ENGINEERING OF SELECTIVE BIOHYDROXYLATION CATALYSTS VIA DIRECTED ENZYME EVOLUTION FOR GREEN AND SUSTAINABLE CHEMICAL PRODUCTION

YANG YI

(B.S. (Hons.), ECUST, China)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMICAL & BIOMOLECULAR ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2015

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Yang Yi

8 May 2015

To My Family

Acknowledgements

This thesis would not have been possible without my supervisor, Associate Professor Li Zhi, whose constant guidance, supervision, encouragement and support throughout my entire PhD candidature has enabled me to make continuous progress of this project and become an explorer in chemical and pharmaceutical research field. I also thank to my Thesis Advisory Committee members, Dr. Lee Dong-Yup, and Dr. Xie Jianping and my co-supervisors from GlaxoSmithKline, Dr. Joseph P. Adams and Dr. Radka Snajdrova for their constructive advices throughout the whole period.

Since the first day of being a member of Prof. Li lab, I have been fortunate to be around a group of wonderful and helpful colleagues. I particularly owe many thanks to Dr. Pham Quang Son for teaching me the basics of molecular biology and protein engineering. I would also like to thank Dr. Wang Zunsheng, Dr. Yan Jinyong, Dr. Wang Tianwen, Dr. Liu Ji, Dr. Li Aitao, Dr. Zhang Jiandong, Dr. Ye Lijuan, Dr. Wang Wen, Dr. Ngo Nguyen Phuong Thao, Dr. Zillillah, Dr. Wu Shuke, Dr. Gao Pengfei, Dr. Priscilia Adrian Limadinata, Mr. Zeng Shichao, Mr. Akbar Vahidi Khalfekandi, and Mr. Tian Kaiyuan for their friendship and valuable discussion during the study. I am also grateful to all the final year project students whom I had the fortune to work with, Ms. Jane Lilis Andryani, Ms. Lisa Lau Li Ling, Mr. Chi Yu Tse, Ms. Toh Swee Hoon, and Mr. Toh Hui Hung.

This thesis has benefited by many other people's efforts. The kind helps from Ms. Li Fengmei, Ms. Li Xiang, Dr. Yang Liming, Mr. Lim Hao Hiang, Joey,

Mr. Lim You Kang and Mr. Ang Wee Siong are really appreciated. Without their help, this thesis would never have been so successful. I am also grateful to the financial support provided by GlaxoSmithKline (GSK) and Singapore Economic Development Board (EDB) through a Green and Sustainable Manufacturing grant.

Last but not least, I would like to thank my parents for their endless love, support and encouragement. I also want to thank my wonderful wife You Fang, who is an immense help throughout my entire PhD candidature and source of constant encouragement. Without their heartily support and encouragement, I could not have completed this thesis.

Table of Contents

Abstract

Regio- and stereoselective hydroxylation is a very useful and important reaction in green chemistry as well as in chemical and pharmaceutical synthesis. Despite of some progresses, this type of reaction remains as a challenge for classic chemistry. Alternatively, nature provides with a useful solution for this reaction by using monooxygenases as catalysts and molecular oxygen as a cheap and green oxidant. However, there are still several problems that limit the practical application of monooxygenases, such as narrow substrate scope and unsatisfied activity and selectivity toward a given non-natural substrate. Previously, our group discovered a novel P450pyr monooxygenase from *Sphingomonas sp.* HXN-200 as a unique catalyst for biohydroxylation with great potential for pharmaceutical and fine chemical synthesis. In this thesis, we successfully engineered this P450pyr monooxygenase for the highly regio- and enantioselective subterminal hydroxylation of alkanes, an easily available and abundant feedstock. We started with the development of a powerful regio- and enantioselective colorimetric high-throughput screening (HTS) assay. Three alcohol dehydrogenases, CpSADH, PfODH, and YAD, which are highly specific for the oxidation of *(S)*-2-octanol, *(R)*-2-octanol, and 1-octanol, respectively, were identified and used in three parallel experiments to oxidise the product mixture from biohydroxylation of *n*-octane. The concentrations of produced alcohols were determined by NBT-PMS assay, and therefore allowed quick identification of mutants with improved regio- and enantioselectivity by comparing the ratio of absorbance against that of the parental enzyme. Based on the *n*-octane-P450pyr docking model, 22 amino acids residues within 6 Å of

n-octane, in the substrate access channel, and in the "big loop" were chosen for iterative saturation mutagenesis(ISM). After six rounds of evolution, several P450pyr mutants with enhanced regio- and enantioselectivity were obtained. One sextuple mutant, P450pyrSM1, showed >99% subterminal selectivity and 98% (*S*)-enantioselectivity for the hydroxylation of *n*-octane; whereas the P450pyr only gave 1-octanol due to its terminal selectivity. With a K_m of 2.187 mM and a k_{cat} of 5.9 min^{-1}, this P450pyrSM1 showed nearly the same catalytic efficiency (k_{cat}/K_m) for the subterminal hydroxylation as that of P450pyr for the terminal hydroxylation. This engineered P450pyr is the first enzyme for this type of highly selective alkane hydroxylation, and the generation of P450pyrSM1 in this study is the first successful example of the full alteration of enzyme regioselectivity and simultaneous establishment of high enantioselectivity for biohydroxylation by directed evolution.

P450pyr monooxygenase showed excellent terminal hydroxylation toward hydrophobic substrates such as alkanes, but no activity towards hydrophilic molecules such as alcohols. In order to extend its substrate range, we then engineered this P450pyr monooxygenase for the terminal hydroxylation of *n*-butanol to produce 1,4-butanediol, a useful chemical in polymer synthesis and chemical production. So far, no chemical- or biocatalyst was reported for this reaction. We started with the development of a colorimetric HTS assay for directed evolution by using 2-methoxyethanol, which is structurally similar to *n*-butanol, as a surrogate substrate. Terminal hydroxylation of 2-methoxyethanol generated an unstable diol which would decompose to ethylene glycol and formaldehyde. The amount of formaldehyde was measured at 550 nm by adding purpald to give a purple compound. All the positive mutants

selected by using the surrogate substrate-based HTS assay were further examined for biohydroxylation of *n*-butanol. In the two rounds of evolution, P450pyr single mutant I83M and double mutant I83M/I82T were found to show excellent terminal regioselectivity for the hydroxylation of *n*-butanol, with relatively high activity and no by-product formation. This gives a unique example of engineering a hydroxylase to accept a hydrophilic substrate from the original preference of a hydrophobic substrate by directed evolution.

Finally, we explored the synthetic potential of all the engineered P450pyr variants generated by directed evolution for the hydroxylation of fluoro- and other halo-substituted toluenes to produce benzyl alcohols, which are important and widely used intermediates for the productions of pharmaceuticals, fine chemicals, and agrochemicals. A P450pyr triple mutant, named as P450pyr3M, was discovered as the first enzyme with excellent activity and chemo- and regioselective for the benzylic hydroxylation of single and multiple fluoro-substituted toluenes. This enzyme also showed a broad substrate range, high activity and high conversion for the hydroxylation of chloro- and bromo-toluenes. These hydroxylations provided a simple access to the corresponding halo-benzyl alcohols, which cannot be prepared thus far by using other bio- or chemical catalysts.

List of Tables

List of Figures

List of Symbols

Ala	δ-aminolevulinic acid hydrochloride
Amp	ampicillin
cdw	cell dry weight
epPCR	error-prone polymerase chain reaction
FAD	flavin adenine dinucleotide
Fdx	ferredoxin
FdR	ferredoxin reductase
GC	gas chromatography
HPLC	high-performance liquid chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
LB	Luria-Bertani
MS	mass spectrometry
NAD(P)H	nicotinamide adenine dinucleotide phosphate
NBT	nitroblue tetrazolium
OD	optical density
ISM	Iterative saturation mutagenesis
HTS	high-throughput screening
RID	refractive index detector
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TB	Terrific Broth
WT	wild type

Chapter 1: Introduction

1.1 Green Chemistry

Chemistry is a key scientific discipline which has a great impact on everyone's daily life. It is estimated that in 2010 the chemical industry contributes about US$ 4900 billion sales worldwide and 7.7 % of the world gross domestic product (GDP), within which approximately US$ 875 billion sales are come from pharmaceuticals.[1] It has been widely acknowledged that the big impact of chemistry calls for responsibility. In order to implement the sustainable development in chemical and pharmaceutical industry, the term of "Green Chemistry" was first used at the beginning of the 1990s.[2] After that, similar concepts such as "Atom Economy" and "E Factor" were developed.[3] In 1998, Anastas and Warner introduced the "Twelve Principles of Green Chemistry": Prevention; Atom Economy; Less Hazardous Chemical Syntheses; Designing Safer Chemicals; Safer Solvents and Auxiliaries; Design for Energy Efficiency; Use of Renewable Feedstocks; Reduce Derivatives; Catalysis; Design for Degradation; Real-time analysis for Pollution Prevention; Inherently Safer Chemistry for Accident Prevention.[4] Since then, the concepts and principles of Green Chemistry have been international recognised and adopted by chemists to design and develop new products, processes and services that use less toxic and inherently safer chemicals, as well as alternative reaction conditions and solvents for improved selectivity and/or energy minimisation. The rapidly growing interest in green chemistry is partly witnessed by the establishment of the US Presidential Green Chemistry Challenge Awards in 1995, the foundation of Green Chemistry Institute in 1997, the publication of the first volume of the

Green Chemistry journal in 1999, and the growth of Green Chemistry relevant conferences and events be held in the USA, Europe and Asia.[5,6]

Although many new greener products and processes have been developed over the last 20 years, there are still various challenges ahead. Currently, many important pharmaceutical synthesis processes are no longer acceptable or still not green enough in the Green Chemistry conscious era. In 2005, the ACS Green Chemistry Institute (GCI) and seven global pharmaceutical companies (including: AstraZeneca, Eli Lilly, GSK, Johnson & Johnson, Merck, Pfizer, and Schering–Plough) organised a roundtable conference to develop a list of key research areas for future improvement.[7] One of the highlighted key challenges is: Oxidation/epoxidation methods without the use of chlorinated solvents.

1.2 Biocatalysis

Biocatalysis refers to the use of natural catalysts such as enzymes to perform complex chemical reactions of organic molecules.[8] It is a major part of biotechnology, which offers a sustainable production of food (e.g., bread, cheese, beer, vinegar), fine chemicals (e.g., amino acids, vitamins), and pharmaceuticals to meet various human needs.[9] Over the last 20 years, increasing interest has been generated in biocatalysis mainly because it is a greener and more environmental friendly approach with reduced energy consumption, waste generation and greenhouse gas emission.[10] Most of the biocatalytic processes are well aligned with the twelve principles of Green Chemistry (Table 1.1).[11]

Table 1.1 Green Chemistry *vs* Biocatalysis[11]

Green Chemistry Principle	Biocatalysis
1. Prevention	Enables more sustainable routes to APIs effectively reducing level of waste.
2. Atom economy	Enables more efficient synthetic routes.
3. Less hazardous chemical syntheses	Generally low toxicity.
4. Designing safer chemicals	No impact.
5. Safer solvents and auxiliaries	Often performed in water; when solvents are used they are generally Class I or II.
6. Design for energy efficiency	Usually performed slightly above room temperature.
7. Use of renewable feedstocks	Renewable.
8. Reduce derivatives	Chemo-, regio-, enantioselective enzymes obviates need for protecting groups.
9. Catalysis	Catalytic.
10. Design for degradation	No impact.
11. Real-time analysis for pollution prevention	No impact.
12. Inherently safer chemistry for accident prevention	Generally performed under mild conditions where risk of explosions is minimal.

Although the word "biocatalysis" is relatively new, but the concept has been utilised in practice for thousands of years in the production of beer, wine, vinegar, yoghurt and cheese.[12] The Egyptians and Babylonians produced alcoholic beverages from barley since 800 BC and the early Christian and Sanskrit writings described fermented dairy products.[13] However, throughout the centuries no one really knew that biocatalysts were involved in these processes. With the advent of the pharmaceutical industry, since 1940s biocatalysts have been widely used in the production of antibiotics (e.g., penicillin).[14] Over the past decades, major advances in microbiology and molecular biology have increased the range of available biocatalysts and their applications.[8,15,16] In particular, the development of protein engineering, including rational design and directed evolution, has enabled scientists to

effectively modify the properties of biocatalysts for targeted chemical reactions.[17-23] In the present day era, almost all the enzyme properties, including stability,[24-26] activity,[27-29] selectivity,[30-32] and substrate specificity [33-35] could be engineered routinely in the laboratory.

According to a recent business report from BBC Research, the global market for industrial enzymes is about US$ 3.3 billion in 2010. This market is expected to reach US$ 4.4 billion by 2015 with an annual growth rate of 6 %.[11] The use of lipases, esterases, and proteases is now widely established throughout industry. It is estimated that currently more than 130 different biocatalytic processes are applied in chemical, pharmaceutical, agricultural, and food industries.[36,37]

Other types of enzymes, such as oxynitrilases, aldolases, alcohol oxidases, alcohol dehydrogenases, nitrilases and cytochrome P450 monooxygenases, are starting to become more and more recognised and some of them have already been commercially available. Two particular groups of enzymes that have seen growing interest and demand are alcohol dehydrogenases for asymmetric ketone reduction to chiral alcohols; and P450 monooxygenases for selective hydroxylation of desired molecules.[11]

1.2.1 Advantages and Disadvantages of Biocatalysis *vs.* Chemical Catalysis

Similar to chemical catalysis, biocatalysis provides an alternative pathway for the conversion of the substrates to the products, in which less free energy is required to reach the transition state.[38] Nevertheless, biocatalysis offer several special characteristics over chemical catalysis.

The biggest and most important advantage of biocatalysis is its unsurpassed selectivity.[39] This selectivity is often functional (chemoselectivity), positional (regioselectivity) and chiral (enantioselectivity) specific.[40] There are numerous successful examples have shown that >99 % of a desired selectivity could be achieved routinely by using biocatalysis,[41] and in many cases >99 % of two or even three selectivities could be reached simultaneously.[42] This high selectivity becomes increasingly important, useful and desirable in the chemical and pharmaceutical industries, as it could offer numerous benefits including reduced protection/deprotection steps, minimised side reactions and by-products, easier downstream purification and separations; as well as much fewer environmental problems.[10,37]

Other advantages of biocatalysis, such as high catalytic activity, working in aqueous medium, and high active at mild reaction conditions, ambient temperature and atmospheric pressure are also attractive. Since all these benefits are in line with the goals and requirements for modern industrial processing including "Green Chemistry", "Sustainable Development", as well as "Environmentally Benign Manufacturing".[43,44]

Compared with chemical catalysis, insufficient stability in a desired media, limited biocatalysts available for the desired reactions, and serious substrate or product inhibition are the three major drawbacks of biocatalysts.[39] However, all these weaknesses are mainly due to a lack of a detailed understanding of the biocatalysis, or advances in biocatalysis have been improved steadily. Just about twenty years ago, only a few biocatalysts were commercially available, however in the present day, this number has increased more than a few hundred-fold.[8] Many of these commercially used enzymes show excellent stability with half-

lives of days, months or even years,[45] and some of them can also accept non-natural substrates and convert them into wanted products.[46,47] Besides, by using immobilisation technology, the stability of biocatalysts in the reaction mixture could be significantly improved and the separation of biocatalyst from products becomes easier at the same time.[48-50] In addition, thanks to the rapid advances in genetic engineering, all of the biocatalyst characteristics, in theory, could be tailored with rational design, directed evolution and metabolic engineering methods to meet any specific reaction criteria.

1.3 Regio- and Enantioselective Hydroxylation

Saturated hydrocarbons, such as alkanes from natural gas and crude oil, are among the world's most abundant and cheapest feedstocks.[51] Currently, hydrocarbons are mainly used as an energy source and their potential as raw materials to be transformed into more important and valuable chemicals is underutilised. This is because the C-C bonds and C-H bonds in saturated hydrocarbons are energetically stable and usually resistant to reactions with acids and bases.[52]

The advances of direct activation and functionalisation of the C-H bonds could lead to new paradigms in material and energy technologies that are more efficient and environmental friendly. Among them, the development of regio- and stereoselective hydroxylation is of particular interest, from both scientific and economic viewpoints.[53,54] Octane, for example, is one of the main constituents of petroleum and gasoline and could be considered as one of the cheapest raw materials in the world.[55] However, its subterminal hydroxylation

product, 2-octanol, is a valuable chemical reagent and useful substrate and intermediate for several industrial processes.[56-61] Besides, since this selective hydroxylation allows a functional group to be added directly at an inactive position, it is especially useful and applicable in the pharmaceutical synthesis, in which usually multi-steps are involved.[62]

Although the regio- and stereoselective hydroxylation is a very important and useful reaction for chemical industry, there are three main drawbacks in the current technologies: [63-67] First, harsh conditions and high temperatures are required; Second, the reactivity for the functionalisation of C-H bonds is insufficient; last but not least, the by-products are corrosive. Therefore the development of novel catalyst, which allows regio- and enantioselective hydroxylation of saturated C-H bond at room temperature, with high yield and few steps, would be particularly important, necessary and interesting.

1.3.1 Regio- and Enantioselective Hydroxylation with Chemical Catalysis

Fundamentally, the hydroxylation is an oxidative process that involves conversion of a C-H group to a C-OH group of an organic compound. During the past two decades, the development of selective chemical catalysts for the regio- and enantioselective hydroxylation has attracted growing interests amongst synthetic chemists. Many homogenous hydroxylations have been reported.[68-70] Within them, the rapid development of metal-complex catalysts is considered as one of the most promising directions.[71] Besides, in contrast to other presently processes, most metal-complexes catalysed reactions could occur at low temperature with high selectivity (Figure 1.1).[72-75]

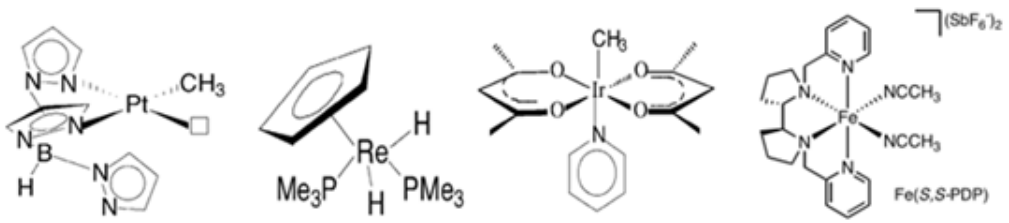


Figure 1.1 Examples of reported C-H activation with metal-complex catalysts.[72-75]

Despite these achievements, only few examples of metal-complex catalysed hydroxylation have been applied in industry since the yield and selectivity are still unsatisfied. To date, regio- and enantioselective hydroxylation remains a great challenge for classic chemistry. Generally, there are two reasons for the poor performance of metal-complex based reactions. First, the regioselectivity and stereoselectivity are difficult to control. There is almost no difference in reactivity between various C-C and C-H bonds in the saturated hydrocarbons. Therefore it is very difficult to design a chemical catalyst which could hydroxylate a desired C-H bond with specific enantioselectivity.[76] Second, the chemoselectivity is also difficult to control. In the methane oxidation process, for example, the C-H bond in methanol is about 11 kcal/mol weaker than that in methane, which implies that methanol is a better substrate for the further oxidation as compared to methane,.[77,78] Therefore, the alkane hydroxylation product usually is a mixture of alcohols, aldehydes, ketones and carboxylic acids.[79]

1.3.2 Regio- and Enantioselective Hydroxylation with Biocatalysis

Although regio- and enantioselective oxidation of C–H to C–OH is a significant challenge for chemical catalysis, but with biocatalysis this reaction is often proceeded in a simple one-step synthesis. By using this approach, so far a large

number of organic substrates have been converted to valuable products such as pharmaceuticals, flavors and fragrances and fine chemicals.[8,80] In biocatalysis, generally this type of reaction is catalysed via monooxygenases.[81-84] Monooxygenase is a sub-class of oxidoreductases which are able to insert an oxygen atom into a C-H bond by receiving an electron from NAD(P)H; while only H_2O is produced as a by-product (Figure 1.2). The necessary oxygen atom comes from molecular oxygen, which is a cheap and environmental friendly oxidant. This biohydroxylation is often highly regio- and enantioselective, which is strongly necessary for the chemical and pharmaceutical industry.[85,86]

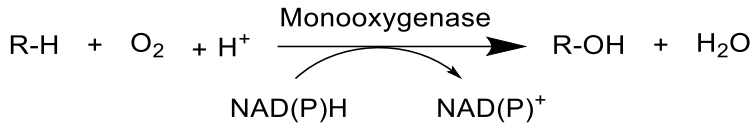


Figure 1.2 Hydroxylation reaction catalysed by monooxygenase

Some monooxygenases have been extensively studied including methane monooxygenases (MMO),[87-89] alkane monooxygenases (alkB)[90-92] and cytochrome P450 monooxygenases (P450cam,[93-95] P450BM-3,[96-101] and P450pyr[42,102-104]). While the MMO only catalyse short-chain alkanes and the alkB are often integral membrane bounded proteins, the cytochrome P450 monooxygenases have received much more attentions due to their broad substrate spectra, regio- and enantiospecific manner as well as efficient recombinant enzyme expression and production, while engineering of MMOs seems almost impossible today. However, most of the cytochrome P450 monooxygenases show terminal hydroxylation selectivity, and only a few of them demonstrate hydroxylation at non-activated subterminal position, and their regio- and stereoselectivity are unsatisfactory.[91,105,106] Besides, most of the

cytochrome P450 monooxygenases show substrate preference of hydrophobic compounds, and only two cytochrome P450 monooxygenases were reported to be capable of catalysing ω-hydroxylation of primary alcohols with very low yield and selectivity.[107] Moreover, most cytochrome P450 monooxygenases show unsatisfactory activity in the hydroxylation of non-natural substrates, such as fluoro-containing compounds.[108,109] All of these drawbacks limit the further applications of P450 monooxygenases for regio- and/or enantioselective hydroxylation.

1.4 Objectives

This thesis aims to explore the potential of creating a set of engineered P450pyr monooxygenases for highly active and efficient regio- and/or enantioselective biohydroxylation. More specifically, the thesis objectives are (Figure 1.3):

- To create P450pyr variants for the highly regio- and enantioselective subterminal hydroxylation of alkanes through directed evolution. The *n*-octane-P450pyr docking model was obtained and used to select key amino acids for iterative saturation mutagenesis (ISM) and screening. A novel, accurate, sensitive, and simple colorimetric high-throughput screening (HTS) assay was developed for the measurement of both the regioselectivity and enantioselectivity of a hydroxylation reaction. Six rounds of evolution were performed, and one P450pyr variant was created for the hydroxylation of *n*-octane to give (*S*)-2-octanol with 98 % *ee* and > 99 % subterminal selectivity. The kinetic properties of this variant was also investigated.

- To create P450pyr variants for highly regioselective terminal hydroxylation of *n*-butanol to 1,4-butanediol through directed evolution. The *n*-butanol was docked into the crystal structure of P450pyr, and key amino acids were selected for ISM. A surrogate substrate-based colorimetric HTS method was developed for the HTS. Two rounds of evolutions were conducted and the P450pyr biohydroxylation activity toward *n*-butanol was successfully created and further improved.
- To discover P450pyr variants for the highly regioselective hydroxylation of fluoro- and other halo-substituted toluenes. All the available P450pyr variants were tested for the enzymatic hydroxylation of 4-fluorotoluene. The P450pyr variant which showed the highest activity as well as regioselectivity was identified. The activity and regioselectivity of this P450pyr variant for the hydroxylation of other fluoro-, and halo-substituted toluenes as well as multiple fluoro-substituted toluenes were also investigated.

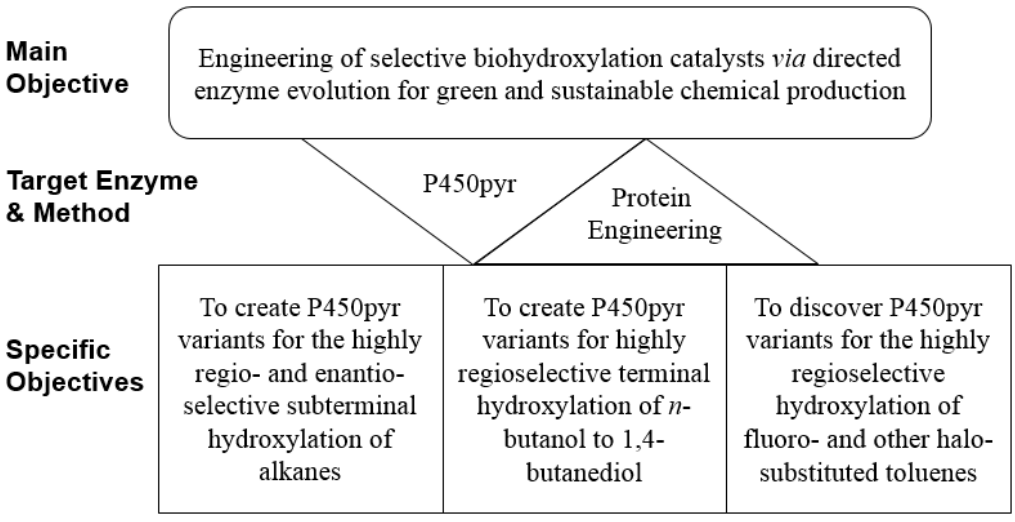
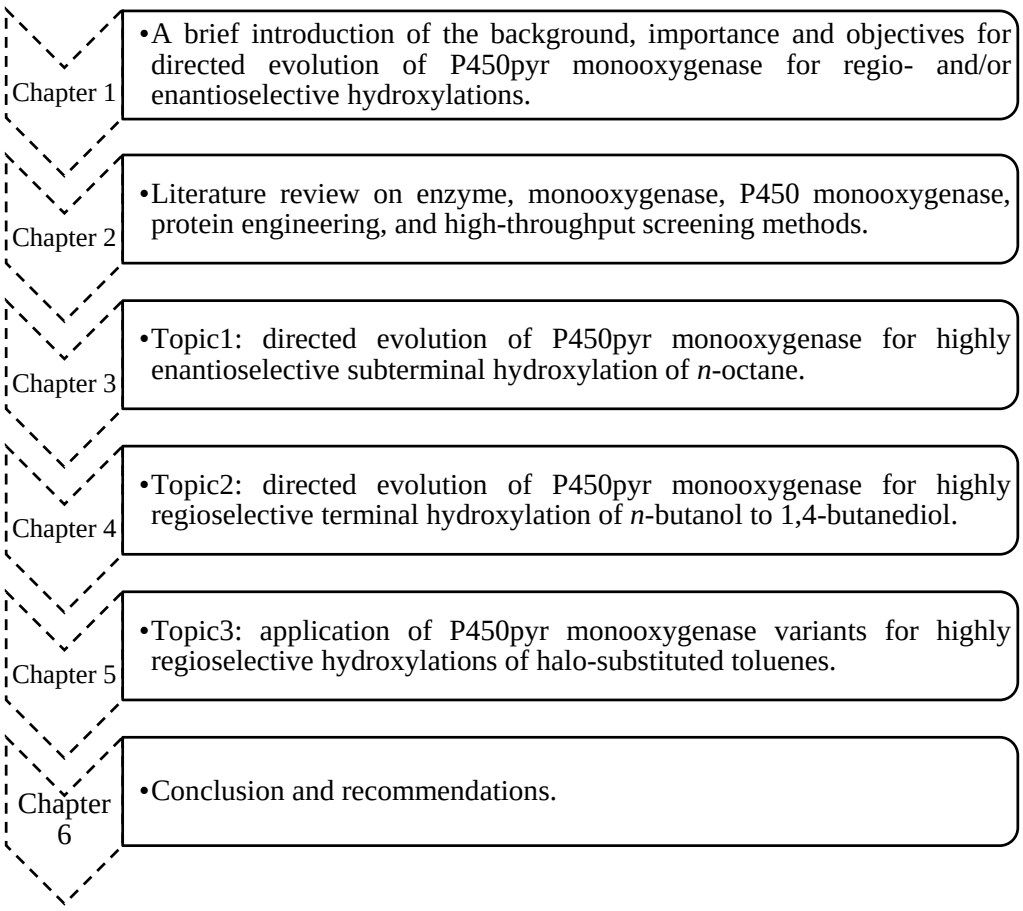


Figure 1.3 Diagram of research objectives

1.5 Outline


Chapter 1
•A brief introduction of the background, importance and objectives for directed evolution of P450pyr monooxygenase for regio- and/or enantioselective hydroxylations.
Chapter 2
•Literature review on enzyme, monooxygenase, P450 monooxygenase, protein engineering, and high-throughput screening methods.
Chapter 3
•Topic1: directed evolution of P450pyr monooxygenase for highly enantioselective subterminal hydroxylation of *n*-octane.
Chapter 4
•Topic2: directed evolution of P450pyr monooxygenase for highly regioselective terminal hydroxylation of *n*-butanol to 1,4-butanediol.
Chapter 5
•Topic3: application of P450pyr monooxygenase variants for highly regioselective hydroxylations of halo-substituted toluenes.
Chapter 6
•Conclusion and recommendations.

Chapter 2: Literature Overview

2.1 Enzymes

Enzymes are essential biological molecules in all living organisms since they are indispensable for cell metabolism, signal transduction, biosynthesis and detoxification.[110,111] Enzymes are also impressive catalysts that can catalyse a wide range of complex chemical reactions with high selectivity. Based on the types of reactions they catalyse, enzymes could be classified into six major categories (Table 2.1). [112]

Table 2.1 Enzyme classification and reactions catalysed

Enzyme Class	Group of Enzyme	Reactions Catalysed	Examples
EC 1	Oxidoreductases	Catalyse oxidation and/or reduction	Dehydrogenases, oxidases
EC 2	Transferases	Transfer of functional group from one molecule to another	Kinases, transaminases
EC 3	Hydrolases	Hydrolysis of substrates	Digestive enzymes
EC 4	Lyases	Addition or removal of functional group to a substrate	Decarboxylases, aldolases
EC 5	Isomerases	Change the molecular form of a substrate	Prolyl isomerase, fumarase
EC 6	Ligases	Joining of two molecules by covalent bond formation	Pyruvate carboxylase

It was estimated that currently about 60% of biotransformations in chemical industry rely on the use of hydrolases, and about 20% reactions using oxidoreductases.[113] With the development of green and sustainable chemical and pharmaceutical production, oxidoreductases have become increasingly popular and important due to its high reaction efficiency and low waste

generation.[114] Oxidoreductases are a class of enzymes that catalyse the transfer of electrons from one molecule (electron donor) to another molecule (electron acceptor), resulting in reactions involving electron transfer, proton abstraction, hydrogen extraction, and hydride transfer or oxygen insertion.[115] Although oxidation/reduction is one of the most useful reactions, it is still problematic by using traditional organic chemical catalysts, and therefore oxidoreductases have become promising alternatives for such challenge.[116] There are mainly four sub-classes of oxidoreductases, namely: reductases, oxidases, oxygenases and dehydrogenases (Figure 2.1).

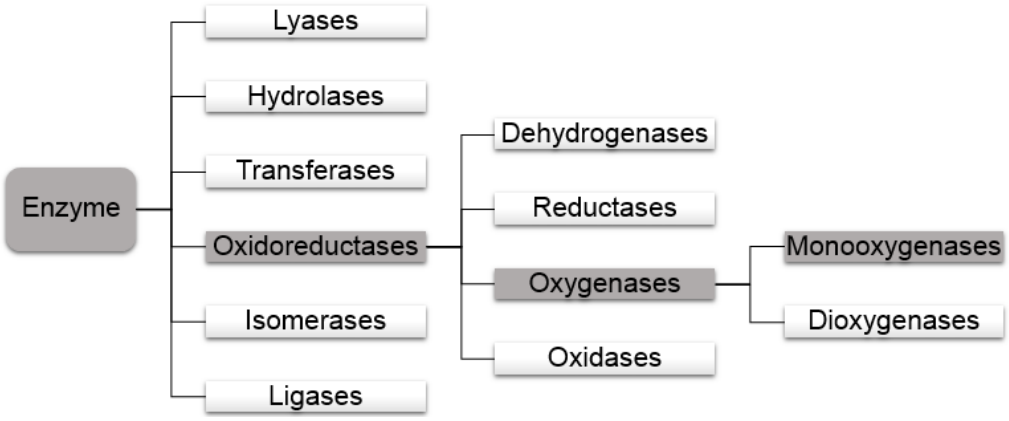


Figure 2.1 Assignment of monooxygenases to enzyme classes. Monooxygenases are associated with the grey coloured subdivisions.

Whereas dehydrogenases catalyse a substrate by transferring one hydrogen to an acceptor; reductases catalyse the reductions of substrates; oxygenases irreversibly incorporate oxygen atoms into substrates; and oxidases catalyse the transfers of electrons between different substrates. Within these enzymes as mentioned, oxygenases are of particular interest as they could directly use molecular oxygen from air as a cheap, green and non-toxic oxidant. However, so far the industrial applications of oxygenases are still limited due to their poor

stability in non-native environments and unsatisfied activity and selectivity toward non-natural substrates. With the advancements in protein engineering, the catalytic properties of several oxygenases have been successfully improved and applied in industrial chemical synthesis.[117]

2.2 Monooxygenases

Monooxygenases belong to a subclass of oxygenases; they incorporate one hydroxyl group into substrate. During this reaction, two atoms of oxygen are reduced to one hydroxyl group and one water molecule by the concomitant oxidation of NAD(P)H.[118] It is a very big and diverse group of enzymes that found in all living organisms as monooxygenases are critical for detoxification and clearance of xenobiotics which are excreted in urine after hydroxylation reactions.[116] Furthermore, in organic synthesis, monooxygenases is the general solution for carbon functionalisation, especially for the regio- and enantioselective hydroxylation, which is still problematic for classic chemistry.[119] Many monooxygenases, such as P450s (P450BM3, P450cam and P450pyr),[120-122] methane monooxygenases (MMO),[123,124] and membrane-bound alkane hydroxylase (alkB),[125,126] have been discovered and extensively investigated to exploit their functions as biocatalysis, and some of them have already been engineered for different target reactions.[127-129]

2.2.1 Methane Monooxygenase (MMO)

Methane monooxygenase (MMO) belongs to the class of oxidoreductases which is capable of oxidising the C-H bond in methane and other small

alkanes.[130] There are two forms of MMO, one is soluble methane monooxygenases (sMMO) which contains a Fe-O-Fe centre,[131] and the other form is particulate methane monooxygenases (pMMO) whose active site utilises copper.[132] These two well-known MMO were identified from *Methylococcus capsulatus* Bath [133-136] and *Methylosinus trichosporium* OB3b, respectively.[137-140] Both of the two MMO could hydroxylate a series of alkanes to alcohols by using NADH as cofactor.[141] However, their highest activity is only obtained with methane as the substrate, and their activity will significantly decrease when the substrate chain length increases.

2.2.2 Alkane Hydroxylase (AlkB)

Similar to MMO, alkane hydroxylases (AlkB) is a group of enzymes which are responsible for the hydroxylation of medium to long chain alkanes. It is an integral membrane-bound non-heme iron monooxygenase, usually consists of three components: AlkB, rubredoxin and rubredoxin reductase. For instance, GPo1 from *Pseudomonas oleovorans* is one of the most famous AlkB enzyme which catalyses a wide range of substrates from alkanes to fatty acids.[125,126] Although the substrate scope of AlkB is wide, the industrial application of this class of enzymes is still limited due to its complexity and poor solubility. As a membrane-bound enzyme, the solubility of AlkB in the aqueous medium is very low, which strongly limit its industrial applications. Furthermore, membrane-bound enzymes are generally not suitable for protein engineering as it is difficult to obtain its crystal structure and also challenging to design rationally for directed evolution.[142,143]

2.2.3 Cytochrome P450 Monooxygenase

Cytochrome P450 monooxygenase (CYP or P450s) is one of the largest and oldest superfamily of heme-containing enzymes that utilises molecular oxygen and NAD(P)H to carry out the oxidative insertion of a single oxygen atom into X-H bonds (X: C, N or S) of an organic substrate with the concomitant reduction of the other oxygen atom to water.[144,145] Interestingly, these enzymes are named not based on their functions but rather due to their characters: "cytochrome" represents a hemoprotein, "P" stands for pigment, and "450" reflects the typical absorption spectra of the reduced CO-bound CYP complex which occurs at 450 nm.[146] Remarkably, P450s have a wide substrate scope including alkanes, terpenes, alkaloids, fatty acids; and they also catalyse oxidation (hydroxylation, epoxidation, dealkylation, etc) often in a regio- and/or enantioselective manner under mild conditions.[127,147,148] To date, 267 families with more than 7000 P450s have been identified and sequenced. Many soluble bacterial P450s are overexpressed in heterologous hosts for synthesis, and some of them have already been characterised with 3D structures.[149,150]

2.2.3.1 Cytochrome P450 Monooxygenase Classes

Based on how electrons from NAD(P)H are delivered to the catalytic site, P450s could be mainly classified under four classes (Figure 2.2), although other subclasses are occurring as well: bacterial type, mitochondrial type, microsomal type and one special self-sufficient type.

The bacterial type P450s is a three-component system with flavin adenine dinucleotide (FAD) containing a reductase, an iron-sulphur redoxin and the heme-containing P450 enzyme (Figure 2.2A). This type of P450s is involved in

the catabolism of compounds used as carbon source,[151,152] and production of biologically active secondary metabolites such as antibiotics.[84,153]

Similar to bacterial type, mitochondrial type P450s also contains three separate parts: a FAD-containing reductase, which transfers reduction equivalents from NAD(P)H; a ferredoxin, which in turn reduces the cytochrome P450 itself. However, in bacterial type P450s all three proteins are soluble (Figure 2.2A); while in mitochondrial type P450s only the ferredoxin is a soluble protein, whereas both the reductase and the cytochrome P450 are bounded to the inner of mitochondrial membrane (Figure 2.2B). In mammals, this type of P450s plays an essential role in the biosynthesis of the cholesterol-derived steroid hormones, vitamin D and bile acids.[154] One example of mitochondrial P450s is CYP11B1, which responses for the formation of cortisol by the 11β-hydroxylation of its precursor 11-deoxy-precursors.[155-157]

The microsomal type P450s is the most common groups of P450s that located in the endoplasmic reticulum (ER) of mammals. This type of P450s contains two integral membrane proteins (Figure 2.2C): the P450 and the microsomal NADPH-P450 reductase containing FAD and FMN.[158] The substrate range of this type of P450s is extremely diverse: including fatty acids, steroids, as well as exogenous compounds such as therapeutic drugs and environmental toxicants.[159]

The self-sufficient type of P450s is an unusual species where the P450 itself is bound to its reductase component into a single polypeptide chain, and thus is catalytically self-sufficient (Figure 2.2D). The first and most famous member of this self-sufficient group of P450s is CYP102A1 (P450BM3), which was

discovered from soil bacterium *Bacillus megaterium* in 1980s.[160] The self-sufficient type of P450s (e.g. P450BM3 and P450Rhf) has been discovered in both prokaryotes and eukaryotes, and interestingly all of them show similar substrate scope, molecular weight, catalytic turnover, and many other enzymatic characters.[161-164]

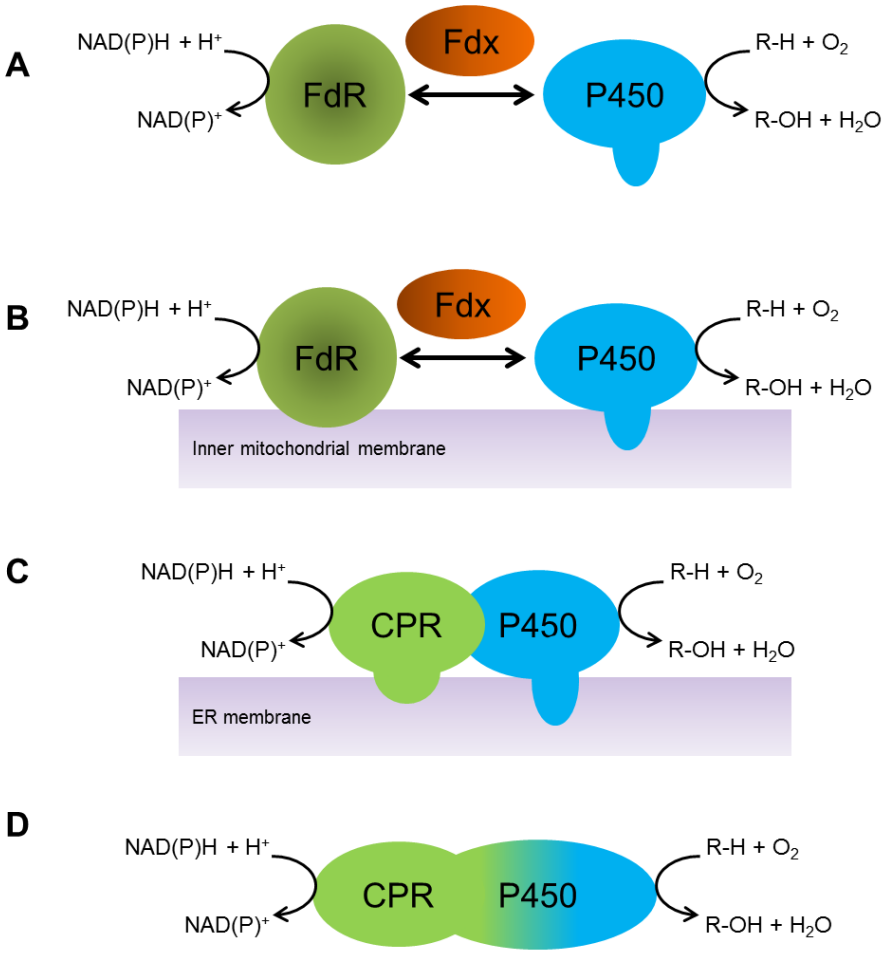


Figure 2.2 Schematic representation of the different cytochrome P450 systems. A: bacterial system, B: mitochondrial system. C: microsomal system, D: self-sufficient system (CPR fusion type).

2.2.3.2 Hydroxylation Mechanism of Cytochrome P450 Monooxygenase

Although the detail information of different P450s catalysed reactions is still growing, the hydroxylation mechanism has been well understood and could be

generally described in Figure 2.3.[165-167] The catalytic cycle is considered to start with the binding of substrate to the active site (heme centre) of the P450s. Then an electron is transferred from NAD(P)H and ferric iron is reduced at the same time. Subsequently, a molecular oxygen binds to the ferrous iron group, and another electron is accepted to produce a negatively charged iron-peroxo intermediate. This intermediate is unstable and will quickly be protonated to create a highly reactive iron-hydro-peroxo complex, which also known as P450 Compound 1 [Fe(V)=O]. This intermediate allows substrate attack to generate a hydroxylated product which would detach from the heme centre.[168] Finally, a water molecule binds to the P450s heme centre, and the enzyme returns to the initial state for the next cycle of reaction.

