Improving the operational stability of the alkane hydroxylating cytochrome P450: CYP153A6

By

Chéri Louise Jacobs

Submitted in fulfilment of the requirements for the degree

Magister Scientiae

In the Faculty of Natural and Agricultural Sciences

Department of Microbial, Biochemical and Food Biotechnology

University of the Free State

February 2015

Supervisor: Prof. M. S. Smit

Co-supervisor: Dr. D. J. Opperman

Acknowledgements

I would like to express my sincerest gratitude towards:

My Heaven Father, for blessing me with this opportunity and giving me strength to complete this degree. Through Him all things are possible.

My supervisor, **Prof. M. S. Smit**, for her guidance, support and patience – without her this study would not have been possible.

My co-supervisor, **Dr. D. J. Opperman**, for his constant motivation and enthusiasm.

DST-NRF Centre of Excellence in Catalysis (c*change) for the financial support of this project. Opinions and expressed and conclusions arrived at, are those of the author and are not necessarily attributed to c*change

My father **Shaun**, my mother **Moira** and my sister **Tarryn** who stood behind me very step of the way, thank you for your unchanging love and support and to my beloved departed sister **Meagan** I know that you are always watching over me, thank you for encouraging me to follow my dreams.

My **friends** and lab colleagues at the Biocatalysis research group, thank you for all your help, encouragement and laughter, especially when times were hard.

Table of Contents

List of Abbreviations I

Chapter 1: Literature review: Bacterial alkane hydroxylases – the CYP153 family 1

1.1 Introduction to cytochrome P450s 1

1.2 Microbial alkane hydroxylases involved in n-alkane degradation 6

1.2.1 C_1-C_4 (Methane monooxygenases) 6

1.2.2 C_5-C_{16} (Integral membrane non-heme iron or cytochrome P450 monooxygenases) 7

1.2.3 C_{17}< (Flavin-containing oxygenases) 7

1.3 CYP153 Hydroxylases 7

1.3.1 Identification and expression of CYP153 8

1.3.1.1 Discovery of the first CYP153 hydroxylase 8

1.3.1.2 Identification, cloning and expression of CYP153 genes 8

1.3.2 Characterisation and reactions catalysed 15

1.3.2.1 Reactions catalyses by CYP153 hydroxylases 15

1.3.2.2 CYP153A6 20

1.3.2.3 Substrate binding studies 22

1.3.2.4 Insight into the mechanism of CYP153 hydroxylases 24

1.3.3 Structure of CYP153 hydroxylases 25

1.3.4 Diversity of CYP153 hydroxylases in microorganisms and environments 27

1.3.5 Applications of CYP153 hydroxylases 29

1.4 Concluding remarks 32

1.5 Aim of the Study 33

Chapter 2: Factors affecting the operational stability of CYP153A6 in cell free extracts (CFE) 35

2.1 Material and methods 35

Section A - General methods

2.1.1 Bacterial strains and plasmids 35

2.1.2 Protein expression 36

2.1.3 Cell harvesting and cell disruption 36

2.1.4 Analysis of expression 37

2.1.4.1 Spectroscopic enzyme quantification 37

2.1.4.2 SDS-PAGE analysis 37

2.1.5 Biotransformations of *n*-octane using cell free extracts 38

2.1.6 Sample extraction and product analysis 38

Section B – Different experiments to investigate the effect of reaction conditions (continuation of section 2.1.5)

2.1.5.1 Experiments using glucose dehydrogenase for cofactor regeneration 39

2.1.5.2 BRM prepared with additional Ferredoxin Reductase and Ferredoxin 40

2.1.5.3 Controlling P450 concentration, storing P450 overnight and reactions carried out in the presence of in situ generated H_2O_2 41

2.1.5.4 Biotransformation reaction buffer 42

2.1.5.5 Concentration of P450 in biotransformation reaction mixture (BRM) 43

2.2 Results 44

2.2.1 Analysis of expressed proteins 44

2.2.1.1 Spectroscopic quantification of P450 content 44

2.2.1.2 SDS-PAGE analysis 44

2.2.2 *n*-Octane biotransformations using cell free extracts containing CYP153A6 47

2.2.2.1 Biotransformations containing glucose dehydrogenase as a cofactor regenerating system for CYP153A6 47

2.2.2.2 Biotransformations containing additional FdR/Fdx to aid electron transfer from NADH to CYP153A6 51

2.2.2.3 Evaluating the effect of H_2O_2, temperature and storage of CFEs on CYP153A6 stability and activity 52

2.2.2.4 Evaluating different buffers, buffer concentrations and buffer pH to optimise biotransformation of *n*-octane using CYP153A6 56

2.2.2.5 Evaluating different concentrations of CYP153A6 in the biotransformation reaction mixture to optimise biotransformation of *n*-octane 60

2.2.3 Conclusion 62

Chapter 3: Design, construction and evaluation of CYP153A6 mutants for improved operational stability 64

3.1 Materials and methods 64

3.1.1 Bacterial strains and plasmids 64

3.1.2 Site-directed mutagenesis to construct CYP153A6 mutants 65

3.1.2.1 Designing CYP153A6 mutants 65

3.1.2.1.1 3DM Information System for Cytochrome P450s created by Bio-Prodict 65

3.1.2.1.2 YASARA (Yet Another Scientific Artificial Reality Application) 66

3.1.2.1.3 Primer design 66

3.1.2.2 Polymerase Chain Reaction (PCR) amplification 67
3.1.2.2.1 Megaprimer PCR amplification 67
3.1.2.3 Visualisation and Purification of PCR and restriction enzyme digestion products 69
3.1.2.4 Restriction enzyme digestion and Ligation 69
3.1.2.5 DNA sequencing 72
3.1.2.6 Transformations 72
3.1.2.7 Plasmid proliferation and extraction 72
3.1.2.8 Transforming *E. coli* BL21-Gold (DE3) for protein expression 72
3.1.3 Protein expression 73
3.1.4 Cell harvesting and cell disruption 73
3.1.5 Analysis of expression 73
3.1.6 Biotransformations 73
3.1.6.1 Hydroxylation of *n*-octane using whole cells 74
3.1.6.2 Hydroxylation of *n*-octane using cell free extracts (CFE) 74
3.1.7 Sample extraction and product analysis 74
3.2 Results 75
3.1.1 3DM Bio-prodict results and Yasara 75
3.1.2 Site directed mutagenesis to construct CYP153A6 mutants 75
3.1.3.1 Quickchange PCR amplification 75
3.2.2 Analysis of expression of CYP153A6 mutants 78
3.2.2.1 Spectroscopic quantification of P450 content 79
3.2.2.2 SDS-PAGE analysis of CYP153A6 mutants 81
3.2.3 Bitransformations 82
3.2.3.1 Hydroxylation of *n*-octane by CYP153A6 mutants using whole cells 82
3.2.3.2 Hydroxylation of *n*-octane by CYP153A6 mutants using cell free extracts 85
3.2.4 Evaluation of the CYP153A6 mutants based on protein expression and enzyme activity .. 87
3.2.4.1 CYP153A6 mutant groups 89
3.2.4.1.1 Mutants that displayed poor protein expression and no/poor activity 92
3.2.4.1.2 Mutants that displayed poor protein expression and unaffected or improved activity 95
3.2.4.1.3 Mutants that displayed acceptable protein expression and poor activity 96
3.2.4.1.4 Mutants that displayed acceptable protein expression and unaffected enzyme activity 97

3.2.4.1.5 Mutants that displayed acceptable protein expression and improved enzyme activity 98

3.3 Conclusion 99

Chapter 4: Conclusions and future outlook 101

Summary 103

Reference 105

Appendix 118

List of Abbreviations

°	Degrees
°C	Degrees Celsius
°C/min	Degrees Celsius per minute
x *g*	Times acceleration due to gravity
[2Fe-2S]	Iron-sulphur cluster
3'	Three-prime
5'	Five-prime
δ-ALA	5-Aminolevulinic acid hydrochloride
ε	Extinction coefficient
µg.mL^{-1}	Microgram(s) per millilitre
µL	Microlitre(s)
µm	Micrometre(s)
µM	Micromolar
A420	Absorbance at 420 nanometres
A450	Absorbance at 450 nanometres
A490	Absorbance at 490 nanometres
AciA/	AciA heme domain – P450RhF PFOR domain fusion
BLAST	Basic Local Alignment Search Tool
bp	Basepair(s)
BRM	Biotransformation reaction mixture
CO-difference	Carbon monoxide-difference
CPR	NADPH-cytochrome P450 reductase
CYP	Cytochrome P450
Cys	Cysteine
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleoside triphosphates
EDTA	Ethylenediaminetetraacetic acid

FAD	Flavin adenine dinucleotide
FdR	Ferredoxin reductase
Fdx	Ferredoxin
FdR/Fdx	Ferredoxin reductase and Ferredoxin partial operon
$FeCl_3$	Ferric chloride
FID	Flame ionisation detector
FMN	Flavin mononucleotide
$g_{WCW}\ L_{BRM}^{-1}$	grams wet cell weight per litre of biotransformation reaction mixture
$g_{DCW}\ L_{BRM}^{-1}$	grams dry cell weight per litre of biotransformation reaction mixture
GC	Gas chromatography
GDH	Glucose dehydrogenase
GOx	Glucose oxidase
h	hours
H_2	Hydrogen
HCl	Hydrochloric acid
kb	Kilobasepair(s)
kDa	Kilodalton(s)
kpsi	Kilopound(s) per square inch
l	Path length
L	Litre
LB	Luria-Bertani
m	Metre(s)
M	Molar
Met	Methionine
mg	Milligram(s)
Mg^{2+}	Magnesium ions
$MgCl_2$	Magnesium chloride
$mg.L^{-1}$	Milligram(s) per litre

min	Minute(s)
min^{-1}	Per minute
mL	Millilitre(s)
mm	Millimetre(s)
mM	Millimolar
$mmol.L_{BRM}^{-1}$	Millimole per litre of biotransformation reaction mixture
NaCl	Sodium chloride
NADH	β-Nicotinamide adenine dinucleotide-reduced
NADPH	β-Nicotinamide adenine dinucleotide phosphate-reduced
NaOH	Sodium hydroxide
NCBI	National Centre for Bioinformatics
ng	Nanogram(s)
nm	Nanometre(s)
No.	Number
ORF	Open-reading frame
P420	Pigment 420
P450	Pigment 450
P450balk/PFOR(P450RhF)	P450balk heme domain – P450RhF PFOR domain fusion
P450cam/PFOR(P450RhF)	P450cam heme domain – P450RhF PFOR domain fusion
PCR	Polymerase chain reaction
Pd	Putidaredoxin
PdR	Putidaredoxin reductase
PFOR	Phthalate family oxygenase reductase
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
sec	Seconds
sec/kb	Second(s) per kilobasepair
sp.	Species

SRSs	Substrate recognition sites
TAE	Tris-Acetate-EDTA
Tm	Melting temperature
Tris	2-Amino-2-(hydroxymethyl)-1,3-propandiol
Tris-HCl	2-Amino-2-(hydroxymethyl)-1,3-propandiol, hydrochloric acid
U	Units
v/v	Volume per volume
w/v	Weight per volume
w/w	Weight per weight
YASARA	Yet Another Scientific Artificial Reality Application

Chapter 1

Bacterial alkane hydroxylases – the CYP153 family

1.1 Introduction to Cytochrome P450s

Cytochrome P450 monooxygenases (CYPs or P450s) were first discovered by Martin Klingenberg in 1958 (Klingenberg, 1958). He observed that the microsomal membrane-bound heme-proteins of rat liver gave a unique 450 nm absorption peak when reduced and bound to carbon monoxide (CO) and it was originally reported as "microsomal CO-binding pigment". In 1962, Omura and Sato proposed the name "cytochrome P450" after characterization of this protein in microsomes from rabbit liver, but it was only in 1964 that the name was accepted (T. Omura & Sato, 1964). Today, P450s are one of the largest and oldest superfamily of enzymes and are distributed throughout all the biological kingdoms of life. In addition, there are currently over 20 000 cytochrome P450 sequences that have been identified and named (http://drnelson.uthsc.edu/cytochromeP450.html).

All P450s contain a heme-prosthetic group in the active site which is bound to the protein through the anionic, thiolate sulphur of an absolutely conserved cysteine residue (Lamb & Waterman, 2013) which is responsible for the characteristic 450 nm absorption peak (Omura, 1999; Stern & Peisach, 1974). The coordination of the thiolate group was confirmed when the first P450 structure (P450cam/CYP101) was solved in 1985 using X-ray crystallography (Poulos *et al*., 1987; Poulos *et al*., 1985). From the structure, Poulos and co-workers observed that the overall shape of P450cam was an asymmetrical triangular prism consisting of twelve α-helices and five anti-parallel beta sheets with the heme located between two α-helices at the bottom of a large binding pocket (Figure 1.2). Despite sharing less than 20% sequence identity, all P450 enzymes share this common overall fold and topology (Denisov *et al*., 2005).

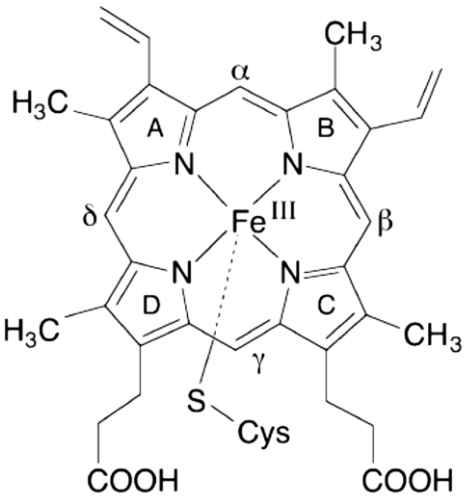

Figure 1.1 The central haem iron (Fe) atom binds to the protein via the thiolate sulphur (S) group of a conserved cysteine (Cys) residue. (Meunier *et al*., 2004)

The principle function of P450s is the oxygenation of various substrates. This reaction requires molecular oxygen and the supply of reducing equivalents nicotinamide adenine dinucleotide

phosphate (NADPH) or nicotinamide adenine dinucleotide (NADH) (Omura, 1999). The mechanism which these enzymes use to insert molecular oxygen into the substrate, termed the "cytochrome P450 catalytic cycle", was first proposed in 1971. However, updated descriptions of the cycle have been published throughout the years (Denisov *et al*., 2005; Estabrook *et al*., 1971; Gunsalus *et al*., 1975; Meunier *et al*., 2004; Sono *et al*., 1996). A generalised catalytic cycle for P450s is illustrated in Figure 1.3; upon the substrate (RH) binding to the active site ①, an electron is transferred to the heme from a cofactor NADPH/NADH ② which reduces the iron allowing an oxygen molecule (O_2) to bind ③. A second electron is transferred to the heme ④ followed by a protonation (H^+) step ⑤ and hydrolysis of the O - O bond allowing one oxygen atom to leave the reaction in the form of water (H_2O) ⑥ and the second oxygen atom to be inserted into the substrate forming the hydroxylated product (-ROH) ⑨ which is then released from the binding site (Danielson, 2002; Denisov *et al*., 2005; Isin & Guengerich, 2007).

Figure 1.2 Crystal structure of P450cam (2CPP) imported from the Protein Data Bank. (http://www.rcsb.org/pdb/explore/images.do?structureId=2CPP)

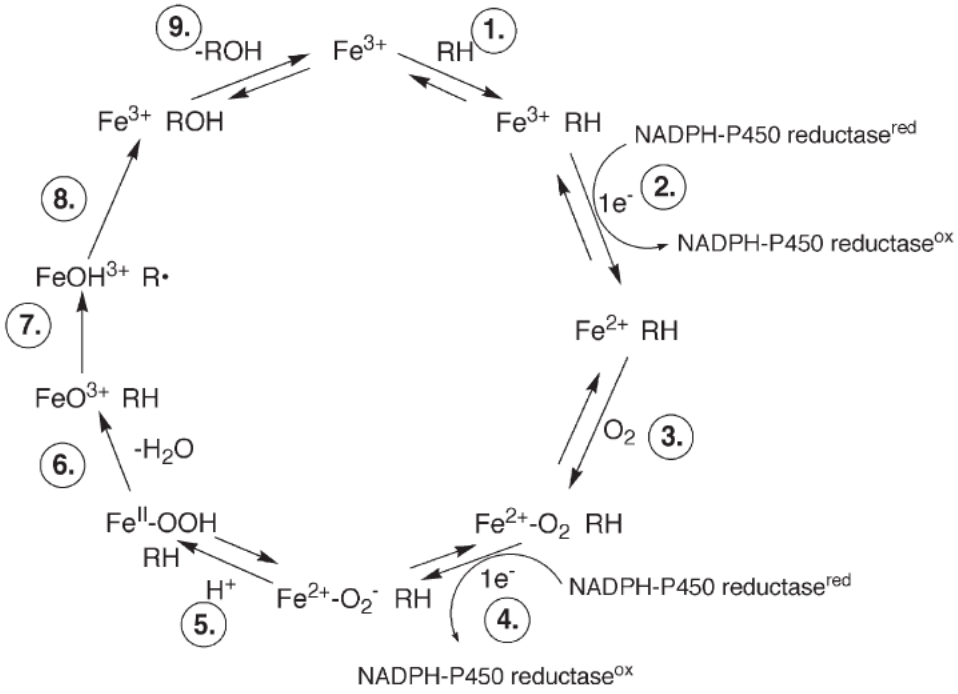

Figure 1.3 A generalised schematic representation of the P450 catalytic cycle (Isin & Guengerich, 2007).

Cytochrome P450 enzymes make use of a variety of redox systems to supply the electrons for catalysis (② and ④). Initially P450s were divided into two classes based on the P450-redox systems thought to exist in nature; P450cam (the first bacterial P450 to be enzymatically and structurally characterised) was found to be soluble and its redox partners (flavin adenine dinucleotide (FAD)-containing putidaredoxin reductase and 2Fe-2S cluster-containing putidaredoxin) were also found to be soluble proteins. In contrast, mammalian P450 enzymes from the liver were found to be integral membrane proteins in the endoplasmic reticulum (ER) which were receiving electrons from the membrane anchored NADPH-cytochrome P450 reductase (CPR). Thus, it was assumed that these two systems were the only two distinct classes: the bacterial system (Class I) composed of three soluble proteins and the eukaryotic system (Class II) composed of membranous P450 and CPR enzymes (McLean *et al*., 2005). However, as more P450s were discovered so did our knowledge of redox partner proteins and the various P450-redox systems grow. Hanneman and co-workers (2007) eventually arranged P450s into ten classes according to their redox partner proteins. Most of the bacterial P450 systems (Figure 1.4 (a)) as well as the eukaryotic, mitochondrial P450 systems (Figure 1.4 (b)) belong to Class I which include three-component systems with FAD-containing reductases (FdR) transferring electrons from NADPH/NADH to ferredoxins (Fdx), the second components of these systems which in turn reduce the P450s. In bacteria, all three proteins are soluble but in eukaryotes, only Fdx is soluble, whereas FdR and the P450 are membrane-associated and membrane-bound, respectively (Hannemann *et al*., 2007).

The P450-redox system of Class II P450s is found in the ER and is the system most commonly used by eukaryotes. NADPH-cytochrome P450 reductase (CPR) contains both FAD and flavin adenine mononucleotide (FMN) prosthetic groups which transfer electrons from NADPH to the heme domain of P450s (Hannemann *et al*., 2007). In this system both P450 and CPR are integral membrane proteins (Figure 1.4 (c)).

Bacterial P450s belonging to Class VII and Class VIII are particularly interesting enzymes because they are naturally fused to their redox partner proteins. This means that the P450 enzyme and its redox partner occur in a single polypeptide chain expressed as a single protein. These P450s are considered to be catalytically self-sufficient because they do not require additional electron transfer proteins. The first P450 fusion discovered was P450BM3 (CPY102A1) which was discovered naturally fused to a cytochrome P450 reductase (Narhi & Fulco, 1986) as illustrated in Figure 1.4 (d). It was later classified as a Class VIII P450 enzyme. The first Class VII P450 reported was CYP116B2 (P450RhF) from *Rhodococcus sp*. (Roberts *et al*., 2002). The phthalate dioxygenase reductase domain (PFOR) of the fusions is not usually associated with P450 enzymes making this a novel class of P450-redox systems (Hannemann *et al*., 2007). The PFOR reductase consists of three parts (Figure 1.4 (e)): a NADH-binding

domain which accepts electrons from NADH and transfers them to the FMN-binding domain which lastly shuttles the electrons to a [2Fe-2S] ferredoxin domain which in turn reduces the P450 (Hannemann *et al*., 2007). Perhaps the most interesting class of the P450-redox system is Class IX (Figure 1.4 (f)), characterized as a nitric oxide reductase (P450nor/CYP55) which is the first and thus far the only known soluble eukaryotic P450 enzyme (Kizawas *et al*., 1991; Takayaa *et al*., 1999). This remarkable enzyme uses NADH as electron donor and does not require other electron transfer proteins (Hannemann *et al*., 2007).

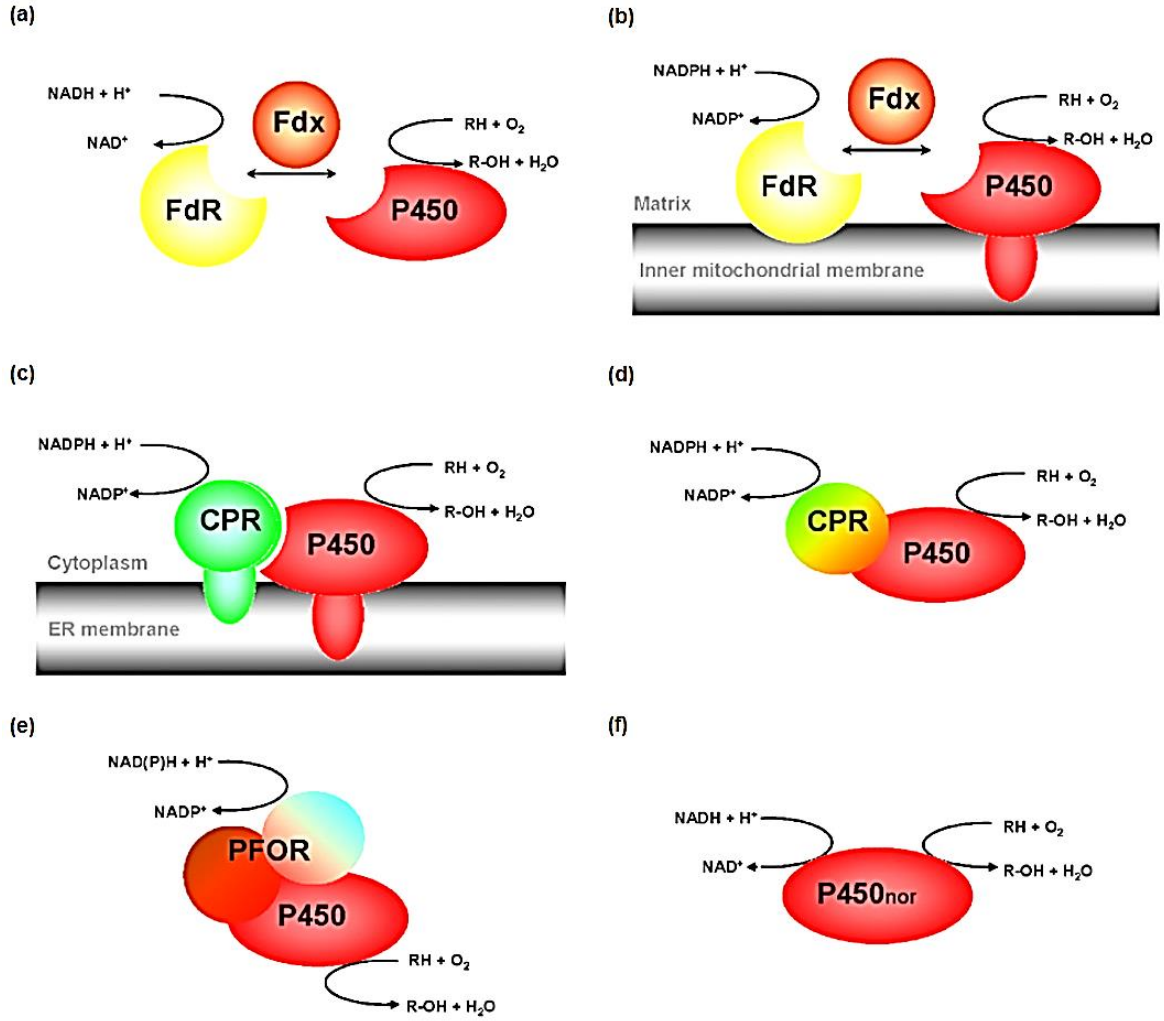

Figure 1.4 Schematic representation of various electron transport systems used by P450 enzymes. (a) Class I, soluble bacterial P450-redox system; (b) eukaryotic mitochondrial P450-redox system; (c) Class II, eukaryotic microsomal P450-redox system; (d) Class VIII, soluble bacterial [CPR]-[P450] fusion system; (e) Class VII, soluble bacterial [PFOR]-[P450] fusion system; (f) Class IX, only soluble independent eukaryotic system (Hannemann *et al*., 2007).

Although P450s share a common catalytic cycle and the P450 fold is highly conserved, there is enough structural diversity which allows binding of different size substrates to different P450s with varying degrees of specificity (Denisov *et al*., 2005). Thus, P450s are very versatile enzymes which catalyse a variety of reactions such as hydroxylations, epoxidations, Baeyer-Villiger oxidations, isomerisations,

dehydrogenations and N-, S-, and O-dealkylations to name a few (Chefson & Auclair, 2006; Isin & Guengerich, 2007; Sono *et al*., 1996). In addition, their versatility enables them to have a wide range of catalytic functions (Figure 1.5) including the metabolism and synthesis of endogenous compounds (e.g. hormones, fatty acids, antibiotics, mycotoxins, steroids etc.) and the conversion of other chemicals foreign to organisms, so called "xenobiotics" (pharmaceutical drugs, pesticides, herbicides, etc.) (Denisov *et al*., 2005; Mot & Parret, 2002; Sono *et al*., 1996) During the metabolism of xenobiotics, in mammals, reactive intermediates are produced during phase I of a three-phase detoxification process and if these reactive intermediates are not deactivated in phase II they may cause damage to other proteins, RNA and DNA in the cell (Danielson, 2002; Liska, 1998).

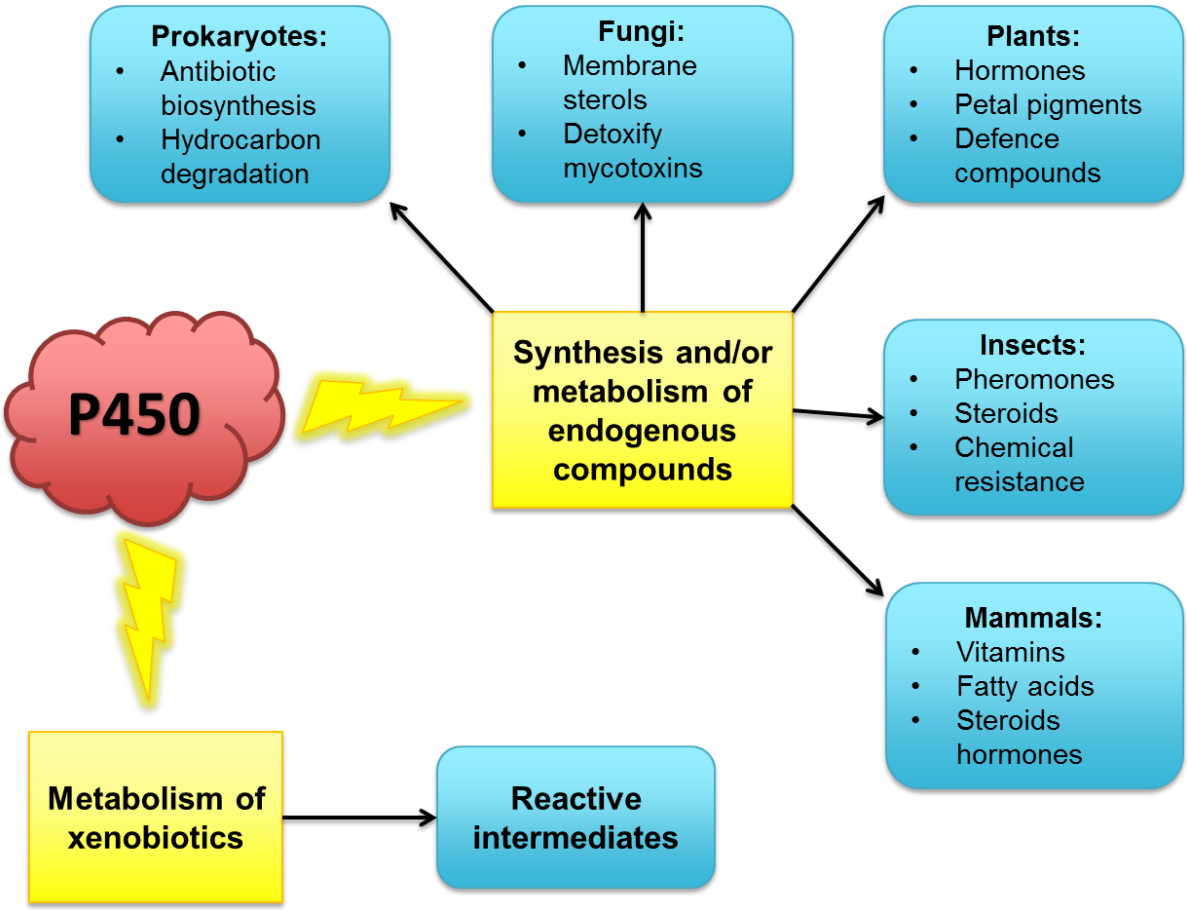

Figure 1.5 Schematic representation of the functions of cytochrome P450 enzymes in various organisms.(Denisov *et al*., 2005; Mot & Parret, 2002; Sono *et al*., 1996)

Many reactions catalysed by P450s are often too difficult to accomplish by means of organic chemistry therefore the development of industrial processes using P450s has become of great interest (Funhoff *et al.*, 2006; Van Beilen & Funhoff, 2005). Of particular interest are enzymes that predominantly perform terminal hydroxylation of hydrocarbons. The bacterial CYP153 family catalyse the terminal hydroxylation of aliphatic, alicyclic and alkyl-substituted compounds with high region- and stereo-

selectivity as well as the epoxidation of aliphatic and alicyclic alkenes (Funhoff *et al.*, 2007; Li & Chang, 2004; Van Beilen *et al*., 2006). The enzymatic and functional characteristics of the CYP153 enzymes as well as their applications in industry will be the focus of this review.

1.2 Microbial alkane hydroxylases involved in *n*-alkane degradation

Although hydrocarbons are known to be the "simplest" organic compounds - because they consist of only carbon and hydrogen – they are also considered as energy-rich organic compounds. Hydrocarbons can be saturated meaning that they contain only single C-H (carbon-hydrogen) bonds or they can be unsaturated meaning that they contain either one or more double or triple C-H bonds. Alkanes are saturated hydrocarbons that occur in nature in various forms such as linear (*n*-alkanes), cyclic (cyclic-alkanes) or branched (iso-alkanes). Alkanes are naturally produced in the environment due to geochemical processes such as decaying plants and algal material and biological processes whereby living organisms release alkanes into the environment via waste products, defence compounds or pheromones (van Beilen *et al.*, 2003). Unfortunately, the largest contributor of alkanes into the environment is anthropogenic, for instance marine oil spills, municipal and industrial waste and runoff and leaks in industrial pipelines and storage tanks (Cappelletti, 2009).

Nearly a century ago, Söhngen published an article on microbes that were responsible for the disappearance of oil slicks on water surfaces; he was able to isolate bacteria that could degrade methane and longer alkanes (Söhngen, 1913). Since then research on microbial alkane degradation has flourished and many microorganisms from different genera have been isolated and identified as alkane degraders because they contain multiple alkane hydroxylases.

Hydrocarbon activation by alkane hydroxylases is the first step in biodegradation: alkanes are usually activated by terminal hydroxylation to produce a primary alcohol which is then further oxidised by alcohol and aldehyde dehydrogenases to produce fatty acids which are then metabolised in the β-oxidation cycle (Van Beilen *et al.*, 2003). Alkane hydroxylases can be divided into three groups based on the chain-length of the alkane substrate they oxidise (Rojo, 2009; Van Beilen & Funhoff, 2007).

1.2.1) C_1-C_4 (Methane monooxygenases)

All known methanotrophs produce membrane-bound particulate methane monooxygenase (pMMO) while only some produce soluble methane monooxygenase (sMMO) which catalyse the hydroxylation of C_1-C_5 and C_1-C_8 alkanes and alkenes, respectively. Most sMMOs also catalyse the

hydroxylation of C_1-C_8 cycloalkanes (Lieberman & Rosenzweig, 2005; Merkx *et al*., 2001). These metalloproteins mediate hydroxylation via a dicopper or diiron catalytic centre.

1.2.2) C_5-C_{16} (Integral membrane non-heme iron or cytochrome P450 monooxygenases)

Several particulate alkane hydroxylases (pAH) that were isolated from various *Pseudomonas sp*. belong to this group including the well-characterised *P. putida* GPo1 an integral-membrane non-heme diiron monooxgenase (AlkB) which oxidises C_5-C_{12} alkanes (Van Beilen *et al.*, 1994). The microsomal P450s of the CYP52 family (yeast) are also well known for their ability to grow on C_{10}-C_{16} alkanes (Van Beilen *et al*., 2003). Another group of P450 enzymes which also belong to this group is the CYP153 family which catalyse the hydroxylation of C_5-C_{12} *n*-alkanes, cyclic-alkanes, alkylbenzenes (Maier *et al.*, 2001; Van Beilen *et al*., 2006).

1.2.3) C_{17} < (Flavin-containing oxygenases)

A Flavin-containing dioxygenase isolated from *Acinetobacter sp.* M-1 was found to hydroxylate alkanes ranging from C_{10}-C_{30} (Maeng *et al.*, 1996). In addition, a soluble alkane hydroxylase also belonging to this species was found to assist growth on C_{13}-C_{44} alkanes (Tani *et al.,* 2001). A flavin-containing alkane monooxygenase called LadA which belongs to the bacterial luciferase family, catalyses the terminal oxidation of linear alkanes up to at least C_{36} (Li et al., 2008).

Monooxygenases catalyse the hydroxylation of alkanes by inserting a single oxygen atom into a C-H bond. Only three types of monooxygenase enzymes can hydroxylate *n*-alkanes at the terminal position namely cytochrome P450 monooxygenases (CYP450s), diiron monooxygenases and the flavin containing very long chain alkane hydroxylase called LadA. Of these three classes of enzymes cytochrome P450 monooxygenases have received the most attention because they are currently giving the highest space-time yields of the alcohols (Bordeaux *et al.*, 2012).

1.3 CYP153 hydroxylases

CYP153s are very versatile cytochrome P450 enzymes because they catalyse hydroxylation and epoxidation reactions over a wide range of substrates. These soluble monooxygenases have the ability to hydroxylate the terminal carbon of aliphatic alkanes, the most unreactive carbon (Funhoff *et al*., 2006). In addition to aliphatic alkanes, CYP153s also catalyse the terminal hydroxylation of other alicyclic and alkyl-substituted substrates as well as the epoxidation of linear and cyclic double bond containing substrates (Funhoff *et al*., 2007). Such reactions are often difficult to accomplish by means

of organic chemistry which is one of the main reasons why CYPs make such attractive tools for the biocatalytic synthesis of useful chemicals on an industrial level (Olaofe *et al.*, 2013).

1.3.1 Identification and expression of CYP153

In the early 1980's Asperger and co-workers identified an *Acinetobacter* stain (EB104) that could grow on and metabolise alkanes through terminal hydroxylation (Asperger, 1981). Müller and co-workers then proceeded to purify the enzyme which they found to have a relative molecular mass (M_r) of 52 000 Da. CO-difference spectra confirmed that the enzyme belonged to the cytochrome P450 superfamily because the enzyme formed a peak at 448 nm when reduced and bound to carbon monoxide (Müller *et al.*, 1989).

1.3.1.1 Discovery of the first CYP153 hydroxylase

The *n*-alkane hydroxylating cytochrome P450 from the *Acinetobacter sp*. (EB104) sparked the interest of Maier and co-workers. They managed to determine the complete nucleotide sequence of the P450 enzyme by performing short walks on genomic DNA via single specific primer –polymerase chain reaction (SSP-PCR). Sequence analysis confirmed that the amino acid (AA) sequence of the P450 corresponded to the typical features exhibited by other P450s, particularly all the amino acid residues and motives that were conserved among P450 enzymes. However, its length of 497 AA (M_r = 56 kDa) was longer than other bacterial P450s. Furthermore, sequence analysis determined that the highest sequence similarity was a mere 33% with CYP111. The guidelines for P450 nomenclature clearly state that members of the same family must have more than 40% sequence identity (Nelson *et al*., 1996) confirming that the enzyme was in fact the first member of a new P450 family, the CYP153 family and was designated CYP153A1.

1.3.1.2 Identification, cloning and expression of CYP153 genes

Jan van Beilen and co-workers had been conducting research on alkane hydroxylases in Gram-positive and Gram-negative bacteria for many years (Smits *et al.*, 2002; Smits *et al.*, 1999; Van Beilen *et al*., 2003; Van Beilen & Panke, 2001; Van Beilen & Smits, 2002; Van Beilen *et al*., 1994). Initially, their research focused on the integral-membrane, non-heme, diiron monooxygenase (AlkB) isolated from *Pseudomonas putida* GPo1 where they have cloned several closely related alkB genes from *P. aeruginosa* and *P. putida* (Van Beilen & Smits, 2002). In addition they observed that other alkane hydroxlase systems were isolated from strains belonging to *Acinetobacter* (Maier *et al*., 2001; Tani *et al*., 2001), *Sphingomonas* (Chang *et al.*, 2002), *Mycobacterium* (Van Beilen & Smits, 2002) and *Rhodococcus* (Van Beilen & Smits, 2002) amongst others. Smits and van Beilen also went on to construct what they termed "alkane-responsive expression vectors", pKKPalk and pCom.

Table 1.1 Enzymes involved in the oxidation of alkanes

Enzyme class	Enzyme Composition and cofactors	Substrate range	Microorganism	Reference
Particulate methane monooxygenase (pMMO)	α3β3γ3 hydroxylase trimer composed of PmoA, PmoB, PmoC; mononuclear copper and dinuclear copper in PmoB	C_1-C_8 (halogenated) alkanes, alkenes	*Methylococcus, Methylosinus, Methylocystis, Methylomicrobium, Methylomonas*, etc.	(Lieberman & Rosenzweig, 2005; McDonald, 2006)
Soluble methane monooxygenase (sMMO)	α2β2γ2 hydroxylase; dinuclear iron reductase, [2Fe–2S], FAD, NADH regulatory subunit	C_1-C_8 (halogenated) alkanes, alkenes, cycloalkanes	*Methylococcus, Methylosinus, Methylocystis, Methylomonas, Methylocella*, etc.	(McDonald, 2006; Merkx *et al*., 2001)
AlkB-related alkane hydroxylases	Membrane hydroxylase; dinuclear iron rubredoxin; mononuclear iron rubredoxin reductase, FAD, NADH	C_5–C_{16} alkanes, fatty acids, alkylbenzenes, cycloalkanes, etc.	*Acinetobacter, Alcanivorax, Burkholderia, Mycobacterium, Pseudomonas, Rhodococcus*, etc.	(Van Beilen *et al*., 2003)
Eukaryotic P450 (CYP52)	Microsomal oxygenase; P450 heme reductase; FAD, FMN, NADPH	C_{10}–C_{16} alkanes, fatty acids	*Candida maltosa, Candida tropicalis, Yarrowia lipolytica*	(Iida *et al*. , 2000; Van Beilen *et al.*, 2003)
Bacterial P450 (CYP153)	P450 oxygenase; P450 heme ferredoxin; iron–sulfur ferredoxin reductase, FAD, NADH	C_5–C_{16} alkanes, (cyclo)-alkanes, alkylbenzenes, etc	*Acinetobacter, Alcanivorax, Caulobacter, Mycobacterium, Rhodococcus, Sphingomonas*, etc.	(Van Beilen *et al.*, 2006)
Dioxygenase	Homodimer , non-heme iron/copper; copper, FAD/FMN, NADPH	C_{10}-C_{44} alkanes	*Geobacillus thermodenitrificans, Acetobacter pasteurianus 386B, Acinetobacter sp*. M-1	(Li *et al.*, 2008; Maeng *et al*., 1996)

[Information from (an Beilen *et al*., 2003)]

Both these vectors were constructed with the *alkB* promoter (*PalkB*) from *Pseudomonas putida* GPo1 which controls the expression of the alkB operon, and allowed the successfully expression of target proteins in both *E. coli* and *Pseudomonas*, (Smits *et al.*, 2001). Modified versions of the pCom vector would later be used to clone many CYP153 genes.

The production of perillyl alcohol was of interest because of its anti-carcinogenic properties and the only known enzyme known to produce perillyl alcohol was found in *Bacillus stearothermophilus* BR388. This enzyme however was not sufficiently regio-selective, producing a mixture of products that were laborious and expensive to separate (Van Beilen *et al*., 2005). In his search to find a limonene hydroxylase with high regio-selectivity to produce perillyl alcohol, van Beilen and co-workers screened 1 800 bacterial strains grown on a range of substrates such as toluene, naphthalene and various alkanes. The screening analysis determined that *Mycobacterium sp.* stain HXN-1500, originally isolated from a trickling-bed reactor (Smits *et al*., 2002), gave the best specific activity for the selective production of perillyl alcohol. The bacteria also grew well on linear alkanes ranging from *n*-hexane to *n*-dodecane and the best result with regards to limonene hydroxylation were obtained when cells were grown on *n*-octane (Van Beilen *et al*., 2005). These results indicated that enzyme responsible was likely to be an alkane hydroxylase

The protein was purified and N-terminal sequencing revealed 52% sequence identity with the N-terminal sequence of a putative P450 gene from the *Caulobacter crescentus* genome sequence. In addition, the *Caulobacter* P450 sequence had a sequence identity of greater than 50% with CYP153A1 from *Acinetobacter calcoaceticus* EB104 (Maier *et al*., 2001). Multiple sequence alignments of the *A. calcoaceticus* EB104 with its homologs from *C. crescentus* and *Bradyrhizobium japonicum,* revealed two regions where the sequences were perfectly conserved: the oxygen-binding stretch in the I helix (GGNDTTRN) and the sequence ending with the heme-binding site (HLSFGFGIHRC). These two sequences were used to design primers to amplify the partial target gene encoding the alkane hydroxylase .The partial gene encoded a peptide that had 70% sequence identity to CYP153A1 and 74% sequence identity to the *C. crescentus* P450. The gene fragment was subsequently used to probe restriction fragments from chromosomal DNA through Southern blotting.

Sequencing of a 6.2 kb DNA fragment revealed a P450 gene with high full-length homology to CY153A1. It was classified as a new member of the CYP153 family and designated CYP153A6 (Nelson *et al*., 1996). Further sequence analysis of the DNA fragment indicated that van Beilen and co-workers were able to clone the complete operon (Figure 1.6) which not only included the CYP153A6 P450 gene (1 260 bp ahpG), but also the genes encoding the redox partner proteins ferredoxin reductase (ahpH)and ferredoxin (*ahpI).*

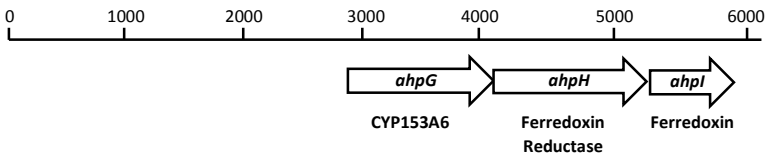

Figure 1.6 Organisation of the cytochrome P450 operon and flanking genes. The open reading frames (ORFs) are indicated by arrows [Modified from (Van Beilen *et al*., 2005)]

Unfortunately, the HXN-1500 stain was not suitable for large-scale application because when cultivated on a 2 L scale with *n*-octane the cells became hydrophobic, forming clumps in the octane layer. Therefore, van Beilen and co-workers cloned the complete operon into pCom8-PFR1500 (Smits *et al*., 2001) and expressed the constructs in *E. coli* GEc137 and *P. putida* GPo12. Although both hosts contain all the necessary genes for growth on *n*-octane, the recombinant *E. coli* did not grow whereas the recombinant *P. putida* grew slowly. Subsequently, growth of the recombinant *P. putida* was tested on other *n*-alkane substrates. Although *n*-heptane supported faster growth (Figure 1.7a), growth on *n*-octane supported higher specific activity for perillyl alcohol production.

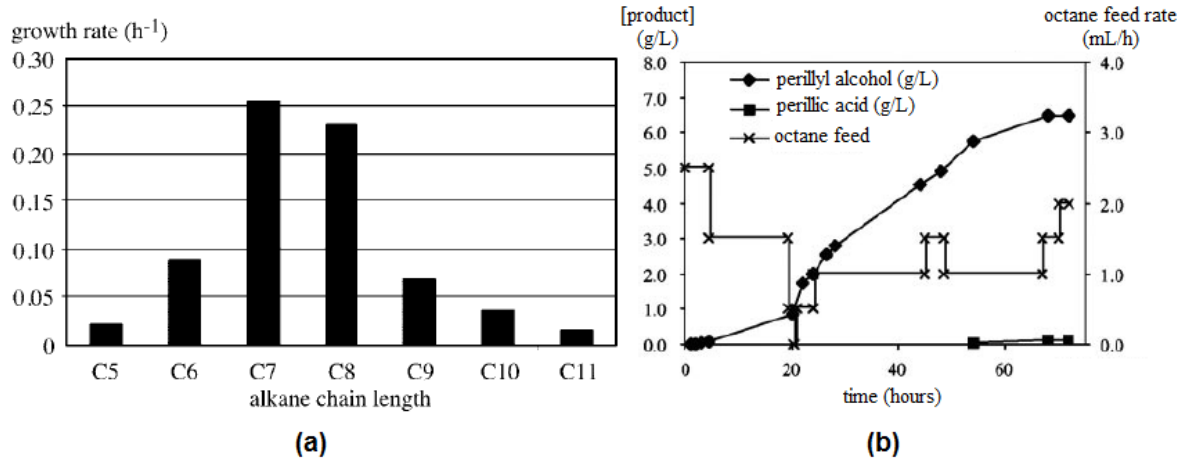

Figure 1.7 (a) Growth rates of *P. putida* GPo12 (pGEc47_B)(pCom8-PFR1500) on *n*-alkanes. C_5, n-pentane; C_6, *n*-hexane; C_7, *n*-heptane;C_8, *n*-octane; C_9, *n*-nonane; C_{10}, *n*-decane; C_{11}, *n*-undecane. (b) Biotransformation of limonene with *P. putida* GPo12 (pGEc47_B)(pCom8-PFR1500) at the 1.5-liter scale. Symbols:◆, perillyl alcohol concentration in the organic phase; ■, perillic acid concentration in the water phase; x, n-octane feed rate. (Van Beilen *et al*., 2005)

In a separate study, the same degenerate primers used to amplify the CYP153A6 gene from *Mycobacterium sp*. HXN-1500, were used to amplify CYP153 gene fragments from several other strains of bacteria (Van Beilen *et al*., 2006). Most stains from *Mycobacteria*, *Rhodococcus erythropolis* and several *Proteobacteria* yielded PCR products. Sequence analysis revealed that all PCR fragments encoded peptides sharing >60% sequence identity with CYP153A1 and CYP153A6. Following a similar experimental strategy as for the elucidation of CYP153A6, five CYP153 enzymes were cloned from

Sphingomonas sp. HXN-200 (Van Beilen *et al*., 2006): CYP153A7, CYP153D3, CYP153A8, CYP153A11 and CYP153D2 (Table 1.2).

