Directed Evolution of Cytochrome P450 for Small Alkane Hydroxylation

Thesis by

Mike Ming Yu Chen

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2011

(Defended April 28, 2011)

ACKNOWLEDGEMENTS

It has been a privilege to be a part of the Arnold lab at Caltech for the last six years. I want to thank my advisor, France Arnold, for providing a great work environment and allowing me the freedom to learn so many new techniques and test new ideas. I am grateful to the members of my committee, Mark Davis, John Bercaw, and Jay Labinger, for advice throughout this process. I am also grateful to the National Science Foundation for the graduate research fellowship that supported my graduate work.

Since the very first days of being a member of the Arnold lab, I have been fortunate to be around a group of wonderful researchers and helpful labmates who are far too numerous for me to name. I particularly owe many thanks to Matt Peters, Peter Meinhold, Michelle Meyer, and Marco Landwehr for teaching me the basics of molecular biology and protein engineering. I am also grateful to the many collaborators that I had the fortune to work with, Rudi Fasan, Daniel Koch, Andrew Sawayama, Chris Snow, Christina Vizcarra, Jorge Rodriguez, Jared Lewis, and Pedro Coelho, amongst many others.

My family has been a constant source of love and support. I want to thank my parents Newton and Stella Chen for encouraging me to pursue my studies and allowing me to be away from them for so long during a health crisis. I want to thank my sister and brother-in-law, Matty and Barry, for their support, encouragement, and providing me a place to visit when I wanted to get away from Caltech. I also want to thank my wonderful girlfriend Sabine, who was an immense help throughout my thesis writing process and source of constant encouragement. Finally, I want to thank my grandmother, Ai Qin Chen, who will always be an inspiration to me for her convictions about learning and self-improvement.

ABSTRACT

Methane is an ideal alternative to petroleum refining as a chemical feedstock source since it is highly abundant an inexpensive. However, the lack of selective methane oxidation catalysts has limited such utilization. Starting from cytochrome P450 CYP102A1 (BM3) from *Bacillus megaterium*, which prefers C_{12}-C_{20} fatty acids as its substrates, I investigated several protein engineering approaches to shift the enzyme's substrate specificity toward small gaseous alkanes, with the ultimate goal of methane. By continuing previous directed evolution efforts in our group, a variant with wild-type-like affinity and catalytic efficiency for propane, $P450_{PMO}$, was isolated. To alleviate the loss of protein thermostability (~ 10 ^{o}C) as a result of this approach, mutations were targeted to the BM3 active site with site saturation mutagenesis, targeted mutagenesis with a reduced set of amino acids, and computationally guided library designs. From these enzyme libraries, variants were identified that replicated much of the $P450_{PMO}$ activities with a minimal number of mutations while maintaining wild-type thermostability.

Continuing the protein engineering with a high throughput ethane hydroxylation screen, variants with improved *in vitro* ethane hydroxylation activity were obtained. However, in whole-cell ethane bioconversions, BM3-derived variants could not match the activity of a natural P450 alkane hydroxylase, CYP153A6. To investigate the oxidation capability of the P450 oxo-ferryl porphyrin radical intermediate directly, I employed a variety of terminal oxidants to support P450 alkane hydroxylation reactions abridging the P450 catalytic cycle. In this study, the CYP153A6 oxo-ferryl intermediate was able to oxidize methane in reactions using iodosylbenzene, which demonstrated that direct methane-to-methanol conversion by a P450 heme porphyrin catalyst at ambient conditions is possible and does not necessarily require the use of additional effectors to alter the active site geometry.

THESIS SUMMARY

Selective hydroxylation of small alkanes is a long-standing problem for which few practical catalysts are available. The lack of catalysts that can efficiently convert gaseous alkanes into transportable liquid commodities has been a barrier to broader utilization of these resources. In particular, methane, the principal component of natural gas, is an ideal alternative to petroleum as a chemical feedstock source since it is highly abundant and inexpensive. Currently, methane is converted to methanol via an energy intensive, endothermic, and costly process that first converts methane into synthesis gas, followed by methanol synthesis from this intermediate. This process is economically feasible only on a large scale, which in combination with general transportation limitations of a gas commodity prevent methane recovery from many sources.

The selective oxidation of small alkanes is difficult because of the inertness of the alkane C-H bond, which requires highly reactive radical or ionic species to cleave. However, as the desired partial oxidation products, alcohol and aldehyde, have weaker C-H bonds compared to the alkane, they are susceptible to further oxidation to CO_2. While this transformation has been achieved only by a limited set of transition-metal-based catalyst systems, a variety of alkane hydroxylases found in alkanotrophic microorganisms support selective alkane oxidation at ambient conditions using oxygen as the oxidant. Chapter 1 of this thesis provides an introduction to enzymatic alkane oxidation by monooxygenases, highlighting the structure and mechanisms of three major enzyme classes, methane monooxygenases (MMOs), non-heme (di-iron) monooxygenases, and cytochrome P450s (P450s).

MMOs enable methanotrophic bacteria to use methane as their sole carbon source, while non-heme (di-iron) monooxygenases and P450s enable microorganisms to grow on medium- and long-chain alkanes. These monooxygenases selectively oxidize alkanes into alcohols as the first

step in hydrocarbon metabolism. Unfortunately, since the majority of these hydroxylases function as a part of a larger enzyme complex and are membrane associated, their potential for industrial applications is limited. For these reasons, we have been engineering well-expressed, soluble, bacterial P450s, in particular CYP102A1 (BM 3) isolated from *Bacillus megaterium,* for small alkane hydroxylation.

In addition to being well-expressed and soluble, BM 3 was chosen as the starting point for this protein engineering effort because it is a rare self-sufficient P450 with its heme (hydroxylase) and reductase domains fused on a single polypeptide. This unique domain architecture has been credited for BM3's high catalytic rate acting on its preferred C_{12} to C_{20} fatty acids substrates. Previous work in our laboratory (by Ulrich Schwaneberg, Edgardo T. Farinas, Anton Glieder, Matthew Peters, and Peter Meinhold) aimed at converting BM3 into a methane monooxygenase applied directed evolution –iterations of mutagenesis, recombination, and screening, to generate BM3 variants with improved alkane hydroxylation activity. Their overall strategy for shifting BM3's substrate specificity was to enhance the promiscuous alkane hydroxylation activity of BM3 and subsequent variants starting with octane as the target substrate. Using a colorimetric screen with *p*-nitrophenoxy octane as a surrogate substrate in combination with monitoring cofactor consumption in the presence of octane, variants with improved activity for octane hydroxylation were identified. Limited activity toward propane, a substrate not hydroxylated by wild-type BM3 was also observed in later generation variants. At this point, selection pressure was shifted toward propane hydroxylation by screening for dimethyl ether demethylation in high-throughput as the surrogate activity. Using this screen and a combination of mutagenesis techniques, ethane hydroxylation activity was obtained with variant 35E11. At this point, I took over the P450 alkane hydroxylase project.

Chapter 2 describes the continuation of laboratory evolution efforts aimed at converting BM3 into a small alkane hydroxylase, starting from variant 35E11. As a result of the previous ten rounds of mutagenesis and screening, variant 35E11 displayed a significantly lower thermostability (ΔT_{50} = – 11.6 ^{o}C) compared to the wild-type enzyme. To reverse this loss in thermostability, which is known to reduce the ability of a protein to acquire beneficial mutations that are destabilizing, known stabilizing mutations from a P450 peroxygenase were grafted onto variant 35E11 singly and in combination. The resulting thermostablized variant was subjected to a domain-based protein-engineering strategy (developed by Rudi Fasan), in which the three domains of BM3 were mutated individually using both random and site-saturation mutagenesis. Beneficial mutations identified through high-throughput screening for dimethyl ether demethylation were verified to improve propane and ethane hydroxylation in the context of the holoenzymes. Using this strategy, re-specialization of BM3 for propane hydroxylation was achieved with variant $P450_{PMO}$, a proficient P450 propane monooxygenase. This variant displays substrate affinity and coupling of cofactor consumption rivaling those of the natural P450s with their preferred substrates. In addition, we were able to demonstrate *in vivo* propane hydroxylation using these BM3 variants in resting *E. coli* cells reaching activities surpassing those reported for natural alkane hydroxylases acting on their preferred substrates.

In Chapter 3, we explored alternative mutagenesis approaches to engineer BM3 for small alkane hydroxylation. Instead of gradually shifting the BM3's substrate specificity by enhancing its promiscuous alkane hydroxylation activity as done previously, we applied several semi-rational library design approaches to mutate the BM3 active site in an attempt to acquire activity for small alkane hydroxylation directly from the wild-type enzyme. From screening of mutagenesis libraries created by combinatorial active site saturation with a reduced set of amino

acids and two structure-based computational library design approaches, we identified variants supporting both propane and ethane hydroxylation. Although, none of the obtained variants reached the level of specialization that was previously obtained with $P450_{PMO}$, the range of obtained propane TON and coupling of cofactor consumption corresponds to those values of generalist intermediates of $P450_{PMO}$ lineage obtained after 10 – 12 rounds of mutagenesis and screening. These results suggest semi-rational library design can be an effective strategy to move away from a specialist enzyme toward generalist enzymes, but functional specialization still requires optimization through several rounds of random mutagenesis and screening.