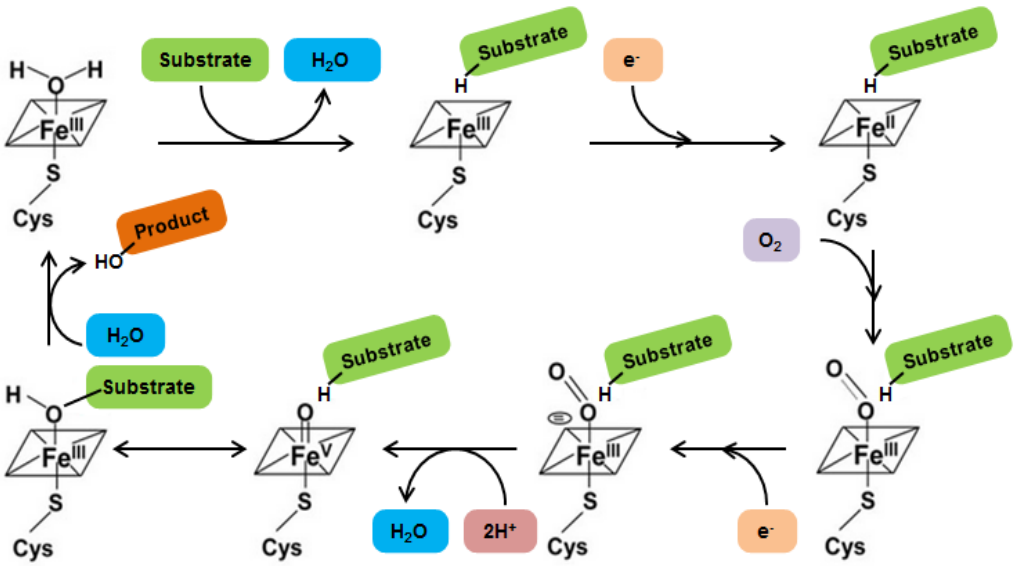


Figure 2.3 The catalytic cycle of cytochrome P450

2.2.3.3 Cytochrome P450cam Monooxygenase

P450cam is a typical bacterial type of P450s which consists of three separate parts: putidaredoxin, P450cam and iron-sulphur ferredoxin which transfer

electrons between putidaredoxin and P450cam.[169] This enzyme was isolated from *Pseudomonas putida* and it converts camphor to 5-*exo*-hydroxycamphor natively, hence the name P450cam.[170]

The substrate scope of wild-type P450cam, however, is quite narrow due to the existence of complementary interactions in P450cam activity site. Since the identification of P450cam crystal structure,[171] protein engineering has been extensively used to increase the activity and selectivity of P450cam for the oxidation of non-natural substrates (Figure 2.4). Many novel mutated enzymes had been created that showed selective hydroxylation activity toward phenyl derivatives, while wild-type P450cam did not accept any of these compounds as substrate for hydroxylation.[172] Other valuable mutants, which showed very high activity toward small and medium chain alkanes, were also generated via the replacement of certain key amino acid residues with bulky amino acid residues.[173-176]

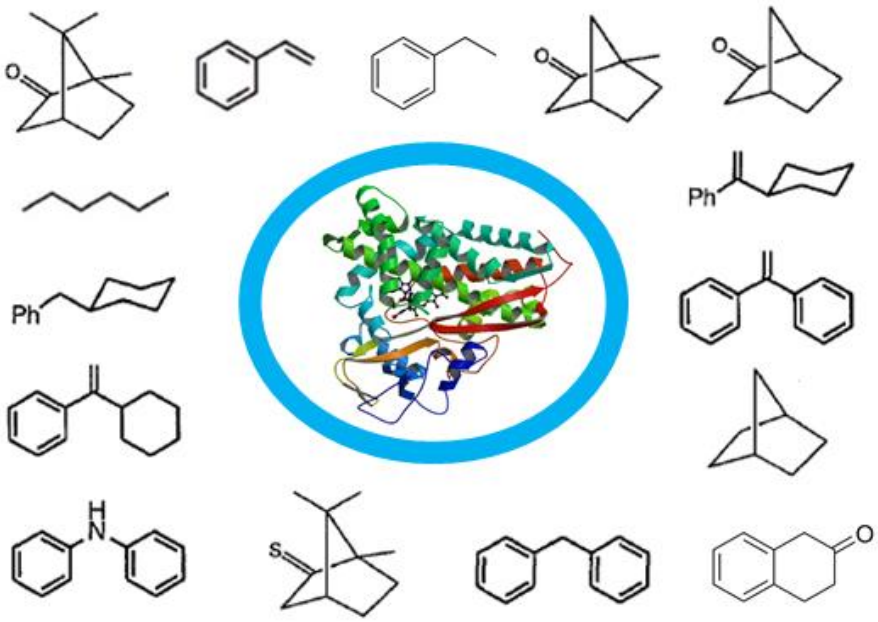


Figure 2.4 The substrate scope of P450cam and its variants.

2.2.3.4 Cytochrome P450BM3 Monooxygenase

P450BM3(CYP102A1), a fully soluble fatty acid hydroxylase with a molecular weight of 119 kDa, is the most extensively examined and engineered P450s owing to the following unique properties.[177] [178] Firstly, high activity, the hydroxylation of fatty acid catalysed by P450BM3 is the most rapid P450-catalysed hydroxylation reactions known thus far (k_{cat} = 17,000 min^{-1}).[161] Secondly, solubility, P450BM3 is not membrane-bound but is conveniently fused to its reductase, which makes the expressing better.[127] Last but most importantly, P450BM3 tolerates mutations which allows extensive protein engineering studies to be carried out. To date, many interesting P450BM3 mutants have been (and are still being) created which are capable of: accepting non-natural substrates (Figure 2.5); exhibiting very high activity with very high regio- and enantioselectivity toward target substrates.[63,179-183] The natural substrate of P450BM3 is fatty acids containing 12-22 carbons, and the products are mixture of ω-1, ω-2 and ω-3 hydroxylated fatty acids.[184] With directed evolution, a library of P450BM3 variants with measurable activity on various linear terpenes, cyclical monoterpenes, and cyclical sesquiterpenes were successfully created.[185] Further work on P450BM3 evolution had expanded its substrate specificity from fatty acids to other non-natural substrates such as alkanes, aromatics and steroids.[121,168] In addition, the stability of P450BM3 was also improved when six amino acids mutation were introduced in its active site. This new created thermos-stabilised P450BM3 could increase the half-life from 3 min to 136 min at 50 ^{o}C.[186]

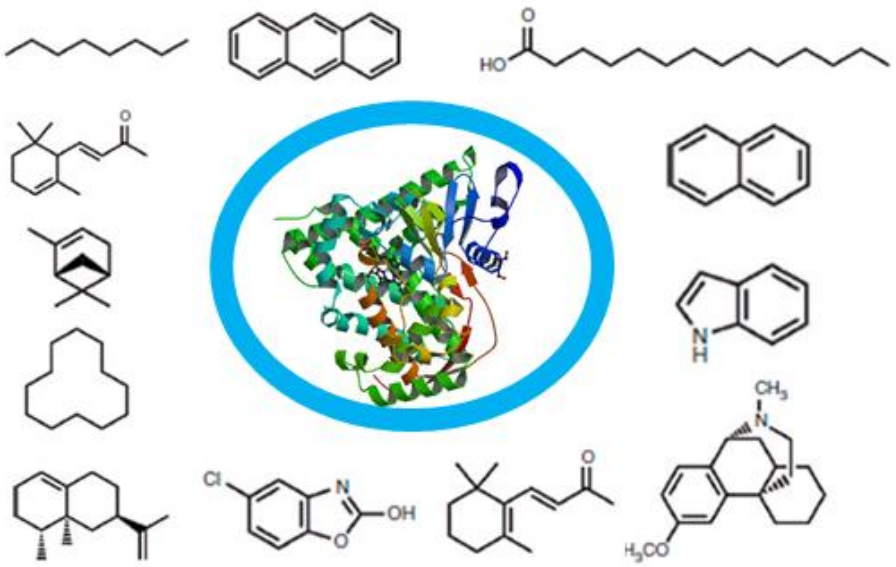


Figure 2.5 The substrate scope of P450BM3 and its variants.

2.2.3.5 Cytochrome P450pyr Monooxygenase

The cytochrome P450pyr monooxygenase was isolated from *Sphingomonas* sp. HXN-200, which was discovered by Li *et al*. in 1999 for the regio- and stereoselective hydroxylation of *N*-benzyl pyrrolidine.[187,188] This soluble P450pyr monooxygenase is a member of the cyp153 family and belongs to the class I of P450s which need an extra electron-delivering system (ferredoxin and ferredoxin reductase protein) (Figure 2.6). Besides, this family of P450s can also be functionally fused to the CPR domain of P450BM3.[107]

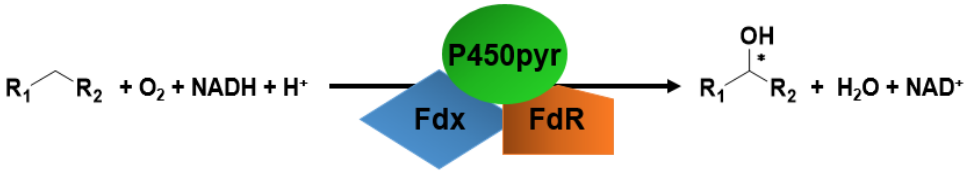


Figure 2.6 Regio- and stereoselective hydroxylations with P450pyr monooxygenase system. P450:P450pyr hydroxylase, Fdx: ferredoxin, and FdR: ferredoxin reductase

The discovery of P450pyr is of great interest due to its high activity and excellent regio- and stereoselectivities on a variety of alicyclic substrates such as piperidines, azetidines, 2-pyrrolidinones and 2-piperidinones.[189-191] Medium chain alkanes from pentane to nonane are also accepted as substrates, and most importantly, P450pyr catalyses the hydroxylation of these alkanes at terminal position exclusively. Besides, this P450pyr is also a good epoxide hydrolase which converts *N*-benzyloxycarbonyl-3,4-epoxy-pyrrolidine and cyclohexene oxide to their corresponding vicinal trans-diols in high *ee* and high yield.[192] Its wide substrate range makes P450pyr an suitable target enzyme for directed evolution.

While the substrate scope of P450pyr is diverse, it is confined to mostly medium-sized hydrophobic substrates (Figure 2.7). No activity towards smaller molecules, such as *n*-butane or hydrophilic molecules such as alcohols has been discovered. This phenomenon is not surprising as the primary role of P450s is detoxification by hydroxylation of hydrophobic molecules to more hydrophilic alcohols. Recently, several new P450pyr mutants were generated by directed evolution with improved *S*-enantioselectivity or *R*-enantioselectivity for the hydroxylation of *N*-benzyl pyrrolidine, enhanced regioselectivity for the hydroxylation of *N*-benzyl pyrrolidinone, and increased enantioselectivity for the hydroxylation of *N*-benzyl piperidinone, respectively.[102,103] One engineered P450pyr variant was also found to catalyse the asymmetric epoxidation of *para*-substituted styrenes, unconjugated 1,1-disubstituted alkenes and cyclic alkenes, being the first enzyme to give high enantioselectivity, activity and high conversion.[192]

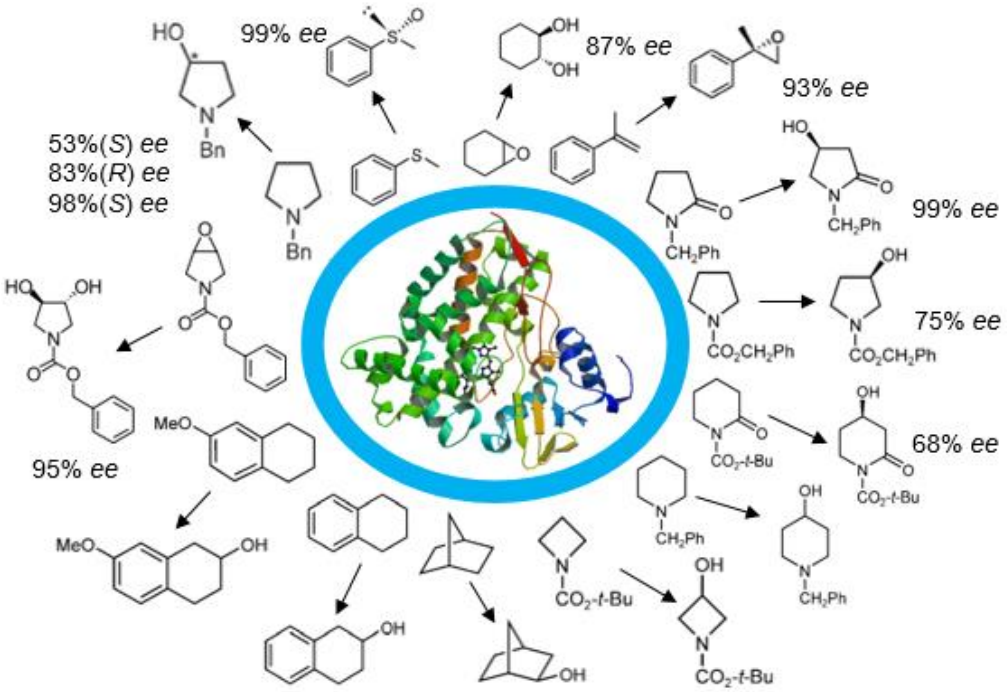


Figure 2.7 The substrate scope of P450pyr and its variants.

2.3 Protein Engineering

Most of the natural enzymes are generally unsuitable as catalysts for industrial processes, which often occur at harsh and extreme conditions that will cause the enzymes to degrade and denature; furthermore, natural enzymes also have weak activity and selectivity toward non-natural substrates.[193] Since 1990s, protein engineering has become one of the most powerful and promising tools to adapt natural enzymes for industrial applications. Based on the use of recombinant DNA technology, protein engineering could create new enzymes with desirable properties, including: substrate specificity, improved kinetics, high stability, good solvent tolerance and substrate/product inhibition. In general, protein engineering can be accomplished through two experimental routes - rational design and directed evolution.

2.3.1 Protein Engineering by Rational Design

The most classical approach in protein engineering is via rational design, which always relies on the information of the protein structure, knowledge about protein folding, dynamics as well as structure-function relationship.[194] It utilises computational design to predict optimal mutations at specific target sites in the protein, followed by site-directed mutagenesis to confirm the prediction. It is an effective and less labour intensive approach, since only certain key amino acid residues have to be isolated and changed to other amino acids with different properties.[195-197]

The catalytic performance of many enzymes have been successfully improved through rational design approach.[198-200] For example, by using a computer model to illustrate key mutations in the active site and substrate access channel, Keasling *et al* developed of a four-point P450BM3 mutant, which increased the coupling efficiency for epoxidation of amorphadiene from 35% to 63% and the epoxidation rate from 8 to 30 per min.[201] However, in many cases, the crystal structure and catalytic mechanism of the enzyme of interest is not available. Under this limited or absent information condition, protein engineering can still be accomplished through another approach: directed evolution.

2.3.2 Protein Engineering by Directed Evolution

Directed evolution is an alternative approach to bridge the gap between natural enzymes and desired enzymes. It involves a creation of genetic diversity, followed by screening and selection for mutants with the desired features.[202] It is an iterative Darwinian optimisation process, where only the "fittest" variants are iteratively selected from a mutant library.[25,203-205]

Compared with millions years of natural evolution, directed evolution is certainly a fast approach that can generate desired enzymes from a greater number of parents within several months' time. In contrast to rational design, directed evolution does not require the knowledge of protein structural nor the mechanism of catalysis, and can be carried out with just the knowledge about the enzyme's genetic sequence.[206] A general comparison of these two methods is summarised in Figure 2.8.

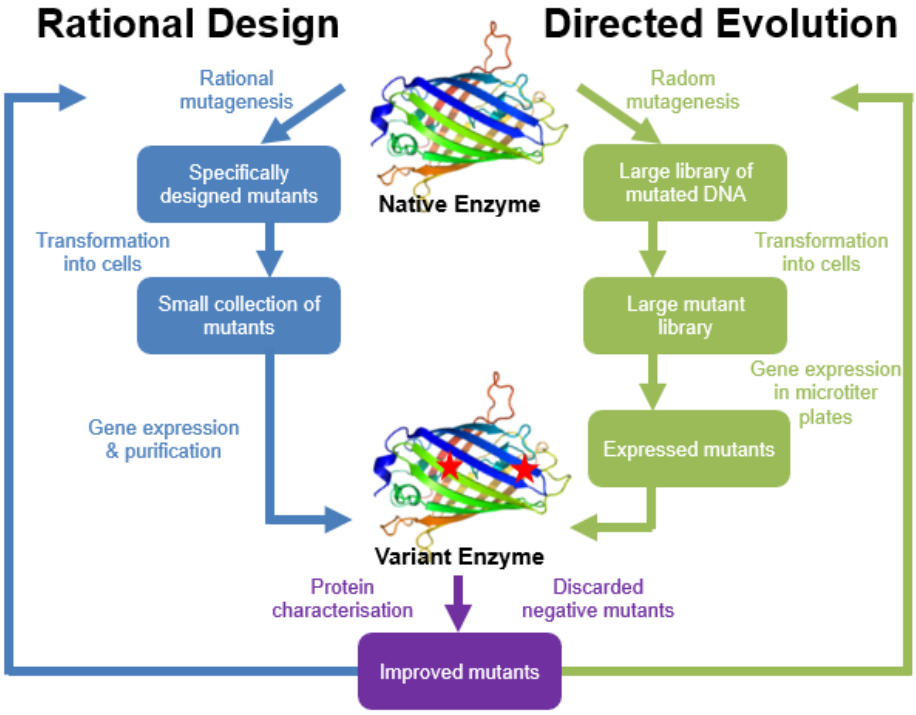


Figure 2.8 A comparison of directed evolution and rational design processes.

In the present day, directed evolution has become increasingly popular with the development of molecular biology technologies and high-throughput screening (HTS) techniques. Figure 2.9 summarises the three main mutant library generation methods: random mutagenesis, semi-rational design, and gene shuffling. Random mutagenesis could introduce mutations at any positions throughout the whole gene sequence, while semi-rational design only mutates

genes at specific position of interest. Moreover, gene shuffling could exchange the DNA fragments from different parental DNA sequences. As molecular biology technologies continue to improve, the boundary between rational design and directed evolution has become more and more indistinct. Researchers have combined both methods to generate many new enzymes with better catalytic properties.[207,208] One of such techniques that combines rational design and directed evolution is iterative saturated mutagenesis(ISM). In this method, key amino acids are identified based on the protein structure and then exchanged with all other amino acids generating a smaller library size compared with gene shuffling or error-prone PCR. Besides, this method has been proven to be more efficient than gene shuffling for the directed evolution of β-Galactosidase.[209]

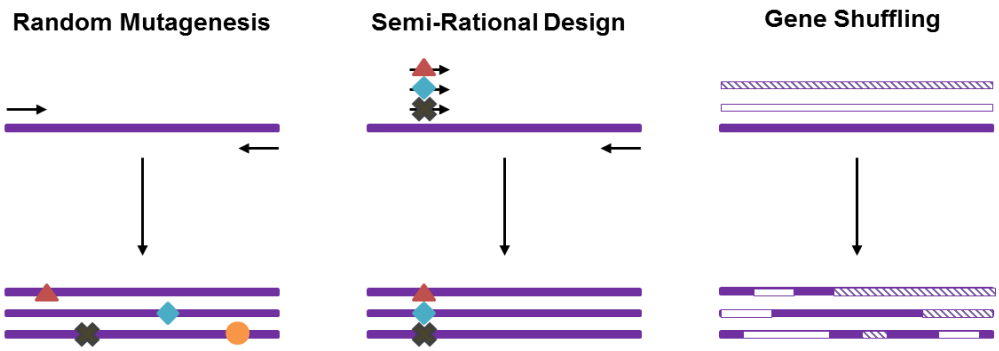


Figure 2.9 Three main mutant library creation strategies: random mutagenesis; semi-rational and gene shuffling.

2.3.2.1 Iterative Saturation Mutagenesis (ISM)

In order to address the limitations of both rational design and directed evolution, in 2007 Reetz *et al* introduced a new and efficient "semi-rational" approach for the protein engineering, iterative saturation mutagenesis (ISM).[205] Based on the available protein structure, iterative cycles of saturation mutagenesis are performed at rationally selected sites; this method could reduce the necessary

mutant library generation and the screening effort drastically. In general, there are five steps in the ISM process (Figure 2.10):

1. Based on the enzyme structural data, identify the key amino acids that may be important for improving certain property;
2. Perform individual site-saturation mutagenesis of all selected residues;
3. Screen all the mutants with HTS and then identify the best hit;
4. The best mutant is used as the template for another round of saturation mutagenesis at the remaining unmutated sites, and repeat steps 2-4 until no further improvement can be achieved.
5. Analyse the relationship between mutated amino acids and created properties.

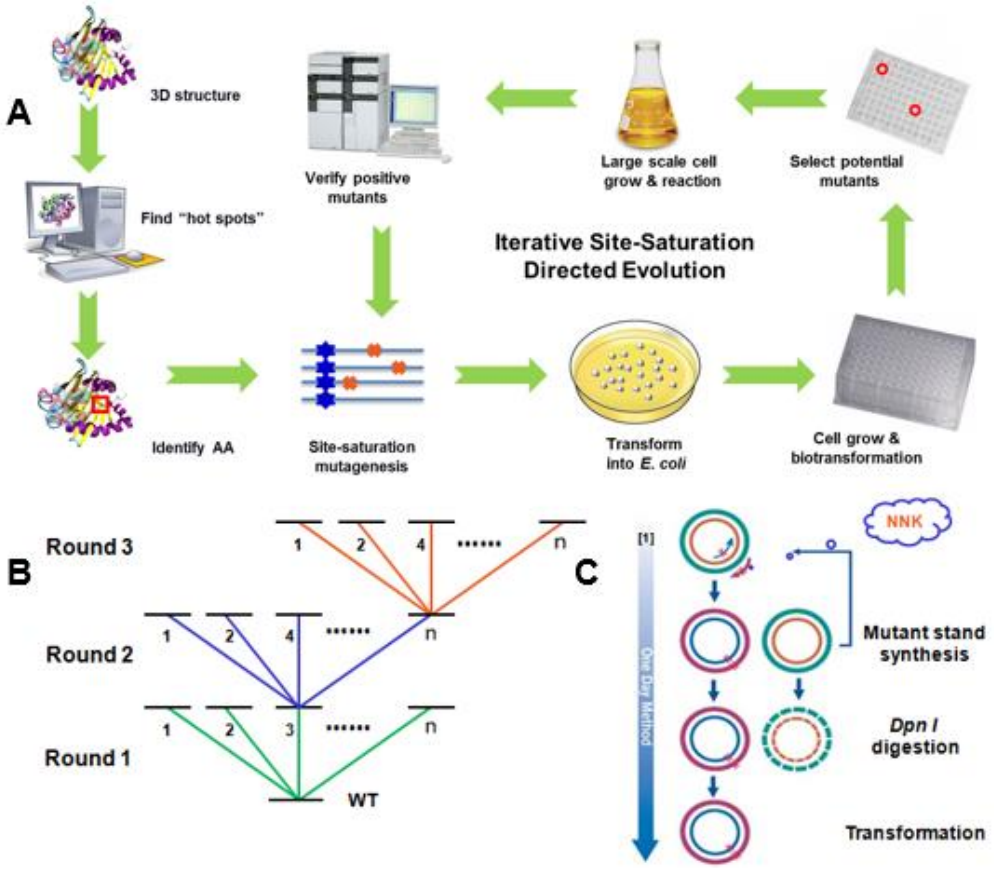


Figure 2.10 General process (A) and mechanism (B&C) of ISM

Since it was reported, the catalytic performances of many enzymes have been successfully engineered through ISM method, such as: increased stability of lipase from *Bacillus subtilis*;[24] broadened substrate scope of the old yellow enzyme homologue YqjM with improved enantioselectivity for bioreduction of substituted cyclopentenone and cyclohexenone derivatives.[210] This ISM method has two unique advantages. First, it ensures all important amino acid substitutions at each critical site. Degenerated primers were designed at certain amino acid residues with the sequence of NNK and other 9-12 bases on the left and right (Figure 2.10C). The NNK sequence encodes for all 20 amino acids (32 possible codon substitutions). Second, it maximises the probability of obtaining cooperative effects of newly introduced mutations in a minimised mutant library size. This ISM process (Figure 2.11B) is significantly different from the common method of using multiple cycles of epPCR or gene shuffling. For the latter two methods, they address the whole enzyme sequence, which means that in each cycle all the amino acids are covered over and over again, though maybe only very few sites are really useful for certain properties. However, a clear limitation of this ISM approach will become obvious when attempting to engineer an enzyme of which no structural data (X-ray or homology model) is available. In such cases epPCR or DNA shuffling may be the method of choice.

One specific form of the ISM is combinatorial active site saturation test (CASTing), which was originally developed with the purpose of expanding the range of substrate acceptance of an enzyme.[211] Usually there are only three steps in the CASTing:

1. Based on the 3D structure of the enzyme, two or three amino acids, whose side chains reside next to the binding pocket, are identified.
2. Each positions are randomised simultaneously with the creation of relatively small libraries of mutants.
3. Screen all the mutants with HTS and identify the best hit.

Compare to typical ISM, usually two or three amino acids (rather than only one) are simultaneously randomised in CASTing, and therefore this approach theoretically allows the possibility of cooperative effects.[212] Currently this method is considered as a useful alternative to epPCR as the starting point of directed evolution studies.

2.3.3 Screening and Selection

There are two steps in protein engineering: introduction of genetic diversification to create the variation, and selection of variants with desirable properties. Normally the later step is more challenging as it usually calls for the screening of thousands of mutants, which is tedious and time consuming.[213] Therefore, many effective HTS and selection techniques have been developed to quickly "get a needle in a haystack."

One of the most commonly used methods is microplate based screening due to their flexibility and automation.[214] Variants from the mutant library are inoculated and grown in a 96-well plate, where enzymes of interest are expressed and subjected to an HTS assay based on UV-absorption, colorimetric, luminescence, or fluorescence.[215] Many standardised equipment are commercially available to quantitatively analyse each screened mutant in a 96-well plate, and mutants with desired absorbance/fluorescence could be

identified and selected efficiently. Many other HTS and selection methods have also been reported, such as: agar plate growth-based and flow cytometry-based selection and screening, and whole-cell based compartmentalization screening approach.[216] However, all these reported methods are highly dependent on the substrate/enzyme combination, and therefore understanding the advantages, disadvantages and limitations of each method at different conditions is crucial to choose the best HTS and selection approach.

In addition, many HTS methods selection processes involve a balance between substrate selection, assay complexity, and detection limit. Sometimes a "surrogate" substrate has to be used in HTS, though the screening process becomes more convenient, this route might tailor the enzyme towards the "surrogate" substrate not the "target" substrate.[217] In general, a well-designed HTS screening system is essential for successful protein engineering, which reduce experimental efforts and increase the chances to identity a mutant with desirable properties.

2.3.3.1 Screening Methods to Determine the Enantioselectivity

There are a few screening and selection methods reported for enzyme enantioselectivity determination.[218-222] The most direct and traditional method to analyse an enzyme's enantioselectivity is to determine the product *ee* by either chiral GC or HPLC. Though both methods could provide precise data, these two methods are still considered as "medium throughput", since normally only about 200 samples could be measured per day per machine. Several indirect methods have been developed to quickly estimate the product *ee* in a more efficient way. For instance, product *ee* could also be determined by further

transformation of product enantiomers with enantioselective enzymes and then detected by UV spectroscopy or IR thermography.[223,224]

Based on this idea, Li *et al* created an HTS method for the determination of alcohol *ee* by utilising two complementary alcohol dehydrogenases.[225] Recently, NBT-PMS colorimetric assay was then introduced to make this concept more easily applicable in directed evolution, and a two-enzyme based colorimetric HTS assay was developed for the asymmetric biohydroxylation.[226] By using this assay, the enantioselectivity of P450pyr for the biohydroxylation of *N*-benzyl-3-pyrrolidine was inverted from *ee* of 43% (*S*) to 83% *ee* (*R*) with three rounds of evolutions (Figure 2.11). [226]

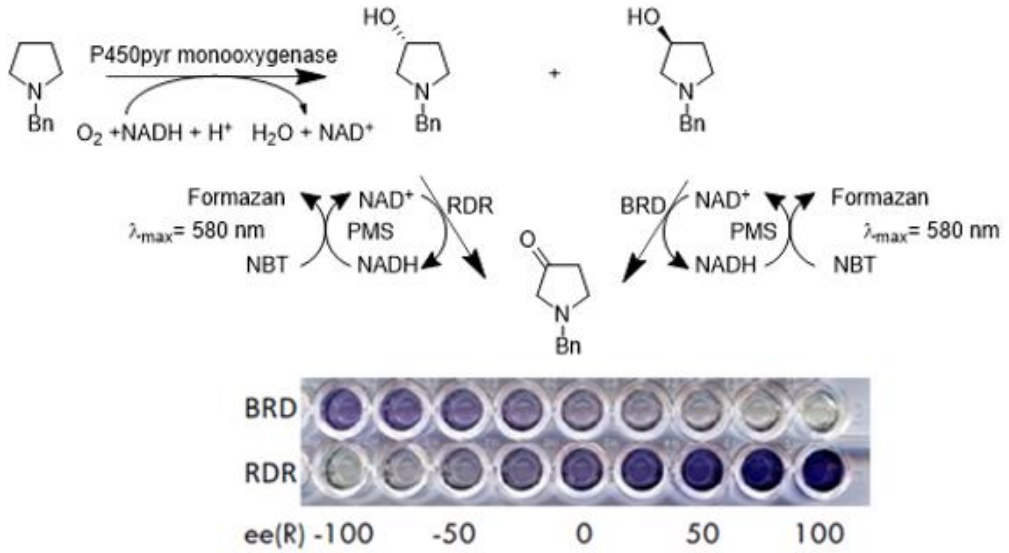


Figure 2.11 Principle of high-throughput screening for the screening of *(S)*- and *(R)*-1-benzyl-3-pyrrolidinol. [226]

Recently, another interesting high-throughput MS screening method for enantioselectivity was reported.[227] Different with enzyme-based assay described above, this method was adopted by using two enantiomers of isotopically labeled substrate for parallel enzymatic hydroxylations, coupled

with MS detection. The principle is shown in Figure 2.12, (*S*)-D- and (*R*)-D-substrate are prepared and used for the hydroxylation, respectively. The O-D bond in the hydroxylation product changes quickly to O-H in the aqueous medium to give a mass of M, and other hydroxylation products still have a mass of M+1. The peak ratio of M and M+1 in MS can be easily determined and then used to calculate the product *ee* based on Equation (1) and (2). By using this method, the enantioselectivity of P450pyr for the biohydroxylation of *N*-benzyl-3-pyrrolidine was improved from *ee* of 43% (*S*) to 98% (*S*) with 80% relative activity as compared to the wild-type P450pyr.[103]

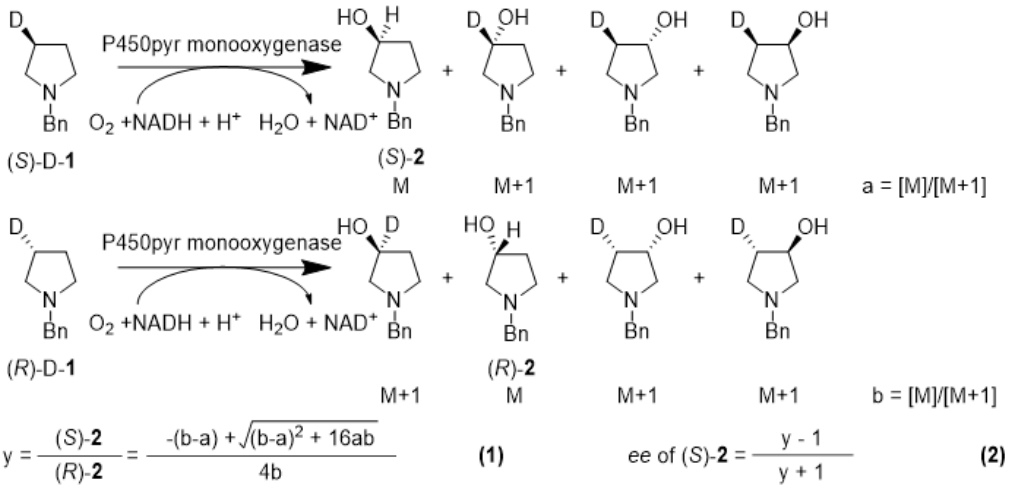


Figure 2.12 Principle of mass-spectrometry based high-throughput screening using deuterated substrates. [103]

2.3.3.2 Screening Methods to Determine the Regioselectivity

The principle of some reported regioselective HTS method is very similar to the two-enzyme based enantioselective HTS assay. For instance, Lentz *et al* developed an HTS procedure that is capable of discriminating between terminal and other position carbon chains hydroxylation products.[228] As shown in Figure 2.13, the yeast alcohol dehydrogenase, which has an oxidation rate that is 10

times higher for primary alcohols than for secondary ones, was used and the process was monitored through NADH formation with a spectrophotometer.[228]

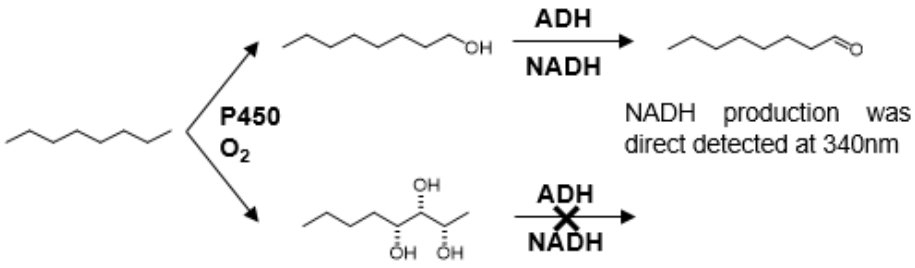


Figure 2.13 Principle of high-throughput screening for the terminal selective hydroxylation of *n*-octane.

Apart from these types of enzyme-based HTS methods, other chemical based direct regioselective colorimetric HTS assays were also reported. For example, the "surrogate" substrate hexyl methyl ether (HME) was used in the evolution of P450BM3 for the terminal alkane hydroxylation.[229] As shown in Figure 2.14, terminal hydroxylation of HME would produce a formaldehyde, which turned purple colour in the presence of Purpald reagent. By using ISM and this screening method, a P450BM3 mutant with 52% terminal selectivity of octane was discovered.[229] Based on the similar mechanism, some other "surrogate" substrates like: nitro-4-(octyloxy)benzene (NOB) and dimethylether (DME) were also used in the engineering of highly regioselective enzyme.[230]

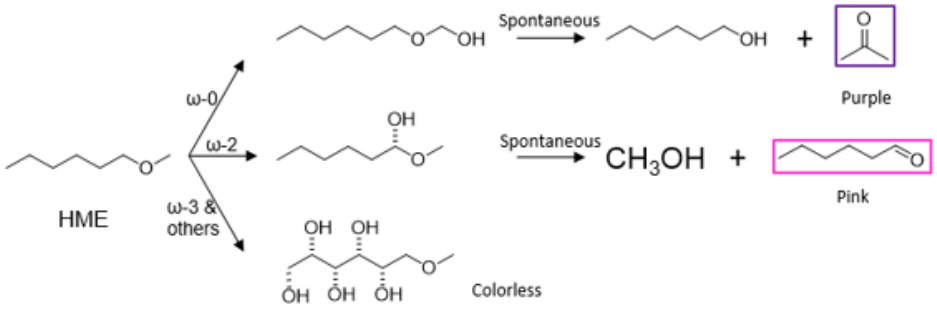


Figure 2.14 Screening for terminal alkane hydroxylation using hexyl methyl ether (HME).

2.3.4 Engineering P450s for New Substrate Specificity and Higher Catalytic Activity and Selectivity

Over the past two decades, several P450 monooxygenases have been sucessfully engineered via directed evolution for the highly ative regio- and/or enantioselective hydroxylation of certain target substrates.

By using a combination of directed evolution and site-directed mutagenesis, Frances H. Arnold and co-workers sucessfully engineered P450BM3 for the alkane hydroxylation. The wild-type P450BM3 shows a very poor chemo-, regio- and enantioselectivities in the oxidising of *n*-octane, the oxidation product is a mixture of 4-octanol, 3-octanol, 2-octanol, 4-octanone, and 3-octanone, besides the activity is also much lower than for the hydroxylation of its natual substrate C_{12} to C_{18} fatty acids (Figure 2.15).[148] By using a surrogate substrate-based colorimetric HTS assay, several variants with more than 5 times improved activity were identified after two rounds of evolutions. After that, a NADPH consumption HTS assay with *n*-octane as the substrate was combined with this colorimetric HTS assay, and two active mutants were generated after several rounds of error-prone PCR, DNA shuffing and site-directed mutagenesis. One variant hydroxylates *n*-octane at the subtreminal 2-position to produce (*S*)-2-octanol (40% *ee*). Another variant, also hydroxylates *n*-octane at the 2-position but forms (*R*)-2-alcohols (40-55% *ee*, Figure 2.15).[106] Besides, these two biocatalysts are highly active, and the maximum turnover rate for the hydroxylation of *n*-octane could reach 400 min^{-1}.

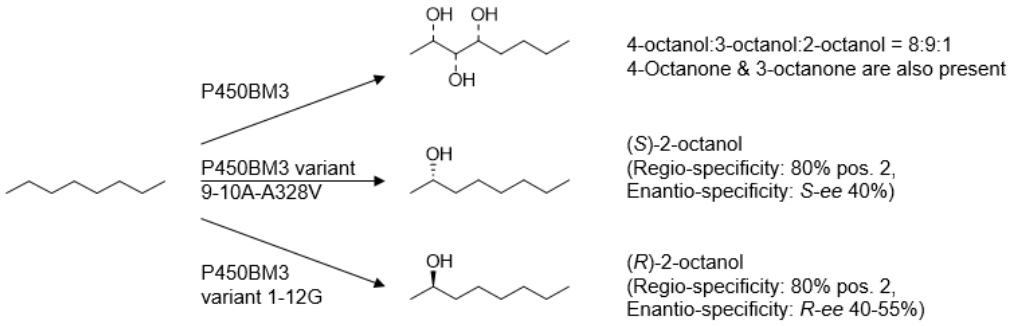
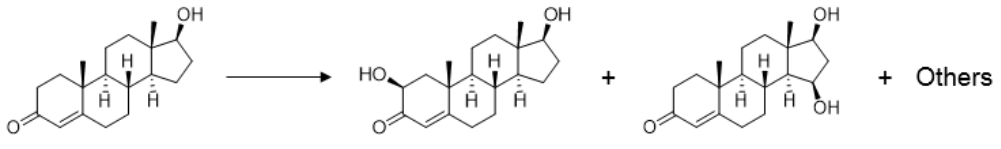


Figure 2.15 Regio- and enantio-selective hydroxylation of *n*-octane with P450BM3 and engineered P450 BM3.

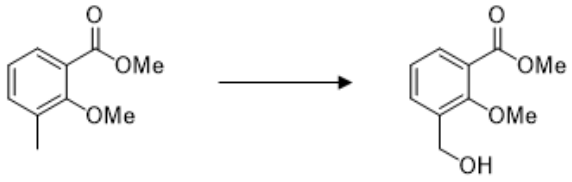
Recently, this P450BM3 has also been successfully engineered by Manfred T. Reetz and co-workers through laboratory evolution for the regio- and stereoselective hydroxylation of steroids. In their inital experiment, the P450BM3(F87A) was discovered to be active for the hydroxylation of testosterone, but resulting in a 1:1 mixture of 2β- and 15β-alcohols. To maximise the probability of cooperative effects within a given site and between sites, selected residues were grouped first and ISM was applied to generate P450BM3 variants. After screening 8,700 mutants with HPLC, variants with 96–97 % selectivity for either 2β- and 15β-isomers were obtained, both of them showed complete diastereoselectivity (Figure 2.16).[231] Moreover, these two mutants also showed increased product formation and NADPH consumption.



P450BM3 (F87A)	52%	45%	3%
P450BM3 (A330W/F87A)	97%	3%	0%
P450BM3 (R47Y/T49F/V78L/A82M/F87A)	3%	96%	1%

Figure 2.16 Regio- and enantio-selective hydroxylation of testosterone with engineered P450 BM3.

To investigate the activity and selectivity of P450BM3 and its variants for the hydroxylation of aromatic compounds, Ulrich Schwaneberg and co-workers generated a designed P450BM3 library by site-directed mutagenesis. Methyl 2-methoxy-3-methylbenzoate was selected as a model substrate and NADPH consumption assay was used to discover potential variant. One mutant, P450BM3(F87A/L188C) showed 535-fold improved catalytic activity towards methyl 2-methoxy-3-methylbenzoate compared to that of wild-type P450BM3.[182]Besides, this variant was also found to be able to selectively hydroxylate several carbonyl- or carboxyl-substitute toluene derivatives. In addition, this variant also showed some activity toward 3-fluorotoluene (but did not show any activity toward chloro- and bromotoluenes).



	oxidation rate [$M^{-1}s^{-1}$]	Improvement
P450BM3	0.1	1
P450BM3 (F87A/L188C)	63.6	535

Figure 2.17 Benzylic hydroxylation of 2-methoxy-3-methylbenzoate with wild-type P450BM3 and engineered P450BM3.

