leemann, P. Beaune and P. Dayer, *Eur. J. Clin. Pharmacol.*, 1996, **51**, 79-85.
350. R. Bort, K. Mace, A. Boobis, M.-J. Gomez-Lechon, A. Pfeifer and J. Castell, *Biochem. Pharmacol.*, 1999, **58**, 787-796.
351. J. C. Moore, D. J. Pollard, B. Kosjek and P. N. Devine, *Acc. Chem. Res.*, 2007, **40**, 1412-1419.
352. A. D. Rodrigues, M. J. Kukulka, E. M. Roberts, D. Ouellet and T. R. Rodgers, *Drug Metab Dispos*, 1996, **24**, 126-136.
353. M. A. Rude, T. S. Baron, S. Brubaker, M. Alibhai, S. B. Del Cardayre and A. Schirmer, *Appl. Environ. Microbiol.*, 2011, **77**, 1718-1727.
354. J. Belcher, K. J. McLean, S. Matthews, L. S. Woodward, K. Fisher, S. E. J. Rigby, D. R. Nelson, D. Potts, M. T. Baynham, D. A. Parker, D. Leys and A. W. Munro, *J. Biol. Chem.*, 2014, **289**, 6535-6550.
355. D.-S. Lee, A. Yamada, H. Sugimoto, I. Matsunaga, H. Ogura, K. Ichihara, S.-i. Adachi, S.-Y. Park and Y. Shiro, *J. Biol. Chem.*, 2003, **278**, 9761-9767.
356. R. W. Wang, D. J. Newton, T. D. Scheri and A. Y. H. Lu, *Drug Metab. Dispos.*, 1997, **25**, 502-507.
357. T. Van der Hoeven, *Anal. Biochem.*, 1984, **138**, 57-65.
358. C. J. C. Whitehouse, W. Yang, J. A. Yorke, B. C. Rowlatt, A. J. F. Strong, C. F. Blanford, S. G. Bell, M. Bartlam, L.-L. Wong and Z. Rao, *ChemBioChem*, 2010, **11**, 2549-2556.
359. A. W. Wood, D. C. Swinney, P. E. Thomas, D. E. Ryan, P. F. Hall, W. Levin and W. A. Garland, *J. Biol. Chem.*, 1988, **263**, 17322-17332.

360. P. J. Landolt, C. S. Smithhisler, H. C. Reed and L. M. McDonough, *J. Econ. Entomol.*, 2000, **93**, 1613-1618.
361. O. V. Olesen and K. Linnet, *Pharmacology*, 1997, **55**, 235-243.
362. N. Lassen and J. Perregaard, *Acta Chem. Scand.*, 1983, **37**, 335-340.
363. S. Rendic, *Drug Metab. Rev.*, 2002, **34**, 83-448.
364. G. D. Breck and W. F. Trager, *Science*, 1971, **173**, 544-546.
365. M. MacCoss and T. A. Baillie, *Science*, 2004, **303**, 1810-1813.
366. A. Nadin, C. Hattotuwagama and I. Churcher, *Angew. Chem., Int. Ed.*, 2012, **51**, 1114-1122.
367. J. P. Hughes, S. Rees, S. B. Kalindjian and K. L. Philpott, *Br. J. Pharmacol.*, 2011, **162**, 1239-1249.
368. R. Macarron, M. N. Banks, D. Bojanic, D. J. Burns, D. A. Cirovic, T. Garyantes, D. V. S. Green, R. P. Hertzberg, W. P. Janzen, J. W. Paslay, U. Schopfer and G. S. Sittampalam, *Nat. Rev. Drug Discovery*, 2011, **10**, 188-195.
369. S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg and A. L. Schacht, *Nat. Rev. Drug Discovery*, 2010, **9**, 203-214.
370. L. Pehlivan, E. Metay, D. Delbrayelle, G. Mignani and M. Lemaire, *Eur. J. Org. Chem.*, 2012, **2012**, 4689-4693.
371. P. D. Leeson and B. Springthorpe, *Nat. Rev. Drug Discovery*, 2007, **6**, 881-890.
372. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 1997, **23**, 3-25.