Alcanivorax borkumensis SK2 was found to encode two closely related CYP153 genes (Hara *et al*., 2004) therefore in order to amplify the full-length CYP153 genes from the two closely related strains *A. borkumensis* AP1 and *A. borkumensis* SK2, primers were designed based on the sequence of *A. borkumensis* SK2. Two CYP153 genes were cloned from *A. borkumensis* AP1 namely CYP153A12 and CYP153A13 which shared 94% sequence identity; in addition, CYP153A13 was almost identical (99.6% sequence identity) to the two CYP153 genes encoded by *A. borkumensis* SK2 (Van Beilen *et al*., 2006).

Furthermore, the flanking regions of five CYP153 genes namely CYP153A1, CYP153A6, CYP153A10, CYP153A13, CYP153A14 encoded homologs of ferredoxin and ferredoxin reductase confirming that CYP153 enzymes use the P450-redox system of Class I cytochrome P450s (Van Beilen *et al*., 2006). In addition, ferredoxin genes were identified in the flanking regions of three CYP153 genes belonging to strain HXN-200 namely, CYP153A7, CYP153A8 and CYP153A11.

In total, van Beilen and co-workers were able to clone eight full-length CYP153 genes. They also managed to retrieve 11 full-length sequences from other finished/unfinished microbial genome projects (Table 1.2). Since the CYP153A6 operon (all three components) was successfully cloned into the pCom8 plasmid and expressed in *P. putida* GPo12, van Beilen and co-workers decided to use the same host for expressing of the newly cloned CYP153s. CYP153A7 was selected first for heterologous expression, the CYP153A7 gene was cloned into pCom together with the downstream ferredoxin resulting in a pCom8-$P_{A7}F_{200}$ plasmid which was expressed in *P. putida* GPo12. However, the recombinant strain showed poor grown on *n*-octane. The plasmid was modified to include the ferredoxin reductase gene from *Mycobacterium sp.* HXN-1500 (Van Beilen *et al*., 2005) resulting in a pCom8-$P_{A7}F_{200}R_{1500}$ plasmid which significantly increased the growth of recombinant *P. putida* on *n*-octane. However, this plasmid did not support the expression of the other CYP153 genes therefore a derivative of pCom-$P_{A7}F_{200}R_{1500}$ was constructed using the pCom12 plasmid (pCom12-

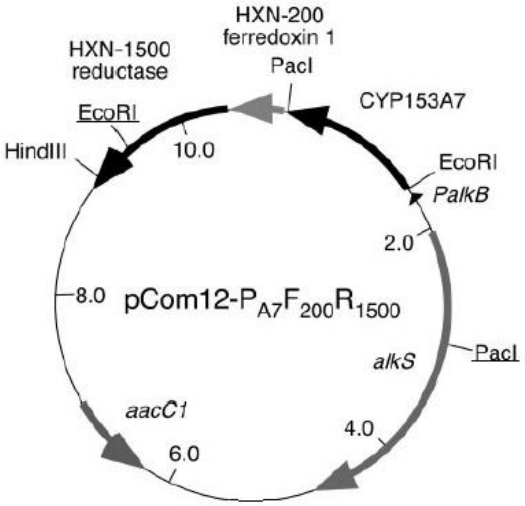

Figure 1.8 Map of pCom12-$P_{A7}F_{200}R_{1500}$. The plasmid is based on pCom8, which contains the PalkB promoter, a gene encoding the positive regulator alkS of PalkB, a gentamicin resistance gene (aacC1), origins of replication for E. coli and Pseudomonas sp. In pCom12 vectors, an EcoRI site is located between the ribosome-binding site and the ATG start codon of the CYP153A7 gene and the PacI site is located between the CYP153A7 gene and the downstream ferredoxin gene. (Van Beilen *et al.*, 2006)

$P_{A7}F_{200}R_{1500}$) which allowed easy insertion of P450 genes as EcoRI-PacI fragments (Figure 1.8). This plasmid was used to clone CYP153A1 (*Acinetobacter sp*. EB104); CYP153A5 (*R. palustris* CGA009); CYP153A8, -A11, -D2 and -D3 (*Sphingomonas sp*. HXN-200); CYP153A12 and -A13 (*A. borkumensis* AP1) and CYP153A14 (*M. marinum M*). Most of the CYP153 genes were successfully expressed in *P. putida* GPo12 with CYP153A5, -A8, -A12 and -D3 as the exceptions. A year later Funhoff and co-workers were able to successfully express CYP153D3 in *P. fluorescens* KOB2Δ1. This resulted in CYP153D3 having the highest expression levels (up to 11% total protein content) among the CYP153 proteins that were expressed in *P. fluorescens* (Funhoff *et al*., 2007).

During the same time, Kubota and co-workers constructed a vector that would enable various P450 genes from Class I to be functionally expressed in *E. coli* (Figure 1.9). The vector, designated pRED, contained genes encoding a linker sequence and the reductase domain of P450RhF (CYP116B2) isolated from *Rhodococcus sp.* (Narhi & Fulco, 1986) which was inserted into a pET21a plasmid (Kubota *et al*., 2005). The various P450 genes were then inserted into the pRED plasmid via specific restriction sites to produce artificial self-sufficient P450 enzymes (Nodate *et al*., 2006). Kubota and co-workers used this plasmid to fuse 16 chimeric P450 genes that were isolated from various environments as well as the CYP153A13a gene to the P450RhF reductase domain. The plasmids were then expressed in *E. coli* BL21-A1 at lower temperatures to favour folding of the fusion proteins, but only 8 of the 16 chimeric P450 fusions were functionally expressed. In addition, a self-sufficient CYP153A13a enzyme was produced which was also functionally expressed in *E. coli*. Fujita and co-workers later compared the functional expression and activity of CYP153A13a in *E. coli* BL21(DE3) as a fusion protein in pRED (Kubota *et al*., 2005) with a vector developed by the Mercian corporation (pT7NS-camAB (Fujita *et al*., 2009)) that contained the putidaredoxin reductase (camA) and putidaredoxin (camB) redox partner proteins of P450cam (CYP101) (Poulos *et al*., 1987). Better specific activity towards cyclohexane, *n*-butylbenzene and 2-*n*-butylbenzofuran was achieved when CYP153A13a were expressed as an artificial self-sufficient system using the pRED vector. A close homolog of CYP153A13a, aciA, however showed better results as a three-component system with the redox partner proteins from CYP101 with *n*-octane as substrate.

More recently, the complete open reading frame of the artificial CYP153A13a fusion (Nodate *et al*., 2006) was sub-cloned into pET-28b(+) (Novagen) and expressed in *E. coli* BL21(DE3) (Pennec *et al*., 2014). This artificial fusion was then directly compared to the three component system CYP153A6, ferredoxin reductase and ferredoxin from *Mycobacterium sp*. HXN-1500 also expressed from pET-28b(+) in *E. coli* BL21 (DE3) (Gudiminchi *et al*., 2012).

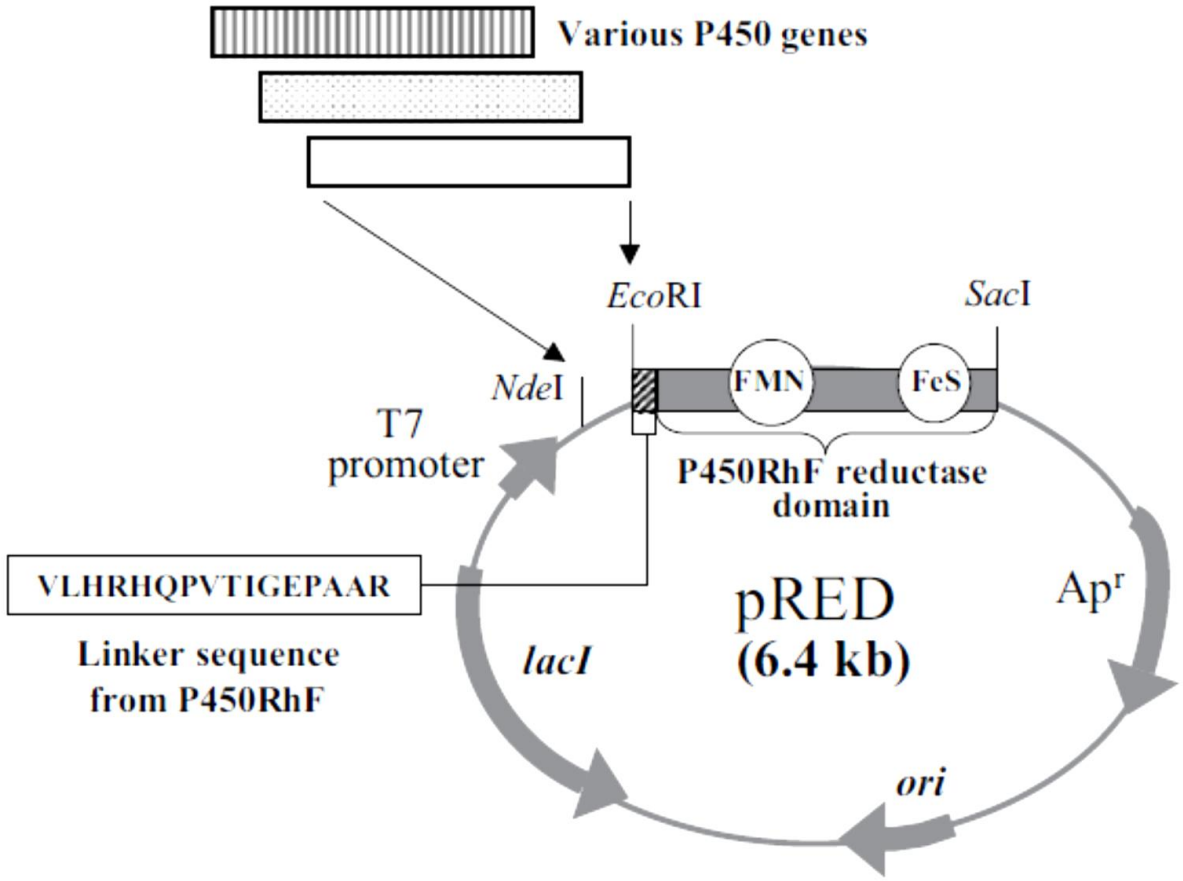

Figure 1.9 Structure of vector pRED for functionally expressing a variety of class I P450s. The DNA fragment encoding the linker sequences and the reductase domain of P450RhF amplified by PCR was inserted into pET21a to generate expression vector pRED. Various class I P450 genes without their stop codon could be inserted into the NdeI‑EcoRI site to produce hybrid P450s. (Nodate *et al.*, 2006)

Similary, other members of the CYP153A family, including CYP153A7 (Pham *et al.*, 2014), CYP153C1 and CYP153D1 (Bell & Wong, 2007), have been cloned into pET expression vectors and successfully expressed in *E. coli.* The various bacterial strains and plasmids used to clone and express CYP153 genes are summarised in Table 1.3.

Over the years the CYP153 family has expanded and now consists of five sub-families (A, B, C, D and E) however phylogenetic analysis indicates that most of the CYP153 genes belong to the CYP153A family (Alonso-Gutiérrez *et al*., 2011; Kubota *et al*., 2005; Nie *et al*., 2014; Van Beilen *et al*., 2006). The number of CYP153 sequences is estimated to be 164 but only 155 encode proteins (http://www.cyped.uni-stuttgart.de/cgi-bin/CYPED5/index.pl?page=fam) which have not all been named (Nelson *et al*., 1996) and only a few have been enzymatically and functionally characterised thus far.

1.3.2 Characterisation and reactions catalysed

CYP153s are soluble proteins therefore, they are characterised as Class I cytochrome P450 enzymes. This means they require redox partner proteins ferredoxin reductase and ferredoxin to transport electrons from NADH/NADPH to the heme domain (Figure 1.4 (a)) in order for catalysis to occur. CYP153s were the first described soluble P450 enzymes that specifically displayed hydroxylating activity towards the terminal carbon of alkanes (Funhoff *et al*., 2006). Based on their bond strengths, terminal methyl C-H bonds are inherently more difficult to hydroxylate than the secondary or tertiary C-H bonds in the hydrocarbon chain (Bordeaux *et al*., 2012; Johnston *et al.*, 2011).

1.3.2.1 Reactions catalysed by CYP153 hydroxylases

Terminal hydroxylation of linear alkanes to produce 1-alkanols is the first step in alkane degradation therefore microorganisms are able use these substrates as their sole carbon source because the alcohol is further metabolised by alcohol and aldehyde dehydrogenases before entering the β-oxidation pathway. Initially growth studies were used to characterise the substrate range of different CYP153 enzymes (Figure 1.10). By growing *P. putida* recombinants expressing the different CYP153 genes on various substrates van Beilen and co-workers were able to determine that CYP153A1, -A6, -A7, -A11, -A13, -A14 and -D2 were able to hydroxylate C_5-C_{10} linear alkanes (Van Beilen *et al*., 2006). Today we know that CYP153s catalyse the hydroxylation and epoxidation of various substrates with high regio-selectivity for the terminal carbon atom (Figure 1.12).

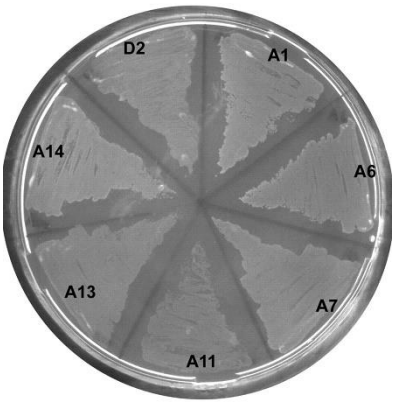

Figure 1.10 Growth on *n*-octane vapour, of *P. putida* GPo12 recombinants (pCom12-$P_xF_{200}R_{1500}$) expressing seven different CYP153 genes. (Van Beilen *et al*., 2006)

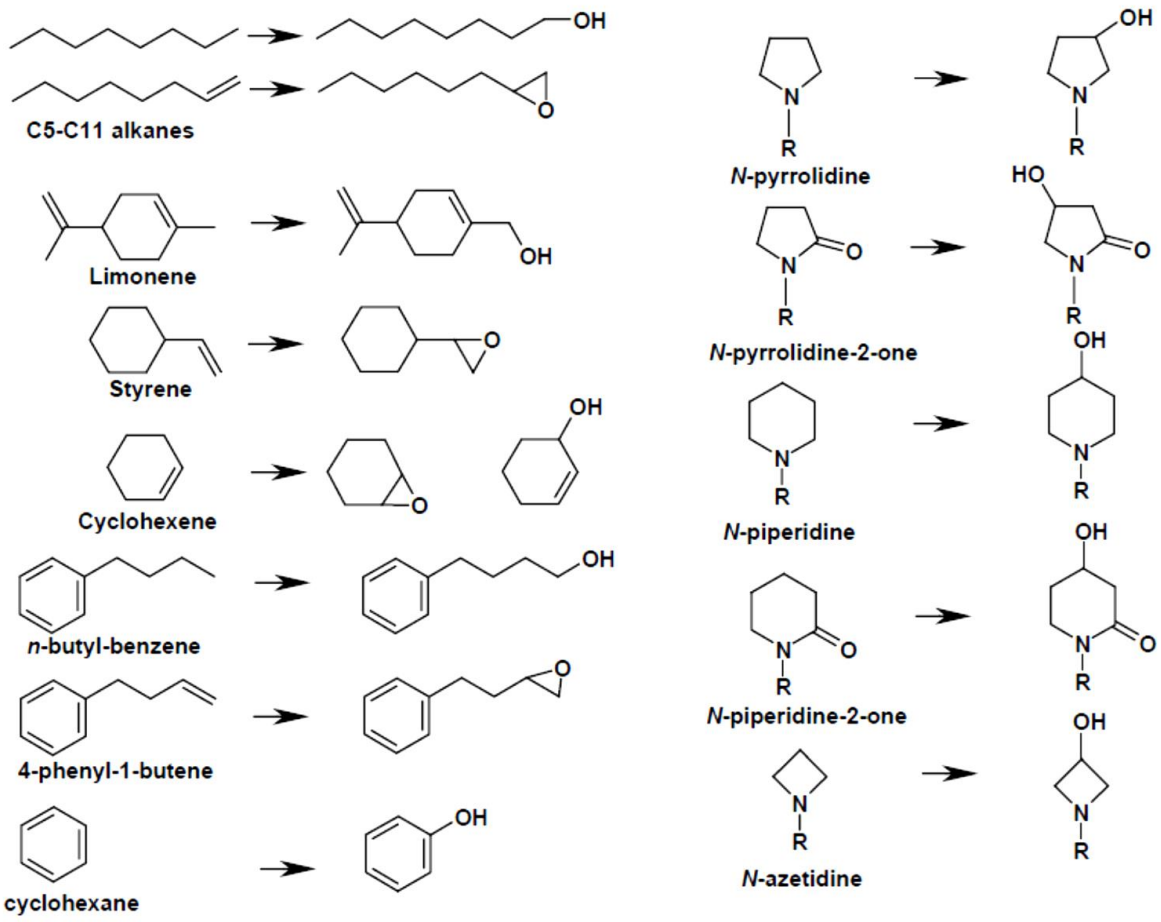

Figure 1.12 Reactions catalysed by CYP153 enzymes. (Funhoff & Van Beilen, 2007)

Funhoff and co-workers evaluated the hydroxylation and epoxidation reactions catalysed by eight CYP153 enzymes; in addition, they were able to better characterise the substrate specificity and regio-selectivity of these enzymes. Reactions were carried out using cell free extracts which were obtained using the French press because in previous experiments they observed this method yielded higher hydroxylation activities (Funhoff *et al*., 2006). The turnover numbers for each enzyme catalysing the various substrates in summarised in Table 1.4 (Funhoff *et al*., 2007). Besides CYP153A6 and CYP153A7, all the other CYP153s had negligible activity towards *n*-hexane. CYP153A13, D2 and D3 had the lowest turnover numbers for octane hydroxylation while CYP153A6 had the highest (58 min^{-1}). Decane was hydroxylated at the terminal carbon by CYP153A1, A6, A7 and A14 while CYP153A11 showed no activity towards this substrate and CYP153A13, D2 and D3 had significantly lower turnover numbers. CYP153A11 gave the highest turnover numbers for limonene and cyclohexane hydroxylation.

Table 1.2 Bacterial strains from which CYP153 genes were isolated, cloned and expressed for identification [Modified from (Van Beilen *et al.*, 2006)]

Bacterial stain	No. of CYP153 genes present[a]	P450 name	Accession no.	References (s)
Acinetobacter sp. strain EB104	1 (cloned)	CYP153A1	AJ311718	(Maier *et al.*, 2001)
Alcanivorax borkumensis AP1	2 (cloned)	CYP153A13	AJ844909	(Van Beilen *et a*l., 2004)
			AJ844908	
Alcanivorax borkumensis SK2	*2 (genome)*	*CYP153A13a*	AY505118	(Hara *et al*., 2004; Kubota *et al*., 2005; Nodate *et al*., 2006)
		CYP153A13b	CAL15649	
Bradyrhizobium japonicum USDA 110	*3 (genome)*	*CYP153A9*	*blr1853 NP_768493*	(Gottfert *et al*., 2001; Kaneko *et al*., 2002)
		CYP153A4	*blr7243 NP_773883*	
		CYP153A3	blr7242 NP_773882	
Burkholderia fungorum	*1 (genome)*	*CYP153A10*	*ZP_00028060 (obsolete version)*	
Caulobacter crescentus CB15	*1 (genome)*	*CYP153A2*	*CC0063 NP_418882*	(Nierman *et a*l., 2001)
Erwinia chrysanthemi 3937	*1 (genome)*	*CYP153E1*	*Unfinished genome*	
Novosphingobium aromaticivorans DSM1244	*2 (genome)*	*CYP153D1*	*ZP_00096754*	(Bell & Wong, 2007; Zhou *et al*., 2011)
		CYP153C1	ZP_00094181	
Mycobacterium marinum M	*1 (genome)*	*CYP153A14*	*Unfinished genome*	(Ramakrishnani & Falkow, 1994)
Mycobacterium sp. strain HXN-1500	*1 (PCR)*	*CYP153A6*	*AJ783967*	(Funhoff *et al*., 2006; Gudiminchi *et al*., 2012; Kubota *et a*l., 2005; Van Beilen *et al*., 2005)
Rhodopseudomonas palustrisCGA009	*1 (genome)*	*CYP153A5*	*ZP_00011192*	(Larimer *et al*., 2004)
Sphingomonas sp. strain HXN-200	*5 (cloned)*	*CYP153A7*	*AJ850057*	(Chang *et al*., 2002; Li *et al.*, 1999)
		CYP153A8	AJ850058	
		CYP153A11	AJ850059	
		CYP153D2	AJ850060	
		CYP153D3	AJ850057	

[a] Genome, the genome sequence contains one or more CYP153 sequences; PCR, the degenerate CYP153 or AlkB primers yielded one or more PCR fragments; cloned, the complete CYP153 genes were cloned and expressed in *P. putida* GPo12(pGEc47ΔB).

Table 1.3 Bacterial strains and plasmids used for the cloning and expression of CYP153 genes

Strain or plasmid	Relevent genotype or characterists	Reference
***P. putida* GPo12(pGEc47ΔB)**	*P. putida* GPo1 cured of the OCT plasmid	(Smits *et al*., 2002)
***E. coli* GEc137(pGEc47ΔB)**	DH-1 *thi fadR*	(Smits *et al*., 2002)
P. fluorescens KOB2Δ1	*alkB knockout mutant of CHA0*	(Funhoff *et al*., 2007; Smits *et al*., 2002)
***E.coli* BL21-A1**		(Kubota *et al*., 2005)
***E.coli* BL21 (DE3)**	*E. coli* B F^- *ompT hsdS*($r_B^-m_B^-$) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte	(Fujita *et al*., 2009; Gudiminchi *et al*., 2012; Pham *et al*., 2012; Yang *et al*., 2014)
pCom8	Broad-host-range expression vector with P*alkB*, Gmr *oriT alkS*	(Smits *et al*., 2001)
pKKPalk	*E. coli* expression vector with P*alkB*, Apr	(Van Beilen *et al*., 2005)
pCom8-$P_{A6}F_{1500}R_{1500}$	Broad-host-range expression vector with P*alkB*, Gmr, *oriT alkS* containing the HXN1500 *Mycobacterium* sp. CYP153A6, ferredoxin reductase and ferredoxin	(Van Beilen *et al*., 2005)
pCom8-$P_{A7}F_{200}$	pCom8 containing the HXN200 CYP153A7 and ferredoxin	(Van Beilen *et al*., 2006)
pCom8-$P_{A7}F_{200}R_{1500}$	As pCom8-PA7F200, including the HXN1500 ferredoxin reductase	(Van Beilen *et al*., 2006)
pCom12-$P_xF_{200}R_{1500}$	As pCom8-$P_{A7}F_{200}R_{1500}$, with the CYP153A1, -A5, -A7, -A8, -A11, -A12, -A13, -A14, -D2, or -D3 gene cloned between the EcoRI and PacI sites	(Funhoff *et al*., 2007; Van Beilen *et al*., 2006)
pRED	pET21a containing P450RhF linker and reductase domain	(Fujita *et al.,* 2009; Kubota *et al*., 2005; Nodate *et al*., 2006)
pT7NS-camAB	pT7NS plasmid containing P450cam reductase domain	(Fujita *et al*., 2009)
pET-28a (+)	Kan^R, T7*lac* promotor, T7 terminator, f1 ori, N-terminal His-tag and thrombin configuration (5 369 bp)	(Pham *et al*., 2012; Yang *et al*., 2014)
pET-28b (+)	Kan^R, T7*lac* promotor, T7 terminator, f1 ori, N-terminal His-tag and thrombin configuration (5 368 bp)	(Gudiminchi *et al*., 2012; Olaofe *et al*., 2013; Pennec *et al*., 2014)

Most of the CYP153s preferred aliphatic alkenes except CYP153A7 and A11 which showed higher activity towards cyclic substrates. In addition, styrene was predominantly epoxidised to *S*-styrene oxide by all the CYP153s except CYP153D2 which produced slightly more *R*-styrene oxide (Funhoff *et al*., 2007).

CYP153s catalysed the hydroxylation of linear alkanes with > 95% regio-selectivity for the terminal carbon and limonene conversion produced perillyl alcohol 100% of the time. Hydroxylation of cyclohexene was favoured by CYP153 enzymes over epoxidation with regio-selectivities of 75-80% and

20-25%, respectively. In addition, styrene-oxide was produced 100% of the time. However, the selectivity for epoxidation of octene varied between 8-100% depending on the enzyme (Funhoff *et al*., 2007).

Table 1.4 Hydroxylation and epoxidation activities of CYP153 enzymes with various substrates (Funhoff *et al*., 2007).

	Turnover number (min^{-1})							
	Hydoxylation					**Epoxidation**		
Enzyme	**Hexane**	**Octane**	**Decane**	**Limonene**	**Cyclo-hexane**	**Styrene**	**Octene**	**Cyclo-hexene**
CYP153A1[1]	0	34	10	3	2	2	6	> 0.1
CYP153A6	8	58	26	30	22	4	4	6
CYP153A7	8	28	19	12	37	40	20	7
CYP153A11	1	23	0	47	38	74	21	10
CYP153A13	< 0.1	2	7	2	10	1>	6	3
CYP153A14	< 0.1	25	21	8	13	18	31	3
CYP153D2	< 0.1	< 0.1	< 0.1	0.2	0.3	0.06	1.8	0.1
CYP153D3	0.1	< 0.1	0.6	< 0.1	0.3	0.1	1.5	0.06

Although, CYP153A7 is able to hydroxylate linear and cyclic alkanes and epoxidise linear alkenes and cycloalkenes, the CYP153s from *Sphingmonas sp*. HXN-200 (CYP153A7, A8, A11, D2 and D3) are better known for their ability to hydroxylate *N*-substituted substrates such as pyrrolidines, pyrrolidones, piperidines, piperidones and azetidines (Figure 1.12). The activities of CYP153D2 and D3 might have been so poor, because these enzymes prefer different substrates than were evaluated by Funhoff and co-workers (2007). In addition, these CYP153s have significant stereo-selectivities (e.g. > 99%) (Chang *et al*., 2004).

CYP153A13 from *Alcanivorax burkholderia* also had lower activity towards linear and cyclic alkanes which could be the result of insufficient electron transfer from its partner proteins, ferredoxin (*Sphingmonas sp*. HXN-200) and ferredoxin reductase (*Mycobacterium sp.* HXN-1500) (Funhoff *et al*., 2007; Van Beilen *et al*., 2006). In contrast, the activity of CYP153A13 from *Alcanivorax borkumensis* SK2 which was fused to the reductase domain of P450RhF had significantly improved activity towards the hydroxylation of *n*- alkanes, cyclic alkanes and *n*-butyl-benzene as well as epoxidation of 1-octene and 4-phenyl-1-butene (Kubota *et al*., 2005). Furthermore, by expressing the CYP153A13 fusion protein in *E. coli* BL21 (DE3), its activity towards *n*-octane, cyclohexane and *n*-butyl-benzene improved even more (Fujita *et al*., 2009). In addition, the fusion protein catalysed the hydroxylation of *n*-butyl-benzofuran to produce 4-benzofuran-2-yl-butan-1-ol which was further converted to 4-benzofuran-2-

[1] Mutated

yl-butyric acid (Figure 1.13). This was the first report that CYP153 could catalyse the terminal hydroxylation of *n*-butyl-benzofuran (Fujita *et al*., 2009).

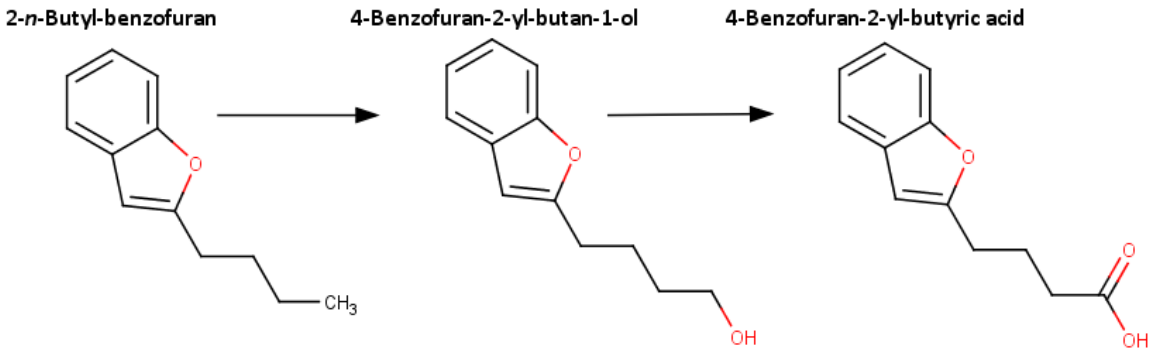

Figure 1.13 Hydroxylation of 2-*n*-Butyl-benzofuran to 4-benzofuran-2-yl-butan-1-ol by CYP153A13 fusion. HRESI-MS confirmed that the 4-benzofuran-2-yl-butan-1-ol was further converted to 4-benzofuran-2-yl-butyric acid. [Modified from (Fujita *et al*., 2009) using MarvinSketch 14.12.1.0]

Bordeaux and co-workers as well as Pennec and co-workers both reported improved activity for the self-sufficient CYP153A13 fusion protein toward *n*-octane hydroxylation therefore it's clear that CYP153A13 expressed as fusion protein (CYP153A13-RhFred)) clearly has higher turnover numbers for the hydroxylation of *n*-octane than when expressed as a three component system (CYP153A13FdRFdx) (Bordeaux *et al*., 2011; Pennec *et al*., 2014).

1.3.2.2 CYP153A6

CYP153A6 is currently the best characterised member of the CYP153 family. It catalyses terminal hydroxylation of C_6-C_{11} alkanes with *n*-octane shown to be the preferred substrate of this enzyme (Funhoff *et al*., 2006). It also has a regio-selectivity of 95% for the terminal carbon atom which is notably higher than other bacterial CYPs performing the same function; the other 5% in most cases is the product of subterminal hydroxylation (Funhoff *et al*. 2007). In addition, it catalyses the hydroxylation of limonene to produce perillyl alcohol which has been shown to have anti-carcinogenic properties and is a key component in cancer treatment (Van Beilen *et al*., 2005; Funhoff *et al*. 2006).

The terminal hydroxylation of *n*-alkanes by the CYP153A6 operon (three component system) has been evaulated in whole cells (WC) and cell free extracts (CFE). The activity of CYP153A6 in CFE of *P. putida* GPo12 for the hydroxlation of *n*-octane to 1-octanol was reported by Funhoff and co-workers (2006) to be 60 min^{-1}. In a subsequent experiment conducted by the same research group the turnover value for the production of 1-octanol by CYP153A6 was reported to be 58 min^{-1}. Therefore these values corresponded to the previous experiment (Funhoff *et al*., 2006; Funhoff & Van Beilen, 2007). However, Pennec and co-workers reported lower turnover numbers (22 min^{-1}) for 1-octanol production by CYP153A6 using CFE of *E. coli* (Pennec *et al*., 2014). For WC reactions the CYP153A6 operon was

expressed in *E. coli* BL21 (DE3). After using a biomass concentration of 7 g_{DCW} L_{BRM}^{-1}, Gudiminchi and co-workers reported that the concentration of 1-octanol after 15 h was 3.4 g L_{BRM}^{-1}. The same research group optimised conditions for biotransformation with CYP153A6 by changing the type of buffer from sodium phosphate to Tris-HCl, the buffer pH from 7.2 to 8 and increasing the glycerol concentration to 100 mM. This resulted in a 1-octanol concentration 2.6 g L_{BRM}^{-1} produced after only 8 h with a biomass concentration of 5 g_{DCW} L_{BRM}^{-1}. Thus the activity of CYP153A6 improved significantly under these conditions (Gudiminchi *et al*., 2012; Olaofe *et al*., 2013; Pennec *et al*., 2014). Because CYP153A6 uses a three component system, higher conversion of *n*-octane to 1-octanol is still currently observed in WC than CFE. This could be due to a lower effective concentration of ferredoxin and ferredoxin reductase in CFE when compared with WC where the electron transport proteins are kept in close proximity to the P450 by the cell membrane. Therefore investigating ways to overcome cofactor regeneration and bypass the redox partner proteins could improve the cell free system.

Currently there is no structure available for CYP153A6 and researchers have made use of homology models to understand why this enzyme hydroxylates C_6-C_{11} alkanes with >95% specificity for the terminal carbon atom. Funhoff and co-workers were able to build a homology model of CYP153A6 using the three dimentional structures of other bacterial P450 enymes as all P450 proteins share a specific fold. CYP153A6 shared 29% sequence identity with the template used and all conserved residues shared by P450 proteins were observed over the full gene sequence (Funhoff *et al*., 2006). According to the model, the active site of CYP153A6 was lined with hydrophobic residues which corresponds to the fact that the substrates prefered by this enzyme are non-polar in nature (Figure 1.14). Substrate docking simulations indicated C_5-C_{11} alkanes all coordinated in the same way, with the terminal carbon in a bent conformation towards the heme iron of the active site. This could be due to the location of Phe407 which apears to be contributing to this bent conformation of the alkane chain toward the active site.

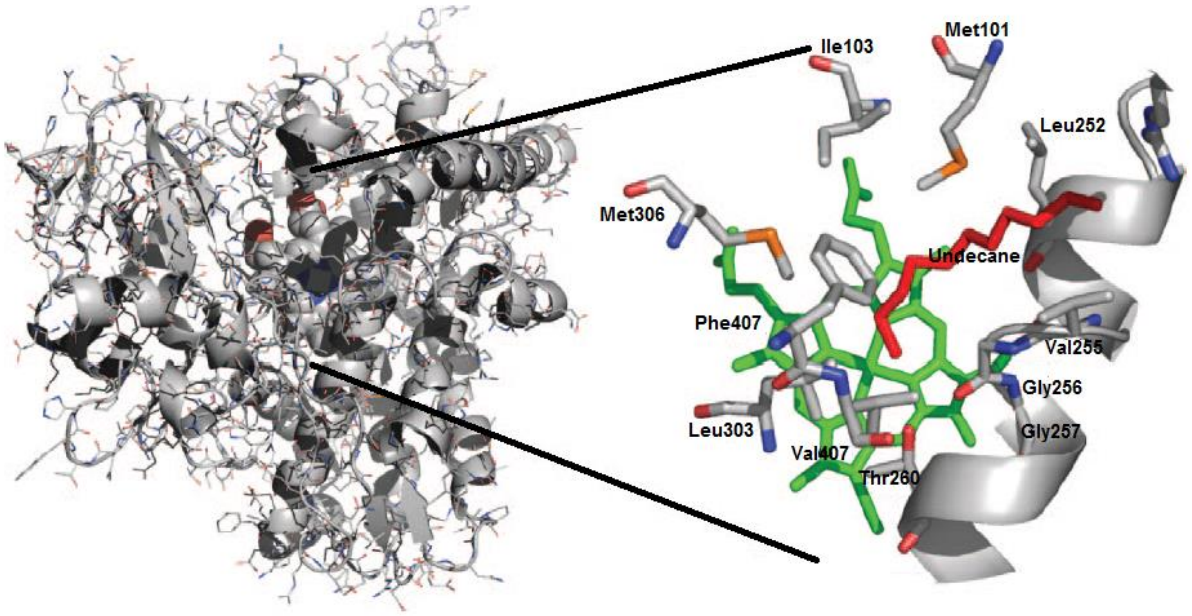

Figure 1.14 Homology model of CYP153A6 (Funhoff *et al*., 2006): Left panel: cartoon representation of the structural model of CYP153A6 generated with Modeler. The heme cofactor is presented space filled. Alkane substrates (pentane to undecane) were docked in the protein and shown to coordinate exclusively in the active site in a terminal position. Right panel: the coordination of undecane, in red, towards the heme cofactor (green) in CYP153A6. The central I-helix is shown, while residues coordinating the active site and substrates are depicted as stacks. (Funhoff *et al*., 2006)

1.3.2.3 Substrate binding studies

When P450s are in a substrate-free form, a water molecule forms the sixth ligand of the hexa-coordinated heme iron. The enzyme is therefore said to have a low-spin state and the soret absorption maximum in a UV spectrum is approximately 420 nm. However, when a substrate binds in the active site, it replaces the water molecule and the heme iron becomes penta-coordinated. The enzyme spin state shifts to a high spin state which results in maximum in the UV spectrum changing from 420 nm to approximately 390 nm (Poulos *et al*., 1987). This type of substrate binding is referred to Type 1 substrate binding and is a common characteristic shared by cytochrome P450 enzymes (Figure 1.15).

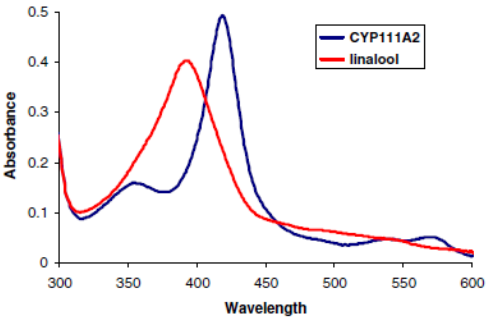

Figure 1.15 Example of the spin-state shift of CYP11A2 when unbound (blue) or bound (red) to linalool. (Bell & Wong, 2007)

This type of substrate binding studies have been used to determine if alkanes and other substrates induce this kind of spin-state shift when bound to the active site of CYP153 enzymes. Funhoff and co-workers employed this characteristic of P450 enzymes to determine the substrate scope of CYP153A6 (Figure 1.16). Most of the *n*-alkanes induced this signature shift in the absorbance spectrum; the

largest change was observed for *n*-octane and *n*-nonane while *n*-hexane and *n*-undecane induced minimal changes in the position of the soret band (Funhoff *et al*., 2006). This indicated that the amount of high-spin formation was possibly affected and/or dependant of chain length. Substrates with methyl side groups such as 2-methyl-octane also induced a high-spin shift in the absorbance spectrum, whereas substrates with dimethyl side groups did not. In addition, substrates like limonene, *p*-cymene and ethyl-cyclohexane did not induce strong spectral shifts either. Overall, alkanes appeared to bind more tightly than other substrates indicating that the binding pocket of CYP153A6 is shaped in such a way that linear alkanes are preferred over cyclic and aromatic substrates (Funhoff *et al*., 2006).

Bell and Wong also evaluated the substrate specificity of CYP153C1 and CYP153D1 from *N. aromaticivorans* DSM12444 using spectroscopy. They evaluated substrates such as short and medium chain (C_6-C_{12}) linear alkanes, polycyclic aromatic hydrocarbons, substituted phenols, polychlorinated benzenes, and steroidal compounds to name a few. Despite the broad range of substrates evaluated, none of these substrates induced a spin-state shift with CYP153D1; on the contrary, CYP153C1 like CYP153A6 had a preference for linear alkanes with optimal binding observed for *n*-heptane, *n*-octane and *n*-nonane which all induced a 80% high-spin state (Bell & Wong, 2007). In addition, 1-octene also induced an 80% high-spin state indicating that in addition to hydroxylating linear alkanes, CYP153C1 also catalyses the epoxidation of linear alkenes.

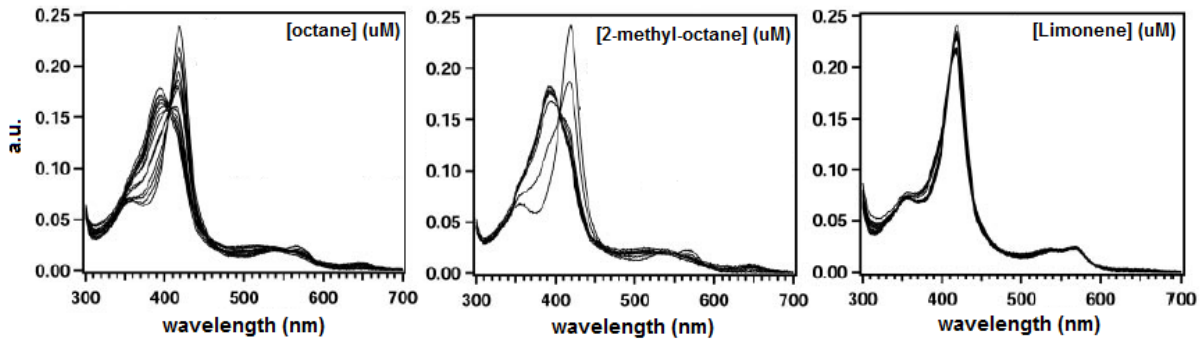

Figure 1.16 Example of the substrate titration curves used to evaluate which substrates induced a Type I spin-state shift in CYP153A6. Octane and 2-methyl-octane yielded almost 100% high spin formation while limonene did not induce a strong spectral shift. (Funhoff *et al*., 2006)

Substrate binding studies were also performed by Bordeaux and co-workers, they focused their research on the self-sufficient CYP153A13, a fusion protein (Kubota *et al*., 2005). Titration experiments were performed with *n*-octane, *n*-decane, *n*-dodecane, and cyclohexane. Spectral shifts similar to CYP153A6 (Funhoff *et al*., 2006) were observed for CYP153A13a which also interacted with alkanes however to a lesser extent (Bordeaux *et al*., 2011). Titration curves revealed that *n*-dodecane bound more tightly in the active site of CYP153A13a than *n*-octane and *n*-decane; in addition, cyclohexane also induced a high-spin state for the production of 1-cyclohexanol (Figure 1.17). These results

indicated that both linear and cyclic alkanes have high affinity for the active site of the CYP153A13a-fusion

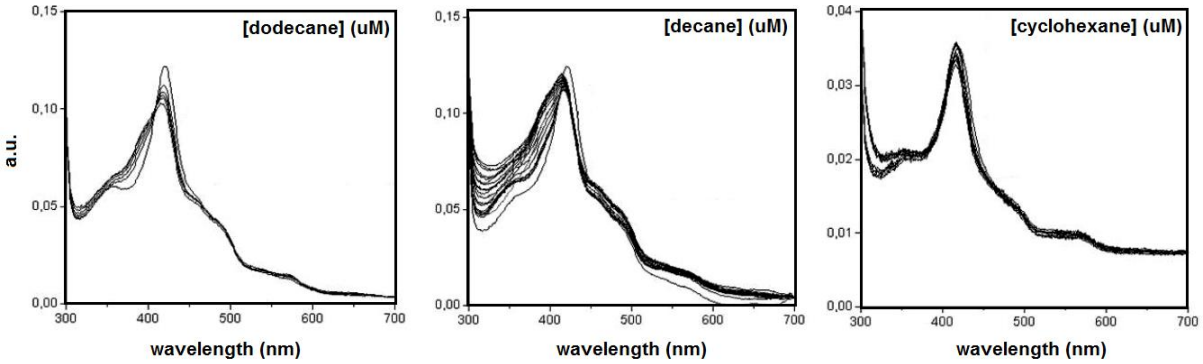

Figure 1.17 Titration curves used to evaluate which substrates induced a Type I spin-state shift in CYP153A13 (fusion protein). Dodecane yielded a stronger binding affinity than octane and decane; in addition, cyclohexane also yielded maximal binding affinity in the active site. (Bordeaux *et al*., 2011)

1.3.2.4 Insight into the mechanism of CYP153 enzymes

In addition to generating a homology model of CYP153A6 to gain some insight regarding the unique characteristics of this enzyme, Funhoff and co-workers subsequently conducted a multiple sequence alignment of eight CYP153s with P450cam (Funhoff *et al*., 2007). Sequence analysis suggested that all presumed residues in the active site of CYP153s are completely conserved suggesting that these residues do not play a role a role in determining substrate preference (Funhoff *et al*., 2007). However CYP153D2 and D3 were expections because in one location (β'-helix) of the active site D2 and D3 contained a proline and arginine residue, respectively whereas the other CYP153s contained a negatively charged residue (Table 1.5). In addition, the residue after proline in D2 was methionine and in D3 the residue next to arginine was glutamine whereas the remaining CYP153s all contained hydrophobic residues (alanine, valine, isoleucine and leucine) in the corresponding postions. Another interesting difference observed between the CYP153s was that CYP153A7 contained an aparagine residue (Asn100) where all other CYP153s contained a methionine residue in the corresponding position (β'-helix). Futhermore, the homology model of CYP153A6 (discussed previously) suggested that position of phenylalanine (Phe407) was resposible for directing linear alkanes towards the heme iron in the active site; all the CYP153s evaluated here also contained a phenylalanine residue in the corresponding position except CYP153A11 which contained a leucine residue (β5). This could indicate why cyclic compounds were found to be the prefered substrates of CYP153A11 because leucine is not a "bulky" residue like phenylalanine resulting in more space in the active site of A11 to accommodate cyclic compounds. In P450cam, polar residues in the active site were mutated to hydrophobic residues in order to change the enzymes substrate preference from camphor to alkanes (Tyr96Phe and Thr101Leu).

1.3.3 Structure of CYP153 hydroxlyases

Although information regarding the function and catalytic mechanisms of various CYP153 enzymes has increased, only one structure has been determined since the discovery of the first CYP153 enzyme in 2001 (Maier *et al*., 2001).

Zhou and co-workers came close to solving the structure of CYP153C1 from *N. aromaticinorans*. Initially they managed to obtain several dark red crystals (Figure 1.18 (a)) however, none of them were suitable for X-ray diffraction because of their non-single thin plate morphology (Zhou *et al*., 2011). The conditions for crystallisation were optimised by varying the buffer pH, precipitate concentration, protein concentration, and temperature. After several weeks, long single rod-shaped crystals were obtained (Figure 1.18 (b)) and the diffraction quality of these crystals was significantly better than before. They proceeded to use several P450 structures to solve the phases of CYP153C1 but unfortunately, none of their attempts were successful.