Chapter 3: Engineering of P450pyr Hydroxylase for the Highly Regio- and Enantioselective Subterminal Hydroxylation of Alkanes

3.1 Introduction

Regio- and enantioselective hydroxylation at a non-activated carbon atoms is a very useful reaction for the functionisation of alkanes, an abundant and cheap feedstock for chemical synthesis, and produce enantiopure alcohols that are useful and valuable synthons and pharmaceutical intermediates.[119,189,232-234] Despite of some progress with transition metal catalysts,[235-237] this type of reaction still remains as a significant challenge in classic chemistry. On the other hand, nature has found a general solution for this reaction via monooxygenase using molecular oxygen as the non-toxic oxidant.[238,239] Many monooxygenases such as the soluble cytochrome P450 monooxygenases (P450cam, P450BM-3, P450pyr),[120,240,241] the soluble methane monooxygenases (sMMO),[124,242,243] and the membrane-bound alkane hydroxylase (alkB)[125,244] are known for this type of transformation. However, most monooxygenases such as MMO, alkB, and P450pyr only showed terminal hydroxylation selectivity,[188,242,245,246] while P450BM-3 and P450cyp102A3 showed ω-1, ω-2 and ω-3 selectivity but the regio- and enantioselectivity are very poor.[148,247]

Recently, directed evolution has become a useful tool to create new enzymes with improved catalytic performance.[248-254] Using this approach, many monooxygenases have been engineered with new substrate specificity[183,255-258] and several with enhanced regio- and/or stereoselectivity[179,229,247,259] for the

target hydroxylations. For instance, P450BM-3 mutants were engineered with excellent regio- and diastereoselectivity for the hydroxylation of testosterone [260] and Artemisinin,[261] respectively. Nevertheless, it is a challenge to develop an enzyme for non-terminal hydroxylation of alkane with excellent regio- and enantioselectivity: engineered sMMO and alkB mutants demonstrated some subterminal selectivity;[125,262] engineered P450BM-3 mutants gave enhanced subterminal selectivity, with 89% regioselectivity and 65% enantioselectivity as the best results for the subterminal hydroxylation of *n*-octane;[63,106,263] and engineered P450cam mutants showed subterminal hydroxylation of *n*-butane.[264] Nevertheless, there is no enzyme catalysing non-terminal hydroxylation of alkane with excellent regio- and enantioselectivity.

We are interested in developing a set of new enzymes that are capable of subterminal hydroxylations with good regio- and enantioselectivities. We previously discovered a soluble P450pyr monooxygenase from *Sphingomonas* sp. HXN-200 as a powerful and best biocatalyst known thus far for terminal hydroxylation of *N*-substituted pyrrolidines, piperidines, azetidines, 2-pyrrolidinones, and 2-piperidinones with high activity, high yield, excellent regioselectivity, and good to excellent enantioselectivity. [187,265-269] Recently, we succeeded in engineering P450pyr monooxygenase by directed evolution to create P450pyr mutants for the hydroxylations of *N*-benzylpyrrolidine with higher and inverted enantioselectivity.[102,103] Our P450pyr is much better than P450BM-3 regarding the regio- and enantioselectivity in hydroxylations and shows different substrate specificity. Herein we wanted to engineer this P450pyr by directed evolution to create new P450pyr mutants for regio- and enantioselective subterminal hydroxylation. In this project, monooxygenase-

catalysed subterminal hydroxylation of *n*-octane was selected as the target reaction.

In order to achieve the aim of this project, we first developed a suitable colorimetric high-throughput screening (HTS) assay. We previously developed a HTS method for the determination of the *ee* of a chiral alcohol by the use of two enantioselective alcohol dehydrogenases. This method was further extending to determine both regio- and enantioselectivity at same time by using three regio- and enantioselective alcohol dehydrogenases. After set up the HTS method, protein engineering started with the identification of key amino acid residues that could influence the regio- and enantioselectivity. Based on *n*-octane-P450pyr docking structure model, all residues within 6Å of the bound substrate were identified and subjected to ISM. The P450pyr was mutated at all the selected amino acid positions by using primers with mixed bases for the appropriate codon. The mutant genes were expressed in *E. coli* co-expressing Fdx and fdR. Biohydroxylations of *n*-octane with these colonies were examined based on the new developed HTS assay. Mutants with a higher subterminal selectivity and higher enantioselectivity were then selected as template for the next round of evolution. Besides, the roles of the mutated residues on changing the regioselectivity and improving enantioselectivity were investigated based on the experimental results and simulation modelling, which could give insight into the understanding of the relationships between selectivity and structure of P450pyr and its mutants.

3.2 Experimental Section

3.2.1 Strains, Media, and Materials

T4 DNA quick ligase kit and deoxynucleotide (dNTP) Solution Mix were purchased from New England Biolabs (NEB). iProof DNA Polymerase was bought from Bio-Rad. DNA loading dye and DNA ladders were purchased from Thermo Scientific. Oligonucleotides were synthesised by AIT biotech, Singapore. LB Broth, bacto agar, tryptone, and yeast extract were purchased from Biomed Diagnostics. Antibiotics ampicillin and kanamycin, as well as YAD (Alcohol Dehydrogenase from *Saccharomyces cerevisiae*, lyophilized powder, ≥300 units/mg protein, Protein, ≥90%) were from Sigma-Aldrich.

Table 3.1 Strains and plasmids used in this study

Strain or plasmid	Characteristic	Reference
E. coli DH5α	Cloning strain	Invitrogen
E. coli BL21(DE3)	Expressing strain	Novagen
pETDuet1	Expressing vector	Novagen
pRSFDuet1	Expressing vector	Novagen
pET28a	Expressing vector	Novagen

Following chemicals were purchased from Sigma-Aldrich and used without further purification: *n*-octane (≥ 99%), 1-octanol (≥ 99%), *(S)*-2-octanol (98% *ee*) and *(R)*-2-octanol (98% *ee*), nitro blue tetrazolium (NBT) (≥ 98%), phenazine methosulfate (PMS) (≥ 90%), Isopropyl β-D-thiogalactopyranoside

(ITPG) (≥ 99%), D-Glucose (≥ 99.5%), and NAD^+ (≥ 96.5%). Other required salts and reagents were purchased from Fisher Scientific or Sigma-Aldrich.

3.2.2 Generation of Mutant Library

The mutant library at selected sites was generated using PCR with primers containing NNK codons covering all 20 possible amino acids. Each PCR reaction tube contained 2 μL each of forward and reverse degenerate primer of a particular target site, 10 μL iProof HF Buffer, 2 μL dNTP Solution Mix, 0.8 μL $MgCl_2$ solution, 0.1 μL (10 ng) template plasmid pRSFDuet-P450pyr, 0.5 μL iProof DNA Polymerase, and 36.6 μL ultra-pure H_2O. PCR amplification was carried out on MJ Research Thermal Cycler using the following thermal cycling protocol: 98 ℃ 3 min, (98 ℃ 10 s, 58 ℃ 30 s, 72 ℃ 4 min) ×30 cycles, 72 ℃ 10 min.

Table 3.2 Primers for site-directed mutagenesis and their melting temperatures

Site	Primers	T_m*
A77	A77-F: CTC GTC CGA TNN KGG ATA TGG CG A77-R: CGC CAT ATC CMN NAT CGG ACG AG	54.6 54.6
I82	I82-F: GGA TAT GGC GGC NNK ATA ATC GAT GAC I82-R: GTC ATC GAT TAT MNN GCC GCC ATA TCC	55.4 55.4
I83	I83-F: GGC GGC ATC NNK ATC GAT GAC G I83-R: CGT CAT CGA TMN NGA TGC CGC C	54.4 54.4
I88	I88-F: CGA TGA CGG CNN KCA AAA AGG I88-R: CCT TTT TGM NNG CCG TCA TCG	50.2 50.2
Q89	Q89-F: GAC GGC ATT NNK AAA GGT GGC G Q89-R: CGC CAC CTT TMN NAA TGC CGT C	52.5 52.5
L98	L98-F: GCG GAC TGG ATN NKC CCA ATT TC L98-R: GAA ATT GGG MNN ATC CAG TCC GC	52.8 52.8

P99	P99-F: GGA CTG GAT CTT NNK AAT TTC ATC GCG P99-R: CGC GAT GAA ATT MNN AAG ATC CAG TCC	53.9 53.9
N100	N100-F: GGA TCT TCC CNN KTT CAT CGC N100-R: GCG ATG AAM NNG GGA AGA TCC	50.2 50.2
I102	I102-F: CCC AAT TTC NNK GCG ATG GAT C I102-R: GAT CCA TCG CMN NGA AAT TGG G	50.7 50.7
A103	A103-F: CCA ATT TCA TCN NKA TGG ATC GGC C A103-R: GGC CGA TCC ATM NNG ATG AAA TTG G	53.4 53.4
S182	S182-F: CTT ACC CGC NNK TCG GAT GTG AC S182-R: GTC ACA TCC GAM NNG CGG GTA AG	54.6 54.6
D183	D183-F: GCT GGT CGN NKG TGA CAA CC D183-R: GGT TGT CAC MNN CGA CCA GC	51.8 51.8
T185	T185-F: GGA TGT GNN KAC CGC AGC T185-R: GCT GCG GTM NNC ACA TCC	48.6 48.6
T186	T186-F: GAT GTG ACA NNK GCA GCA CC T186-R: GGT GCT GCM NNT GTC ACA TC	49.7 49.7
L251	L251-F: GTA CTT NNK CTG ATC GTT GGC G L251-R: CGC CAA CGA TCA GMN NAA GTA C	50.7 50.7
V254	V254-F: CCT GAT CNN KGG CGG GAA C V254-R: GTT CCC GCC MNN GAT CAG G	51.4 51.4
G255	G255-F: CCT GAT CGT TNN KGG GAA CG G255-R: CGT TCC CMN NAA CGA TCA GG	49.7 49.7
D258	D258-F: CGG GAA CNN KAC CAC ACG D258-R: CGT GTG GTM NNG TTC CCG	48.6 48.6
T259	T259-F: CGG GAA CGA TNN KAC ACG C T259-R: GCG TGT MNN ATC GTT CCC G	49.2 49.2
L302	L302-F: GCA AAC ACC GNN KGC TCA TAT GC L302-R: GCA TAT GAG CMN NCG GTG TTT GC	52.8 52.8
M305	M305-F: CTT GCT CAT NNK CGC CGC ACG M305-R: CGT GCG GCG MNN ATG AGC AAG	54.1 54.1
F403	F403-F: CGT TCA AAT NNK GTG CGC GG F403-R: CCG CGC ACM NNA TTT GAA CG	49.7 49.7

* values were provided by AITbiotech, Singapore

PCR products were subjected to DNA gel electrophoresis and the location of DNA in gel was detected using gel imaging system. The band containing the product was excised from the gel, purified using QIAquick Gel Extraction Kit (Qiagen), and the purified DNA was digested with methylated restrictive enzyme *Dpn I* at 37℃ overnight. Finally, the mutant library creation was completed by transferring the digested DNA into *E. coli* BL21*(DE3)* competent cells containing the pETDuet-Fdx-FdR vector.

3.2.3 General Procedure for Surrogate Substrate-based Colorimetric HTS Assay

Mutants were inoculated and grown at 37 ºC in 96-deep well plate with 600 μL TB medium per well. The expressing cultures were inoculated by transferring 100 μL overnight culture into 900 μL expressing medium containing 50 μg/mL kanamycin, 100 μg/mL ampicillin, 500 μM IPTG, and 500 μM ALA. After expressing at 22 ℃ for 8 h, the cells were harvested by centrifugation at 3220 g and 4 ℃ for 15min. The cell pellets were resuspended in potassium phosphate buffer (100 mM, pH 8.0) containing 5 mM surrogate substrate 4-nitrophenetole and 2 % (w/v) glucose.

Biotransformation was performed for 4 hours at 30 ℃ and 800 rpm. The 96-deep well plates were covered with gas-permeable seals during the whole cell growth and biotransformation process. The reaction was stopped by centrifugation at 3220 g for 15 min at room temperature. 100 μL of the supernatant from each well was transferred to a new 96-well plate and the absorbance at 410 nm was measured using Tecan Infinite M200 Microplate Reader. The principle and procedure is given in Figure 3.1.

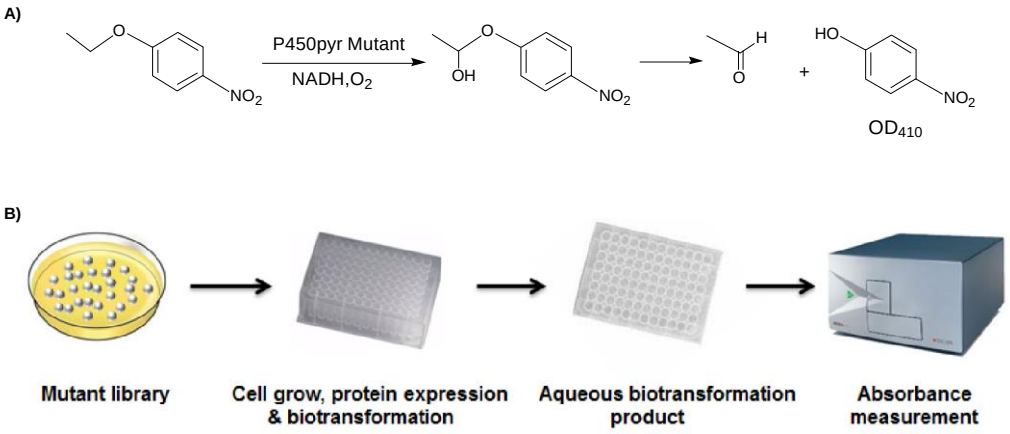


Figure 3.1 The principle and procedure of surrogate substrate-based colorimetric HTS assay. (a) The principle: Biohydroxylation of 4-nitrophenetole at subterminal position generated an unstable hemi-acetal intermediate that decomposed to 4-nitropheol. The formation of 4-nitrophenol can be determined by using a microplate reader based on UV absorption at 410 nm. (b) The procedure of the surrogate-substrate based colorimetric HTS assay.

3.2.4 Genetic Engineering of *E. coli* Expressing CpSADH and *E. coli* Expressing PfODH

The CpSADH and PfODH were cloned into pET28a plasmid (Figure 3.2). To construct these two plasmids, four primers were designed as listed in Table 3.3, both of the CpSADH and PfODH genes were amplified by PCR using MJ Research PTC-200 thermal cycler with the following program: 95 ℃ 5 min, (95 ℃ 30 s, 55 ℃ 30 s, 72 ℃ 1min) × 30 cycles, 72 ℃ 10 min.

The PCR amplified CpSADH and PfODH fragments as well as expressing vector pET28a were then digested with their respective enzymes. The digestion products were purified again over 1% agarose gel in Tris-acetate-EDTA buffer. The ligation was conducted using T4 quick ligase kit. The ligation solution was then transformed into chemical competent *E. coli* DH5α cells, which were then plated on LB agar plates with appropriate antibiotics (50 μg/mL kanamycin and 100 μg/mL ampicilin). The constructed plasmids were confirmed by DNA

sequencing. Finally, the purified plasmids were transformed into chemical competent *E. coli* BL21*(DE3)* cells and plated on LB agar plates containing the appropriate antibiotics.

Table 3.3 Construction of pET28a-CpSADH and pET28a-PfODH plasmids.

Gene to be amplified	Primers (Restriction sites underlined)	Cloned into
CpSADH	CpSADH-F (NdeI): GGGCA TATGT CAATC CCGTC CAGCC CpSADH-R (XhoI): GGGCT CGAGT TACGG GTTAA AGACG AC	pET28a
PfODH	PfODH-F (NdeI): GGGCA TATGA GCTAC AACTT CCATA ATA PfODH-R (XhoI): GGGCT CGAGTT ATTGA GCCGT GTAAC	pET28a

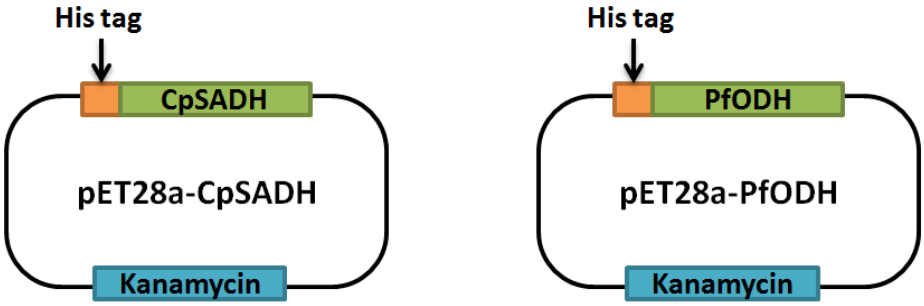


Figure 3.2 pET28a-PfODH and pET28a-CpSADH expression vector

3.2.5 Expressing and Purification of CpSADH and PfODH

E. coli BL21*(DE3)* pET28a-CpSADH and *E. coli* BL21*(DE3)* pET28a-PfODH were cultured in LB medium containing 50 μg/mL kanamycin at 37 °C, respectively. At OD_{600} of 0.6-0.8, IPTG was added to a final concentration of 500 μM and the temperature was reduced to 30 °C to induce the expression of his-tagged CpSADH and PfODH. After 12 h of expression, the cells were harvested by centrifugation at 3220 g and 4 °C for 15 min. The cell pellet was then re-suspended in binding buffer (50mM potassium phosphate buffer, 0.3 M

NaCl, 10 mM imidazole, pH 8.0) at OD_{600} of 30. Cells were broken by passing through cell disruptor at 21 psi for 2 times, followed by centrifugation (20376 g, 4 ^{o}C, and 60 min). The his-tagged PfODH was expressed, harvested and broken under same condition.

The his-tagged CpSADH and PfODH were purified by using ÄKTA purifier (GE Healthcare) system at 4 ^{o}C with UV at 280 nm to monitor protein concentration. The cell free extracts were loaded onto a Ni-NTA column, and non-his-tagged proteins were washed out by using 50 mM potassium phosphate buffer (pH 8.0) containing 0.3 M NaCl and 50 mM imidazole. The his-tagged protein was eluted by using 50 mM potassium phosphate buffer (pH 8.0) containing 0.3M NaCl and 500 mM imidazole. The fractions containing his-tagged protein (CpADH or PfODH) were concentrated using Millipore Amicon Ultra-4 centrifugal filter device (10 kDa) at 3220 g and 4 ℃, washed three times with 50 mM potassium phosphate buffer (pH 8.0).

3.2.6 Colorimetric HTS Assay Validation for the Determination of Regio- and Enantioselectivity in the Hydroxylation of *n*-octane

The developed HTS assay was used to determine the terminal alcohol and the *ee* of subterminal alcohol by using samples containing different amount of *(S)*-2-octanol, *(R)*-2-octanol and 1-octanol. In the first group of experiments, 81 samples containing 1-octanol, (*S)*-2-octanol, and *(R)*-2-octanol at different ratio with the total concentration of 1.0 mM were used. In the second group, 15 samples with different total concentration and ratio (1-octanol was at 0.05 mM, 0.5 mM, or 5 mM, and (*S*)- and (*R*)-2-octanol was at totally 0.5 mM but with different ratio) were examined. All samples were put in three microtiter plates,

and reacted with YAD, CpSADH, and PfODH separately, in the presence of 2 mg/mL NBT, 0.1 mg/mL PMS, and 1.0 mM NAD^+. UV absorption at 580nm was recorded using a microtiter plate reader.

3.2.7 General Procedure for Using the Colorimetric HTS assay in the Evolution of P450pyr Hydroxylase

Mutants were grown and expressed under the same condition as given in section 3.2.3. After harvesting, cells were re-suspended in potassium phosphate buffer (100 mM, pH 8.0) containing 5 mM *n*-octane, 2 % (w/v) glucose, and 2 (v/v) % DMSO. Biotransformation was performed in 96-deep well plate at 30 ^oC and 800 rpm for 4 h.

After reaction, the deep well plates were centrifuged, and 80 μL aqueous biotransformation mixtures were added into three 96-well plates, respectively. 10 μL NBT-PMS solution (containing 2 mg/mL NBT, 0.1 mg/mL PMS and 1.0 mM NAD^+, final concentration) and 10 μL YAD, CpSADH or PfODH solution (0.5 U/mL) were added to each well of the microtiter plates, respectively. After incubation in darkness for 1 h at room temperature, OD_{580} was measured by using a microtiter plate reader. Mutants with improved subterminal selectivity and enantioselectivity were identified by comparing the two ratios of the corresponding absorbencies on three plates with the ratios for the parent enzyme put on the same plates (Figure 3.3).

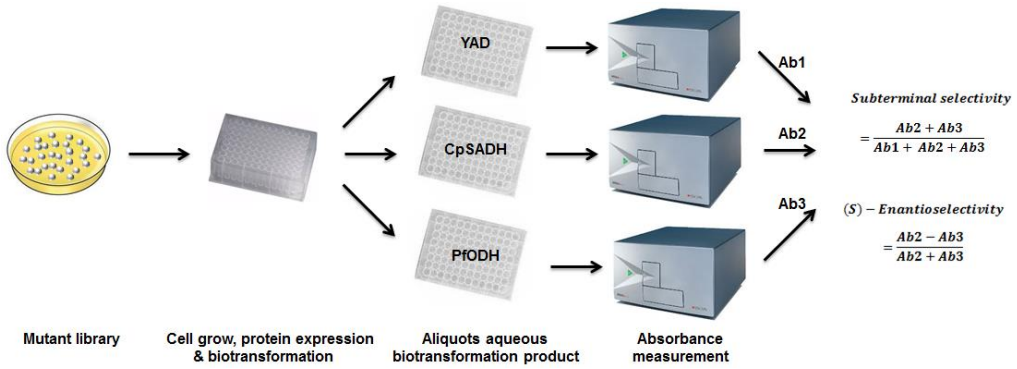


Figure 3.3 The process of the new developed regio- and enantioselective colorimetric based colorimetric HTS assay.

3.2.8 Whole-cell Biotransformation in Shaking Flask

Positive mutants were selected and re-cultured on LB agar plate containing 50 μg/mL kanamycin and 100 μg/mL ampicillin for overnight at 37 ℃. The colonies were picked and inoculated in 3 mL LB medium and then shaken at 250 rpm and 37 ℃ for 4 hours. The cells were then transferred in a 125 mL flask containing 25 mL TB medium with appropriate antibiotics. The culture was shaken at 250 rpm and 37 ℃ until the OD_{600} research 0.8. The cells were then induced for protein expression under 500 μM IPTG and 500 μM δ-ALA and the culture was shaken for additional 6 h at 250 rpm and 22 ℃.

Cell pellets were obtained by centrifugation at 3220 g for 15 min at 4 ℃, washed with 30 mL distilled water, and centrifuged again at 3220 g for 15 min. The cells were re-suspended to 8 g cdw L^{-1} in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 5 mM *n*-octane, 2% (w/v) glucose, and 2 % (v/v) DMSO. The biotransformation was performed in the shaking flask at 250 rpm and 30 ℃ for 4 h.

3.2.9 Analytic Method

The activity and subterminal selectivity of the positive mutants for the hydroxylation of *n*-octane were determined by using Agilent 7890A gas chromatograph with HP-5 column (30 m × 0.32 mm × 0.25 mm). Temperature program: 80 °C for 1 min, then to 180 °C at 20 °C min^{-1}, and finally to 230 °C at 50 °C min^{-1} for 1 min. Retention times: 3.87 min for 1-octanol and 3.32 min for 2-octanol.

The *ee* value of 2-octanol was determined by using Agilent 7890A gas chromatograph with Macherey-Nagel Hydrodex-β-TBDAc column chiral column (25 m × 0.25 mm) at 70 °C. Retention times: 73.3 min for *(R)*-2-octanol and 74.8 min for *(S)*-2-octanol.

The activity and subterminal selectivity of positive mutants for the hydroxylation of propylbenzene were determined by using Agilent 7890A gas chromatograph with HP-5 column (30 m × 0.32 mm × 0.25 mm). Temperature program: initial temperature 100 °C, then to 280 °C at 20 °C min^{-1}, and finally hold at 280 °C for 2 min. Retention times: 3.29 min for 1-phenyl-2-propanol and 3.85 min for 3-phenyl-1-propanol.

The *ee* value of 1-phenyl-2-propanol was determined by using Agilent 7890A gas chromatograph with Macherey-Nagel Hydrodex-β-TBDAc column chiral column (25 m × 0.25 mm) at 100 °C. Retention times: 40.1 min for *(S)*-1-phenyl-2-propanol and 42.6 min for *(R)*-1-phenyl-2-propanol.

3.3 Results and Discussion

3.3.1 Identification of Suitable Residues of P450pyr for ISM

To identify the suitable residues for iterative saturation mutagenesis (ISM), *n*-octane was docked onto the 3D structure of P450pyr (PDB ID 3RWL, Figure. 3.4).[103] The P450 enzyme was set to the ferry-oxo intermediate known as Cpd I state to mimic the actual binding geometry. The active pose of the substrate was concluded from the reported catalytic geometry.

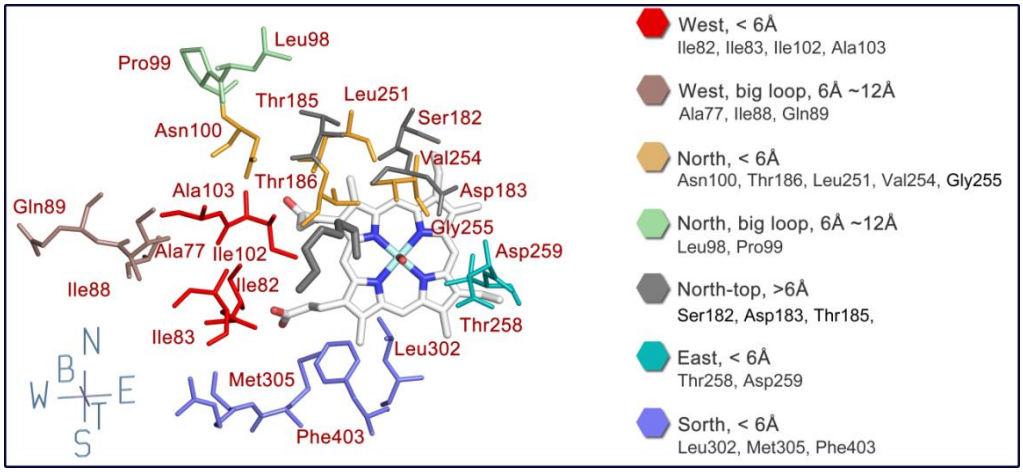


Figure 3.4 Spatial overview of the target sites for ISM based on the active docking pose of *n*-octane (grey stick) in P450pyr hydroxylase. The heme was represented as ferryl-oxo complex known as CpdI state. 22 amino acid residues from seven major geometric areas were chosen for the mutagenesis.

Based on the docking results, 14 residues located within 6 Å of the substrate molecule were selected and grouped in 4 parts, west (Ile82, Ile83, Ile102 and Ala103), north (Asn100, Thr186, Leu251, Val254 and Gly255), east (Asp258 and Thr259) and south (Leu302, Met305 and Phe403). Three north-top residues (Ser182, Asp183 and Thr185) were also chosen due to their abrupt position in the substrate-accessing channel. In the previous study, a pocket reshaping was observed with the inward movement of the 'big loop' (Ser74-Asp105).This was

due to the increased rigidity contributed by the mutations that resulted in the big improvement in the enantioselectivity. As such, two groups of residues within 12 Å from the docked *n*-octane, west big-loop (A77, I88 and Q89) and north big-loop (L98 and P99), were appended to the above-mentioned target sites.

3.3.2 Creation of P450pyr Mutant Library for Directed Evolution

For saturation mutagenesis at each targeted amino acid site, random mutations at the desired position of the P450pyr gene were generated by PCR using NNK degenerated codon. PCR products were then transformed into *E. coli* BL21(DE3) containing ferredoxin and ferredoxin reductase genes to create P450pyr mutant library. Figure 3.5 showed a typical result of DNA gel electrophoresis seen under a gel imaging system. The weak bands in the centre showed some presence of primer dimers, but the amplified DNA fragments were sufficient to create a mutant library.

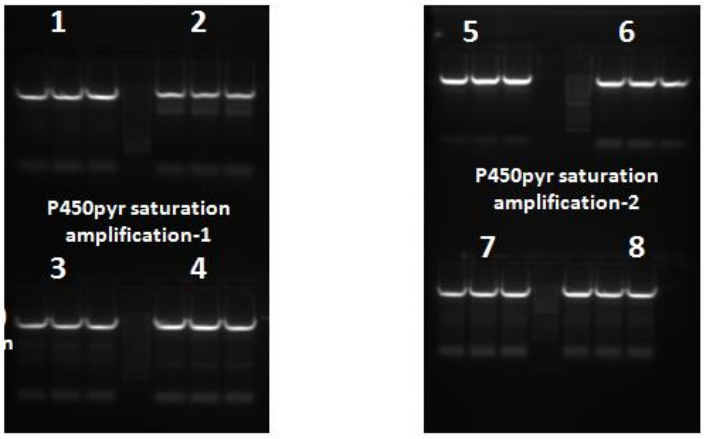


Figure 3.5 Mutant library generation with saturation mutagenesis. Imaging of gel electrophoresis following PCR amplification.

3.3.3 P450pyr Mutant Library Quality Control

In order to check the coverage and quality of the created mutant library, 12 clones were randomly selected from site I82 mutant library. The cells were grown in culture tube with 5 mL LB medium at 37 ℃ for 8 h. The cell pellet was collected by centrifugation (2 min, 20376 g) and the plasmid was extract using Miniprep Kit (Qiagen). All the purified plasmid samples were sequenced by AITBiotech. As shown in Figure 3.6, there were 10 different types of amino acids within 12 randomly selected mutants. This confirms that the quality of the mutant library.

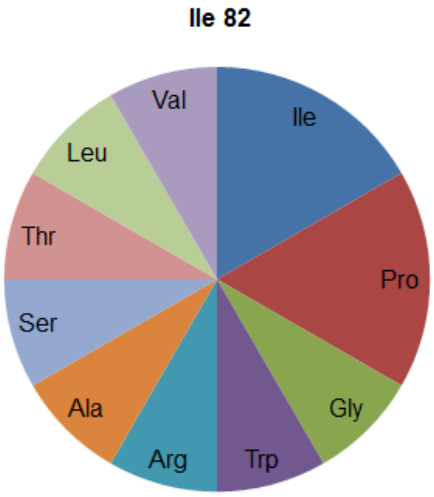


Figure 3.6 The quality of mutant library I82 created with NNK degeneracy. The original amino acid was given on the top of the circle. Ten different created amino acids were represented with different colours.

3.3.4 Development of Surrogate Substrate-based Colorimetric HTS Assay

The first HTS assay for the subterminal hydroxylation screening was developed based on the mechanism described in Figure 3.1. This method allows screening of hundreds of mutants simultaneously within a very short time, which enable

one to analyse a mutant library within a reasonably time span. During reaction, P450pyr catalyses the subterminal hydroxylation of the surrogate substrate, 4-nitrophenetole, into an unstable hemi-acetal intermediate, which in turn decomposes to a yellow compound, 4-nitrophenol, which could be monitored by taking absorbance reading at OD_{410}. Improved mutants can be easily identified by comparing the absorbance value to that of the parent mutant. In order to find a suitable substrate concentration for the colorimetric HTS, samples containing different concentrations of both substrate and product were prepared and their absorbance at 410nm were detected. As shown in Table 3.4, the product absorbance at 410 nm was highly depending on solute concentration while the substrate was almost colourless at low concentration. The 100 μM was found to be the best substrate concentration for the colorimetric HTS assay.

Table 3.4 The absorbance spectra for different concentrations of 4-nitrophenetole and 4-nitrophenol were measured at 410 nm.

410nM		**Blank**	**10μM**	**100μM**	**1mM**	**10mM**
4-nitrophenetole (Substrate)	Sample1	0.043	0.044	0.045	0.053	0.078
	Sample2	0.044	0.043	0.045	0.052	0.078
4-nitrophenol (Product)	Sample1	0.044	0.750	3.626	3.692	3.892
	Sample2	0.043	0.761	3.662	3.685	3.894

3.3.5 Directed Evolution with Surrogate Substrate-based Colorimetric HTS Assay

The mutants' subterminal hydroxylation activities were correlated to the amount of yellow 4-nitrophenolate produced, hence potential mutants could be identified by comparing the ratio of absorbance compared to wild-type P450pyr. A typical 96-well plate is shown below in Figure 3.7. Potential mutants with

subterminal hydroxylation activity could produce yellow even visible to the naked eye. In the first round of ISM, using wild-type P450pyr as a template, we screened 188 mutants per site, which translates into a total of 4136 mutants (22 sites) screened. Finally, 6 potential mutants were identified from the first round of screening.

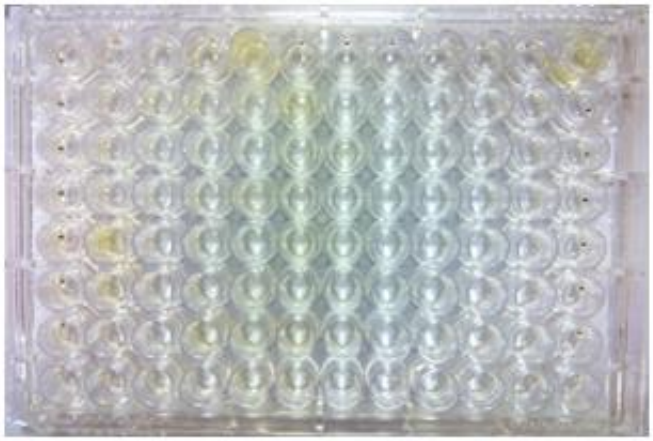


Figure 3.7 A 96-well plate containing a mutant library for colorimetric HTS. Mutants with subterminal hydroxylation activity produce yellow 4-nitrophenolate visible to naked eye.

All of the identified potential mutants were respectively inoculated to undergo further experiment to confirm their activity and regioselectivity of target substrate, *n*-octane. Within these six mutants, one mutant N100S displayed the best subterminal selectivity (5%) and activity. It was very interesting that only one amino acid changed at position 100 (replacing asparagine with serine) could create a subterminal hydroxylation selectivity.

To further improve the subterminal regioselectivity and activity, a second round of evolution was applied by using N100S mutant as template. In this round, all left 21 sites were subjected to the ISM, and finally 20 potential mutants were selected after screening a 3948-mutant library. Mutant N100S/F403I was found to give the highest subterminal regioselectivity (33%) with significantly

improved relative activity (from 49 % to 110 %). This mutant was then used as template for the third round evolution, a triple mutant, N100S/F403I/T186I, was found to be the best mutant in this round, which gave 40% regioselectivity but the activity decreased to 44 % of the activity of wild-type P450pyr for terminal hydroxylation.

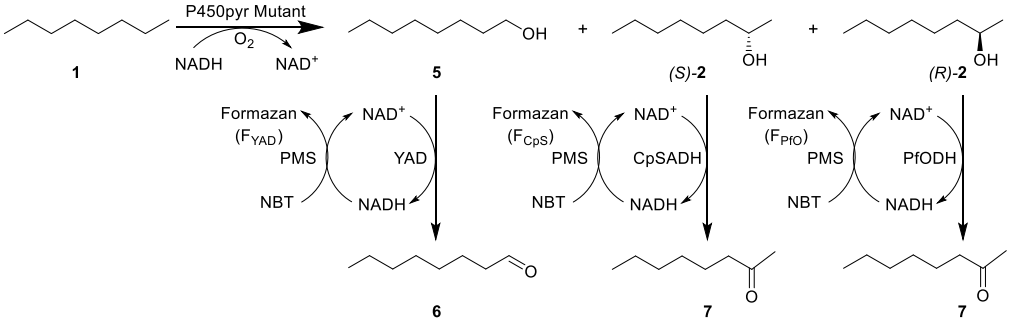
3.3.6 Principle of Regio- and Enantioselective Colorimetric HTS Assay

Limited improvement in past three rounds of mutagenesis suggests potential issues with the surrogate substrate based HTS method. Firstly, the surrogate substrate is structurally dissimilar to the target substrate, and therefore the activity and selectivity on the surrogate substrate is not equivalent to the activity and selectivity on the target substrate. We might have missed many mutants that have high subterminal regioselectivity toward *n*-octane but low or no activity of 4-nitrophenetole. Secondly, this project requires the enzyme to have high regio- and enantioselectivity. While the current surrogate substrate-based colorimetric HTS assay does provide some regioselectivity information, it offers no information on enantioselectivity. Lastly, the absorbance measured in this assay is the product of activity and regioselectivity. In other words, a mutant with higher activity but lower regioselectivity can give a higher absorbance than another mutant with lower activity but higher regioselectivity. In order to solve all these limitations, a novel real substrate-based regio- and enantioselective colorimetric HTS assay was needed.

The development of a real-substrate based regio- and enantioselective colorimetric HTS method was not an easy task. Most of the existing HTS

methods found in literature were mainly focused on improving the enzyme enantioselectivity in the kinetic resolution. Previously several novel, accurate, sensitive and simple HTS methods were established in our lab, and recently an enantioselective colorimetric HTS was developed based on two enantioselective alcohol dehydrogenases. Based on all these achievements, we decided to develop a novel colorimetric HTS assay to determine both regio- and enantioselectivity at the same time for the subterminal hydroxylation of *n*-octane.

Figure 3.8 illustrated the principle of the regio- and enantioselective HTS assay. Three NAD^+-dependent alcohol dehydrogenases, that are highly specific for the oxidation of 1-octanol, *(S)*-2-octanol, and *(R)*-2-octanol, respectively, were used to separately oxidize the product mixture from a P450pyr variant-catalysed hydroxylation of *n*-octane in three parallel experiments. As shown in Figure 3.8, in the presence of NAD^+, NBT, and PMS, the concentrations of 1-octanol, *(S)*-2-octanol, and *(R)*-2-octanol can be obtained based on the UV absorption of Formazan at 580 nm. The subterminal selectivity and the enantioselectivity can be calculated by using equation (1) and (2), respectively. To prove the concept, a number of alcohol dehydrogenases were screened, and CpSADH from *Candida parapsilosis*,[270] PfODH from *Pichia finlandica*,[271] and YAD from *Saccharomyces cerevisiae* [228] were found to be highly specific for the oxidation of *(S)*-2-octanol, *(R)*-2-octanol and 1-octanol, respectively.



$$\text{Subterminal selectivity} = \frac{[(S)\text{-}2] + [(R)\text{-}2]}{[5] + [(S)\text{-}2] + [(R)\text{-}2]} \times 100\% = \frac{OD_{580}(F_{CpS}) + OD_{580}(F_{PfO})}{OD_{580}(F_{YAD}) + OD_{580}(F_{CpS}) + OD_{580}(F_{PfO})} \times 100\% \quad (1)$$

$$\text{Enantioselectivity}(S) = \frac{[(S)\text{-}2] - [(R)\text{-}2]}{[(S)\text{-}2] + [(R)\text{-}2]} \times 100\% = \frac{OD_{580}(F_{CpS}) - OD_{580}(F_{PfO})}{OD_{580}(F_{CpS}) + OD_{580}(F_{PfO})} \times 100\% \quad (2)$$

Figure 3.8 Principle of the colorimetric HTS assay to measure both subterminal selectivity and enantioselectivity of P450pyr variant for the hydroxylation of *n*-octane. The concentrations of the possible hydroxylation products 1-octanol, *(S)*-2-octanol, and *(R)*-2-octanol are determined based on UV absorption of Formazan at 580 nm in the oxidation of the biohydroxylation mixture with highly selective alcohol dehydrogenase YAD, CpADH, and PfODH, respectively. The selectivities of the P450pyr variant are calculated using Eq (1) and (2).

3.3.7 Genetic Engineering, Expressing and Purification of CpSADH and PfODH

The CpSADH and PfODH genes were synthesised and codon optimised by Genscript Corp. (Piscataway, NJ). The sequences of the codon optimised CpSADH and PfODH are listed in Figure 3.9 and Figure 3.10, respectively. Both CpSADH and PfODH genes were cloned into the *NdeI* site and *XhoI* site of pET28a vector, respectively. Since both the CpSADH and PfODH genes had been codon optimised, after transformation of these two constructed plasmids into *E. coli* BL21(DE3) respectively, both CpSADH and PfODH genes were highly expressed with IPTG as inducer.

```
   1  ATGTCAATCC CGTCCAGCCA GTATGGTTTC GTTTTTAATA AGCAAAGCGG TCTGAATCTG
  61  CGTAATGATC TGCCGGTCCA CAAGCCGAAG GCCGGCCAGC TGCTGCTGAA AGTGGATGCA
 121  GTTGGTCTGT GTCATTCTGA CCTGCACGTT ATTTATGAAG GCCTGGATTG CGGTGACAAC
 181  TACGTCATGG GCCATGAAAT TGCGGGCACC GTTGCCGCGG TGGGTGATGA CGTGATCAAC
 241  TATAAAGTTG GTGATCGTGT TGCATGTGTC GGCCCGAATG GTTGCGGCGG TTGTAAATAC
 301  TGCCGCGGCG CTATCGATAA CGTGTGCAAA AATGCGTTTG GTGATTGGTT CGGCCTGGGT
 361  TATGACGGCG GTTATCAGCA ATACCTGCTG GTTACCCGTC CGCGCAACCT GAGCCGTATT
 421  CCGGATAATG TGTCTGCTGA CGTTGCAGCT GCGAGTACCG ATGCGGTGCT GACGCCGTAC
 481  CACGCCATCA AAATGGCACA GGTTTCACCG ACCTCGAACA TTCTGCTGAT TGGTGCCGGC
 541  GGTCTGGGCG GTAATGCAAT TCAAGTGGCC AAGGCATTTG GTGCCAAAGT CACGGTGCTG
 601  GATAAAAAGA AAGAAGCTCG CGACCAGGCG AAGAAACTGG GCGCTGATGC GGTTTATGAA
 661  ACCCTGCCGG AAAGCATTTC TCCGGGTAGT TTTTCCGCCT GTTTTGATTT CGTTTCAGTC
 721  CAGGCAACGT TCGACGTCTG CCAAAAGTAC GTGGAACCGA AAGGCGTCAT CATGCCGGTG
 781  GGTCTGGGTG CTCCGAACCT GTCGTTTAAT CTGGGTGATC TGGCGCTGCG TGAAATTCGC
 841  ATCCTGGGCA GCTTCTGGGG CACCACGAAT GACCTGGATG ACGTTCTGAA ACTGGTCTCC
 901  GAAGGCAAGG TGAAACCGGT GGTTCGTTCA GCGAAGCTGA AAGAACTGCC GGAATACATT
 961  GAAAAACTGC GTAACAACGC CTATGAAGGT CGTGTCGTCT TTAACCCGTA A
```

Figure 3.9 Codon optimized sequence of the CpSADH gene.

```
   1  ATGAGCTACA ACTTCCATAA TAAGGTGGCT GTCGTTACGG GTGCGCTGTC GGGCATCGGT
  61  CTGTCAGTGG CAAAGAAATT CCTGCAACTG GGCGCCAAAG TGACCATTAG TGATGTTTCC
 121  GGTGAAAAGA AATATCATGA AACCGTGGTT GCCCTGAAAG CACAGAACCT GAATACGGAC
 181  AACCTGCACT ACGTGCAAGC TGATAGCTCT AAAGAAGAAG ACAACAAGAA GCTGATCTCT
 241  GAAACCCTGG CCACGTTTGG CGGTCTGGAT ATTGTTTGCG CGAACGCCGG CATCGGTAAA
 301  TTTGCACCGA CCCATGAAAC GCCGTTCGAT GTTTGGAAAA AGGTTATCGC CGTCAACCTG
 361  AATGGCGTCT TTCTGCTGGA CAAACTGGCA ATTAACTATT GGCTGGAAAA AAGCAAGCCG
 421  GGCGTCATCG TGAATATGGG TTCAGTCCAT AGCTTTGTGG CCGCACCGGG CCTGGCGCAC
 481  TACGGTGCAG CTAAAGGCGG TGTGAAGCTG CTGACCCAGA CGCTGGCTCT GGAATATGCG
 541  AGCCACGGCA TTCGTGTGAA CTCTGTTAAT CCGGGTTACA TTAGTACCCC GCTGATCGAT
 601  GAAGTCCCGA AAGAACGCCT GGACAAGCTG GTGTCCCTGC ACCCGATTGG CCGTCTGGGT
 661  CGCCCGGAAG AAGTTGCAGA TGCTGTCGCG TTCCTGTGTA GCCAAGAAGC GACCTTTATC
 721  AATGGCGTTT CCCTGCCGGT GGATGGCGGT TACACGGCTC AATAA
```

Figure 3.10 Codon optimized sequence of the PfODH gene.