The BM3 variants we obtained with high activity on small alkanes hydroxylate propane and longer chain alkanes predominantly at the more energetically favorable subterminal position. In contrast, sMMO and other alkane hydroxylases utilized by microorganisms for alkane metabolism selectively oxidize at the terminal carbon to produce 1-alcohols. Since selective terminal hydroxylation has been difficult to achieve by engineering BM3, a sub-terminal hydroxylase, we investigated in Chapter 4 whether a small-alkane terminal hydroxylase could be obtained by directed evolution of a longer-chain alkane hydroxylase that exhibits this desirable regioselectivity. For this study, we engineered two alkane hydroxylases that prefer medium-chain-length alkanes (C_6 – C_{10}), AlkB from *P. putida* GPo1 and CYP153A6 from *Mycobacterium* sp. HXN-1500, for enhanced butane hydroxylation activity using an *in vivo* growth-based selection system (developed by Daniel Koch). This system enabled selection for terminal alkane hydroxylase activity based on enhanced growth complementation of an adapted strain of *P. putida*. The resulting enzymes, AlkB-BMO1, -BMO2, and CYP153A6-BMO1 conferred improved growth on butane as the sole carbon source and exhibited higher rates of 1-butanol production in whole-cell butane bioconversions while maintaining their preference for

terminal hydroxylation. These results demonstrated the usefulness of this *in vivo* selection system, which could be generally applied to directed evolution of enzymes for small alkane hydroxylation.

To apply selection pressure for the main goal of this research, selective hydroxylation of ethane and methane, we developed a high-throughput screen to directly assay for P450 alkane hydroxylation, described in Chapter 5. With the use of a pressurizable 96-well reactor, the P450 alkane hydroxylation reaction was conducted in high throughput and the alcohol product was quantified spectroscopically by a coupled enzyme assay. Applying this screen to BM3 variants generated in our laboratory, we identified variant E31 as the best candidate for further engineering, since it displayed both the highest activity in the screen and wild-type-like thermostability. Subsequent rounds of site-saturation and random mutagenesis resulted in improved variants demonstrating the efficacy of the screen. However, none of the identified BM3 variants were able to produce ethanol or methanol in whole-cell alkane bioconversions using growth-arrested *E. coli* cells. In contrast, CYP153A6, a natural terminal alkane hydroxylase, was able to produce ethanol in whole-cell alkane bioconversions. The inability of BM3 variants to produce ethanol *in vivo* reflects their poor affinity for ethane and indicates they still lag behind a natural P450 alkane hydroxylase in terminal hydroxylation of small alkanes.

The complete absence of methane oxidation activity in numerous BM3 variants evolved for propane and ethane hydroxylation activity led us to question if the P450 oxo-ferryl porphyrin radical intermediate, compound I, can oxidize the 105 kcal/mol methane C-H bond. In Chapter 6, we separated the substrate binding problem presented by the small size of methane from the challenge of the higher activation barrier of the reaction presented by the methane C-H bond by assaying the reactivity of compound I directly through terminal oxidant-supported P450

reactions. Using iodosylbenzene, 3-chloroperoxybenzoic acid, and hydrogen peroxide as oxidants, we investigated the ability of the compound I of five P450s (BM3, $P450_{PMO}$, $P450_{cam}$, CYP153A6, and CYP153A6-BMO1) to hydroxylate alkanes ranging from methane to octane. From these terminal oxidant-supported P450 reactions, we found the compound I of CYP153A6, and CYP153A6 BMO-1 to be able to break the methane C-H bond using PhIO as the oxidant. This demonstrates both the feasibility of P450 methane oxidation and the use of terminal oxidant-supported P450 reactions as an assay to investigate the compatibility of P450 active sites for small alkane oxidation. By chemically generating the active radical, we eliminated the requirement for substrate binding to initiate P450 catalysis, which enabled us to determine the innate substrate range of each P450 active site.

Although the BM3 variants we generated could not hydroxylate methane, we have found other applications for which they excel, such as regioselective hydroxylation of non-activated carbon centers. The involvement of human P450s in the degradation of most drug compounds made us wonder if BM3 variants can be used to replicate or predict the metabolism patterns of human P450s and produce these metabolites on a preparative scale. In Chapter 7, we demonstrated that a small panel of BM3 variants covers the breadth of reactivity of human P450s by producing 12 of 13 mammalian metabolites for two marketed drugs, verapamil and astemizole, and one research compound. The most active enzymes could support preparation of individual metabolites for preclinical bioactivity and toxicology evaluations. Underscoring their potential utility in drug lead diversification, engineered BM3 variants also produce novel metabolites by catalyzing reactions at carbon centers beyond those targeted by animal and human P450s. Finally, we enhanced the production of a specific metabolite by directed evolution of the enzyme catalyst.

Chapter 8 details experimental procedures and materials used throughout the studies described in this thesis.

TABLE OF CONTENTS

FIGURES AND TABLES

ABBREVIATIONS

MMO	Methane monooxygenase
M. c. Bath	*Methyloccus capsulatus* Bath
M. t. OB3b	*Methylosinus trichorium* OB3b
E. coli	*Escherichia coli*
BM3	Cytochrome P450 BM3 (CYP102A1)
A6	CYP153A6
CAM	CYP101
PMO	$P450_{PMO}$
ET	Electron transfer
PCET	Proton coupled electron transfer
KIE	Kinetic isotope effect
SRS	Substrate recognition site
TON	Turnover number
EPPCR	Error-prone polymerase chain reaction
SOEPCR	Splicing by overlap extension polymerase chain reaction
SSM	Site-saturation mutagenesis
CAST(ing)	Combinatorial active site saturation test
NADH	Nicotinamide adenine dinucleotide, reduced form
NAD^+	Nicotinamide adenine dinucleotide, oxidized form
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced form
$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidized form
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
DME	Dimethyl ether
BDE	Bond dissociation energy
DFT	Density functional theory

Chapter 1

Introduction: Enzymatic Alkane Oxidation by Monooxygenases

A. Introduction

Petroleum and natural gas are the primary energy resources currently utilized to meet the world's energy needs (*1*). In addition to its use as a fuel source, the conversion of crude oil to olefins and aromatics through refining has also allowed petroleum to act as a major feedstock for the chemical industry. This ability to generate chemical precursors—through processes such as cracking, dehydrogenation, and reforming—differentiates petroleum from natural gas, which has been limited to usage as a fuel. However, as the world's known reserves of crude oil are shrinking (*2*), the need to find alternative sources for chemical feedstocks, such as natural gas, is becoming more pressing. This search for alternative feedstocks is also motivated by the environmental impact of petroleum refining. As the reactions to produce olefins and aromatics from petroleum are endothermic, CO_2 is released during both the generation of these chemical precursors and in the subsequent partial oxidation steps to produce the desired oxygenated compounds (e.g., aldehydes, alcohols, carboxylic acids).

Methane, the principal component of natural gas, is an ideal alternative to petroleum refining, since it fulfills all the requirements for a chemical feedstock, including high abundance, low cost, and lower carbon footprint (CO_2 emission) compared to petroleum refining. In addition to the methane available in known natural gas and coal sources, it can also be produced via biogas (*3*), by fermentation of organic matter (*3*), and vast quantities are stored as methane hydrates at the ocean floor (*4*). There are also economic incentives to convert methane into oxygenated products, as it is less expensive than petroleum-generated olefins and aromatics. Finally, the methane oxidation reaction is exothermic. Therefore replacing the highly endothermic petroleum refining processes with methane oxidation would also result in concurrent energy production with the chemical products instead of energy consumption. Despite

all these favorable factors, methane is still underutilized as a feedstock owing to a lack of economical and sustainable strategies for its selective oxidation (*5*).

The selective oxidation of methane to oxygenated products represents a significant challenge, as the methane C-H bond is extremely inert (105 kcal/mol) (*6*). Therefore, highly reactive radical or ionic species are required to cleave the methane C-H bond. However, as the desired partial oxidation products, methanol and formaldehyde, have weaker C-H bonds compared to methane, they are susceptible to further oxidation to CO_2. To overcome these challenges, research toward partial methane oxidation and improved methane utilization has taken several different approaches: (1) the one-step oxidation of methane to methanol or formaldehyde, (2) oxidative and non-oxidative coupling of methane, (3) Fischer-Tropsch synthesis of hydrocarbons from synthesis gas (syngas), generated from steam reformation of methane. Currently, industrial conversion of methane to methanol falls into the latter category, utilizing an energy intensive, endothermic, and costly process to first convert methane into syngas, followed by methanol synthesis from this intermediate (*1, 7*). While there is a variety of mixed metal-oxide heterogeneous catalysts capable of the desired methane partial oxidation (*8*) and coupling reactions (*9 – 10*), these catalysts currently lack the reactivity and selectively necessary for commercialization (*5*).