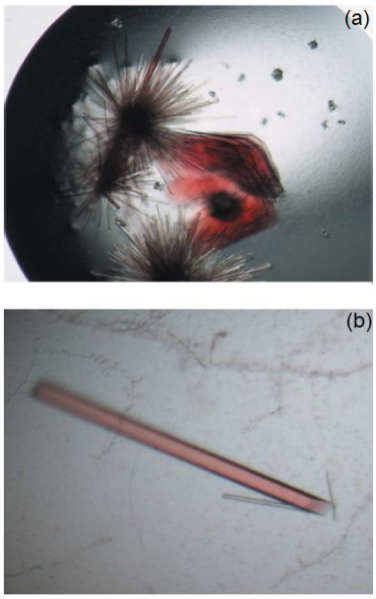

Figure 1.18 Crystals of CYP153C1 after initial screening (a) Crystals of CYP153C1 after optimisation (b). (Zhou *et al*., 2011)

The following year, the structure of CYP153A7 (also known as P450pyr hydroxylase) from *Sphingomonas sp.* HXN-200 was determined by Pham and co-workers. An *E. coli* Top10 strain carrying plasmid pTr99A-His-P450pyr expressed P450pyr with a poly-histidine tag at the N-terminus. Fast protein liquid chromatography (FPLC) was used to purify the his-tagged P450pyr together with His-tagged ferredoxin and His-tagged ferredoxin reductase using a Ni-NTA column and a size-exclusion column. His-tagged P450pyr was then crystallised and its structure determined by X-ray diffraction (PDB ID: 3RWL) (Figure 1.19 (a)). The structure of CYP153A7 exhibits an overall fold of P450 hydroxylases with the heme located at the bottom of a gorge leading to the active site. The large size of the pocket also explains the broad substrate range of these enzymes (Pham *et al*., 2012).

Figure 1.19 illustrates the difference in structure between three *n*-alkane hydroxylatng monooxygenases: (a) CYP153A7 [PDB ID: 3RWL] which catalyses terminal oxidation of medium-chain alkanes (C_6-C_{10}), N-substituted pyrolidines and the epoxidation of cyclic-alkanes; (b) Particulate methane monooxygenase (pMMO) [PDB ID: 1YEW] from *Methylococcus capsulatus* which catalyses terminal oxidation of small-chain alkanes (C_1-C_5) and (c) Long-chain alkane monooxygenase (LadA) [PDB ID: 3B9N) from *Geobacillus thermodenitrificans* which catalyses terminal oxidation of linear alkanes up to C_{36}.

Table 1.5 Alignment of residues lining the active site of CYP153s with residues in P450cam. [Information from (Funhoff *et a*l., 2007)]

Second structure	A1[a]	A6	A7	A11	A13	A14	D2	D3	P450cam
B'-helix	**Asp167**	**Ser98**	**Asp97**	**Glu105**	**Ser140**	**Asp132**	**Pro96**	**Arg97**	*Tyr96Phe*
	Val168	**Leu99**	**Leu98**	**Leu106**	**Val140**	**Ile133**	**Met97**	**Gln98**	Asp97
	Met170	Met101	Asn100	Met108	Met143	Met135	Met99	Met100	Ile99
	Ile172	Ile103	Ile102	Ile110	Ile145	Ile137	Ile101	Ile102	*Thr101Leu*
	Val266	**Ile195**	**Val194**	**Val202**	**Phe239**	**Ser231**	**Leu193**	**Pro193**	Pro187
I-helix	Asn327	Asn249	Asn248	Asn256	Asn300	Asn291	Asn247	Asn247	Met241
	Leu328	**Ile250**	**Val249**	Leu257	Leu301	Leu292	Leu248	Leu248	Cys242
	Val329	**Val251**	**Leu250**	**Leu258**	**Ala302**	**Val293**	**Ile249**	**Val249**	Gly243
	Leu330	Leu252	Leu251	Leu259	Leu303	Leu294	Leu250	Leu250	*Leu244Met*
	Leu331	Leu253	Leu252	Leu260	Leu304	Leu251	Leu251	Leu195	Leu245
	Ile332	Ile254	Ile253	Ile261	Ile305	Ile295	Ile252	Ile252	Leu246
	Val333	Val255	Val254	Val262	Val306	Val296	Val253	Val253	*Val247Leu*
	Gly334	Gly256	Gly255	Gly263	Gly307	Gly198	Gly254	Gly254	*Gly248Ala*
	Gly335	Gly257	Gly256	Gly264	Gly308	Gly199	Gly255	Gly255	Gly249
	Asp337	Asp259	Asp258	Asp266	Asp310	Asp301	Asp257	Asp257	Asp251
	Thr338	Thr260	Thr259	Thr267	Thr311	Thr202	Thr258	Thr258	Thr252
β3	Pro380	Pro302	Pro301	Pro309	Pro353	Pro344	Pro302	Pro300	*Leu294Met*
	Leu381	Leu303	Leu302	Leu310	Leu354	Leu345	Leu303	Leu301	Val295
	Met384	Met306	Met305	Met313	Met357	Met306	Met348	Met304	Gly298
β5	Phe482	Phe407	Phe403	**Leu411**	Phe455	Phe446	Phe404	Phe403	Ile395
	Val483	Val408	Val404	Val412	Val456	Val447	Val405	Val404	Val396

Residues in bold are not conserved, if compared with the predominant residue at this position in the CYP153 family, or are not conserved at all. Residues in P450cam, including those mutated, that affect the substrate specificity (from camphor hydroxylation towards alkane hydroxylation), are underlined and in italics. (Xu *et al*., 2005)

[a] Mutated

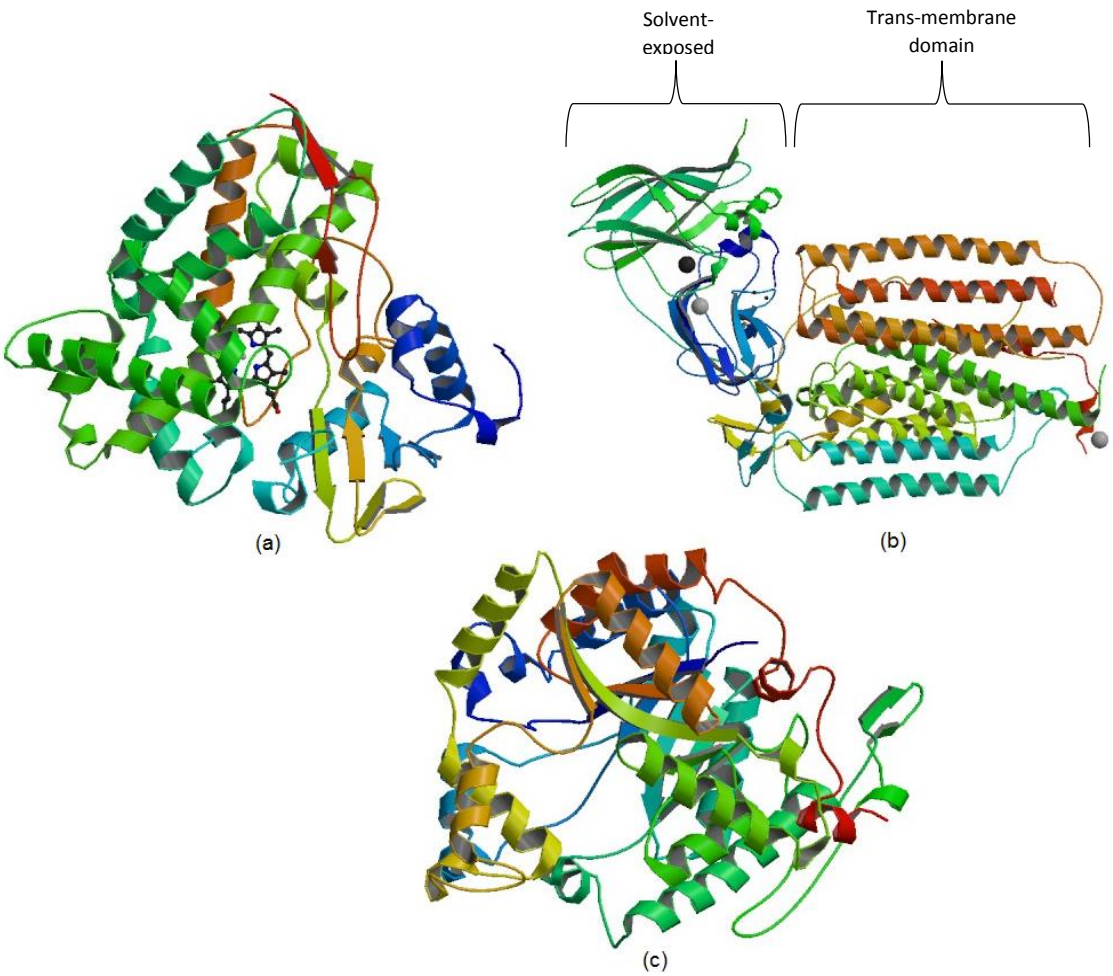

Figure 1.19 Ribbon structures of three alkane hydroxylating monooxygenases: (a) CYP153A7 [PDB:3RWL] from *Shingomonas sp*. HXN-200 which catalyses terminal oxidation of medium-chain alkanes (C_6-C_{10}); (b) Particulate methane monooxygenase (pMMO) [PDB:1YEW] from *M. capsulatus* which catalyses terminal oxidation of small-chain alkanes (C_1-C_5) and (c) Long-chain alkane monooxygenase (LadA) [PDB:3B9N] from *Geobacillus thermodenitrificans* which catalyses terminal oxidation of linear alkanes up to C_{36}. (Li *et a*l., 2008; Lieberman & Rosenzweig, 2005; Pham *et al*., 2012)

At first glance, the difference in size between the binding pockets of CYP153A7 and LadA can clearly be seen. The structure of LadA reveals a large and deep binding pocket with the active site buried deep within the structure resulting in the terminal carbon of *n*-alkanes smaller than C_{14} unable to reach the active site (Li *et al*., 2008). The ability of LadA to degrade long chain alkanes does not rely on its catalytic abilities but more its ability to capture and accommodate these long substrates. Even though the binding pocket of CYP153A7 is not as large as that of LadA, it does not need to be because long-chain alkane hydroxylation is not one of its functions. Also, the heme located in the active site can clearly be seen indicating that the active site is not buried too deep, these structural features are therefore ideal for the terminal carbon of

medium-chain alkanes to reach the active site. However, the binding pockets of both LadA and CYP153A7 are lined with hydrophobic residues which enable their respective non-polar substrates to bind.

The structure of pMMO is very different to that of CYP153A7 and LadA since they are both soluble proteins and pMMO is membrane-bound. Particulate methane monooxygenase consists of three subunits pmoA, pmoB and pmoC, a solvent exposed domain above the trans-membrane domain which consists of α-helices (Himes & Karlin, 2009). The three subunits form a cylindrical trimer with a hole in the centre which extends into the membrane where it is lined with hydrophobic residues which would supports the binding of the non-polar substrate, methane (Lieberman & Rosenzweig, 2005). However, the dicopper centre located in the soluble domain of pMMO was found to be the active site, with the transmembrane regions speculated to be involved in electron transfer to the active site, as well as facilitating substrate entry through the lipid bilayer as methane is more soluble in the membrane (Balasubramanian *et al*. 2010).

Determining the X-ray structure of CYP153A7 is considered to be a huge breakthrough because it will allow more accurate modelling of other members of the CYP153 family such as CYP153A6.

1.3.4 Diversity of CYP153 hydroxylases in microorganisms and environments

The CYP153 family consists of five subfamilies (CYP153A, CYP153B, CYP153C, CYP153D and CYP153E) which were identified from numerous bacterial strains isolated from various environments. For instance, CYP153A6 from *Mycobacterium sp.* HXN-1500 a genus from Acinetobacteria and CYP153A7 from *Sphingmonas sp*. HXN-200 a genus from Proteobacteria were isolated from a trickling-bed air filter in Stuttgart, Germany (Van Beilen *et al*., 2005) but can also be found in various terrestrial, freshwater and marine environments (Nie et al., 2014); CYP153A13 from *Alcanivorax borkumensis* a marine bacterium was isolated from petroleum-contaminated soil and groundwater (Kubota *et al*., 2005) and *Novosphingobium aromaticivorans* from which CYP153C1 and CYP153D1 were identified is characterised as an oligotrophic bacterium which means it can thrive in environment's where the nutrient source is significantly low such as deep soil and water sediments (Bell & Wong, 2007).

Oil-polluted environments such as oil spills in the ocean that also spread to the coast contain alkane-degrading bacteria. To demonstrate the diversity of CYP153 genes spread throughout various bacterial genera, Wang and co-workers did a phylogenetic study of 117 bacterial strains that were isolated from 18 different oil-degrading consortia. From 63 of the 117 strains, they derived 82 sequences belonging to the CYP153A subfamily; most of the strains belonged to Proteobacteria and included genera's such as *Alcanivorax*, *Erythrobacter*, *Martelella*, *Parvibaculum* and *Salinisphaera* (Wang *et al*., 2010). After

conducting phylogenetic analysis based on amino acid sequence alignment of CYP153 family members, it was clear that CYP153 genes isolated from different strains belonging to the same genera, grouped together in the phylogenetic tree, indicating that the phylogeny of the CYP153 genes correlated with the phylogeny of the host. In addition, the sequences derived from *Alcanivorax strains* formed a large group that span three clusters of the phylogenetic tree. These sequences shared 75-100% sequence identity with CYP153A13 from *Alcanivorax burkumensis* SK2 indicating that they belong to the CYP153A subfamily (Nelson *et al*., 1996). Wang and co-workers also identified isolates that were not known to contain P450 enzymes and was demonstrated here for the first time: *Brachybacterium*, *Halomonas*, *Idiomarina*, *Leifsonia, Mesorhizobium*, *Martelella*, *Ochrobactrum*, *Salinisphaera*, *Sphingobium*, *Solimonas*, *Tetrathiobacter* and *Tistrella.* Phylogenetic analysis indicated that P450 enzymes identified in these isolates represented novel members of the CYP153A subfamily (Wang *et al*., 2010).

Another research group were able to identify 130 CYP153 genes from 87 genomes of bacterial strains belonging to the Proteobacteria, Actinetobacteria and Planctomycetes which were isolated from various terrestrial, freshwater and marine environments (Nie *et al*., 2014). In addition, they also determined that CYP153 genes were the most abundant in terrestrial bacteria (585 homologs) and the least abundant in bacteria isolated from freshwater (43 homologs).

Although most genes coding for CYP153s are found in AlkB-lacking stains (Van Beilen *et al*., 2006) the co-existence of multiple alkane hydroxylases in the same host seems common in *Alcanivorax* strains (Nie *et al*., 2014; Nie *et al*., 2013). This could be why *Alcanivorax* strains are reported to be dominant members of oil-degrading consortia because the presence of multiple alkane hydroxylases enables them to degrade alkanes of various chain lengths and allows them to adapt better in different environments (Nie *et al*., 2014).

1.3.5 Applications of CYP153 hydroxylases

Alkane hydroxylases play an important role in tackling oil pollution because they convert oil components especially xenobiotics compounds that are hazardous to more valuable oxygenated derivatives (Van Beilen & Funhoff, 2007).

Initially, the CYP153s became the focus of biotransformation processes for the synthesis of pharmaceutical intermediates. This included perillyl alcohol, a putative agent in cancer treatments, which is produced from limonene by CYP153A6 (Van Beilen *et al*., 2005), while *Sphingomonas sp* HXN-200 which contains five CYP153s (CYP153A7, A8, A11, D2 and D3) have been used in hydroxylation of pyrrilodines,

piperdines and azetidines to prepare pharmaceutical intermediates for antibacterial and hypersensitive drugs (Chang *et al*., 2002; Li & Chang, 2004; Van Beilen & Funhoff, 2007). In addition, the use of microorganisms or enzymes as biocatalysts in the petrochemical industry has become more attractive for the production of value added chemicals because they are able to efficiently use non-conventional substrates such as linear alkanes as an inexpensive hydrocarbon feedstock and are more environmentally friendly, in contrast to chemical catalysts which are not only expensive but also harmful to the environment (Olaofe *et al*., 2013). CYP153 hydroxylases have the ability to hydroxylate the terminal carbon of aliphatic alkanes, the most unreactive carbon in these molecules (Funhoff *et al*., 2007). Selective terminal hydroxylation of *n*-alkanes has become of great interest in industry, because no chemical catalyst has yet been developed that can give absolute selective terminal hydroxylation of *n*-alkanes.

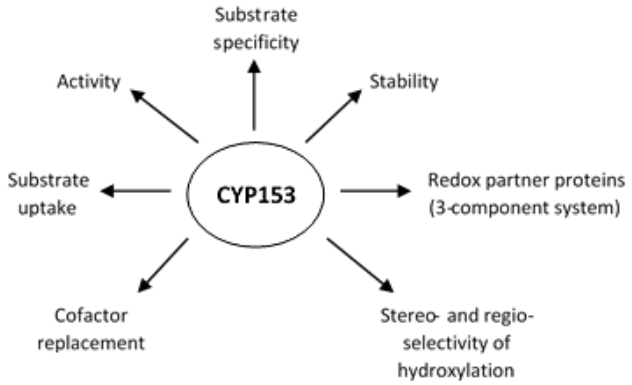

Figure 1.20 Factors hampering CYP153 hydroxylases for large-scale application. [Modified from (Bernhardt, 2006)]

However, the use of CYP153 hydroxylases for industrial processes is hampered by factors such as low enzyme activity, biocatalyst instability and the requirement for reduced (expensive) co-factors as well as the fact that they use a multi-component system for electron transfer (Figure. 1.20) (Funhoff & Van Beilen, 2007; Olaofe *et al*., 2013).

Whole cells are used most often to overcome the limitations of cofactor regeneration and a three-component system. The cell keeps the ferredoxin and ferredoxin reductase in a confined space with CYP153 to allow more efficient electron transfer. In addition, the cellular metabolism regenerates the necessary cofactors. However, substrate uptake, product toxicity and oxygen transfer are often problems associated with whole cell systems (Olaofe *et al*., 2013). Researchers have been tirelessly investigating ways to replenish cofactors in a cell free system using for example electrochemical (Kochius *et al*., 2014) or enzymatic (e.g. glucose dehydrogenase) (Pennec *et al*., 2014; Weckbecker & Hummel, 2005) methods

(Figure 1.21). Processes using the peroxide shunt pathway eliminate the use of redox partner proteins and cofactors altogether because hydrogen peroxide acts as source of electrons and oxygen for heme proteins. Therefore, researchers are also using protein engineering to give hydroxylases peroxygenase properties (Cirino & Arnold, 2002, 2003). However, this has yet to be achieved for CYP153 hydroxylases.

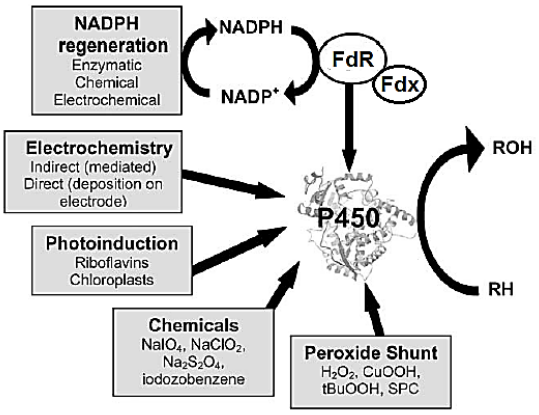

Figure 1.21 Natural cofactors of CYP153 hydroxylases (FdR/Fdx and NADPH/NADH) and diffierent statergies in place to regenerate or replace cofactors. [Modified from (Chefson & Auclair, 2006)]

CYP153A6 has been extensively studied in whole cells and cell free extracts because the high regio-selectively of this enzyme for the terminal carbon atom makes it a good candidate for industrial application (Funhoff *et al*., 2006; Gudiminchi *et al*., 2012; Olaofe *et al*., 2013; Pennec *et al*., 2014). However, the instability of this enzyme appears to be its biggest limitation.

Construction of the self-sufficient CYP153A13 from *Alcanivorax burkumensis* SK2 has also shown to improved activity towards the hydroxylation of n-octane compared to the CYP153A13 in a three-component system (Fujita *et al*., 2009; Kubota *et al*., 2005; Nodate et al., 2006). Recently, CYP153A16 from *M. marinum sp* was fused to the reductase domain of P450BM3 (NADPH-cytochrome P450 reductase (CPR)) for the terminal hydroxylation of C_{12} fatty acids (Scheps *et al*., 2013). The fusion protein yielded ω-hydroxy-dodecanoic acid with high regio-selctivity (> 95%) for the terminal carbon.

After determining the structure of CYP153A7, Pham and co-workers were able to develop a mutant that improved the enantio-selectively of CYP153A7 towards the production of S-N-benzyl-3-hydroxpyrrilodine (98%). Recently, CYP153A7 has been engineered for sub-terminal hydroxylation of n-alkanes (at C-2) with high regio- and enantio-selectivity (Yang *et al*., 2014).

The conversion of methane to methanol has become a potential solution for to get “standard gas” to consumers (Van Beilen & Funhoff, 2007) however, product inhibition, enzyme instability and poor heterologous expression are major problems when using methane monooxygenases (Lee *et al*., 2004). CYP153A6 has been engineered for the terminal hydroxylation of small chain alkanes as a possible solution for this problem (Koch *et al.*, 2009); in addition, other P450 enzymes such as P450cam (Bell *et al*., 2002)

and P450BM3 (Meinhold *et al.*, 2006) have also been engineered for terminal hydroxylation of short-chain alkanes. However, their regio-selectivity is still quite poor.

Protein engineering appears to be having the most success for overcoming the limitations observed for CYP153 enzymes however, it is important to remember that only few of the enzymes from this family (i.e. CYP153A6, CYP153A7 and CYP153A13) have been extensively studied. In addition, only one structure (CYP153A7) for these enzymes has been solved thus far, which limits the full potential of what can be achieved using protein engineering.

1.4 Concluding remarks

CYP153 hydroxylases were the first soluble bacterial enzymes identified that possessed the ability to catalyse hydroxylation of the terminal carbon of aliphatic alkanes, the most unreactive carbon in the molecules. Since the discovery of the first CYP153 by Maier and co-workers in 2001, our knowledge regarding this enzyme family has grown but at the same time, it is still very limited.

It is clear from phylogenetic analysis that genes encoding CYP153 hydroxylases are widespread in alkane degrading bacteria from various environments. Although, many CYP153 sequences have been identified from various strains of bacteria, only a few have been named and functionally characterised, which limits the understanding of the full potential of this enzyme family. Sequence analysis has also shown that most of the unidentified CYP153 genes belong to the CYP153A subfamily which is already/currently the largest subfamily of the five subfamilies.

Only one structure (CYP153A7, PDB ID 3RWL) is available for the CYP153 family. Other research groups have attempted solving the structures of other CYP153s but have been unsuccessful thus far which could be due to the complexity of these enzymes. The structure of CYP153A7 clearly shows that the enzyme has a fairly large binding pocket which clarifies why these enzymes are able to accept a broad range of substrates ranging from linear alkanes to aromatic substrates. In addition, since the structure of CYP153A7 became available, protein engineering has really taken off for CYP153 hydroxylases not only for CYP153A7 - which has now been engineered for subterminal hydroxylation of alkanes - but for other CYP153s as well, because it has allowed more accurate modelling. It's also possible that the structure of CYP153A7 will provide insight for solving other structures in the future.

Properties limiting the industrial application of the CYP153s include their dependence on two electron transfer proteins and their instability. The former has been addressed by the construction of artificial fusions, but these are also not very successful. The instability problem has not yet received much attention.

1.5 Aim of the Study

CYP153s which are found in bacteria are very versatile cytochrome P450 enzymes which hold great promise for industrial application, because of the unique regioselectivity displayed during conversion of *n*-alkanes and because they accept a broad range of substrates. However, the possibility of using CYP153s for large-scale applications is hampered by their three main limitations: (1) they are cofactor dependant; (2) they require redox partner proteins – Ferredoxin (Fdx) and Ferrodoxin reductase (FdR) - for electron transfer during catalysis and (3) they are unstable.

Biocatalyst instability among P450s is largely caused by hydrogen peroxide (H_2O_2) which is produced due to protein uncoupling (Loida & Sligar, 1993). Protein uncoupling is usually observed in the absence of substrate or in presence of compounds that do not fit the active site. Reduction of molecular oxygen without subsequent substrate oxidation results in formation of H_2O_2 (Van Beilen *et al.*, 2003). Sulphur-containing amino acids – cysteine and methionine – are the most susceptible to oxidation by various reactive oxygen species such as H_2O_2 (Opperman & Reetz, 2010a) causing enzyme instability. This phenomenon has been demonstrated by different groups when substituting Cys and Met residues with more oxidatively stable amino acids yielded enzymes with improved stability (Heinzelman *et al*., 2009; Lin *et al*., 2006; Miyazaki-Imamura *et al*., 2003; Opperman & Reetz, 2010a; Slavica *et al.*, 2005).

CYP153A6 was selected to investigate problems associated with operational stability of CYP153s, because it is a well characterized member of this family: It catalyses the terminal hydroxylation of C_6-C_{11} alkanes with *n*-octane shown to be the preferred substrate of this enzyme. It also has a high regio-selectivity for the terminal carbon atom producing 95% 1-alkanols and 5% 2-alkanols (Funhoff *et al*., 2006; Funhoff *et al*., 2007).

The overall aim of this study was to create mutants of CYP153A6 that would improve operational stability without compromising enzyme activity or product specificity (1-octanol).

The first aim of this study was to establish factors affecting the operational stability of the enzyme and conditions that could be used to challenge mutants. In order to achieve this octane hydroxylation

reactions with wild-type CYP153A6 were performed under various conditions using cell free extracts. These conditions included: (1) Incorporating glucose dehydrogenase from *Bacillus megaterium* to facilitate cofactor regeneration; (2) adding additional FdR and Fdx to facilitate electron transfer, (3) adding additional NADH to ensure sufficient reducing equivalents for electron transfer; (4) varying buffer concentration and pH and (5) varying P450 concentration. After evaluating the effects of these various conditions on the activity and stability of CYP153A6, a final set of conditions was selected for testing mutants.

The second aim of this study was to design and construct mutants of CYP153A6 using site directed mutagenesis to substitute specific cysteine and methionine residues with amino acids that are more oxidatively stable in the presence of H_2O_2. Programs such as 3DM Bio-Prodict (http://www.bio-prodict.nl/) and Yasara (Krieger & Vriend, 2002) were used to carefully select surface exposed as well as internally located cysteine and methionine residues for mutation. The 3DM database also aided selection of amino acid to use for substitutions based on multiple sequence alignments. In addition, the availability of the CYP153A7 (3RWL) structure facilitated construction of a homology model of CYP153A6 used for mapping the positions of cysteine and methionine residues that were to be substituted.

The third aim of this study was to evaluate the operational stability of the CYP153A6 mutants. This was carried out by performing biotransformations of *n*-octane - under the conditions that were selected from aim one – using cell free extracts. In addition, the oxidative stability of the mutants were challenged by adding glucose oxidase to biotransformation reactions for *in situ* generation of H_2O_2.

Chapter 2 focuses on the evaluation of the operational stability of the wild-type CYP153A6 (first aim) while chapter 3 focuses on the design and construction of the CYP153A6 mutants (second aim) as well as evaluating the operational stability of the mutants compared to wild-type (third aim).

Chapter 2

Factors affecting operational stability of CYP153A6 in cell free extracts (CFEs)

CYP153A6 is a well-characterised member of the CYP153 family. It catalyses the terminal hydroxylation of C_6-C_{11} alkanes with regio-selectivity of 95% for the terminal carbon atom (Funhoff *et al*., 2006) however, CYP153A6 has three main limitations: (1) it uses a three component system i.e. it requires redox partner proteins, ferredoxin (Fdx) and ferredoxin reductase (FdR) for electron transport, (2) it is NADH-dependant and (3) it is unstable. Whole cell biotransformations used most often to overcame these problems because the cell keeps the proteins (CYP153A6, FdR and Fdx) in a confined space allowing more efficient transfer of electrons and the cell metabolism is able to take care of cofactor regeneration. However, the instability of CYP153A6 is observed in whole cells nonetheless (Gudiminchi *et al*., 2012; Olaofe *et al*., 2013) therefore, this chapter focuses on evaluating various conditions to try to improve the operational stability of CYP153A6 (wild type) by performing biotransformations of *n*-octane (the preferred substrate of CYP153A6) using cell free extracts (CFE).

2.1 Materials and Methods

Section A - General methods

2.1.1 Bacterial strains and plasmids

All bacterial strains and plasmids are listed in Table 2.1. The complete operon of CYP153A6 from *Mycobacterium sp.* HXN-1500, also encoding ferredoxin reductase (FdR) and ferredoxin (Fdx) had previously been cloned into pET-28b(+) (Gudiminchi *et al*., 2012). The genes encoding glucose dehydrogenase (BmGDH) from *Bacillus megaterium sp.*and the partial operon encoding FdR and Fdx (FdR/Fdx) from *Mycobacterium sp.* had each been cloned into separate pETDuet-1 vectors.

Competent *Escherishia coli* (*E. coli*) Top 10 cells (Life technologies™) were used as host for transformation to generate the expression constructs. The cells were plated on Luria-Bertani (LB) agar (Table A1) plates supplemented with 30 $\mu g.mL^{-1}$ kanamycin (CYP153A6) or 100 $\mu g.mL^{-1}$ ampicillin (GDH and FdR/Fdx) and incubated at 37°C for 12-16 h. All the expression constructs, that is for, CYP153A6 operon, GDH, and

FdR/Fdx partial operon were used to transform competent *E. coli* BL21-Gold (DE3) cells which was used as host for expression.

2.1.2 Protein Expression

Pre-cultures were prepared by inoculating approximately 20 µL of transformed *E.coli* BL21-Gold (DE3) cells into 5 mL of LB media supplemented with 30 $\mu g.mL^{-1}$ kanamycin (CYP153A6), 100 $\mu g.mL^{-1}$ ampicillin (GDH and FdR/Fdx) and cultivated at 37°C for 12-16 h on a rotary shaker (160 rpm). For expression, 2 mL of pre-culture was used to inoculate 100 mL of ZYP5052 auto-induction media (Table A1), also containing the appropriate antibiotics, 1 mM δ-aminolevulinic acid hydrochloride (as heme precursor for P450 expression) and 50 µM $FeCl_3 \cdot 6H_2O$ (to compensate for the lack of iron in *E. coli* when expressing CYP and/or Fdx) in 500 mL Erlenmeyer flasks. Protein expression was carried out at 20°C for 48 h on an orbital shaker (200 rpm).

Table 2.1 Bacterial strains and plasmids used in this study

Strain/plasmid	Description	Manufacturer
Echerishia coli **Top 10**	One shot® TOP10 chemically competent cells F⁻ *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Life Technologies™
Echerishia coli **BL21-Gold(DE3)**	*E. coli* B F⁻ *ompT hsdS*($r_B^- m_B^-$) *dcm*⁺ Tet^r *gal* λ(DE3) *endA* Hte	Agilent technologies
pET-28b(+) (5368bp)	Kan^R, T7*lac* promotor, T7 terminator, f1 ori, N-terminal His-tag and thrombin configuration	Novagen®
pETDuet-1 (5420bp)	Amp^R, T7/lac promotor-1, T7/lac promotor-2, T7 terminator, His-tag, S-tag and f1 ori	Novagen®

2.1.3 Cell harvesting and cell disruption

After 48 h of expression, cells were harvested by centrifugation (7000 x *g* for 10 min at 4°C) using an Eppendorf centrifuge 5430 R. Cells were resuspended in a ratio of one gram of cells (wet weight) to 2 mL of 200 mM Tris-HCl buffer (pH 8) containing either 100 mM glycerol and 100 mM Glucose or 100 mM glycerol only. Therefore, the initial biomass concentration was 0.500 g_{WCW} mL^{-1}.

The cells were lysed using a One Shot Cell disrupter System (Constant Systems Ltd, United Kingdom). This high-pressure homogeniser forces the cell suspension through a narrow orifice and into a collecting beaker (Chisti & Moo-Young, 1986). Cell free extracts (CFE) were obtained by a single passage through the One Shot Disruptor at a pressure of 30 kilo pounds per square inch (kPSi) and the cell debris was removed from the soluble fraction by centrifugation (12000 x *g* for 30 min at 4°C).

2.1.4 Analysis of Expression

2.1.4.1 Spectroscopic enzyme quantification

CYP153A6 was quantified using Carbon monoxide-difference (CO-difference) spectra (Omura & Sato, 1964) recorded in F8 Maxisorp Nunc-ImmunoTM Modules (Nunc A/s, Denmark). Spectra were performed using both whole cells and cell free extracts, in which 200 µL of cell suspension was added to wells of two different microtitre strips. The samples were reduced by adding a few grains of sodium dithionate and one well of each sample was saturated with carbon monoxide. The absorbance spectrum was recorded between 350 nm and 500 nm using Spectramax M2 Microtiter Plate Reader (Molecular Devices Corporation, USA). The P450 concentration (c) was calculated using the Beer-Lambert Law ($A = \varepsilon c l$), where the $A_{450nm-490nm}$ (A) difference was corrected for a pathlength (l) of 0.5 cm and an extinction coefficient (ε) of 91 $mM^{-1}cm^{-1}$.

2.1.4.2 SDS-PAGE Analysis

Protein expression levels in cell free extracts were visually analysed with sodium dodecyl polyacrylamide electrophoresis (SDS-PAGE) using the Mini-PROTEAN® 3 system (Bio-Rad). The gel used for protein separation was comprised of an 8% resolving gel and a 4% stacking gel, according to the protocol described by Laemmli (Laemmli, 1970). Gels were run at 100 V for approximately 1 h using TGS buffer [Tris (25 mM), glycine (192 mM), SDS (0.1% w/v), pH 8.3] (Bio-Rad). PageRuler™ Unstained Protein Ladder (Thermo Scientific) or PageRuler™ Prestained Protein Ladder (Thermo Scientific) were used as molecular weight markers (Figure 2.1). The staining and de-staining protocol used was that described by Fairbanks and co-workers in which Comassie Brilliant Blue R-250 was used to visualise the proteins on the gel (Fairbanks *et al.*, 1971).

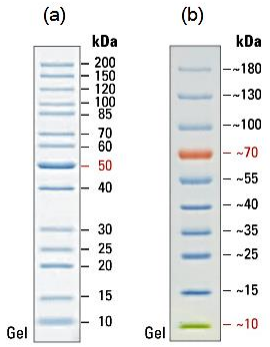

Figure 2.1 Protein molecular weight markers. (a) PageRuler™ Unstained Protein Ladder (Thermo Scientific) and (b) PageRuler™ Prestained Protein Ladder (Thermo Scientific)

2.1.5 Biotransformations of *n*-octane using cell free extract

In this part of the study, biotransformations were performed under various conditions to evaluate the stability and activity of the wild type CYP153A6 in order to determine which of these conditions would be optimal for evaluating the CYP153A6 mutants (Chapter 3).

Cell free extract biotransformations were performed using a 1.2 mL reaction volume – of which 200 µL was used for CO-difference spectra - in 40 mL amber vials with non-permeable screw caps. The stability of CYP153A6 was monitored by measuring the concentration of P450 at different times using CO-difference spectra (Section 2.1.4.1). Although the initial biomass concentration from which CFEs were prepared was 500 g_{WCW} L^{-1}, the soluble protein in the different experiments and in some cases different biotransformation reactions represented different biomass concentrations. Reactions were started by adding 250 µL (1.2 mol L_{BRM}^{-1}) of *n*-octane to the BRM before the vials were placed in an orbital shaker (200 rpm).

2.1.6 Sample extraction and product analysis

Before sample extractions, CO-difference was performed using 200 µL of biotransformation reaction mixture (BRM) from each vial (sections 2.1.4.1 and 2.1.5). Five hundred microliters of ethyl acetate (EtAc) containing 2 mM 2-decanol as internal standard was added to the remaining 1 mL of BRM to stop the reaction. The mixture was transferred to a 2 mL Eppendorf microtube and an additional 500 µL of ethyl acetate (EtAc) containing internal standard, was used to rinse any remaining product from the vial and this was also transferred to the 2mL microtube. The mixture was vortexed for 5 min and centrifuged at 10

000 x *g* for 5 min. Two hundred microliters of the upper organic layer was transferred to GC vials for analysis.

A Shimadzu GC-2010 gas chromatograph, equipped with a flame ionisation detector (FID) and a FactorFour™ VF-5ms column (60 m x 0.32 mm x 0.25 µm) was used to analyse 1 µL samples. The flow of hydrogen (carrier gas) through the column was 30 mL min^{-1} at a pressure of 45.4 kPa. The temperature program used was as follows: the initial oven temperature was 60°C held for 5 min. The temperature was increased at a rate of 25°C min^{-1} to 180°C held for 3.2 min and the inlet temperature was 200°C.

Table 2.2 Retention times of internal standard, substrate and product separated from cell free extract biotransformations using GC analysis

Compound name	Type	Retention time (min)
***n*-octane**	Substrate	7.10
1-octanol	Product (target)	10.94
2-decanol	Internal standard	12.28

Section B - Different experiments to investigate effect of reaction conditions (continuation of section 2.1.5)

Reaction parameters that were evaluated for their effect on activity and stability of the CYP153A6 system included enzyme concentration, addition of NADH, concentration of added NADH, substrate concentration, product toxicity, type of buffer used, buffer concentration, concentration of glucose in the buffer, pH of the buffer and reaction temperature. The oxidative stability of CYP153A6 was evaluated using hydrogen peroxide (H_2O_2) produced *in situ* by glucose oxidase (Sigma) added to the biotransformation reaction mixture (BRM). All these conditions will be discussed in more detail in the following subsections.

2.1.5.1 Experiments using glucose dehydrogenase for cofactor regeneration

Glucose dehydrogenase from *Bacillus megaterium* (GDH) catalyses the oxidation of *D*-glucose to *D*-gluconolactone using NAD^+ or $NADP^+$ as cofactor (Figure 2.2). In this study, GDH expressed in *E. coli* BL21-Gold (DE3) (section 2.1.1) was evaluated as a possible co factor regenerating system for the CYP153A6 using a variety of conditions in two separate experiments.

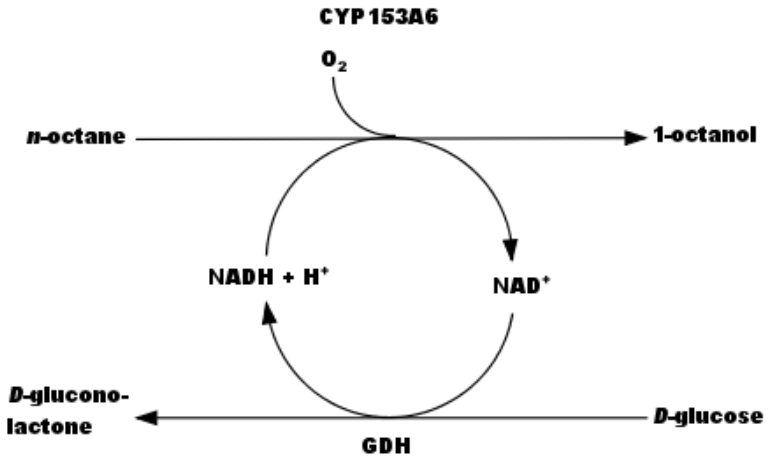

Figure 2.2 Glucose dehydrogenase regenerating reduced cofactor for CYP153A6.

All reactions contained 200 mM Tris HCl (pH 8) buffer containing 100 mM glycerol and 100 mM glucose (section 2.1.3). Each reaction was carried out in duplicate for 1, 2, 4 and 8 h and incubated at 20°C on a shaker (200 rpm). Parameters that were varied were P450, GDH, octane and added NADH concentration as well as the effect of product toxicity (octanol added to BRM) (Experiment 1) and of replacing GDH containing CFE with CFE from a strain transformed with empty vector (Experiment 2).

In the first experiment, (Table A2) the effects of P450 and GDH concentration were evaluated by adding either 200 or 400 µL of enzyme containing CFE to the BRM. NADH concentration was also evaluated by adding either 10 or 20 µL of a 120 mM NADH stock solution so that the final concentration of NADH in the BRM would be either 1 or 2 mM. In reactions where substrate concentration was evaluated, either 25 µL (0.12mol L_{BRM}^{-1}) or 250 µL (1.2mol L_{BRM}^{-1}) of *n*-octane was added to the BRM. Additionally product toxicity was also evaluated by dissolving 4.8 µL of 1-octanol in 1 mL of *n*-octane. This solution (250 µL) was then added to the BRM so that the final initial concentration of 1-octanol in the BRM was 7.7 mmol L_{BRM}^{-1}.

In the second experiment the effect of replacing GDH with buffer or CFE from a strain containing an empty pET-28b (+) vector was evaluated in reactions containing 400 µL P450 containing CFE and 1 mM NADH (Table A3).

2.1.5.2 BRM prepared with additional Ferredoxin Reductase and Ferredoxin

CYP153A6, FdR and Fdx work as a three component system with electrons transferred from NADH to CYP153A6 via FdR and Fdx (Figure 2.3). Although the strain containing the CYP153A6 operon also expresses FdR and Fdx so that P450 containing CFEs also always contain FdR and Fdx, in this experiment,

the effects of additional FdR/Fdx were evaluated. The FdR and Fdx from *Mycobacterium sp*. were expressed from a partial operon cloned into pETDuet and transformed into *E. coli* BL21-Gold (DE3).

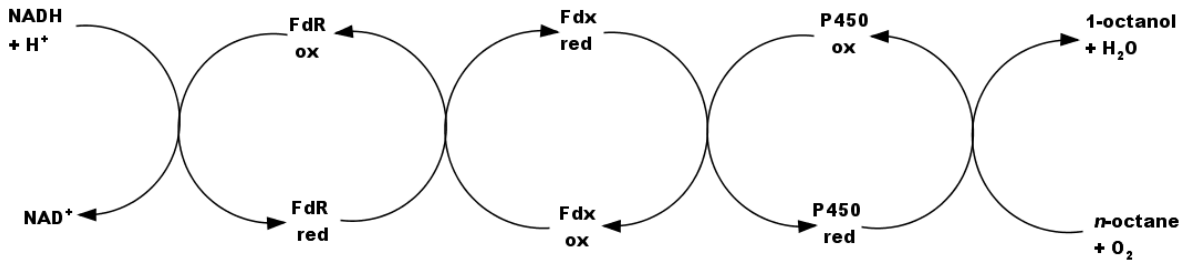

Figure 2.3 Electron transfer via Ferredoxin reductase and Ferredoxin to cytochrome P450

All reactions contained 200mM Tris-HCl (pH 8) buffer containing 100mM glycerol and 250 µL of P450 containing CFE (also containing FdR and Fdx), while different volumes (125 or 250 µL) of CFE containing only FdR and Fdx were added. The latter was replaced or supplemented with CFE from an empty vector containing strain (Table A4). The role of glucose was also evaluated in this experiment with one set of reactions containing no glucose. Each reaction was carried out in duplicate for 1, 2, 4, 8 and 18 h and incubated at 20°C on a shaker (200 rpm)

2.1.5.3 Controlling P450 concentration, storing P450 overnight and reactions carried out in the presence of in situ generated H_2O_2

All reactions contained 200 mM Tris-HCl buffer (pH 8) containing 100 mM glycerol and 100 mM glucose. The final concentration of P450 in the BRM was controlled by calculating the initial P450 concentration of CYP153A6 using CO-difference spectra, and then diluting the P450 containing CFE using CFE from an empty vector strain to a concentration that would give the desired final concentration in the BRM. To elaborate, the initial P450 concentration in this experiment was 23 µM, the desired final concentration in the BRM was 4 µM therefore the initial concentration was diluted to 19 µM. Reactions were carried out in triplicate for 2, 4, 8 and 18 h.

In all the previous experiments biotransformations were performed on the same day that cells were harvested, disrupted and the CFEs containing the different proteins prepared. In other words, the enzyme was “fresh”. In this experiment, the effect of storing the enzyme overnight on ice in a 4°C fridge was evaluated - therefore the enzyme was “one day old”. Additional NADH was added to these reactions in the absence of glucose oxidase and reactions were carried out for 2 and 18 h at 20°C.

Glucose oxidase (GOx) isolated from *Aspergillus niger sp* (Sigma-Aldrich) catalyses the oxidation of *D*-glucose to *D*-gluconolactone by utilising oxygen as an electron acceptor, simultaneously producing hydrogen peroxide (H_2O_2). In the presence of water, *D*-gluconolactone is converted to *D*-gluconic acid (Figure 2.4). The effect of H_2O_2 on the stability and activity of CYP153A6 was evaluated using two different concentrations of glucose oxidase (GOx) which was added to reactions by dissolving the enzyme powder in buffer to obtain a final concentration of either 400 or 600 U/L in the BRM (Table A5). Thermal stability was evaluated simultaneously by incubating biotransformation reactions at either 20°C or 30°C.

Figure 2.4 Catalytic reaction of glucose oxidase. *D*-glucose is oxidised to *D*-gluconic acid in the presence of water (H_2O), utilising oxygen (O_2) as an electron acceptor, simultaneously producing hydrogen peroxide (H_2O_2).

2.1.5.4 Biotransformation reaction buffer

The buffer added to biotransformations in all the previous experiments was a 200 mM Tris-HCl buffer (pH 8) which contained 100 mM glycerol and 100 mM glucose or no glucose. In the following two experiments, different buffers were evaluated at different concentrations and different pH to try to optimise cell free extract biotransformations of *n*-octane using CYP153A6. These conditions were evaluated in two separate experiments which were carried out simultaneously.

In the first experiment the effect of two Tris buffer concentrations (200 and 400 mM) were evaluated at different pH, 8 and 8.5. Therefore, three buffers were prepared: 200 mM Tris (pH 8), 200 mM Tris (pH 8.5) and 400 mM Tris (pH 8.) all containing 100 mM glycerol and no glucose. The cell biomass was harvested in three separate 50 mL Falcon tubes and resuspended in each of the buffers. Glucose (final concentration 100 mM) was only added once the BRM was prepared. These buffers were evaluated using the same conditions at 20 and 30°C (Table A6). Again, the initial P450 concentration (30 µM) was diluted with empty vector CFE to obtain a desired final concentration of 3 µM in the BRM and a total protein content as from a cell suspension with biomass concentration of 0.208 g_{WCW} mL_{BRM}^{-1}. Each reaction was carried out in triplicate for 2, 8 and 18 h.

In the second experiment, the effect of two different buffers was evaluated. The first buffer was 300 mM Tris (pH 8) containing 100 mM glycerol and the second buffer was a mixture of MES, MOPS and Tris buffers (each 300 mM) containing 100 mM glycerol which was prepared at pH 6, 6.5, 7, 7.5, 8 and 8.5. Harvested

cells were resuspended in 20 mM Tris buffer (pH 8) containing 100 mM glycerol as opposed to the usual 200 mM Tris buffer to ensure that the final buffer concentration (300 mM) would not be affect once cell suspensions were added to the BRM and to ensure that the desired pH for each reaction containing it's respective buffer would not change. Therefore, initial pH readings were taken to ensure that the desired pH in the BRM did not change. A final P450 concentration of 3µM in the BRM was targeted and again a total protein content as from a cell suspension with final biomass concentration of 0.208 g_{WCW} mL_{BRM}^{-1}. Each reaction was carried out in triplicate for only 2 h at 20°C (Table A7).