The CpSADH and PfODH enzymes were purified separately by FPLC with a Ni-NTA column. His-tagged CpSADH as well as his-tagged PfODH were obtained in high purity checked by SDS-PAGE (Figure 3.11).

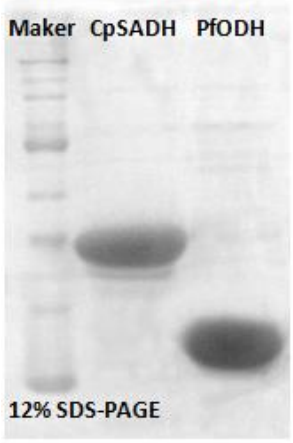
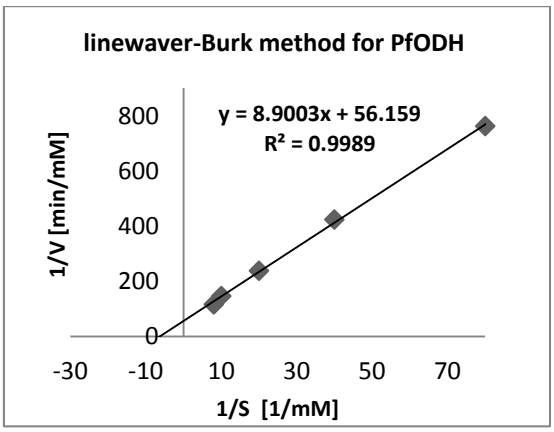


Figure 3.11 SDS-PAGE of purified N-terminal histag CpSADH and PfODH.

3.3.8 Kinetic Study of YAD, CpSADH and PfODH

The kinetic parameters of his-tagged YAD, CpSADH and PfODH for the oxidation of 1-octanol, *(S)*-2-octanol, and *(R)*-2-octanol, were investigated respectively by using a NBT-PMS assay in the presence of NAD^+. Reactions were performed with substrate concentration from 0.125 mM to 1.25 mM in the present of 1 μM of YAD, CpSADH and PfODH, respectively, and the K_m and V_{max} values of each enzyme were obtained from the Lineweaver-Burk plot (Figure 3.12 and Table 3.5).



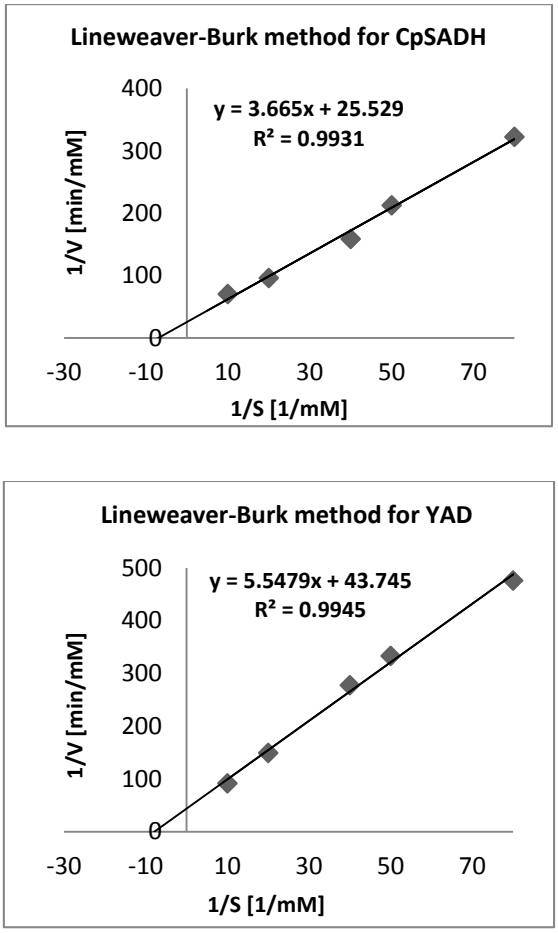


Figure 3.12 Lineweaver-Burk curves of PfODH-catalysed oxidation of (*R*)-2-octanol, CpSADH-catalysed oxidation of (*S*)-2-octanol, and YAD-catalysed oxidation of 1-octanol.

Table 3.5 Kinetic data for oxidation of 1-octanol, *(R)* and *(S)*-2-octanol catalysed by YAD, PfODH and CpSADH, respectively.

Substrate	**Enzyme**	**K_m [µM]**	**V_{max} [µM/min]**	**K_{cat} [1/min]**
1-octanol	YAD	127	23	23
(S)-2-octanol	CpSADH	143	39	39
(R)-2-octanol	PfODH	158	18	18

3.3.9 Validation of the Regio- and Enantioselective colorimetric HTS assay

In the literatures, the authors mentioned that CpSADH and PfODH were highly specific to the *(S)*- and *(R)*-2-octanol and displayed no activity towards 1-

octanol. To confirm this phenomenon and to evaluate the accuracy and sensitivity of our assay under complex conditions, we prepared several samples containing different mixtures of *(S)*-/*(R)*-2-octanol and 1-octanol. We assayed each sample with purified CpSADH and PfODH separately, coupled with the NBT-PMS assay.

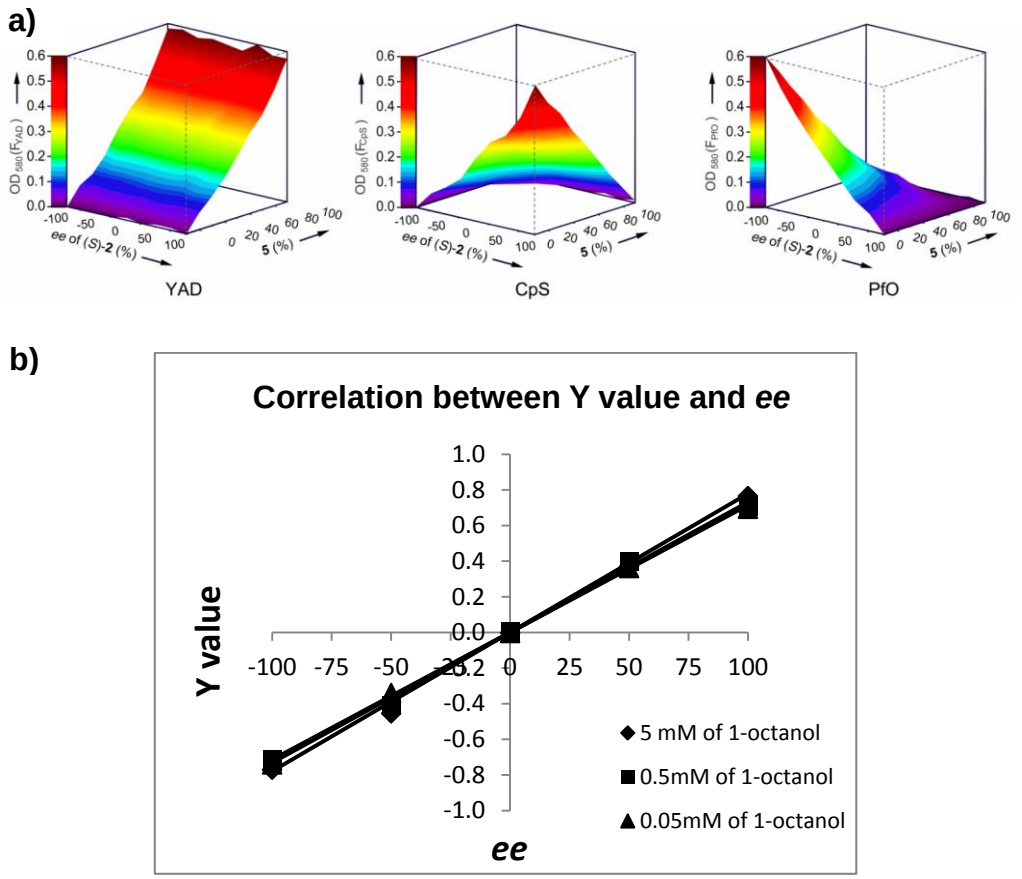


Figure 3.13 Validation of the colorimetric HTS assay. (a) OD_{580} values measured for the oxidation of the first group samples with YAD, CpADH, and PfODH, respectively. (b) Linear correlation between determined *ee* values (Y) and real *ee* values of the second group samples. $Y = \frac{OD_{580}(F_{CpS}) - OD_{580}(F_{PfO})}{OD_{580}(F_{CpS}) + OD_{580}(F_{PfO})}$

As shown in Figure 3.13(a), the UV absorption was proportional to the concentration of 1-octanol, (*S*)-2-octanol, and (*R*)-2-octanol, respectively. The determined concentration of 2-octanol and *ee* of (*S*)-2-octanol were very close to the real values of these compounds in the mixture samples. As shown in Figure 3.13(b), the determined *ee* values of (*S*)-2-octanol were also very close

to the real *ee* values of these compounds in the mixture samples. More importantly, all of these values were independent from the total concentration of the samples.

3.3.10 Directed Evolution with Real Substrate-based Regio- and Enantioselective Colorimetric HTS Assay

The mutant library creation, mutants' growth and protein induction procedures involved here were as same as that of the previous three rounds of mutagenesis. P450pyr (N100S/F403L/T186I) was used as template to generate a library of mutants for the 4th round of evolution, and the remaining 19 target residues were to be subjected to mutagenesis respectively. After screening a total of around 3500 variants, one interesting mutant, P450pyr(N100S/F403L/T186I/L302V)**,** with very good regio- and enantioselectivity was identified. As shown in the Figure 3.14, based on the GC and chiral GC analysis, this mutant significantly increased the subterminal selectivity of *n*-octane to 92% and gave (*S*)-2-octanol with 72% *ee*.

Motivated by the great improvement in the fourth round, the fifth round of evolution was conducted, and the best mutant in this round P450pyr I83F/N100S/F403I/T186I/L302V, displayed > 99 % subterminal selectivity with 95 % (*S*)-enantioselectivity of *n*-octane. Finally, an excellent mutant was created from the sixth round, P450pyr A77Q/I83F/N100S/F403I/T186I/L302V (P450pyrSM1), which gave > 99% subterminal selectivity and 98% (*S*)-enantioselectivity (Figure 3.15). All the key results of total six rounds of directed evolution was summarised in Table 3.6.

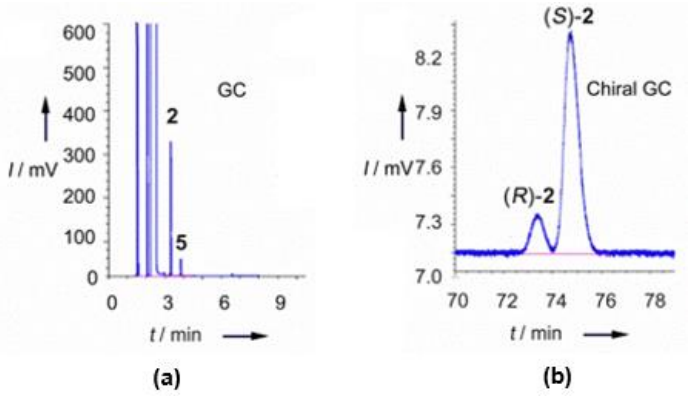


Figure 3.14 Chromatogram of biohydroxylation of octane by recombinant *E. coli* P450pyr mutant [P450pyr(N100S/L302V/F403L/T186I)] with (a) normal GC and (b) chiral GC.

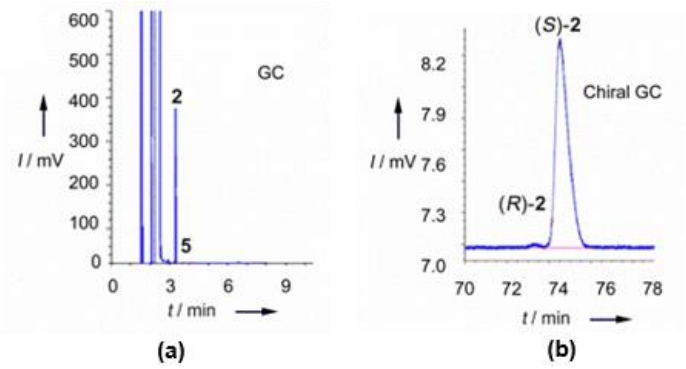


Figure 3.15 Chromatogram of biohydroxylation of octane by recombinant *E. coli* P450pyr mutant [P450pyr(A77Q/I83F/N100S/F403I/T186I/L302V)] with (a) normal GC and (b) chiral GC.

Table 3.6 Directed evolution of P450pyr hydroxylase for regio- and enantioselective subterminal hydroxylation of *n*-octane to *(S)*-2-octanol.

Round	No. of sites saturated	No. of clones screened	No. of positive clones identified	Best mutant	Subterminal selectivity [%][a]	*Ee* of (*S*)-2 [%][b]	Activity [U/g cdw][c]	Relative activity [%][d]
WT	Nil	Nil	Nil	Nil	0	Nil	1.8	100
1	22	4136	10	N100S	5	ND[e]	0.9	49
2	21	3948	20	N100S/F403I	33	ND	2.0	110
3	20	3760	40	N100S/T186I/F403I	40	56	0.8	44
4	19	3572	40	N100S/T186I/L302V/F403I	92	72	1.7	94
5	18	3384	100	I83F/N100S/T186I/L302V/F403I	>99	95	1.8	97
6	17	3196	120	A77Q/I83F/N100S/T186I/L302V/F403I	>99	98	1.7	90

[a] Subterminal selectivity of the best mutant in each round was determined by GC analysis of the products from the biotransformation of 5 mM *n*-octane with 2 g cdw/L of *E. coli* cells expressing the P450pyr mutant in 10 mL potassium phosphate buffer (100 mM; pH 8.0) containing 2% (w/v) glucose at 30 ºC and 250 rpm for 4 h. [b] *ee* of (*S*)-2-octanol was determined by chiral GC analysis of the products from the biotransformation described in [a]. [c] Activity is the specific activity determined for the first 30 min of the biotransformation described in [a]. [d] Activity refers to the activity of *E. coli* cells expressing P450pyr for the terminal hydroxylation of *n*-octane under the same condition. [e] ND: not determined.

3.3.11 Evaluation of New Developed Colorimetric HTS Assay

An example of using new developed colorimetric HTS assay to determine the regio- and enantioselectivity of P450pyr mutant (N100S/T186I/L302V/F403I) in the fourth-round screening was listed in Figure 3.16. Based on the Equation (1) and Equation (2) from Figure 3.8 and the absorbance at OD_{580} in Figure 3.16, the estimated subterminal selectivity and enantioselectivity(*S*) of this mutant based on HTS is 91 % and 70 % respectively. Compared to the real regio- and enantioselectivity data from Figure 3.14 (subterminal selectivity = 92 %; enantioselectivity(*S*) = 72 %), the results estimated from colorimetric assay were very close and equivalent, which also indicated that our new developed three-enzymes based colorimetric regio- and enantioselective HTS screening assay had a very good accuracy.

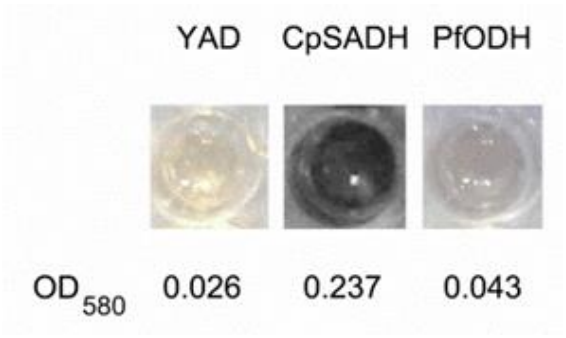


Figure 3.16 Using the colorimetric HTS assay to determination the regio- and enantio-selectivity of P450pyr mutant (N100S/T186I/L302V/F403I) for the hydroxylation of *n*-octane.

3.3.12 Cell Growth and Specific Activity of Recombinant *E. coli* (P450pyrSM1)

Recombinant *E. coli* (P450pyrSM1) was grown in TB medium, and the expression of P450pyr monooxygenase was induced by adding 500 μM of IPTG

and 500 μM of ALA. As shown in Figure 3.17, a cell density of 4.0-4.5 g cdw/L was researched after 10 h. Cells taken at different time points showed different hydroxylation activity toward *n*-octane. The highest specific activity was observed for cells grown after 7 h, in the middle of exponential grow phase. Thus, harvested cells at this time point were used in the following biohydroxylations.

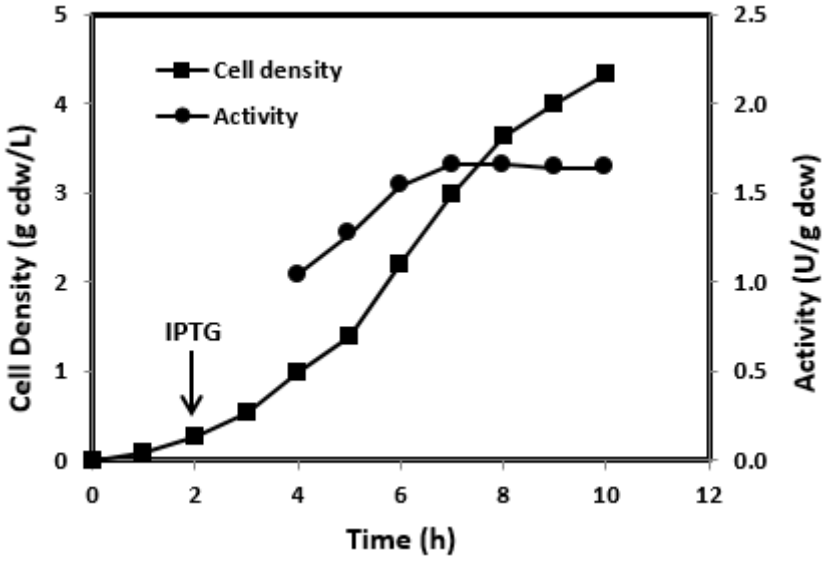


Figure 3.17 Cell growth and specific activity for the hydroxylation of *n*-octane of *E. coli* (P450pyrSM1). Cells were initially grown at 37 ℃ for 2 h, induced by the addition of IPTG and ALA, and then grown at 22 ℃.

3.3.13 Kinetic Study of wild-type P450pyr and P450pyr SM1

The his-tagged P450pyrSM1 was cloned, expressed and purified according to the method described previously for his-tagged P450pyr. To study the kinetics, 5 μM of purified his-tagged wild-type P450pyr and P450pyrSM1 were, respectively, mixed with excess Ferredoxin (Fdx) and Ferredoxin reductase (FdR) in 100 mM phosphate buffer (pH 8.0) containing 2% (w/v) glucose. Biotransformation of *n*-octane at different concentration (0.125-1.25 mM) in the

presence of 2 mM NADH was performed at 30 °C and 300 rpm for 30 min. The product (1-octanol or 2-octanol) concentration was determined by GC analysis, and the results were shown in Figure 3.18 and Table 3.7.

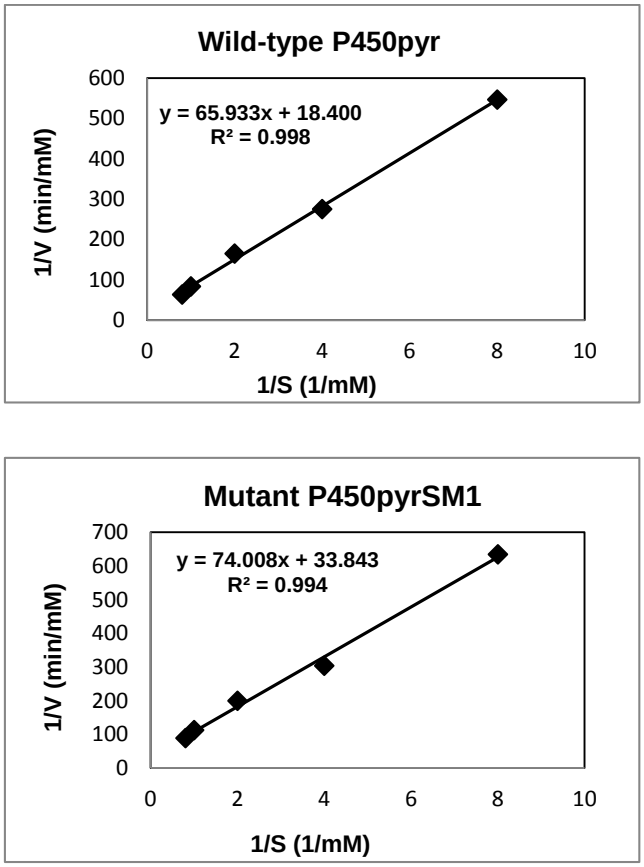


Figure 3.18 Lineweaver-Burk curve of the hydroxylation of *n*-octane with his-tagged P450pyr and his-tagged P450pyr SM1.

Table 3.7 Kinetic data of wild-type P450pyr and P450pyrSM1 for the hydroxylation of *n*-octane

	K_m (mM)	V_{max} (μM/min)	k_{cat} (1/min)	k_{cat}/K_m(1/min.mM)
WT P450pyr	3.583	54.3	10.86	3.031
P450pyrSM1	2.187	29.5	5.9	2.698

3.3.14 Improvement of Product Concentration with *E. coli* (P450pyrSM1-GDH)

Although the mutant, A77Q/I83F/N100S/F403I/T186I/L302V, was found to have comparable activity to wild-type P450pyr for the biohydroxylation of *n*-octane, the produced alcohol concentration still not satisfied. In order to make our P450pyr mutant as an efficient biocatalyst, we aim to improve its activity by following our previously reported procedure for P450pyrTM - adding glucose dehydrogenase (GDH) to our current cofactor regeneration system. *E. coli* (P450pyrSM1-GDH) co-expressing P450pyrSM1 monooxygenase and a glucose dehydrogenase (GDH) was engineered. The cells were grown and used for biohydroxylation in the same procedures described for *E. coli* (P450pyrSM1). As illustrated in Figure 3.19, the new constructed recombinant with cofactor recycling system improved product yield by 1.7 times.

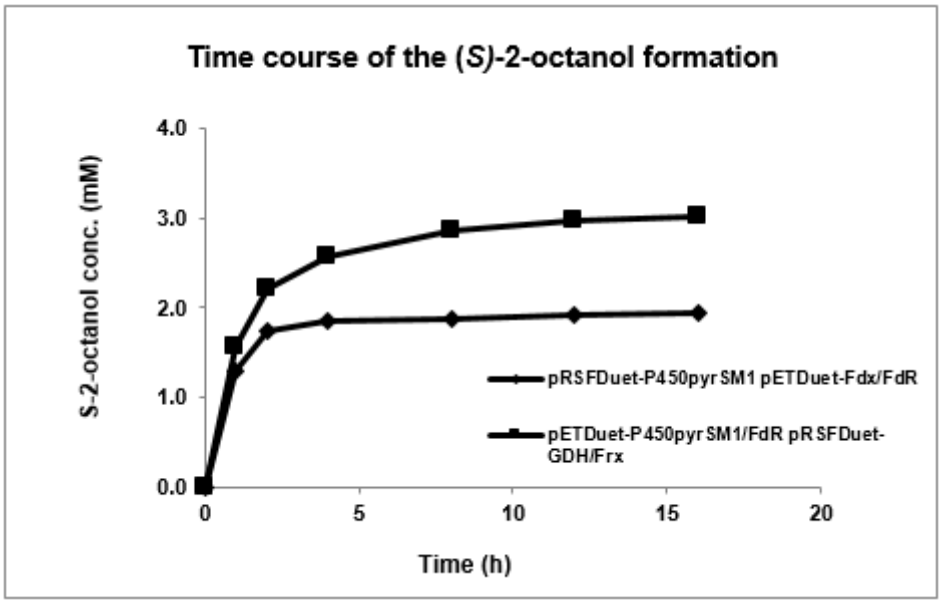


Figure 3.19 Time course of the formation of *(S)*-2-octanol in the biohydroxylation of 10 mM *n*-octane with resting cells of *E. coli* BL21*(DE3)* (P450pyrSM1 with or without GDH) in K buffer (100 mM; pH 8.0) containing glucose (2%; w/v) at 30 ^{o}C; pRSFDuet-P450pyrSM1 + pETDuet-Fdx/FdR (♦); pETDuet-P450pyrSM1/FdR + pRSFDuet-GDH/Fdx (■).

3.3.15 Molecular Dynamics and Docking Simulation

To understnd the stereochemistry outcome, the hydroxylation of *n*-octane with P450pyr and P450pyrSM1 were investigated by molecular dynamics and docking simulation. A reshaping of the binding pocket that is responsible for the subterminal and enantioselectivity was observed for the enzyme mutant. As shown in the Figure 3.20A, the binding pocket of P450pyr is very compact (9.3 Å long, I102 to T259 and 7.1 Å wide, G255 to L302), the substrate stacks with a big hydrophobic cluster comprising of the northwest and south residues, and *n*-octane hence takes a vertical binding pose. The distance between the terminal C-H and the heme-O is the shortest (2.8 Å), giving rise to terminal hydroxylation (Figure 3.20C). In P450pyrSM1, the hydrophobic cluster is disrupted by the F403I and L302V mutation (Figure 3.20B). The binding pocket consequently extends southwards reaching 10.3 Å of width. The substrate adopts a horizontal binding pose along the heme plane, in north-south orientation. This special geometry of the binding pocket gives clearly catalytic preference to (*S*)-subterminal hydroxylation (Figure 3.20D): the distance between the subterminal C-H_S and the heme-O is 2.7 Å, while the distance between the subterminal C-H_R and the heme-O is 3.8 Å.

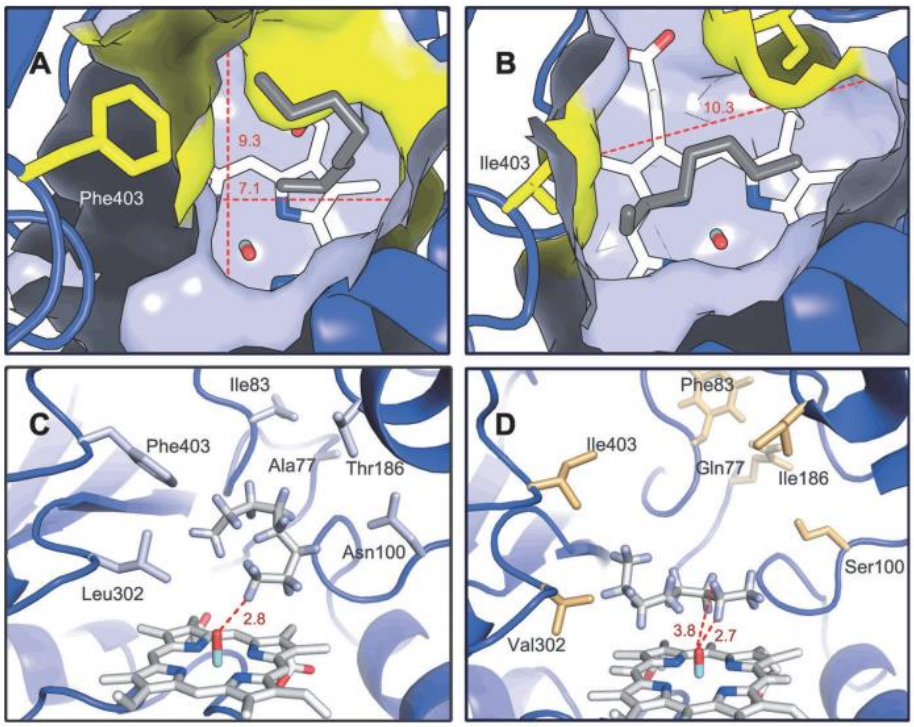


Figure 3.20 Substrate *n*-octane-P450pyr enzyme binding pose. (A) - (B): Top-view of the surface structure of (A) P450pyr and (B) P450pyrSM1. Hydrophobic clusters are shown in yellow. (C) - (D): Side view of the binding pose of (C) P450pyr and (D) P450pyrSM1. The distances between the heme-oxygen atom and the nearby hydroxgen atom of *n*-octane are denoted by dashed lines.

3.3.16 Regio- and Enantioselective Hydroxylation of Propylbenzene

The positive variants obtained in six rounds of evolution of P450pyr for the subterminal hydroxylation of *n*-octane were examined for the subterminal hydroxylation of propylbenzene to produce (*S*)-1-phenyl-2-propanol. The best mutant obtained from the 6th round, I83F/N100S/T186I/L251V/L302V/F403I (P450pyrSM2), gave 98% subterminal selectivity and 95% (*S*)-enantioselectivity for the hydroxylation of propylbenzene. These selectivities are clearly confirmed in Figure 3.21.

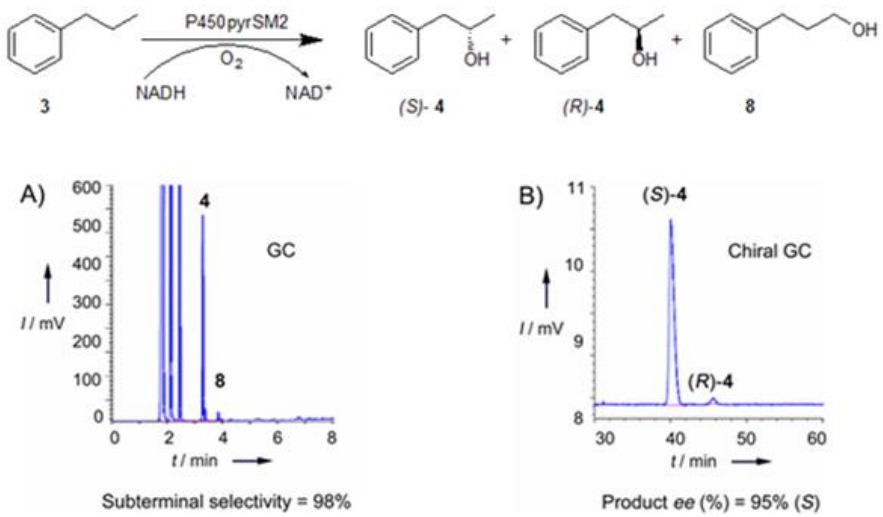


Figure 3.21 Analysis of the products from regio- and enantioselective subterminal hydroxylation of propylbenzene with P450pyrSM2. (A) GC chromatogram. (B) Chiral GC chromatogram.

3.4 Conclusion

Regio- and enantioselective hydroxylation of *non*-activated carbon atoms is a useful reaction for the production of valuable compounds such as enantiopure alcohols and steroids. Compared with classical organic chemical method, the monooxygenase-catalysed hydroxylation is highlighted because it is more green, clean, selective and efficient. However, most monooxygenases show terminal selectivity, only P450BM-3 shows some subterminal hydroxylation activity, but the regio- and enantio-selectivity is poor. In this project, we solve this problem by engineering the P450pyr with directed evolution to create a set of new enzymes for enantioselective hydroxylation of *non*-activated carbon atoms at subterminal position.

In this chapter, a terminal selective cytochrome P450pyr has been successfully evolved for the subterminal hydroxylation of alkane at *non*-activated carbon atom, giving a sextuple mutant P450pyrSM1 with excellent subterminal and

enantioselectivity towards *n*-octane and providing with the first enzyme for this type of highly selective alkane hydroxylation. This represents also the first success of fully altering the regioselectivity of a monooxygenase in biohydroxylation by evolution. Subterminal hydroxylation with the engineered P450pyr hydroxylase allows for regio- and enantioselective functionalization of alkane, a useful and challenging reaction in classic chemistry. P450pyrSM1 catalyses the hydroxylation of *n*-octane to produce (*S*)-2-octanol in 98% *ee* with >99 % subterminal selectivity. Another sextuple mutant P450pyrSM2 hydroxylates propylbenzene to give *(S)*-1-phenyl-2-propanol in 95% *ee* with 98 % subterminal selectivity. Both subterminal alcohols (*S*)-2-octanol and *(S)*-1-phenyl-2-propanol are useful and valuable intermediates for chemical and pharmaceutical synthesis. A novel, accurate, sensitive, and simple colorimetric HTS assay is developed for measuring both regioselectivity and enantioselectivity for a hydroxylation reaction. This provides with a solid basis for the successful evolution of regio- and enantioselective P450pyr hydroxylase for subterminal hydroxylation. The colorimetric HTS assay could be generally applicable for the discovery of other type of regio- and enantioselective enzymes for hydroxylations. The molecular modelling on the hydroxylations of *n*-octane with P450pyr and P450pyrSM1 mutant gives insight into the structure basis and the role of key mutations for fully altering the terminal selectivity to subterminal selectivity and gaining excellent enantioselectivity. The general knowledge and information obtained might be useful for guiding future engineering of other selective P450 enzymes.

Chapter 4: Evolving P450pyr Monooxygenase for Highly Regioselective Terminal Hydroxylation of *n*-butanol to 1,4-butanediol

4.1 Introduction

1,4-butanediol is a very important and useful chemical that is used predominantly as intermediate to synthesise a wide variety of chemical compounds.[272-274] It is able to react with many different bifunctional reagents to produce different types of polymers, e.g. polyurethane and polybutylene terephthalate.[275-277] It could also be further converted to gamma butyrolactone (GBL) and gamma hydroxybutanoic acid (GHB), both of which are important feedstock for the production of various valuable pharmaceuticals and fine chemicals.[272-274] Attributed to the continuous expansion in manufacturing polymers, fine chemicals and pharmaceuticals, the global demand for 1,4-butanediol grows steadily over the years. The global demand was approximately 2.4 million tonnes in 2010, and is expected to grow at an annual growth rate of 5% from 2012 to 2018.[278] Currently, 1,4-butanediol is produced from petrol chemicals involving multi-step synthesis, high temperature and pressure, and the generation of toxic and flammable by-products and carbon footprint.[279-281]

We are interested in developing a green and sustainable route to produce 1,4-butanediol directly from *n*-butanol, which can be produced by fermentation from renewable resource,[282,283] *via* hydroxylation. Thus far, no chemical- or bio-catalyst has been reported for this reaction with the exception of a CYP153 application under high pressure described in patent application WO

2014079683 A1 from Evonik Industries describing the principal feasibility of this oxidation from alkanes to terminal diols by a CYP153 enzyme.[284] Although many monooxygenases, such as P450,[120-122] sMMO,[123,124] and alkB,[125,126] were known for the hydroxylation of alkanes, only two cytochrome P450 enzymes, CYP153A16 from *Mycobacterium marinum* and CYP153A *P*. sp. from *Polaromonas* sp. strain JS666, were reported for the ω-hydroxylation of C_8-C_{12} primary alcohols, demonstrating the potential for P450s to catalyse terminal hydroxylation of alcohols.[107] While CYP153A *P*. sp. could only oxidise 1-octanol and 1-nonanol to the corresponding α,ω-diols with very low yields, CYP153A16 was able to ω-hydroxylate C_8-C_{12} primary alcohols, displaying highest activity for 1-nonanol with 46.5% conversion. However, the selectivity of these enzyme are both poor, with only 40–58 % and 74-96 % of the product was corresponding α,ω-diols, and the rest consisted of aldehydes, fatty acids and some other by-products. Besides, both enzymes do not show any hydroxylation activity toward small alcohols.[107] Directed evolution has become a useful tool for engineering enzymes with new substrate acceptance, improved activity, regio-, and stereo-selectivity, or enhanced enzyme stability.[18,285,286] Here we reported the development of the first enzyme for highly regioselective terminal hydroxylation of *n*-butanol to 1,4-butanediol by directed evolution of P450pyr monooxygenase.

As mentioned above, P450pyr from *Sphingomonas* sp. HXN-200, a class I P450 monooxygenase, is able to catalyse the regio- and stereoselective hydroxylations.[189-191] Directed evolution of P450pyr was also successfully performed to achieve excellent enantioselectivity for the hydroxylation of *N*-benzylpyrrolidine,[226,227] as well as excellent regio- and enantioselectivivity for

the subterminal hydroxylation of alkanes in Chapter 3. Although the substrate scope of P450pyr is diverse, including *n*-alkanes, cyclic alkanes and *N*-heterocycles, it is still limited to medium-sized hydrophobic substrates, and initial experiments showed that P450pyr did not accept *n*-butanol as the substrate for any hydroxylation. Thus, the target of this project was to create new P450pyr enzymes for terminal hydroxylation of *n*-butanol by switching the substrate acceptance from a hydrophobic to hydrophilic compound (Figure 4.1).

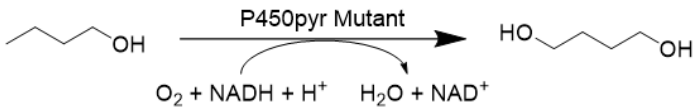


Figure 4.1 P450pyr Monooxygenase-catalysed regioselective terminal hydroxylation of *n*-butanol to 1,4-butanediol.

In order to achieve this goal, we first developed a colorimetric HTS assay for the enzyme directed evolution. It was a very difficult task, because both substrate and product are small colourless molecules with no UV absorbance and no distinguishing functional difference. To solve all these problems, instead of a real substrate-based HTS assay we developed a surrogate substrate-based colorimetric HTS method in this project. The new developed HTS assay would be used to investigate all the P450pyr mutants which were generated from ISM based on *n*-butanol-P450pyr docking structure model analysis. All the good mutants identified from this surrogate substrate-based HTS would be tested with actual substrate *n*-butanol and analysed with various analytical tools. Several rounds of directed evolution would be performed and mutants which capable of producing 1,4-butanediol directly from *n*-butanol would be created and identified. Besides, the roles of the mutated amino acid residues on changing

the substrate scope and creating activity toward hydrophilic molecules would also be investigated based on simulation modelling and docking structure, which gave us a better understanding of the relationships between substrate scope and structure of P450pyr and its mutants.

4.2 Experimental Section

4.2.1 Chemicals

Following chemicals were purchased from Sigma-Aldrich and used without further purification: 2-methoxyethanol (≥ 99 %), *n*-butanol (≥ 99 %), 1,2-butanediol (≥ 98 %), 1,3-butanediol (≥ 99 %), 1,4-butanediol (≥ 99 %), purpald (≥ 99 %), δ- aminolevulinic acid hydrochloride (ALA) (≥ 97 %), Isopropyl β-D-1-thiogalactopyranoside (IPTG) D-Glucose (≥ 99.5%), dimethylformamide (DMF) (≥ 99 %), ethanol (≥ 99 %), ammonium sulphate (≥ 99 %), BSTFA (contains 1% TMCS, 99%), Sodium phosphate dibasic (≥ 99 %), and Sodium phosphate monobasic (≥ 99 %).

4.2.2 Strains and Biochemicals

Escherichia coli BL21(*DE3*) strain was purchased from Novagen. Plasmids pRSFDuet-P450pyr and pETDuet-Fdx-FdR were obtained from our own laboratory collections. Restriction enzyme *Dpn* I, Deoxynucleotide (dNTP) Solution Mix, and Q5 Hot Start High-Fidelity DNA Polymerase were purchased from New England Biolabs (NEB). Tris-acetate-EDTA (TAE) buffer, DNA loading dye, and DNA ladder were purchased from Thermo Scientific. Oligonucleotides were synthesised by AIT biotech, Singapore. LB Broth, Bacto

Yeast Extract, and Bacto Tryptone were purchased from Biomed Diagnostics. Antibiotic ampicillin and kanamycin were purchased from Sigma-Aldrich. QIAquick Gel Extraction Kit (Qiagen) and QIAprep spin plasmid miniprep Kit were purchased from Qiagen.

4.2.3 Generation of P450pyr Monooxygenase Mutant Library

To generate P450pyr monooxygenase mutant library, PCR was carried out on each of the selected amino acid site for ISM using the designed primers shown in Table 4.1. Each PCR reaction tube contained the following mixture: 2.5 μL forward and reverse primers of a particular target site, 10 μL Q5 reaction buffer, 1 μL 10 mM dNTP mixture, 0.5 μL (10 ng) template DNA, 0.5 μL Q5 Hot Start High-Fidelity DNA polymerase and 33 μL of nuclease-free H_2O. PCR amplification was carried out on Bio-Rad S1000 Thermal Cycler using the following thermal cycling program: 98 ℃ for 30 s, 25 cycles [98 ℃ for 10 s, annealing temperatures for each pair of primers (see Table S1 Tm*) for 30 s, 72 ℃ for 3 min], and 72 ℃ for 5 min.