The most hopeful strategy for selective methane oxidation is through electrophilic activation by late transition metal ions, such as Pt(II) (*11*), Pd (II) (*12*), Rh (*13*), and Hg(II) (*14*). These systems are derived from the landmark study by Shilov demonstrating the production of alcohol and alkyl-chloride using Pt(II) salts in aqueous solution (*11*) (see equation (1)). These systems have been shown to be capable of both stoichiometric and catalytic oxidation of methane. Their most attractive feature is a high selectivity for the partially oxidized product: i.e.,

the reactivity for the methane C-H bond is substantially greater than that of a product C-H bond, such as H-CH_2OH or H-CH_2SO_4H (*15*). The mechanism of the Shilov systems occurs in three steps: (1) electrophilic activation of the R-H bond by Pt(II) to form a Pt(II)-alkyl intermediate, (2) oxidation of the Pt(II)-alkyl complex by $[PtCl6]^{2-}$ to give a Pt(IV)-alkyl species, (3) nucleophilic S_N2 attack of water at Pt-C bond results in the formation of the alcohol product and regenerates the Pt(II) catalyst.

$$R\text{-}H + PtCl_6^{2-} \xrightarrow{PtCl_4^{2-}} R\text{-}OH + R\text{-}Cl + PtCl_4^{2-} + 2HCl \quad (1)$$

Advancement of the original system has been made by Periana et al., which has replaced the oxidant $[PtCl_6]^{2-}$ with sulfuric acid (*15*). Using an Hg^{2+}complex in sulfuric acid, a one-pass yield of 40% conversion of methane to methyl hydrogensulfate was obtained at > 90% selectivity (*14*). An improved system utilizing Pt(II) chelated by 2,2'-bipyrimidine, which is more thermodynamically robust, resulted in a one-pass yield of greater than 73% (*15*). While these yields are the highest reported for direct partial oxidation of methane, several key disadvantages have prevented commercialization: low turnover frequency (*16*), costly methanol recovery from concentrated sulfuric acid, and catalyst poisoning by water and oxidation products (*5*).

In contrast to the difficulties for transition metal catalysts to selectively oxidize methane, metalloenzymes, specifically methane monooxygenases (MMOs) with metal centers composed of abundantly available metals (iron and copper) are able to convert methane to methanol at room temperature, atmospheric pressure, in water, and using O_2 as the oxidant (*17*). Alkane hydroxylases, including MMOs, are discussed in detail in the next section. As the structures of these metalloenzymes have become available, they have inspired chemists to make "biomimetic" catalysts (*18*) in attempts to capture the metal-centers in a functional form using a variety of

scaffolds. The synthesis and characterization of multiple di-iron Fe^{IV}=O complexes modeled after the Q intermediate of MMOs have been reported (*19 – 22*). To date, these complexes have been shown to activate C-H bonds as strong as 100 kcal/mol, however, the obtained reaction rates were much lower than those observed with metalloenzymes (*20, 22*).

B. Alkane Oxidizing Enzymes

B.1. Methane monooxygenases (MMOs)

While a catalyst that supports efficient conversion of methane to methanol has so far eluded transition metal chemistry, Nature found a solution to utilize methane as an energy source long ago. Methanotrophic bacteria found in a variety of environments including methane vents in the deep sea, gastrointestinal tracts of cows, and landfills are unique in their ability to utilize methane as their sole carbon and energy source (*23*). Methanotrophs, comprising 13 different genera within the α and γ protobacteria (*24*), are defined by their expression of a methane monooxygenase (MMO) that directly converts methane to methanol. The methanol product is further oxidized to formaldehyde by a methanol dehydrogenase and is used both for biomass synthesis (*23*) and as a source of ATP through further oxidation reactions (*23*).

Most studies of MMOs have been focused on enzymes from *Methyloccus capsulatus* Bath (*M. c.* Bath) and *Methylosinus trichorium* OB3b (*M. t.* OB3b) (*25*). There are two types of MMOs employed by methanotrophs, soluble MMO (sMMO) (*17*) and membrane-bound or particulate MMO (pMMO) (*26*). All but one genus of methanotrophic bacteria express pMMO, and a small subset produces both MMOs (*24*). In methanotrophs expressing both MMO forms, sMMO is expressed when less than 0.8 μM copper is present in the growth medium, whereas

with ~ 4 μM copper present, pMMO is expressed along with the developments of extensive, intracytoplasmic membranes (*27 – 28*).

B.2. pMMO

Particulate MMOs are integral membrane metalloenzymes produced in nearly all methanotrophs and are composed of the three subunits pmoA, pmoB, and pmoC (*26*). The three protomers are arranged in an $\alpha 3 \cdot \beta 3 \cdot \gamma 3$ trimeric complex, Figure 1.1 (*25*). The soluble region of the enzyme complex extends ~ 45 Å from the membrane and is composed of six β-barrels. A significant opening spans the length of the pMMO trimer at its center; this pore is ~ 11 Å wide in the soluble portion and expands to ~ 22 Å within the membrane. Despite decades of research and the availability of two crystal structures (*29*), only recently has the location of the copper active site been identified (*30*). Balasubramanian et al. demonstrated that expression of only the soluble domain of pMMO, pmoB, from *M.t.*OB3b (*31*) was sufficient for methane oxidation (*30*). This study conclusively identified the active site to be a dicopper center with a Cu-Cu distance of 2.5 – 2.7 Å coordinated by three highly conserved His residues (*32*). In light of this discovery, it is puzzling why Nature chose such a large enzyme complex for this reaction, when a soluble sub-domain of pMMO is fully capable of the transformation. One theory forwarded by the authors suggests that the membrane portions may play an important role in increasing the local methane concentration as methane preferentially partitions between the aqueous solution and the membrane (*30*).

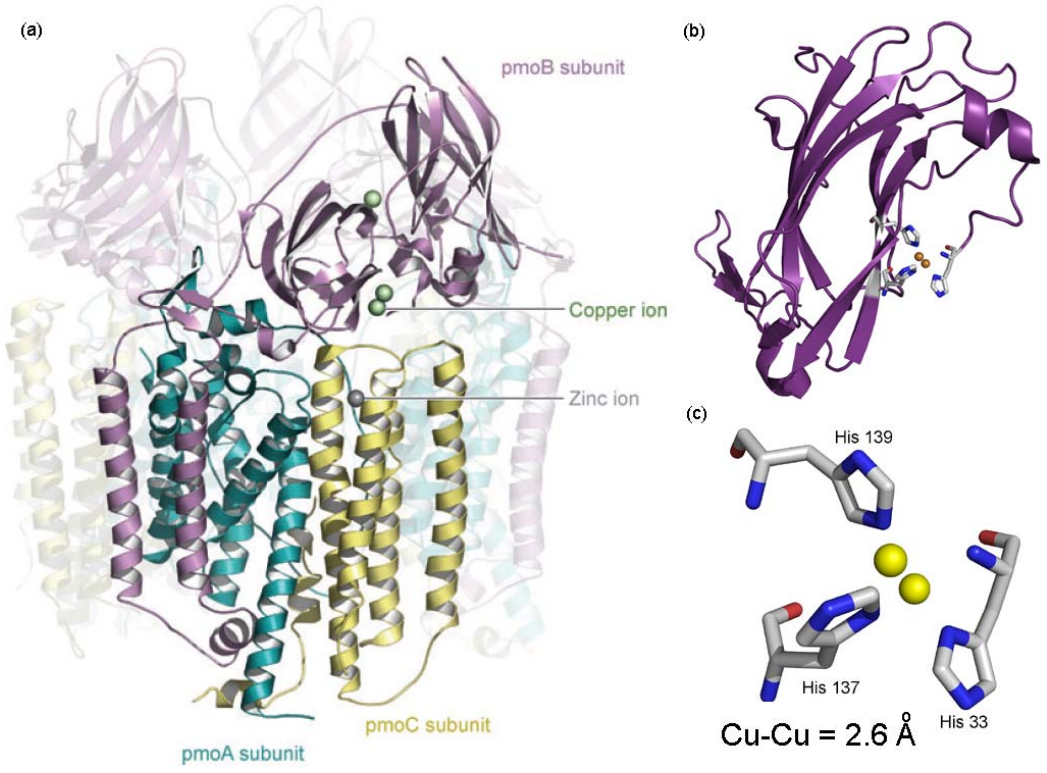


Figure 1.1: The pMMO (*M. c.* Bath) structure (pdb: 1YEW); (a) the full structure with one protomer highlighted, reproduced from ref 29; (b) the soluble domain pmoB; (c) the first coordination sphere of the dicopper metal center

Although pMMO is much more prevalent than sMMO in methanotrophs, difficulties in its characterization due to the fact that it is an integral membrane enzyme have resulted in far less understanding of its biochemistry as compared to sMMO. In fact, the conditions for isolating catalytically active pMMO have been the subject of extensive research, and the optimal conditions still remain unclear. Copper concentration in the growth medium, anaerobicity of the growth condition, and the detergent-protein ratio are among the many conditions that have been shown to affect the measured enzyme activity (*33 – 34*). The *in vitro* characterization of pMMO is further complicated by the absence of a known physiological reductant. Typically, purified pMMO is assayed for propylene oxidation activity using either NADH or duroquinol as the reductant (*35*). Activities ranging from 0.002 to 0.126 U/mg (1 U = 1 µmol propylene oxidized per min) have been reported from various preparations (*32 – 33, 36 – 37*).

pMMO has been shown to oxidize only alkanes and alkenes up to five carbons in length (*38 – 39*). Interestingly, for these multi-carbon substrates, sub-terminal oxidation at the C-2 position is preferred (*40*). Studies using chiral alkanes have given evidence to suggest the pMMO mechanism for oxygen insertion occurs in a concerted fashion rather than involving radical or cationic intermediates, as with sMMO or cytochrome P450s (*41 – 42*). In addition, an absence of a carbon kinetic isotope effect in the oxidation of propane also suggests little or no structural rearrangement occurs at the carbon center during the rate-limiting step (*43*). Unfortunately, attempts to determine the pMMO mechanism have been sparser compared to similar efforts with sMMO, and much of the mechanism is still not well understood.