2.1.5.5 Concentration of P450 in biotransformation reaction mixture (BRM)

The concentration of P450 used in the BRM was evaluated more thoroughly in this experiment to try to optimise biotransformation of *n*-octane to 1-octanol by CYP153A6. Concentration was evaluated to determine if CYP153A6 would be more stable at higher or lower concentrations. The concentration was evaluated in terms of volume of P450 containing CFE added to the BRM therefore biotransformation reaction mixtures were prepared by adding either 50, 100 or 250 µL of P450 containing CFE. This meant that the final P450 biomass concentration in the BRM would be different for each of these volumes. After calculating the initial concentration of CYP153A6 using CO-difference spectra, P450 was diluted with empty vector CFE so that reactions containing 50 µL of P450 would have a final concentration of 1 µM, reactions containing 100 µL of P450 would have a final concentration of 2 µM and reactions containing 250 µL of P450 would have a final concentration of 4 µM in the BRM.

Cells were resuspended in 200 mM Tris-HCl (pH 8) buffer containing 100 mM glycerol. The final concentration for the buffer used in the BRM had to be 300 mM Tris-HCl (pH8) containing 100 mM glycerol and 100 mM glucose therefore three buffers were prepared: 333 mM Tris, 341 mM Tris and 371 mM Tris, each containing 100 mM glycerol and 100 mM glucose which was added to reactions containing 50, 100 and 250 µL of P450, respectively to obtain a final concentration of 300 mM Tris. Furthermore, all reactions received an additional 1 mM NADH and were carried out for 2, 8 and 20 h at 20 °C and 30°C (TableA8).

2.2 Results

2.2.1 Analysis of expressed proteins

2.2.1.1 Spectroscopic quantification of P450 content

The P450 content of the CYP153A6 was quantified by performing CO-difference spectra using both whole cells and cell free extract. CO-difference spectra is a useful application to determine if cyctochrome P450 proteins have folded properly and are therefore catalytically active (soret band at 450nm) or if the protein has folded incorrectly and is therefore catalytically inactive (soret band at 420nm). If both functional and non-functional forms of the protein are present, soret bands are observed at 420 nm and 450 nm (Figure 2.5 (a))(Omura & Sato, 1964). A soret band at 450 nm was obtained for CYP153A6 in both whole cells and cell free extracts (soluble fraction) (Figure 2.5 (b)). When comparing concentrations of the CYP153A6 in whole cells and cell free extracts it was clear that a significant percentage of P450 was lost after disrupting the cells and retrieving the soluble protein by centrifugation. This could be due to the fact that the cell suspension of 0.5 g_{wcw} mL^{-1} was too concentrated and would have required more than one passage through the cell disruptor to ensure that all the cells were broken in order to maximise the recovery of the soluble CYP153A6 protein. Futhermore, the expression levels of CYP153A6 varied between experiments in whole cells which also affected expression levels in cell free extracts but the reason for this difference in protein expression is not clear (Figure 2.5 (c) and (d)). The initial P450 concentrations calculated for CYP153A6 in CFE varied between 13 and 19 µM for the experiments using the different cofactor regenerating enzymes (Figure 2.5 (c)). In the last experiments the intial P450 concentration of CYP153A6 was calculated the day *E. coli* cells expressing the protein were haravested and dirupted to obtain the CFE (day 1). Thereafter, CYP153A6 was stored overnight and the P450 concentration was recalculated the next day (day 2). Needless to say, the initial P450 concentration of CYP153A6 decreased overnight (Figure 2.5 (d)).

2.2.1.2 SDS-PAGE analysis

Expression of proteins was analysed on 8% SDS-PAGE resolving gels using the crude cell lysates and cell free extracts (soluble fraction).

A band corresponding the to the expected molecular weight of 48 kDa for CYP153A6, which consists of 439 amino acids, was observed in the crude cell lysates (Figure 2.6, lane 1) and cell free extracts (Figure 2.8, lane 3). No band corresponding to the molecular weight of 45kDa for Ferredoxin reductase (424 amino acids) could not be observed by SDS-PAGE analysis in crude cell lysates or cell free extracts. In addition,

Ferredoxin (11 kDa) could not visualised on the gel because based on the protein ladder this would not be possible.

The expression of the proteins used in experiments for cofactor regeneration and electron transfer are illustrated in Figure 2.7. Bands corresponding to the expected molecular weights were observed for Glucose dehydrogenase (Lane 1) which consists of 261 amino acids and has an expected molecular weight of 30 kDa but in crude cell lysates (a) and cell free extracts (b) the molecular weight was observed as 36 kDa, Nagao and co-workers demostrated different migration patterns for glucose dehydrogenase which could be affected by the composition of the gel (Laemmli, 1970; Nagao *et al.*, 1992). Lane 2 contained cell lysates for ferredoxin reductase and ferredoxin which were co-expressed as a partial operon but no bands corresponding to the molecular weights of either of these proteins were observed by SDS-PAGE analysis. We suspect that ferredoxin and ferredoxin reductase express in fairly low amounts because other researchers in our research group did not visualise either ferredoxin or ferredoxin reductase when performing SDS-PAGE analysis Lanes 3 contained an empty pETDuet-1 plasmid which was used as a negative control.

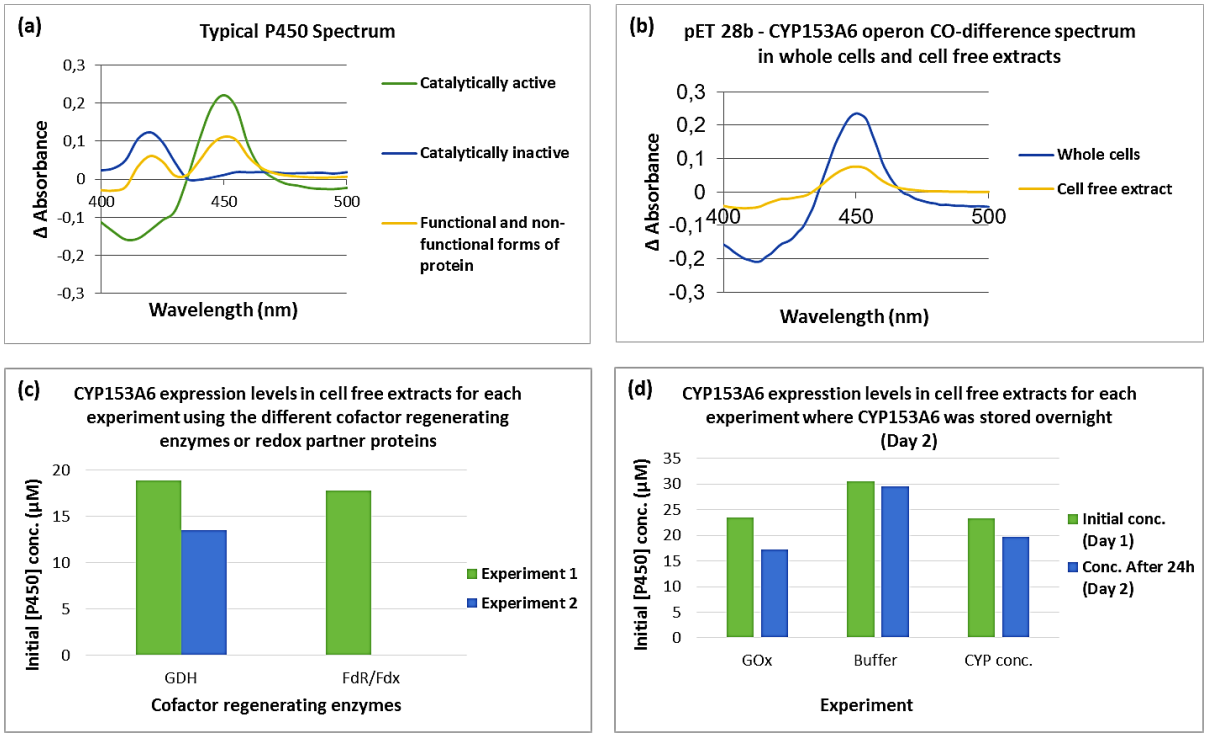

Figure 2.5 CO-difference spectra demonstrating a typical CO-difference spectrum (a). CO-difference spectra obtained using whole cells and cell free extracts of CYP153A6, illustrating the poor recovery of CYP153A6 after *E. coli* cells have been disputed to obtain the cell extract (b).

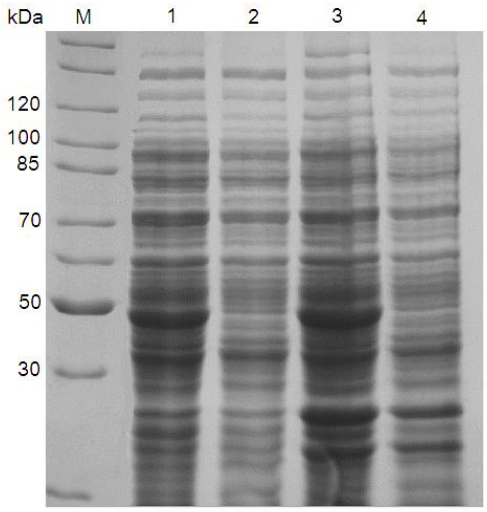

Figure 2.6 SDS-PAGE analysis of the crude cell lysates and cell free extracts of *E. coli* BL21-Gold (DE3) strains, transformed with pET-28b (+) plasmids for the expression of CYP153A6 operon (48 kDa) and an empty pET-28b(+) plasmid. Lane 1, crude cell lysate of CYP153A6; lane 2, empty pET-28b (+) plasmid; lane 3, cell free extract of CYP153A6 operon and lane 4, empty pET-28b (+) plasmid. M = PageRuler™ Unstained Protein Ladder (Thermo Scientific).

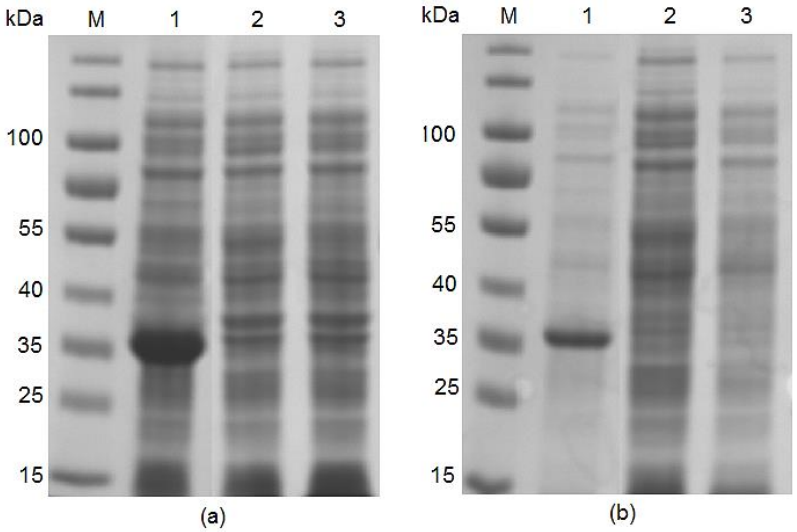

Figure 2.7 SDS-PAGE analysis of the crude cell lysates (a) and cell free extracts (b). Lane 1, GDH (36 kDa); lane 2, FdR (45 kDa) and Fdx (11 kDa) partial operon and lane 3, empty pETDuet-1 plasmid (control). M = PageRuler™ Prestained Protein Ladder (Thermo Scientific).

2.2.2 *n*-Octane biotransformations using cell free extracts containing CYP153A6

CYP153A6 is a soluble protein that functions as a three component system with ferredoxin reductase and ferredoxin to transfer electrons from NADH to the heme iron of the P450 (Funhoff *et al*., 2006; Hannemann *et al*., 2007). Generally, this type of electron transfer system is assumed to not be as efficient as P450s with electron transfer proteins fused to their heme domain, therefore this characteristic of the CYP153A6 is considered to be one of its limitations. Other limations of the CYP153A6 is that it is NADH dependant and it is highly unstable.

The instabilty of the CYP153A6 in whole cells has been reported by a number of research groups including ours (Gudiminchi *et al*., 2012; Olaofe *et al*., 2013). It was for this reason that cell free extracts from *E. coli* BL21-Gold (DE3) expressing the CYP153A6 was selected for the hydroxylation of *n*-octane to investigate the effect on enzyme stability.

Instabilty of P450 based systems can be the result of a number of factors such as cofactor stability, temperature, amount of substrate used, cofactor regeneration system used, product toxicity and pH to name a few. Therefore, many factors that were thought to effect the operational stabilty of the CYP153A6 were evaluated using various biotransformations under different conditions.

2.2.2.1 Biotransformations containing glucose dehydrogenase as a cofactor regenerating system for CYP153A6

Glucose dehydrogenase was used for cofactor regeneration for CYP153A6 in two experiments in which the effects of P450, GDH, NADH and substrate concentration were investigated as well as adding the product (1-octanol) at the beginning of the reaction. The effects on the enzyme's operational stability and activity were monitored.

Enzyme concentration plays an important role in biocatalyst stability. The protein content in the reactions differed depending on the volumes of cell free extracts added so that a total of 400 µL CFE corresponded to a total biomass concentration of 0.166 g_{WCW} mL_{BRM}^{-1}, 600 µL to 0.249 g_{WCW} mL_{BRM}^{-1} and 800 µL to 0.332 g_{WCW} mL_{BRM}^{-1}. Figure 2.8, illustrates the effect of increasing the concentration of P450, GDH and NADH by doubling the volume in the BRM. Reaction 1 was the control in this experiment which contained 200 µL (~5 µM) of P450, 200 µL of GDH, 1 mM of NADH and 250 µL of substrate (Table A2). The activity of reaction 1 levels off after 2 h and the remaining P450 is less than 1 µM after 8 h. When twice the volume of GDH was present in the BRM (reaction 4) and a lower concentration of P450 was used, the activity of CYP153A6

decreased however when the BRM contained a higher concentration of P450 (~ 15 µM) and a lower concentration of GDH (reaction 5), not only did the initial activity of CYP153A6 increase, but the reaction continued for 8 h. Increasing the concentration of additional NADH to 2mM (20 µL) in the BRM did not significantly improve activity (reaction 3) however, when twice the volume of P450, GDH and NADH was present in the BRM (reaction 2), the activity of CYP153A6 increased but not higher than that of reaction 5 this could be due to the higher concentration of GDH present in the reaction. Furthermore, none of these conditions significantly improvement the stability of CYP153A6 in terms of remaining P450.

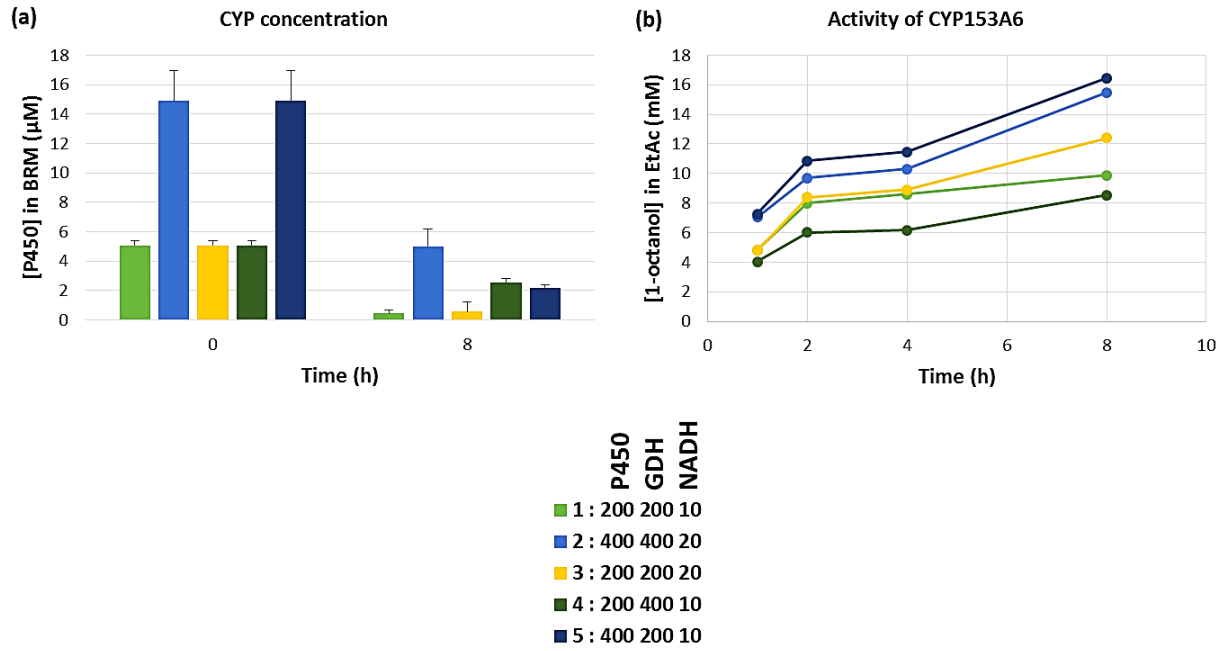

Figure 2.8 Reactions illustrating the effect on CYP153A6 stability (a) and activity (b) when increasing the concentration of P450 (reaction 5), GDH (reaction 4), NADH (reaction 3) or all three (reaction 2) in the BRM. Reaction 1 (control).

Figure 2.9, illustrates the effect of different substrate concentrations on the stability (a) and activity (b) of CYP153A6. The BRM of both reactions (5 and 7) were the same but reaction 5 contained 250 µL of substrate and reaction 7 contained only 25 µL of substrate in a total reaction volume of 1.2 mL. It was clear that lower substrate concentration resulted in lower enzyme activity, but interestingly the lower activity resulted in less strain on the enzyme so that more P450 remained after 8 h. (Figure 2.9 (a)). Pennec and co-workers also observed this kind of results when evaluating the effect of substrate concentration on enzyme activity, they tested several enzymes including CYP153A6 in CFE and all enzymes had better activity with 250 µL of substrate (Pennec *et al*., 2014).

Product toxicity was also evaluated in this experiment by adding 1-octanol (7.7 mmol L_{BRM}^{-1}) to the initial BRM of reaction 8 (Figure 2.10). The P450 concentration for reaction 8 decreased more rapidly than reaction 5 which contained no additional 1-octanol. Although it was clear that the product produced by CYP153A6 had a negative effect on its stability, further investigation would be required to evaluate the full extent to which product toxicity plays a role in CYP153A6 instability, however this was not further investigated in this study.

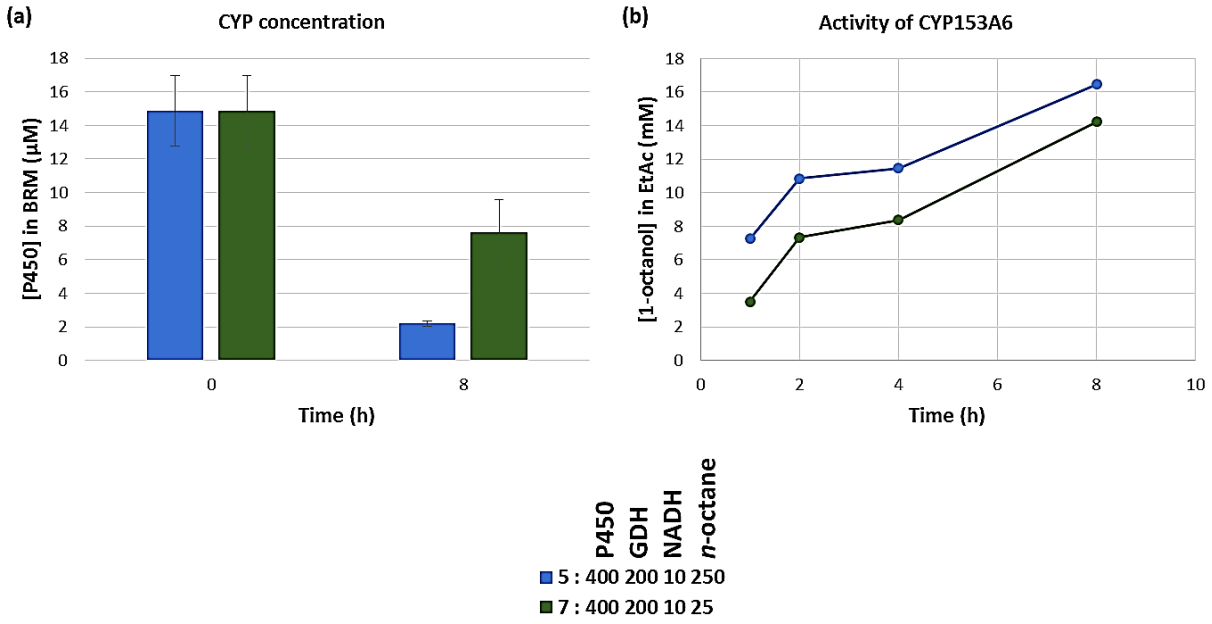

Figure 2.9 Reactions illustrating the effect of substrate concentration on CYP153A6 stability and activity. Reaction 5 contains 250 µL of *n*-octane and reaction 7 contains 25 µL of *n*-octane.

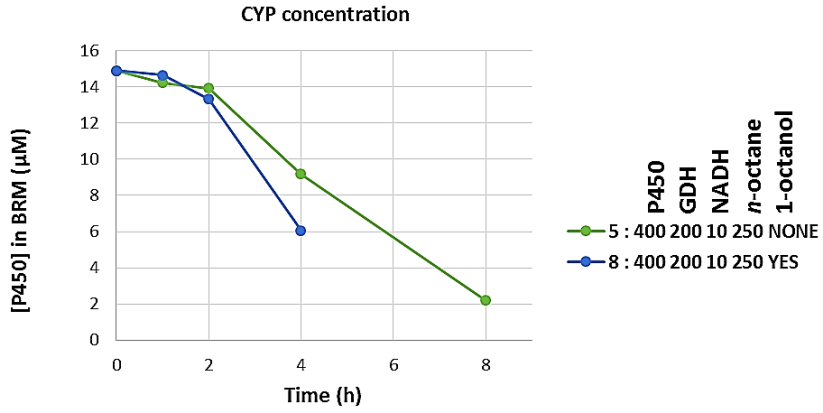

Figure 2.10 Reactions illustrating the effect of additional 1-octanol added on CYP153A6 stability to investigate the possibility of product toxicity. Reaction 8 received additional 1-octanol its initial biotransformation reaction mixture.

The experimental design of the next experiment was based on the results of the previous experiment. Here, all reactions contained 250 µL of substrate and 400 µL of P450[2] and 200 µL or 400 µL of GDH containing cell free extract or cell free extract from an empty vector control strain (corresponding to an initial total biomass concentration in the reactions of 0.249 or 0.332 g_{WCW} mL_{BRM}^{-1}). Thus, some reactions were carried out in the absence of GDH (Table A3).

Figure 2.11 illustrates the results of the conditions tested. Reaction 1 which contained 400 µL (~ 9.5µM) of P450, 200 µL of GDH containing CFE and 1 mM NADH in the BRM was the control in this experiment. Reaction 2 contained double the volume of GDH than reaction 1 which decreased the activity of CYP153A6. The BRM of reactions 3 was the same as reaction 1 but 200 µL of empty vector CFE had also been added to the BRM and although this did not improve the activity of CYP153A6, the empty vector CFE had a stabilising effect on the enzyme in terms of remaining P450 after 8 h (Figure 2.11 (a)). In reaction 4, GDH was excluded from the BRM and replaced with 200 µL of empty vector CFE, this resulted in a significant increase in enzyme activity but more importantly the enzyme was also more stable. When the volume of empty vector CFE extract was doubled in reaction 5 the CYP153A6 also remained relatively stable, but there was a decrease in enzyme activity.

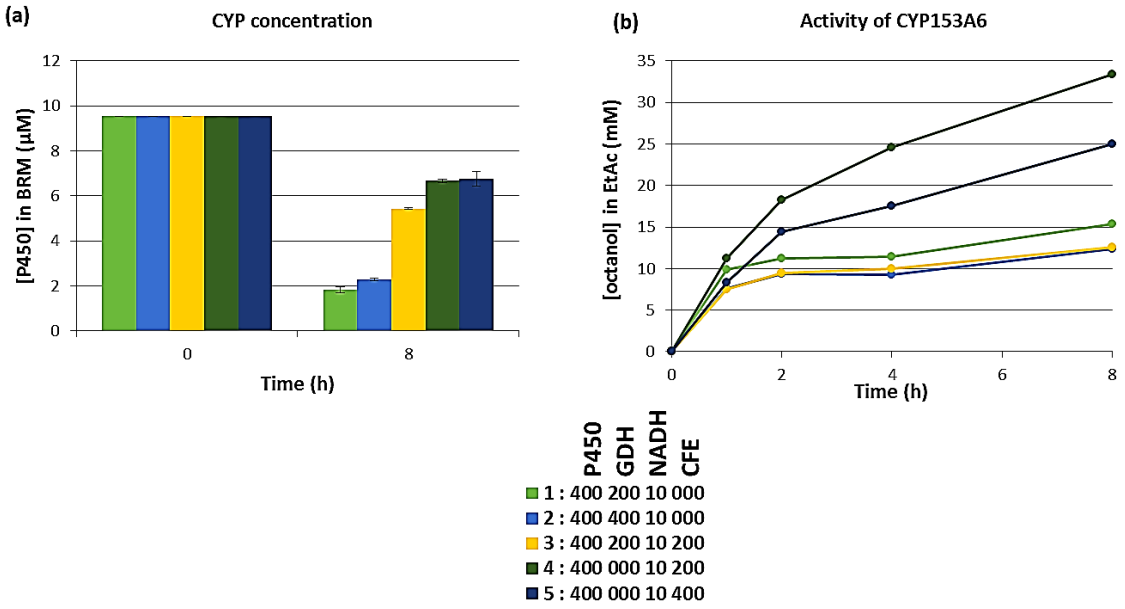

Figure 2.11 Reactions illustrating the negative effect of GDH on CYP153A6. Reaction 1 (control); reaction 2 contained double the volume of GDH than reaction 1. Reaction 3 contained GDH and empty vector cell free extract. Reaction

[2] CYP expression levels in experiment 1 differed from the CYP expression levels in experiment 2. Refer to section 2.2.1.1.

4 and 5 contained no GDH but reaction 5 contained double the volume of empty vector cell free extract as reaction 3. All reactions received additional NADH.

Glucose dehydrogenase oxidises *D*-glucose to *D*-gluconolactone (Figure 2.2), which is converted to gluconic acid in the presence of water. It's possible that the surrounding environment in the BRM became acidic which could be having a negative effect on the activity and stability of CYP153A6. Pham *et al*., 2013 suggested that this could explain the decrease in CYP153A6 activity in the presence of GDH. Furthermore, the activity appeared to be better when higher biomass concentrations of P450 were used. In addition, adding empty vector CFE appeared to have positive effect on the stability of CYP153A6 however increasing this concentration resulted in a decrease in activity therefore the effect of adding empty vector CFE to reactions was also evaluated in subsequent experiments.

2.2.2.2 Biotransformations containing additional FdR/Fdx to aid electron transfer from NADH to CYP153A6

Because the previous experiments had indicated that the addition of cofactor regenerating enzymes did not benefit activity or stability, it was decided to leave such enzymes out in future experiments so that we only relied on the cofactor regenerating enzymes present in the crude CFEs. It was also decided to keep the total CFE constant at a total volume of 500 µL representing a total biomass concentration of 0.208 g_{WCW} mL_{BRM}^{-1}.

In this experiment the volume of P450 containing CFE (which also always contain FdR and Fdx) was reduced to 250 µL while the effect of adding CFE from a strain expressing only FdR and Fdx from a partial operon was investigated. The effect of additional FdR/Fdx on CYP153A6 stability (Figure 2.12 (a)) and activity (Figure 2.12 (b)) were evaluated under different conditions for an extended period of 18 h (Table A4). The role of glucose on CYP153A6 stability and activity was also evaluated in this experiment with no NADH added.

In Figure 2.12, reaction 1 was the control and contained 250 µL (~ 4.2 µM) of P450, no additional FdR/Fdx, 250 µL of empty vector CFE and glucose (100 mM). Although the activity levelled off after 4 h, CYP153A6 remained relatively stable because the P450 content only decreased by 25% after 18 h. In reaction 2, additional FdR/Fdx was added to the reaction which increased enzyme activity but the remaining P450 content decreased. In reaction 3, more FdR/Fdx was added than reaction 2 and this increased the enzyme activity even more, producing up to 50 mM 1-octanol while the reaction was still not levelling off. However, when comparing the enzyme activity to the enzyme concentration, it was clear that more active

the enzyme was, the more unstable it became. The BRM of the last reaction (reaction 4) was the same as reaction 2 but it contained no glucose, clearly demonstrating the negative effect that the absence of glucose had on enzyme activity and stability with an 83% decrease in the P450 concentration from the initial concentration (Figure 2.12 (a)).

Compared to GDH, additional FdR/Fdx added to biotransformations significantly improved enzyme activity. Also, no additional NADH was added to any of these reactions which indicated that CYP153A6 can efficiently take care of cofactor regeneration using glucose and glycerol as a source of electrons and that the NADH present in CFEs of *E. coli* was sufficient.

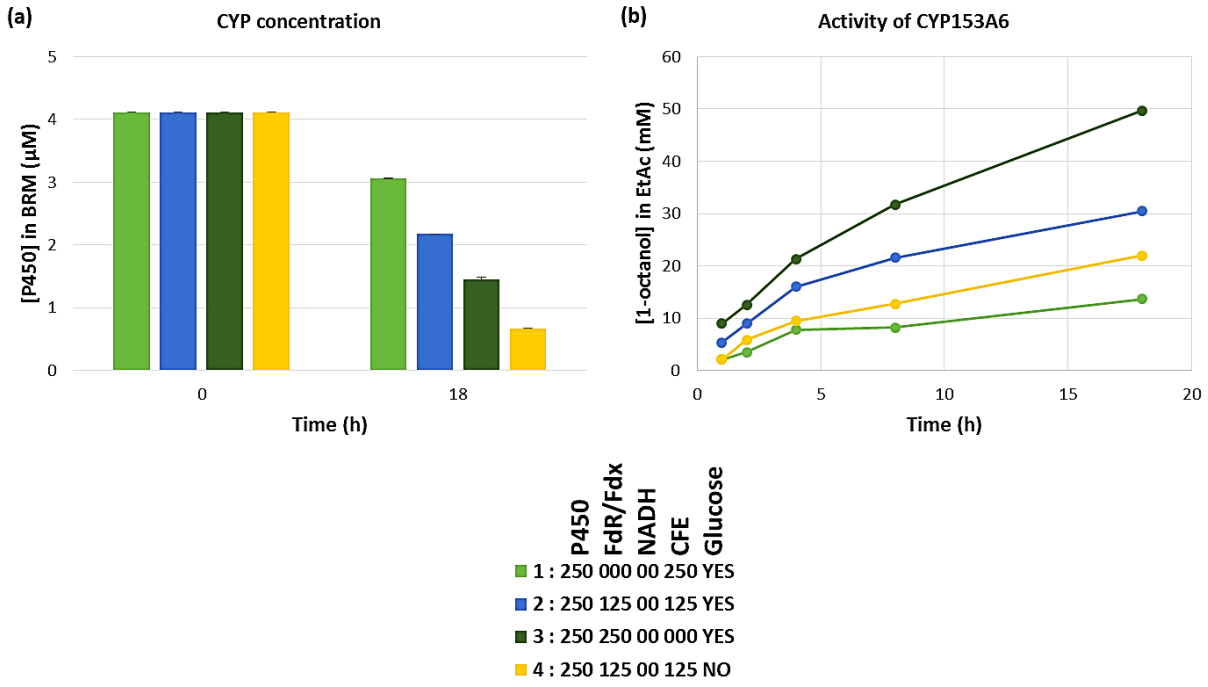

Figure 2.12 Reactions illustrating the effect of additional FdR/Fdx on CYP153A6 stability (a) and activity (b). Reaction 2 contained an additional 125 µL of FdR/Fdx; reaction 3 contained an additional 250 µL of FdR/Fdx and reaction 4 contained an additional 125 µL FdR/Fdx but it contained no glucose, while reactions 1, 2 and 3 contained 100 mM glucose. Reaction 1 (control).

2.2.2.3 Evaluating the effect of H_2O_2, temperature and storage of CFEs on CYP153A6 stability and activity

Under the conditions used in the previous experiment, the CYP153A6 system was surprisingly stable with 20 mM octanol still produced between 8 h and 18 h. In this experiment, the stability of the system was challenged through the in situ generation of hydrogen peroxide, by increasing the reaction temperature to 30°C and by storing the CFEs overnight.

Glucose oxidase (GOx) catalyses the oxidation of *D*-glucose to *D*-gluconolactone and simultaneously produces H_2O_2 (Figure 2.4). For this reason, GOx was added to reactions to evaluate the effect H_2O_2 had on CYP153A6 stability and activity. Two concentrations of GOx were evaluated, 400 and 600 U/L but reactions were also carried out in the absence of GOx in order to make a clear comparison of the effects. The initial P450 concentration calculated for CYP153A6 in this experiment - using CO-difference spectra - was 23 µM which was diluted with empty vector CFE to try and achieve a final concentration of 4 µM in the BRM, but this was not easily achieved because the final concentration obtained was 6 µM (Figure 2.13 (a)). The reason for wanting to obtain a final concentration of 4 µM in the BRM was based on the promising results obtained in the previous experiment where the final P450 concentration had been 4 µM (section 2.2.2.2, Figure 2.12). Furthermore, none of the reactions received additional NADH and biotransformations were carried out at 20 and 30°C for 2, 8 and 18 h (Table A5).

The control reaction containing only P450 and additional FdR/Fdx displayed lower activity than the equivalent reaction in the previous experiment, but octanol production still appeared linear after 18 h. (Reaction 3, Figure 2.13). The effect of GOx on CYP153A6 stability and activity when biotransformations were incubated at 20°C is evident. Reactions 1 and 2 which contained 400 and 600 U/L GOx respectively, showed a significant increase in enzyme activity compared to reaction 3 which contained no GOx, however the activity of reactions 1 and 2 began to decrease after 8 h whereas the activity of reaction 3 was still continuing after 18 h (Figure 2.13 (b)). Furthermore, no P450 was detected (*[3]) after 18 h for reactions that contained GOx, therefore it appeared as though the presence of H_2O_2 had a negative effect on CYP153A6 stability in terms of remaining P450 content (Figure 2.13 (a)).

The conditions of reactions illustrated in Figure 2.14 were the same as in Figure 2.13 but these reactions were incubated at 30°C instead of 20°C to evaluate not only the oxidative stability of CYP153A6 but its thermal stability as well. Reactions 4 and 5 contained 400 and 600 U/L GOx respectively, and although an increase in enzyme activity was again observed, it appeared as though reactions levelled off much quicker at higher temperature, even reaction 6 which contained no GOx began to level off. However, reactions 4 and 5 levelled off after 4 h and reaction 6 levelled off after 8 h. Furthermore, it was clear that higher temperature had a significant impact on the stability of CYP153A6 in terms of remaining P450 because no P450 was detected (*) for all three reactions after 18 h.

[3] An asterisk displayed on a graph indicates that no P450/CYP was detected at that particular time for that specific reaction.

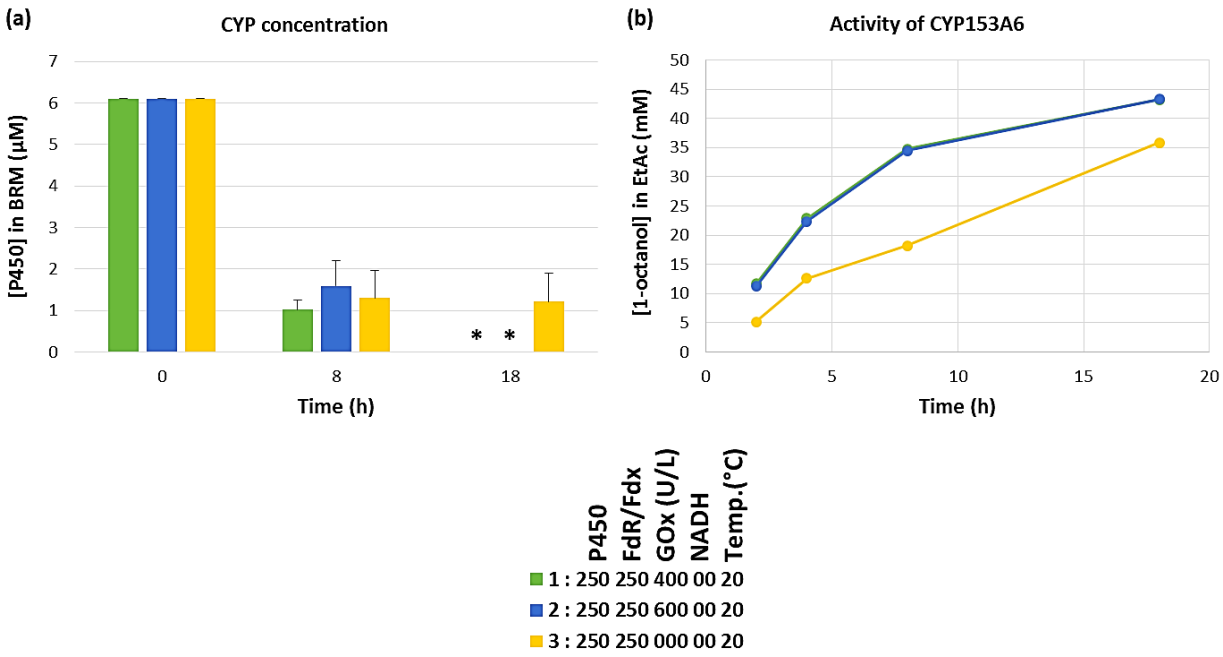

Figure 2.13 Reactions illustrating the effect of H_2O_2 produced by glucose oxidase on CYP153A6 stability (a) and activity (b). Reaction 1 contained 400 U/L GOx, reaction 2 contained 600 U/L GOx and reaction 3 contained no GOx.

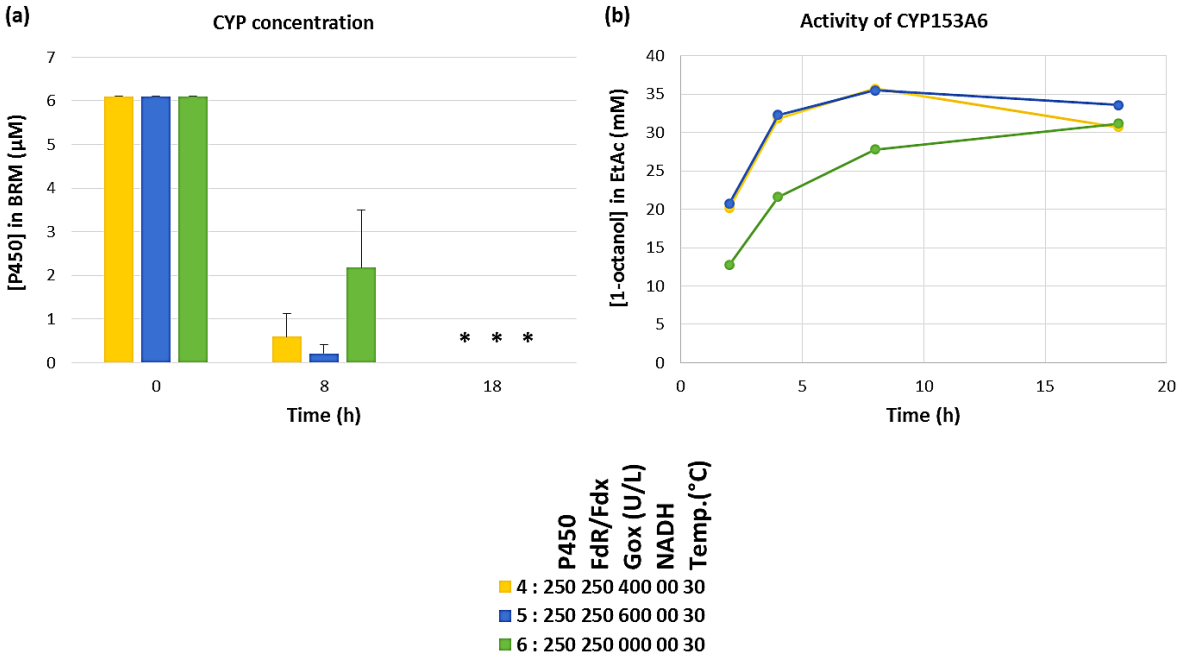

Figure 2.14 Reactions illustrating the effect of H_2O_2 produced by glucose oxidase on CYP153A6 stability (a) and activity (b) at a higher temperature (30 °C). Reaction 4 contained 400 U/L GOx, reaction 5 contained 600 U/L GOx and reaction 6 contained no GOx.

Reactions 1 to 6 were performed on the same day that *E. coli* cells expressing CYP153A6 and FdR/Fdx were harvested, disrupted and cell free extracts containing these soluble proteins were obtained therefore, these enzymes could be called "fresh" and these reactions were referred to as "Day 1" reactions. The initial P450 concentration for CYP153A6 in CFE on day 1 was 23 µM. Reactions 7 and 8 contained the same enzymes used in reactions 1-6 but enzymes were stored overnight, on ice, in a 4°C fridge and biotransformations were only started the following day therefore, the enzymes were "1 day old" and these reactions were referred to as "Day 2" reactions. After storing CYP153A6 overnight the initial P450 concentration decreased to 17 µM. Reactions 7 and 8 contained no GOx but because the enzyme was stored overnight, these reactions received additional NADH (Figure 2.15). When comparing the final P450 concentration in the BRM for reaction 3 (day 1) and reactions 7 and 8 (day 2), we observed that there was a 36% decrease in P450 concentration when the enzyme was stored overnight (Figure 2.15 (a)) and it appeared as though this decrease in P450 concentration ultimately also led to a decrease in enzyme activity (Figure 2.15 (b)). However, when we compare reactions 7 and 8, we observed that the activity of reaction 7 (no additional NADH) levelled off after 8 h and no P450 was detected after 18 h (*). On the other hand, the additional NADH that was added to reaction 8 increased enzyme activity and it appeared to have a stabilising effect on CYP153A6 in terms of remaining P450. In fact, when comparing the concentration of remaining P450 after 18 h which was calculated for reactions 3 and 8 relative to the initial P450 concentration in the BRM for each reaction, reaction 8 had 36% P450 remaining while reaction 1 only had 20%. Furthermore, the total turnover numbers calculated for reaction 8 was significantly higher than for reactions 3 and 7 (Figure 2.15 (c)).

Although the presence of GOx increased the activity of CYP153A6, the activity levelled off much more quickly than when no GOx was present. In addition, no P450 was detected after 18 h in reactions that contained GOx at 20 °C therefore it clearly had a negative effect on the stability of CYP153A6. A higher temperature, such as 30°C, also appeared to have a negative effect on stability and activity. Furthermore, when comparing reactions that contained 400 U/L GOx and 600 U/L GOx there was no significant difference between these two concentrations at both 20 and 30°C. Although storing CYP153A6 overnight decreased the initial P450 concentration which resulted in decreased enzyme activity, it was decided that for all subsequent experiments (in Chapter 3 as well), enzymes would be stored overnight and additional NADH would be added to all reactions because of its stabilising effect on CYP153A6.

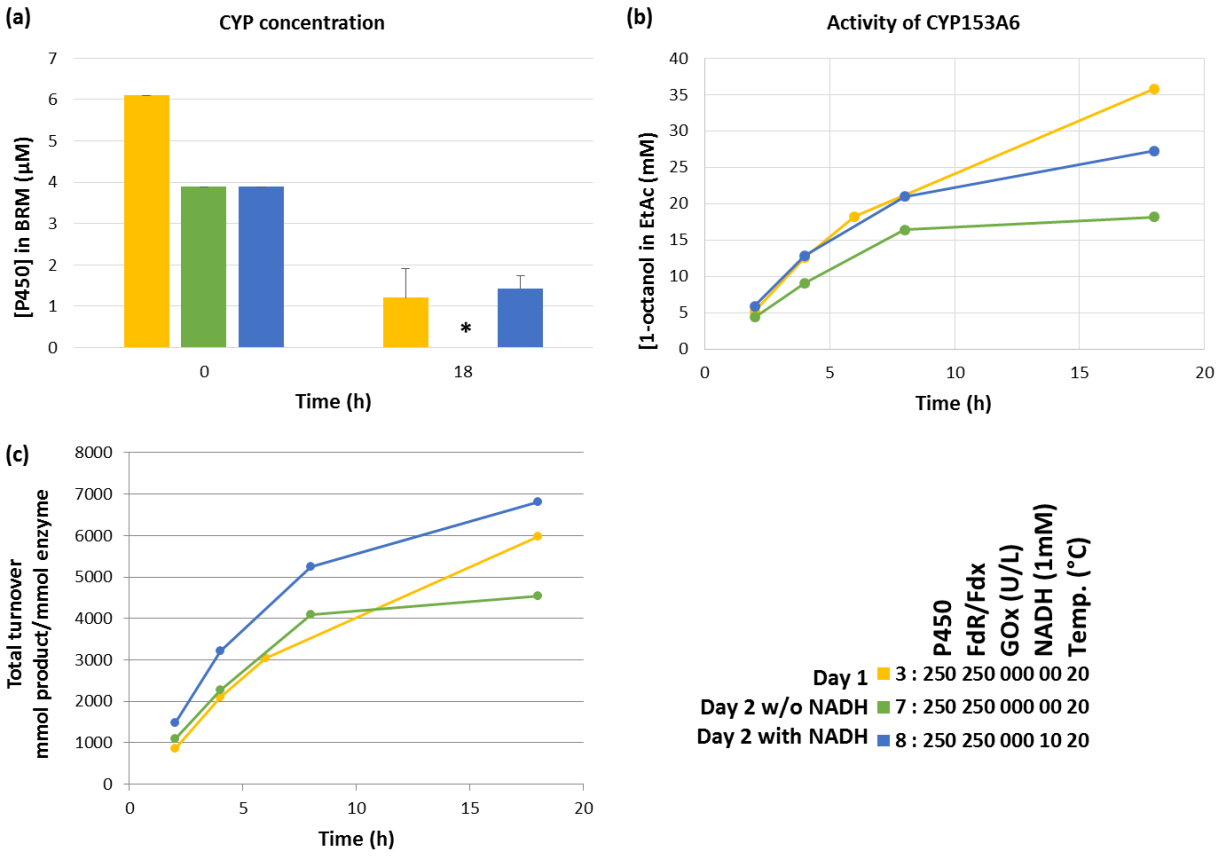

Figure 2.15 Reactions illustrating the effect on CYP163A6 stability (a) and activity (b) when CYP153A6 and FdR/Fdx were stored overnight, on ice, in a 4°C fridge. Reaction 3 (Day 1) contained enzymes that were not stored overnight, the enzymes were used on the same day that *E.coli* expressing the protein was harvested and the cell free extract was obtained. Reactions 7 and 8 (Day 2) contained the same enzyme as was used in reaction 3 but it had been stored overnight. Reaction 8 received additional NADH while reaction 7 did not. Besides P450 concentration and enzyme activity, total turnover (c) was also calculated in this experiment.