Table 4.1 Primer sequences used for site-directed mutagenesis

Amino acid		Primer Sequence (5' to 3') [a]	Tm* (℃) [b]
A77	A77-F	C TCG TCC GAT NNK GGA TAT GGC G	69
	A77-R	C GCC ATA TCC MNN ATC GGA CGA G	
I82	I82-F	GGA TAT GGC GGC NNK ATA ATC GAT GAC	69
	I82-R	GTC ATC GAT TAT MNN GCC GCC ATA TCC	
I83	I83-F	GGC GGC ATC NNK ATC GAT GAC G	70
	I83-R	C GTC ATC GAT MNN GAT GCC GCC	
I88	I88-F	C GAT GAC GGC NNK CAA AAA GG	65
	I88-R	CC TTT TTG MNN GCC GTC ATC G	
Q89	Q89-F	GAC GGC ATT NNK AAA GGT GGC G	68
	Q89-R	C GCC ACC TTT MNN AAT GCC GTC	
L98	L98-F	GC GGA CTG GAT NNK CCC AAT TTC	67
	L98-R	GAA ATT GGG MNN ATC CAG TCC GC	
P99	P99-F	GGA CTG GAT CTT NNK AAT TTC ATC GCG	67
	P99-R	CGC GAT GAA ATT MNN AAG ATC CAG TCC	

N100	N100-F N100-R	G GAT CTT CCC NNK TTC ATC GC GC GAT GAA MNN GGG AAG ATC C	63
I102	I102-F I102-R	CCC AAT TTC NNK GCG ATG GAT C G ATC CAT CGC MNN GAA ATT GGG	65
A103	A103-F A103-R	CC AAT TTC ATC NNK ATG GAT CGG CC GG CCG ATC CAT MNN GAT GAA ATT GG	68
S182	S182-F S182-R	CTT ACC CGC NNK TCG GAT GTG AC GT CAC ATC CGA MNN GCG GGT AAG	69
D183	D183-F D183-R	GC TGG TCG NNK GTG ACA ACC GGT TGT CAC MNN CGA CCA GC	67
T185	T185-F T185-R	G GAT GTG NNK ACC GCA GC GC TGC GGT MNN CAC ATC C	64
T186	T186-F T186-R	GAT GTG ACA NNK GCA GCA CC GG TGC TGC MNN TGT CAC ATC	64
L251	L251-F L251-R	GTA CTT NNK CTG ATC GTT GGC G C GCC AAC GAT CAG MNN AAG TAC	64
V254	V254-F V254-R	C CTG ATC NNK GGC GGG AAC GTT CCC GCC MNN GAT CAG G	67
G255	G255-F G255-R	C CTG ATC GTT NNK GGG AAC G C GTT CCC MNN AAC GAT CAG G	64
D258	D258-F D258-R	C GGG AAC NNK ACC ACA CG CG TGT GGT MNN GTT CCC G	64
T259	T259-F T259-R	C GGG AAC GAT NNK ACA CGC GCG TGT MNN ATC GTT CCC G	64
L302	L302-F L302-R	G CAA ACA CCG NNK GCT CAT ATG C G CAT ATG AGC MNN CGG TGT TTG C	68
M305	M305-F M305-R	CTT GCT CAT NNK CGC CGC ACG CGT GCG GCG MNN ATG AGC AAG	71
F403	F403-F F403-R	CGT TCA AAT NNK GTG CGC GG CC GCG CAC MNN ATT TGA ACG	66

[a] N represents any possible bases and K represents either guanine (G) or thymine (T)
[b] Tm* is the value of optimal annealing temperatures calculated with NEB Tm Calculator https://www.neb.com/tools-and-resources/interactive-tools/tm-calculator

PCR products were subjected to DNA gel electrophoresis and the location of DNA in gel was detected using gel imaging system. The product band was excised from the gel and then purified using QIAquick Gel Extraction Kit (Qiagen). The purified DNA was digested with methylated restrictive enzyme *Dpn I* at 37 ℃ overnight. P450pyr mutant library was created by transferring the digested DNA into competent *E. coli* BL21*(DE3)* cells containing the pETDuet-Fdx-FdR vector. According to the ISM theoretical calculations, a total of 94 mutant clones are required for each amino acid residue to have 95% coverage of all 20 possible amino acids. Therefore, 188 clones for each site were taken to

two 96-deep well plates in our experiment to ensure the coverage of > 95% for all possible 20 amino acids.

4.2.4 General Procedure of Screening P450pyr mutants for Terminal Hydroxylation of Alcohol by Using Surrogate Substrate-Based Colorimetric HTS Assay

The generated P450pyr mutants were picked from LB agar plate and inoculated into 96-deep well plate with 600 μL TB medium containing 50 μg/mL kanamycin and 100 μg/mL ampicillin. The deep-well plate was shaken at 900 rpm and 37 ℃ for 8 h. 100 μL cell cultures were taken out and mixed with 100 μL 50% glycerol to prepare cell stock which was stored in -80 ℃ refrigerator. After that, 800 μL TB medium containing 0.5 mM IPTG, 0.5 mM ALA, 50 μg/mL kanamycin and 100 μg/mL ampicillin were added to the remaining cell cultures in each well to induce the P450pyr protein expression. After expressing at 22 ℃ for 8 h, *E. coli* cells were harvested by centrifugation at 3220 g and room temperature for 15 min. 500 μL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose and 5 mM 2-methoxyethanol were added to the cell pellets in each well. The biotransformation was performed at 900 rpm and 30 ℃ for 4 h and stopped by centrifugation at 3220 g for 10 min.

100 μL aliquot from each well were transferred to a microliter plate that contained 50 μL aqueous solution of 1 M NaOH and 100 mM purpald. After incubating at room temperature for 10 min, the absorbance of each well was measured at 550 nm using spectrophotometer (Thermo Scientific Multiskan Go). Mutants with hydroxylation activity towards the surrogate substrate 2-methoxyethanol will produce formaldehyde, which react with purpald to give

purple colour compound. The principle of this surrogate substrate-based colorimetric HTS assay was shown in Figure ***4.6***4.2A. The general procedure of this assay was shown in Figure 4.2B.

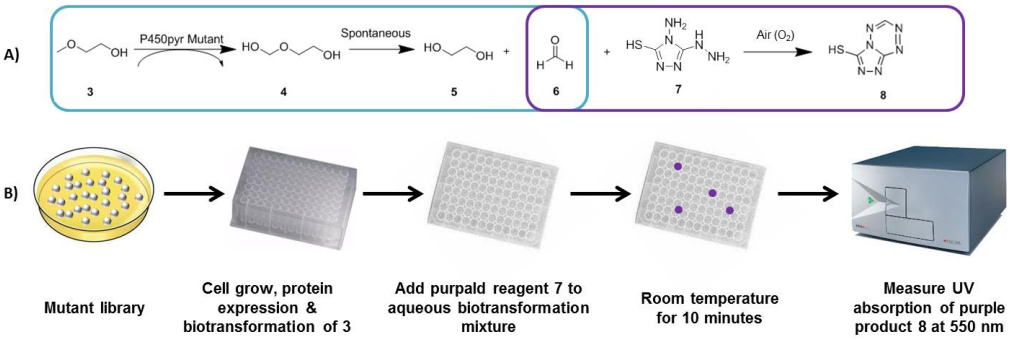


Figure 4.2 The principle and procedure of surrogate substrate-based HTS assay. A) Principle of surrogate substrate-based colorimetric HTS assay B) Procedure of screening E. coli (P450pyr mutants) for the terminal hydroxylation of n-butanol by using surrogate substrate-based colorimetric HTS assay.

4.2.5 General Procedure for Biohydroxylation of *n*-butanol with *E.coli* Expressing Positive P450pyr Mutants in Shaking Flask

E. coli strains containing positive P450pyr mutants were inoculated into 50 mL TB containing 50 μg/mL kanamycin and 100 μg/mL ampicillin in multiple shaking flasks, respectively. The mixtures were shaken at 250 rpm and 37 ℃ until OD_{600} researched 0.8. The proteins expression were induced by adding IPTG and ALA to a final concentration of 0.5 mM, respectively, and the mixture was shaken for additional 6 h at 250 rpm and 22 ℃.

The cell pellets were collected by centrifugation at 3220 g and room temperature for 10 min. The cell were resuspended to a density of 24 g cdw L^{-1} in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose

and 5 mM of *n*-butanol. Biotransformation was performed at 300 rpm and 30 ℃ for 4 h, and the formation of 1,4-butanediol was analysed by GC.

4.2.6 GC Analysis of 1,4-butanediol

GC analysis was performed by using Agilent GC 6890 with Agilent HP-INNOWAX column (25 m × 0.32 mm) in splitless injection mode with inlet temperature of 280 ℃ and detector temperature of 220 ℃. Temperature program: 60 ℃ for 3 min, increased to 180 ℃ at 30 ℃/min, and further increased to 220 ℃ at 20 ℃/min. Retention times: 6.86 min for 2-hexnaol (internal standard), 8.58 min for 1,2-butanediol, 8.93 min for 1,3-butanediol, and 9.47 min for 1,4-butanediol.

To analysis the biotransformation mixture, the supernatant was first concentrated at room temperature with Eppendorf Concentrator Plus. 32 % (w/w) ethanol and 16 % (w/w) ammonium sulfate were then added to the concentrated supernatant to form two-liquid phases to extract 1,4-butanediol into ethanol phase. 100 µL ethanol phase were taken and used for GC analysis.

4.2.7 MS Analysis of the Derivative of 1,4-butanediol Made by BSTFA-TMCS Method

To further confirm the identity of the biotransformation product is 1,4-butanediol, the mixture from biohydroxylation of *n*-butanol was derivatised by BSTFA-TMCS method. 1 mL biotransformation supernatant was dried at 45 ℃ with Eppendorf Concentrator Plus, followed by the addition of 20 µL DMF and 100 µL BSTFA-TMCS regent. The derivatisation reaction was conducted at 70 ℃ for 30 min. The derivatised mixtures were centrifuged at 20376 g for 5

min, and finally 60 µL sample was used for GC-MS analysis. Similarly, the standard 1,4-butanediol was derivatised by using the same procedure.

GC-MS analysis was performed using Agilent GC-MSD system 6890-5973A with GC-MSD HP-5 MS column in splitless injection mode with inlet temperature of 250 ℃ and the MS interface transfer line temperature of 270 ℃. Temperature program: 80 ℃ for 1.5 min, increased to 140 ℃ at 10 ℃/min, hold at 140 ℃ for 3 min, increased to 300 ℃ at 100 ℃/min and hold at 300 ℃ for 5 min. Retention time for the derivatised (1,4-butanediol bis (trimethylsilyl)ether) is 6.25 min.

4.2.8 Molecular Modelling of Substrates Docking in P450pyr and Its Mutants

The structure of P450pyr was obtained based on x-ray structure and enhanced using MODELLER[287] *via* CHIMERA[288] interface to supplant seven non-critical missing residues via the high precision Discrete Optimized Protein Energy score (DOPE-HR) algorithm.[289] Amino acid residue mutations were performed *via* ACCELRYS Discovery Studio. Ramachandran plots for P450pyr, P450pyr I83M and P450pyrI83M/I82T models were generated before and after molecular dynamics (MD) simulations, and all revealed > 90 % of residues within allowed regions.

MD simulations of wild type and mutant enzymes were all conducted with GROMACS 4.5[290] using GROMOS 53a6 force field at a 0.2 ps recording interval in a water system with 150 mM NaCl. Long-range electrostatic interactions were modelled *via* Particle Mesh Ewald (PME) method, and short range van der Waals interactions were localized at 10 Å. Bond lengths were

constrained *via* LINCS algorithm.[291] Energy minimization was performed with Steepest descent algorithm, in which V-rescale thermostat and Parrinello-Rahman barostat were used to stabilize temperature at 300 K and pressure at 1 atm, respectively. 5 ns production runs coupled to NPT were conducted, and the recorded enzyme conformers were clustered using Gromos algorithm with RMSD cutoff at 0.18 nm.[292] The main cluster centroids were docked with *n*-butanol using Autodock VINA,[293,294] where the catalytic heme was modified as the Cpd I ferryl-oxo-complex state.[120] Docking results were evaluated via binding energy scores with docked clustering at 1 Å non-fitted RMSD cutoff. Catalytically relevant active binding postures for terminal hydroxylation of *n*-butanol required the terminal substrate-C to be at closest approach to heme-O atom at a distance of < 6 Å, and also with Fe-O-H angles in the optimal range of 110^{o} to 160^{o} as part of established geometric criteria for substrate-heme binding in P450 enzymes. [295,296]

4.3 Results and Discussion

4.3.1 Development of Colorimetric HTS Method

In this project, develop a suitable colorimetric HTS assay for discovery of P450pyr mutants for terminal hydroxylation of *n*-butanol was indeed a challenge. There is no distinguishing functional difference between *n*-butanol and 1,4-butanediol except for the one more -OH group, neither the substrate nor the product are UV or fluorescence sensitive, and no enzyme or chemical catalyst were reported to react with butanol or 1,4-butanediol exclusively. Therefore, using a surrogate substrate instead of real substrate would be a more

practical choice, so far many enzymes have already been successfully engineered by using surrogate substrates.[297]

After extensive literature search, a surrogate substrate, 2-methoxyethanol, was selected to carry out the assay in place of *n*-butanol. As illustrated in Figure 4.3A, terminal hydroxylation of 2-methoxyethanol produced an unstable hemiacetal functional group, which is unstable would quickly become formaldehyde and ethylene glycol. The amount of produced formaldehyde would further react with an uncoloured compound, purpald, to produce an intense purple product in the presence of air within 10 minutes. Although purpald could react with both aldehydes and ketones, only the reaction with aldehydes could produce a coloured production, reaction with ketones produces a colourless product because it cannot undergo further oxidation to form an extended heterocyclic conjugated system.

The performance of the assay was determined by standard calibration. It was found that the assay produced good linearity and excellent sensitivity from 0.05 mM to 1 mM within 10 minutes (Figure 4.3B and Figure 4.3C). This proved the effectiveness of the assay as an efficient and sensitive method to quantify the amount of formaldehyde present in a particular sample mixture.

Purpald's reaction with various aldehydes proceeds at different rate, with formaldehyde completing the reaction within 10 minutes.[298] Other aldehydes such as acetaldehyde or even the aldehyde group in open chain form of glucose proceed at a slower rate than formaldehyde. It is important to note that glucose is present in significant amount in the reaction mixture. 2 % w/v glucose is equivalent to about 100 mM concentration. This high concentration of glucose may interfere with the assay. An experiment was conducted to determine the

effect of 2% w/v glucose on the assay and it was found that the influence of glucose is insignificant until more than 2 hours after the addition of purpald. This result allowed us to assay of formaldehyde concentration in a quick and selective manner.

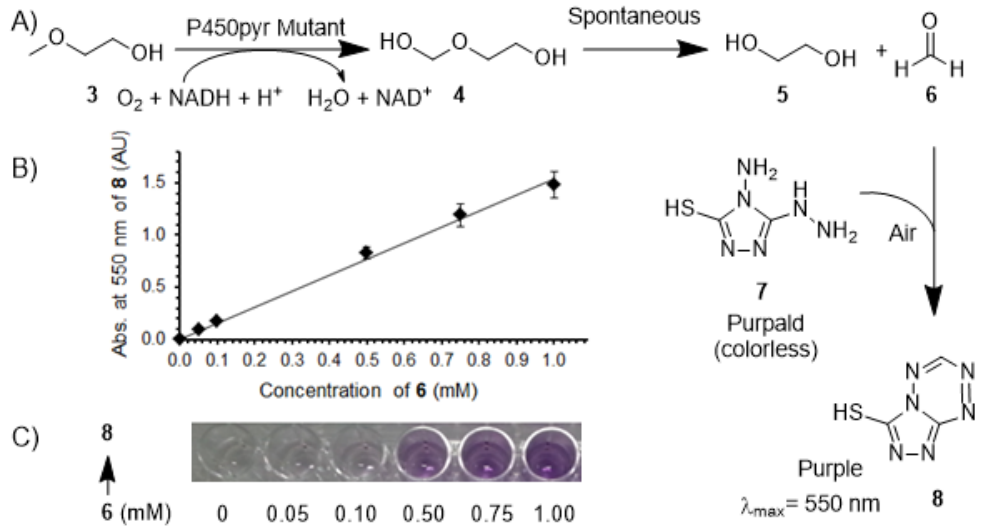


Figure 4.3. Principle and sensitivity of the colorimetric HTS assay using surrogate substrate and purpald. A) Principle of the colorimetric HTS assay for terminal hydroxylation of surrogate substrate. B) Calibration curve of the concentration of formaldehyde and the UV absorption at 550 nm of purple compound produced by the treatment of formaldehyde and purpald. Data are the mean values with standard deviations of three replicates. C) Photos of purple produced from formaldehyde at different concentration by the addition of purpald for 10 min.

4.3.2 Identification of suitable amino acid residues of P450pyr for Iterative Saturation Mutagenesis (ISM)

To identify the suitable amino acid residues for ISM, *n*-butanol was docked into the x-ray structure of P450pyr (PDB ID: 3RWL) with the catalytic heme modified as the Cpd I ferryl-oxo-complex. The catalytically relevant active binding posture for *n*-butanol was obtained based on established geometric criteria for substrate-heme binding in P450 enzymes.[295] As shown in Figure 4.4, totally twenty two amino acids residues were selected for ISM. Elven amino

acids were located within ~6.5 Å of the docked substrate: (i) important hydrophobic residues: Ile82, Ile83, le102, Leu302, Met305, and Phe403; and (ii) residues on the heme-proximal helix: Leu251, Val254, Gly255, Asp258, and Thr259. Four amino acids located along another helix (Ser182, Asp183, Thr185 and Thr186) were also chosen due to their relatively abrupt position in the substrate access channel. Seven important amino acids from the 'big loop' (Ala77, Ile88, Gln89, Leu98, Pro99, Asn100, and Ala103) were also chosen for ISM, since in our previous study an improvement in P450pyr enantioselectivity was attributed to a reshaped binding pocket *via* inward movement of this "big loop" (Ser74-Asp105).[103]

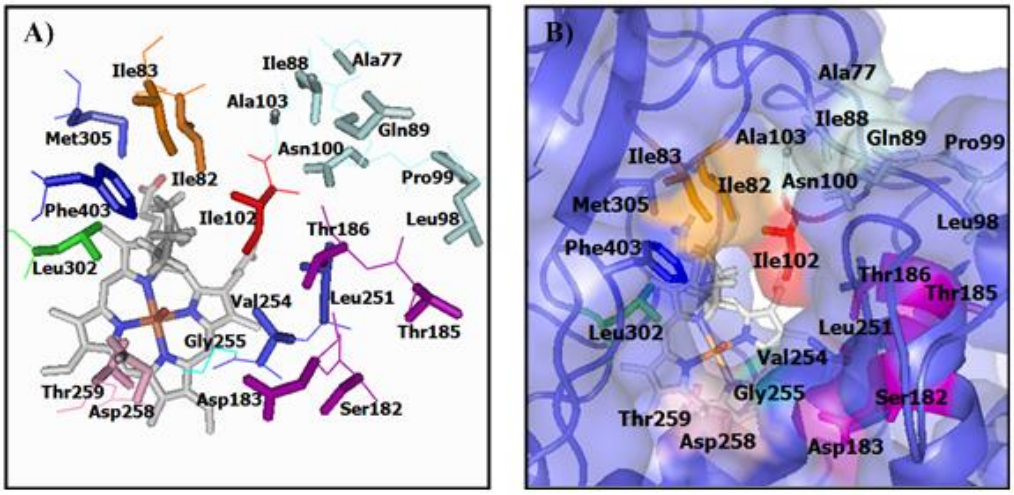


Figure 4.4 Selection of suitable amino acid residues for ISM. (A) Spatial overview of 22 selected amino acid residues for ISM, in relation to the active binding posture of *n*-butanol (grey cartoon stick with hydrogen atoms) in P450pyr. (B) Surface presentation of the same overview with additional structural motifs for a more aesthetic appreciation of the enzyme-substrate binding posture.

4.3.3 Generation of P450pyr Monooxygenase Mutant Library

The mutant library of each selected site was generated by site-saturation mutagenesis in which all possible 20 amino acids will be generated in the mutant library. Complementary primers were designed for each amino acid residue by

using NNK sequence with 9-12 nearby bases. N represents any possible bases and K represents either guanine (G) or thymine (T). Generally, there three advantages of using "NNK" sequence: firstly it covers all possible 20 amino acids; secondly the amount of analysis effort is greatly reduced with only 32 possible codons; more importantly, the distribution of degenerate codons is more even in NNK.[205]

PCR amplification was done using the designed primers shown in Table 4.5. In order to make the PCR efficiency, it is crucial to use the appropriate annealing temperature for each pair of primes. Figure 4.5 shows an example of PCR products of four different sites, the PCR product bands as well as primer dimers can be seen clearly in this image.

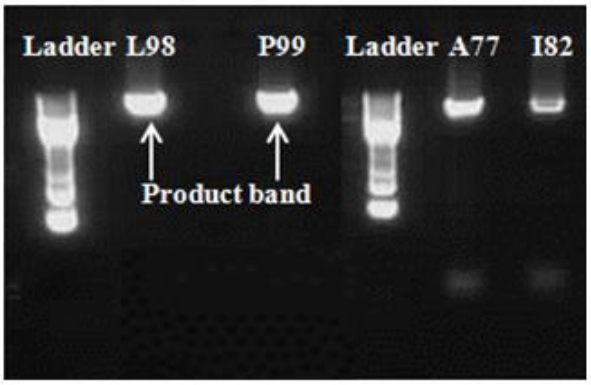


Figure 4.5 Image of gel electrophoresis after PCR amplification. Site L98, P99, A77 and I82 are shown.

4.3.4 Screening of P450pyr Mutants by Using Surrogate Substrate-Based Colorimetric HTS Assay

Figure 4.6 showed a sample 96-well plate that was in the process of colorimetric high-throughput screening assay. Mutants with high hydroxylation activity towards the surrogate substrate produce sufficient formaldehyde that showed an

intense purple colouration visible to naked eyes. Exact quantification of absorbance was performed with a spectrophotometer.

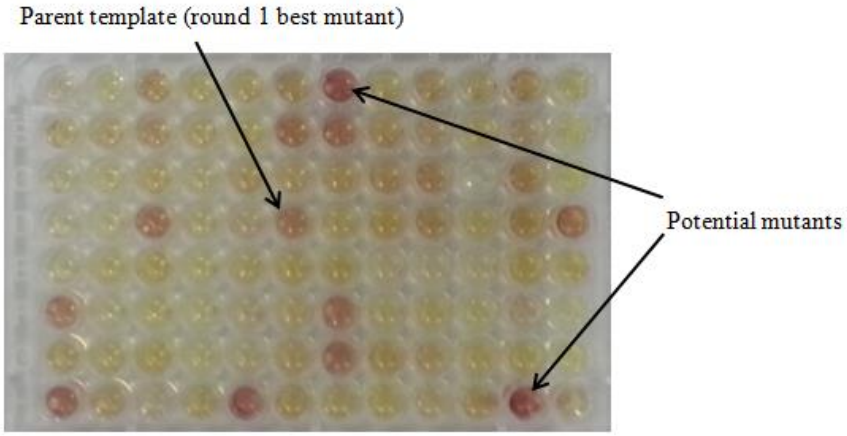


Figure 4.6 A sample 96-well plate in the process of HTS assay (2nd round).

4.3.5 Selection of Identification and Quantification Method for 1,4-butanediol

HPLC is generally used for analysing non-volatile chemicals, as well as proteins and amino acids. It is also widely used in identification and quantification of polar compounds to bypass the effort of samples extraction. For 1,4-butanediol, however, its structure is too simple and does not have conjugated system for UV detector, the most common HPLC detector. Though universal detectors such as refractive index (RID) detector are available, it is about 1000 times less sensitive compare with UV detector. Under our lab condition, the detection limit of 1,4-butanediol using RID is about 1 mM. As such, we consider HPLC analysis is not practical until the product concentration is high enough after several rounds of directed evolution.

GC is another frequently used analytical method for detection of organic molecules. Compare with HPLC, most GC columns are sensitive to moisture and thus GC analysis requires organic extraction since most biotransformation reaction are carried out in aqueous phase. Since 1,4-butanediol is a very hydrophilic compound, which is almost insoluble in most organic solvents such as chloroform and hexane, the organic extraction of 1,4-butanediol from biotransformation mixture is very challenge. An ideal organic solvent suitable to extract as much 1,4-butanediol as possible is the one with the highest relative polarity and lowest solubility in water. After an extensive literature search and testing with several organic solvents, including: cyclohexanol, ethanol, 1-propanol, 1-butanol, 1-octanol and acetone, an ethanol-ammonium sulphate method adapted from Li *et. al* was identified as the most efficient system for 1,4-butanediol extraction. [299]

4.3.6 Directed Evolution with Surrogate Substrate-based Colorimetric HTS Assay

In the first round of evolution, all of the 22 selected amino acid positions were subjected to ISM. P450pyr variants were generated at each selected amino acid site by PCR using NNK degenerate codon. The digested DNAs were transferred into competent *E. coli* BL21(*DE3*) cells containing the pETDuet-Fdx-FdR vector. The created P450pyr variants were inoculated, grown, expressed, harvested, and subjected to the surrogate substrate-based colorimetric HTS assay on microtiter plate by the use of 5 mM 2-methoxyethanol and 4 h

biotransformation, treatment with NaOH and purpald for 10 min, and measurement of OD_{550} to determine the formation of purple product.

After screening of 4136 clones, 234 clones were found to show an OD_{550}* (after deduction with the OD vale from the wild-type enzyme) of >0.3 (Table 4.2). An example with P450pyr mutant I83M showing OD_{550}* of 0.462 was given in Figure 4.7A. The 234 promising clones were selected for individual biotransformation of *n*-butanol on a 10 mL scale. The product after 4 h reaction was extracted and then analysed by GC. P450pyr mutant I83M was found to give the highest productivity with excellent regioselectivity for the terminal hydroxylation. 0.25 mM 1,4-butanediol were produced (Figure 4.7B), and no hydroxylation product at other position could be found in GC chromatogram.

Table 4.2 Directed evolution of P450pyr hydroxylase for terminal hydroxylation of *n*-butanol to 1,4-butanediol

Round	No. of sites saturated	No. of clones screened	No. of promising clones [a]	Best mutant [b]	Product Conc. (mM) [c]
Wild-type	Nil	Nil	Nil	Nil	0
1	22	4136	234	I83M	0.25
2	21	3948	122	I83M/I82T	0.40

[a] Surrogate substrate based colorimetric HTS assay was used. Clones gave product OD_{550}* of > 0.3 in the 1st round and of > 0.5 in the 2nd round were identified as promising clones. [b] All promising clones were tested for the biohydroxylation of *n*-butanol, and the formation of product was determined by GC analysis. [c] Biotransformation of 5 mM *n*-butanol was performed with the resting cells of *E. coli* (P450pyr or best mutant) (24 g cdw L^{-1}) on 10 mL scale for 4 h. The product concentration was determined by GC analysis.

The P450pyr mutant I83M was used as a template for the second round of evolution by ISM at the remaining 21 selected amino acid residues. In this round, only the 122 clones which gave OD_{550}* of > 0.5 by using colorimetric HTS assay were selected for further investigation on the hydroxylation of *n*-butanol with GC analysis. P450pyr Mutant I83M/I82T was identified to produce 0.4

mM of 1,4-butanediol in 4 h biotransformation of *n*-butanol. As showed in Figure 4.7B, no by-product was formed, suggesting the excellent terminal hydroxylation selectivity.

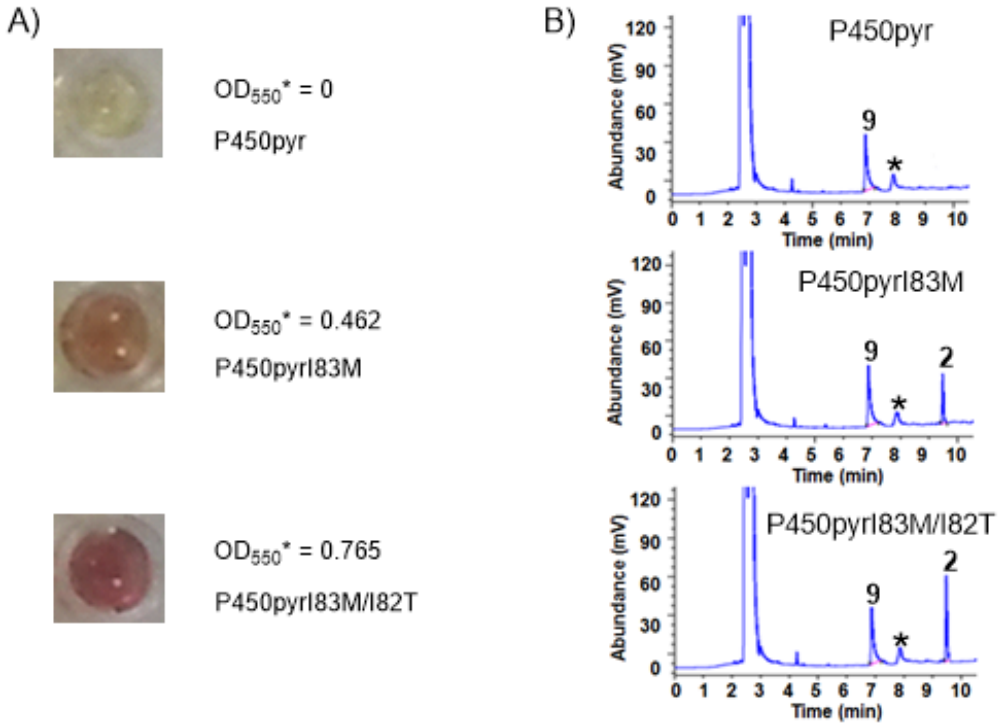


Figure 4.7 Photos and chromatogram of biohydroxylation of 2-methoxyethanol and *n*-butanol by recombinant *E. coli* P450pyr and its variants. A) Photos and OD_{550}* values of the purple product from the biohydroxylation of surrogate substrate with *E. coli* (P450pyr) and its variants, respectively. OD_{550}* = OD_{550} (P450pyr mutant) - OD_{550} (P450pyr). B) GC chromatograms of the product from biohydroxylation of *n*-butanol with *E. coli* (P450pyr), and its variants, respectively. **9**: 2-hexanol (internal standard); **2**: 1,4-butanediol; *****:unrelated cell metabolite that existed also in the control samples where no transformation of *n*-butanol was observed.

4.3.7 Comparison of Product Concentrations Determined by Colorimetric Assay and GC Analysis

The comparison of using the P450pyr and its mutant I83M and I83M/I82T for the hydroxylation of surrogate substrate 2-methoxyethanol followed by colorimetric HTS assay and for the hydroxylation of *n*-butanol followed by GC analysis, respectively, was performed, and similar product concentration was

detected as shown in Figure 4.8. This suggested the surrogate substrate-based colorimetric HTS assay was reliable for the real screening.

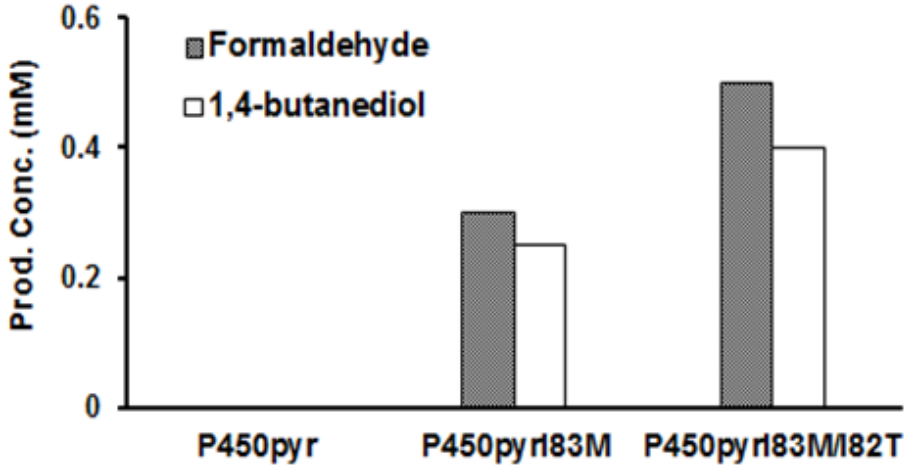


Figure 4.8 Comparison of product concentrations determined by colorimetric HTS assay as and GC analysis.

4.3.8 MS Analysis of the Derivative of 1,4-butanediol Made by BSTFA-TMCS Method

The biotransformation product from biohydroxylation of *n*-butanol with *E. coli* (P450pyrI83M), and *E. coli* (P450pyrI83M/I82T) were derivatised with BSTFA-TMCS to give 1,4-butanediol bis(trimethylsilyl)ether, and the MS analysis confirmed its structure. The retention time for the standard derivatised (1,4-butanediol bis(trimethylsilyl)ether) is 6.25 min. As shown in Figure 4.9, the MS peak of the derivatised biohydroxylation product peak at 6.25 min was as same as standard derivatised 1,4-butanediol, which further confirmed that the biotransformation product is 1,4-butanediol.

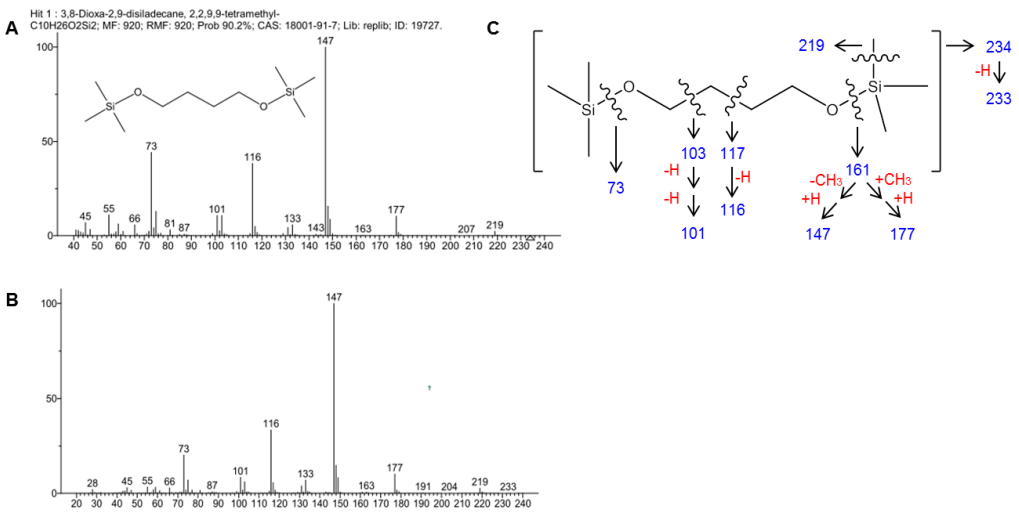


Figure 4.9 MS analysis of peak at 6.25 min in the GC chromatogram for BSTFA-TMCS derivative of 1,4-butanediol. (A) derivative from standard 1,4-butanediol; (B) derivative from biotransformation product with induced cell of *E. coli* (P450pyrI83M). (C) Proposed fragmentation of the derivatised 1,4-butanediol.

4.3.9 Molecular Modelling of Substrates Docking in P450pyr and its Mutants

To explore the molecular basis of the observed hydroxylation reactions, *in silico* modelling of docking *n*-butanol in the X-ray structure of P450pyr and the structure models of P450pyr mutants I83M and I83M/I82T were performed, respectively. As shown in *n*-butanol-P450pyr binding conformation (Figure 4.10A and 4.10D), the hydrophobic end of the substrate is located between the two hydrophobic residues Ile102 and Leu302 and the hydroxyl group of the substrate is oriented towards the hydrophilic Thr 259. In this conformation, the hydroxyl group is very close (2.0 Å) to the heme oxygen atom (heme-O), thus being inadequate for any hydroxylation activity.

For the single mutant (Figure 4.10B and 4.10E), I83M introduced a longer residue Met83 which influences the nearby amino acid residues to give a shorter

distance between the two hydrophobic residues Ile102 and Leu302 (5.7 Å in I83M vs 8.5 Å in P450pyr). As a result, the hydrophobic end of the substrate cannot be located between Ile102 and Leu302 and thus took another orientation. The hydroxyl group of the substrate was still oriented towards the hydrophilic Thr 259 but far away from the heme-O. In this conformation, the distance between heme-O and C(2), C(3), and C(4) was 4.4 Å, 4.0 Å, and 3.7 Å, respectively, thus giving rise to the regioselective terminal hydroxylation of *n*-butanol.

For the double mutant (Figure 4.10C and 4.10F), I82T introduced a hydrophilic Thr82 residue, which removed the hydrophobic interaction between Ile82 and Ile102 and enabled the nearby hydrophilic residue Ser75 encroached upon the binding pocket. The hydroxyl group of the substrate was oriented towards the hydrophilic Ser75 and Thr82, and the hydrophobic end of the substrate was located close to heme-O, with a distance of 3.4 Å. This distance was shorter than the one for the single mutant I83M (3.7 Å), suggesting higher terminal hydroxylation activity for the double mutant I83M/I82T. In the substrate binding pose of the double mutant, the distance between heme-O and C(2) and C(3) was 4.8 Å and 5.4 Å, respectively, which explained the high regioselectivity for the terminal hydroxylation.

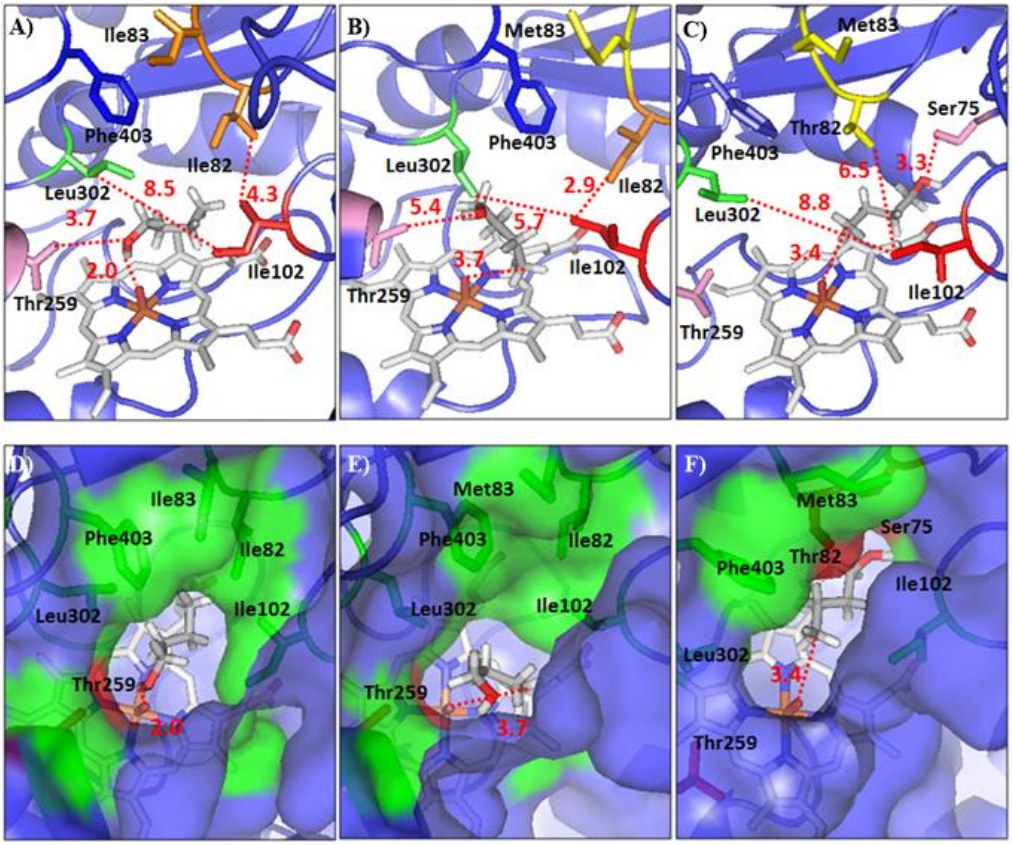


Figure 4.10 Enzyme-substrate binding pose for *n*-butanol in P450pyr or mutant. A-C) Enzyme structures in cartoon ribbon style with mutated residues in yellow. A) P450pyr, B) P450pyr I83M, C) P450pyr I83M/I82T. D-F) Enzyme structures in surface style with hydrophobic residues in green and hydrophilic residues in red. D) P450pyr, E) P450pyr I83M, F) P450pyr I83M/I82T. Distances are denoted by a red dashed line.

4.4 Conclusion

P450pyr mutants (I83M and I83M/I82T) were successfully engineered by directed evolution to enable highly regioselective terminal hydroxylation of *n*-butanol to produce 1,4-butanediol. To our knowledge, this is the first report of this hydroxylation reaction by either chemical or enzymatic method. Although the current enzyme mutant is not active enough for practical synthesis of 1,4-butanediol from the bio-based *n*-butanol, our research provides with a solid basis for further engineering the enzyme to improve the activity. It is also an unique example of evolving a hydroxylase to switch the substrate acceptance

from hydrophobic to a hydrophilic compound. The surrogate substrate-based colorimetric HTS assay was proven to be useful and reliable for the screening of enzymes for the terminal hydroxylation of alcohol. Molecular docking provided some insight into the substrate binding pose, the reshaping of binding pocket by the mutations, and the role of key mutations for creating terminal hydroxylation activity and regioselectivity. This information could be very useful for further engineering more active mutants for terminal hydroxylation of *n*-butanol and potentially useful for switching substrate acceptance of a hydroxylase from hydrophobic to hydrophilic compounds.