B.3. sMMO

Due to both its unique ability to oxidize methane as well as its high substrate promiscuity, i.e., the ability to hydroxylate more than 50 different compounds including aromatics (*17, 35*), sMMO has been a favored target for research (*39*). sMMO has been purified from *M. t.* OB3b (*44*), *M. c.* Bath (*45*), and several other strains of methanotrophs (*46*). It belongs to the family of bacterial multi-component monooxygenases (BMMs) (EC.1.14.13.25), which includes toluene monooxygenase, phenol hydroxylase, and alkene monooxygenase (*47*), that enable their hosts to utilize a variety of hydrocarbons as their sole carbon and energy source (*47 – 48*). Using a common carboxylate-bridged di-iron center in their hydroxylase, BMMs are able to activate oxygen for formal insertion into the substrate C-H bond, which initiates the metabolism of these hydrocarbons.

Typical of BMM family members, sMMO is comprised of three components; a hydroxylase (MMOH), which houses the di-iron active site, a reductase (MMOR), which contains a flavin adenine dinucleotide (FAD), and a [2Fe-2S]-ferredoxin (Fd) cofactor that

shuttles electrons from the NADH cofactor to the MMOH active site, and a regulatory protein (MMOB), which is required for methane oxidation (*17*). The MMOH subunit consist of three polypeptides arranged as an α2β2γ2 dimer, Figure 1.2 (a). The di-iron active site is embedded in a four-helix bundle and coordinated by four carboxylates and two imidazoles from two E(D/H)XXH binding motifs.

The resting state of the hydroxylase (H_{ox}) active site is a di(μ-hydroxo)-(μ-carboxylato)diiron (III) species. The catalytic cycle (Figure 1.2 (b)) is initiated by a two-electron reduction to the di-iron (II) form (H_{red}). The reduction occurs simultaneously with a carboxylate shift of the terminally coordinated glutamate (E243), which results in protonation and displacement of both bridging hydroxyl ligands. Rapid reaction of H_{red} with O_2 in the presence of MMOB results in a peroxodiiron (III) intermediate (H_{peroxo}). In the absence of electron-rich substrates, H_{peroxo} rapidly decays into the intermediate Q, a diiron(IV) oxo intermediate with a short Fe-Fe distance of 2.5 Å (*49*). The Q intermediate has been shown to be responsible for the oxidation of a variety of substrates (*50 – 54*) including methane. In the absence of substrate, Q decays slowly to H_{ox} by acquiring two electrons and two protons through a still unknown process. Both H_{peroxo} and Q intermediates have well-defined Mossbauer and optical spectroscopic properties (*53, 55*). The conversion of H_{peroxo} to Q has been shown to be both pH-dependent and exhibiting a solvent kinetic isotope effect (KIE), which indicates that the O-O bond cleavage occurs heterolytically via a proton-promoted mechanism (*56 – 57*).

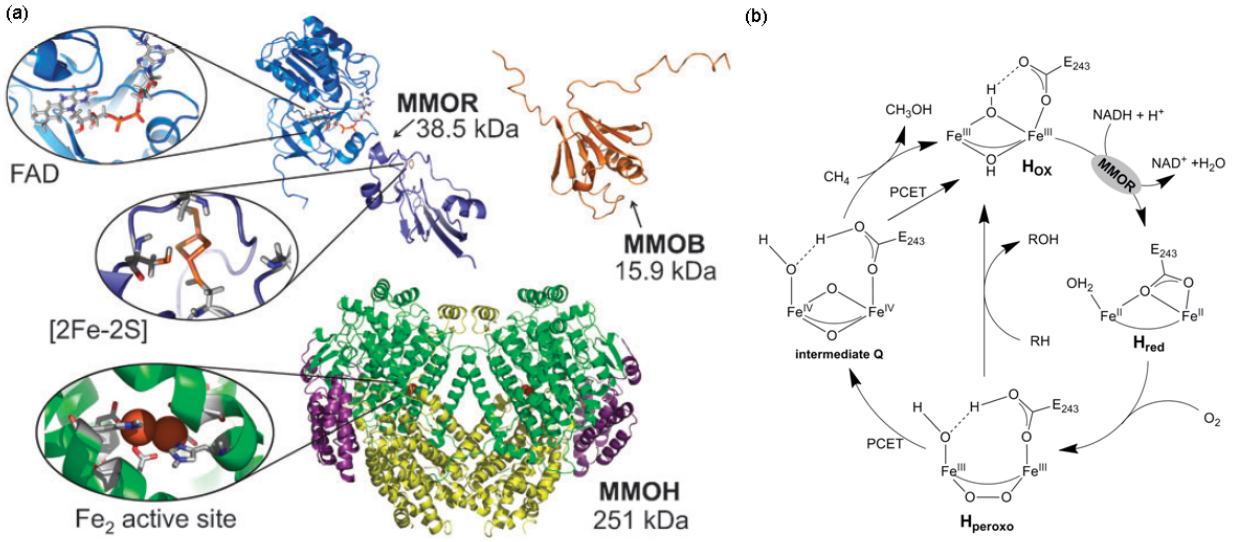


Figure 1.2: The sMMO structure and mechanism. (a) The structure of MMOH (pdb: 1MTY), MMOB (pdb: 1CKV), and MMOR (pdb: 1JQ4) with the cofactors highlighted, reproduced from ref 17. (b) The sMMO catalytic cycle, see text for details, (PCET: proton coupled electron transfer)

Based on density functional theory (DFT) calculation with ~ 100 atoms (*58 – 60*), methane initially approaches the Q intermediate in the ~ 185 $Å^3$ hydrophobic binding site distal to the histidine ligands. The bridging oxygen atom abstracts a hydrogen atom from methane in an outer-sphere, proton-coupled electron transfer reaction, during which one of the iron atoms is reduced to Fe (III). The electron is taken from a C-H σ-orbital, leaving behind a bound methyl radical. The C-O bond formation along with a second electron transfer to the other iron center from the methyl radical occurs either through a rebound mechanism with very short distances (H-O---C of 1.97 Å) or a concerted mechanism with the methyl fragment tightly bound to the hydroxyl group (*61*). These two proposed pathways have comparable activation barriers from DFT calculations, therefore the reaction most likely has a mixed character. The catalytic cycle is then completed with the release of the methanol product, returning the enzyme to its di-iron (III) resting state.

MMOH is only active in the presence of a protein cofactor, MMOB, which when complexed with MMOH changes its structure and reactivity. For example, MMOH from *M. t.* OB3b oxidizes alkanes and nitrobenzene to form secondary alcohols and *m*-nitrophenol products in the absence of MMOB (*62*). Upon MMOB addition, the product ratios shift such that mostly primary alcohols and *p*-nitrophenol are formed. In addition, MMOB must be present for efficient generation of MMOH intermediates in the reaction cycle, which suggests that binding of MMOB initiates the electron transfer (ET) and O_2 binding steps (*63 – 64*). The presence of MMOB has been generally reported to enhance ET between MMOH and MMOR (*63*), but when chemically reduced MMOR was added to premixed solutions of MMOH and MMOB the same ET between MMOH and MMOR was inhibited (*65*). These apparently conflicting results have led investigators to suggest that slow structural changes associated with MMOB and MMOR binding to MMOH may result in hysteresis in MMOH activity (*62*). A current hypothesis is that the interaction of one hydroxylase component of MMOH with MMOR or MMOB could be dependent on the presence of MMOR or MMOB bound to the other component of MMOH (*66*). As a consequence of this dependence, the oxidative phase of the catalytic cycle may only occur at one of the two active sites at a time. This hypothesis has been experimentally verified by observing a ~ 50% maximal conversion of the initial di-iron (II) protein during reactions of MMOH with oxygen (*53*).

The complexity of the interactions between these three enzyme components could be necessary to facilitate and coordinate the transport of the four substrates, hydrocarbon, oxygen, electrons, and protons of the sMMO reaction. The selective trafficking of these substrates to the diiron active site of the hydroxylase is also aided by the presence of biologically well-engineered substrate tunnels and pockets (*67*). Co-crystallization of MMOH with halogenated alkanes, Xe

(*68*), and ω-halogenated primary alcohols (*69*) has revealed the presence of multiple hydrophobic substrate binding pockets that trace a contiguous pathway from the protein surface to the di-iron center. The entry of the substrate appears to pass through several such cavities in its path from aqueous solution to the enzyme's active site (*69*). Finally, as many as eleven binding sites have been identified with Xe, which has similar polarity, water solubility, and van der Waals radius as methane. The binding of these surrogate substrates of methane did not induce significant side-chain displacement in the enzyme; therefore it appears that methane and other sMMO substrates are bound in pre-formed hydrophobic pockets.

Kinetic studies of the oxidation of hydrocarbon substrates by intermediate Q monitored through stopped-flow spectroscopy have shown three distinct substrate classes. The first class of substrates, including ethane, methanol, ethanol, and some ethers, displays a linear dependence of reaction rate on substrate concentration. In addition, a kinetic isotope effect (KIE) of near unity was observed, which suggests that the breaking of the substrate C-H bond is not the rate-determining step. The second class of substrates, including methane and diethyl ether, also displays a linear dependence of reaction rate with substrate concentration but display a KIE > 1, suggesting that C-H bond activation reaction is rate-determining. Finally, the last class of substrates includes nitromethane, acetonitrile, and acetaldehyde, and displays normal Michaelis-Menten kinetics with hyperbolic dependence of reaction rate with substrate concentration and a KIE > 1. For many of the hydrocarbon substrates discussed above, with the exception of methane, the H_{peroxo} intermediate is also a viable oxidant. However, when the H_{peroxo} intermediate is used as the oxidant rather than Q, only class II and III kinetic behavior is observed. This has led some investigators to conclude that reactions with H_{peroxo} proceed through a classical

hydrogen atom transfer mechanism, whereas those of Q are extensively non-classical and involve hydrogen atom tunneling.