2.2.2.4 Evaluating different buffers, buffer concentrations and buffer pH to optimise biotransformation of n-octane using CYP153A6

In all the previous experiments, biotransformation reaction mixtures were made up using 200 mM Tris-HCl buffer (pH 8) which contained 100 mM glycerol and 100 mM glucose or no glucose because preliminary experiments performed to evaluate different buffers and pH, indicated that 200 mM Tris-HCl (pH 8) containing 100 mM glucose and glycerol improved the activity of CYP153A6 the most. This was also demonstrated by Pennec and workers (2014) when they performed whole cell biotransformations using the same buffer and obtained a 1-octanol concentration 2.6 g L_{BRM}^{-1} after only 8 h with a biomass

concentration of 5 g_{DCW} L_{BRM}^{-1} compared to Gudiminchi and co-workers (2012) who obtained a 1-octanol concentration of 3.4 g L_{BRM}^{-1} after 15 h using a biomass concentration of 7 g_{DCW} L_{BRM}^{-1} and a 200 mM sodium phoshate buffer (pH 7.2) containing 100 mM glucose and 0.8% glycerol. However, 200 mM Tris was the highest buffer concentration evaluated in preliminary experiements therefore, we decided to evaluate higher buffer concentrations at a variety of pH values as an attempt to improve the activity of CYP153A6 in biotransformations using cell free extracts (Table A6 and A7). These conditions were evaluated in two separate experiments which were carried out simultaneously.

After the proteins (CYP153A6 and FdR/Fdx) were expressed in *E. coli*, the cells were harvested and resuspended into 4 buffers, 200 mM Tris-HCl, pH 8 (buffer 1), 200 mM Tris-HCl, pH 8.5 (buffer 2), 400 mM Tris-HCl, pH 8 (buffer 3) and 20 mM Tris-HCl, pH 8 (buffer 4). CFE from cells resuspended in the first three buffers were used in an experiment using only Tris-HCl buffer at 200 mM and 400 mM and at pH 8 and 8.5 while CFE from cells resuspended in the fourth buffer were used in an experiment using different buffers at different pH ranging from 6 to 8.5 (Table A6). The cells were then disrupted and the cell free extracts obtained for biotransformations.

The initial P450 concentration calculated for CYP153A6 in buffer 1 was 30 µM, in buffer 2, 9 µM and in buffer 3, 12 µM therefore, before biotransformations even began it was clear that different Tris concentrations and different pH had a huge effect on CYP153A6. The desired final concentration of CYP153A6 in the BRM was 3 µM therefore CYP153A6 in buffer 1 (30 µM) was diluted with empty vector CFE to a concentration of 14 µM to try to achieve this final concentration in the BRM but this was not easily achieved hence the large error bars for reactions 1 and 4 at time 0 h (Figure 2.16 (a)). The initial concentration for CYP153A6 in buffers 2 and 3 were not high enough therefore the CYP was not diluted with empty vector CFE. Furthermore, all reactions received 100 mM glucose which was added to the respective buffers which made up the BRM and all reactions received additional 1 mM NADH (Table A6). Reactions were carried out at 20 and 30°C for 2, 8 and 18 h.

Results in Figure 2.16 illustrate the effect of different Tris buffer concentrations at either pH 8 or 8.5 on CYP153A6 stability (a) and activity (b). Reactions 1 and 4 which contained 200 mM Tris buffer (pH 8) had the highest activity of their respective temperatures and reaction 1 had the highest activity overall producing up to 20 mM 1-octanol after 18 h but both reactions levelled off after 8 h. Unfortunately, the CO-difference spectra for reaction 1 had too much noise therefore it was difficult to calculate the P450 concentration at 18 h. Reactions 2 and 5 which contained 400 mM Tris buffer (pH 8) had no activity and reactions 3 and 6 which contained 200 mM Tris buffer (pH 8.5) had very poor activity. When comparing

reactions 1, 2 and 3, it was clear that buffer concentration and pH plays an important role in enzyme activity.

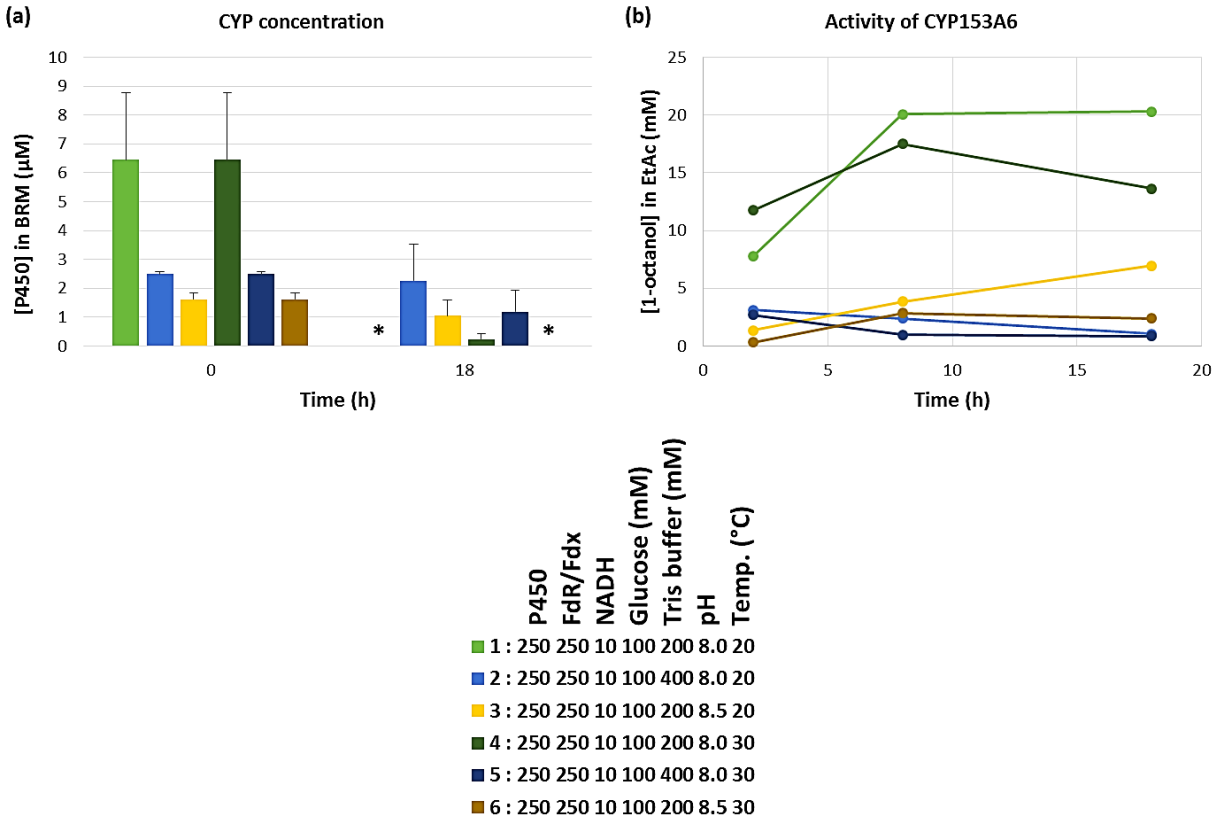

Figure 2.16 Reactions illustrating the effect of different Tris buffer concentrations and pH on CYP153A6 stability (a) and activity (b). Reactions 1 and 4 contained 200 mM Tris buffer (pH 8), reactions 2 and 5 contained 400 mM Tris buffer (pH 8) and reactions 3 and 6 contained 200 mM Tris buffer (pH 8.5). Reactions 1, 2 and 3 were incubated at 20°C and reactions 4, 5 and 6 were incubated at 30°C.

In the second experiment, the initial P450 concentration calculated for CYP153A6 in buffer 4 was 19 µM which was diluted with empty vector CFE to 14 µM so that the final concentration in the BRM would be 3 µM but this was again not easy to achieve (Figure 2.17 (a)).

Two buffers were prepared that would be used to make up the BRM for each reaction according to the experimental design (Table A7). The first buffer was 300 mM Tris (pH 8) and the second buffer was a mixture of 300 mM MES, 300 mM MOPS and 300 mM Tris buffers. This buffer was prepared at six different pH:

1. 300 mM MES-MOPS-Tris (pH 6)
2. 300 mM MES-MOPS-Tris (pH 6.5)
3. 300 mM MES-MOPS-Tris (pH 7)
4. 300 mM MES-MOPS-Tris (pH 7.5)
5. 300 mM MES-MOPS-Tris (pH 8)
6. 300 mM MES-MOPS-Tris (pH 8.5)

Cell free extracts of CYP153A6 and FdR/Fdx were prepared in 20 mM Tris (pH 8) buffer to avoid changes to the final buffer concentration (300 mM) in the BRM as well as changes to the desired pH for each of the respective buffers used for biotransformations. Again, Glucose (final concentration, 100 mM) was added to each of the buffers once preparations for the biotransformation began. In addition, initial pH readings were taken to ensure that the pH in the BRM was at the desired pH as stated above. Each reaction was carried out in triplicate for only 2 hours at 20°C.

Reactions in Figure 2.17 illustrate the effect of these different buffers and their pH on CYP153A6 stability (a) and activity (b). Reaction 1 which contained 300 mM Tris (pH 8) had the highest activity after 2 h producing up to 10 mM 1-octanol. Interestingly, reaction 1 in the previous experiment which contained 200 mM Tris (pH 8) produced less 1-octanol (7 mM) after 2 h. The ideal pH for the 300 mM MES-MOPS-Tris buffer was at pH 7 (reaction) because CYP153A6 had the second highest activity at this pH and enzyme was relatively stable. The P450 concentration was the lowest at pH 6 (reaction 2) and pH 6.5 (reaction 3) most probably due to the buffer being too acidic, this reflected in the activity of reaction 2 but interestingly the activity of reaction 3 was quite high after 2 h. as for Reactions 5, 6 and 7, although CYP153A6 was quite stable, its activity was lower.

Although reaction 1 did not have the best result in terms of P450 concentration after 2 h, CYP153A6 had the best activity overall with the 300 mM Tris (pH 8) buffer. The best results for the MES-MOPS-Tris buffer was observed for reaction 4 MES buffer and MOPS buffer are quite expensive therefore using the conditions of reaction 5 would not be feasible. Therefore, the biotransformations of all subsequent experiments were made up with 300 mM Tris buffer (pH 8).

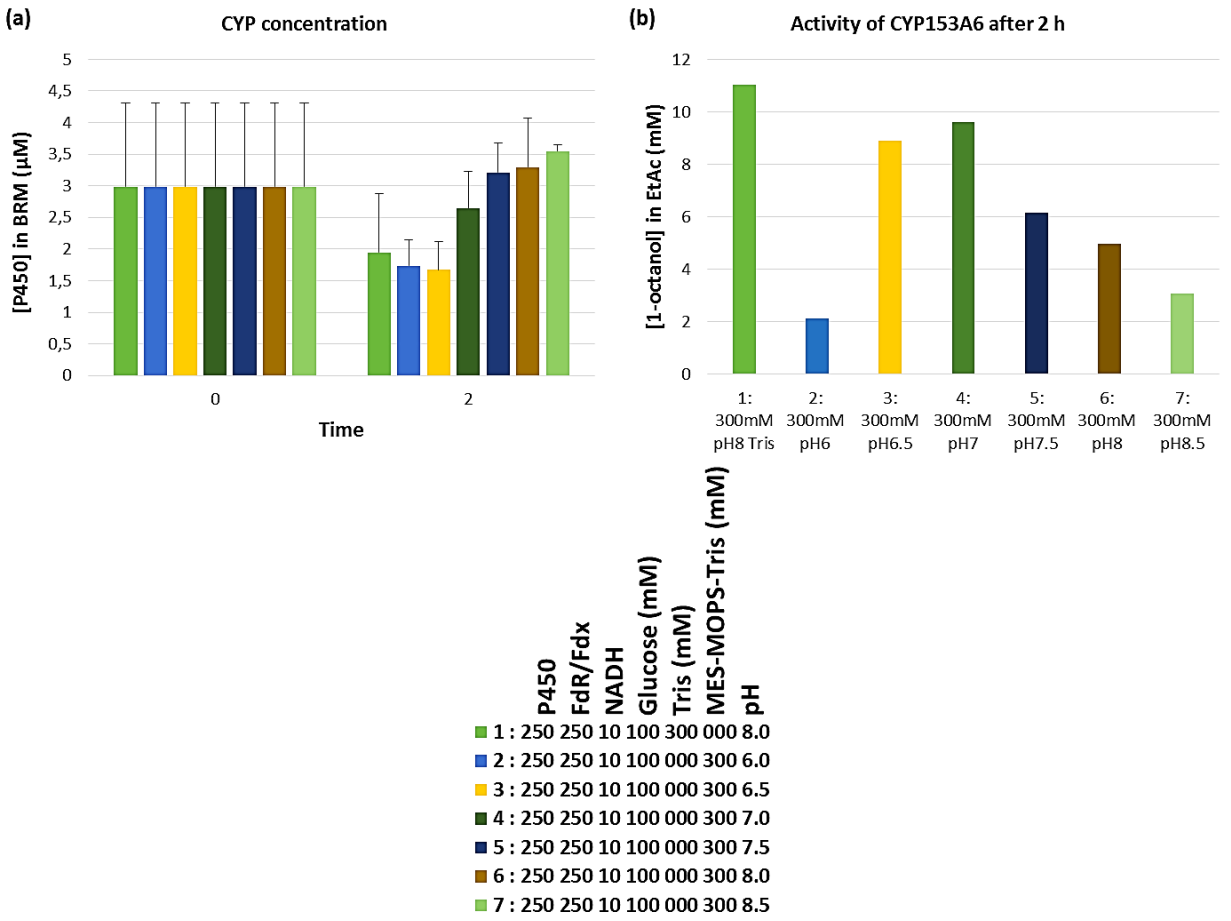

Figure 2.17 Reactions illustrating the effects of 300 mM Tris buffer and 300 mM MES-MOPS-Tris buffer on CYP153A6 stability (a) and activity (b). Reaction 1 contained 300 mM Tris (pH 8), reaction 2 contained 300 mM MES-MOPS-Tris (pH 6), reaction 3 contained 300 mM MES-MOPS-Tris (pH 6.5), reaction 4 contained 300 mM MES-MOPS-Tris (pH 7), reaction 5 contained 300 mM MES-MOPS-Tris (pH 7.5), reaction 6 contained 300 mM MES-MOPS-Tris (pH 8 and reaction 7 contained 300 mM MES-MOPS-Tris (pH 8.5).

2.2.2.5 Evaluating different concentrations of CYP153A6 in the biotransformation reaction mixture to optimise biotransformation of n-octane

In this experiment, the biotransformation reaction mixtures were made up using three different concentrations of CYP153A6 to evaluate if the enzyme would be more stable at a higher or lower concentration. Concentration was evaluated by adding different volumes of CYP153A6 containing CFE to the BRM, these were, 50 µL, 100 µL and 250 µL therefore the final biomass contribution in the BRM for each of these volumes of CYP153A6 would be 0.021 g_{WCW} mL_{BRM}^{-1}, 0.041 g_{WCW} mL_{BRM}^{-1} and 0.104 g_{WCW} mL_{BRM}^{-1} respectively. In order to keep the enzyme as stable as possible because it was being stored overnight, CFE were prepared in 200 mM Tris (pH 8) buffer containing 100 mM glycerol as usual but

because the final concentration of the buffer had to be 300 mM Tris (pH 8), three different buffers had to prepared at a higher concentration in order to achieve this final concentration:

1. A 333 mM Tris buffer (pH 8) containing 100 mM glycerol and 100 mM glucose was prepared which was added to biotranformations containing 50 µL of P450.
2. A 341 mM Tris buffer (pH 8) containing 100 mM glycerol and 100 mM glucose was prepared which was added to biotranformations containing 100 µL of P450.
3. A 371 mM Tris buffer (pH 8) containing 100 mM glycerol and 100 mM glucose was prepared which was added to biotranformations containing 250 µL of P450.

The initial P450 concentration calculated was 19 µM and the desired final concentration in biotransformations for reactions containing 50, 100 and 250 µL of P450 was 1, 2 and 4 µM respectively. However, the initial P450 concentration was too low to be diluted with empty vector CFE in order to achieve the desired final concentrations. Therefore, CYP153A6 was not diluted with empty vector CFE because the initial P450 concentration had to at least be 24 µM in order to obtain the desired final P450 concentrations in a 1.2 mL BRM. Furthermore, all reactions received an additional 1 mM NADH and were carried out for 2, 8 and 20 h at 30°C (Table A8).

Results in Figure 2.18 illustrate the effects of different P450 concentration on the stability (a) and activity (b) of CYP153A6. Surprisingly, the final P450 concentrations for each of the respective volumes of P450 added to the BRM were higher than expected (Figure 2.18 (a)) despite low initial P450 concentration (19 µM). The results clearly indicated that enzyme activity increased proportionally with increasing P450 concentration (Figure 2.18 (b)). Reactions 2 and 3 clearly illustrate that with lower P450 concentrations then reaction levels off more quickly. For reaction 1, we observed that there was a decrease in activity after 8 h, however the reaction had not completely levelled off. In terms of remaining P450 after 20 h, we observed that CYP153A6 was more unstable at higher temperature.

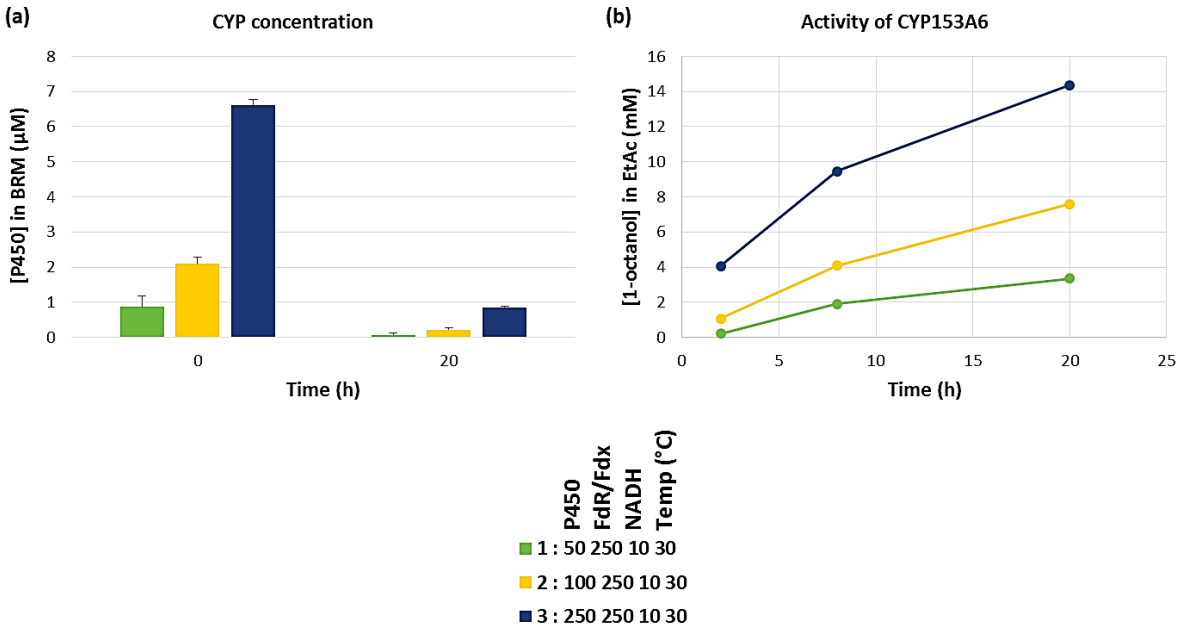

Figure 2.18 Reactions illustrating the effects of different P450 concentrations on the stability (a) and activity (b) of CYP153A6. Reactions 1 contained 50 µL P450; reaction 2 contained 100 µL P450 and reaction 3 contained 250 µL P450. All reactions were incubated at 30°C

2.2.3 Conclusion

Wild-type *E. coli* enzymes in CFEs and wild-type enzymes supplemented with glucose dehydrogenase were evaluated for cofactor regeneration. The stability and activity of CYP153A6 was negatively affected by glucose dehydrogenase (GDH). This was most probably due to the production of gluconic acid which caused the environment surrounding CYP153A6 to become too acidic. When additional ferredoxin reductase and ferredoxin (FdR/Fdx) were added to biotransformations, there was a significant increase in the activity of CYP153A6. However, as the activity of CYP153A6 increased it became more unstable because the P450 concentration decreased proportionally.

After much evaluation it was also clear that glucose played an important role in the stability of CYP153A6 and although there was a decrease in initial P450 concentration when storing CYP153A6 overnight, additional NADH helped increase enzyme stability and activity, even the total turnover numbers were higher at each time interval.

When glucose oxidase (GOx) was added to biotransformations there was an unexpected increase in the activity of CYP153A6 but these reactions levelled off after 8 h while reactions which did not contain GOx were still continuing. Furthermore, no P450 was detected after 18 h in reactions that contained GOx

therefore it was clear that H_2O_2 produced by GOx was having a negative effect on the stability of CYP153A6.

The thermal stability of CYP153A6 was challenged by performing biotransformations at 30°C and in most cases, the activity levelled off after 8 h whereas at 20°C the activity still continued. Unstable proteins also tend to denature faster at higher temperatures which is possibly the reason why no P450 was detected for most of the reactions incubated at 30°C. However, it appeared as though using higher P450 concentrations helped with this issue.

The buffer in which *E. coli* cells expressing CYP153A6 were resuspended also played an important role in the stability of the enzyme especially when the enzyme was stored overnight. Although glucose in the BRM had a stabilising effect on CYP153A6, it was also important that glucose only be added to the buffer once the biotransformation reaction mixture was being prepared. Furthermore, higher buffer concentrations had a negative effect on the activity of CYP153A6 as well as buffers which had a pH of less than 7 or higher than 8.

The conditions selected to evaluate the operational stability of the wild type CYP153A6 were not only to try to optimise biotransformation of *n*-octane, but also to find conditions that will challenge the enzyme because these conditions would be used when performing biotransformations of *n*-octane using mutants of CYP153A6 (Chapter 3). Therefore, the final conditions selected for evaluating the operational stability of the CYP153A6 mutants were:

- Storing the enzymes (CYP153A6 and FdR/Fdx) overnight in 200 mM Tris-HCl buffer (pH 8) containing 100 mM glycerol and performing biotransformation the following day
- Making up the BRM with 300 mM (final concentration) Tris-HCl buffer (pH 8) containing 100 mM glycerol and 100 mM (final concentration) glucose.
- Diluting CYP153A6 with empty vector CFE to achieve a final concentration of 1.5 µM in the BRM using CFE from a P450 containing cell suspension with effectively a final biomass concentration of 0.104 g_{WCW} mL_{BRM}^{-1}
- Adding additional FdR/Fdx to a final biomass concentration of 0.104 g_{WCW} mL_{BRM}^{-1} to all reactions
- Adding additional 1 mM NADH to all reactions
- Incubating biotransformation reactions at 30 °C
- Adding a higher concentration (> 600 U/L) of glucose oxidase to all reactions

Chapter 3

Design, construction and evaluation of CYP153A6 mutants for improved operational stability

Biocatalyst instability among P450 enzymes is largely caused by hydrogen peroxide (H_2O_2) which is produced due to protein uncoupling (Loida & Sligar, 1993). Uncoupling mostly occurs in the absence of substrate, causing the ferre-hydroperoxide compound to decay forming H_2O_2 (Denisov *et al.*, 2005). In turn, H_2O_2 oxidises sulphur-containing amino acids -cysteine (Cys) and methionine (Met) - residues in the protein because these amino acids are the most susceptible to oxidation by reactive oxygen species (Opperman & Reetz, 2010). However, many research groups have found that substituting cysteine and methionine residues with more oxidatively stable amino acids results in a more stable enzyme (Kemble *et al.*, 2001; Opperman & Reetz, 2010; Slavica *et al*., 2005).

Therefore, this chapter focuses on the design and construction of CYP153A6 mutants using site directed mutagenesis to substitute specific cysteine and methionine residues in CYP153A6 with more oxidatively stable amino acids. The activity of the mutants was evaluated by performing biotransformations of *n*-octane using cell free extracts.

3.1 Materials and methods

3.1.1 Bacterial strains and plasmids.

All bacterial strains and plasmids are listed in Table 2.1. CYP153A6 from *Mycobacterium sp*. HXN-1500 was expressed as an operon from the pET28b(+) vector (Figure 3.1) which included the ferredoxin reductase (FdR) and ferredoxin (Fdx) (Gudiminchi *et al*., 2012). In this study, 25 mutants of CYP153A6 were evaluated for improved operational stability, 11 of the mutants were previously created by Dr. D.J. Opperman and 14 additional mutants were designed and constructed as part of this study. Expression was done in *E. coli* BL21-Gold (DE3).

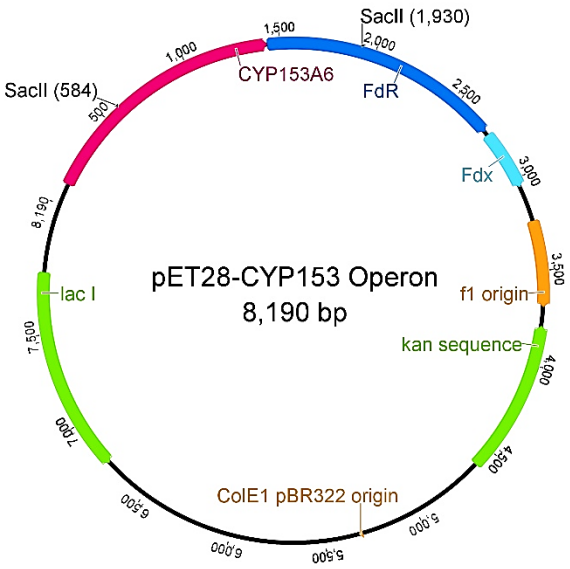

Figure 3.1 Vector map illustrating the complete operon from Mycobacterium sp. HXN-1500 encoding CYP153A6, ferredoxin reductase (FdR) and ferredoxin (Fdx) cloned into pET-28b (+). The kanamycin resistance gene, lacI gene and the f1 origin are also indicated in the vector map. Figure generated with Geneious R6

3.1.2 Site-directed mutagenesis to construct CYP153A6 mutants

3.1.2.1 Designing CYP153A6 mutants

3.1.2.1.1 3DM Information System for Cytochrome P450s created by Bio-Prodict

CYP153A6 mutants were designed using the 3DM Information System for Cytochrome P450s created by Bio-Prodict (https://3dm.bio-prodict.nl/index.php). The 3DM Information System relies on structure based multiple sequence alignments to group protein sequences belonging to a superfamily into subfamilies centred on available X-ray structures. The 3DM Information System then uses a unified numbering scheme for residues belonging to the structurally conserved core of the superfamily to collect and compare amino acid information for residues at structurally equivalent positions.

The process for mutant design using 3DM Information System was as follows:

1. A protein BLAST (Basic Local Alignment Search Tool) search was performed using the amino acid sequence of CYP153A6 as the query sequence. The search found 335 sequences with at least 50% amino acid identity to CYP153A6. The majority (327) of these sequences belonged to the CYP153A7 subfamily based on the X-ray structure of CYP153A7 (Pham *et al.*, 2012).

2. These 335 sequences were used to create a 3DM Information System subset and amino acid distribution and correlated mutation statistics were automatically calculated for this subset.

3. Positions in the core for which CYP153A6 has Met or Cys were selected and the system displayed the amino acid distribution for those specific positions in the consensus sequence alignment of the subset in the form of graphs (Figure A1).

3.1.2.1.2 YASARA (Yet Another Scientific Artificial Reality Application)

Yasara (Krieger *et al.*, 2002) was used to create a homology model of CYP153A6 using P450pyr (Pham *et al*., 2012, PDB: 3RWL) as template. The positions of Cys and Met residues that were to be mutated were mapped on this structure: 3RWL

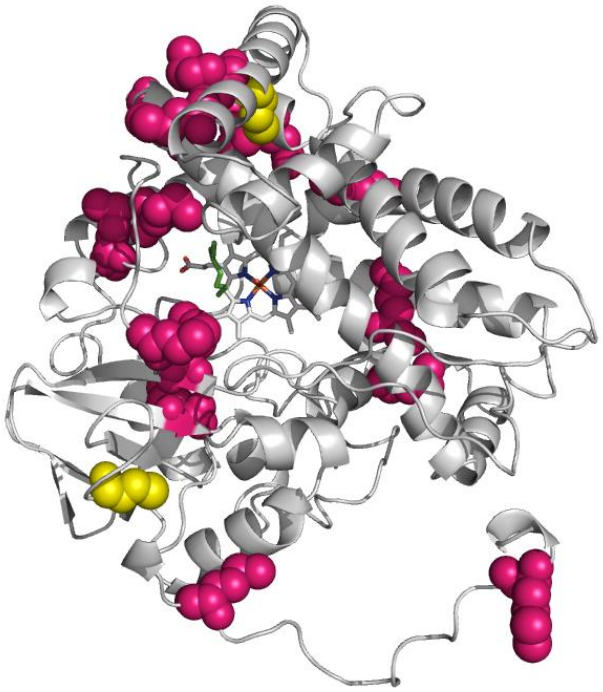

Figure 3.2 Homology model of CYP153A6 mapping the positions of the Cysteine (yellow) and Methionine (pink) residues that were mutated in this study to improve the operational stability of CYP153A6. The substrate (n-octane) is shown in green at the centre of the active site above the heme propionyl group (blue/orange/grey) Figure generated with Pymol.

3.1.2.1.3 Primer design

All primers used in constructing CYP153A6 mutants are listed in Table 3.1. The underlined codons in each primer indicated where the mutation was introduced. Only codons that were known to express well in *E. coli* genes were selected for mutagenesis. Primers were designed using Geneious R6 (Java version

1.6.0_35-b10 (64 bit), Biomatters Ltd, Auckland, New Zealand [http://www.geneious.com/]) and the Integrated DNA Technologies OligoAnalyzer 3.1 online tool (https://eu.idtdna.com/calc/analyzer); Integrated DNA Technologies also supplied the primers for this study.

3.1.2.2 Polymerase Chain Reaction (PCR) amplification

Site-directed directed mutagenesis was performed using the Quickchange (Stratagene) PCR method. PCR reaction mixtures (50 µL) consisted of 10X KOD Hot Start DNA Polymerase buffer (5 µL), $MgSO_4$ (1.5 mM), deoxynucleoside triphosphates (dNTPs, 0.2 mM each), KOD Hot Start DNA Polymerase (1 U), plasmid DNA (20 ng) and both forward and reverse primers (Table 3.1, 0.1 µM).

PCR amplification was carried out using the T100™ Thermal Cycler (Bio-Rad). The thermal cycling conditions consisted of an initial incubation step at 95°C for 2 min to activate the polymerase. This was followed by 18 cycles[4] of denaturing at 95°C for 20 s, annealing of the primer to the template for 10 s (Table 3.1) and an elongation step at 70°C for 4 min (25 s/kb). Thermal cycling concluded with a final elongation step at 70°C for 10 min to ensure complete elongation of the amplicons.

3.1.2.2.1 Megaprimer PCR amplification

CYP153A6 mutations C52V and M164T were the only two mutations which were not achieved using the standard Quickchange PCR method. To overcome this problem both mutations were introduced simultaneously using a Megaprimer-method (Sanchis *et al*., 2008). Site-directed mutagenesis using the 'megaprimer' PCR method was first introduced by Kammann and co-workers (1989) and describe the use of two rounds of PCR; in the first PCR, the megaprimer containing the desired mutation(s) is constructed and amplified, the "megaprimer" is then purified from an agarose gel and used as the primer in the second PCR to insert the mutation(s) into the DNA sequence (Kammann *et al.,* 1989). The method used here (Sanchis *et al*., 2008), consisted of only one round of PCR with two thermal cycling stages (without a 'megaprimer purification' intermediate step). The PCR reaction mixture was prepared as described previously but in order to insert both mutations (C52V and M164T) simultaneously, a combination of the two primer sets were used (C52V_Forward and M164T_Reverse primers or C52V_Reverse and M164T_Forward). Table 3.2 illustrates the thermal cycling conditions used: The first stage of PCR generates the double-stranded megaprimer containing the mutations introduced by the mutagenic

[4] Mutants M101S, M105R and M129F were amplified by reducing the number of cycles to 16 and increasing the final elongation time to 5 min.

primers. The second stage of PCR extends the megaprimer using the wild type plasmid DNA template, this yields the complete DNA sequence containing the two mutations (Figure A2).

Table 3.1 Primers used for site-directed mutagenesis by the Quickchange (Stratagene) PCR method

CYP153A6 mutation	Forward primer sequence (5'- 3')	Reverse primer sequence (5'- 3')	T_m (°C)	PCR annealing temperature
C52V	CCG GTT CAC TAC GTG AAG AGC AGC ATG	CAT GCT GCT CTT CAC GTA GTG AAC CGG	62.9	69
C208A	TG GAG GCC GCG ACG TAT TTC ACC G	C GGT GAA ATA CGT CGC GGC CTC CA	62.5	68
M56G	AAG AGC AGC GGC TTC GGC CCC TAC TG	CA GTA GGG GCC GAA GCC GCT GCT CTT	68.1	68
M91L	C GGC ATC ACC ATC CTG GAC GAC AAC	GTT GTC GTC CAG GAT GGT GAT GCC G	63.6	68
M101N	GCA TCG CTT CCC AAC TTC ATC GCC ATG	CAT GGC GAT GAA GTT GGG AAG CGA TGC	63.8	68
M101S	GCA TCG CTT CCC AGC TTC ATC GCC ATG	CAT GGC GAT GAA GCT GGG AAG CGA TGC	65.5	69
M105R	CC ATG TTC ATC GCC CGT GAT CCA CCG AAA C	G TTT CGG TGG ATC ACG GGC GAT GAA CAT GG	65.7	69
M129F	GAA AAC CTT GCC ACC TTC GAA TCG GTG ATT CG	CG AAT CAC CGA TTC GAA GGT GGC AAG GTT TTC	63.9	60
M164T	GAA TTG ACC ACG AAG ACC CTG GCG ACG	CGT CGC CAG GGT CTT CGT GGT CAA TTC	64.7	69
M202R	GAA GAA CAG CGC CGT GCC GAG CTG ATG	CAT CAG CTC GGC ACG GCG CTG TTC TTC	66.9	68
M206V	CGA GCT GGT GGA GTG CGC GAC GTA TTT	AAA TAC GTC GCG CAC TCC ACC AGC TCG	66	68
M241D	CA ACA CGA CAC GAT GCG CCC GAG GAA TAT C	G ATA TTC CTC GGG CGC ATC GTG TCG TGT TG	65.3	68
M306G	CT CTT TCG CAC GGT CGT CGT ACC GC	GC GGT ACG ACG ACC GTG CGA AAG AG	65.2	68
M329L	GGC GAC AAA GTC GTG CTG TGG TAC GTG TCC	GGA CAC GTA CCA CAG CAC GAC TTT GTC GCC	66.9	68

Underlined codons in primers indicate introduced mutation

Table 3.2 Megaprimer PCR thermal cycling conditions

Cycling stage	Step	Temperature (°C)	Time	Number of cycles
Synthesis of megaprimer	Initial denaturation	95	2 min	1
	Denaturation	95	20 sec	6
	Annealing	58	10 sec	
	Extension	70	10 sec	
Nicked DNA sequence containing mutations	Denaturation	95	20 sec	20
	Annealing	70	10 sec	
	Extension	70	4 min	

3.1.2.3 Visualisation and Purification of PCR and restriction enzyme digestion products

PCR amplification was confirmed by DNA gel electrophoresis with 0.8% (w/v) agarose gels in TAE [Tris (40 mM), Acetic acid (20 mM), EDTA (1 mM), pH 8.3, Bio-Rad] electrophoresis buffer for approximately 30 min at 100 V. Ethidium bromide was added to agarose gels to facilitate DNA visualisation using a ChemiDox™ XRS+ Gel Documentation system (Bio-Rad). The size of DNA fragments were determined by loading GeneRuler™ DNA ladder mix (Thermo Scientific) adjacent to PCR samples (Figure 3.3). PCR products for which the size of the DNA fragments had been confirmed were digested with *Dpn*I at 37°C overnight, to ensure removal of the template plasmid DNA. The *Dpn*I-digested PCR product was then separated on a second agarose gel (as described above) and visualised on a DarkReader™ transilluminator (Bio-Rad) to facilitate the excision of the desired DNA band from the agarose gel using a sterile surgical blade. The DNA was purified using the BioSpin Gel Extraction Kit (BioFlux), according to the manufacturer's instructions.

3.1.2.4 Restriction enzyme digestion and Ligation

After completing the Megaprimer PCR and the desired PCR products had been identified using gel electrophoresis and purified (section 3.1.2.3), a restriction enzyme digestion was performed to separate the two mutations, C52V and M164T. Restriction enzyme, Cfr241 (SacII) (Fermentas) was used to digest the plasmid DNA carrying both mutations as well as the plasmid DNA carrying the wild type CYP153A6 operon, as illustrated in Figure 3.4. Restriction digest reactions were performed in volumes of 10 µL as indicated in Table 3.3 and incubated for 3 h at 37°C, yielding two DNA fragments each (Figure 3.4 (b)). Therefore, the plasmid DNA containing the two mutations was cut in such a way that one mutation (C52V)

was part of the “backbone” of the plasmid while the other mutation (M164T) was part of the insert. The digested plasmids were separated on an agarose gel and both fragments were purified.

The mutations were finally separated by ligating the insert containing mutation M164T into the “backbone” of the wild-type CYP153A6 operon plasmid and ligating the insert of the wild-type CYP153A6 operon plasmid in the “backbone” of the plasmid containing the C52V mutation (Figure 3.4 (c)). Ligation of DNA fragments was performed using T4 DNA ligase (Thermo Scientific); reaction mixtures were prepared as indicated in Table 3.3 and incubated overnight at 4°C.

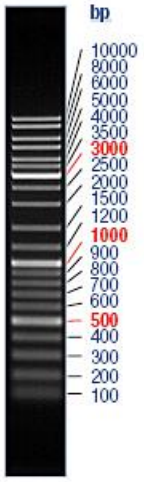

Figure 3.3 GeneRuler™ DNA ladder mix (Thermo Scientific) was used as reference to determine if PCR amplification of CYP153A6 mutants was successful.

Table 3.3 Composition of reaction mixtures to perform restriction enzyme digestion and ligation

Restriction Enzyme Digest		**Ligation**	
Plasmid DNA	6 µL	T4 DNA ligase	0.5 µL
10X Buffer B	1 µL	10X Buffer R	1 µL
Restriction enzyme	0.5 µL	“Backbone”	4 µL
Deionised water	3.5 µL	Insert	4.5 µL
Total volume	10 µL	Total volume	10 µL

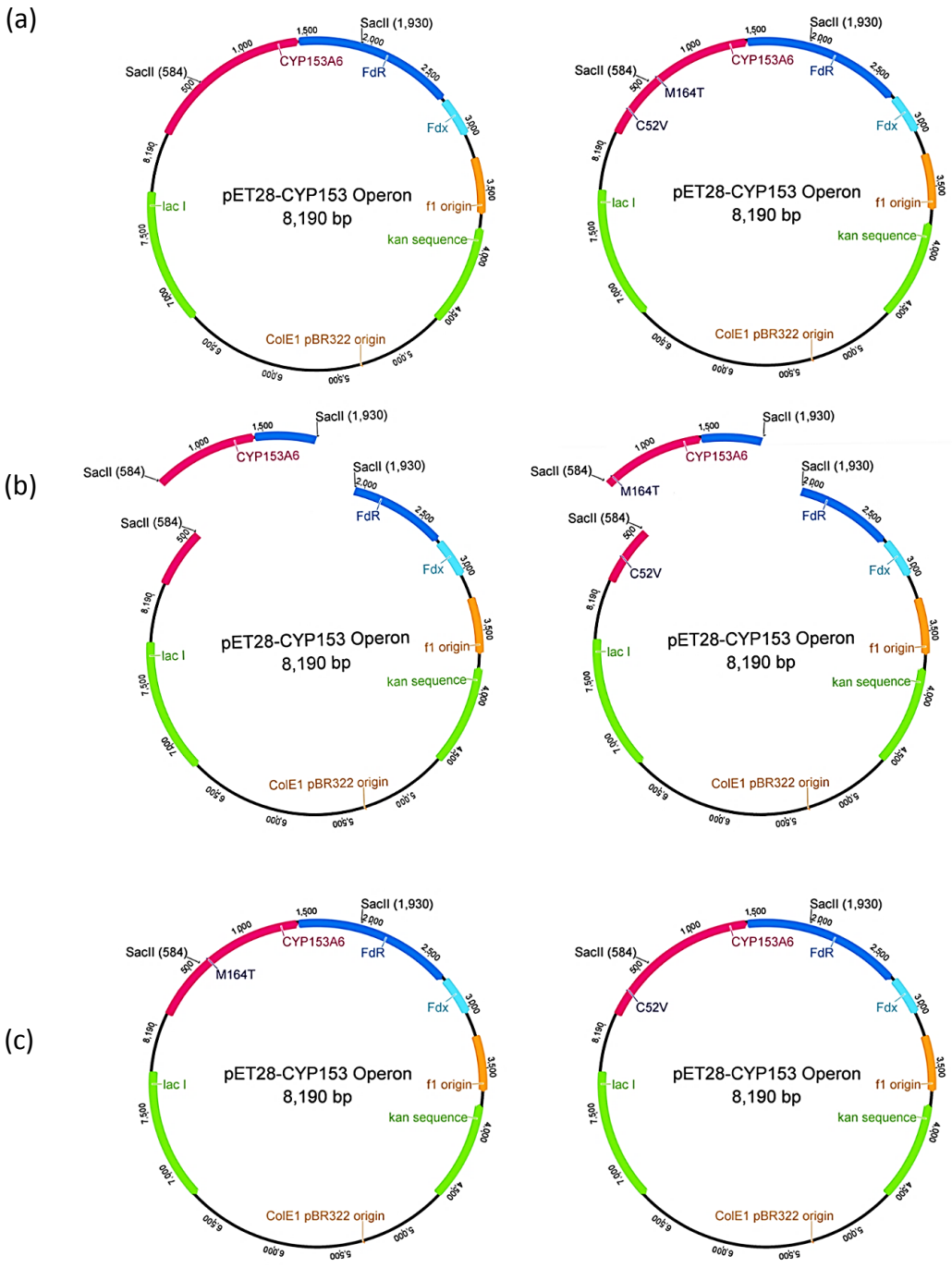

Figure 3.4 Schematic representation of restriction enzyme digestion and ligation reaction used to separate mutations C52V and M164T. Restriction enzyme Cfr241 (SacII) was used to digest the plasmid DNA carrying both mutations as well as the plasmid DNA carrying the wild type CYP153A6 operon (a). Restriction digest reactions yielded two DNA fragments each therefore, the plasmid DNA containing the two mutations was cut in such a way that mutation C52V was part of the "backbone" of the plasmid while the mutation M164T was part of the insert (b). The insert containing mutation M164T was ligated into the "backbone" of the wild-type CYP153A6 operon plasmid while the insert of the wild-type CYP153A6 operon plasmid was ligated into the "backbone" of the plasmid containing the C52V mutation thus separating the two mutations into their own pET-28b (+) plasmids (c).

3.1.2.5 DNA sequencing

To verify that all mutations were correctly inserted using the Quickchange PCR method, the plasmids were submitted for DNA sequencing using a Hitachi 3130*xl* Genetic Analyser (Applied Biosystems). The BigDye® Terminator (v 3.1) Cycle Sequencing Kit (Applied Biosystems®) was used to amplify the insert for sequencing according to the manufacturer's instructions, using the primers listed in Tables A9 and A10 (appendix). A final concentration of sequencing primer of 3.2 pmol was used in a total reaction volume of 10 µL. The nucleotide sequences of CYP153A6 mutants were analysed using Geneious R6 by performing a sequence alignment with the wild type CYP153A6.

3.1.2.6 Transformation

Transformation reactions were performed by thawing 50 µL of competent *E. coli* Top10 cells (Life Technologies™) on ice before adding 4 µL of plasmid DNA. After the cells had been incubated for 30 min on ice, they were heat shocked at 42°C for 40 s and transferred back to the ice for another 2 min where after 250 µL of SOC media (Table A1) was added to the cells and incubated at 37°C for 1 h. Using sterile techniques, the transformation mixture was plated out on LB plates containing kanamycin (30 µg mL^{-1}) and incubated at 37°C overnight.

3.1.2.7 Plasmid proliferation and extraction

E. coli that had been successfully transformed presented as single colonies on LB plates. A few of these single colonies were selected and inoculated into 5 mL LB media also supplemented with kanamycin (30 µg mL^{-1}) and incubated at 37°C on a shaker (160 rpm) overnight. Plasmids were then isolated from *E. coli* cells and purified using the BioSpin Plasmid DNA Extraction Kit (BioFlux) according to the manufacturer's instructions.

3.1.2.8 Transforming E. coli BL21-Gold (DE3) for protein expression

After DNA sequencing confirmed that each of the 14 CYP153A6 mutations were successful, mutants were transformed into *E. coli* BL21-Gold (DE3) for protein expression as described above. A loopful of the transformed cells was inoculated into 5 mL LB media and incubated overnight at 37°C on a shaker (160 rpm). Glycerol stocks were prepared for each of the cultures by mixing 1 mL of culture with 1 mL of 50% (v/v) glycerol in sterile 2 mL cryotubes and stored at -80°C.

3.1.3 Protein Expression

Pre-cultures were prepared by inoculating approximately 20 µL of transformed *E. coli* BL21-Gold (DE3) cells into 5 mL of LB media supplemented with 30 $\mu g.mL^{-1}$ kanamycin (CYP153A6 (WT) and mutants) or 100 $\mu g.mL^{-1}$ ampicillin (FdR/Fdx) and cultivated at 37°C for 12-16 h on a rotary shaker (160 rpm).

For expression, 2 mL of pre-culture was used to inoculate 100 mL of ZYP5052 auto-induction media (Table A1), also containing the appropriate antibiotics in 500 mL Erlenmeyer flasks. One millimolar δ-aminolevulinic acid hydrochloride (heme precursor) and 50 µM $FeCl_3 \cdot 6H_2O$ (to compensate for the lack of iron in *E.coli*) was also added to aid the expression of CYP153A6 (WT) and the 25 CYP153A6 mutants. Protein expression was carried out at 20°C for 48 h on an orbital shaker (200 rpm).