Chapter 5. Benzylic Hydroxylation of Fluoro- and Other Halo-Substituted Toluenes with Engineered P450pyr Monooxygenase

5.1 Introduction

Fluoro- and other halo-substituted benzyl alcohols are important and widely used intermediates for the productions of pharmaceuticals and fine chemicals, of pharmaceuticals, fine chemicals, agrochemicals.[300-302] For instances: 2-fluorobenzyl alcohol is an intermediate for synthesizing novel acid pump antagonists,[303] and glucokinase activators.[304] 3-fluorobenzyl alcohol is useful for preparation of boron-containing antifungal agent for the potential treatment of onychomycosis[305] and neuronal nitric oxide synthase inhibitors.[306] 4-fluorobenzyl alcohol is the intermediate for C-2-symmetric diol-based HIV-1 protease inhibitors[307] and 2,4-diaminoquinazoline derivatives as SMN2 promoter activators for the potential treatment of spinal muscular atrophy.[308] Currently, this group of compounds is mainly produced from pre-functionalised toluenes, which involving not only multi-steps of pre-function and also the generation of toxic and corrosive by-products such as HCl.[309-311] Direct C-H bond activation of fluoro- and other halo-substituted toluenes presents the simplest route to produce their corresponding alcohols. Although tremendous progresses of C-H bond activation have been reported using transition-metal catalysts,[312,313] to date only one catalyst cerium (IV) triflate was reported for the hydroxylation of benzylic aromatics.[314] Nevertheless, this catalyst only showed excellent chemo- and regioselectivity toward 4-fluorotoluene, to other substrates (including 2-fluorotoluene, 2-chlorotoluene, 3-chlorotoluene),

however, the catalyst's chemo- and regioselectivity and activity were relatively poor. Besides, this reaction also suffered from long reaction time, limited substrate scope, high catalyst loading, as well as requiring expensive and hazardous reactants.[314]

On the other hand, nature offers an alternative and attractive solution for this reaction in the use of a monooxygenase, as a green and efficient catalyst and molecular oxygen from atmosphere as a cheap and green oxidant.[121,315,316] However, only very few monooxygenases have been reported for this type of reaction, and their catalytic performances（substrate acceptance, selectivity and activity) are unsatisfactory for industry applications. Xylene oxygenase from *Pseudomonas putida* is the first enzyme reported for this type of reaction, although a good activity was showed, its substrate scope was still limited, no activity toward 2-position substituted chloro- and bromotoluenes and no activity of any fluorotoluenes.[109] *p*-cymene monooxygenase from *Pseudomonas putida* F1, however, was able to hydroxylate 4-fluorotoluene, 3-chlorotoluene and 4-chlorotoluene, but the activity and conversion were extremely low.[108] Recently, a double mutant of P450 BM3 was reported for the benzylic hydroxylation of some toluene derivatives, but with no activity toward bromo- and chlorotoluene, it showed activity toward 3-fluorotoluene, however no reliable measurement of the produced 3-fluorobenzyl alcohol.[182]

We are interesting in developing new enzymes for the hydroxylation of fluoro- and other halo-substituted toluenes with high chemo- and regioselectivity to prepare useful and valuable alcohols. In this project, we explored the synthetic application of engineered P450pyr monooxygenases generated in chapter 3 and

4 for this challenge reaction. All the available P450pyr variants generated by directed evolution as well as the wild-type P450pyr were used for the enzymatic hydroxylation of 4-fluorotoluene first. The mutant which showed the best selectivity and highest activity towards 4-fluorotoluene would then be use for the hydroxylation of other fluoro- and halo-substituted toluenes. Moreover, the best mutant was also explored for the hydroxylation of multiple fluoro-substituted toluenes since the hydroxylation of strong electron-withdrawing groups containing substrate is still problematic in classic chemistry. In addition, the relationship between mutated amino acids and high benzylic hydroxylation activity were also investigated based on the simulation model and docking structure. The obtained information and knowledge is useful for further engineering of P450pyr for other hydroxylations and oxidations.

5.2 Experimental Section

5.2.1 Chemicals

All of the following chemicals were purchased from Sigma-Aldrich and used without further purification: δ-aminolevulinic acid hydrochloride (ALA) (≥ 97 %), Isopropyl β-D-1-thiogalactopyranoside (IPTG), D-Glucose (≥ 99.5%), Sodium phosphate dibasic (≥ 99 %), and Sodium phosphate monobasic (≥ 99 %), 2-fluorotoluene (≥ 99 %), 3-fluorotoluene (99 %), 4-fluorotoluene (97 %), 2-chlorotoluene (99 %), 3-chlorotoluene (98 %), 4-chlorotoluene (98 %), 2-bromotoluene (99 %), 3-bromotoluene (98 %), 4-bromotoluene (98 %), 2-fluorobenzyl alcohol (98 %), 3-fluorobenzyl alcohol (98 %), 4-fluorobenzyl alcohol (97 %), 2-chlorobenzyl alcohol (99 %), 3-chlorobenzyl alcohol (98 %),

4-chlorobenzyl alcohol (99 %), 2-bromobenzyl alcohol (99 %), 3-bromobenzyl alcohol (99 %), 4-bromobenzyl alcohol (99 %). 3,4-difluorotoluene (> 98.0 %), 2,3,4,5,6-pentafluorotoluene (> 98.0 %), 3,4-difluorobenzyl alcohol (≥ 98 %) and pentafluorobenzyl alcohol (> 96.0 %) were purchased from Tokyo Chemical Industry. LB Broth, Bacto Yeast Extract, Bacto Tryptone were purchased from Biomed Diagnostics.

5.2.2 Strains and Biochemicals

Escherichia coli BL21(*DE3*) was used as host for enzyme expression. The recombinant *E. coli* (P450pyr) or *E. coli* (mutated P450pyr) with dual plasmids, pETDuet containing ferredoxin (Fdx) and ferredoxin reductase (FdR) genes, and pRSFDuet containing P450pyr or mutated P450pyr gene, were generated from directed evolution as described in chapter 3 and 4. Oligonucleotides were synthesised by AIT biotech, Singapore. LB Broth, Bacto Yeast Extract, and Bacto Tryptone were purchased from Biomed Diagnostics. Antibiotic ampicillin and kanamycin were purchased from Sigma-Aldrich. QIAquick Gel Extraction Kit (Qiagen) and QIAprep spin plasmid miniprep Kit were purchased from Qiagen.

5.2.3 Analytical methods

The concentrations of all the compounds (halo-substituted toluenes and their corresponding alcohols) were determined using a Shimadzu prominence HPLC system (reverse phase) with an Agilent Poroshell 120 EC-C18 column (150 × 4.6 mm, 2.7 μm) and UV detection at 210 nm. HPLC analysis condition: 40% water with 60% acetonitrile. Flow rate: 0.4 mL/min. Retention times of

different compounds were listed as follow: 12.4 min for 2-fluorotoluene and 5.8 min for its alcohol; 12.3 min for 3-fluorotoluene and 5.8 min for its alcohol; 12.2 min for 4-fluorotoluene and 5.7 min for its alcohol; 16.5 min for 2-chlorotoluene and 6.5 min for its alcohol; 16.5 min for 3-chlorotoluene and 6.5 min for its alcohol; 16.3 min for 3-chlorotoluene and 6.4 min for its alcohol; 18.3 min for 2-bromotoluene and 6.8 min for its alcohol; 18.3 min for 3-bromotoluene and 6.8 min for its alcohol; 18.2 min for 4-bromotoluene and 6.7 min for its alcohol; 12.6 min for 3,4-difluorotoluene and 6.0 min for its alcohol; 16.3 min for 2,3,4,5,6-pentafluorotoluene and 7.0 min for its alcohol.

5.2.4 Biohydroxylation of 4-fluorotoluene with Resting Cells of Recombinant *E. coli* (P450pyr) and *E. coli* (mutated P450pyr)

The recombinant *E. coli* BL21 (DE3) strains expressing P450pyr [*E. coli* (P450pyr)] or mutated P450pyr [*E. coli* (mutated P450pyr)] were inoculated respectively into 5 mL LB medium containing 50 mg/L of kanamycin and 100 mg/L ampicillin. 2 mL overnight inoculum was transferred to 50 mL TB medium containing 50 mg/L of kanamycin and 100 mg/L ampicillin in a 250 mL shaking flask. Cells were grown at 37 ℃ and 250 rpm until OD_{600} research 0.6~0.8 and then induced by adding IPTG and ALA to a final concentration of 0.5 mM. Cells were then grown at 22 ℃ for another 12 h and then harvested by centrifugation at 5000 g for 5 min. The cells were re-suspended to a density of 8 g cdw L^{-1} in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose and 5 mM of 4-fluorotoluene. Biotransformation was performed at 300 rpm and 30 ℃ for 4 h, and the formation of 4-fluorobenzyl alcohol was analysed by reverse HPLC.

5.2.5 Cell Growth, Hydroxylation Activity and Protein Expression of *E. coli* (P450pyr3M)

The cell growth and specific hydroxylation activities of the P450pyr3M at different time points were examined by using 4-fluorotoluene as substrate. The cell was grown in TB medium, and the expression of P450pyr monooxygenase was induced by adding 500 μM of IPTG and 500 μM of ALA. The induction was carried out at 22 ^{o}C for additional 8 h. Samples were taken at different time points for the determination of optical density at 600 nm. At the same time, induced cells were harvested and re-suspended for the hydroxylation of 4-fluorotoluene for 30 min to examine the specific activity. A SDS-PAGE (12% resolving gel and 4% stacking gel) was applied to check the purity of the proteins.

5.2.6 Biohydroxylation of Halo-toluenes with Resting Cells of Recombinant *E. coli* (P450pyr3M)

The recombinant *E. coli* (P450pyr3M) was inoculated, grown, expressed and harvested as described above. For specific activity determination, the cells were resuspended to a density of 4 g cdw L^{-1} in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose, 10% DMSO and 10 mM of different halo-toluenes and reaction was performed at 300 rpm and 30 ℃ for 30 min; while for product concentration and yield detection, the cells were re-suspended to a density of 8 g cdw L^{-1} in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose, 10% DMSO and 5 mM of different halo-toluenes and reaction was performed at 300 rpm and 30 ℃ for 4 h.

5.2.7 Biohydroxylation of 3-bromotoluene with resting cells of *E. coli* (P450pyr3M)

The recombinant *E. coli* (P450pyr3M) was inoculated, grown, expressed and harvested as described above. The cell were re-suspended to a density of 8 g cdw L^{-1} in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose, 10% DMSO and 25 mM of different 3-bromotoluene. Biotransformation was performed at 300 rpm and 30 ℃ for 12 h.

5.2.8 Molecular Modelling of Halo-Toluenes Docking in P450pyr and P450pyr3M

The molecular dynamics and docking simulation were performed as described in Chapter 3 and 4. Ramachandran plots for wild-type P450pyr and P450pyr3M models were generated before and after MD simulations. The main cluster centroids were docked with halo-toluene substrates using Autodock VINA,[293,294] with the heme set as ferryl-oxo-heme complex state known as Cpd I. Docking results were evaluated *via* the criteria of binding energy scores with docked-posture clustering *via* a 1 Å non-fitted RMSD cutoff. Catalytically relevant active binding postures of halo-toluene substrates were taken as those where substrate methyl-C was within a distance of < 6 Å from the heme-O atom, and also consistent with optimized geometric criteria for substrate-heme binding in P450 enzymes.[295,296]

5.3 Results and discussion

5.3.1 Biohydroxylation of 4-fluorotoluene with Resting Cells of Recombinant *E. coli* (P450pyr) and *E. coli* (mutated P450pyr)

Totally 29 engineered P450pyr monooxygenases were generated in Chapter 3 and Chapter 4 (Table 5.1). The recombinant *E. coli* BL21 (DE3) strains expressing P450pyr or mutated P450pyr were all used for the biohydroxylation of 4-fluorotoluene. So far only one enzyme, *p*-cymene monooxygenase from *Pseudomonas putida* F1 was reported capable of direct hydroxylate 4-fluorotoluene to 4-fluorobenzyl alcohol, however the activity and conversion was very low, only 3.5 μM of 4-fluorobenzyl alcohol was produced from 5 mM 4-fluorotoluene after 3 hours.

Table 5.1 P450pyr mutant library: mutant numbers and their corresponding mutations

No.	Mutant	No.	Mutant
1	N100S	16	N100S/F403M
2	N100S/F403I	17	N100S/M305Q
3	N100S/F403I/T186I	18	N100S/F403I/M305Q
4	N100S/F403I/T186I/L302V	19	N100S/F403I/T186I/M305Q
5	N100S/F403I/T186I/L302V/I83F	20	N100S/F403I/T186I/L302V/M305Q
6	N100S/F403I/T186I/L302V/I83F/A77Q	21	M305Q/T186I
7	N100S/F403I/T186I/L302V/I83F/L251V	22	N100S/T186I/M305Q
8	N100S/F403I/T186I/L302V/I83F/I102P	23	N100S/F403I/L302V
9	I82S	24	N100S/F403I/T186I/L302V/T259A
10	T185K	25	N100S/F403I/T186I/L302V/T259A/305Q
11	T259A	26	N100S/F403I/T186I/D183E
12	I83L	27	N100S/F403I/T185Y
13	T185V	28	I83M
14	V404A	29	I83M/I82T
15	M305Q		

Most of the P450pyr variants as well as the wild-type P450pyr were found to be capable of direct converting 4-fluorotoluene to its alcohol. Among all these investigated P450s, more than half of the P450pyr variants showed a very clean reaction and higher product concentration compare with P450pyr. A triple mutant N100S/F403I/M305Q (named as: P450pyr3M, Figure 5.1) was found to show the highest 4-fluorobenzyl alcohol concentration: 2.9 mM, which around 10 fold compared to the wild-type P450pyr. The rate of production in 4 h is about 1.51 U/g, which is more than 600 times higher than *p*-cymene in 3 h. Therefore, this variant, P450pyr3M, was selected a promising catalyst for other benzylic hydroxylation.

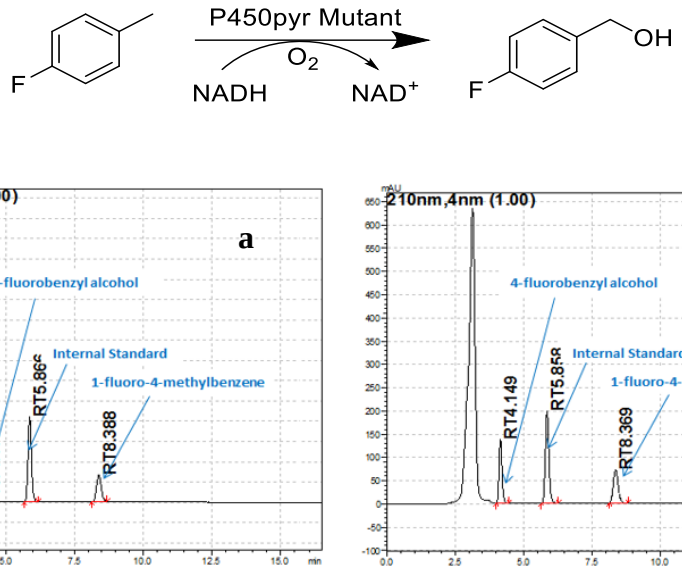


Figure 5.1 Hydroxylation of 4-fluorotoluene by recombinant *E. coli* (a) P450pyr, (b) P450pyr mutant N100S/F403I/M305Q (P450pyr3M).

5.3.2 Cell Growth, Hydroxylation Activity and Protein Expression of *E. coli* (P450pyr3M)

The cell growth and specific hydroxylation activities of the P450pyr3M at different time points were examined by using 4-fluorotoluene as substrate. As

shown in Figure 5.2A, a cell density of 4.0 was researched after 10 h. Cells taken at different time points showed different hydroxylation activity towards 4-fluorotoluene and the highest specific activity was observed for cells grown after 7 h, in the middle of exponential grow phase. Thus, harvested cells at this time point were used in the following biohydroxylation reactions. The SDS–PAGE of cell-free extracts of *E. coli* P450pyr3M was prepared. As showed in Figure 5.2B, the P450pyr3M enzyme was clearly expressed in induced cells compare with *non*-induced cells, its expressing level was even higher than wild-type P450pyr, thus enabling efficient and active biohydroxylation.

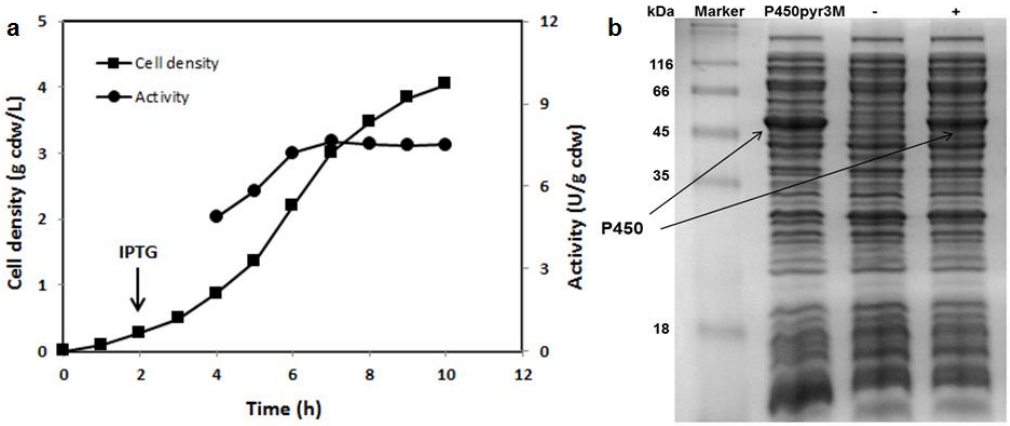


Figure 5.2 Cell growth, specific activity and protein expression of *E. coli* (P450pyr3M). (a) Growth curve and specific activity for the hydroxylation of 4-fluorotoluene of *E. coli* (P450pyr3M). (b) SDS–PAGE of cell-free extracts of *E. coli* recombinant strains. 1: induced *E. coli* (P450pyr3M); 2: *non*-induced *E. coli* (P450pyr3M) as a negative control, and 3: induced *E. coli* (P450pyr) as a positive control.

5.3.3 Biohydroxylation of Other Fluoro-Substituted Toluenes and Multi Fluoro-Substituted Toluenes with Resting Cells of Recombinant *E. coli* (P450pyr3M)

After examining the cell growth and specific hydroxylation activities, *E.coli* (P450pyr3M) was investigated for biohydroxylation of 2- and 3-fluorotoluenes.

As shown in Figure 5.3, P450pyr3M exhibited very high hydroxylation activity (8.3 and 10.9 U g cdw^{-1}) toward 2- and 3-fluorotoluenes. No by-product was formed in these two reactions, suggesting the excellent chemo- and benzylic hydroxylation selectivity.

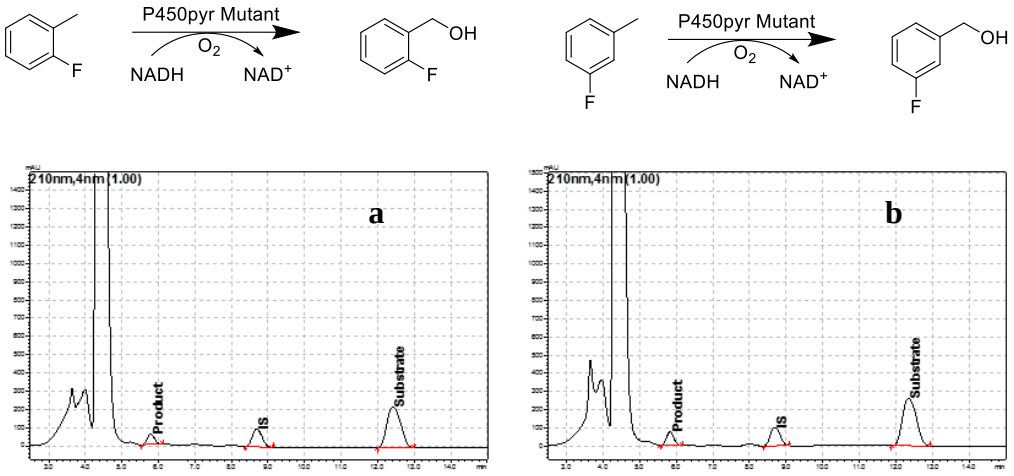


Figure 5.3 Hydroxylation of 2-fluorotoluene (a) and 3-fluorotoluene (b) by recombinant *E. coli* (P450pyr3M). Determined from 30 min biotransformation of 10 mM substrate with 4 g/cdw of cells at 30 ^{o}C and 300 rpm.

Encouraged by the success in the biohydroxylation of single fluoro-substituted toluenes, P450pyr3M was used for the hydroxylation of more complicated substrates, multi fluoro-substituted toluenes. 3,4-difluorotoluene and 2,3,4,5,6-pentafluorotoluene were selected since their hydroxylate products were important pharmaceutical intermediates, for example: 3,4-difluorobenzyl alcohol could be used for synthesizing DNA-dependent protein kinase inhibitors,[317] while 2,3,4,5,6-pentafluorobenzyl alcohol is the key intermediate for the preparation of inhibitors of the kinase domain of vascular endothelial growth factor receptor-2.[318] Although useful and important, so far no chemical- and bio-catalyst was reported capable of directed hydroxylation of these two compounds. As showed in Figure 5.4, interestingly P450pyr3M showed

comparative hydroxylation activity and conversion toward these two multi fluoro-substituted toluenes while wild-type P450pyr showed lower even lost its hydroxylation ability. These results confirmed that compare with wild-type P450pyr and other P450s, this P450pyr3M could accept the substrates containing bigger and stronger electron-withdrawing groups. To the best of our knowledge, P450pyr3M is the first enzyme for this type of hydroxylation to produce multi fluoro-substituted benzylic alcohols.

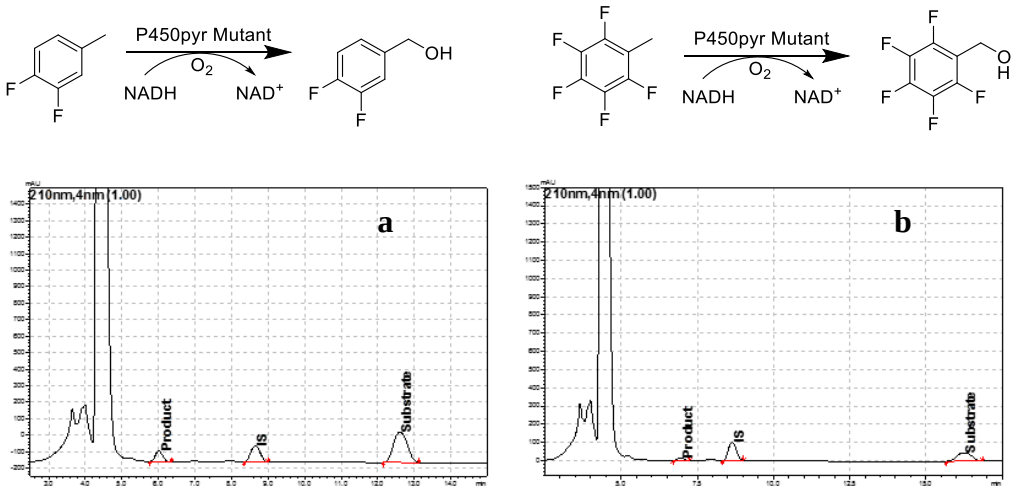


Figure 5.4 Hydroxylation of 3,4-difluorotoluene (a) and 2,3,4,5,6-pentafluorotoluene (b) by recombinant *E. coli* (P450pyr3M). Determined from 30 min biotransformation of 10 mM substrate with 4 g/cdw of cells at 30 ºC and 300 rpm.

5.3.4 Biohydroxylation of Chloro- and Bromo-Substituted Toluenes with Resting Cells of Recombinant *E. coli* (P450pyr3M)

Besides the fluoro-substituted toluenes, P450pyr3M was also found to catalyse the benzylic hydroxylation of *ortho*-, *meta*- and *para*-substituted chloro- and bromotoluenes at the methyl group giving the corresponding alcohols as the only product (Figure 5.5). Chloro- and bromo-substituted benzyl alcohols are two groups of useful intermediates. For example, 2-chlorobenzyl alcohol is an intermediate for synthesizing sodium iodide symporter inhibitors,[319] 3-

chlorobenzyl alcohol is used for producing coumarin monoamine oxidase B inhibitors[320], 4-chlorobenzyl alcohol is an intermediate for advanced glycation end products (RAGE) Inhibitors.[321] 2-bromobenzyl alcohol is used for producing macrocyclic antagonists to the human motilin receptor,[322] 3-bromobenzyl alcohol is an intermediate for synthesizing metabotropic glutamate receptor,[323] and 4-bromobenzyl alcohol could be used for making vanilloid 1 receptor antagonists for the treatment of pain.[324]

As showed in Figure 5.5, P450pyr3M also showed very high activity toward these chloro- and bromotoluenes compare to any other reported catalysts. For instance, P450pyr3M showed 13.0 U g cdw/L specific activity for 3-chlorotoluene that is around 4.3 times higher than xylene oxygenase,[109] and the rate of production of catalysed by P450pyr3M is about 1.8 U g cdw/L for 4 h which is 850 times higher than that (2.10 nanomoles product/minute) achieved by *p*-cymene for 3 h.[108]

Interestingly, compared to the toluenes with a substitution at the 2- and 4-position (*ortho*-, and *para*-position), the halo-toluenes with a substitution at the 3-position (*meta*-position) were hydroxylated faster (higher specific activity and productivity). This preference is significantly different from other oxygenases, such as *p*-cymene monooxygenase and xylene oxygenase that both showed 4-position favorite and no activity on 2-position substitution, indicating the unique substrate specificity and special synthetic application of P450pyr3M.

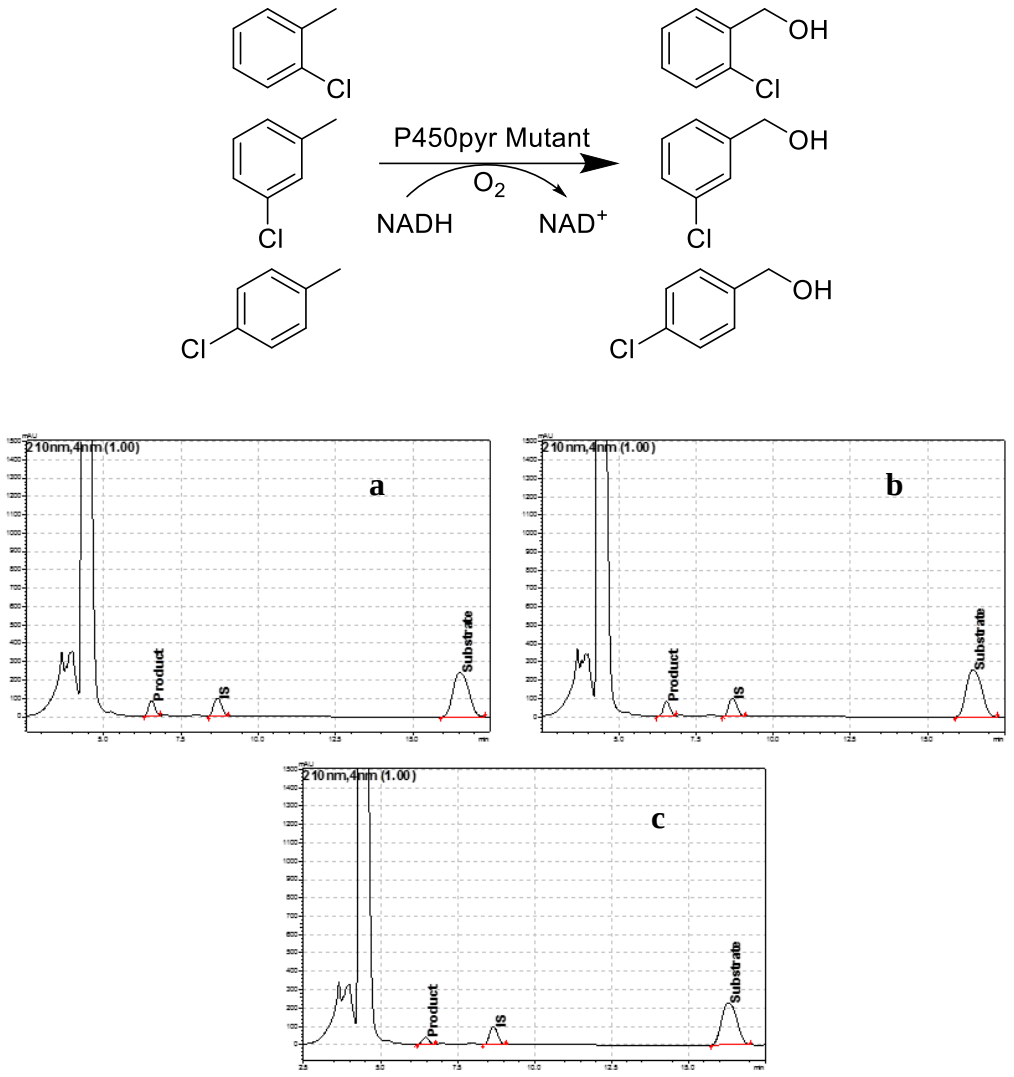


Figure 5.5 Hydroxylation of 2-chlorotoluene (a) 3-chlorotoluene (b) and 4-chlorotoluene (c) by recombinant *E. coli* (P450pyr3M). Determined from 30 min biotransformation of 10 mM substrate with 4 g/cdw of cells at 30 ºC and 300 rpm.

Table 5.2 Hydroxylation of single fluoro-, chloro-, bromo- and multi fluoro-substituted toluenes with resting cells of *E.coli* P450pyr3M and *E.coli* wild-type P450pyr

Substrate	Product	Enzyme	Activity [U (g cdw)$^{-1}$] [a]	Prod. conc. (mM) [b]	Yield.(%)[b]
2-fluorotoluene	2-fluorobenzyl alcohol	P450pyr3M	8.3	3.1	62
		P450pyr	ND[c]	0.5	10
3-fluorotoluene	3-fluorobenzyl alcohol	P450pyr3M	10.9	3.3	66
		P450pyr	ND[c]	0.7	14
4-fluorotoluene	4-fluorobenzyl alcohol	P450pyr3M	7.6	2.9	58
		P450pyr	ND[c]	0.3	6
2-chlorotoluene	2-chlorobenzyl alcohol	P450pyr3M	10.6	3.3	66
		P450pyr	ND[c]	1.1	22
3-chlorotoluene	3-chlorobenzyl alcohol	P450pyr3M	13.0	3.5	70
		P450pyr	ND[c]	1.2	24
4-chlorotoluene	4-chlorobenzyl alcohol	P450pyr3M	8.1	3.1	62
		P450pyr	ND[c]	1.0	20
2-bromotoluene	2-bromobenzyl alcohol	P450pyr3M	10.8	3.3	66
		P450pyr	ND[c]	1.1	22
3-bromotoluene	3-bromobenzyl alcohol	P450pyr3M	13.4	3.7	74
		P450pyr	ND[c]	1.3	26
4-bromotoluene	4-bromobenzyl alcohol	P450pyr3M	8.2	3.1	62
		P450pyr	ND[c]	1.0	20
3,4-difluorotoluene	3,4-difluorobenzyl alcohol	P450pyr3M	5.1	1.8	36
		P450pyr	ND[c]	0.3	6
2,3,4,5,6-pentafluorotoluene	2,3,4,5,6-pentafluorobenzyl alcohol	P450pyr3M	2.0	1.1	22
		P450pyr	ND[c]	0.1	2

[a] activity is the specific activity determined for the first 30 min of the biotransformation, reaction were conducted with 10 mM substrate at 4 g cdw/L of *E.coli* expressing the P450pyr3M in potassium phosphate buffer (100 mM; pH 8) containing glucose (2%, w/v) at 30 °C and 250 rpm for 30 min. [b] reaction were conducted with 5 mM substrate 10% DMSO at 8 g cdw/L of *E.coli* expressing the P450pyr3M at 30 °C and 250rpm for 4 h. Concentration and yield were determined by HPLC analysis. [C] ND: not determined.

5.3.5 Biohydroxylation of 3-bromotoluene with resting cells of *E. coli* (P450pyr3M)

To further improve the product concentration and test its stability, P450pyr3M was examined for hydroxylation at a higher substrate concentration, and 3-bromotoluene was chosen as a model substrate. 25 mM of substrate was added

at the beginning, and product was quickly formed and reached 11.6 mM at 2 h and 16.5 mM at 4 h, finally 17.8 mM of product was produced after 12 h (Figure 5.6). This is the highest concentration we have ever achieved in a simple flask condition by using *E. coli* (P450pyr) cells. According to our previous experiment results, both substrate and product inhibition are serious under this condition, which will prevent further improve the productivity.

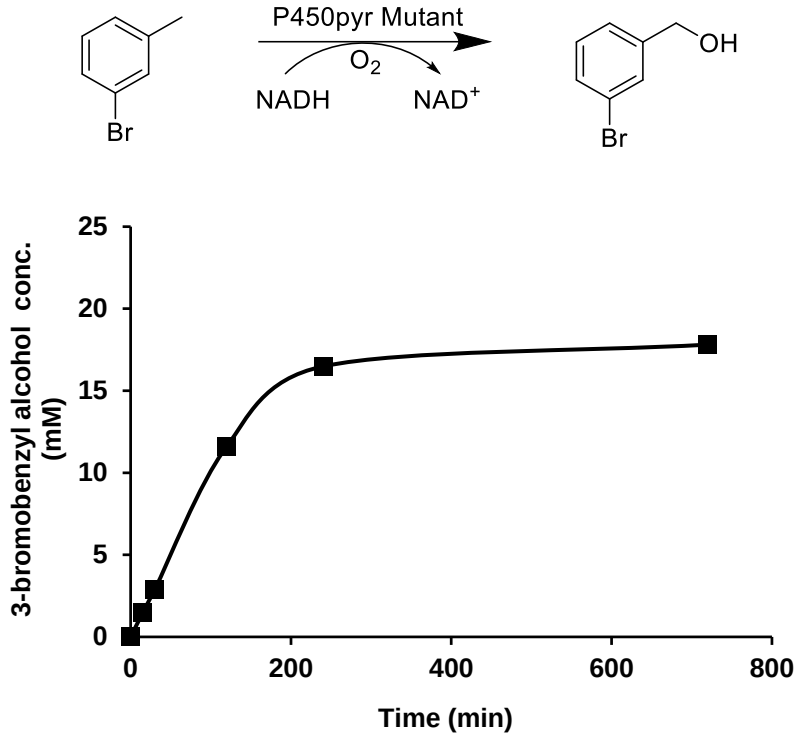


Figure 5.6 Time course of biohydroxylation of 25 mM 3-bromotoluene by resting cells of *E. coli* (P450pyr3M) (8.0 g cdw/L) in potassium phosphate buffer (100 mM, pH 8) containing 10% DMSO and glucose (2%, w/v) at 30 ºC.

5.3.6 Biohydroxylation of 3-fluorotoluene with resting cells of *E. coli* (P450pyr3M) and *E.coli* (P450pyr3M-GDH)

Follow the same process described in chapter 3, *E. coli* (P450pyr3M-GDH) co-expressing P450pyr3M monooxygenase and a glucose dehydrogenase (GDH) was engineered. The cells were grown and used for biohydroxylation in the

same procedures described for *E. coli* (P450pyr3M). As illustrated in Figure 5.7, the new constructed recombinant with cofactor recycling system significantly improved the initial biohydroxylation activity up to 3 fold and also improved the final production yield about 20%.

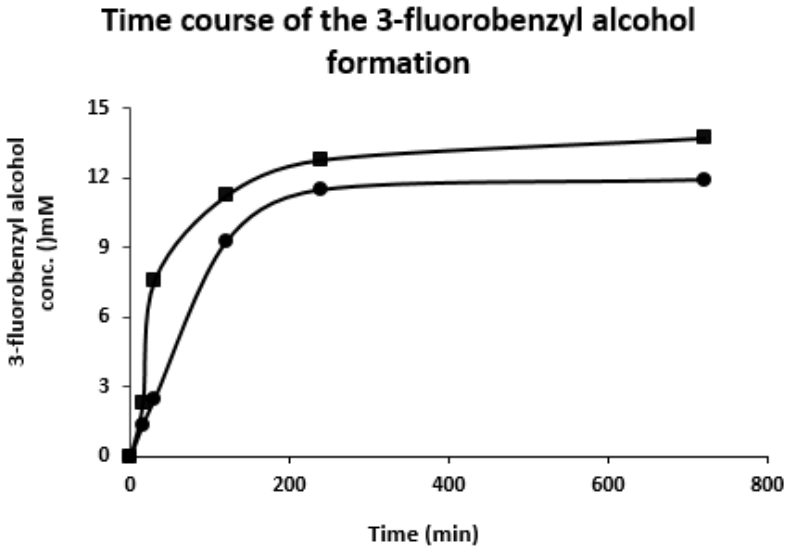


Figure 5.7 Time course of biohydroxylation of 25 mM 3-fluorotoluene by resting cells of *E. coli* BL21(DE3) (P450pyr3M with or without GDH) (8.0 g cdw/L) in potassium phosphate buffer (100 mM, pH 8) containing 10% DMSO and glucose (2%, w/v) at 30 ^{o}C. pRSFDuet-P450pyr3M + pETDuet-Fdx/FdR (•); pETDuet-P450pyr3M/FdR + pRSFDuet-GDH/Fdx (■).

5.3.7 Molecular Dynamics Simulations of P450pyr and P450pyr3M

To shed light on the possible molecular basis of the stereo-chemical outcome of the reactions, *in silico* modelling and docking simulations were performed. Ramachandran plots were generated respectively for P450pyr and P450pyr3M models before and after MD simulations (Figure 5.8).

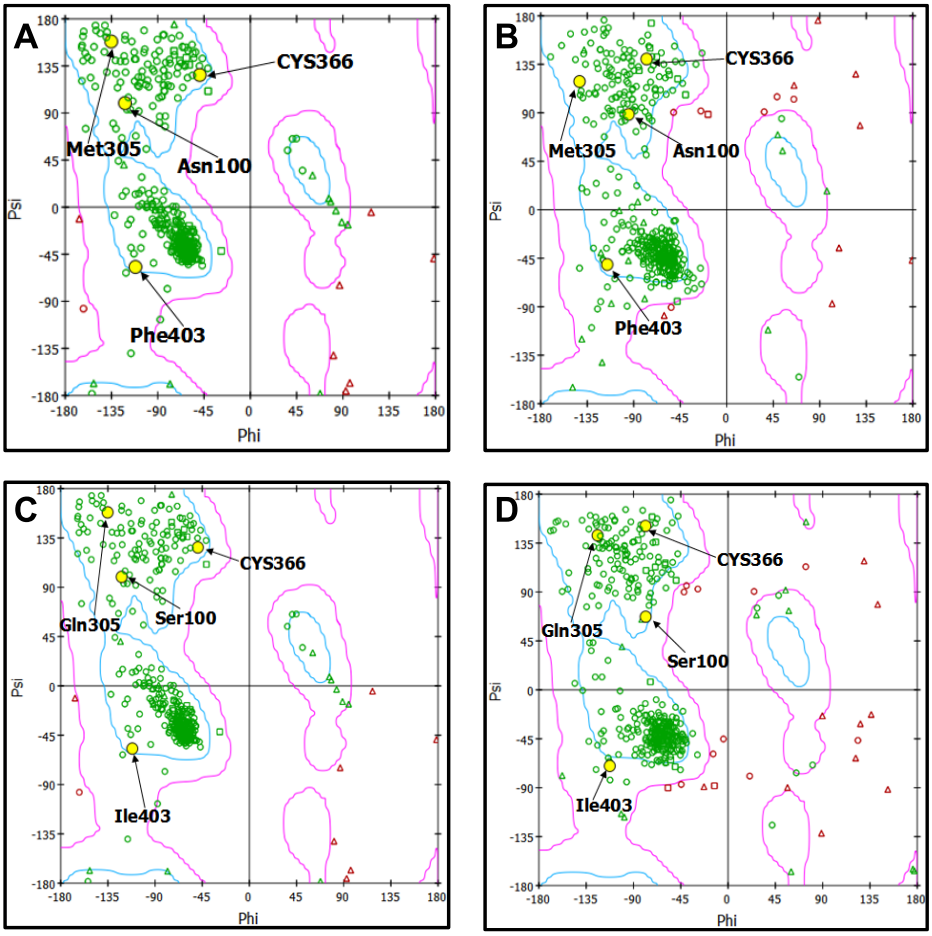


Figure 5.8 Ramachandran plots for (a) P450pyr model before MD simulation (b) P450pyr model after MD simulation (c) P450pyr3M before MD simulation (d) P450pyr3M after MD simulation. All plots have > 90 % of residues within allowed regions, inclusive of the main residues of interest: the mutation sites at residues 100, 305, 403; and CYS366 which is bonded to the catalytic heme. Flexible glycine residues are depicted as triangles.

As shown in Figure 5.8, more than 90% of residues to be within allowed regions. The well-structured residues included the main residues of interest: the mutation sites at residues 100, 305, 403; and CYS366, which was bonded to the catalytic heme.

As can be seen from the generated P450pyr and P450pyr3M models, the binding pocket trunk of the P450pyr was predominantly hydrophobic, and it can be suitably characterized by two core hydrophobic residues Ile102 (shown in red) and Leu302 (shown in light green). Figure 5.9A showed the spatial relation

between these two core hydrophobic residues (Ile102, Leu302) and three residues Asn100, Met305, Phe403 in P450pyr model, while Figure 5.9B showed the spatial relation between the same core and the three mutated residues Ser100, Gln305, Ile403 in the triple mutant P340pyr3M model.

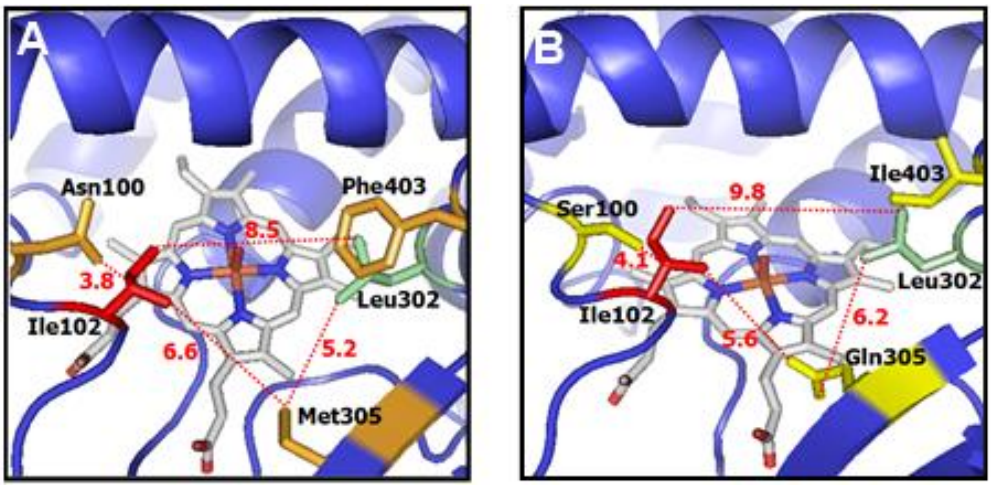


Figure 5.9 Positional relation between the two core hydrophobic residues Ile102, Leu302, and the three residues of interest (100, 305, 403) in (A) wild-type P450pyr (B) triple mutant P450pyr3M.