This difference could be particularly important for methane oxidation, as methane is kinetically stable with a large barrier height for its oxidation. For the reaction with the Q intermediate, tunneling across this barrier could lead to progression along the reaction coordinate, whereas the reaction with the H_{peroxo} intermediate may not proceed due to absence of tunneling. While this explanation could resolve why sMMO homologs cannot activate methane while possessing nearly the same di-iron active site, unfortunately, KIE studies for the sMMO methane reaction which would determine if tunnel effects were present have yielded varied results. Under single-turnover conditions, KIE values of 23 to 50 have been reported (*50, 70*), which indicates proton tunneling in the transition state. However, under steady-state conditions, a KIE of only 1.7 was observed, when comparing V_{max} (or k_{cat}) values (*70 – 71*), which suggests an absence of tunneling.

Further complicating the sMMO reaction mechanism is the fact that, while the rate-determining step is thought to be the hydrogen atom transfer, multiple studies have revealed that there is no correlation between the reaction rate of a given substrate with the Q intermediate and its homolytic bond dissociation energy (BDE). For example, the oxidation rates of sMMO for methane and ethane are nearly identical despite a BDE difference of ~ 4 kcal/mol. Another example would be a comparison between acetonitrile and nitromethane, which have similar homolytic and heterolytic BDEs, but display a 62-fold difference in reaction rates at 4 ^{o}C (*72*). Reconciliation of the KIE results that indicate the hydrogen abstraction to be rate limiting and the lack of correlation between substrate BDE and oxidation rate remain a challenge.

B.4. Using methanotrophs/MMOs for methanol synthesis

While methanotrophs and MMOs have been focus of extensive research over the past decades, successful attempts to use either the organisms or enzymes for methanol synthesis have been sparse. The inability to express either pMMO or sMMO in a heterologous host severely limits their utilization in industrially relevant organisms as well as the ability to use standard molecular biology methods to engineer desired protein properties. In addition, the multi-component nature of MMOs is also a hindrance to evolving more active or more stable variants. One successful strategy for methanol biosynthesis using methanotrophs is to inhibit the downstream enzyme in methanol metabolism, methanol dehydrogenase (MDH). Using NaCl as a MDH inhibitor, 7.7 mM of methanol were accumulated in *M. t.* OB3b cultures after 20 hours (*73*). Optimization of the growth conditions as well as the addition of ethylene diamine tetra-acetic acid to further inhibit MDH resulted in 13.2 mM methanol accumulation after 12 h batch fermentations with an overall activity of 0.036 U/mg cell mass (1 U = 1 μmol methanol/min). While this strategy is successful in producing methanol, significant yield improvements and reduction of the product loss to the natural methanol metabolism of the methanotroph host are hard to envision.

Studies of the sMMO mechanism as well as its crystal structure have also inspired researchers to make biomimetic catalysts replicating the same carboxylate bridged di-iron core as sMMO stabilized with a variety of ligands (*19, 74*). While advances in ligand design have led to catalysts which can reach the equivalent H_{peroxo} and Q intermediate states in the sMMO catalytic cycle, the obtained reactivity with alkane substrates has been modest, with no reported methane activity (*19, 74*). A key obstacle in reaching methane oxidation activity for these biomimetics could be an intrinsic inaccuracy in the structural model they are attempting to emulate. As all

available crystal structures of MMOH have been solved in the absence of MMOB, which modulates the MMOH tertiary structure directly affecting both substrate access and the first coordination sphere of the diiron center. It is therefore questionable if the observed active site configurations reflect that of the active configuration during methane oxidation.

B.5. AlkB and non-heme di-iron alkane monooxygenases

Expanding the search for potential methane biocatalysts beyond MMOs, two other class of enzymes, non-heme di-iron alkane monooxygenases and cytochrome P450s, are also able to activate oxygen and perform O-atom insertion into inert alkane C-H bonds. The family of non-heme di-iron alkane hydroxylases has been identified in bacteria and fungi utilizing $C_5 - C_{16}$ *n*-alkanes as their sole carbon source (*75*). Exemplified by the most studied alkane hydroxylase isolated from *Pseudomonas putida* GPo1, the non-heme di-iron alkane hydroxylase is a three-component system consisting of (1) a soluble NADH-rubredoxin reductase (AlkT) (*76*), (2) a soluble rubredoxin (AlkG) (*77*), and (3) the integral membrane oxygenase (AlkB) (*78 – 79*). Although AlkB can be functionally expressed in *Escherichia coli* as lipoprotein vesicles, purification and maintenance of activity in the purified state is difficult, which has limited its mechanistic and structural analysis (*80*).

Through alanine scanning mutagenesis, an eight-histidine motif has been shown to be necessary for AlkB function and presumably is responsible for coordination the di-iron core (*81*). This motif represents a class of di-iron centers that is shared with desaturases, epoxidases, decarbonylases, and methyl oxidases, and differs from the carboxylate bridged di-iron center of sMMO (*81*). However, Mossbauer studies of the AlkB metal center revealed similar features as sMMO, with characteristics of an antiferromagnetically coupled pair of Fe (III) ions in its resting state (*82*). The di-iron cluster also becomes high-spin diferrous following reduction and can be

quantitatively oxidized back to its resting state by enzymatic turnover in the presence of substrate and oxygen (*82*). Further evidence for the similarities between the AlkB and sMMO mechanisms has been provided through studies with the use of norcarane as a chemical probe (*83*). From these studies, the AlkB reaction has been shown to be consistent with an oxygen-rebound mechanism via a substrate-centered radical, analogous to the proposed P450 and sMMO mechanisms, exhibiting limited rearranged products (*83*).

The ability to functionally express AlkB heterologously in *E. coli* certainly makes it a potentially better industrial biocatalyst compared to MMOs and also more amenable to enzyme engineering. However, the integral membrane nature of AlkB limits the enzyme's expression to the available membrane surface area. In addition, the lack of a crystal structure and knowledge of both the second coordination sphere of the diiron center and the component interactions are significant hindrances to directed evolution efforts to shift the AlkB substrate range from C_5-C_{16} alkanes to methane.

B.6. Cytochome P450s

Cytochrome P450s, which utilize a thiolate-ligated heme (iron protoporphyrin IX) prosthetic group in their active sites (*84*), represent an entirely different solution to diiron centers for catalytic oxygen insertion into C-H bonds. Unlike MMOs and non-heme diiron alkane hydroxylases, which are only found in methanotrophs and alkanotrophs, the superfamily of cytochrome P450s is one of the most prevalent enzyme families found across all three domains of life. To date, over 10,000 P450 enzymes have been identified (data source: http://drnelson.utmem.edu/CytochromeP450.html). P450s are involved in the metabolism of xenobiotics and the biosynthesis of signaling molecules. In the first role, P450s serve as a protective mechanism for the degradation of exogenous compounds by introducing polar

functional groups to facilitate further metabolism or excretion. This defense mechanism is particularly prominent in plants, which require P450s to break down herbicides due to their immobile nature (*85 – 86*). This is exemplified by the presence of over 400 P450 genes in rice (*87*). In their other role, P450s are responsible for synthesis of a variety of steroid hormones and the conversion of polyunsaturated fatty acids to biologically active molecules implicated in development and homeostasis.

The defining reaction P450s is the reductive activation of molecular oxygen as it is one of the few oxygenases possessing the requisite "$Fe^{IV}=O^{\cdot+}$" state for alkane C-H bond activation. In this reaction, one oxygen atom is inserted into the substrate while the other is reduced to water. The overall equation for the reaction is $RH + NAD(P)H + O_2 + H^+ \rightarrow ROH + NAD(P)^+ + H_2O$, where RH is the substrate. In addition to this canonical reactivity, due to the existence of multiple oxidants in the P450 catalytic cycle, P450s can also catalyze epoxidation, dealkylation, sulfoxidations, desaturation, carbon-carbon bond scission, and carbon-carbon bond formation among other known reactivities (*88 – 89*).

Most P450s are membrane bound just as MMOs and alkane hydroxylases and thus are relatively difficult to manipulate. Fortunately, many bacterial P450s are soluble, monomeric proteins, and as a result, they have been the focus of early research. In particular, the prototypical enzymes CYP101 ($P450_{cam}$) from *Pseudomonas putida* (*90 – 91*) and CYP102A1 (BM3), a natural fusion enzyme from *Bacillus megaterium* in which the flavoproteins required for electron transfer and the hemeprotein are on a single polypeptide chain (*92*), provided much of the structural and mechanistic information of P450s. Recent interest in developing industrially useful P450 catalysts has also focused on enzymes from thermophilic organisms, including CYP119

(*93*), CYP174A1 (*94*), and CYP231A2 (*95*) as well as BM3 for its unique self-sufficiency and high catalytic rates (*96 – 100*).