3.1.4 Cell harvesting and cell disruption

For experiments using cell free extracts, cells were harvested, resuspended and disrupted as described in Chapter 2 section 2.1.3.

For experiments using whole cells, after the cells were harvested by centrifugation, 1 g of cells (wet weight) were resuspended in 10 mL of 200 mM Tris-HCl buffer (pH 8) containing 100 mM glycerol and 100 mM Glucose. Therefore, the initial biomass concentration was 100 $g_{WCW}\ L^{-1}$.

3.1.5 Analysis of Expression

The expression of CYP153A6 mutants were quantified using CO-difference spectra as described in Chapter 2, section 2.1.4.1 and their expression levels were visually analysed using SDS-PAGE as described in Chapter 2, section 2.1.4.2.

3.1.6 Biotransformations

The stability and activity of the mutants were evaluated in whole cells as well as cell free extracts. Reactions sacrificed at different time intervals were used to monitor the stability of the mutants by calculating the concentration of remaining P450 at different times from CO-difference spectra (section 2.1.4.1) For both whole cell and cell free extract experiments, biotransformations were performed in 40 mL amber vials with non-permeable screw caps.

3.1.6.1 Hydroxylation of n-octane using whole cells

All whole cell biotransformations were performed using a 1 mL reaction volume - of which 200 µL was used for CO-difference spectra - where the final biomass concentration of P450 containing cells in the biotransformation reaction mixture (BRM) was 0.05 g_{WCW} mL^{-1}. The BRM also contained 100 mM glycerol and 100 mM glucose in 200 mM Tris-HCl buffer (pH 8). Two hundred and fifty microliters of *n*-octane was added to each reaction vial before it was placed on an orbital shaker and incubated at 20°C with shaking at 200 rpm. All reactions were carried out in duplicate for 2 and 24 h.

3.1.6.2 Hydroxylation of n-octane using cell free extracts (CFE)

Cell free extract (CFE) biotransformations were performed using a 1.2 mL reaction volume – of which 200 µL was used for CO-difference spectra. The CFE in these reactions came from cell suspensions containing 0.500 g_{WCW} mL^{-1} *E. coli* cells expressing only FdR and Fdx (250 µL) and a mixture of CFEs from cell suspensions containing 0.500 g_{WCW} mL^{-1} *E. coli* cells expressing the CYP153A6 operon and cells transformed with empty vector (pET-28b (+)) so that the final P450 concentration in the BRM should have been 1.5 µM and the total volume of the mixture 250 µL. Thus, the total volume of CFE added to the BRM was 500 µL and the protein content representative of a biomass concentration of 0.208 g_{WCW} mL^{-1}. The BRM also contained 100 mM glycerol and 100 mM glucose in 300 mM Tris-HCl buffer (pH 8), and additional 1 mM NADH. All reactions were started by adding 250 µL of *n*-octane to the reaction vial before being placed on an orbital shaker (200 rpm) at 30°C and all reactions were carried out in triplicate for 2, 8 and 18 h. It is also important to note that enzymes were stored on ice in a 4°C fridge, overnight before commencing with biotransformations the following day.

All 25 mutants were evaluated using two rounds of experiments under the conditions described above. In the first round, all the mutants were evaluated by dividing them into four experiments and 800 U/L (final concentration) of glucose oxidase (GOx) (Figure 2.4) was added to all reactions to evaluate the oxidative stability of the mutants in the presence of *in situ* generated H_2O_2. In the second round, all mutant were evaluated by dividing them into five experiments and the GOx concentration in all reactions was increased to 1600 U/L (final concentration) to further challenge the oxidative stability of the mutants.

3.1.7 Sample extraction and product analysis

Samples were extracted and analysed using gas chromatography (GC) as described in Chapter 2, section 2.16.

3.2 Results

The aim of this study was to create mutants of CYP153A6 that would improve the operational stability of the enzyme without compromising enzyme activity and product specificity. The methionine residues of the 11 mutants that were previously created by Dr. Opperman, were mostly surface exposed whereas, the methionine and cysteine residues of the 14 mutants created in this study were located on the inside of protein as well as the outside (Table 3.4). Therefore, we were able to investigate the effects of mutating residues located within CYP153A6 as well as on the surface of the protein. In addition, the 3DM Bio-Prodict database was used to carefully select which of the internally located methionine and cysteine residues should be substituted for mutation.

3.2.1 3DM Bio-prodict results and Yasara

The graphs generated by 3DM demonstrated the amino acid distribution at the cysteine and methionine positions that were selected for mutation (Figure A1). From the graphs, we were able to determine which amino acid residues had the highest occurrence in nature out of the select cytochrome P450 proteins that were found to be similar to CYP153A6. These differences in amino acid residues at the same position indicate how natural selection has caused these proteins to evolve over time. The only amino acid that was 100% conserved was methionine at position 59 (306)[5]indicating that this amino acid was most likely located in the active site of the protein.

Yasara was used to map the positions cysteine and methionine residues that were to be substituted with more oxidatively stable amino acids. The amino acid distributions were then used to select which amino acid would best for substitution based on their characteristics (polar/non-polar, big/small, charged/uncharged) relative to the position they would occupy in the protein (surface exposed/internal).

3.2.1 Site directed mutagenesis to construct CYP153A6 mutants

3.2.1.1 Quickchange PCR amplification

To create the additional 14 CYP153A6s mutants site-directed mutagenesis was performed using the Quickchange (Stratagene) PCR method. Plasmid DNA of the wild type CYP153A6 was used as template in the PCR reactions and mutations were inserted into the DNA sequence by the use of mutagenic primers. Twelve of the fourteen PCR reactions using the standard Quickchange PCR method successfully amplified

[5] This corresponds to position 306 in the homology model of CYP153A6.

the plasmid. Reactions for mutations C52V (lane 1) and M129F (lane 8) (Figure 3.5) did not yield any product, while DNA sequencing revealed that the primers used to amplify mutations M101S (lane 6), M105R (lane 7) and M164T (lane 9) had been inserted into the plasmid therefore new plasmids had to be amplified for these mutants.

Table 3.4 Twenty five mutants of CYP153A6 that were evaluated for improved operational stability

CYP153A6 mutants	
Mutants designed and previously constructed	**Mutants designed and constructed in this study**
pET28: CYP153A6_operon_M4G	pET28: CYP153A6_operon_C52V
pET28: CYP153A6_operon_M18I	pET28: CYP153A6_operon_C208A
pET28: CYP153A6_operon_M70V	pET28: CYP153A6_operon_M56G
pET28: CYP153A6_operon_M105L	pET28: CYP153A6_operon_M91L
pET28: CYP153A6_operon_M129L	pET28: CYP153A6_operon_M101N
pET28: CYP153A6_operon_M241L	pET28: CYP153A6_operon_M101S
pET28: CYP153A6_operon_M292A	pET28: CYP153A6_operon_M105R
pET28: CYP153A6_operon_M292F	pET28: CYP153A6_operon_M129F
pET28: CYP153A6_operon_M265I	pET28: CYP153A6_operon_M164T
pET28: CYP153A6_operon_M202R_M206L	pET28: CYP153A6_operon_M202R
pET28: CYP153A6_operon_M231L_M232L	pET28: CYP153A6_operon_M206V
	pET28: CYP153A6_operon_M241D
	pET28: CYP153A6_operon_M306G
	pET28: CYP153A6_operon_M329L

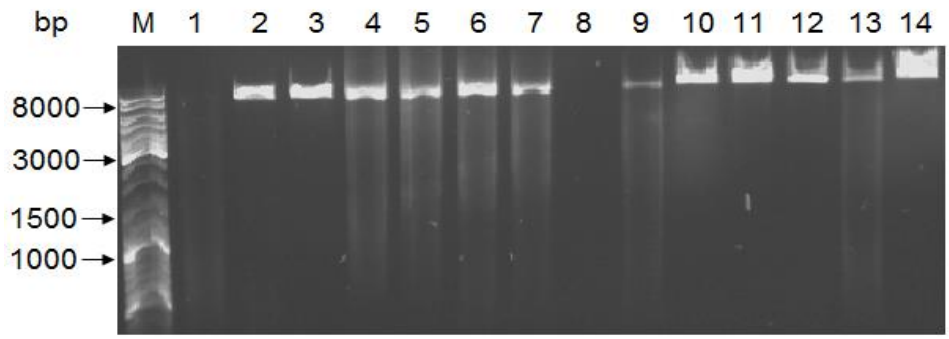

Figure 3.5 Agarose gel electrophoresis of CYP153A6 mutants which were amplified using the Quickchange (Stratagene) PCR method. C52V (lane 1), C208A (lane 2), M56G (lane 3), M91L (lane 4), M101N (lane 5), M101S (lane 6), M105R (lane 7), M129F (lane 8), M164T (lane 9), M202R (lane 10), M206V (lane 11), M241D (lane 12), M306G (lane 13) and M329L (lane 14). M=GeneRuler™ DNA ladder mix (Thermo Scientific).

These three mutations together with C52V and M129F were PCR amplified again but the number of cycles used for the thermal cycling conditions (section 3.1.2.2) was reduced to 16 cycles and the final elongation time was increased to 5 min (Figure 3.6). Adjusting the thermal cycling conditions as such worked for mutations M101S (lane 1), M105R (lane 2) and M129F (lane 5) because they were all PCR amplified and successfully transformed into *E. coli* Top10 cells. Again, mutant M164T (lane 4) was PCR amplified but could not be transformed into *E. coli* Top10 cells and mutant C52V (lane 3) was not amplified at all.

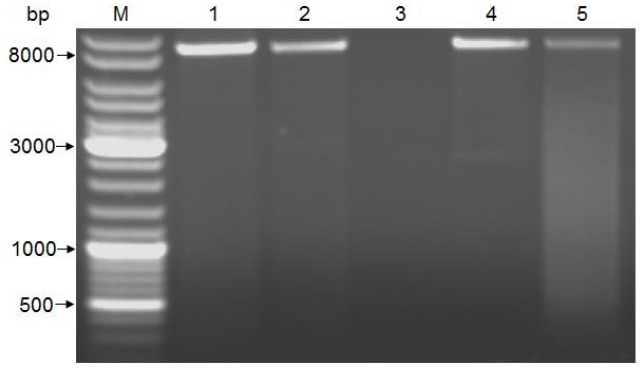

Figure 3.6 Agarose gel electrophoresis of CYP153A6 mutants which were amplified using the Quickchange (Statagene) PCR method. M101S (lane 1), M105R (lane 2), C52V (lane 3), M164T (lane 4) and M129F (lane 5). M=GeneRuler™ DNA ladder mix (Thermo Scientific).

At this point, plasmids for 12 out of 14 mutants were amplified and transformed and mutations were confirmed through DNA sequencing.

Mutations C52V and M164T were subsequently obtained using the Megaprimer PCR method. Primers for both mutations were combined, that is, the forward primer for mutation C52V was combined with the reverse primer for mutation M164T and vice versa resulting in both mutations inserted into the DNA sequence simultaneously (Figure 3.7). The two smaller bands (~360bp (lane 1)) and ~600bp (lane 2)) observed on the agarose gel are the megaprimers that were generated from the mutagenic primers in the first stage of the PCR reaction. In the second PCR stage, these megaprimers were extended using the wild type CYP153A6 plasmid DNA template to generate the two larger bands (~9000bp (lane 1) and ~8000bp (lane 2)) observed on the agarose gel which are the complete DNA sequences containing both mutations. Because the PCR product observed in lane 2 contained a lot of non-specific binding, only the DNA from the PCR product in lane 1 was purified and digested with restriction enzyme Cfr241 (SacII) to separate the two mutations as described previously in section 3.1.2.4. These last two mutants were again confirmed using DNA sequencing.

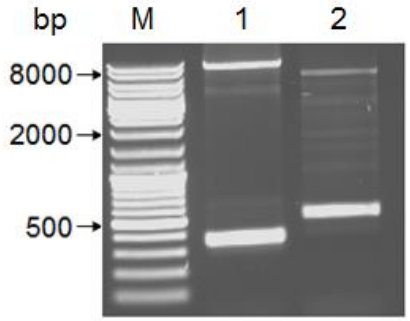

Figure 3.7 Agarose gel electrophoresis of CYP153A6 mutants which were amplified using the Megaprimer PCR method. Lane 1 shows PCR products of the PCR reaction containing the C52V_Forward and M164T_Reverse primer combination. Lane 2 shows the PCR products of the PCR reaction containing the C52V_Reverse and M164T_Forward primer combination. The two smaller bands (~360bp (lane 1)) and ~600bp (lane 2)) are the megaprimers that were generated in each reaction and the two larger bands (~9000bp (lane 1) and ~8000bp (lane 2)) are the complete DNA sequences containing both mutations.

3.2.2 Analysis of expression of CYP153A6 mutants

The expression of CYP153A6 (WT) and all 25 mutants was carried out in ZYP5052 auto-induction media (Table A1), supplemented with 30 $\mu g.mL^{-1}$kanamycin. To aid the expression of the P450 proteins, 1 mM δ-aminolevulinic acid hydrochloride (heme precursor) and 50 µM $FeCl_3 \cdot 6H_2O$ (to compensate for the lack of iron in *E. coli*) was also added to the auto-induction media. Furthermore, protein expression was carried out for 48 h on an orbital shaker (200 rpm) at a relatively low temperature of 20°C to ensure proper folding of P450 proteins during expression.

After the proteins had been expressed, the cells were harvested and resuspended in 200 mM Tris-HCl buffer (pH 8) containing 100 mM glycerol. For whole cell experiments 1 g of cells (wet cell weight) was resuspended in 10 mL of buffer therefore, the initial biomass concentration was 0.100 g_{WCW} mL^{-1}. For experiments using CFE, 1 g of cells (wet cell weight) was resuspended in 2 mL of buffer before disrupting the cells therefore, the initial biomass concentration was 0.500 g_{WCW} mL^{-1}.

3.2.2.1 Spectroscopic quantification of P450 content

The P450 content of the CYP153A6 mutants was quantified by performing CO-difference spectra using both whole cells and cell free extracts. The CO-difference spectrum immediately indicated how the various mutations had affected the expression of CYP153A6. An example of this is illustrated in Figure 3.8 which shows a comparison between mutants where protein expression was not affected siginificantly (Figure 3.6 (a)) and mutants where only a soret band at 420 nm was observed, meaning that these proteins had folded incorrectly and were therefore catalytically inactive (Figure 3.8 (b)). Similar results were also observed for CO-difference spectra using CFE.

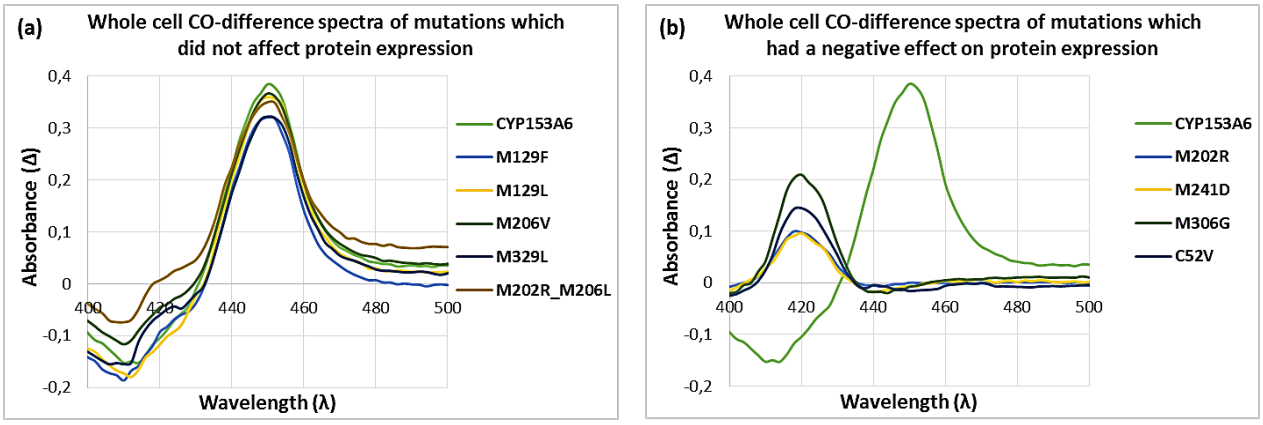

Figure 3.8 A comparison of whole cell CO-difference spectra between mutants that did not a significant effect on the expression of CYP153A6 (a) and mutants that had a negative effect on the expression of CYP153A6 (b).

The initial P450 concentrations calculated for the CYP153A6 mutants in whole cells (initial biomass concentration, 0.100 g_{WCW} mL^{-1}) varied between the mutants and between experiments (Figure 3.9). The results in Figure 3.9 were obtained by calculating the average P450 concentration of two independent experiments carried out in duplicate. Mutants M91L, M129F, M129L, M206V and M329L had similar expression to the wild type CYP153A6 indicating that substituting the amino acids at these specific positions did not significantly affect protein expression. The rest of the mutants all displayed lower expression levels compared to the wild type. Specifically in the case of mutants M101N, M105R, M164T,

M292F and C208A, P450 concentrations in whole cells were always less than 6 µM. In addition, no properly folded CYP[6] could be detected for mutations M202R, M241D, M306G and C52V indicating that substituting the amino acids at these specific positions with the selected amino acids .

The initial P450 concentrations were also calculated for CYP153A6 mutants in CFE (initial biomass concentration, 0.500 g_{WCW} mL^{-1}) as demonstrated in Figure 3.8. The results shown in Figure 3.10 are the average P450 concentrations calculated from triplicate measurements of CO-difference spectra for two independent experiments designated as Round 1 and Round 2. Again, we observed that there was variation in the expression levels between mutants and experiments. Since mutants M202R, M241D, M306G and C52V displayed no evidence of properly folded CYP in whole cells it was expected that this would be same results in CFE. CO-difference spectra also revealed that once the cells had been broken to obtain the CFE, CYP recovery became a challenge especially for mutants which demonstrated low expression levels in whole cells such as mutants M101N and M105R. However, despite high expression levels displayed by mutant M329L in whole cells, the CYP recovery for this mutant in CFE was very poor. This was also the case for mutants M129F, M129L and M206V although the CYP recovery for these mutants was not as poor as with mutant M329L. The rest of the mutants had similar comparative levels as was observed for whole cells. Unfortunately, mutants M56G, M202R_M206L and M231_M232L were only evaluated once in CFE.

[6]CYPs which displayed a soret band at 420 nm in their CO-difference spectra were marked with an asteriks (*) on the particular graph in question to indicate that properly folded CYP was not detected.

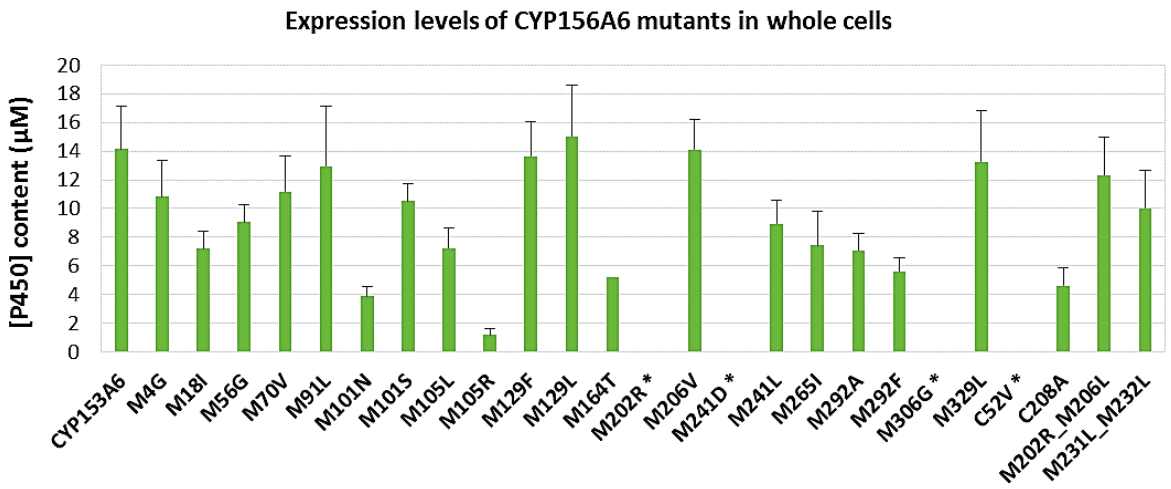

Figure 3.9 Comparison of the expression levels of CYP153A6 mutants in whole cells. Mutants marked with an asterisk displayed no evidence of properly folded protein, after using CO-difference spectra to determine the P450 concentration. Averages and standard deviations are for two independent experiments with measurements done in duplicate.

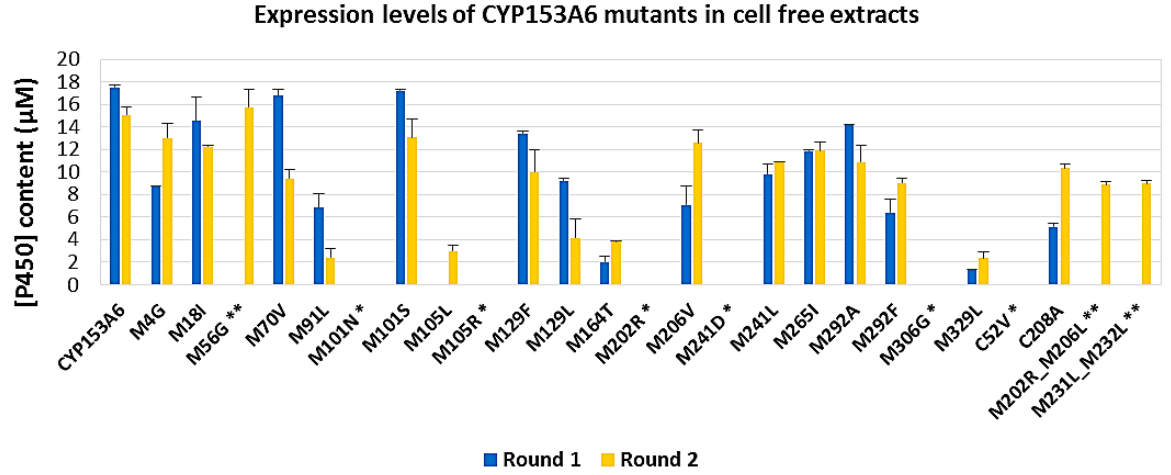

Figure 3.10 Comparison of the expression levels of CYP153A6 mutants in cell free extracts. Mutants marked with a single asterisk (*) displayed no evidence of properly folded CYP being present after analysing the CYPs using CO-difference spectra. Averages and standard deviations are for CO-difference readings carried out in triplicate for each mutant in two independent experiments (Round 1 and Round 2). Mutants marked with double asterisk (**) were only evaluated once.

3.2.2.2 SDS-PAGE analysis of CYP153A6 mutants

Expression of proteins was also evaluated on 8% SDS-PAGE resolving gels using the cell free extracts (soluble fraction). Bands corresponding to the expected molcular weight of 48 kDa were observed for the wild type CYP153A6 and most of the CYP153A6 mutants. Expression levels of the mutants observed on SDS-PAGE gels corresponded to what was observed for the concentrations calculated using CO-difference

spectra (Figure 3.11), therefore mutants C52V (lane 1(b)), M105R (lane 3 (d)), M202R (lane 6 (e)) and M241D (lane 6 (d)) were not observed on the SDS-PAGE gels. Acceptable expression levels were observed for mutants M129L (lane 2 (a)), M129F (lane 3 (a)), M101S (lane 2 (b)), M261I (lane 4 (c)), M206V (lane 5 (d)) and M18I (lane 3 (e)) corresponding to the CO-difference spectra results. As in the CO-difference spectra a significant difference was observed in expression between mutants M241D (lane 6 (d)) and M241L (lane 7 (d)) even though both mutations occur at the same position. Subtituting methioine with aspartic acid and leucine had opposite effects on protein expression, indicating the importance of amino acid selection when creating mutants. Interestingly, a band corresponding to the expected molcular weight (48 kDa) was observed for mutant M101N (lane 4 (e)) despite the fact that no CYP was detected in cell free extract using CO-difference spectra.

3.2.3 Biotransformations

The activity of the CYP153A6 mutants were evaluated by performing biotransformations using whole cells and cell free extracts. Whole cell biotransformations were performed as preliminary experiments to confirm whether these mutants still had activity towards *n*-octane and if the regioselectivity was affected.

3.2.3.1 Hydroxylation of n-octane by CYP153A6 mutants using whole cells

When carrying out experiments using soluble proteins that function as a three component system such as CYP153A6, ferredoxin reductase and ferredoxin, whole cell biotransformations are used most often because the cell walls keep all the proteins in a confined space so that electron transfer can occur more easily between the proteins. Since CYP153A6 is also NADH dependent, whole cell biotransformations are also more convenient because the cellular metabolism of *E. coli* is able to regenerate NADH when additional glucose and glycerol are added to the reactions. Experiments using whole cells were carried out with the CYP153A6 mutants to evaluate their ability to hydroxylate *n*-octane before conducting experiments using cell free extracts.

The samples were extracted and submitted for GC analysis. All reactions for each mutant were carried out in duplicate.

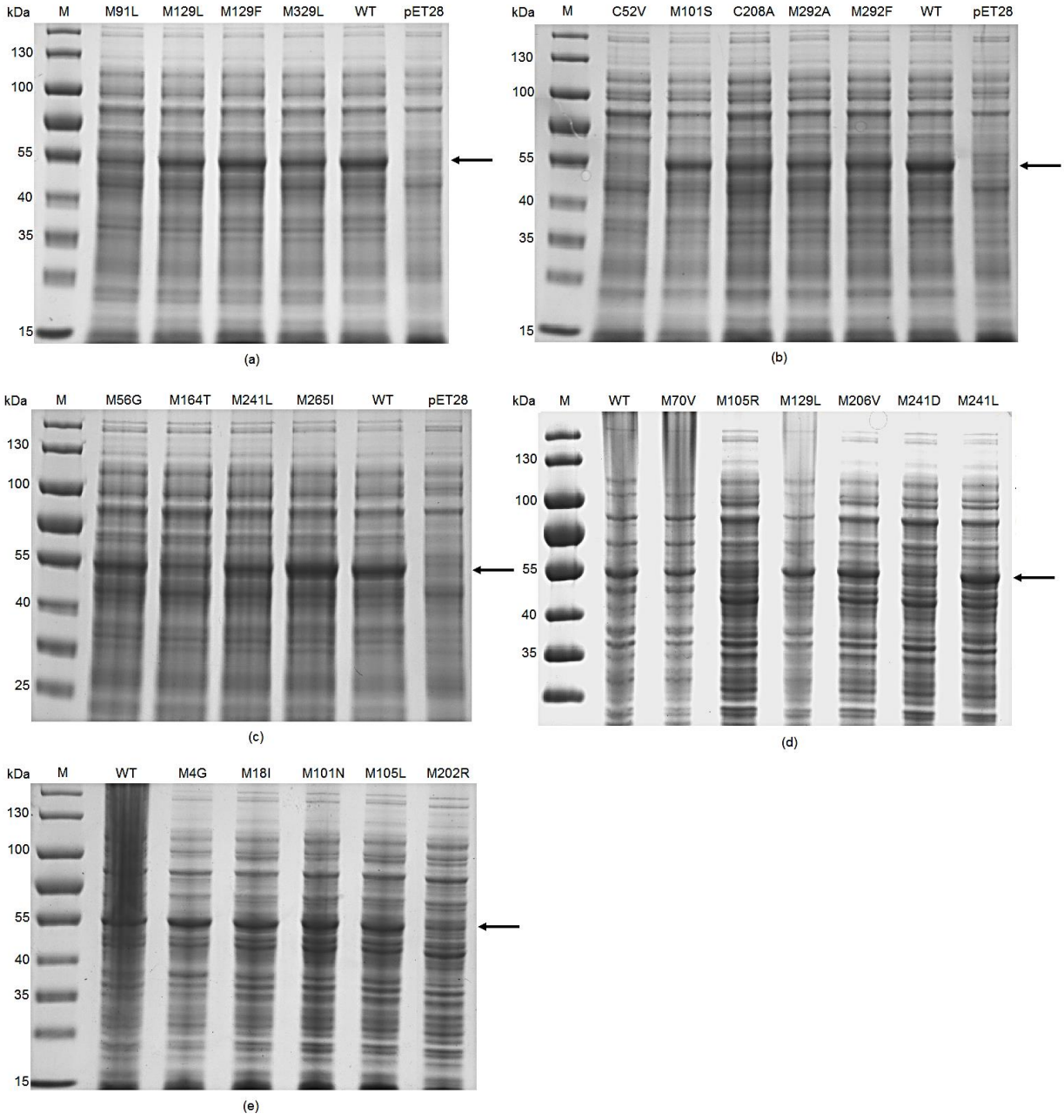

Figure 3.11 SDS-PAGE analysis of cell free extracts (soluble fraction) of *E. coli* BL21-Gold (DE3) stains, transformed with pET-28b (+) plasmids for the expression of CYP153A6 operon (WT) and the CYP153A6 mutants. M = PageRuler™ Prestained Protein Ladder (Thermo Scientific).

At the time when whole cell biotransformation were performed, mutants C52V and M164T [marked with an asterisks (*) on the graph] were still in the process of being constructed therefore no whole cell activity data was collected for these two mutants at this point. The data shown in Figure 3.12 is the average 1-

octanol (mM) produced for two independent experiments where the reactions were carried out in duplicate.

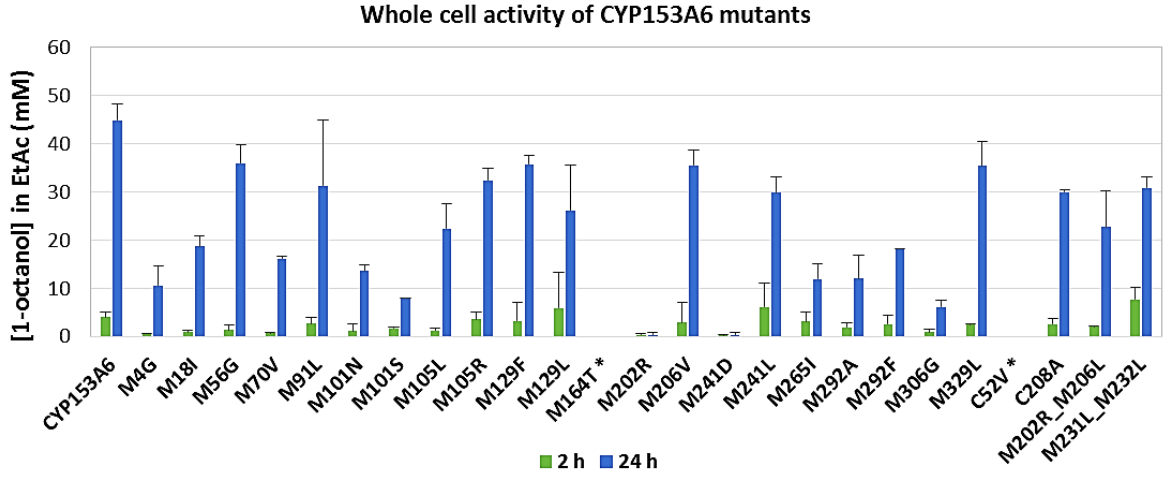

Figure 3.12 Production of 1-octanol by CYP153A6 mutants after conducting whole cell biotransformations for 2 and 24 h Error bars are for two independent experiments with reactions done in triplicate.

Whole cell biotransformations were carried out using a 1 mL reaction volume which contained 0.5 mL buffer [200 mM Tris-HCl (pH 8) buffer containing 100 mM glycerol and glucose] and 0.5 mL of a 0.1 g_{wcw} mL^{-1} cell suspension therefore the final biomass concentration in the BRM was 0.05 g_{wcw} mL^{-1}. Reactions were started by adding 250 µL of *n*-octane to each reaction vial before placing the vials on an orbital shaker and incubating them at 20°C. Each reaction was stopped after 2 and 24 h after which

Mutants M56G, M91L, M105R, M129F, M206V, M329L, M232L_M232L and C208A all produced more than 30 mM of 1-octanol after 24h. However, the wild type CYP153A6 still had the best activity producing more than 40 mM of 1-octanol. The activities of mutants M105R and C208A were impressive given that their CYP concentrations in the original cell suspensions were less than 2 µM and 6 µM respectively while that of the others were between 7 to 15 µM. On the other hand, despite having a relatively high CYP concentration, M101S displayed poor activity producing less than 10 mM of 1-octanol after 24 h. Although no catalytically active P450 was detected for M306G using CO-difference spectra, production of 1-octanol by this mutant was detected with GC analysis. This however, was not the case for M202R and M241D both of which displayed no activity towards *n*-octane.

Variation in activity was observed for all the mutants, but M91L, M129L and M202R_M206L displayed the most significant differences in activity between the two experiments. For this reason, it was decided to

continue biotransformtions using CFE because CFE reactions would also allow adjustment of CYP concentrations to the same level and to challenge the CYPs by *in situ* generation of H_2O_2.

3.2.3.2 Hydroxylation of n-octane by CYP153A6 mutants using cell free extracts

When biotransformations using whole cells were performed the cellular metabolism of glycerol and glucose was relied on for cofactor regeneration. However, in the case of CFE, conversion of *n*-octane to 1-octanol appeared to be limited when CYP153A6 interacted solely with co-expressed FdR and Fdx therefore additional FdR/Fdx expressed as a partial operon was added to CFE reactions.

Chapter 2 describes all the various parameters that were evaluated in order to determine which conditions would not only favour the conversion of *n*-octane by the CYP153A6 mutants but challenge them as well. From these experiments, a set of conditions was chosen to use as the starting point for evaluating the CYP153A6 mutants for improved operational stability.

Cell free extract (CFE) biotransformations were performed using a 1.2 mL reaction volume which contained 0.7 mL buffer [300 mM Tris-HCl (pH 8) buffer containing 100 mM glycerol and glucose], 0.25 mL CFE from a 0.5 g_{WCW} mL^{-1} cell suspension of *E. coli* expressing only FdR and Fdx (FdR/Fdx) and 0.25 mL CFE from a mixture of cell suspensions of *E. coli* expressing the CYP together with FdR and Fdx and *E. coli* transformed with empty vector. These two cell suspensions were mixed to obtain a final CYP concentration of 1.5 µM in the BRM. Additional NADH (1 mM) was also added to each reaction as well as glucose oxidase (GOx) which oxidises glucose to glucono-lactone and produces H_2O_2 simultaneously (Figure 2.4). Glucose oxidase was added to reactions to challenge the oxidative stability of the mutants in the presence of H_2O_2. Again, reactions were started by adding 250 µL of *n*-octane to each reaction vial. Biotransformations were carried out on an orbital shaker (200 rpm) at 30°C to evaluate relative thermostability in addition to oxidative stability. All reactions for each mutant were carried out in triplicate and each reaction was stopped after 2, 8 and 20 h after which the samples were extracted and submitted for GC analysis.

The mutants and the wild type (WT) were evaluated over two rounds of experiments, termed Round 1 and Round 2. Round 1 consisted of four experiments with the mutants divided randomly between the experiments, in addition, 800 U/L GOx was added to all reactions in Round 1 (Figure A3). Round 2, consisted of five experiments; the concentration of GOx added to all reactions was doubled to 1600 U/L (Figure A4).

After reviewing the data collected for the two rounds of experiments (Figures A3 and A4), it was clear that there was a lot of variation in activity of each of the mutants and the wild type. Therefore, we had to find a way to normalise the data so that the results could be more comparable. We began by comparing the activity data of the wild type from both rounds of experiments (Figure 3.13). Firstly, we observed that there was more variation in the activity data for Round 1 because the range of 1-octanol produced was between 0.5 mM and 10 mM (Figure 3.13 (a)) whereas the activity data for Round two was between a range of 2.5 mM and 6 mM (Figure 3.13 (b)) therefore the data from Round two was considered more comparable. The second problem was that the wild type was unfortunately not tested in the second run of experiments for Round 1 (Figure A3: Round 1, experiment 2) therefore there were only three sets of data to compare for Round 1 regarding the wild type activity whereas for Round 2 there was five sets of data that could be used for comparison. The data collected for the wild type in round 2 was therefore considered to be more reliable for comparison.

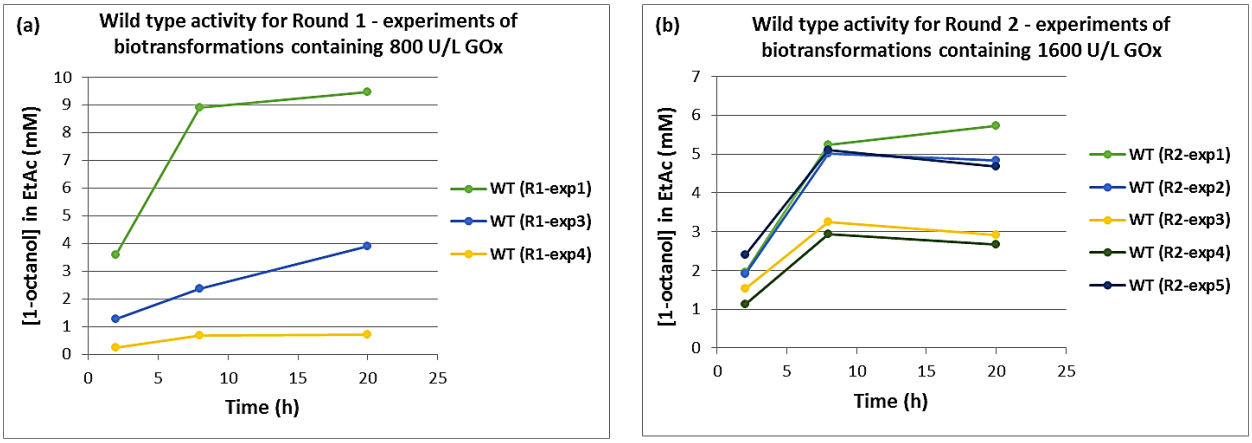

Figure 3.13 Wild type activity data collected for three and five independent experiments carried out in Round 1 (a) and Round 2 (b) of cell free extract biotransformations, respectively.

We then had to determine how we were going to normalise the data collected for the mutants so that it could be more comparable to evaluate the performance of the mutants. If we refer to the activity data for mutant M70V which was tested in three independent experiments (twice in Round 1 and once in round 2) we clould clearly observe the variation in activity for this mutant (Figure 3.14 (a)). However, when we calcualted the activity of M70V as percentage of the wild type activity for corresponding experiments, we observed that the percentage activity data was similar for each of the three experiments (Figure 3.14 (b)). Therefore, we descided to report the activity data as a percentage of wild type activity however, we only used the data collected in Round 2 because these results were more reliable and because the wild type

was not tested in experiment 2 (Round 1) therefore the perecentage activity could not be calculated for these experiments (Figure A3: Round 1 experiment 2)).

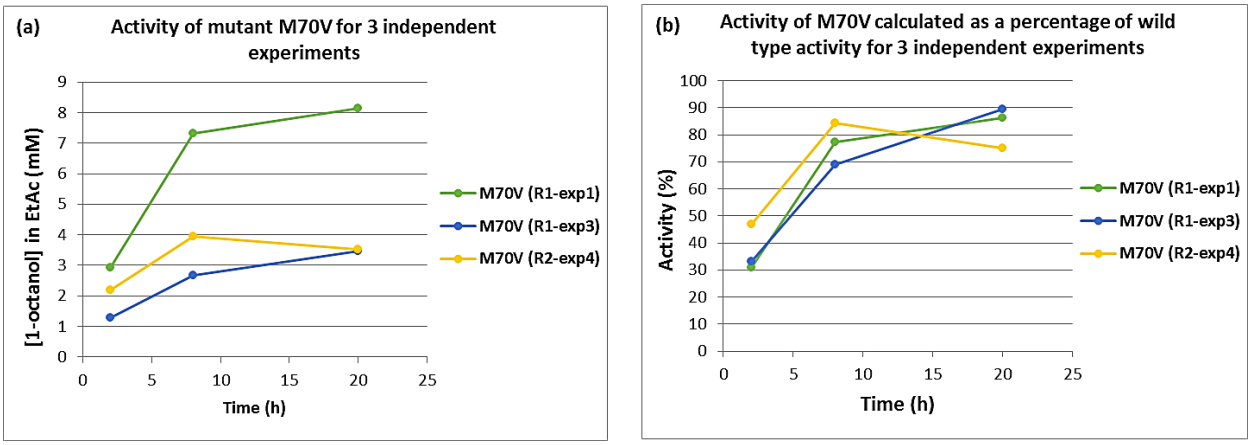

Figure 3.14 Activity data collected for mutant M70V for three independent experiments (a). Activity of M70V calculated as a percentage of the wild type activity in corresponding experiments (b).

3.2.4 Evaluation of the CYP153A6 mutants based on protein expression and enzyme activity

The results of mutant expression levels in whole cells and cell free extracts were combined into one graph (Figure 3.15 (a)). It should be noted that the biomass concentration of whole cell suspensions were 0.1 g_{wcw} mL^{-1} whereas the biomass concentration of cell free extracts was 0.5 g_{wcw} mL^{-1}) In order to determine which mutants had either acceptable[7] expression or poor expression it was decided to use 7.2 µM as a "cut-off" point (indicated by a red line on the graph) to rate expression levels as acceptable or not. This value was decided upon because in order to obtain a final P450 concentration of 1.5 µM in the BRM, the initial P450 concentration had to be at least 7.2 µM. Mutants with higher initial P450 concentrations (for cell free extracts) were diluted with an equally concentrated (0.5 g_{wcw} mL^{-1}) cell suspension of *E. coli* expressing an empty pET-28b (+) vector. Therefore, if the expression level of a mutant was above 7.2 µM, the mutant was considered to have acceptable expression but if expression levels were below 7.2 µM the mutant was considered to have poor expression.

[7] "Acceptable expression" means that the mutation did not have a negative effect on protein expression however, expression levels were not as high as the wild type CYP153A6.

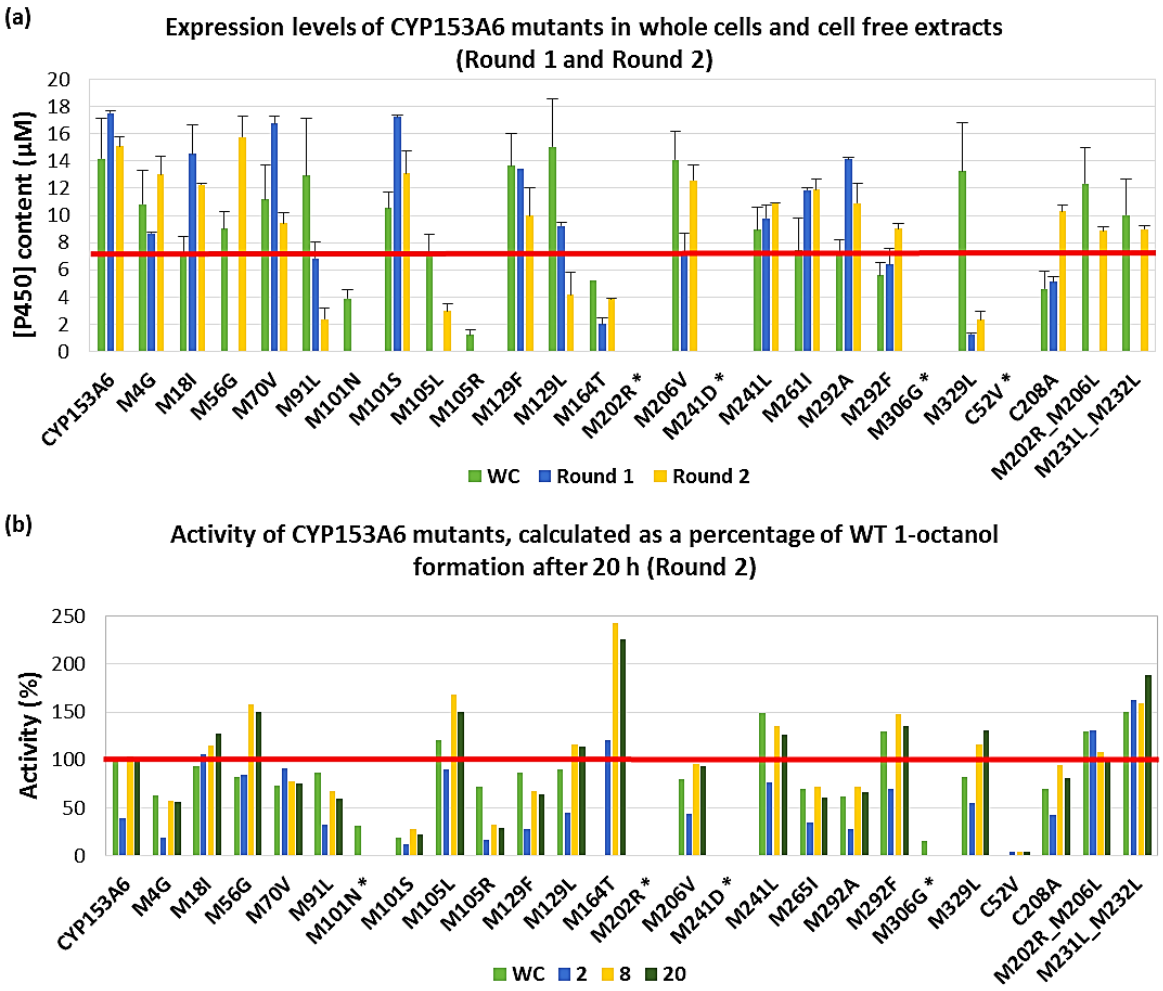

Figure 3.15 A summary of the P450 levels of the CYP153A6 mutants in whole cells (WC) (biomass concentration 0.1 g_{wcw} mL^{-1})and cell free extracts (Round 1 and Round 2: biomass concentration 0.5 g_{wcw} mL^{-1}). A "cut off" point was implemented at 7.2 µM in order to determine which mutants had acceptable expression or poor expression; no catalytically active P450 was detected for mutants marked with asterisk (a). The activity of the CYP153A6 mutants was calculated as a percentage of WT 1-octanol formation after 20 h. Therefore, WT activity (100%) was implemented as the "cut-off" point in order to determine which mutants had improved activity, unaffected activity, poor activity or no activity; no 1-octanol was detected using GC analysis for mutants marked with an asterisk (b).

In order to make the data collected for activity more comparable, the activity of the mutants were calculated as a percentage of WT 1-octanol (product) formation after 20 h using the results obtained for Round 2 as previously stated, these results are shown in Figure 3.15 (b). The results from round 2 were chosen because these results were deemed more reliable than results obtained for Round 1 for the reasons mentioned previously and a higher concentration (1600 U/L) of GOx was used in Round 2 biotransformations therefore, the mutants were under higher oxidative stress than in Round 1. The

thermostability of the mutants was also challenged because biotransformations were carried out at 30°C. Because the activity was calculated as a percentage of WT activity this meant the WT had 100% activity, which was implemented as the "cut-off" point (indicated by a red line on the graph) to determine which mutants had improved activity, unaffected activity, poor activity or no activity. "Improved activity" meant that the activity of the mutant was significantly higher (>100%) than the activity of the WT, "unaffected activity" meant that the activity of the mutant was similar to or within close range (75-100%) of the activity of the WT and "poor activity' meant that the activity of the mutant was significantly lower (< 75%) of the activity of the WT.