5.3.8 Molecular Modelling of Halo-Toluenes Docking in P450pyr and P450pyr3M

Figure 5.10A and 5.10B showed the docked posture of 4-fluorotoluene in P450pyr and in P450pyr3M, respectively. The triple mutation N100S/M305Q/ F403I enlarged the binding pocket, where F403I removed the relatively bulky phenyl ring to allow the predominantly hydrophobic 4-fluorotoluene to dock with reduced hydrophobic-hydrophilic constrain between the nearby shorter residue Ser100 that was due to the mutation N100S. This was in contrast to the more constrained binding posture in the wild-type enzyme where 4-fluorotoluene docked between the bulky phenyl ring of Phe403 and the longer

Asn100 residue in a position where the substrate methyl moiety was oriented further away from the activated heme-bound oxygen at 5.3 Å. Moreover, the mutation M305Q resulted in a more hydrophilic side chain Gln305, to mediate the binding posture of the predominantly hydrophobic substrate, such that the substrate methyl moiety was more directly oriented towards the activated heme-bound oxygen at 3.8 Å. This explained the improved hydroxylation activity of P450pyr3M over the wild type when both were treated with 4-fluorotoluene, as experimentally observed. As an additional note, for both cases, the hydrophilic fluoro substituent was well oriented towards the hydrophilic region near ASP258 at the helix as shown.

Additional docking simulations with 3-fluorotoluene and 4-bromotoluene were also established for P450pyr3M. As shown in Figure 5.10C and 5.10D respectively, with 3-fluorotoluene as substrate, the substrate methyl-C to heme-O distance is 3.6 Å; while the corresponding distance is shorter at 3.3 Å with 4-bromotoluene as the substrate. Comparing the docked postures of 3-fluorotoluene and 4-fluorotoluene in the P450pyr3M, both postures were remarkably similar, indicating an overall optimal binding pose of fluoro-toluene substrates in P450pyr3M. The fluoro-substituent repositioning from *para*- to *meta*- position was sufficient to tilt the substrate methly-C closer to the heme-O from 3.8 Å to 3.6 Å when docked (Figure 5.10B and 5.10C).

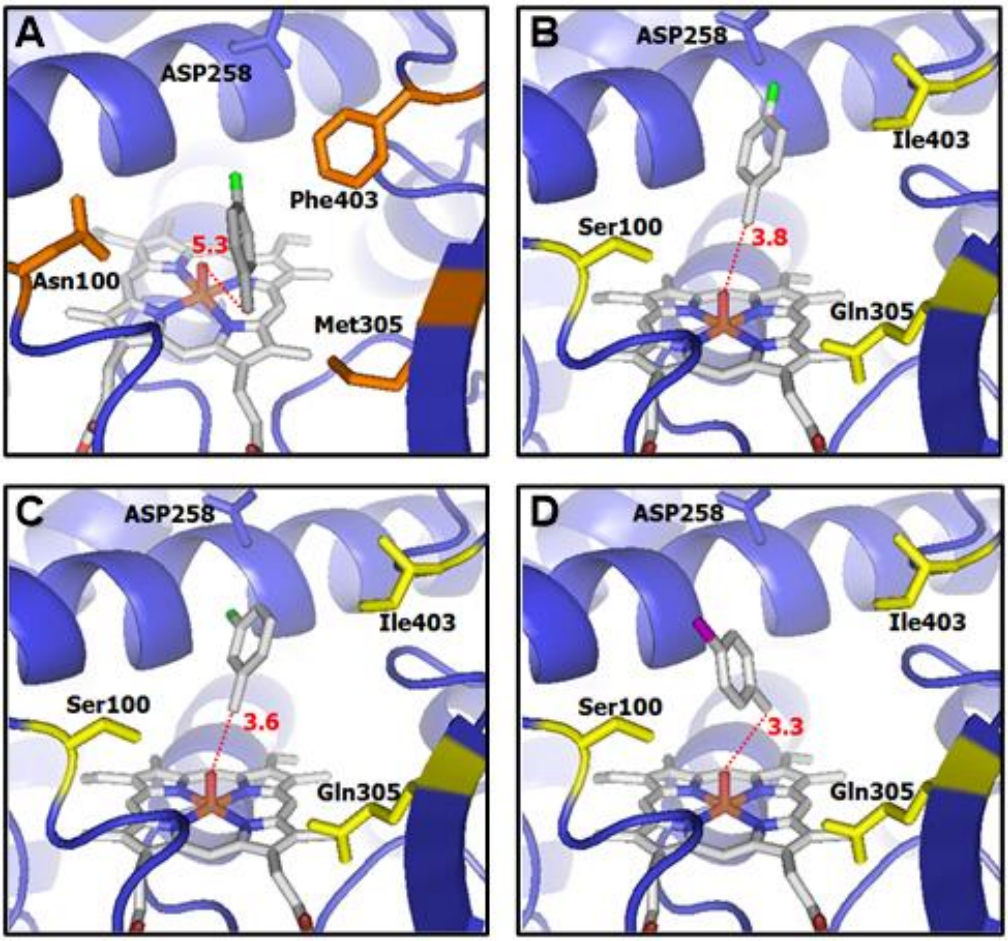


Figure 5.10 Substrate-enzyme binding pose. (a) Wild-type P450pyr with 4-fluorotoluene (b) P450pyr3M with 4-fluorotoluene, (c) P450pyr3M and 3-fluorotoluene, and (d) P450pyr3M and 4-bromotoluene. Mutation sites are illustrated in yellow. The distance (in angstrom) between the methyl-C atom of substrate and the heme-oxygen atom is denoted by a red dashed line.

For 4-bromotoluene, steric repulsion existed between the much larger bromo-substituent and the hydrophilic region near ASP258. The enlarged binding pocket space between Ser100 and Ile403 allowed a reorientation of 4-bromotoluene to avoid the aforementioned steric repulsion, while subject to hydrophobic-hydrophilic constrain mediated by Gln305. The resulting docked posture had substrate methly-C to heme-O distance reduced to 3.3 Å (Figure 5.9D). Thus the preference of inserting an O-atom into a halo-toluene substrate was improved with mutations over the wild-type enzyme, along with *meta*-substituent positions over *para*-substituent positions (for the same halo group);

and bromo-substituents over flouro-substituents (for the same substituent position). Overall, the *in silico* modelling and docking simulations can help account for the higher hydroxylation activity in the mutant enzymes with the differences in halogen substituent choices.

An alternative appreciation of the effect of the triple mutation N100S/M305Q/F403I on the substrate binding postures was presented in Figures 5.11. They depicted the surface representation of the docked substrates for the P450pyr and P450pyr3M. The binding posture of 3-fluorotoluene in the P450pyr with the narrower binding pocket space was illustrated in Figure 5.11A, in contrast to that of same substrate in the widened pocket space of the mutant enzyme (Figure 5.11B). In Figure 5.11C, the docked posture of 3-fluorotoluene in the mutant enzyme was remarkably similar to that for 4-fluorotoluene (Figure 5.11B), where in both cases, the hydrophilic fluoro-substituent was well oriented towards hydrophilic region near ASP258 (represented as light pink patch). A significant change in docked posture occured for 4-bromotoluene as shown in Figure 5.11D where steric hindrance between the bromo-substituent and the region near ASP258 forced the substrate to re-orientate along the increased space near Ser100-Ile102, thereby reducing the methyl-C to heme-O distance down to 3.3 Å (distance as shown in Figure 5.10).

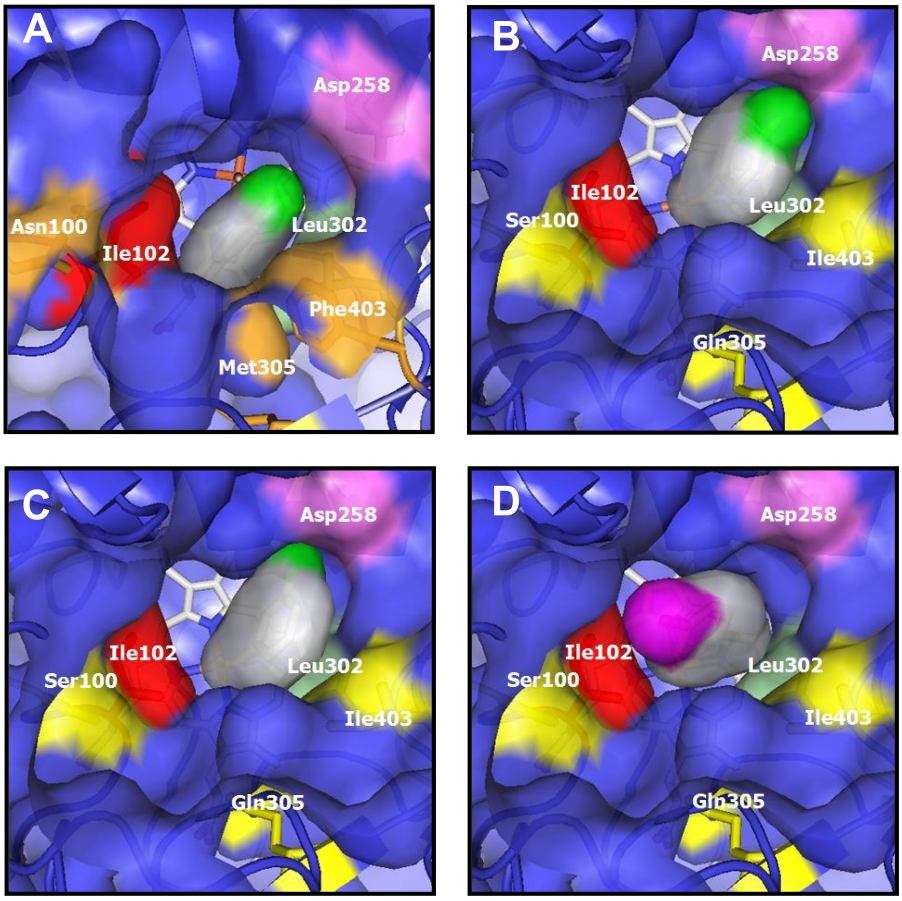


Figure 5.11 Surface representation of enzyme-substrate binding pose (a) 4-fluorotoluene in wild type P450pyr (b) 4-fluorotoluene in triple mutant P450pyr3M (c) 3-fluorotoluene in P450pyr3M (d) 4-bromotoluene in P450pyr3M. Mutated residues are shown in yellow.

5.4 Conclusions

In summary, total 29 recombinants *E. coli* P450pyr variants as well as wild-type P450pyr were applied for the biohydroxylation of fluoro-substituted toluene. One triple mutant, P450pyr3M, was discovered as the first enzyme with excellent activity and chemo- and regioselective for the benzylic hydroxylation of fluoro-, chloro- and bromo-substituted toluenes. This enzyme also worked well with the substrates containing electron-withdrawing groups, such as 3,4-difluorotoluene and 2,3,4,5,6-pentafluorotoluene. These hydroxylation reactions provide a simple access to the corresponding halo-benzyl alcohols,

which are useful pharmaceutical intermediates and cannot be prepared thus far by using other enzymatic or chemical systems. *In silico* modelling and substrate docking provided with some insight into the influence of three important amino acid mutations of the engineered P450pyr mutant on the enhanced activity. The obtained information and knowledge is useful for further engineering of P450pyr for other hydroxylations and oxidations.

Chapter 6. Summary and recommendations

6.1 Summary

Regio- and enantioselective hydroxylation is a useful but challenging reaction in classic chemistry. Alternatively, nature finds a useful solution for this type of reaction by using monooxygenases as catalysts. Many monooxygenases have been used for the hydroxylations of special types of substrates and cytochrome P450 monooxygenases form the largest sub-family. Nevertheless, there are still many problems left that limit the further application of P450 monooxygenases for regio- and enantioselective hydroxylations, such as the narrow substrate scope and unsatisfied activity and selectivity toward a given non-natural substrate. In this thesis, we used the directed evolution method to develop a set of novel P450 monooxygenases with better regio- and/or enantioselectivity, new substrate acceptance and improved activity for certain targeted hydroxylations.

Firstly, a terminal selective P450pyr was successfully engineered through directed evolution for the subterminal hydroxylation of alkanes with excellent enantioselectivity. A novel colorimetric HTS assay was developed for the measurement of both the regioselectivity and the enantioselectivity of a hydroxylation reaction. By using this HTS assay and ISM, sextuple-mutant P450pyrSM1 was created for the hydroxylation of *n*-octane to give (*S*)-2-octanol with 98 % *ee* and > 99% subterminal selectivity. It is a breakthrough in the directed evolution: 1) the engineered P450pyr mutant is the first enzyme for the highly selective alkane hydroxylation. 2) the generation of P450pyrSM1 is the first example of the full alteration of enzyme regioselectivity and

simultaneous establishment of high enantioselectivity by directed evolution. 3) the new developed colorimetric HTS assay could be generally applicable to the discovery of other types of catalyst for regio- and enantioselective hydroxylations.

Secondly, this P450pyr was successfully engineered for highly regioselective terminal hydroxylation of *n*-butanol. Based on a surrogate substrate-based colorimetric HTS assay, in two rounds of evolution, P450pyr single mutant I83M and double mutant I83M/I82T were generated to show excellent terminal regioselectivity for the hydroxylation of *n*-butanol. This is the first report of this hydroxylation by chemical or enzymatic method. This is also a unique example of evolving a monooxygenase to switch the substrate acceptance from a hydrophobic to hydrophilic compound.

Thirdly, an engineered P450pyr mutant (P450pyr3M) was discovered as the best catalyst for benzylic hydroxylation of single and multiple fluoro-substituted toluenes. This mutant also showed a broader substrate range, higher activity and higher conversion for the hydroxylation of chloro-, and bromo-toluenes compare with all the reported chemical- and bio-catalysts. These hydroxylations provide a simple access to the corresponding halo-benzyl alcohols, which are useful pharmaceutical intermediates.

In conclusion, this thesis demonstrated successful examples of using directed evolution to create new enzymes with better catalytic performances, including: better regio- and/or enantioselectivity, new substrate acceptance and improved activity for the green, clean and sustainable production of chemicals and pharmaceuticals. Molecular modelling on the hydroxylations of different

substrates with P450pyr and P450pyr mutants gave some insights on the unsderstanding of the roles of several key amino acid mutations for the better catalytic performances. This information could be useful for further evolution of P450pyr to achive high catalytic performance for other desired hydroxylations or oxidations.

6.2 Recommendations

The further development of biocatalysts for the green and sustainable production of chemicals could focus on the following aspects.

Besides activity, substrate scope and selectivity, the P450pyr stability could also be improved through directed evolution. Except the ISM method used in this thesis, other methods such as: epPCR and DNA shuffling could also be used to improve the enzyme's stability. Besides, based on the crystal structure and the knowledge of enzyme stability, we may rationally change some literature-known key residues which located not in the activity site but at the enzyme surface to improve the enzyme stability. With increased stability, the P450pyr enzyme would have longer survival times, improved reaction kinetics and more tolerant of the destabilising effects of mutations.

Besides ISM, other methods such as synthetic biology, whereby increasingly large sequences of DNA can be synthesised *de novo*, could be used to further engineer P450pyr with novel functions. Based on other protein sequence-structure-function relationships, more judicious strategies could be used to

design (for binding, specificity and active site modelling) the best amino acid sequence to develop novel biocatalysts that are both highly active and robust.

The regio- and enantioselective HTS assay was proven to be very simple, fast and powerful in our study. This method can be further applied in P450pyr engineering for the regio- and enantioselective hydroxylation of other more useful substrates, such as steroids, which have important clinical applications but industrial productions are limited due to complexity of the molecules. The knowledge of the P450pyr and P450pyr-substrate structures will provide useful information on possible amino acid sites that become important for substrate binding, which could be employed as a basic for the screening of new substrates. What's more, as we mentioned several times in this thesis, a few amino acid residues in the active site have been repeatedly observed to play key roles during substrate binding and regio- and enantioselectivity controlling. ISM on these sites would save a lot of mutant library generation and HTS screening work.

Although the new developed HTS assays work well in the experiment, these assays still suffers from the low product concentration at the end of reaction. This is possibly because of low cell amount and mass transfer inhibition of substrate molecules in each well. In order to generate more cells in the 96-deep well plate, rich medium such as TB or SOC could be used, while the substrate mass transfer problem could be solved by treatment using solvent, detergent, freezing and thawing, both method could further increase the biohydroxylation conversion in the HTS.

Several organisms and enzymes have been discovered to produce *n*-butanol through hydroxylation of butane. Co-expression of butane monooxygenase and

P450pyr monooxygenase could be used to attempt to produce 1,4-butanediol directly from butane in a one pot reaction. This can be done by either transforming the P450pyr, Fdx and FdR gene into organisms that can convert butane to butanol or coupled two permeabilised microorganisms.

The productivity of P450pyr and its mutants were still not good enough for practical application. The P450-catalysed biohydroxylations suffer from the immiscibility of substrate and/or substrate and product inhibitions, and thus the productivities are generally lower than other types of enzyme-catalysed reaction. To enhance the biohydroxylation productivity, several biphasic systems such as: aqueous/ organic biphasic system, aqueous/resin biphasic system and aqueous/ionic liquid biphasic system could be used. Besides, the cell grow condition, proteins (P450, Fdx and FdR) expression level/ratio and substrate concentration could also be studied. What's more, the expressing *E.coli* strain, dissolved oxygen content, IPTG and ALA concentrations also have to be optimised to achieve a higher productivity.

Finally, in order to understand the biochemistry mechanism of energy transfer within P450pyr system, it could be necessary to crystallise its redox factors: ferredoxin (Fdx) and ferredoxin reductase (FdR). The X-ray structure of Fdx and FdR together with P450pyr, would provide an integral structure-function relationship of P450pyr complex system, which could be useful for the improvement of P450pyr activity and productivity, as well as developing a P450pyr-Fdx-FdR fusion protein for enzyme immobilization. Besides, it would also be interesting to crystallise P450pyr mutants together with their substrates. These structures would greatly enhance our understanding of the structure-function relationship between P450pyr and substrates, and these structures

could also provide an insight into the connection between selectivity and key residues in P450pyr active pocket, which would be very useful and helpful to further improve the design of new amino acid mutations within the binding pocket that are important for improving P450pyr hydroxylation performance.

Bibliography

1. Cayuela Valencia, R. *The future of the chemical industry by 2050*, xiv, 322 pages (Wiley-VCH, Weinheim, 2013).
2. Anastas, P.T. & Williamson, T.C. Green chemistry: An overview. in *Green Chemistry: Designing Chemistry for the Environment*, Vol. 626 (eds. Anastas, P.T. & Williamson, T.C.) 1-17 (1996).
3. Sheldon, R.A. The E factor: fifteen years on. *Green Chemistry* 9, 1273-1283 (2007).
4. Anastas, P.T. & Warner, J.C. *Green chemistry : theory and practice*, xi, 135 p. (Oxford University Press, Oxford England ; New York, 1998).
5. Anastas, P. & Eghbali, N. Green Chemistry: Principles and Practice. *Chemical Society Reviews* 39, 301-312 (2010).
6. Clark, J.H. Green Chemistry: challenges and opportunities. *Green Chemistry* 1, 1-8 (1999).
7. Constable, D.J.C. et al. Key green chemistry research areas - a perspective from pharmaceutical manufacturers. *Green Chemistry* 9, 411-420 (2007).
8. Schmid, A. et al. Industrial biocatalysis today and tomorrow. *Nature* 409, 258-68 (2001).
9. Buchholz, K., Kasche, V. & Bornscheuer, U.T. *Biocatalysts and enzyme technology*, xx, 606 p. (Wiley-Blackwell, Weinheim, 2012).
10. Gavrilescu, M. & Chisti, Y. Biotechnology-a sustainable alternative for chemical industry. *Biotechnol Adv* 23, 471-99 (2005).
11. Whittall, J. & Sutton, P. *Practical methods for biocatalysis and biotransformations*, (John Wiley & Sons Inc., Hoboken, 2012).
12. Buchholz, K. & Collins, J. The roots--a short history of industrial microbiology and biotechnology. *Appl Microbiol Biotechnol* 97, 3747-62 (2013).
13. Farnworth, E.R. *Handbook of fermented functional foods*, xviii, 581 p. (CRC Press, Boca Raton, 2008).
14. Adlercreutz, P. & Straathof, A.J.J. *Applied biocatalysis*, xvi, 443 p., 9 p. of plates (Harwood Academic Publishers, Amsterdam, 2000).
15. Burton, S.G., Cowan, D.A. & Woodley, J.M. The search for the ideal biocatalyst. *Nat Biotechnol* 20, 37-45 (2002).
16. Faber, K. *Biotransformations in organic chemistry*, (Springer, New York, 2011).
17. Petrounia, I.P. & Arnold, F.H. Designed evolution of enzymatic properties. *Curr Opin Biotechnol* 11, 325-30 (2000).
18. Turner, N.J. Directed evolution drives the next generation of biocatalysts. *Nat Chem Biol* 5, 567-73 (2009).

19. Rubin-Pitel, S.B. & Zhao, H. Recent advances in biocatalysis by directed enzyme evolution. *Comb Chem High Throughput Screen* 9, 247-57 (2006).

20. Jaeger, K.E. & Eggert, T. Enantioselective biocatalysis optimized by directed evolution. *Curr Opin Biotechnol* 15, 305-13 (2004).

21. Turner, N.J. Directed evolution of enzymes for applied biocatalysis. *Trends Biotechnol* 21, 474-8 (2003).

22. Powell, K.A. et al. Directed Evolution and Biocatalysis. *Angew Chem Int Ed Engl* 40, 3948-3959 (2001).

23. Zhao, H., Chockalingam, K. & Chen, Z. Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr Opin Biotechnol* 13, 104-10 (2002).

24. Reetz, M.T., Soni, P., Fernandez, L., Gumulya, Y. & Carballeira, J.D. Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem Commun (Camb)* 46, 8657-8 (2010).

25. Eijsink, V.G., Gaseidnes, S., Borchert, T.V. & van den Burg, B. Directed evolution of enzyme stability. *Biomol Eng* 22, 21-30 (2005).

26. Hoseki, J. et al. Increased rigidity of domain structures enhances the stability of a mutant enzyme created by directed evolution. *Biochemistry* 42, 14469-75 (2003).

27. Chica, R.A., Doucet, N. & Pelletier, J.N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr Opin Biotechnol* 16, 378-84 (2005).

28. Altamirano, M.M., Blackburn, J.M., Aguayo, C. & Fersht, A.R. Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold. *Nature* 403, 617-22 (2000).

29. Cherry, J.R. & Fidantsef, A.L. Directed evolution of industrial enzymes: an update. *Curr Opin Biotechnol* 14, 438-43 (2003).

30. May, O., Nguyen, P.T. & Arnold, F.H. Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine. *Nat Biotechnol* 18, 317-20 (2000).

31. Bornscheuer, U.T. & Pohl, M. Improved biocatalysts by directed evolution and rational protein design. *Curr Opin Chem Biol* 5, 137-43 (2001).

32. Reetz, M.T., Wilensek, S., Zha, D. & Jaeger, K.E. Directed Evolution of an Enantioselective Enzyme through Combinatorial Multiple-Cassette Mutagenesis. *Angew Chem Int Ed Engl* 40, 3589-3591 (2001).

33. Carr, R. et al. Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity. *Angewandte Chemie-International Edition* 42, 4807-4810 (2003).

34. Yano, T., Oue, S. & Kagamiyama, H. Directed evolution of an aspartate aminotransferase with new substrate specificities. *Proc Natl Acad Sci U S A* 95, 5511-5 (1998).

35. Lingen, B. et al. Alteration of the substrate specificity of benzoylformate decarboxylase from Pseudomonas putida by directed evolution. *Chembiochem* 4, 721-6 (2003).

36. Liese, A., Seelbach, K. & Wandrey, C. *Industrial biotransformations*, xiv, 556 p. (Wiley-VCH, Weinheim, 2006).

37. Straathof, A.J.J., Panke, S. & Schmid, A. The production of fine chemicals by biotransformations. *Current Opinion in Biotechnology* 13, 548-556 (2002).

38. Faber, K. *Biotransformations in organic chemistry : a textbook*, xi, 454 p. (Springer-Verlag, Berlin ; New York, 2004).

39. Bommarius, A.S. & Riebel, B.R. *Biocatalysis*, xxiii, 611 p. (Wiley-VCH, Weinheim ; Cambridge, 2004).

40. Patel, R.N. Synthesis of chiral pharmaceutical intermediates by biocatalysis. *Coordination Chemistry Reviews* 252, 659-701 (2008).

41. Wohlgemuth, R. Asymmetric biocatalysis with microbial enzymes and cells. *Curr Opin Microbiol* 13, 283-92 (2010).

42. Yang, Y., Liu, J. & Li, Z. Engineering of P450pyr Hydroxylase for the Highly Regio- and Enantioselective Subterminal Hydroxylation of Alkanes. *Angewandte Chemie-International Edition* 53, 3120-3124 (2014).

43. Wohlgemuth, R. Biocatalysis--key to sustainable industrial chemistry. *Curr Opin Biotechnol* 21, 713-24 (2010).

44. Coad, J. *Green technology*, 48 p. (Raintree, Chicago, Ill., 2012).

45. Polizzi, K.M., Bommarius, A.S., Broering, J.M. & Chaparro-Riggers, J.F. Stability of biocatalysts. *Curr Opin Chem Biol* 11, 220-5 (2007).

46. Riva, S. Biocatalytic modification of natural products. *Curr Opin Chem Biol* 5, 106-11 (2001).

47. Hibbert, E.G. et al. Directed evolution of transketolase activity on non-phosphorylated substrates. *J Biotechnol* 131, 425-32 (2007).

48. Ngo, T.P.N., Li, A., Tiew, K.W. & Li, Z. Efficient transformation of grease to biodiesel using highly active and easily recyclable magnetic nanobiocatalyst aggregates. *Bioresource Technology* 145, 233-239 (2013).

49. Huang, R., Wu, S., Li, A. & Li, Z. Integrating interfacial self-assembly and electrostatic complexation at an aqueous interface for capsule synthesis and enzyme immobilization. *Journal of Materials Chemistry A* 2, 1672-1676 (2014).

50. Wang, W., Xu, Y., Wang, D.I.C. & Li, Z. Recyclable Nanobiocatalyst for Enantioselective Sulfoxidation: Facile Fabrication and High Performance of Chloroperoxidase-Coated Magnetic Nanoparticles with Iron Oxide Core and Polymer Shell. *Journal of the American Chemical Society* 131, 12892-+ (2009).

51. Olah, G.A. & Molnár, A.r.d. *Hydrocarbon chemistry*, xxiv, 871 p. (Wiley-Interscience, Hoboken, N.J., 2003).

52. Sanford, M.S. C-H bond functionalization in organic synthesis. *Abstracts of Papers of the American Chemical Society* 233(2007).

53. Kumar, S. Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert Opinion on Drug Metabolism & Toxicology* 6, 115-131 (2010).

54. Periana, R.A. et al. Perspectives on some challenges and approaches for developing the next generation of selective, low temperature, oxidation catalysts for alkane hydroxylation based on the CH activation reaction. *Journal of Molecular Catalysis a-Chemical* 220, 7-25 (2004).

55. Speight, J.G. *The chemistry and technology of petroleum*, xxvi, 927 pages (CRC Press, Taylor & francis Group, Boca Raton, 2014).

56. Vasishtha, A.K., Trivedi, R.K. & Das, G. SEBACIC ACID AND 2-OCTANOL FROM CASTOR-OIL. *Journal of the American Oil Chemists Society* 67, 333-337 (1990).

57. Volante, R.P. A NEW, HIGHLY EFFICIENT METHOD FOR THE CONVERSION OF ALCOHOLS TO THIOLESTERS AND THIOLS. *Tetrahedron Letters* 22, 3119-3122 (1981).

58. Davis, B.H. & Brey, W.S. DEHYDRATION AND DEHYDROGENATION OF 2-OCTANOL BY THORIUM OXIDE. *Journal of Catalysis* 25, 81-& (1972).

59. Chokkaram, S., Srinivasan, R., Milburn, D.R. & Davis, B.H. Conversion of 2-octanol over nickel-alumina, cobalt-alumina, and alumina catalysts. *Journal of Molecular Catalysis a-Chemical* 121, 157-169 (1997).

60. van Woezik, B.A.A. & Westerterp, K.R. The nitric acid oxidation of 2-octanol. A model reaction for multiple heterogeneous liquid-liquid reactions. *Chemical Engineering and Processing* 39, 521-537 (2000).

61. Agulyansky, A., Agulyansky, L. & Travkin, V.F. Liquid-liquid extraction of tantalum with 2-octanol. *Chemical Engineering and Processing* 43, 1231-1237 (2004).

62. Shioiri, T., Izawa, K. & Konoike, T. *Pharmaceutical process chemistry*, xxiv, 502 p. (Wiley-VCH, Weinheim, Germany, 2011).

63. Meinhold, P., Peters, M.W., Chen, M.M., Takahashi, K. & Arnold, F.H. Direct conversion of ethane to ethanol by engineered cytochrome P450 BM3. *Chembiochem* 6, 1765-8 (2005).

64. Hermans, I., Spier, E.S., Neuenschwander, U., Turra, N. & Baiker, A. CATL 22-Selective oxidation catalysis: Opportunities and challenges. *Abstracts of Papers of the American Chemical Society* 238(2009).

65. Ge, H., Leng, Y., Zhou, C. & Wang, J. Direct hydroxylation of benzene to phenol with molecular oxygen over phase transfer catalysts: Cyclodextrins complexes with vanadium-substituted heteropoly acids. *Catalysis Letters* 124, 324-329 (2008).

66. Zhang, X. Direct hydroxylation of benzene to phenol. *Progress in Chemistry* 20, 386-395 (2008).

67. Niwa, S. et al. A one-step conversion of benzene to phenol with a palladium membrane. *Science* 295, 105-107 (2002).

68. Kakiuchi, F. & Chatani, N. Catalytic methods for C-H bond functionalization: Application in organic synthesis. *Advanced Synthesis & Catalysis* 345, 1077-1101 (2003).

69. Davies, H.M.L. & Beckwith, R.E.J. Catalytic enantioselective C-H activation by means of metal-carbenoid-induced C-H insertion. *Chemical Reviews* 103, 2861-2903 (2003).

70. Wendlandt, A.E., Suess, A.M. & Stahl, S.S. Copper-Catalyzed Aerobic Oxidative C?H Functionalizations: Trends and Mechanistic Insights. *Angewandte Chemie-International Edition* 50, 11062-11087 (2011).

71. Park, Y.J., Park, J.W. & Jun, C.H. Metal-organic cooperative catalysis in C-H and C-C bond activation and its concurrent recovery. *Accounts of Chemical Research* 41, 222-234 (2008).

72. Periana, R.A. et al. Platinum catalysts for the high-yield oxidation of methane to a methanol derivative. *Science* 280, 560-564 (1998).

73. Jones, W.D. Hydrocarbon chemistry - Conquering the carbon-hydrogen bond. *Science* 287, 1942-1943 (2000).

74. Wick, D.D. & Goldberg, K.I. C-H activation of Pt(II) to form stable Pt(IV) alkyl hydrides. *Journal of the American Chemical Society* 119, 10235-10236 (1997).

75. Chen, M.S. & White, M.C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science* 318, 783-787 (2007).

76. Cowan, A.J. & George, M.W. Formation and reactivity of organometallic alkane complexes. *Coordination Chemistry Reviews* 252, 2504-2511 (2008).

77. Graham-Lorence, S., Amarneh, B., White, R.E., Peterson, J.A. & Simpson, E.R. A three-dimensional model of aromatase cytochrome P450. *Protein Sci* 4, 1065-80 (1995).

78. Berkowitz, J., Ellison, G.B. & Gutman, D. 3 METHODS TO MEASURE RH BOND-ENERGIES. *Journal of Physical Chemistry* 98, 2744-2765 (1994).

79. Crabtree, R.H. Aspects of Methane Chemistry. *Chemical Reviews* 95, 987-1007 (1995).

80. Reetz, M.T. Biocatalysis in organic chemistry and biotechnology: past, present, and future. *J Am Chem Soc* 135, 12480-96 (2013).

81. Reetz, M.T. et al. Directed evolution as a method to create enantioselective cyclohexanone monooxygenases for catalysis in Baeyer-Villiger reactions. *Angewandte Chemie-International Edition* 43, 4075-4078 (2004).

82. Lipscomb, J.D. Biochemistry of the soluble methane monooxygenase. *Annu Rev Microbiol* 48, 371-99 (1994).

83. Balasubramanian, R. & Rosenzweig, A.C. Structural and mechanistic insights into methane oxidation by particulate methane monooxygenase. *Accounts of Chemical Research* 40, 573-580 (2007).

84. Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *Journal of Biotechnology* 124, 128-145 (2006).

85. Huisman, G.W. & Gray, D. Towards novel processes for the fine-chemical and pharmaceutical industries. *Current Opinion in Biotechnology* 13, 352-358 (2002).

86. Pollard, D.J. & Woodley, J.M. Biocatalysis for pharmaceutical intermediates: the future is now. *Trends in Biotechnology* 25, 66-73 (2007).

87. Whittington, D.A., Sazinsky, M.H. & Lippard, S.J. X-ray crystal structure of alcohol products bound at the active site of soluble methane monooxygenase hydroxylase. *Journal of the American Chemical Society* 123, 1794-1795 (2001).

88. Chen, K.H.C., Chan, S.I., Yu, S.S.F., Chen, C.L. & Kuo, S.S.J. Toward delineating the structure and function of the particulate methane monooxygenase (pMMO) from methanotrophic bacteria. *Faseb Journal* 18, C145-C145 (2004).

89. Chan, S.I. & Yu, S.S.F. Controlled oxidation of hydrocarbons by the membrane-bound methane monooxygenase: The case for a tricopper cluster. *Accounts of Chemical Research* 41, 969-979 (2008).

90. Kotani, T., Kawashima, Y., Yurimoto, H., Kato, N. & Sakai, Y. Gene structure and regulation of alkane monooxygenases in propane-utilizing Mycobacterium sp TY-6 and Pseudonocardia sp TY-7. *Journal of Bioscience and Bioengineering* 102, 184-192 (2006).

91. Glieder, A., Farinas, E.T. & Arnold, F.H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nature Biotechnology* 20, 1135-1139 (2002).

92. Whyte, L.G. et al. Gene cloning and characterization of multiple alkane hydroxylase systems in Rhodococcus strains Q15 and NRRL B-16531. *Applied and Environmental Microbiology* 68, 5933-5942 (2002).

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

94. Wirtz, M., Klucik, J. & Rivera, M. Ferredoxin-mediated electrocatalytic dehalogenation of haloalkanes by cytochrome P450(cam). *Journal of the American Chemical Society* 122, 1047-1056 (2000).

95. Kazlauskaite, J., Westlake, A.C.G., Wong, L.L. & Hill, H.A.O. Direct electrochemistry of cytochrome P450cam. *Chemical Communications*, 2189-2190 (1996).

96. Li, H. & Poulos, T.L. The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nat Struct Biol* 4, 140-6 (1997).

97. Fulco, A.J. P450BM-3 AND OTHER INDUCIBLE BACTERIAL P450 CYTOCHROMES - BIOCHEMISTRY AND REGULATION. *Annual Review of Pharmacology and Toxicology* 31, 177-203 (1991).

98. Wong, T.S., Arnold, F.H. & Schwaneberg, U. Laboratory evolution of cytochrome P450BM-3 monooxygenase for organic cosolvents. *Biotechnology and Bioengineering* 85, 351-358 (2004).

99. Nazor, J. & Schwaneberg, U. Laboratory evolution of P450BM-3 for mediated electron transfer. *Chembiochem* 7, 638-644 (2006).

100. Li, H.Y. & Poulos, T.L. The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nature Structural Biology* 4, 140-146 (1997).

101. Landwehr, M. et al. Enantioselective alpha-hydroxylation of 2-arylacetic acid derivatives and buspirone catalyzed by engineered cytochrome P450BM-3. *Journal of the American Chemical Society* 128, 6058-6059 (2006).

102. Tang, W.L., Li, Z. & Zhao, H.M. Inverting the enantioselectivity of P450pyr monooxygenase by directed evolution. *Chemical Communications* 46, 5461-5463 (2010).

103. Pham, S.Q., Pompidor, G., Liu, J., Li, X.D. & Li, Z. Evolving P450pyr hydroxylase for highly enantioselective hydroxylation at non- activated carbon atom. *Chemical Communications* 48, 4618-4620 (2012).

104. Pham, S.Q., Gao, P. & Li, Z. Engineering of recombinant E. coli cells co-expressing P450pyrTM monooxygenase and glucose dehydrogenase for highly regio- and stereoselective hydroxylation of alicycles with cofactor recycling. *Biotechnol Bioeng* 110, 363-73 (2013).

105. Urlacher, V. & Schmid, R.D. Biotransformations using prokaryotic P450 monooxygenases. *Current Opinion in Biotechnology* 13, 557-564 (2002).

106. Peters, M.W., Meinhold, P., Glieder, A. & Arnold, F.H. Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *Journal of the American Chemical Society* 125, 13442-13450 (2003).

107. Scheps, D., Malca, S.H., Hoffmann, H., Nestl, B.M. & Hauer, B. Regioselective omega-hydroxylation of medium-chain n-alkanes and primary alcohols by CYP153 enzymes from Mycobacterium marinum and Polaromonas sp. strain JS666. *Org Biomol Chem* 9, 6727-33 (2011).

108. Nishio, T., Patel, A., Wang, Y. & Lau, P.C.K. Biotransformations catalyzed by cloned p-cymene monooxygenase from Pseudomonas putida F1. *Applied Microbiology and Biotechnology* 55, 321-325 (2001).

109. Wubbolts, M.G., Reuvekamp, P. & Witholt, B. TOL PLASMID-SPECIFIED XYLENE OXYGENASE IS A WIDE SUBSTRATE

RANGE MONOOXYGENASE CAPABLE OF OLEFIN EPOXIDATION. *Enzyme and Microbial Technology* 16, 608-615 (1994).

110. Hunter, T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80, 225-36 (1995).

111. Mackie, R.I. & White, B.A. Recent advances in rumen microbial ecology and metabolism: potential impact on nutrient output. *J Dairy Sci* 73, 2971-95 (1990).

112. Tipton, K. & Boyce, S. History of the enzyme nomenclature system. *Bioinformatics* 16, 34-40 (2000).

113. Straathof, A.J., Panke, S. & Schmid, A. The production of fine chemicals by biotransformations. *Curr Opin Biotechnol* 13, 548-56 (2002).

114. Liese, A., Seelbach, K. & Wandrey, C. *Industrial biotransformations*, 423 p. (Wiley-VCH, Weinheim ; New York, 2000).

115. Frey, P.A. & Hegeman, A.D. *Enzymatic reaction mechanisms*, xviii, 831 p. (Oxford University Press, Oxford ; New York, 2007).

116. Guengerich, F.P. Cytochrome p450 and chemical toxicology. *Chem Res Toxicol* 21, 70-83 (2008).

117. Cirino, P.C. & Arnold, F.H. Protein engineering of oxygenases for biocatalysis. *Current Opinion in Chemical Biology* 6, 130-135 (2002).

118. Harayama, S., Kok, M. & Neidle, E.L. FUNCTIONAL AND EVOLUTIONARY RELATIONSHIPS AMONG DIVERSE OXYGENASES. *Annual Review of Microbiology* 46, 565-601 (1992).

119. Labinger, J.A. & Bercaw, J.E. Understanding and exploiting C-H bond activation. *Nature* 417, 507-514 (2002).

120. Denisov, I.G., Makris, T.M., Sligar, S.G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chemical Reviews* 105, 2253-2277 (2005).

121. Whitehouse, C.J.C., Bell, S.G. & Wong, L.-L. P450(BM3) (CYP102A1): connecting the dots. *Chemical Society Reviews* 41, 1218-1260 (2012).

122. Choi, K.-Y. et al. Cloning, expression and characterization of CYP102D1, a self-sufficient P450 monooxygenase from Streptomyces avermitilis. *Febs Journal* 279, 1650-1662 (2012).

123. Merkx, M. et al. Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: A tale of two irons and three proteins. *Angewandte Chemie-International Edition* 40, 2782-2807 (2001).

124. Smith, T.J., Slade, S.E., Burton, N.P., Murrell, J.C. & Dalton, H. Improved system for protein engineering of the hydroxylase component of soluble methane monooxygenase. *Appl Environ Microbiol* 68, 5265-73 (2002).

125. Koch, D.J., Chen, M.M., van Beilen, J.B. & Arnold, F.H. In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Appl Environ Microbiol* 75, 337-44 (2009).

126. Groves, J.T., Austin, R.N., Chang, H.K. & Zylstra, G.J. The non-heme diiron alkane monooxygenase of Pseudomonas oleovorans (AlkB) hydroxylates via a substrate radical intermediate. *Journal of Inorganic Biochemistry* 86, 54-54 (2001).

127. Jung, S.T., Lauchli, R. & Arnold, F.H. Cytochrome P450: taming a wild type enzyme. *Curr Opin Biotechnol* 22, 809-17 (2011).

128. Kumar, S. Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert Opin Drug Metab Toxicol* 6, 115-31 (2010).

129. Urlacher, V.B. & Girhard, M. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol* 30, 26-36 (2012).

130. Lipscomb, J.D. Biochemistry of the soluble methane monooxygenase. in *Annual Review of Microbiology*, Vol. 48 (ed. Ornston, L.N.) 371-399 (1994).

131. Jiang, Y., Wilkins, P.C. & Dalton, H. ACTIVATION OF THE HYDROXYLASE OF SMMO FROM METHYLOCOCCUS-CAPSULATUS (BATH) BY HYDROGEN-PEROXIDE. *Biochimica Et Biophysica Acta* 1163, 105-112 (1993).

132. Murrell, J.C., McDonald, I.R. & Gilbert, B. Regulation of expression of methane monooxygenases by copper ions. *Trends in Microbiology* 8, 221-225 (2000).

133. Pilkington, S.J. & Dalton, H. PURIFICATION AND CHARACTERIZATION OF THE SOLUBLE METHANE MONOOXYGENASE FROM METHYLOSINUS-SPORIUM-5 DEMONSTRATES THE HIGHLY CONSERVED NATURE OF THIS ENZYME IN METHANOTROPHS. *Fems Microbiology Letters* 78, 103-108 (1991).

134. Green, J. & Dalton, H. A STOPPED-FLOW KINETIC-STUDY OF SOLUBLE METHANE MONO-OXYGENASE FROM METHYLOCOCCUS-CAPSULATUS (BATH). *Biochemical Journal* 259, 167-172 (1989).

135. Green, J. & Dalton, H. STEADY-STATE KINETIC-ANALYSIS OF SOLUBLE METHANE MONO-OXYGENASE FROM METHYLOCOCCUS-CAPSULATUS (BATH). *Biochemical Journal* 236, 155-162 (1986).

136. Carlsen, H.N., Joergensen, L. & Degn, H. INHIBITION BY AMMONIA OF METHANE UTILIZATION IN METHYLOCOCCUS-CAPSULATUS (BATH). *Applied Microbiology and Biotechnology* 35, 124-127 (1991).