B.7. P450 structure

The overall P450 fold (Figure 1.3 (a)) is retained across the enzyme superfamily even though members can share less than 20% sequence identity (*101*). The core four-helix bundle composed of three parallel helices (D, L, and I) and the antiparallel E helix are conserved in all P450s (*102*). The prosthetic heme group is ligated to the absolutely conserved cysteine located on a loop containing a highly conserved FxxFx(H/R)xCxG binding motif. This thiolate ligation gives rise to the 450 nm Soret absorbance maximum for the ferrous-CO complex for which P450s were named (*103*). The other common feature among P450s is a kink at the center of the I helix, which contains the amino acid sequence (A/G)Gx(E/D)T that has been implicated in oxygen binding and protonation (*104 – 105*).

Although the P450 fold is highly conserved, there is sufficient structural diversity to accommodate the binding of significantly different substrates ranging from ethanol in CYP2E1 (*106*) to large peptide antibiotics in CYP165C1 (*107*). In addition, since as few as one mutation can alter enzyme reactivity and selectivity, P450 family members (sharing at least 60% sequence identity) can have very different reactivities (*108*). P450 substrate binding occurs in an induced-fit mechanism accompanied with large (~ 10 Å) shifts in the flexible protein regions (*109*). As the substrate is embedded in the protein core, it interacts with various protein regions, which results in a large set of substrate recognition sites (SRS). Six SRSs have been found to be common to P450s (*110*): the B' helix region (SRS1), parts of the F and G helices (SRS2 and SRS3), a part of the I helix (SRS4), the K helix β2 connecting region (SRS6), and the β4 hairpin (SRS5).

B.8. P450 catalytic mechanism

The P450 hydroxylation mechanism is well understood and can be described as depicted in Figure 1.3 (b). The P450 catalytic cycle is initiated by substrate binding, which displaces the distal water ligand of the resting low-spin (LS) state of the Fe (III) heme (**1**) resulting in a high-spin (HS) substrate bound complex (**2**). The HS Fe (III) has a more positive reduction potential, which triggers electron transfer from the P450 reductase producing a ferrous intermediate (**3**) (*111*). Oxygen readily binds to the ferrous iron center leading to the formation of an oxy-P450 complex (**4**), which is the last stable intermediate in this cycle. A second electron transfer, usually the rate-limiting step of the catalytic cycle, results in a ferric hydroperoxo anion (**5**), which after protonation yields a ferric hydroperoxo complex (**6**). A second protonation at the distal oxygen followed by heterolytic cleavage of the O-O bond leads to the release of water and the formation of the oxo-ferryl porphyrin radical intermediate referred to as "Compound I" (CMP I) (**7**). CMP I then transfers an oxygen atom to the substrate, following a hydrogen abstraction-radical rebound mechanism (*112*) generating the alcohol product and returning to the Fe (III) resting state. The intermediates of this catalytic cycle have common features with peroxidases, cytochrome oxidases, and non-heme di-iron oxidases.

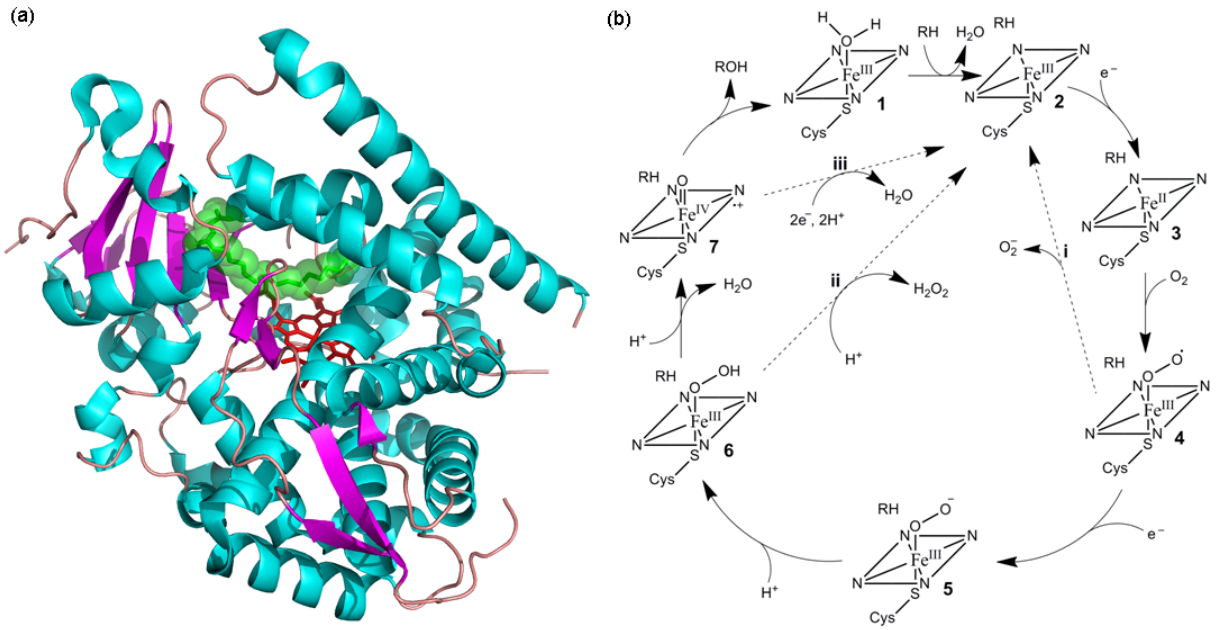


Figure 1.3: Cytochrome P450 structure and catalytic cycle; (a) the structure of the hydroxylase domain of CYP102A1 (BM3) with the heme shown in red and the substrate N-palmitoyl glycine substrate shown in green (pdb: 1JPZ); (b) the P450 catalytic cycle (see text for details)

In addition to having multiple distinct intermediates, many of which are also viable electrophilic and nucleophilic oxidants (*113*), the P450 catalytic cycle contains three branch points (*114*). These three abortive reactions are (**i**) autooxidation of the oxy-ferrous intermediate (**4**) with the release of a superoxide anion and returning the enzyme to its resting state (**2**), (**ii**) a peroxide shunt, where the coordinated hydroperoxide anion (**6**) dissociates, completing an unproductive two-electron reduction of oxygen, and (**iii**) oxidase uncoupling, where the CMP I (**7**) is oxidized to water instead of product formation, which results in a four-electron reduction of oxygen with the formation of two water molecules. These processes are generally referred to as uncoupling, which often occurs in reaction with non-natural substrates that are bound insufficienly to properly regulate solvent/proton access to the active site (*89 – 90*). These pathways are also prominent in eukaryotic P450s involved in host defense responses to xenobiotics through reactive oxygen species generation.

The electrons for the reduction step of the P450 catalytic cycle are provided by either (a) cytochrome P450 reductase (CPR), a soluble flavoprotein with FAD and FMN prosthetic groups, or (b) an iron-sulfur protein that shuttles electrons from a flavoprotein with a single FMN prosthetic group, or (c) a P450 reductase-like domain fused to the P450 heme domain. In each case, the electron donor uncouples the two electrons provided by NAD(P)H and transfers them singly to the P450 enzyme. Since the final reducing agent for the catalytic cycle is NAD(P)H, which has a midpoint potential of -320 mV (*115*), the resting state of the heme iron with a midpoint potential of ca. -300 mV (*116 – 117*) is reduced slowly in the absence of substrate. The substrate binding event triggers a change in the spin state of the heme iron from LS to HS, which induces a positive shift of 100 to 300 mV in the heme reduction potential allowing for rapid electron transfer (*118*). This mechanism clearly acts as a safeguard against the unproductive consumption of NAD(P)H and the formation of superoxide and peroxides. This substrate-induced initiation of electron transfer represents a specific P450 regulatory mechanism and is a clear departure from the initiation of the MMO catalytic cycle through binding of MMOB.

The P450 proton relay mechanism composed of several water molecules stabilized in the P450 active site as well as an acid-alcohol pair of amino acids (CYP101:Thr252, Glu366 CYP102: T268, Glu409) is equally important to P450 catalysis. This relay along with the electron transfer mechanism regulates the production of reactive intermediates and controls the flux of species into the branching points between productive and nonproductive pathways (*119*). For example, mutation of the conserved threonine in $P450_{cam}$ to a hydrophobic residue resulted in near normal rates of cofactor oxidation, but was accompanied only by the release of hydrogen peroxide as the mutant could not effectively cleave the O-O bond without proper protonation (*120*). Coupling of product formation with cofactor consumption was restored by mutating this

position to amino acids capable of hydrogen-bonding interactions. The function of this proton delivery network is also dependent on substrate binding. For non-natural, poorly fitting substrates, their binding is insufficient to expel excess water from the active site, and protonation of the hydroperoxide anion (**6**) can occur at the proximal position, resulting in peroxide release. In fact, the uncoupling of proton and electron transfer does not even require a poorly fitting substrate; simply blocking the site of hydroxylation with fluoro-groups is sufficient to result in normal cofactor consumption with only water or peroxide production (*121*).

B.9. P450 substrate binding and substrate specificity

As mentioned previously, P450 substrate binding occurs by an "induced fit" model as proposed by Koshland (*122*) in which the enzyme accommodates different substrates in its active site by virtue of having a high level of flexibility to undergo appropriate conformational changes. Comparison of the X-ray structures of cytochrome P450s crystallized in substrate-free and substrate-bound forms (*109*) shows large structural rearrangements induced by substrate binding, which suggests that the SRSs are quite flexible and can provide a variety of substrates access to the heme. The absence of charged and hydrogen-bonding groups in the typical P450 substrate, as well as in the active sites of most P450 enzymes, requires such binding mechanisms as an alternative means to stabilize the substrate-enzyme complex. In many cases, different substrate analogues bind tightly to P450 enzymes simply due to their poor solubility in water rather than the presence of specific interactions with active site residues (*123*).