3.2.4.1 CYP153A6 mutant groups

After analysing the results summarised in Figure 3.15, the mutants were divided into five groups based on protein expression and percentage activity (Figure 3.16). These five groups include:

1. Poor expression and no/poor activity
2. Poor expression and unaffected/improved activity
3. Acceptable expression and poor activity
4. Acceptable expression and unaffected activity
5. Acceptable expression and improved activity

Structural homology models of CYP153A6 (Figures 3.16 and 3.17) were created for each mutant group to highlight the positions of each mutant in the protein structure and the scatter plot indicates the expression and activity of each mutant comparative to the other mutants as well as the wild type CYP153A6. The following sub-sections zoom in on the mutant groups and discuss how each individual mutation could have affected protein expression and activity relative to the properties of the substituent amino acid and its position in the protein structure.

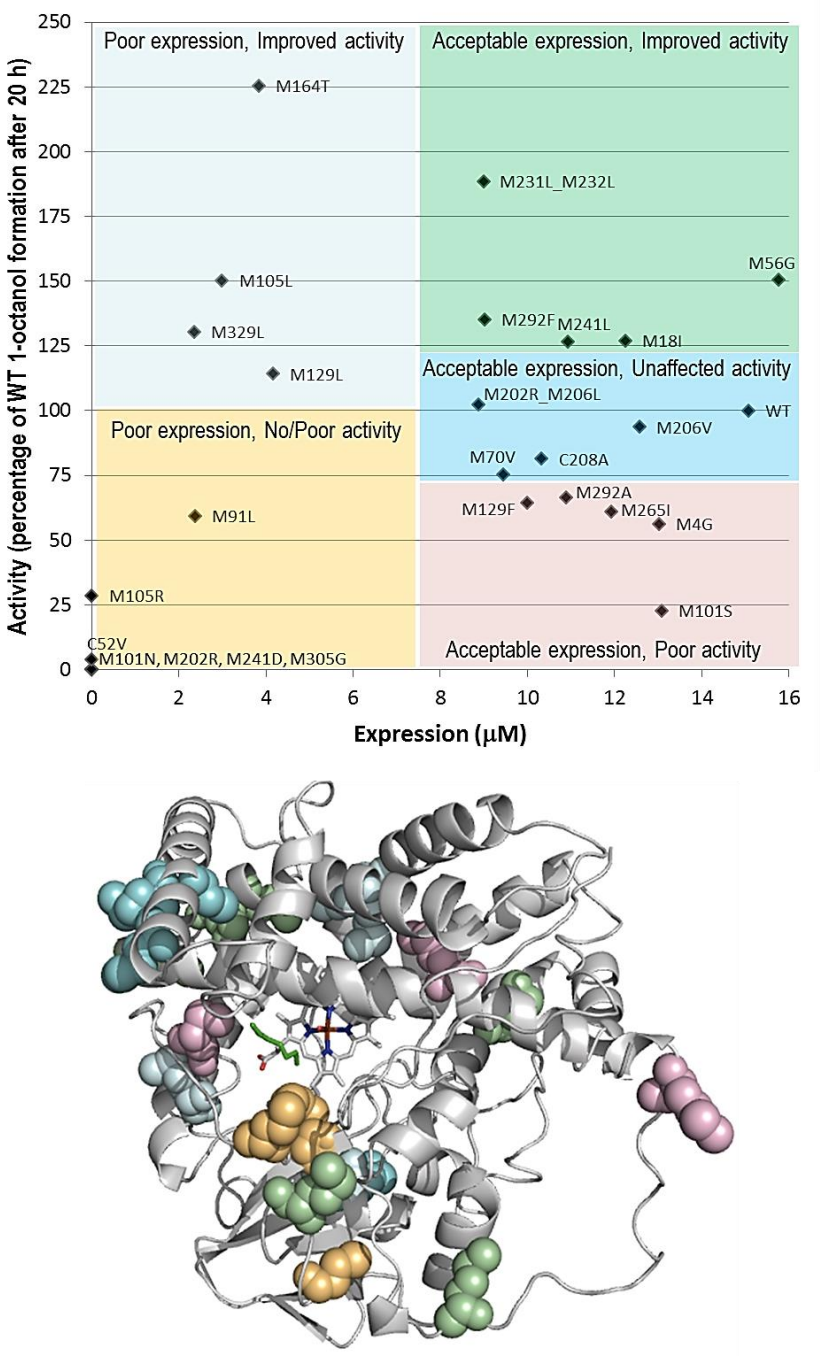

Figure 3.16 Structural homology model of CYP153A6 indicating the position of the mutants, corresponding to five groups based on protein expression and percentage activity relative to the wild type CYP153A6. Mutants M91L, M101N, M105R, M202R, M241L, M306G and C52V had poor expression and no/poor activity (orange). Mutants M105L, M129L, M164T and M329L had poor expression and acceptable activity (pale blue). Mutants M4G, M101S, M129F, M292A and M265I had acceptable expression and poor activity (pale pink). Mutants M70V, M206V, M202R_M206L and C208A had acceptable expression and activity similar to that of the WT (cyan). Mutants M18I, M56G, M292F, M241L and M231L_M232L had acceptable expression and improved activity (green). The active site (heme propionyl group) is illustrated at the centre of the protein, with the substrate (octane) is illustrated as a green "stick" figure. Homology model generated by PyMol.

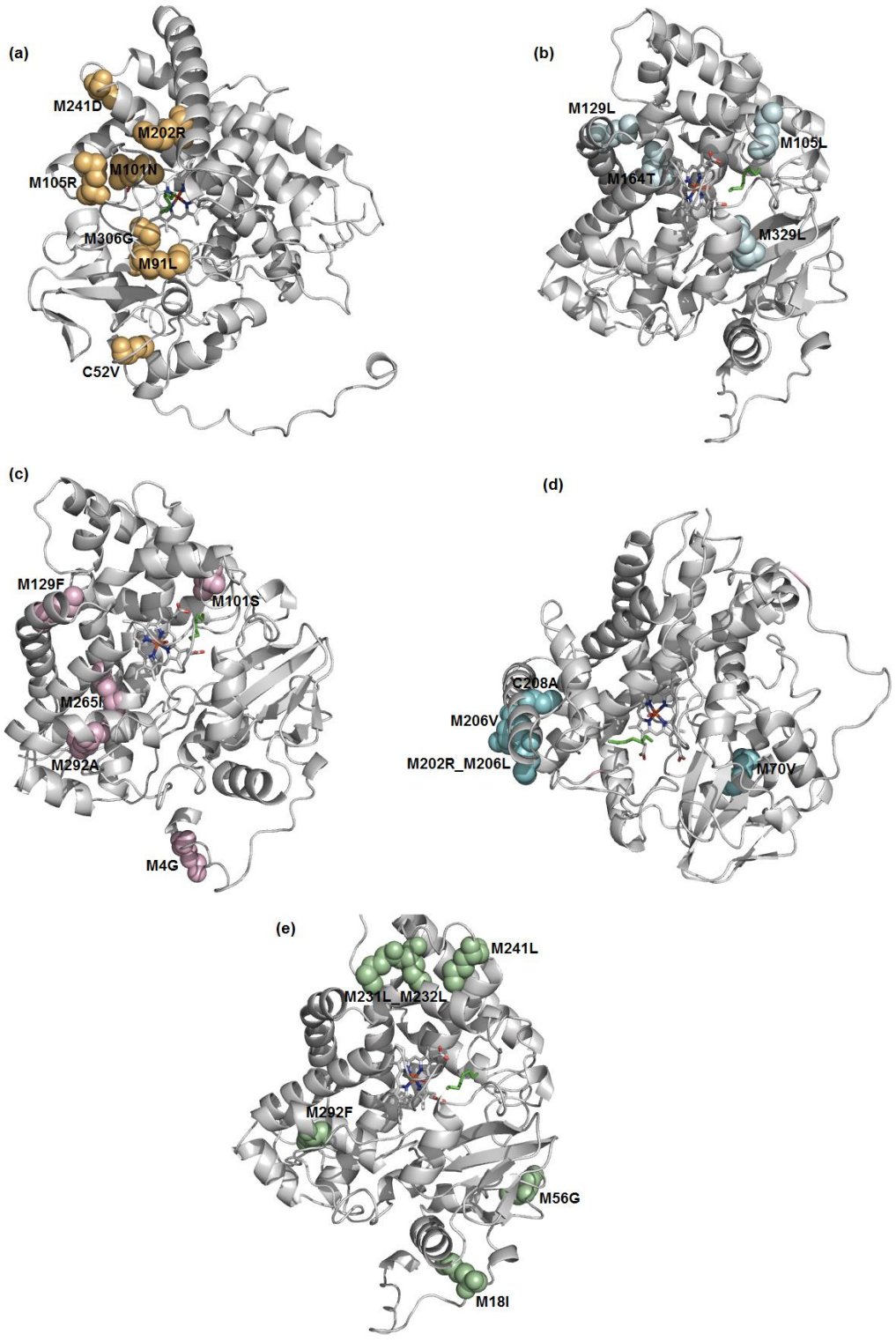

Figure 3.17 Structural homology models of CYP153A6 generated for each of the five groups indicating the position of the mutants: Poor expression and poor/ no activity (a); Poor expression and unaffected or improved activity (b); Acceptable expression and poor activity (c) Acceptable expression and unaffected activity (d) and Acceptable expression and improved activity (e).

3.2.4.1.1 Mutants that displayed poor protein expression and no/poor activity

The seven mutants that were allocated to this group namely, C52V, M91L, M101N, M105R, M202R, M241D and M306G, had either no expression or P450 levels lower than 2.5 µM in CFEs and activity in Round 2 was less than 60% of WT activity in the same experiment (Figures 3.16 and 3.17 (a)).

M101N, M202R, M241D and M306G had the worst results of all the mutants because no catalytically active CYP was detected for any of these mutants in CFE and GC analysis detected no production of 1-octanol for any of these mutants after 20 h of biotransformations using cell free extracts.

The homology model of CYP153A6 generated here and the homology model created by Funhoff and co-workers, clearly illustrates that M101 and M306 which were substituted with asparagine and glycine respectively, are located in the active site (Funhoff *et al*., 2006). This could be the reason why mutating these residues gave a negative result because amino acids in the active site are most often conserved. Structural alignments performed by the 3DM database between CYP153A6, CYP153A7 and other P450s with similar amino acid identity to CYP153A6, determined that M306 (3DM position 283[8]) was 100% conserved (Figure A1 (j)). However, glycine was chosen as the substituent amino acid at position 306 because given the close proximity to the heme propionyl group, glycine is a small and particularly flexible amino acid because of its hydrogen side chain therefore it should be able to fit into parts of the protein that larger amino acids cannot. In retrospect, substituting an amino acid with a size and polarity closer to that of methionine might have given a better result. For example, Pham and co-workers created a triple mutant (I83H/M305Q/A77S) to improve enantioselective hydroxylation by P450pyr hydroxylase (CYP153A7); here they substituted Met305 with glutamine (Q), a polar amino acid which frequently plays a role in polar interactions in the active or binding sites of proteins (Barnes *et al*., 2003;Pham *et al*., 2012).

For position 101 (3DM position 55), serine (23.6%) and asparagine (5.4%) were the two amino acids that occur second and third most frequently after methionine, respectively (Figure A1 (c)) and both mutations were tested. We observed that substituting methionine, a non-polar residue with asparagine a polar residue, led to a decrease in protein expression in whole cells so that no catalytically active P450 was detected for mutant M101N using CO-difference spectra after disrupting the cells to obtain the CFE. In addition, asparagine is also a larger residue than methionine and based on its position in the structure we

[8] Position 283 corresponds to the consensus sequence generated by the 3DM database, 283 is therefore the 3DM position. A structural alignment indicated that position 283 corresponded to position 306 in the homology model of CYP153A6.

observed that this residue extended toward the active site which could have affected the substrates ability to bind These were the most probable reasons why there was no activity for this mutant in CFE, although the specific activity of this mutant in WCs was even higher than that of the WT (average TTN for M101N 4100 and for WT 3300 in the same two experiments).

M202R and M241D displayed no expression or activity in either whole cells or CFE. Arginine was selected firstly because it has the second highest occurrence in nature based on the 3DM structural alignment (Figure A1 (g)) and secondly because it is characterised as an amphipathic amino acid which means that its side chain contains both polar (hydrophilic) and non-polar (hydrophobic) regions. Based on the surface location of M202, arginine seemed to be a good substitution for methionine however this was not the case. As for M241D, the 3DM data revealed that asparagine had a 25.2% occurrence while aspartic acid only had a 14.2% occurrence (Figure A1 (i)). However, aspartic acid was selected based on the correlation mutation analysis (CMA) carried out with 3DM (Figure A1 (i)), because there is a 92 % correlation between mutations at 3DM positions 115 and 200 which corresponds to T162 and M241 in CYP153A6. According to this analysis, there is never an asparagine at 3DM position 200 if there is a threonine at position 115. The T-D combination however occurs in 3 % of these CYPs. When considering the fact that aspartic acid is a charged amino acid, which is usually found on the surface of proteins, position M241 was not an ideal location for aspartic acid because this residue faces towards the interior of the protein.

Although no protein expression was detected for C52V, GC analysis confirmed that 1-octanol was in fact produced by this mutant in CFE. The activity of C52V calculated as a percentage of the wild type activity was however less than 4%. The 3DM data determined that glutamine (Q) had the second highest occurrence in nature after cysteine (C) based on the structural alignment (Figure A1 (a)). Cysteine 52 is located near the surface but with the side-chain facing towards the interior of the protein. Although both cysteine and glutamine are polar amino acids, substituting cysteine with valine (V) - a non-polar amino acid - seemed more rational, especially since cysteine and valine are also more similar in size than cysteine and glutamine. In addition, successful substitutions between cysteine and valine to improve oxidative and thermal stability have been documented (Opperman & Reetz, 2010; Perry & Wetzel, 1987).

Selecting an amino acid to substitute M105 was particularly challenging because according to the 3DM data, 95.4% of amino acids at position 105 are methionine, 1.7% are alanine (A) and 1.3% arginine (R) (Figure A1 (d)). What made the choice even more challenging was that alanine and arginine are two opposite amino acids; alanine is a small non-polar amino acid and arginine is a large, polar amino acid that is positively charged. Arginine being amphipathic in nature (Barnes *et al.*, 2003) was selected to substitute

methionine at position 105 based on the location as it is situated just outside the binding site and although it's not directly on the surface of the protein, it is partially exposed to the solvent. However, this choice had a detrimental effect on protein expression and enzyme activity.

M91L was one of the mutants designed and constructed by Dr. Opperman. Substituting methionine with leucine in order to improve the oxidative stability as well as thermal stability of enzymes is common according to literature (Kim *et al*., 2001; Lin *et al*., 2006; Opperman & Reetz, 2010; Suemori & Iwakura, 2007). M91L performed much better in terms of protein expression and activity in whole cells than in CFE. According to the homology model, M91L is located on the border of the binding site and part of an unstructured loop (not included in the 3DM core) that forms part of the substrate access channel similar to M105R. It is partially exposed to the surrounding aqueous environment and given that leucine is a non-polar amino acid, this could have had a negative effect on the enzyme's solubility and activity.

Lastly, the total turnover number (TTN) was calculated for each mutant as a percentage of the wild type's TTN [WT TTN taken as 100% and indicated on the graph by a red line (Figure 3.18)]. Surprisingly, the TTN of M105R and C52V were 9-fold and 7-fold higher than the wild type respectively, but these results were considered negligible because their P450 concentrations were so low they cannot be seen in Figure 3.15 (a) and the GC results indicated minimal 1-octanol production for both mutants.

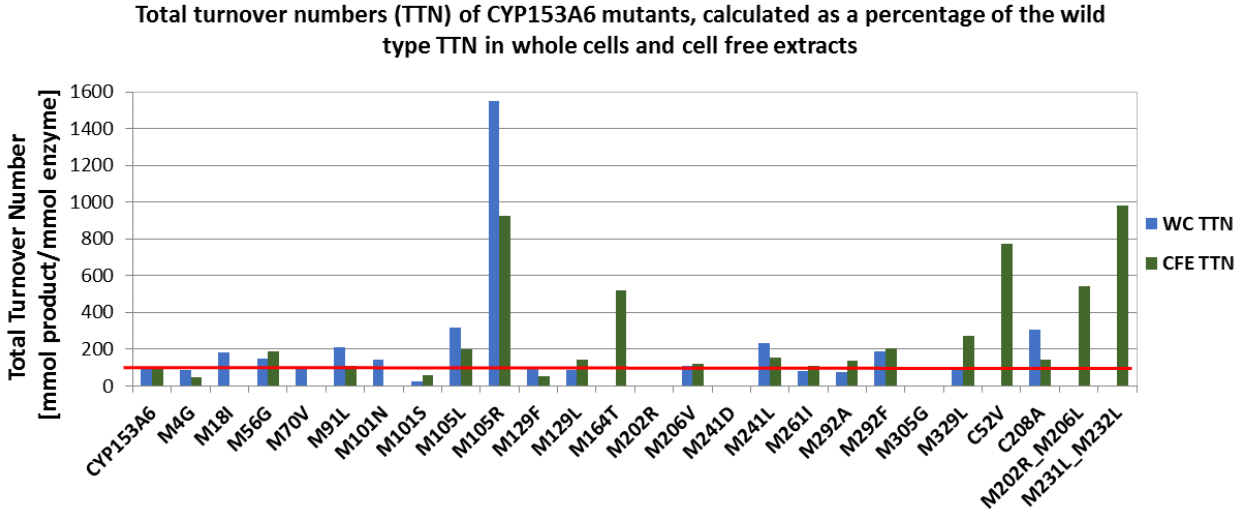

Figure 3.18 Total Turnover Numbers (TTN) of each CYP153A6 mutant, calculated as a percentage of the wild type CYP153A6 TTN (100%) in whole cells (WC) and cell free extracts (CFE).

3.4.1.2.2 Mutants that displayed poor protein expression and unaffected or improved activity

Although these four mutants – M105L, M129L, M164T and M329L – had expression levels of less than 4 µM, their activity was significantly higher than the wild type (Figures 3.16 and 3.17 (b)).

Mutants M105L and M129L were also of those designed and constructed by Dr. Opperman and although M105R and M105L are in the same position, substituting methionine with leucine resulted in this enzyme having an activity 50% and 122% higher than the wild type and M105R, respectively. However, the expression of M105L like M105R was poor in whole cells and CFE but despite these poor expression levels, M105L had a TTN that was 2-fold higher than the TTN of the wild type (Figure 3.18). M129L is buried in the protein but remote from the active site. Protein expression was not affected, in fact, expression levels for this mutant in whole cells was equal to that of the wild type but after disrupting the cells to obtain the CFE, CYP levels decreased significantly indicating that a lot of CYP was lost during the disruption and centrifugation process (see chapter 2, section 2.1.3). In addition, the activity of M129L improved on the wild type's activity by 14%.

Even though the 3DM data showed that only 7.7% of homologues P450s contained a leucine at position 329 (Figure A1 (k)), expression of M329L in whole cell was close to that of the wild type but as in the case of the M129L mutant, CYP recovery after disrupting the cells was very poor resulting in low levels in CFE. However, these low levels did not affect the enzyme's activity which improved on the wild type's activity by 30%. Furthermore, the TTN of M329L was almost 3-fold higher than the TTN of the wild type (Figure 3.18).

Poor expression levels were observed for M164T in whole cells and CFE. M164T is located on the inside of the protein and extends towards the heme propionyl group. According to the 3DM data, M164 had an 88% occurrence in nature among the selected P450s whereas threonine only had an 8.9% occurrence (Figure A1 (f)). Given that threonine is a polar amino acid and methionine is a non-polar amino acid, as well as the sidechain size difference, the 125% improvement on the wild type's activity of M164T were unexpected. Furthermore, the TTN of M164T was 5-fold higher than the TTN of the wild type (Figure 3.18).

Despite the significant improvement in specific activity for all these enzymes, it would not be practical to use them in preparative biocatalysis because of their poor expression levels. However, this proves that poor protein expression doesn't necessarily mean poor activity.

3.2.4.1.3 Mutants that displayed acceptable protein expression and poor activity

Contrary to what was observed in the previous group, M4G, M101S, M129F M265I and M292A displayed acceptable protein expression (expression levels above 7.2 µM) but they had poor activity (Figures 3.16 and 3.17 (c)).

As previously mentioned, serine was selected as a substituent for methionine 101 (3DM position 55) based on the 3DM data which determined that out of 239 sequences (structural based alignment) 23.6% contained a serine residue at this position (Figure A1 (c)). The homology model and studies conducted by Funhoff and co-workers confirmed that M101 resides in the binding site. M101S presented high levels of expression in cell free extracts but decreased enzyme activity was observed, suggesting that the mutation might have interfered with the substrate's ability to bind (serine is a hydrophilic amino acid). Tang and co-workers created a mutant of P450pyr hydroxylase (CYP153A7) to improve the enantioselective hydroxylation of N-benzyl pyrrolidine (Tang *et al.*, 2010). They accomplished this by substituting asparagine at position 100 (3DM position 55) to serine (N100S) and threonine at position 186 to isoleucine (T186I) to create a double mutant N100S/T186I. Therefore, it's possible that M101 can be used to investigate altering the substrate scope of CYP153A6.

In addition to the M129L mutation, phenylalanine was also chosen as a substituent residue for M129 although it showed only a 4.5% occurrence in nature among the selected P450s compared to the 38.1% occurrence of leucine (Figure A1 (e)). The expression of the M129F mutant was not only acceptable in whole cell but also in CFE. Although phenylalanine like methionine is non-polar and M129F was remote from the active site, activity was significantly reduced compared to the WT as well as the M129L mutant. M4G, M265I and M292A had acceptable expression in whole cells and CFE. No 3DM data was available for M4G because it is not in the core, however glycine was selected as the substituent for methionine. From the homology model, we observed that M4 was located on an unstructured "tail" that extended away from the folded protein. However, based on the general structure of cytochrome P450s and the structure of CYP153A7 (Pham *et al.*, 2012) we had reason to believe that this "tail" was incorrectly modelled from its actual position. Although it was difficult to determine the actual position of M4G, the homology model suggested that M4G was surface exposed and it is difficult to explain why a mutation at this position have such a negative impact on activity.

In addition to Leucine, researchers have also found Alanine and Isoleucine to be favourable substituents for methionine with regards to improving the oxidative stability of an enzyme (Johnsen *et al.,* 2000; Kim

et al., 2001; Opperman & Reetz, 2010). According to the 3DM data isoleucine and alanine were both possible choices as substituents of methionine at positions 265 (Figure A1 (m)) and 292 (Figure A1 (n)), respectively. Both these amino acids occurred within the folded core of the protein but despite M292A located further away from the active site than M265I the activity of both M265I and M292A decreased significantly.

These mutants provided evidence that good protein expression doesn't necessarily mean good enzyme activity.

3.2.4.1.4 Mutants that displayed acceptable protein expression and unaffected enzyme activity

C208A, M70V, M206V and double mutant M202R_M206L, were allocated to this group because firstly these mutations did not affect protein expression and their activity was equal to or close to that of the wild type therefore characterised as "unaffected enzyme activity" (Figures 3.16 and 3.17 (d)). What these four mutants all had in common was that the mutations were located away from the active site.

M70V was found buried within the structure. M70V expressed well in whole cells as well as CFE and although the activity of M70V fluctuated in whole cells, it was more consistent in CFE.

C208 was mutated to alanine based on the information supplied by the 3DM data (Figure A1 (b)). What was interesting about position 208 was the fact that out of the 286 P450 sequences in the alignment, 50.3% contained a cysteine residue at this position, 43.7% contained methionine and only 2.8% contained alanine at this position. Nonetheless, alanine was selected to substitute cysteine. Vintém and co-workers (2005) also describe substituting cysteine with alanine in order to prevent inactivation of monoamine oxidase A (MAO A) and monoamine oxidase B (MAO B) from inactivation by cyclopropylamines. They developed two mutants, one for each enzyme, C374A (MAO A) and C365A (MAO B). Although the mutants retained activity, their turnover numbers were lower than that of the wild type (Vintém *et al*., 2005). Our mutant, C208A retained only 81% activity but had a TTN 8-fold higher than the wild type (Figure 3.18). In addition, the monoamine oxidase mutants were surface exposed and alanine is a non-polar amino acid. C208A, although close to the surface of the protein, was buried in the protein away from the surrounding aqueous environment, therefore selecting alanine as the substituent for cysteine proved to be a good choice.

Because M202 and M206 lie so close to one another, a double mutant M202R_M206L was created. Previously, we observed that M202R as a single mutant had no expression and no activity while M206V expressed well in whole cells and CFE and had an activity of 93% of WT in CFE. With arginine selected to

replace M202 (see 3.2.4.1.1), leucine was selected to replace M206 in the double mutant, because according to the 3DM correlation data valine never occurs together with arginine at position 202 (3DM position 149) in these CYPs, while the R-L combination occurs in 4 % of these CYPs. Interestingly the R-M combination occurs in less than 1 % of these CYPs, which might explain the lack of expression in the case of the M202R mutant. With the double mutant M202R_M206L, the activity of the enzyme was equivalent to that of wild type, despite the fact that M206 was substituted with leucine which had only the third highest occurrence at position 206. In the double mutant, the principle remains the same, the more mutations inserted into a single enzyme, the more stable the enzyme becomes. This is the reason why combinatorial mutants are known to be more successful than single mutants (Meinhold *et al.*, 2006b; Müller *et al*., 2013; Opperman & Reetz, 2010; Pennec *et al.*, 2014; Pham *et al*., 2013). Furthermore, the TTN of M202R_M206L was 5-fold higher than the TTN of the wild type (Figure 3.18).

3.2.4.1.5 Mutants that displayed acceptable protein expression and improved enzyme activity

Out of the 25 mutants evaluated, M18I, M56G, M241L, M292F and M231L_M232L were the only mutants that improved on the activity of the wild type (Figures 3.16 and 3.17 (d)) and like the previous group, all these mutations were remote from the active site.

M56G was the only mutant to not only improve on the wild types activity but also had better expression levels than wild type in CFE. Since no 3DM information was available for M56 because it does not belong to the core, selecting a substitute for methionine was quite challenging. Glycine was selected because its small hydrogen side chain gives it more conformational flexibility than other amino acids (Barnes *et al*., 2003) which proved to be a good choice given that M56G improved on the wild types activity by 50%. In addition, the TTN of M56 was double that of the wild type (Figure 3.18).

Like methionine at positions 4 and 56, no 3DM data was available for methionine at position 18 and based on the homology model, M18 was also located on the “tail” outside the ‘’body” of the protein. However, unlike M4G, the homology model suggested that M18 was not surface exposed but extended toward the interior of the protein and based on the fact that M18I expressed well in CFE and improved on the wild type’s activity by 27%, this gave us reason to believe that isoleucine was not exposed to the surrounding aqueous environment.

Although only 2.5% of these CYPs contained a leucine residue at position 241 (Figure A1 (i)), leucine once again proved to be a favourable substitution for methionine because not only did M241L have acceptable

expression in whole cells and CFE, its activity exceeded that of the wild type by 48% in whole cells and 26% in CFE.

In addition to alanine, phenylalanine was also a possible substituent for M292 based on the 3DM data (Figure A1 (n)). The interesting thing about M292A and M292F was that alanine and phenylalanine differ only in that phenylalanine contains a benzene ring however, it was this difference that made M292F a catalytically more active enzyme than M292A.

The 3DM data indicated that 20.5% of these CYPs contained a leucine residue at position 231 (Figure A1 (o)) and at position 232 (Figure A1 (p)), 56.5% of the sequences contained a leucine residue which was higher than methionine (43.5%). Therefore, leucine was a favourable substitution at both positions and although the L-L combination occur in only 8 % of these sequences, out of all 25 mutants, double mutant M231L_M232L had the best result overall in terms of expression and activity. M231L_M232L had acceptable expression in whole cells and CFE and its activity exceeded that of the wild type by 50% in whole cells and 88% in CFE. In addition, the TTN of M231L_M232L exceeded that of the wild type by almost 10-fold (Figure 3.18).

3.3 Conclusion

Biocatalyst instability among cytochrome P450 enzymes has been linked to hydrogen peroxide (H_2O_2), a product of protein uncoupling (Derat *et al*., 2006; Fang *et al*., 1997; Guengerich & Isin, 2008; Meunier *et al*., 2004). Hydrogen peroxide primarily inactivates these enzymes through degradation of the heme prosthetic group. However, many research groups have demonstrated that oxidation of cysteine and methionine residues can cause enzyme instability and how substituting Cys and Met residues with more oxidatively stable amino acids, can overcome this problem (Heinzelman *et al*., 2009; Johnsen *et al*., 2000; Kim *et al*., 2001; Kim *et al*., 2001; Lin *et al*., 2006; Miyazaki-Imamura *et al*., 2003; Nielsen & Borchert, 2000; Opperman & Reetz, 2010; Perry & Wetzel, 1987; Slavica *et al*., 2005; Vintém *et al*., 2005; Yonghua *et al*., 2000).

The instability of CYP153A6 was clearly demonstrated when performing whole cell biotransformations to evaluate the hydroxylation of *n*-octane by *E. coli* expressing the CYP153A6 operon (Gudiminchi *et al*., 2012).In this study, we employed site directed mutagenesis to substitute specific Cys and Met residues in

CYP153A6 as a resort to improve the operational stability of the enzyme in the presence of H_2O_2. However, hydroxylation of *n*-octane was carried out by performing biotransformations using cell free extracts.

Here, we demonstrate that although H_2O_2 does contribute to the instability of CYP153A6 that perhaps it is not the main cause of the enzymes instability. An increase in the initial enzyme activity was observed for the wild type (WT) CYP153A6 and the mutants when H_2O_2 was present during the hydroxylation of *n*-octane using cell free extracts. However, reactions containing H_2O_2 levelled off after 8 h whereas reactions without H_2O_2 did not, indicating that H_2O_2 was in fact affecting the stability of the WT and the CYP153A6 mutants. Mutations which were found to improve operational stability of CYP153A6 in the presence of H_2O_2 were not surface exposed however; they were remote from the active site. Furthermore, we demonstrate that single mutations and double mutations are just a stepping-stone to improve enzyme stability; only with combinatorial mutations will it become possible to create a fully functional and stable enzyme. Furthermore, attempts to crystallise CYP163A6 have not been successful yet. Oxidation of Cys and Met residues also contribute to microheterogeneity which can make protein crystallisation very difficult (Opperman & Reetz, 2010). Bloom and co-workers also found that protein stability promotes evolvability (Bloom *et al*., 2006) therefore, if a stable mutant of CYP153A6 can be created without compromising enzyme activity and product specificity, industrial application could be the next step.

Chapter 4

Conclusions and future outlook

From experiments done with the wild type CYP153A6 and the CYP153A6 mutants, we concluded that:

- **All the mutants - like the wild type - displayed 95% regio-selectively for the terminal carbon atom while the other 5% was the product of subterminal hydroxylation.** Although the graphs indicated in the results for Chapters two and three only refer to the production of 1-octanol, the GC results indicated that ≤ 5% 2-octanol was also produced by the CYP153A6 mutants and the wild type which is consistent with literature (Funoff *et al*., 2007).
- **Although H_2O_2 does contribute to the instability of CYP153A6, it is not the sole cause of the enzyme's instability**. We observed an increase in the initial activity of CYP153A6 (WT) in the presence of H_2O_2, this increase was also observed for the CYP153A6 mutants. However, H_2O_2 caused the activity of the WT and the mutants to level off after 8 h whereas in the absence of H_2O_2 enzyme activity continued to increase well above the 8 h mark. This was an indication that H_2O_2 was affecting the stability of the WT and the mutants however, the increase in the initial activity was not expected and we are not certain why, or how H_2O_2 contributed to the increased activity but this made it clear there must be other factors contributing to instability of CYP153A6. Investigating the reaction conditions used for biotransformation of *n*-octane more extensively studied could be start to determining what other factors are contributing the instability of CYP153A6.

- **Good protein expression doesn't necessarily mean that an enzyme will have high activity and vice versa.** In this study, we identified five mutants that had acceptable protein expression meaning that the amino acid substitutions for these five mutants did not have a significant impact on protein expression compared to the WT however, these substitutions significantly decreased enzyme activity. On the other hand, we identified four mutants that had improved on the activity of the wild type but had significantly low expression levels compared to the WT.

- **Combinatorial mutations might be the solution to creating a completely stable CYP153A6 enzyme.** We identified five mutants (M18I, M56G, M241L, M292F and M231L_M232L) with improved stability compared to the WT. Of these five mutants, the double mutant,

M231L_M232L, was the most stable overall indicating that more mutations, the more stable the enzyme. Therefore, it could be beneficial to make a combinatorial mutant using these five mutations and to create a variety of combinatorial mutants using the other single mutants as wells.

Summary

CYP153 hydroxylases have been found to be a very diverse group of cytochrome P450 enzymes because they have been isolated from a many different genera of alkane-degrading bacteria collected from various environments. This family of enzymes were the first bacterial monooxygenases identified that could catalyse terminal hydroxylation of aliphatic alkanes by regio-selectively inserting a single oxygen atom into the terminal C-H bond. Currently, there are no synthetic chemical catalysts capable of performing such a reaction. Therefore, CYP153s have attracted great interest from the petrochemical industry for the production of value added chemicals. In addition, the cost of using biological catalysts is much less than chemical catalysts and they promote green chemistry. However, many CYP153s are unstable biocatalysts therefore one of the long-term goals of our research group is to create a stable CYP153 hydroxylase for large-scale application. This study was a first attempt at achieving this goal. The first aim of this study was to establish factors affecting the operational stability of CYP153A6 - a well-characterised member of the CYP153 family – in order to establish conditions suitable for challenging mutants.

Our research group had previously cloned the complete operon from *Mycobacterium sp.* HXN-1500 encoding CYP153A6, Ferredoxin reductase and Ferredoxin into pET-28b (+) and expressed it in *E. coli* BL21 (DE3). When evaluating the operational stability of CYP153A6 (wild type), we observed that enzyme activity decreased when glucose dehydrogenase was added to biotransformations to facilitate cofactor regeneration. Glucose dehydrogenase oxidises glucose to glucono-lactone using NAD^+ or $NADP^+$ as cofactor. In the presence of water, glucono-lactone is converted to gluconic acid causing acidification which caused the activity of CYP153A6 to decrease. In cell free extracts limited interaction with FdR and Fdx limited 1-octanol production, so that with additional FdR and Fdx, 1-octanol production increased significantly. We also determined that the wild-type CYP153A6 could efficiently take care of cofactor regeneration using glucose and glycerol as a source of electrons when biotransformations are carried out 20°C. When storing the enzyme overnight, adding additional NADH to biotransformations appeared to have a stabilising effect on CYP153A6. Currently it appears as though CYP153A6 is most stable when biotransformations are carried out using a 300 mM Tris-HCl (pH 8) buffer containing 100 mM glucose and glycerol. We observed that higher buffer concentrations such as 400 mM Tris significantly decreased enzyme activity as well as when the buffer pH was lower than 7 or higher than 8. When evaluating the oxidative stability of the CYP153A6, an increase in the initial enzyme activity was observed when glucose oxidase was added to biotransformations for *in situ* generation of H_2O_2. However, reactions containing

H_2O_2 levelled off after 8 h whereas reactions without H_2O_2 did not, indicating that H_2O_2 was having a negative effect on the stability of CYP153A6. In addition, glucose oxidase also produces gluconic acid in the presence of water which could have simultaneously contributed to the decrease in enzyme activity as observed in the reactions containing glucose dehydrogenase. The reason for evaluating the operational stability of the CYP153A6 under various conditions was in order to select a set of conditions that would not only optimise reaction conditions for evaluating CYP153A6 mutants but the conditions should challenge the mutants as well.

The second aim of the study was to design and construct mutants of CYP153A6 by substituting specific cysteine and methionine residues with more oxidatively stable amino acids. Mutants were designed using the 3DM Bio-Prodict database in conjunction with Yasara. Site-directed mutagenesis was used to construct the mutants using the Quickchange PCR method as well as the Megaprimer PCR method. Once DNA sequencing confirmed that mutagenesis was successful, the plasmid DNA was transformed into *E. coli* BL21 (DE3). A total of 14 CYP153A6 mutants were created in this study.

For the third and final aim of this study, the operational stability of these 14 mutants together with 11 mutants that had previously been constructed by Dr. Opperman was evaluated by performing biotransformations of *n*-octane using cell free extracts. The conditions for the biotransformation were based on the set of conditions that had been selected for the mutants in the first aim of this study. The mutants were thus evaluated using a 300 mM Tris-HCl buffer (pH 8) containing 100 mM glucose and glycerol; additional FdR and Fdx were added to all reactions and as well additional 1 mM NADH. Biotransformations were carried out at 30°C and the mutants were tested in the presence of 800 and 1600 U/L (final concentration) glucose oxidase. After evaluation, mutants were divided into five groups based on expression levels and activity. From these results, it was clear that mutants with mutations M18I, M56G, M241L, M292F and M231L_M232L, all on the periphery of the enzyme, displayed significantly improved activity over the wild type in the presence of H_2O_2. Mutant M231L_M232L displayed the most improved activity indicating that combinatorial mutants will improve enzyme stability more significantly than single mutants.

References

Alonso-Gutiérrez, J., Teramoto, M., Yamazoe, A, Harayama, S., Figueras, a, & Novoa, B. (2011). Alkane-degrading properties of *Dietzia* sp. H0B, a key player in the Prestige oil spill biodegradation. *Journal of Applied Microbiology*, **111**(4), 800–810.

Asperger, O. (1981). Occurrence of cytochrome P-450 in Acinetobacterstrains after growth on *N*-hexadecane. *FEMS Microbiology*, ***1***, 309–312.

Barnes, M. R., Gray, I. C., Southan, C., Semple, C. A., Blake, J. A., Eppig, J., Hamburger, A. (2003). *BIOINFORMATICS FOR GENETICISTS*. (*1st Edition., Vol. 4*) West Sussex: John Wiley & Sons, Ltd, pp. 289-314.

Bell, S. G., & Wong, L. L. (2007). P450 enzymes from the bacterium *Novosphingobium aromaticivorans*. *Biochemical and Biophysical Research Communications*, **360(**3), 666–72.

Bell, S., Stevenson, J., Boyd, H., Campbell, S., Riddle, A., Orton, E., & Wong, L. (2002). Butane and propane oxidation by engineered cytochrome P450(cam). *Chemical Communications*, 490–491.

Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. *Journal of Biotechnology*, **124**(1), 128–145.

Bloom, J. D., Labthavikul, S. T., Otey, C. R., & Arnold, F. H. (2006). Protein stability promotes evolvability. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(15), 5869–5874.

Bordeaux, M., Galarneau, A., & Drone, J. (2012). Catalytic, mild, and selective oxyfunctionalization of linear alkanes: current challenges. *Angewandte Chemie (International Ed. in English)*, **51**(43), 10712–10723.

Bordeaux, M., Galarneau, A., Fajula, F., & Drone, J. (2011). A regioselective biocatalyst for alkane activation under mild conditions. *Angewandte Chemie*, **123**(9), 2123–2127.

Cappelletti, M. (2009). Monooxygenases involved in the n-alkanes metabolism by *Rhodococcus* sp. BCP1: molecular characterization and expression of alkB gene. Ph.D. Thesis, University of Bollogna.

Chang, D. (2003). Regio- and Stereoselective hydroxylations with *Sphingomonas* sp . HXN-200. Ph.D. Thesis, Swiss Ferderal Institute of Technology Zürich.

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

Chefson, A., & Auclair, K. (2006). Progress towards the easier use of P450 enzymes. *Molecular bioSystems*, **2(**10), 462–469.

Chisti, Y., & Moo-Young, M. (1986). Disruption of microbial cells for intracellular products. *Enzyme and Microbial Technology*, **8**(4), 194–204.

Cirino, P. C., & Arnold, F. H. (2002). Regioselectivity and activity of cytochrome P450 BM-3 and mutant F87A in reactions driven by hydrogen peroxide, **9**, 1–6.

Cirino, P. C., & Arnold, F. H. (2003). A self-sufficient peroxide-driven hydroxylation biocatalyst. *Angewandte Chemie (International Ed. in English)*, **42**(28), 3299–3301.

Danielson, P. B. (2002). The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. *Current Drug Metabolism*, **3**(6), 561–597.

Denisov, I. G., Makris, T. M., Sligar, S. G., & Schlichting, I. (2005). Structure and chemistry of cytochrome P450. *Chemical Reviews*, **105**(6), 2253–2277.

Derat, E., Kumar, D., Hirao, H., & Shaik, S. (2006). Gauging the relative oxidative powers of compound I, ferric-hydroperoxide, and the ferric-hydrogen peroxide species of cytochrome P450 toward C-H hydroxylation of a radical clock substrate. *Journal of the American Chemical Society*, **128**(2), 473–484.

Estabrook, R. W., Hildebrandt, A. G., Baron, J., Netter, K. J., & Leibman, K. (1971). A new spectral intermediate associated with cytochrome P-450 function in liver microsomes. *Biochemical and Biophysical Research Communications*, **42**(1), 132–139.

Fairbanks, G., Steck, T. L., & Wallachl, D. F. H. (1971). Electrophoretic analysis. *Biochemistry*, **10**(13), 2606–2617.

Fang, X., Kobayashi, Y., & Halpert, J. R. (1997). Stoichiometry of 7-ethoxycoumarin metabolism by cytochrome P450 2B1 wild-type and five active-site mutants. *FEBS Letters*, **416**(1), 77–80.

Fujita, N., Sumisa, F., Shindo, K., Kabumoto, H., Arisawa, A., Ikenaga, H., & Misawa, N. (2009). Comparison of two vectors for functional expression of a bacterial cytochrome P450 gene in *Escherichia coli* using CYP153 genes. *Bioscience, Biotechnology, and Biochemistry*, **73**(8), 1825–1830.

Funhoff, E. G., Bauer, U., García-Rubio, I., Witholt, B., & van Beilen, J. B. (2006). CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation. *Journal of Bacteriology*, **188(**14), 5220–5227.

Funhoff, E. G., Salzmann, J., Bauer, U., Witholt, B., & van Beilen, J. B. (2007). Hydroxylation and epoxidation reactions catalyzed by CYP153 enzymes. *Enzyme and Microbial Technology*, **40**(4), 806–812.

Funhoff, E. G., & Van Beilen, J. B. (2007). Alkane activation by P450 oxygenases. *Biocatalysis and Biotransformation*, **25**(2-4), 186–193.

Gottfert, M., Rothlisberger, S., Kundig, C., Beck, C., Marty, R., & Hennecke, H. (2001). Potential symbiosis-specific genes uncovered by sequencing a 410-Kilobase DNA region of the *Bradyrhizobium japonicum* chromosome. *Journal of Bacteriology*, **183(**4), 1405–1412.

Gudiminchi, R. K., Randall, C., Opperman, D. J., Olaofe, O. a, Harrison, S. T. L., Albertyn, J., & Smit, M. S. (2012). Whole-cell hydroxylation of *n*-octane by *Escherichia coli* strains expressing the CYP153A6 operon. *Applied Microbiology and Biotechnology*, **96**(6), 1507–1516.

Guengerich, F. P., & Isin, E. M. (2008). Mechanisms of cytochrome P450 reactions. *Acta Chim Slov,* **55***,* 7–19.

Gunsalus, I. C., Pederson, T. C., & Sligar, S. G. (1975). Oxygenase-catalyzed biological hydroxylations. *Annual Review of Biochemistry*, **44**, 377–407.

Hannemann, F., Bichet, A., Ewen, K. M., & Bernhardt, R. (2007). Cytochrome P450 systems--biological variations of electron transport chains. *Biochimica et Biophysica Acta*, **1770**(3), 330–344.

Hara, A., Baik, S., Syutsubo, K., Misawa, N., Smits, T. H. M., Beilen, J. B. Van, & Harayama, S. (2004). Cloning and functional analysis of alkB genes in *Alcanivorax borkumensis* SK2. *Environmental Microbiology*, **6**(3), 191–197.

Heinzelman, P., Snow, C. D., Smith, M. a, Yu, X., Kannan, A., Boulware, K., Arnold, F. H. (2009). SCHEMA recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability. *The Journal of Biological Chemistry*, **284**(39), 26229–26233.

Himes, R. a, & Karlin, K. D. (2009). Copper-dioxygen complex mediated C-H bond oxygenation: relevance for particulate methane monooxygenase (pMMO). *Current Opinion in Chemical Biology*, **13**(1), 119–131.

Iida, T., Sumita, T., Ohta, A., & Takagi, M. (2000). The cytochrome P450ALK multigene family of an *n*-alkane-assimilating yeast, *Yarrowia lipolytica*: cloning and characterization of genes coding for new CYP52 family members. *Yeast*, **16**, 1077–1087.

Isin, E. M., & Guengerich, F. P. (2007). Complex reactions catalyzed by cytochrome P450 enzymes. *Biochimica et Biophysica Acta*, **1770**(3), 314–329.

Johnsen, L., Fimland, G., & Eijsink, V. (2000). Engineering increased stability in the antimicrobial peptide pediocin PA-1. *Applied and Environmental Microbiology*, **66**(11), 4798–4802.

Johnston, J. B., Ouellet, H., Podust, L. M., & Ortiz de Montellano, P. R. (2011). Structural control of cytochrome P450-catalyzed ω-hydroxylation. *Archives of Biochemistry and Biophysics*, **507**(1), 86–94.

Kammann, M., Laufs, J., Schell, J., & Gronenborn, B. (1989). Nucleic Acids. *Nucleic Acids Res.*, **17**(13), 7367.

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Wada, T. (2002). Complete Genomic Sequence of Nitrogen-fixing Symbiotic Bacterium *Bradyrhizobium japonicum* USDA110, **197**, 189–197.

Kemble, D. J., & Sun, G. (2009). Direct and specific inactivation of protein tyrosine kinases in the Src and FGFR families by reversible cysteine oxidation. *Proceedings of the National Academy of Sciences of the United States of America*, **106(**13), 5070–5075.

Kim, Y. H., Berry, a. H., Spencer, D. S., & Stites, W. E. (2001). Comparing the effect on protein stability of methionine oxidation versus mutagenesis: steps toward engineering oxidative resistance in proteins. *Protein Engineering Design and Selection*, **14**(5), 343–347.

Kizawas, H., Tomura, D., Odat, M., Fukamizu, A., Hoshino, T., Gotohll, O., Shounij, H. (1991). Nucleotide Sequence of the Unique Nitrate / Nitrite-inducible Cytochrome P-450 cDNA from *Fusarium oxysporum*. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, **266**(16), 10632–10637.