373. D. G. I. Kingston, G. Samaranayake and C. A. Ivey, *J. Nat. Prod.*, 1990, **53**, 1-12.
374. M. J. Waring, *Expert Opin. Drug Discovery*, 2010, **5**, 235-248.
375. M. J. Waring and C. Johnstone, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 1759-1764.
376. M. J. Waring, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2844-2851.
377. P. D. Dobson and D. B. Kell, *Nat. Rev. Drug Discovery*, 2008, **7**, 205-220.
378. P. D. Dobson, K. Lanthaler, S. G. Oliver and D. B. Kell, *Curr. Top. Med. Chem.* , 2009, **9**, 163-181.
379. S. Oswald, M. Grube, W. Siegmund and H. K. Kroemer, *Xenobiotica*, 2007, **37**, 1171-1195.
380. F. Lovering, J. Bikker and C. Humblet, *J. Med. Chem.*, 2009, **52**, 6752-6756.
381. T. J. Ritchie and S. J. F. MacDonald, *Drug Discovery Today*, 2009, **14**, 1011-1020.
382. S. L. Schreiber, *Science* 2000, **287**, 1964-1969.
383. P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler and S. L. Schreiber, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 18787-18792.
384. R. S. Heath, M. Pontini, S. Hussain and N. J. Turner, *ChemCatChem*, 2016, **8**, 117-120.
385. P. R. Balding, C. S. Porro, K. J. McLean, M. J. Sutcliffe, J.-D. Marechal, A. W. Munro and S. P. de Visser, *J. Phys. Chem. A*, 2008, **112**, 12911-12918.
386. P. R. Graves, J. J. Kwiek, P. Fadden, R. Ray, K. Hardeman, A. M. Coley, M. Foley and T. A. J. Haystead, *Mol. Pharmacol.*, 2002, **62**, 1364-1372.
387. S. Murahashi and T. Shiota, *Tetrahedron Lett.*, 1987, **28**, 2383-2386.
388. S. Yamaori, M. Kushihara, I. Yamamoto and K. Watanabe, *Biochem. Pharmacol.*, 2010, **79**, 1691-1698.
389. F. Lehmann, A. Pettersen, E. A. Currier, V. Sherbukhin, R. Olsson, U. Hacksell and K. Luthman, *J. Med. Chem.*, 2006, **49**, 2232-2240.
390. C. A. Maier and B. Wuensch, *J. Med. Chem.*, 2002, **45**, 438-448.
391. N. Sakai, T. Moriya and T. Konakahara, *J. Org. Chem.*, 2007, **72**, 5920-5922.
392. G. B. Jones, J. M. Wright, G. Hynd, J. K. Wyatt, P. M. Warner, R. S. Huber, A. Li, M. W. Kilgore, R. P. Sticca and R. S. Pollenz, *J. Org. Chem.*, 2002, **67**, 5727-5732.

393. Y. Yu, W. Yang, F. Rominger and A. S. K. Hashmi, *Angew. Chem., Int. Ed.*, 2013, **52**, 7586-7589.
394. C. B. de Koning, I. R. Green, J. P. Michael and J. R. Oliveira, *Tetrahedron*, 2001, **57**, 9623-9634.
395. J. Thien, Part II Thesis, Oxford, 2013.
396. H. Chen and E. Plettner, *Tetrahedron Lett.*, 2012, **53**, 2059-2062.
397. T. He, W.-C. Gao, W.-K. Wang and C. Zhang, *Adv. Synth. Catal.*, 2014, **356**, 1113-1118.
398. J. L. Marshall, J. P. Brooks and G. W. Hatzenbuehler, III, *J. Org. Chem.*, 1969, **34**, 4193-4195.
399. M. Bahgat, M. N. Aboul-Enein, A. A. El Azzouny, A. Maghraby, A. Ruppel and W. M. Soliman, *Acta Pol. Pharm.*, 2009, **66**, 333-340.