Given this general mode of substrate binding, it is unsurprising to find that the P450 specificity for substrate hydroxylation can be readily determined by three factors: (a) the affinity of the substrate for the P450 active site, which is largely determined by the substrate lipophilicity, (b) the intrinsic reactivity of the individual C-H bond in the substrate as determined

by the C-H bond strength, and (c) steric constraints imposed by the active site geometry. While the compatibility of a substrate within a P450 active site and steric constraints of binding modes are case specific, lipophilicity has been shown to be directly correlating to K_M or K_d for sets of similarity structured compounds (*124 – 125*). The general preference of P450 oxidation occurs with the following order of C-H bonds: tertiary>secondary>primary, which was determined using several small molecular probes that minimized the effect of the P450 active site structure in controlling the site of oxidation (*126*). This preference is reinforced by DFT calculations for the activation barriers for hydrogen abstraction, which predict a similar reactivity preference: benzylic or allylic>tertiary >secondary>primary (*127 – 128*).

B.10. P450 reactions using terminal oxidants

In addition to the normal P450 "turnover" conditions utilizing oxygen and NAD(P)H, the P450 catalytic cycle can also be accessed through the branching/shunt pathways using a variety of terminal oxidants including hydrogen peroxide, alkyl peroxides, acyl peroxides, and iodosobenzene. Early studies with alkylperoxides provided evidence for the formation of a ferric alkylperoxo complex (Fe^{III}-OOR) (*129*) as well as a compound II-like ferryl (Fe^{IV}=O) species, which is one oxidation equivalent higher than the resting ferric state (*130 – 131*) but little evidence for the formation of a CMP I-like ferryl porphyrin radical. However, recent works have confirmed the formation of both a compound-II – like and a compound-I – like species as transient intermediates (*132 – 133*).

The obvious advantage between reactions utilizing O_2/NAD(P)H vs. peroxide is oxygen binding to a Fe^{2+} heme center vs. peroxide binding to Fe^{3+} heme center. The difference in the redox state between oxygen and peroxide eliminates the need for two reduction steps in the peroxide-driven pathway. However, the efficiency of this mode of reaction is generally poor due

to the intrinsically destructive nature of peroxides as well as the lack of acid-base catalytic residues in the P450 structure, with the exception of P450s which naturally utilize peroxides as their oxidant (*134*). In natural peroxidases, such as chloroperoxidase, the formation of the ferryl-oxo intermediate involves proton transfer from the proximal to the distal oxygen atom of the bound hydrogen peroxide, which is aided by a conserved His-Arg or His-Asp amino acid pair (*135*). P450s have highly hydrophobic active sites that lack these acid-base catalytic residues in close proximity to the oxygen binding pockets.

The peroxide-driven P450 reactions proceed through the formation of CMP 0, which after protonation and heterolytic O-O cleavage generates CMP I. In contrast, P450 reactions driven with iodosobenzene (PhIO) produce only CMP I as an oxidant without any potential involvement of peroxo-iron species, since PhIO is a single oxygen donor (*136*). The initial finding of solvent oxygen incorporation through experiments with ^{18}O-labeled water in reactions supported by PhIO led researchers to question if the oxidation proceeded via a ferryl intermediate (*137 – 138*). However, subsequent work has shown that the oxygen of PhIO readily exchanges with the medium through a porphyrin-oxidant complex, $[(Porp)Fe^{III}\text{-}OIPh]^{+}$ (*139*). Nevertheless, whether PhIO-meditated reaction is a faithful mimic of the P450 reaction remains contentious due to differences observed in regio- and chemoselectivities (*136*) and kinetic isotope effects (*140*) between reactions supported by PhIO and NAD(P)H/O_2 (*141*).

B.11. H-abstraction and mechanistic comparisons with MMOs

Similarities between the mechanisms of P450s, di-iron non-heme alkane hydroxylases, and sMMO have long been recognized (*142 – 144*), as the active oxidants of these enzymes, a μ-oxo-diiron (IV) intermediate called compound Q for sMMO and an oxo-iron (IV) porphyrin π-cation-radical called CMP I for P450s, share the same net oxidation state. This is unsurprising

considering the energy requirements for breaking the inert alkane C-H bonds. In the consensus radical rebound mechanism for the O-atom insertion step, the ferryl oxygen initially abstracts a hydrogen from the substrate, leaving a carbon radical, which in turn recombines with the oxo radical coordinated to the iron atom (*145*). This mechanism is supported by observed large intramolecular isotope effects as well as the partial loss of stereo-chemistry at the carbon center for reactions with chemical probes (*146 – 148*).

Complexities in the radical rebound mechanisms arising from the mixed-spin nature of the transition state during the H-atom abstraction by the ferryl-oxo intermediate have been explained by DFT calculations (*149*). For the P450 mechanism, the reaction can proceed through a the low (doublet)-spin state, where the unpaired electron residing on the substrate after H-atom abstraction has an opposite spin to the electron in the iron-hydroxyl orbital. In contrast, the oxidant can also be in a high (quartet)-spin state, such that the substrate-based radical has the same spin as the P450 iron-hydroxyl species. As the collapse of the low-spin pathway lacks the spin-inversion barrier of the high-spin state, it proceeds without a barrier, and H-atom abstraction and the radical rebound can be considered to proceed in a concerted fashion. This mixed-spin transition state model with two distinct pathways for the H-atom abstraction-rebound mechanism has been able to reconcile seemingly contradictory radical lifetime experiments as well as differences in reaction KIEs (*149*). In addition to having a LS triplet state and a HS quintet state, the reaction with diiron metal centers (i.e., sMMO, AlkB) is further complicated by two possible angles of approach of the substrate C-H bond. In contrast to P450s, where the presence of the porphyrin prevents side-on or equatorial approaches, both end-on and side-on approaches are possible with di-iron ferryl-oxo species. As a consequence, both linear and bent Fe-O-H geometries are possible for the transition state (so that electrons can be transferred to both σ and

π orbitals), which results in four distinct reaction pathways with intermediates: $^{3}TS_{\pi}$, $^{3}TS_{\sigma}$, $^{5}TS_{\pi}$, and $^{5}TS_{\sigma}$.

Regardless of the spin-state of the reaction pathway, the transition state for the H-abstraction step presents the largest barrier in the reactions involving CMP I, or Q (*58, 127*). For methane, this barrier height is 26.7 kcal/mol for P450 CMP I, which is significantly higher than the ca.19 kcal/mol barrier for known P450 substrate camphor. The barrier heights for other small gaseous alkanes, ethane (21.6 – 21.8 kcal/mol) and propane (terminal: 21.6 kcal/mol, subterminal: 19 kcal/mol) are also much lower than methane (*127, 149 – 151*). As a comparison, this transition state barrier for H-abstraction from a methane C-H bond has been calculated to be as low as 13.8 kcal/mol (*60*) and as high as 23.2 kcal/mol (*152*) for sMMO. While these calculations show the barrier heights for P450 hydroxylation of known substrates and sMMO hydroxylation of methane are comparable, it is difficult to draw meaningful conclusions for the potential of P450s to oxidize methane.

C. Cytochrome P450 Enzyme Engineering

Due to their involvement in drug metabolism and their ability for regio- and stereoselective hydroxylation on a variety of substrates, numerous potential applications have been suggested for P450s (*153 – 155*). While, undoubtedly many industrial whole-organism biosyntheses involve P450s in their pathway (*156*), direct/explicit use of P450s for biotechnology applications has been scarce. The most publicized example of incorporation of P450s into practical products is the generation of transgenic "blue" carnations from petunias (*157*) and for enhanced production of "blue" roses (*158*). Examples for the use of P450s in the synthesis of high value pharmatheutical therapeutics include the production of hydrocortisone (*159*), cortisone (*160*), Pravastatin (*161*), and Artemisinin (*162*).

The limiting factors toward broader utilization of wild-type P450s as biocatalysts include poor expression, being membrane-bound, and lacking known redox partner proteins. For these reasons, protein engineering efforts have been focused on expanding the substrate range of well-expressed bacterial P450s with known redox partners to accept desired target compounds. Much of this work has been focused on $P450_{cam}$ and BM3 (*163*). $P450_{cam}$ is a type I P450 requiring a putidaredoxin containing a [2Fe-2S] cluster and an FAD-containing putidaredoxin reductase for the highly stereoselective hydroxylation of its physiological substrate (1R)-camphor to (1R)-5-*exo*-hydroxycamphor. BM3, as mentioned previously, is a type II P450 with its FAD and FMN containing reductase fused to its heme domain as a single polypeptide chain (*92*). BM3's native function is believed to be the detoxification of polyunsaturated fatty acids (*164 – 165*). Its unique domain architecture has been credited for its high hydroxylation and epoxidation rates (~ 17,000 min^{-1} (*166*)) on long-chain fatty acid substrates (*92, 166*). In addition, these two proteins were also the first P450s of which crystal structures in the presence (*167 – 168*) and absence of substrates (*169 – 173*) were solved, which has also aided protein engineering efforts.