Klingenberg, M. (1958). Pigments of rat liver microsomes. *Archives of Biochemistry and Biophysics*, **75,** 376–386.

Koch, D. J., Chen, M. M., van Beilen, J. B., & Arnold, F. H. (2009). In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Applied and Environmental Microbiology*, **75**(2), 337–344.

Kochius, S., Park, J. B., Ley, C., Könst, P., Hollmann, F., Schrader, J., & Holtmann, D. (2014). Electrochemical regeneration of oxidised nicotinamide cofactors in a scalable reactor. *Journal of Molecular Catalysis B: Enzymatic*, **103**, 94–99.

Krieger, E., & Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA — a Self-Parameterizing Force Field, **402**, 393–402.

Kubota, M., Nodate, M., Yasumoto-Hirose, M., Uchiyama, T., Kagami, O., Shizuri, Y., & Misawa, N. (2005). Isolation and functional analysis of cytochrome P450 CYP153A genes from various environments. *Bioscience, Biotechnology, and Biochemistry*, **69**(12), 2421–2430.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. *Nature*, **227**, 680–685.

Lamb, D. C., & Waterman, M. R. (2013). Unusual properties of the cytochrome P450 superfamily. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **368**(1612), 1-13

Larimer, F. W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Harwood, C. S. (2004). Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nature Biotechnology*, **22**(1), 55–61.

Lee, S., Goo, J., Kim, H., Oh, J., Kim, Y., & Kim, S. (2004). Optimization of methanol biosynthesis from methane using *Methylosinus trichosporium* OB3b. *Biotechnology Letters*, **26**, 947–950.

Li, L., Liu, X., Yang, W., Xu, F., Wang, W., Feng, L., Rao, Z. (2008). Crystal structure of long-chain alkane monooxygenase (LadA) in complex with coenzyme FMN: unveiling the long-chain alkane hydroxylase. *Journal of Molecular Biology*, **376**(2), 453–465.

Li, Z., & Chang, D. (2004). Recent advances in regio- and stereoselective biohydroxylation of non- activated carbon atoms. *Current Organic Chemistry,* **8**, 1647–1658.

Li, Z., Feiten, H., Beilen, J. B. Van, Duetz, W., & Witholt, B. (1999). Preparation of optically active *N* -benzyl-3-hydroxypyrrolidine by enzymatic hydroxylation. *Tetrahedron: Asymmetry*, **11**, 1323–1333.

Lieberman, R. L., & Rosenzweig, A. C. (2005). Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature*, **434(**7030), 177–182.

Lin, L. L., Lo, H. F., Chiang, W. Y., Hu, H. Y., Hsu, W .H., & Chang, C. T. (2006). Replacement of Methionine 208 in a truncated *Bacillus* sp. TS-23 - amylase with oxidation-resistant leucine enhances its resistance to hydrogen peroxide. *Current Microbiology*, **46**, 211–216.

Liska, D. J. (1998). The detoxification enzyme systems. *Alternative Medicine Review*, **3**(3), 187–198.

Loida, P. J., & Sligar, S. G. (1993). Molecula recognition in cytochrme P450: Mechanism for the control of uncoupling reactions. *Biochemistry*, **32**, 11530–11538.

Maeng, J., Sakai, Y., Tani, Y., & Kato, N. (1996). Isolation and characterization of a novel oxygenase that catalyzes the first step of nalkane oxidation in *Acinetobacter* sp strain M-1. *Journal of Bacteriology*, **178,** 3695–3700.

Maier, T., Förster, H. H., Asperger, O., & Hahn, U. (2001). Molecular characterization of the 56-kDa CYP153 from *Acinetobacter* sp. EB104. *Biochemical and Biophysical Research Communications*, **286**(3), 652–8.

McDonald, I. (2006). Diversity of soluble methane monooxygenase- containing methanotrophs isolated from polluted environments. *FEMS Microbiology Letters*, **255**, 225–232.

McLean, K. J., Sabri, M., Marshall, K. R., Lawson, R. J., Lewis, D. G., Clift, D., Munro, a W. (2005). Biodiversity of cytochrome P450 redox systems. *Biochemical Society Transactions*, **33**(Pt 4), 796–801.

Meinhold, P., Peters, M. W., Hartwick, A., Hernandez, A. R., & Arnold, F. H. (2006a). Engineering cytochrome P450 BM3 for terminal alkane hydroxylation. *Advanced Synthesis & Catalysis*, **348**(6), 763–772.

Merkx, M., Kopp, D., Sazinsky, M., Blazyk, J., Muller, J., & Lippard, S. (2001). Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: a tale of two irons and three proteins. *Angewandte Chemie International Edition*, **40**, 2782–2807.

Meunier, B., de Visser, S. P., & Shaik, S. (2004). Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chemical Reviews*, **104**(9), 3947–80.

Miyazaki-Imamura, C., Oohira, K., Kitagawa, R., Nakano, H., Yamane, T., & Takahashi, H. (2003). Improvement of H_2O_2 stability of manganese peroxidase by combinatorial mutagenesis and high-throughput screening using in vitro expression with protein disulfide isomerase. *Protein Engineering Design and Selection*, **16**(6), 423–428.

Mot, R. De, & Parret, A. H. A. (2002). Monooxygenases in prokaryotes, **10**(11), 502–508.

Müller, C. A., Akkapurathu, B., Winkler, T., Staudt, S., Hummel, W., Gröger, H., & Schwaneberg, U. (2013). In Vitro Double Oxidation of *n*-Heptane with Direct Cofactor Regeneration. *Advanced Synthesis & Catalysis*, **355**(9), 1787–1798.

Müller, R., Asperger, O., & Kleber, H. (1989). Purification of cytochrome P450 from *n*-hexadecane-grown *Acinetobacter calcoaceticus*. *Biomedica Biochima Acta*, **48**(4), 243–254.

Nagao, T., Mitamura, T., Wang, X. H., Negoro, S., Yomo, T., Urabe, I., & Okadat, H. (1992). Cloning, nucleotide sequences, and enzymatic properties of glucose dehydrogenase isozymes from *Bacillus megaterium* IAM1030. *Journal of Bacteriology*, **174**(15), 5013–5020.

Narhi, L. O., & Fulco, A. J. (1986). Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P450 monooxygenase induced by barbiturates in *Bacillus megaterium*. *The Journal of Biological Chemistry*, **261(**16), 7160–7169.

Nelson, D., Koymans, L., Kamataki, T., Stegeman, J., Feyereisen, R., Waxman, D., Nebert, D. (1996). P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*, **6**(1), 1–42.

Nie, Y., Chi, C. Q., Fang, H., Liang, J. L., Lu, S. L., Lai, G. L., Wu, X. L. (2014). Diverse alkane hydroxylase genes in microorganisms and environments. *Scientific Reports*, **4**, 1-11.

Nie, Y., Liang, J.-L., Fang, H., Tang, Y.-Q., & Wu, X.-L. (2013). Characterization of a CYP153 alkane hydroxylase gene in a Gram-positive *Dietzia* sp. DQ12-45-1b and its "team role" with alkW1 in alkane degradation. *Applied Microbiology and Biotechnology*.

Nielsen, J. E., & Borchert, T. V. (2000). Protein engineering of bacterial K-amylases. *Biochimica et Biophysica Acta*, ***1543***, 253–274.

Nierman, W. C., Feldblyum, T. V, Laub, M. T., Paulsen, I. T., Nelson, K. E., Eisen, J. a, Eisen, J. (2001). Complete genome sequence of *Caulobacter crescentus*. *Proceedings of the National Academy of Sciences of the United States of America*, **98(**7), 4136–4141.

Nodate, M., Kubota, M., & Misawa, N. (2006). Functional expression system for cytochrome P450 genes using the reductase domain of self-sufficient P450RhF from *Rhodococcus* sp. NCIMB 9784. *Applied Microbiology and Biotechnology*, **71**(4), 455–462.

Olaofe, O. A., Fenner, C. J., Gudiminchi, R. K., Smit, M. S., & Harrison, S. T. L. (2013). The influence of microbial physiology on biocatalyst activity and efficiency in the terminal hydroxylation of *n*-octane using *Escherichia coli* expressing the alkane hydroxylase, CYP153A6. *Microbial Cell Factories*, **12**(1).8

Omura, T. (1999). Forty years of cytochrome P450. *Biochemical and Biophysical Research Communications* **266**, 690–698.

Omura, T., & Sato, R. (1964).The carbon monoxide-binding pigment of liver microsomes. *Journal of Biological Chemistry,* **239**, 2370-2378.

Opperman, D. J., & Reetz, M. T. (2010). Towards practical Baeyer-Villiger-monooxygenases: design of cyclohexanone monooxygenase mutants with enhanced oxidative stability. *Chembiochem : A European Journal of Chemical Biology*, **11(**18), 2589–96.

Pennec, A., Jacobs, C. L., Opperman, D. J., & Smit, M. S. (2014). Revisiting cytochrome P450-mediated oxyfunctionalization of linear and cyclic alkanes. *Advanced Synthesis & Catalysis*, **357**(1), 118–130.

Perry, L. J., & Wetzel, R. (1987). The role of cysteine oxidation in the thermal inactivation of T4 lysozyme. *"Protein Engineering, Design and Selection,"* **1**(2), 101–105.

Pham, S. Q., Gao, P., & Li, Z. (2013). Engineering of recombinant *E. coli* cells co-expressing P450pyrTM monooxygenase and glucose dehydrogenase for highly regio- and stereoselective hydroxylation of alicycles with cofactor recycling. *Biotechnology and Bioengineering*, **110**(2), 363–373.

Pham, S. Q., Pompidor, G., Liu, J., Li, X.-D., & Li, Z. (2012). Evolving P450pyr hydroxylase for highly enantioselective hydroxylation at non-activated carbon atom. *Chemical Communications (Cambridge, England)*, **48**(38), 4618–4620.

Poulos, T. L., Finzel, B. C., & Howard, A. J. (1987). High-resolution crystal structure of cytochrome P450cam. *Journal of Molecular Biology*, **195**(3), 687–700.

Poulos, T. L., Finzelsy, C., Gunsalusll, I. C., Wagner, C., & Kraut, J. (1985). The 2.6-A crystal structure of *Pseudomonas putida* cytochrome P-450. *The Journal of Biological Chemistry*, **90**(30), 16122–16130.

Ramakrishnani, L., & Falkow, S. (1994). *Mycobacterium marinum* persists in cultured mammalian cells in a temperature-restricted fashion. *Infection and Immunity*, **62**(8), 3222–3229.

Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L., & Turner, N. J. (2002). Identification of a new class of cytochrome P450 from a *Rhodococcus* sp. *Journal of Bacteriology*, **184**(14), 3898–3908.

Rojo, F. (2009). Degradation of alkanes by bacteria. *Environmental Microbiology*, **11**(10), 2477–2490.

Sanchis, J., Fernández, L., Carballeira, D. J., Drone, J., Gumulya, Y., Höbenreich, H. Reetz, M. T. (2008). Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. *Applied Microbiology and Biotechnology*, **81**, 387–397.

Scheps, D., Honda Malca, S., Richter, S. M., Marisch, K., Nestl, B. M., & Hauer, B. (2013). Synthesis of ω-hydroxy dodecanoic acid based on an engineered CYP153A fusion construct. *Microbial Biotechnology*, **6**(6), 694–707.

Slavica, A., Dib, I., & Nidetzky, B. (2005). Single-site oxidation, cysteine 108 to cysteine sulfinic acid , in *D* -amino acid oxidase from *Trigonopsis variabilis* and its structural and functional consequences. *Applied and Environmental Microbiology*, **71**(12), 8061–8068.

Smits, T. H. M., Balada, S. B., Witholt, B., & van Beilen, J. B. (2002). Functional analysis of alkane hydroxylases from Gram-negative and Gram-positive bacteria. *Journal of Bacteriology*, **184**, 1733–1742.

Smits, T. H. M., Röthlisberger, M., Witholt, B., & Beilen van, J. B. (1999). Molecular screening for alkane hydroxylase genes in Gram-negative and Gram-positive strains. *Environmental Microbiology*, **1**, 307–318.

Smits, T. H., Seeger, M. a, Witholt, B., & van Beilen, J. B. (2001). New alkane-responsive expression vectors for *Escherichia coli* and *Pseudomonas*. *Plasmid*, **46**(1), 16–24.

Söhngen, N. L. (1913). Benzin, Petroleum, Paraffin01 and Paraffin als Kohlenstoff-und Energiequelle fur Mikroben. *Zentbl. Bakt. ParasitKde*, **37**, 595.

Sono, M., Roach, M. P., Coulter, E. D., & Dawson, J. H. (1996). Heme-containing oxygenases. *Chemical Reviews*, **96**(7), 2841–2888.

Stern, J. O., & Peisach, J. (1974). Study of cytochrome of the CO-adduct. *The Journal of Biological Chemistry*, **249**, 7495–7498.

Suemori, A., & Iwakura, M. (2007). A systematic and comprehensive combinatorial approach to simultaneously improve the activity , reaction specificity , and thermal stability of *p* - hydroxybenzoate hydroxylase. *The Journal of Biological Chemistry*, **282**(27), 19969–19978.

Takayaa, N., Suzukia, S., Kuwazakia, S., Shouna, H., Maruoc, F., Yamaguchid, M., & Takeo, K. (1999). Cytochrome P450nor, a novel class of mitochondrial cytochrome P450 involved in nitrate respiration in the fungus *Fusarium oxysporum*. *Archives of Biochemistry and Biophysics*, **372**(2), 340–346.

Tang, W. L., Li, Z., & Zhao, H. (2010). Inverting the enantioselectivity of P450pyr monooxygenase by directed evolution. *Chemical Communications (Cambridge, England)*, **46**(30), 5461–5463.

Tani, A., Ishige, T., Sakai, Y., & Kato, N. (2001). Gene structures and regulation of the alkane hydroxylase complex in *Acinetobacter* sp. strain M-1. *Journal of Bacteriology*, **183**, 1819–1823.

Van Beilen, J. B., Duetz, W. a, Schmid, A., & Witholt, B. (2003). Practical issues in the application of oxygenases. *Trends in Biotechnology*, **21**(4), 170–177.

Van Beilen, J. B., & Funhoff, E. G. (2005). Expanding the alkane oxygenase toolbox: new enzymes and applications. *Current Opinion in Biotechnology*, **16**(3), 308–314.

Van Beilen, J. B., & Funhoff, E. G. (2007). Alkane hydroxylases involved in microbial alkane degradation. *Applied Microbiology and Biotechnology*, **74**(1), 13–21.

Van Beilen, J. B., Funhoff, E. G., van Loon, A., Just, A., Kaysser, L., Bouza, M., Witholt, B. (2006). Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. *Applied and Environmental Microbiology*, **72**(1), 59–65.

Van Beilen, J. B., Holtackers, R., Lüscher, D., Bauer, U., Witholt, B., & Duetz, W. A. (2005). Biocatalytic production of perillyl alcohol from limonene by using a novel *Mycobacterium* sp. cytochrome P450 alkane hydroxylase expressed in *Pseudomonas putida*. *Applied and Environmental Microbiology*, **71**(4), 1737–1744.

Van Beilen, J. B., Li, Z., Duetz, W. a., Smits, T. H. M., & Witholt, B. (2003). Diversity of alkane hydroxylase systems in the environment. *Oil & Gas Science and Technology*, **58**(4), 427–440.

Van Beilen, J. B., Marin, M. M., Smits, T. H. M., Rothlisberger, M., Franchini, A. G., Witholt, B., & Rojo, F. (2004). Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis*. *Environmental Microbiology*, **6**(3), 264–273.

Van Beilen, J., & Panke, S. (2001). Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes. *Microbiology*, **147**, 1621–1630.

Van Beilen, J., & Smits, T. (2002). Alkane hydroxylase homologues in Gram-positive strains. *Environmental Microbiology*, **4**(11), 676–682.

Van Beilen, J., Wubbolts, M., & Witholt, B. (1994). Genetics of alkane oxidation by *Pseudomonas oleovorans*. *Biodegradation*, **5**, 161–174.

Vintém, A. P. B., Price, N. T., Silverman, R. B., & Ramsay, R. R. (2005). Mutation of surface cysteine 374 to alanine in monoamine oxidase A alters substrate turnover and inactivation by cyclopropylamines. *Bioorganic & Medicinal Chemistry*, **13**(10), 3487–3495.

Wang, L., Wang, W., Lai, Q., & Shao, Z. (2010). Gene diversity of CYP153A and AlkB alkane hydroxylases in oil-degrading bacteria isolated from the Atlantic Ocean. *Environmental Microbiology*, **12**(5), 1230–1242.

Weckbecker, A., & Hummel, W. (2005). Glucose dehydrogenase for the regeneration of NADPH and NADH. *Methods Biotechnol*, **17**(8), 225–237.

Xu, F., Bell, S., Lednik, J., Insley, A., Rao, Z., & Wong, L. (2005). The heme monooxygenase cytochrome P450(cam) can be engineered to oxidize ethane to ethanol. *Angewandte Chemie*, **44**, 4029–4032.

Yang, Y., Jiang, L., Zhu, L., Wu, Y., & Yang, S. (2000). Thermal stable and oxidation-resistant variant of subtilisin E. *Journal of Biotechnology*, **81**, 113–118.

Yang, Y., Liu, J., & Li, Z. (2014). Engineering of p450pyr hydroxylase for the highly regio- and enantioselective subterminal hydroxylation of alkanes. *Angewandte Chemie (International Ed. in English)*, **53**(12), 3120–3124.

Zhou, R., Huang, C., Zhang, A., Bell, S. G., Zhou, W., & Wong, L.-L. (2011). Crystallization and preliminary X-ray analysis of CYP153C1 from *Novosphingobium aromaticivorans* DSM12444. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, **67**(Pt 8), 964–967.

Websites:

Cytochrome P450 Homepage (2009) http://drnelson.utmem.edu/

OligoAnalyser (2009) http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/

3DM Bio-Prodict: https://3dm.bio-prodict.nl/

The Cytochrome P450 Engineering Database: http://www.cyped.uni-stuttgart.de/cgi-bin/CYPED5/index.pl

RCSB Protein Data Bank (PDB): http://www.rcsb.org/pdb/home/home.do

Appendix

Table A1. Composition of media used in this study

Media	Composition	Concentration
Luria-Butani (LB) media	Tryptone	10 $g.L^{-1}$
	Yeast extract	5 $g.L^{-1}$
	NaCl	5 $g.L^{-1}$
Luria-Butani (LB) agar	Tryptone	10 $g.L^{-1}$
	Yeast extract	5 $g.L^{-1}$
	NaCl	5 $g.L^{-1}$
	Bacteriological agar	12 $g.L^{-1}$
ZYP5052 Auto-induction media	Tryptone	10 $g.L^{-1}$
	Yeast extract	5 $g.L^{-1}$
	$(NH^4)_2SO_4$	25 mM
	KH_2PO_4	50 mM
	Na_2HPO_4	50 mM
	Glucose	0.5 $g.L^{-1}$
	Glycerol	5 $g.L^{-1}$
	α- lactose	2 $g.L^{-1}$
	$MgSO_4$	2 mM
SOC media	Tryptone	10 $g.L^{-1}$
	Yeast extract	5 $g.L^{-1}$
	NaCl	10 mM
	MgCl	10 mM
	KCl	2.5 mM
	$MgSO_4$	10 mM
	Glucose	20 mM

Table A2 Conditions used in experiment 1 to evaluate the effect of GDH on the activity and stability of CYP153A6. Each reaction was carried out in triplicate for 1, 2, 4 and 8 hours (h).

Reaction	1	2	3	4	5	6	7	8
P450	200	400	200	200	400	200	400	400
GDH	200	400	200	400	200	200	200	200
NADH	10	20	20	10	10	10	10	10
Buffer	790	380	780	590	590	790	590	590
***n*-octane**	250	250	250	250	250	25	25	250
1-octanol	NONE	NONE	NONE	NONE	NONE	NONE	NONE	YES

Table A3 Conditions used in experiment 2 to evaluate the effect of GDH on the activity and stability of CYP153A6. Each reaction was carried out in duplicate for 1, 2, 4 and 8 h.

Reaction	1	2	3	4	5
P450	400	400	400	400	400
GDH	200	400	200	0	0
NADH	10	10	10	10	10
CFE	0	0	200	200	400
Buffer	590	390	390	590	590
***n*-octane**	250	250	250	250	250

Table A4 Conditions used to evaluate the effect of additional FdR/Fdx on the activity/stability of CYP153A6. Each reaction was carried out in duplicate for 1, 2, 4, 8 and 18 h.

Reaction	1	2	3	4
P450	250	250	250	250
FdR/Fdx	0	125	250	125
CFE	250	125	0	125
Glucose (100 mM)	YES	YES	YES	NONE
Buffer	700	700	700	700
***n*-octane**	250	250	250	250

Table A5 Conditions used to evaluate the effect of H_2O_2 produced by glucose oxidase on CYP153A6 activity/stability. Each reaction was carried out in triplicate for 2, 4, 8 and 18 h.

Reaction	1	2	3	4	5	6	7	8
P450	250	250	250	250	250	250	250	250
FdR/Fdx	250	250	250	250	250	250	250	250
NADH (1 mM)	NONE	NONE	NONE	NONE	NONE	NONE	NONE	YES
Glucose oxidase (U/L)	400	600	0	400	600	0	0	0
Buffer	700	700	700	700	700	700	700	700
***n*-octane**	250	250	250	250	250	250	250	250
Temp. (°C)	20	20	20	30	30	30	20	20

Table A6 Conditions used to optimise cell free extract biotransformations of *n*-octane using CYP153A6 by evaluating the effect of different buffer concentrations. Reactions carried out in triplicate for 2, 8 and 18h.

Reaction	1	2	3	4	5	6
P450	250	250	250	250	250	250
FdR/Fdx	250	250	250	250	250	250
NADH (1 mM)	YES	YES	YES	YES	YES	YES
Glucose	100	100	100	100	100	100
Tris-HCl (mM)	200	400	200	200	400	200
pH	8	8	8.5	8	8	8.5
Buffer	700	700	700	700	700	700
***n*-octane**	250	250	250	250	250	250
Temp (°C)	20	20	20	30	30	30

Table A7 Conditions used to optimise cell free extract biotransformations of *n*-octane using CYP153A6 by evaluating the effect of different buffers and buffer pH. Reactions carried out in triplicate for 2 h only.

Reaction	1	2	3	4	5	6	7
P450	250	250	250	250	250	250	250
FdR/Fdx	250	250	250	250	250	250	250
NADH (1 mM)	YES	YES	YES	YES	YES	YES	YES
Glucose (100 mM)	YES	YES	YES	YES	YES	YES	YES
Tris-HCl buffer (300 mM)	700	0	0	0	0	0	0
MES-MOPS-Tris buffer (300 mM)	0	700	700	700	700	700	700
pH	8	6	6.5	7	7.5	8	8.5
***n*-octane**	250	250	250	250	250	250	250
Temp. (°C)	20	20	20	20	20	20	20

Table A8 Conditions used to evaluate the effect of different P450 concentrations used in the BRM to try to optimise biotransformations of *n*-octane by CYP153A6. Each reaction was carried out in triplicate for 2, 8 and 18 h

Reaction	1	2	3
P450	50	100	250
FdR/Fdx	250	250	250
NADH (1 mM)	YES	YES	YES
Glucose (100 mM)	YES	YES	YES
Tris-HCl buffer (300 mM)	950	850	700
pH	8	8	8
***n*-octane**	250	250	250
Temp. (°C)	30	30	30

Table A9 Sequencing primers used in this study

Primer name	Primer sequence (5'- 3')	Tm (°C)	Manufacturer
T7 promotor	TAA TAC GAC TCA CTA TAG GG	48.3	Integrated DNA technologies Inc.
CYP153A6_R_HindIII	AAG CTT CAG GCG TTG ATG CGC ACG	64.5	Integrated DNA technologies Inc.

Table A10 Sequencing primer used for each CYP153A6 mutation

CYP153A6 mutation	T7 promotor	CYP153A6_R_HindIII
C52V	+	-
C208A	+	-
M56G	+	-
M91L	+	-
M101N	+	-
M101S	+	-
M105R	+	-
M129F	+	-
M164T	+	-
M202R	+	-
M206V	+	-
M241D	+	-
M306G	-	+
M329L	-	+

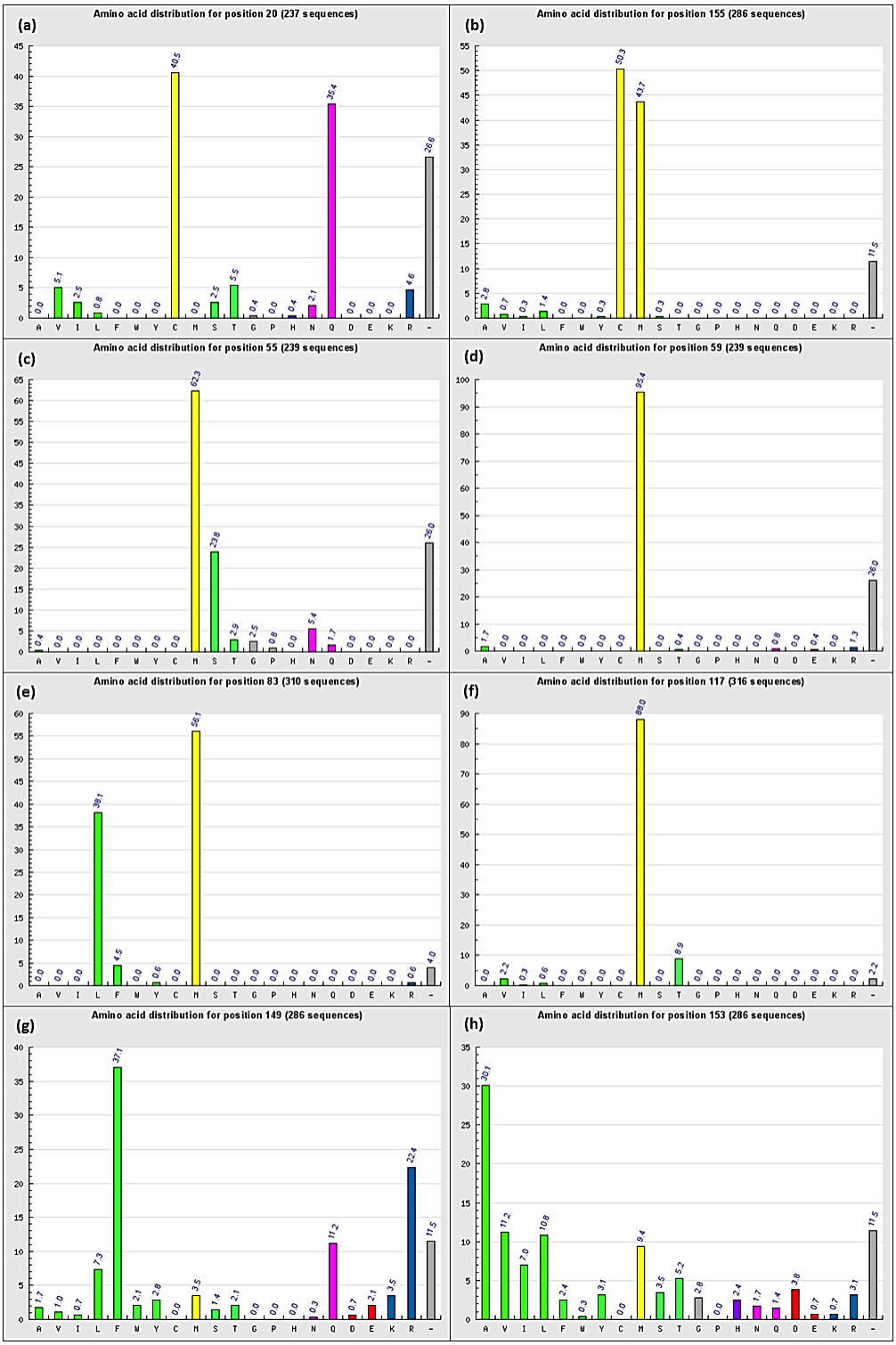

Figure A1 Structural sequence alignment between CYP153A6 and 335 structurally similar P450s was performed by the 3DM-Bioprodict database. Graphs illustrating the amino acid distribution for selected positions in a consensus sequence, in this case positions containing Cysteine and Methionine residues. Position numbers indicated on the graph the "3DM numbers" which correspond to the 3DM consensus sequence. Positions that correspond to the homology model of CYP153A6 are, (a) 52, (b) 208, (c) 101, (d) 105, (e) 129, (f) 164, (g) 202 and (h) 206.

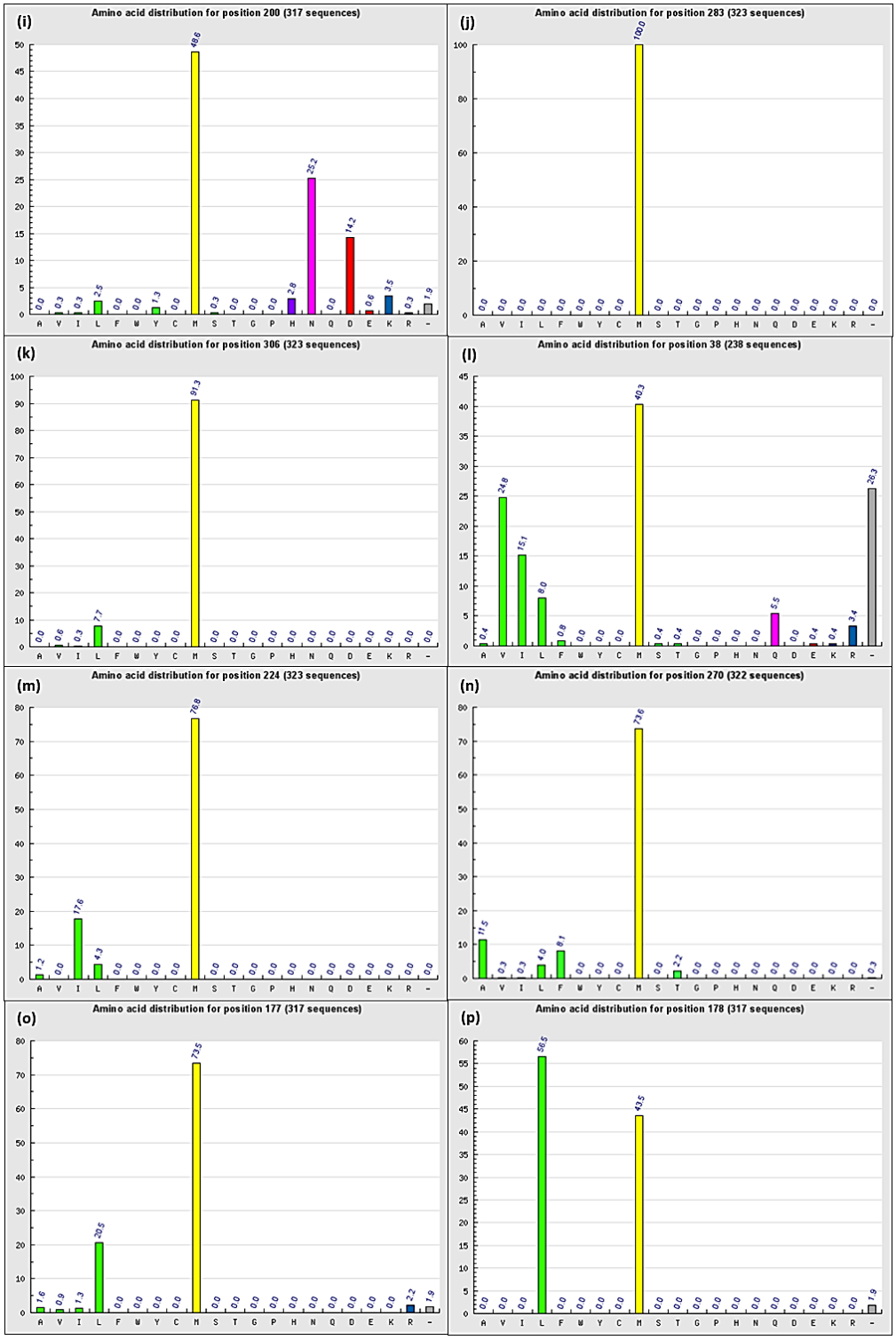

Figure A1 (continued) Position numbers indicated on the graph the "3DM numbers" which correspond to the 3DM consensus sequence. Positions that correspond to the homology model of CYP153A6 are, (i) 241, (j) 306, (k) 329, (l) 70, (m) 265, (n) 292, (o) 231 and (p) 232.

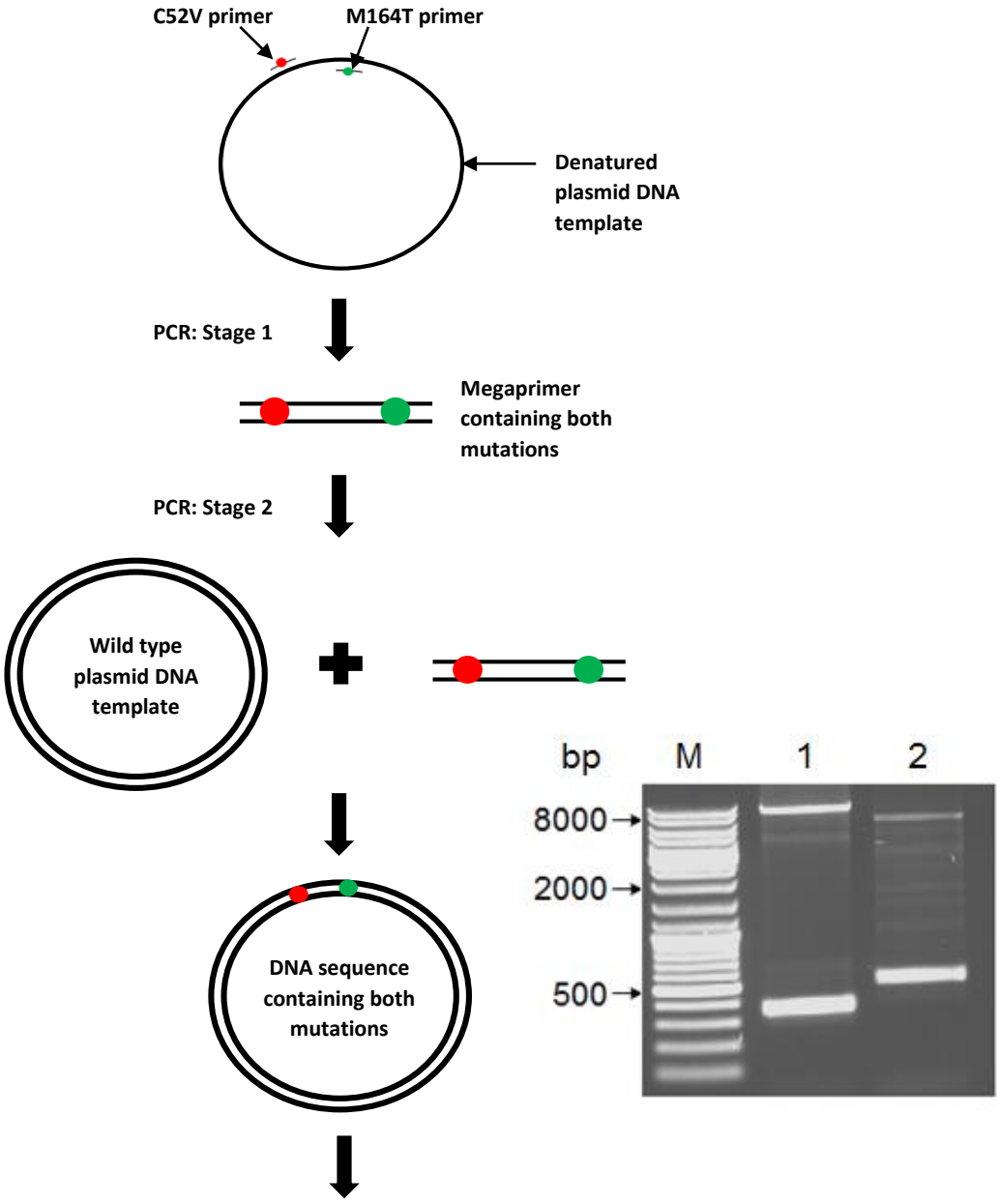

Figure A2 Schematic outline of the mutagenesis protocol to insert two mutations simultaneously using the megaprimer PCR method. In stage 1 of the reaction the megaprimer is generated containing the mutations introduced by the mutagenic primers. In stage 2 of the reaction, the megaprimer is extended using the wild-type plasmid DNA template. The agarose gel illustrates the PCR products of the two reactions carried out in this study, lane 1 shows the PCR products of the reaction mixture that contained the C52V_forward and M164T_Reverse primers and lane 2 shows the PCR products of the reaction mixture that contained the M164T_forward and C52V_Reverse primers. The two most distinct bands observed in lanes 1 and 2 are the complete DNA sequence containing both mutations (8190 bp) and the megaprimer (approximately 360 bp). A restriction digest to separate the two mutations and ligate them into separate pET-28b (+) was carried out using the PCR product from lane 1 because too much non-specific binding was observed from the PCR product in lane 2.

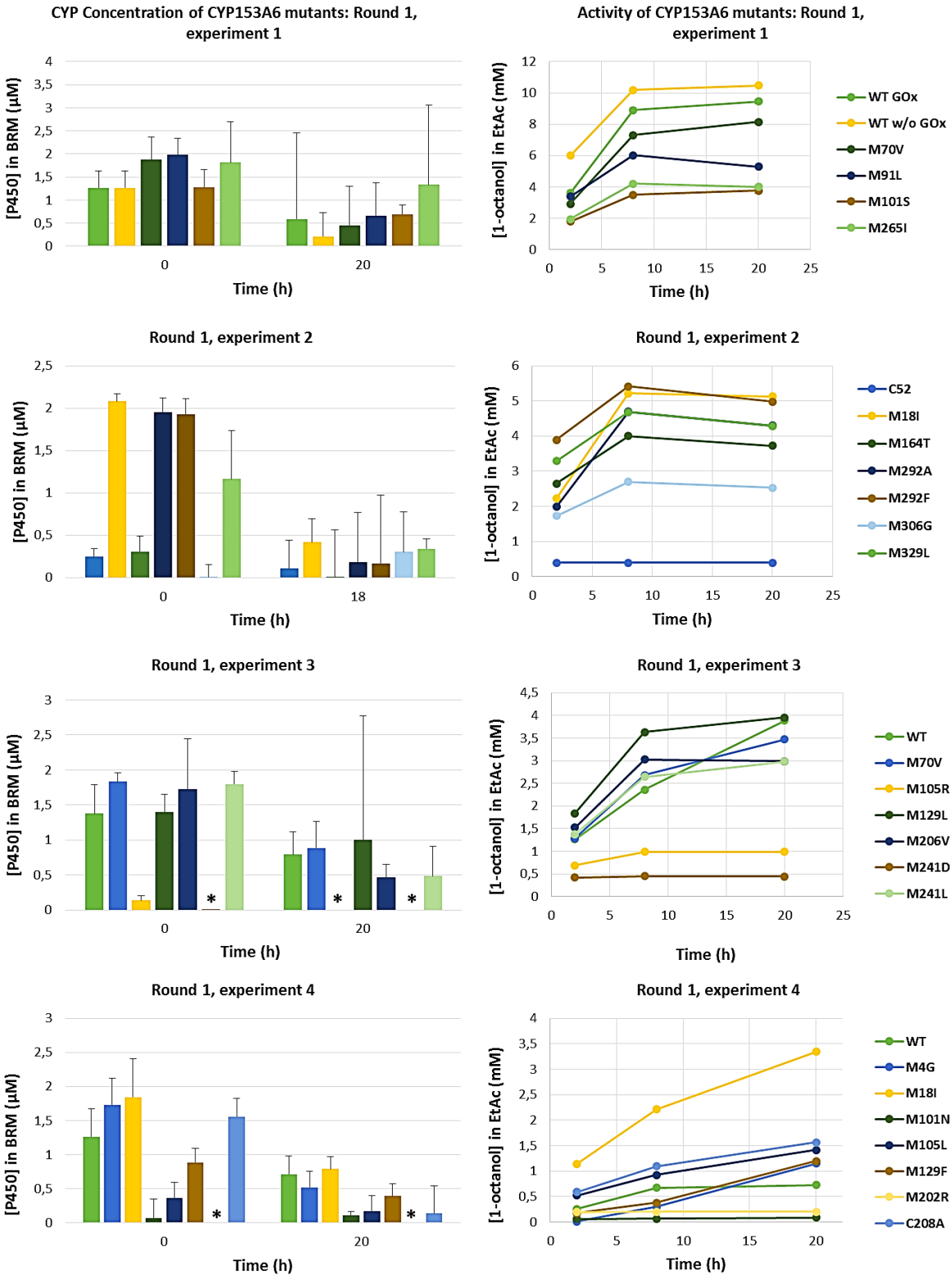

Figure A3 Round 1 consisted of four experiments into which the CYP153A6 mutants were divided at random. Oxidative stability of each mutant was evaluated in the presence of 800 U/L (final concentration) GOx for *in situ* production of H_2O_2. No catalytically active P450 was detected for mutants marked with an asterisk (*).

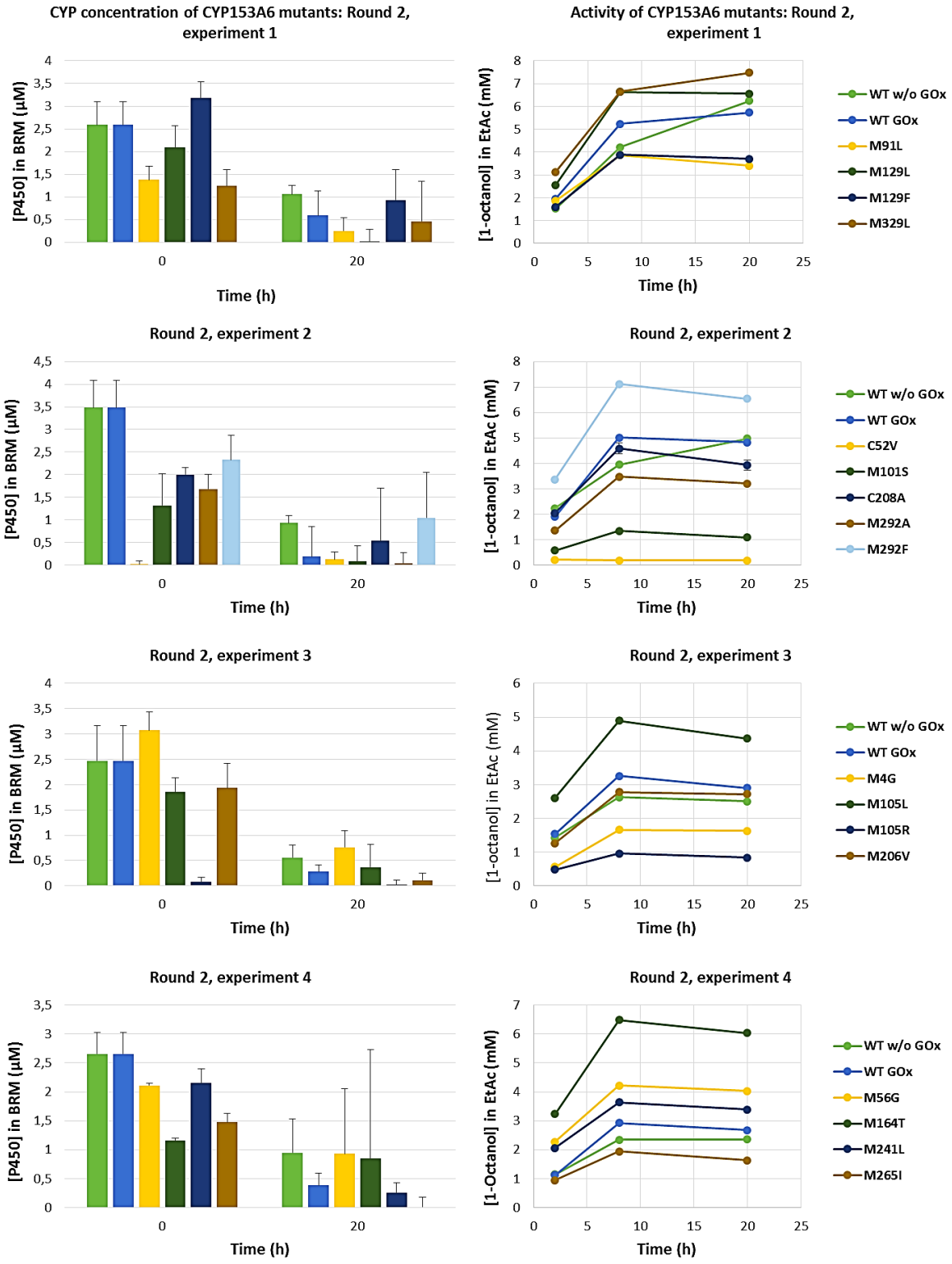

Figure A4 Round 2 consisted of five experiments into which the CYP153A6 mutants were divided at random. Reaction operational stability of each was evaluated in the presence of 1600 U/L (final concentration) GOx for *in situ* production of H_2O_2.

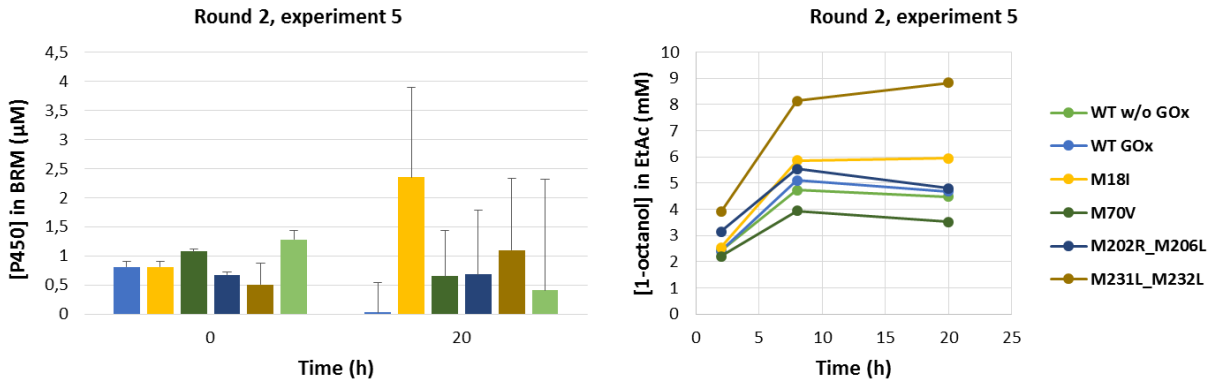

Figure A4 (continued) Round 2 consisted of five experiments into which the CYP153A6 mutants were divided at random. Oxidative stability of each mutant was evaluated in the presence of 1600 U/L (final concentration) GOx for *in situ* production of H_2O_2.