400. O. Tsuge, H. Watanabe, K. Masuda and M. M. Yousif, *J. Org. Chem.*, 1979, **44**, 4543-4547.
401. L. Vasvari-Debreczy, A. H. Beckett and W. Vutthikongsirigool, *Tetrahedron*, 1981, **37**, 4337-4342.
402. R. Breslow and S. H. Gellman, *J. Chem. Soc., Chem. Commun.*, 1982, 1400-1401.
403. R. Breslow and S. H. Gellman, *J. Am. Chem. Soc.*, 1983, **105**, 6728-6729.
404. S. Mathew and H. Yun, *ACS Catal.*, 2012, **2**, 993-1001.
405. N. J. Turner, *Curr. Opin. Chem. Biol.*, 2011, **15**, 234-240.
406. S. C. Dodani, J. K. B. Cahn, T. Heinisch, S. Brinkmann-Chen, J. A. McIntosh and F. H. Arnold, *ChemBioChem*, 2014, **15**, 2259-2267.
407. T. Ohshima, *Kagaku to Seibutsu*, 1992, **30**, 656-665.
408. E. W. Svastits, J. H. Dawson, R. Breslow and S. H. Gellman, *J. Am. Chem. Soc.*, 1985, **107**, 6427-6428.
409. J.-P. Mahy, J. Ciesielski and P. Dauban, *Angew. Chem., Int. Ed.*, 2014, **53**, 6862-6864.
410. R. Singh, M. Bordeaux and R. Fasan, *ACS Catal.*, 2014, **4**, 546-552.
411. T. K. Hyster, C. C. Farwell, A. R. Buller, J. A. McIntosh and F. H. Arnold, *J. Am. Chem. Soc.*, 2014, **136**, 15505-15508.
412. J. A. McIntosh, P. S. Coelho, C. C. Farwell, Z. J. Wang, J. C. Lewis, T. R. Brown and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2013, **52**, 9309-9312.
413. X. Ren, J. A. Yorke, E. Taylor, T. Zhang, W. Zhou and L. L. Wong, *Chem. - Eur. J.*, 2015, **21**, 15039-15047.
414. S. D. Nelson, G. D. Breck and W. F. Trager, *J. Med. Chem.*, 1973, **16**, 1106-1112.
415. A. C. Davis and A. L. Levy, *J. Chem. Soc.*, 1951, 3479-3489.
416. V. J. Hruby, D. Yamashiro and V. Du Vigneaud, *J. Amer. Chem. Soc.*, 1968, **90**, 7106-7110.
417. E. E. Smissman, R. L. Inloes, S. El-Antably and P. J. Shaffer, *J. Med. Chem.*, 1976, **19**, 161-163.
418. Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, V. T. Van, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, 3rd, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, 3rd, J. Kong, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Phys Chem Chem Phys*, 2006, **8**, 3172-3191.

419. G. Han, M. C. McIntosh and S. M. Weinreb, *Tetrahedron Lett.*, 1994, **35**, 5813-5816.
420. G. Han, M. G. LaPorte, M. C. McIntosh, S. M. Weinreb and M. Parvez, *J. Org. Chem.*, 1996, **61**, 9483-9493.
421. B. Eiermann, P. O. Edlund, A. Tjernberg, P. Dalen, M.-L. Dahl and L. Bertilsson, *Drug Metab. Dispos.*, 1998, **26**, 1096-1101.
422. T. Hiroi, T. Chow, S. Imaoka and Y. Funae, *Drug Metab. Dispos.*, 2002, **30**, 970-976.
423. G. Lahm and T. Opatz, *Org. Lett.*, 2014, **16**, 4201-4203.
424. O. Zhurakovskyi, PhD Thesis, University of Oxford, 2013.
425. S. Ellis, Part II Thesis, University of Oxford, 2014.