137. Yoshinari, T. & Shafer, D. DEGRADATION OF DIMETHYL NITROSAMINE BY METHYLOSINUS-TRICHOSPORIUM OB3B. *Canadian Journal of Microbiology* 36, 834-838 (1990).

138. Davis, K.J., Cornish, A. & Higgins, I.J. REGULATION OF THE INTRACELLULAR LOCATION OF METHANE MONO-OXYGENASE DURING GROWTH OF METHYLOSINUS-TRICHOSPORIUM OB3B ON METHANOL. *Journal of General Microbiology* 133, 291-297 (1987).

139. Cornish, A. et al. SUCCINATE AS AN INVITRO ELECTRON-DONOR FOR THE PARTICULATE METHANE MONO-OXYGENASE OF METHYLOSINUS-TRICHOSPORIUM OB3B. *Biotechnology Letters* 7, 319-324 (1985).

140. Cardy, D.L.N., Laidler, V., Salmond, G.P.C. & Murrell, J.C. THE METHANE MONOOXYGENASE GENE-CLUSTER OF METHYLOSINUS-TRICHOSPORIUM - CLONING AND SEQUENCING OF THE MMOC GENE. *Archives of Microbiology* 156, 477-483 (1991).

141. Elango, N. et al. Crystal structure of the hydroxylase component of methane monooxygenase from Methylosinus trichosporium OB3b. *Protein Sci* 6, 556-68 (1997).

142. Yang, C.-G. et al. Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* 452, 961-U4 (2008).

143. Koch, D.J., Chen, M.M., van Beilen, J.B. & Arnold, F.H. In Vivo Evolution of Butane Oxidation by Terminal Alkane Hydroxylases AlkB and CYP153A6. *Applied and Environmental Microbiology* 75, 337-344 (2009).

144. Degtyarenko, K.N. & Archakov, A.I. MOLECULAR EVOLUTION OF P450 SUPERFAMILY AND P450-CONTAINING MONOOXYGENASE SYSTEMS. *Febs Letters* 332, 1-8 (1993).

145. Degtyarenko, K.N. STRUCTURAL DOMAINS OF P450-CONTAINING MONOOXYGENASE SYSTEMS. *Protein Engineering* 8, 737-747 (1995).

146. Hannemann, F., Bichet, A., Ewen, K.M. & Bernhardt, R. Cytochrome P450 systems - biological variations of electron transport chains. *Biochimica Et Biophysica Acta-General Subjects* 1770, 330-344 (2007).

147. Funhoff, E.G. & Van Beilen, J.B. Alkane activation by P450 oxygenases. *Biocatalysis and Biotransformation* 25, 186-193 (2007).

148. Farinas, Edgardo T., Schwaneberg, U., Glieder, A. & Arnold, Frances H. Directed Evolution of a Cytochrome P450 Monooxygenase for Alkane Oxidation. *Advanced Synthesis & Catalysis* 343, 601-606 (2001).

149. Poulos, T.L., Finzel, B.C. & Howard, A.J. HIGH-RESOLUTION CRYSTAL-STRUCTURE OF CYTOCHROME-P450CAM. *Journal of Molecular Biology* 195, 687-700 (1987).

150. Ravichandran, K.G., Boddupalli, S.S., Hasemann, C.A., Peterson, J.A. & Deisenhofer, J. CRYSTAL-STRUCTURE OF HEMOPROTEIN DOMAIN OF P450BM-3, A PROTOTYPE FOR MICROSOMAL P450S. *Science* 261, 731-736 (1993).

151. Gunsalus, I.C. & Sligar, S.G. Oxygen reduction by the P450 monoxygenase systems. *Adv Enzymol Relat Areas Mol Biol* 47, 1-44 (1978).

152. Schiffler, B. & Bernhardt, R. Bacterial (CYP101) and mitochondrial P450 systems-how comparable are they? *Biochem Biophys Res Commun* 312, 223-8 (2003).

153. Fjaervik, E. & Zotchev, S.B. Biosynthesis of the polyene macrolide antibiotic nystatin in Streptomyces noursei. *Applied Microbiology and Biotechnology* 67, 436-443 (2005).

154. Hannemann, F., Bichet, A., Ewen, K.M. & Bernhardt, R. Cytochrome P450 systems--biological variations of electron transport chains. *Biochim Biophys Acta* 1770, 330-44 (2007).

155. White, P.C., Pascoe, L., Curnow, K.M., Tannin, G. & Rosler, A. MOLECULAR-BIOLOGY OF 11-BETA-HYDROXYLASE AND 11-BETA-HYDROXYSTEROID DEHYDROGENASE ENZYMES. *Journal of Steroid Biochemistry and Molecular Biology* 43, 827-835 (1992).

156. White, P.C. et al. A MUTATION IN CYP11B1 (ARG-448- HIS) ASSOCIATED WITH STEROID 11-BETA-HYDROXYLASE DEFICIENCY IN JEWS OF MOROCCAN ORIGIN. *Journal of Clinical Investigation* 87, 1664-1667 (1991).

157. Mornet, E., Dupont, J., Vitek, A. & White, P.C. CHARACTERIZATION OF 2 GENES ENCODING HUMAN STEROID 11-BETA-HYDROXYLASE (P-45011-BETA). *Journal of Biological Chemistry* 264, 20961-20967 (1989).

158. Lewis, D.F.V. *Guide to cytochrome P450 structure and function*, xxiii, 215 p., 8 p. of plates (Taylor & Francis, London ; New York, 2001).

159. Hasler, J.A. et al. Human cytochromes P450. *Molecular Aspects of Medicine* 20, 1-137 (1999).

160. Narhi, L.O. & Fulco, A.J. Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in Bacillus megaterium. *J Biol Chem* 261, 7160-9 (1986).

161. Warman, A.J. et al. Flavocytochrome P450BM3: an update on structure and mechanism of a biotechnologically important enzyme. *Biochemical Society Transactions* 33, 747-753 (2005).

162. Nakayama, N., Takemae, A. & Shoun, H. Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus Fusarium oxysporum. *Journal of Biochemistry* 119, 435-440 (1996).

163. De Mot, R. & Parret, A.H.A. A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. *Trends in Microbiology* 10, 502-508 (2002).

164. Seo, J.A., Proctor, R.H. & Plattner, R.D. Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in Fusarium verticillioides. *Fungal Genetics and Biology* 34, 155-165 (2001).

165. Pazmino, D.E.T., Winkler, M., Glieder, A. & Fraaije, M.W. Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications. *Journal of Biotechnology* 146, 9-24 (2010).

166. Groves, J.T. The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies. *Proc Natl Acad Sci U S A* 100, 3569-74 (2003).

167. Schlichting, I. The Catalytic Pathway of Cytochrome P450cam at Atomic Resolution. *Science* 287, 1615-1622 (2000).

168. Rittle, J. & Green, M.T. Cytochrome P450 Compound I: Capture, Characterization, and C-H Bond Activation Kinetics. *Science* 330, 933-937 (2010).

169. Sevrioukova, I.F. & Poulos, T.L. Structural biology of redox partner interactions in P450cam monooxygenase: a fresh look at an old system. *Arch Biochem Biophys* 507, 66-74 (2011).

170. Harford-Cross, C.F. et al. Protein engineering of cytochrome P450cam (CYP101) for the oxidation of polycyclic aromatic hydrocarbons. *Protein Engineering Design and Selection* 13, 121-128 (2000).

171. Poulos, T.L., Finzel, B.C., Gunsalus, I.C., Wagner, G.C. & Kraut, J. THE 2.6-A CRYSTAL-STRUCTURE OF PSEUDOMONAS-PUTIDA CYTOCHROME-P-450. *Journal of Biological Chemistry* 260, 6122-6130 (1985).

172. Fowler, S.M. et al. CYTOCHROME P-450(CAM) MONOOXYGENASE CAN BE REDESIGNED TO CATALYZE THE REGIOSELECTIVE AROMATIC HYDROXYLATION OF DIPHENYLMETHANE. *Journal of the Chemical Society-Chemical Communications*, 2761-2762 (1994).

173. Stevenson, J.-A., Westlake, A.C.G., Whittock, C. & Wong, L.-L. The Catalytic Oxidation of Linear and Branched Alkanes by Cytochrome P450cam. *Journal of the American Chemical Society* 118, 12846-12847 (1996).

174. Kumar, S., Chen, C.S., Waxman, D.J. & Halpert, J.R. Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. *J Biol Chem* 280, 19569-75 (2005).

175. Bell, S.G. et al. Butane and propane oxidation by engineered cytochrome P450cam. *Chemical Communications*, 490-491 (2002).

176. Bell, S.G. et al. Engineering cytochrome P450cam into an alkane hydroxylase. *Dalton Transactions*, 2133 (2003).

177. Cryle, M.J., Espinoza, R.D., Smith, S.J., Matovic, N.J. & De Voss, J.J. Are branched chain fatty acids the natural substrates for P450BM3? *Chemical Communications*, 2353-2355 (2006).

178. Narhi, L.O. & Fulco, A.J. CHARACTERIZATION OF A CATALYTICALLY SELF-SUFFICIENT 119,000-DALTON CYTOCHROME-P-450 MONOOXYGENASE INDUCED BY BARBITURATES IN BACILLUS-MEGATERIUM. *Journal of Biological Chemistry* 261, 7160-7169 (1986).

179. Venkataraman, H., Beer, S.B.A.d., Geerke, D.P., Vermeulen, N.P.E. & Commandeur, J.N.M. Regio- and Stereoselective Hydroxylation of Optically Active α-Ionone Enantiomers by Engineered Cytochrome P450 BM3 Mutants. *Advanced Synthesis & Catalysis* 354, 2172-2184 (2012).

180. Sideri, A., Goyal, A., Di Nardo, G., Tsotsou, G.E. & Gilardi, G. Hydroxylation of non-substituted polycyclic aromatic hydrocarbons by cytochrome P450 BM3 engineered by directed evolution. *J Inorg Biochem* 120, 1-7 (2013).

181. Sawayama, A.M. et al. A Panel of Cytochrome P450 BM3 Variants to Produce Drug Metabolites and Diversify Lead Compounds. *Chemistry-a European Journal* 15, 11723-11729 (2009).

182. Neufeld, K., Marienhagen, J., Schwaneberg, U. & Pietruszka, J. Benzylic hydroxylation of aromatic compounds by P450 BM3. *Green Chemistry* 15, 2408-2421 (2013).

183. Dennig, A., Lulsdorf, N., Liu, H. & Schwaneberg, U. Regioselective o-hydroxylation of monosubstituted benzenes by P450 BM3. *Angew Chem Int Ed Engl* 52, 8459-62 (2013).

184. Dietrich, M., Do, T.A., Schmid, R.D., Pleiss, J. & Urlacher, V.B. Altering the regioselectivity of the subterminal fatty acid hydroxylase P450 BM-3 towards gamma- and delta-positions. *J Biotechnol* 139, 115-7 (2009).

185. Seifert, A. et al. Rational Design of a Minimal and Highly Enriched CYP102A1 Mutant Library with Improved Regio-, Stereo- and Chemoselectivity. *Chembiochem* 10, 853-861 (2009).

186. Lewis, J.C. et al. Combinatorial Alanine Substitution Enables Rapid Optimization of Cytochrome P450(BM3) for Selective Hydroxylation of Large Substrates. *Chembiochem* 11, 2502-2505 (2010).

187. Chang, D., Witholt, B. & Li, Z. Preparation of (S)-N-substituted 4-hydroxy-pyrrolidin-2-ones by regio- and stereoselective hydroxylation with Sphingomonas sp. HXN-200. *Org Lett* 2, 3949-52 (2000).

188. Zhang, W., Tang, W.L., Wang, Z.S. & Li, Z. Regio- and Stereoselective Biohydroxylations with a Recombinant Escherichia coli Expressing P450(pyr) Monooxygenase of Sphingomonas Sp. HXN-200. *Advanced Synthesis & Catalysis* 352, 3380-3390 (2010).

189. Li, Z. & Chang, D. Recent Advances in Regio- and Stereoselective Biohydroxylation of Non- Activated Carbon Atoms. *Current Organic Chemistry* 8, 1647-1658 (2004).

190. Li, Z., Feiten, H.-J., van Beilen, J.B., Duetz, W. & Witholt, B. Preparation of optically active N-benzyl-3-hydroxypyrrolidine by enzymatic hydroxylation. *Tetrahedron: Asymmetry* 10, 1323-1333 (1999).

191. Zhang, W., Tang, W.L., Wang, Z. & Li, Z. Regio- and Stereoselective Biohydroxylations with a Recombinant Escherichia coli Expressing P450pyr Monooxygenase of Sphingomonas Sp. HXN-200. *Advanced Synthesis & Catalysis* 352, 3380-3390 (2010).

192. Li, A., Liu, J., Pham, S.Q. & Li, Z. Engineered P450pyr monooxygenase for asymmetric epoxidation of alkenes with unique and high enantioselectivity. *Chemical Communications* 49, 11572-11574 (2013).

193. Tracewell, C.A. & Arnold, F.H. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Current Opinion in Chemical Biology* 13, 3-9 (2009).

194. Zhao, H.M., Chockalingam, K. & Chen, Z.L. Directed evolution of enzymes and pathways for industrial biocatalysis. *Current Opinion in Biotechnology* 13, 104-110 (2002).

195. Arnold, F.H. Protein engineering for unusual environments. *Curr Opin Biotechnol* 4, 450-5 (1993).

196. Chen, K. & Arnold, F.H. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc Natl Acad Sci U S A* 90, 5618-22 (1993).

197. Tobin, M.B., Gustafsson, C. & Huisman, G.W. Directed evolution: the 'rational' basis for 'irrational' design. *Current Opinion in Structural Biology* 10, 421-427 (2000).

198. Marti, S. et al. Computational design of biological catalysts. *Chem Soc Rev* 37, 2634-43 (2008).

199. Ward, T.R. Artificial enzymes made to order: combination of computational design and directed evolution. *Angew Chem Int Ed Engl* 47, 7802-3 (2008).

200. Marshall, J. Proteins made to order. *Nature* (2012).

201. Dietrich, J.A. et al. A Novel Semi-biosynthetic Route for Artemisinin Production Using Engineered Substrate-Promiscuous P450(BM3). *Acs Chemical Biology* 4, 261-267 (2009).

202. Turner, N.J. Directed evolution of enzymes for applied biocatalysis. *Trends in Biotechnology* 21, 474-478 (2003).

203. Roodveldt, C., Aharoni, A. & Tawfik, D.S. Directed evolution of proteins for heterologous expression and stability. *Current Opinion in Structural Biology* 15, 50-56 (2005).

204. Hibbert, E.G. et al. Directed evolution of biocatalytic processes. *Biomol Eng* 22, 11-9 (2005).

205. Reetz, M.T. & Carballeira, J.D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protoc* 2, 891-903 (2007).

206. Powell, K.A. et al. Directed evolution and biocatalysis. *Angewandte Chemie-International Edition* 40, 3948-3959 (2001).

207. Akanuma, S., Yamagishi, A., Tanaka, N. & Oshima, T. Further improvement of the thermal stability of a partially stabilized Bacillus subtilis 3-isopropylmalate dehydrogenase variant by random and site-directed mutagenesis. *European Journal of Biochemistry* 260, 499-504 (1999).

208. Cherry, J.R. et al. Directed evolution of a fungal peroxidase. *Nature Biotechnology* 17, 379-384 (1999).

209. Parikh, M.R. & Matsumura, I. Site-saturation mutagenesis is more efficient than DNA shuffling for the directed evolution of beta-fucosidase from beta-galactosidase. *Journal of Molecular Biology* 352, 621-628 (2005).

210. Bougioukou, D.J., Kille, S., Taglieber, A. & Reetz, M.T. Directed Evolution of an Enantioselective Enoate-Reductase: Testing the Utility of Iterative Saturation Mutagenesis. *Advanced Synthesis & Catalysis* 351, 3287-3305 (2009).

211. Reetz, M.T., Bocola, M., Carballeira, J.D., Zha, D.X. & Vogel, A. Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. *Angewandte Chemie-International Edition* 44, 4192-4196 (2005).

212. Reetz, M.T., Wang, L.W. & Bocola, M. Directed evolution of enantioselective enzymes: Iterative cycles of CASTing for probing protein-sequence space. *Angewandte Chemie-International Edition* 45, 1236-1241 (2006).

213. Zhang, L. & Demain, A.L. *Natural products : drug discovery and therapeutic medicine*, xiv, 382 p. (Humana Press, Totowa, N.J., 2005).

214. Wahler, D. & Reymond, J.L. High-throughput screening for biocatalysts. *Current Opinion in Biotechnology* 12, 535-544 (2001).

215. Dar, Y.L. High-throughput experimentation: A powerful enabling technology for the chemicals and materials industry. *Macromolecular Rapid Communications* 25, 34-47 (2004).

216. Leemhuis, H., Kelly, R.M. & Dijkhuizen, L. Directed evolution of enzymes: Library screening strategies. *IUBMB Life* 61, 222-8 (2009).

217. Martinez, R. & Schwaneberg, U. A roadmap to directed enzyme evolution and screening systems for biotechnological applications. *Biol Res* 46, 395-405 (2013).

218. Reetz, M.T. New methods for the high-throughput screening of enantioselective catalysts and biocatalysts. *Angewandte Chemie-International Edition* 41, 1335-1338 (2002).

219. Reetz, M.T. Combinatorial and evolution-based methods in the creation of enantioselective catalysts. *Angewandte Chemie-International Edition* 40, 284-310 (2001).

220. Muller, C.A., Markert, C., Teichert, A.M. & Pfaltz, A. Mass spectrometric screening of chiral catalysts and catalyst mixtures. *Chemical Communications*, 1607-1618 (2009).

221. Goddard, J.P. & Reymond, J.L. Enzyme assays for high-throughput screening. *Current Opinion in Biotechnology* 15, 314-322 (2004).

222. Otten, L.G., Hollmann, F. & Arends, I.W.C.E. Enzyme engineering for enantioselectivity: from trial-and-error to rational design? *Trends in Biotechnology* 28, 46-54 (2010).

223. Millot, N. et al. Rapid determination of enantiomeric excess using infrared thermography. *Organic Process Research & Development* 6, 463-470 (2002).

224. Abato, P. & Seto, C.T. EMDee: An enzymatic method for determining enantiomeric excess. *Journal of the American Chemical Society* 123, 9206-9207 (2001).

225. Li, Z., Butikofer, L. & Witholt, B. High-throughput measurement of the enantiomeric excess of chiral alcohols by using two enzymes. *Angewandte Chemie-International Edition* 43, 1698-1702 (2004).

226. Tang, W.L. National University of Singapore and University of Illinois, at Urbana-Champaign (2011).

227. Chen, Y.Z., Tang, W.L., Mou, J. & Li, Z. High-Throughput Method for Determining the Enantioselectivity of Enzyme-Catalyzed Hydroxylations Based on Mass Spectrometry. *Angewandte Chemie-International Edition* 49, 5278-5283 (2010).

228. Lentz, O. et al. Altering the regioselectivity of cytochrome P450CYP102A3 of Bacillus subtilis by using a new versatile assay system. *Chembiochem* 7, 345-350 (2006).

229. Meinhold, P., Peters, M.W., Hartwick, A., Hernandez, A.R. & Arnold, F.H. Engineering Cytochrome P450 BM3 for Terminal Alkane Hydroxylation. *Advanced Synthesis & Catalysis* 348, 763-772 (2006).

230. Lewis, J.C. & Arnold, F.H. Catalysts on Demand: Selective Oxidations by Laboratory-Evolved Cytochrome P450 BM3. *Chimia* 63, 309-312 (2009).

231. Kille, S., Zilly, F.E., Acevedo, J.P. & Reetz, M.T. Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nature Chemistry* 3, 738-743 (2011).

232. Panke, S. & Wubbolts, M. Advances in biocatalytic synthesis of pharmaceutical intermediates. *Curr Opin Chem Biol* 9, 188-94 (2005).

233. Breuer, M. et al. Industrial methods for the production of optically active intermediates. *Angew Chem Int Ed Engl* 43, 788-824 (2004).

234. Lewis, J.C., Coelho, P.S. & Arnold, F.H. Enzymatic functionalization of carbon-hydrogen bonds. *Chem Soc Rev* 40, 2003-21 (2011).

235. Kamata, K., Yonehara, K., Nakagawa, Y., Uehara, K. & Mizuno, N. Efficient stereo- and regioselective hydroxylation of alkanes catalysed by a bulky polyoxometalate. *Nat Chem* 2, 478-83 (2010).

236. Chen, M.S. & White, M.C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science* 318, 783-7 (2007).

237. Periana, R.A. et al. Platinum catalysts for the high-yield oxidation of methane to a methanol derivative. *Science* 280, 560-4 (1998).

238. van Beilen, J.B. & Funhoff, E.G. Alkane hydroxylases involved in microbial alkane degradation. *Appl Microbiol Biotechnol* 74, 13-21 (2007).

239. van Beilen, J.B. & Funhoff, E.G. Expanding the alkane oxygenase toolbox: new enzymes and applications. *Curr Opin Biotechnol* 16, 308-14 (2005).

240. Whitehouse, C.J.C., Bell, S.G. & Wong, L.-L. P450BM3 (CYP102A1): connecting the dots. *Chemical Society Reviews* 41, 1218-1260 (2012).

241. Choi, K.Y. et al. Cloning, expression and characterization of CYP102D1, a self-sufficient P450 monooxygenase from Streptomyces avermitilis. *FEBS J* 279, 1650-62 (2012).

242. Merkx, M. et al. Dioxygen Activation and Methane Hydroxylation by Soluble Methane Monooxygenase: A Tale of Two Irons and Three Proteins A list of abbreviations can be found in Section 7. *Angew Chem Int Ed Engl* 40, 2782-2807 (2001).

243. Borodina, E., Nichol, T., Dumont, M.G., Smith, T.J. & Murrell, J.C. Mutagenesis of the "leucine gate" to explore the basis of catalytic versatility in soluble methane monooxygenase. *Appl Environ Microbiol* 73, 6460-7 (2007).

244. Austin, R.N., Chang, H.K., Zylstra, G.J. & Groves, J.T. The non-heme diiron alkane monooxygenase of Pseudomonas oleovorans (AlkB) hydroxylates via a substrate radical intermediate. *Journal of the American Chemical Society* 122, 11747-11748 (2000).

245. Bottcher, D. & Bornscheuer, U.T. Protein engineering of microbial enzymes. *Curr Opin Microbiol* 13, 274-82 (2010).

246. Scheps, D., Honda Malca, S., Hoffmann, H., Nestl, B.M. & Hauer, B. Regioselective [small omega]-hydroxylation of medium-chain n-alkanes and primary alcohols by CYP153 enzymes from Mycobacterium marinum and Polaromonas sp. strain JS666. *Organic & Biomolecular Chemistry* 9, 6727-6733 (2011).

247. Lentz, O. et al. Altering the regioselectivity of cytochrome P450 CYP102A3 of Bacillus subtilis by using a new versatile assay system. *Chembiochem* 7, 345-50 (2006).

248. Reetz, M.T. Directed evolution of enantioselective enzymes: an unconventional approach to asymmetric catalysis in organic chemistry. *J Org Chem* 74, 5767-78 (2009).

249. Tee, K.L. & Schwaneberg, U. Directed evolution of oxygenases: screening systems, success stories and challenges. *Comb Chem High Throughput Screen* 10, 197-217 (2007).

250. Bornscheuer, U.T. et al. Engineering the third wave of biocatalysis. *Nature* 485, 185-94 (2012).

251. Reetz, M.T. Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. *Angew Chem Int Ed Engl* 50, 138-74 (2011).

252. Bornscheuer, U. Protein Engineering as a Tool for the Development of Novel Bioproduction Systems. 1-16 (Springer Berlin Heidelberg, 2013).

253. Cirino, P.C. & Arnold, F.H. Protein engineering of oxygenases for biocatalysis. *Curr Opin Chem Biol* 6, 130-5 (2002).

254. Davids, T., Schmidt, M., Bottcher, D. & Bornscheuer, U.T. Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr Opin Chem Biol* 17, 215-20 (2013).

255. Whitehouse, C.J. et al. A highly active single-mutation variant of P450BM3 (CYP102A1). *Chembiochem* 10, 1654-6 (2009).

256. Whitehouse, C.J. et al. Evolved CYP102A1 (P450BM3) variants oxidise a range of non-natural substrates and offer new selectivity options. *Chem Commun (Camb)*, 966-8 (2008).

257. Pazmiño, D.E.T., Snajdrova, R., Rial, D.V., Mihovilovic, M.D. & Fraaije, M.W. Altering the Substrate Specificity and Enantioselectivity of Phenylacetone Monooxygenase by Structure-Inspired Enzyme Redesign. *Advanced Synthesis & Catalysis* 349, 1361-1368 (2007).

258. Lewis, J.C. et al. Combinatorial alanine substitution enables rapid optimization of cytochrome P450BM3 for selective hydroxylation of large substrates. *Chembiochem* 11, 2502-5 (2010).

259. Landwehr, M. et al. Enantioselective alpha-hydroxylation of 2-arylacetic acid derivatives and buspirone catalyzed by engineered cytochrome P450 BM-3. *J Am Chem Soc* 128, 6058-9 (2006).

260. Kille, S., Zilly, F.E., Acevedo, J.P. & Reetz, M.T. Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat Chem* 3, 738-43 (2011).

261. Zhang, K., Shafer, B.M., Demars, M.D., 2nd, Stern, H.A. & Fasan, R. Controlled oxidation of remote sp3 C-H bonds in artemisinin via P450 catalysts with fine-tuned regio- and stereoselectivity. *J Am Chem Soc* 134, 18695-704 (2012).

262. Halsey, K.H., Sayavedra-Soto, L.A., Bottomley, P.J. & Arp, D.J. Site-directed amino acid substitutions in the hydroxylase alpha subunit of butane monooxygenase from Pseudomonas butanovora: Implications for substrates knocking at the gate. *J Bacteriol* 188, 4962-9 (2006).

263. Glieder, A., Farinas, E.T. & Arnold, F.H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat Biotechnol* 20, 1135-9 (2002).

264. Bell, S.G. et al. Butane and propane oxidation by engineered cytochrome P450cam. *Chem Commun (Camb)*, 490-1 (2002).

265. Li, Z. & Witholt, B. Practical biohydroxylation with sphingomonas sp HXN-200, a highly active, regio and stereos elective, and easy to handle biocatalyst. *Abstracts of Papers of the American Chemical Society* 225, U213-U213 (2003).

266. Li, Z. et al. Preparation of (R)- and (S)-N-protected 3-hydroxypyrrolidines by hydroxylation with Sphingomonas sp. HXN-200, a highly active, regio- and stereoselective, and easy to handle biocatalyst. *J Org Chem* 66, 8424-30 (2001).

267. Chang, D.L., Feiten, H.J., Witholt, B. & Li, Z. Regio- and stereoselective hydroxylation of N-substituted piperidin-2-ones with Sphingomonas sp HXN-200. *Tetrahedron-Asymmetry* 13, 2141-2147 (2002).

268. Chang, D. et al. Highly enantioselective hydrolysis of alicyclic meso-epoxides with a bacterial epoxide hydrolase from Sphingomonas sp. HXN-200: simple syntheses of alicyclic vicinal trans-diols. *Chem Commun (Camb)*, 960-1 (2003).

269. Chang, D. et al. Practical syntheses of N-substituted 3-hydroxyazetidines and 4-hydroxypiperidines by hydroxylation with Sphingomonas sp. HXN-200. *Org Lett* 4, 1859-62 (2002).

270. Yamamoto, H., Kawada, K., Matsuyama, A. & Kobayashi, Y. Cloning and expression in Escherichia coli of a gene coding for a secondary alcohol dehydrogenase from Candida parapsilosis. *Bioscience Biotechnology and Biochemistry* 63, 1051-1055 (1999).

271. Yamamoto, H. & Kudoh, M. Novel chiral tool, (R)-2-octanol dehydrogenase, from Pichia finlandica: purification, gene cloning, and application for optically active alpha-haloalcohols. *Appl Microbiol Biotechnol* 97, 8087-96 (2013).

272. Schep, L.J., Knudsen, K., Slaughter, R.J., Vale, J.A. & Megarbane, B. The clinical toxicology of gamma-hydroxybutyrate, gamma-butyrolactone and 1,4-butanediol. *Clin Toxicol (Phila)* 50, 458-70 (2012).

273. Shannon, M. & Quang, L.S. Gamma-hydroxybutyrate, gamma-butyrolactone, and 1, 4-butanediol: a case report and review of the literature. *Pediatric emergency care* 16, 435-440 (2000).

274. Seitz, M. & Reiser, O. Synthetic approaches towards structurally diverse gamma-butyrolactone natural-product-like compounds. *Curr Opin Chem Biol* 9, 285-92 (2005).

275. Barikani, M., Honarkar, H. & Barikani, M. Synthesis and Characterization of Polyurethane Elastomers Based on Chitosan and Poly(epsilon-caprolactone). *Journal of Applied Polymer Science* 112, 3157-3165 (2009).

276. Ding, X.M., Hu, J.L., Tao, X.M. & Hu, C.R. Preparation of temperature-sensitive polyurethanes for smart textiles. *Textile Research Journal* 76, 406-413 (2006).

277. Yi, J., Boyce, M.C., Lee, G.F. & Balizer, E. Large deformation rate-dependent stress-strain behavior of polyurea and polyurethanes. *Polymer* 47, 319-329 (2006).

278. Anonymous. BUTANEDIOL. *Chemical Week* 168, 45 (2006).

279. Rose, J. Industrial Organic Chemistry Klaus Weissermel and Hans-Jurgen Arpe 3rd edn. VCH, Weinheim, 1997 xvii + 464 pages. £70 ISBN 3-527-28838-4. *Applied Organometallic Chemistry* 13, 857-858 (1999).

280. Burdick, D.L. & Leffler, W.L. *Petrochemical in Nontechnical Language*, (PennWell Books, 2010).

281. Pesterfield, L. The 100 Most Important Chemical Compounds: A Reference Guide (by Richard L. Myers). *Journal of Chemical Education* 86, 1182 (2009).

282. Colby, J., Stirling, D.I. & Dalton, H. The soluble methane mono-oxygenase of Methylococcus capsulatus (Bath). Its ability to oxygenate n-alkanes, n-alkenes, ethers, and alicyclic, aromatic and heterocyclic compounds. *Biochem J* 165, 395-402 (1977).

283. Lan, E.I. & Liao, J.C. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metab Eng* 13, 353-63 (2011).

284. Engel, P., Haas, T., Pfeffer, J.C., Thum, O. & Gehring, C. Process for producing alpha, omega-diols from alkanes or 1 -alkanols employing a cyp153 alkane hydroxylase. (Google Patents, 2014).

285. Lauchli, R. et al. High-Throughput Screening for Terpene-Synthase-Cyclization Activity and Directed Evolution of a Terpene Synthase. *Angewandte Chemie-International Edition* 52, 5571-5574 (2013).

286. Reetz, M.T., Kahakeaw, D. & Lohmer, R. Addressing the numbers problem in directed evolution. *Chembiochem* 9, 1797-1804 (2008).

287. Sali, A. & Blundell, T.L. COMPARATIVE PROTEIN MODELING BY SATISFACTION OF SPATIAL RESTRAINTS. *Journal of Molecular Biology* 234, 779-815 (1993).

288. Pettersen, E.F. et al. UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605-1612 (2004).

289. Shen, M.-Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Science* 15, 2507-2524 (2006).

290. Berendsen, H.J.C., Vanderspoel, D. & Vandrunen, R. GROMACS - A MESSAGE-PASSING PARALLEL MOLECULAR-DYNAMICS IMPLEMENTATION. *Computer Physics Communications* 91, 43-56 (1995).

291. Hess, B., Bekker, H., Berendsen, H.J.C. & Fraaije, J. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 18, 1463-1472 (1997).

292. Daura, X. et al. Peptide folding: When simulation meets experiment. *Angewandte Chemie-International Edition* 38, 236-240 (1999).

293. Morris, G.M. et al. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry* 30, 2785-2791 (2009).

294. Trott, O. & Olson, A.J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry* 31, 455-461 (2010).

295. Shaik, S. et al. P450 Enzymes: Their Structure, Reactivity, and Selectivity-Modeled by QM/MM Calculations. *Chemical Reviews* 110, 949-1017 (2010).

296. Zhang, Y., Morisetti, P., Kim, J., Smith, L. & Lin, H. Regioselectivity preference of testosterone hydroxylation by cytochrome P450 3A4. *Theoretical Chemistry Accounts* 121, 313-319 (2008).

297. Peters, M.W., Meinhold, P., Glieder, A. & Arnold, F.H. Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *J Am Chem Soc* 125, 13442-50 (2003).

298. Avigad, G. A simple spectrophotometric determination of formaldehyde and other aldehydes: application to periodate-oxidized glycol systems. *Anal Biochem* 134, 499-504 (1983).

299. Li, Z., Teng, H. & Xiu, Z. Aqueous two-phase extraction of 2,3-butanediol from fermentation broths using an ethanol/ammonium sulfate system. *Process Biochemistry* 45, 731-737 (2010).

300. Veisi, H. & Ghorbani-Vaghei, R. Recent progress in the application of N-halo reagents in the synthesis of heterocyclic compounds. *Tetrahedron* 66, 7445-7463 (2010).

301. Wu, X.Y., Mahalingam, A.K. & Alterman, M. Rapid Mo(CO)(6) catalysed one-pot deoxygenation of heterocyclic halo-benzyl alcohols with Lawesson's reagent. *Tetrahedron Letters* 46, 1501-1504 (2005).

302. Li, Z. et al. Oxidative biotransformations using oxygenases. *Current Opinion in Chemical Biology* 6, 136-144 (2002).

303. Yoon, Y.A. et al. Novel 1H-pyrrolo 2,3-c pyridines as acid pump antagonists (APAs). *Bioorganic & Medicinal Chemistry Letters* 20, 5237-5240 (2010).

304. McKerrecher, D. et al. Discovery, synthesis and biological evaluation of novel glucokinase activators. *Bioorg Med Chem Lett* 15, 2103-6 (2005).

305. Baker, S.J. et al. Discovery of a new boron-containing antifungal agent, 5-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborole (AN2690), for the potential treatment of onychomycosis. *Journal of Medicinal Chemistry* 49, 4447-4450 (2006).

306. Jing, Q. et al. An Accessible Chiral Linker to Enhance Potency and Selectivity of Neuronal Nitric Oxide Synthase Inhibitors. *ACS Med Chem Lett* 5, 56-60 (2014).

307. Pyring, D. et al. Design and synthesis of potent C-2-symmetric diol-based HIV-1 protease inhibitors: Effects of fluoro substitution. *Journal of Medicinal Chemistry* 44, 3083-3091 (2001).

308. Thurmond, J. et al. Synthesis and biological evaluation of novel 2,4-diaminoquinazoline derivatives as SMN2 promoter activators for the potential treatment of spinal muscular atrophy. *Journal of Medicinal Chemistry* 51, 449-469 (2008).

309. Wender, I., Greenfield, H., Metlin, S. & Orchin, M. CHEMISTRY OF THE OXO AND RELATED REACTIONS .6. EXPERIMENTS WITH META-SUBSTITUTED AND PARA-SUBSTITUTED BENZYL ALCOHOLS. *Journal of the American Chemical Society* 74, 4079-4083 (1952).

310. Zhang, Y., Zang, H. & Cheng, B.-W. A green and highly efficient protocol for hydrolysis of substituted benzyl chloride using ionic liquid bmim BF(4) as phase-transfer catalyst. in *Environment Materials and Environment Management Pts 1-3*, Vol. 113-116 (eds. Du, Z.Y. & Sun, X.B.) 1990-1992 (2010).

311. Hu, Y.L., Jiang, H., Zhu, J. & Lu, M. Facile and efficient hydrolysis of organic halides, epoxides, and esters with water catalyzed by ferric sulfate in a PEG(1000)-DAIL BF4 /toluene temperature-dependent biphasic system. *New Journal of Chemistry* 35, 292-298 (2011).

312. Che, C.-M., Lo, V.K.-Y., Zhou, C.-Y. & Huang, J.-S. Selective functionalisation of saturated C-H bonds with metalloporphyrin catalysts. *Chemical Society Reviews* 40, 1950-1975 (2011).

313. Lu, H. & Zhang, X.P. Catalytic C-H functionalization by metalloporphyrins: recent developments and future directions. *Chemical Society Reviews* 40, 1899-1909 (2011).

314. Laali, K.K., Herbert, M., Cushnyr, B., Bhatt, A. & Terrano, D. Benzylic oxidation of aromatics with cerium(IV) triflate; synthetic scope and mechanistic insight. *Journal of the Chemical Society-Perkin Transactions 1*, 578-583 (2001).

315. Lewis, J.C., Coelho, P.S. & Arnold, F.H. Enzymatic functionalization of carbon-hydrogen bonds. *Chemical Society Reviews* 40, 2003-2021 (2011).

316. Urlacher, V.B. & Girhard, M. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends in Biotechnology* 30, 26-36 (2012).

317. Hardcastle, I.R. et al. Discovery of potent chromen-4-one inhibitors of the DNA-dependent protein kinase (DNA-PK) using a small-molecule library approach. *Journal of Medicinal Chemistry* 48, 7829-7846 (2005).

318. Wissner, A. et al. 2-(Quinazolin-4-ylamino)- 1,4 benzoquinones as covalent-binding, irreversible inhibitors of the kinase domain of vascular endothelial growth factor receptor-2. *Journal of Medicinal Chemistry* 48, 7560-7581 (2005).

319. Lacotte, P., Puente, C. & Ambroise, Y. Synthesis and Evaluation of 3,4-Dihydropyrimidin-2(1H)-ones as Sodium Iodide Symporter Inhibitors. *Chemmedchem* 8, 104-111 (2013).

320. Pisani, L. et al. Discovery of a Novel Class of Potent Coumarin Monoamine Oxidase B Inhibitors: Development and Biopharmacological Profiling of 7- (3-Chlorobenzyl)oxy -4-(methylamino)methyl -2H-chromen-2-one Methanesulfonate (NW-1772) as a Highly Potent, Selective, Reversible, and Orally Active Monoamine Oxidase B Inhibitor. *Journal of Medicinal Chemistry* 52, 6685-6706 (2009).

321. Han, Y.T. et al. Ligand-Based Design, Synthesis, and Biological Evaluation of 2-Aminopyrimidines, a Novel Series of Receptor for Advanced Glycation End Products (RAGE) Inhibitors. *Journal of Medicinal Chemistry* 55, 9120-9135 (2012).

322. Marsault, E. et al. Discovery of a new class of macrocyclic antagonists to the human motilin receptor. *Journal of Medicinal Chemistry* 49, 7190-7197 (2006).

323. Nakazato, A. et al. Synthesis, in vitro pharmacology, structure-activity relationships, and pharmacokinetics of 3-alkoxy-2-amino-6-fluorobicyclo 3.1.0 hexane-2,6-dicarboxylic acid derivatives as potent and selective group II metabotropic glutamate receptor antagonists. *Journal of Medicinal Chemistry* 47, 4570-4587 (2004).

324. Gomtsyan, A. et al. Novel transient receptor potential vanilloid 1 receptor antagonists for the treatment of pain: Structure-activity relationships for ureas with quinoline, isoquinoline, quinazoline, phthalazine, quinoxaline, and cinnoline moieties. *Journal of Medicinal Chemistry* 48, 744-752 (2005).

Appendices

List of Publications

- Yi Yang; Ji Liu; Zhi Li. Engineering of P450pyr hydroxylase for highly regio- and enantioselective subterminal hydroxylation of alkane. *Angew. Chem. Int. Ed.*, 2014, *53*, 3120-3124. (Highlighted as a Very Important Paper and at a back cover)
- Yi Yang, Yu Tse Chi, Hui Hung Toh, Zhi Li, Evolving P450pyr monooxygenase for highly regioselective terminal hydroxylation of n-butanol to 1,4-butanediol. *Chem. Commun.*, 2015, *51*, 514-517.
- Yi Yang, Zhi Li, Evolving P450pyr monooxygenase for regio- and stereoselective hydroxylations. *Chimia*, 2015, *69*, 136-141.
- Li Juan Ye, Hui Hung Toh, Yi Yang, Joseph P. Adams, Radka Snajdrova, Zhi Li, Engineering of amine dehydrogenase for asymmetric reductive amination of ketone by evolving Rhodococcus phenylalanine dehydrogenase. *ACS Catal.* 2015, *5*, 1119-1122.
- Akbar Vahidi K., Yi Yang, Thao P.N. Ngo, Zhi Li, Simple and efficient immobilization of extra-cellular his-tagged enzyme directly from cell culture supernatant as active and recyclable nanobiocatalyst: high performance production of biodiesel from waste grease. *ACS Catal.* 2015, *5*, 3157-3161.
- Yi Yang, Hui Hung Toh, Ji Liu, Joseph P. Adams, Radka Snajdrova, Zhi Li, Benzylic hydroxylation of fluoro- and other halo-substituted toluenes with engineered P450pyr. (Manuscript in preparation)

List of Presentation

- Regio- and stereo-selective subterminal biohydroxylation of non-activated carbon atoms, Poster Presentation, *14th Asia Pacific Confederation of Chemical Engineering Congress*, Singapore, 2012
- Directed evolution of P450pyr monooxygenase for regio- and stereo-selective biohydroxylations, Poster Presentation, *3rd Asian Symposium on Innovative Bio-production*, Singapore, 2013
- Engineering of P450pyr monooxygenase to create new enzymes for stereoselective hydroxylation of non-activated carbon atom at subterminal positions, Oral presentation, *GSM Symposium 2013*, Singapore, 2013
- Directed evolution of P450pyr monooxygenase for regio- and stereo-selective biohydroxylations, Poster Presentation, *6th Singapore Catalysis Forum*, Singapore, 2013
- Engineering of P450pyr monooxygenase for regio- and stereoselective subterminal hydroxylation, Poster Presentation, *Biotrans2013*, Manchester, UK, 2013
- Engineering of P450pyr hydroxylase for the highly regio- and enantioselective subterminal hydroxylation of alkanes, Poster Presentation, *Biosystems Design 1.0*, Singapore, 2015
- Evolving P450pyr monooxygenase for highly regioselective terminal hydroxylation of *n*-butanol to 1,4-butanediol, Poster Presentation. *Southeast Asia Catalysis Conference*, Singapore, 2015 (best poster award (silver prize))