C.1. $P450_{cam}$

Protein engineering of $P450_{cam}$ has been successful in switching its substrate specificity (*174 – 179*) through the introduction of point mutations to the active site. This strategy has been particularly successful because the enzyme backbone remains relatively fixed during catalysis (*119*). With just one to three active site mutations, $P450_{cam}$ variants have been generated to hydroxylate (+)-R-pinene, a structural relative to (+)-camphor (*174*), aromatic compounds such as ethylbenzene (*179*), diphenylmethane, phenylcyclohexane, naphthalene, pyrenes (*175, 178, 180 – 182*), phenanthrene, fluoranthene, (*183*), polychlorinated benzenes (*184*), and indole to form indigo (*185*). Although improving the promiscuous activity of $P450_{cam}$ for non-natural

substrates has proved to be relatively easy, i.e., requiring just a few mutations, the resulting variants generally exhibited poor coupling between cofactor consumption and product formation, typically ranging from 5% to 32%.

Far fewer examples exist for $P450_{cam}$ variants with nearly wild-type coupling efficiency for non-native substrates, as these variants require multiple rounds of mutagenesis. For example, simply introducing a Y96F mutation will increase $P450_{cam}$'s ability to epoxidize styrene 25-fold, with 32% coupling compared to only 7% coupling for the wild-type (*186*). Placing an additional V247L mutation in the active site increases the coupling efficiency to 60% (*175*). The best example for the ability of $P450_{cam}$ to be engineered to accept new substrates with high coupling efficiency was provided by Xu et al. *(187)* with the engineering of $P450_{cam}$ to hydroxylate short-chain alkanes. Over the course of multiple studies (*176, 187 – 188*), the authors gradually decreased the active site volume by incrementally introducing bulky, hydrophobic residues. A $P450_{cam}$ quadruple mutant, F87W/Y96F/T101L/V247L, was able to oxidize butane at 750 min^{-1} with 95% coupling compared to 4% for wild type (*176*). To achieve 85% coupling of product formation with NADH consumption for propane oxidation, five additional mutations, L244M/L294M/T185M/L1358P/G248A, were required (*187*). This variant with nine active site mutations was also able to oxidize ethane at 78.2 min^{-1} with 10.5 % coupling (*187*).

Finally, multiple attempts have been made to fuse the electron transfer components of $P450_{cam}$ to its hydroxylase domain to increase enzyme activity (*189 – 190*). In addition, fusion of the $P450_{cam}$ hydroxylase domain with reductases from naturally occurring self-sufficient P450s such as P450RhF isolated from *Rhodococcus sp.* NCIMPB 9784 (*190*) has also resulted in self-sufficient enzymes. However, while these constructs are often superior to free enzymes at a 1:1:1 ratio (hydroxylase: putidaredoxin reductase: putidaredoxin) in rates of electron transfer and

product formation, they fall short of the optimal free enzyme activity obtained at higher ratios of electron transfer components, i.e., 1: ≥ 4 : ≥ 12 (*121*).

C.2. CYP102A1 (P450BM3)

BM3 is one of the most studied and frequently engineered P450s, as it has the fastest known catalytic rate for a P450 and is a natural fusion protein with a type II reductase and its hydroxylase found on a single polypeptide (*134*). Its high rate of catalysis has been shown to be the result of BM3 domain architecture: when the heme and reductase domains are expressed independently and mixed together, the resulting activity is severely diminished (*191*). In addition to the self-sufficiency, BM3 can also be easily manipulated genetically and expressed at levels up to 1 g/L in laboratory strains of *Escherichia coli* (*192*). In contrast to $P450_{cam}$, the BM3 protein backbone undergoes large structural changes (> 10 Å) during catalysis as revealed by differences between the substrate-free crystal structure (*168*) and structures of BM3 complexed with known substrates (*168 – 169*). The large active site volume reflects the poor regioselectivity with which BM3 hydroxylates its preferred C_{12}-C_{20} fatty acids substrates (*193*). All these factors make BM3 well suited for directed evolution experiments and potential use as a biocatalyst.

Much like $P450_{cam}$, BM3 has been engineered to accept a variety of substrates. Most rational engineering efforts of BM3 targeted hydrophobic residues lining the substrate-binding channel as well as active site residues immediate to the heme center. By mutating the two residues (R47, Y51) located at the opening of the substrate pocket that stabilizes the carboxylic acid moiety of the preferred fatty acid substrate, BM3 variants were isolated with increased activity for oxidation of alkyltrimethyl ammonium compounds (*194*), hydroxylation of shorter chain fatty acids (*195*), and epoxidation of the anti-malarial drug-precursor amorphadiene (*196*). Introducing additional active site mutations to these variants yielded a variant (R47L/Y51F

/A264G) with activity toward fluoranthrene (*177*) and alkoxyresorufins (*197*) and a variant (R47L/Y51F/F87A/A264G) with activity toward pyrene (*177*). Screening with a simple NADPH consumption assay to monitor cofactor consumption in the presence of a substrate, directed evolution of BM3 yielded a highly promiscuous variant F87V/L188Q/A74G with enhanced activity for a variety of substrates such as indole, alkanes, arene, and polycyclic, aromatic hydrocarbons (*198 – 200*). Similar engineering efforts generated variants with activity for β-ionone, a carotenoid intermediate (*201*), and valencene for (+)-nootkatone production (*202*).

From these rational engineering efforts, general structural function relationships have emerged, which have aided further engineering efforts. For example, active site residue F87, which is positioned directly between the bound substrate and the heme center, has been shown to affect both the substrate specificity and regioselectivity of fatty acid hydroxylation (*203*). Introduction of the F87V mutation converted BM3 into a regio- and stereoselective arachidonic acid epoxygenase (*203*). This mutation along with F87A also increases the oxidation activity for a variety of aromatic compounds (*204*). This improved affinity for aromatic compounds is consistent with the removal of the F87 phenyl side chain directly adjacent to the heme center. Building on these studies, Pleiss and coworkers constructed a focused library targeting residues F87 and A328, allowing for a restricted set of non-polar amino acids (A,V, F, L, I). From this library, variants with activity for the oxidation of linear terpenes, cyclic monoterpenes, cyclic sesquiterpenes (*99*), cyclo-octane, cyclodecane, and cyclododecane (*205*) were found.

Work in our group has been focused on evolving BM3 to accept small alkanes as substrates with the ultimate target of methane hydroxylation. Starting from wild-type BM3, we enhanced its promiscuous activity for octane hydroxylation by applying a colorimetric screen for the hydroxylation of *p*-nitrophenoxy octane to evaluate enzyme libraries generated by random

mutagenesis (*206*). After multiple rounds of random mutagenesis, variant 139-3 was obtained, supporting 1,000 turnover number (TON) on octane, a 6.7-fold increase compared to wild type and measurable activity for propane oxidation (500 TON) (*207*). Selection pressure was then shifted toward propane hydroxylation through the use of a colorimetric screen based on dimethyl ether demethylation (DME) (*208*). Subsequent rounds of mutagenesis yielded variant 35E11, with 17 total amino acid mutations, supporting 6,000 propane TON and 250 ethane TON (*209*). In addition to small alkane hydroxylation activity, other BM3 variants generated in the 35E11 lineage were found with (1) regioselectivity for terminal hydroxylation of octane (*210*), (2) stereoselective secondary hydroxylation of linear alkanes (*208*), and (3) stereoselective epoxidation of alkenes (*211 – 212*).

D. References

1. Arakawa, H., Aresta, M., Armor, J. N., Barteau, M. A., Beckman, E. J., Bell, A. T., Bercaw, J. E., Creutz, C., Dinjus, E., Dixon, D. A., Domen, K., DuBois, D. L., Eckert, J., Fujita, E., Gibson, D. H., Goddard, W. A., Goodman, D. W., Keller, J., Kubas, G. J., Kung, H. H., Lyons, J. E., Manzer, L. E., Marks, T. J., Morokuma, K., Nicholas, K. M., Periana, R., Que, L., Rostrup-Nielson, J., Sachtler, W. M. H., Schmidt, L. D., Sen, A., Somorjai, G. A., Stair, P. C., Stults, B. R., and Tumas, W. (2001) Catalysis research of relevance to carbon management: Progress, challenges, and opportunities, *Chem. Rev. 101*, 953-996.
2. Bentley, R. W. (2002) Global oil & gas depletion: an overview, *Energy Policy 30*, 189-205.
3. Chynoweth, D. P., Owens, J. M., and Legrand, R. (2001) Renewable methane from anaerobic digestion of biomass, *Renewable Energy 22*, 1-8.
4. O. R. N. Laboratory (2003) *Basic Research Needs To Assure A Secure Energy Future*, Oak Ridge.
5. Hermans, I., Spier, E. S., Neuenschwander, U., Turra, N., and Baiker, A. (2009) Selective oxidation catalysis: opportunities and challenges, *Top. Catal. 52*, 1162-1174.
6. (2003) *CRC Handbook of Chemistry and Physics*, 84th ed., CRC Press, Boca Raton, FL.
7. Lunsford, J. H. (2000) Catalytic conversion of methane to more useful chemicals and fuels: a challenge for the 21st century, *Catal. Today 63*, 165-174.
8. Krylov, O. V. (1993) Catalytic reactions of partial methane oxidation, *Catal. Today 18*, 209-302.
9. Forlani, O., and Rossini, S. (1992) Rare-earths as catalysts for the oxidative coupling of methane to ethylene, *Mater. Chem. Phys. 31*, 155-158.
10. Maitra, A. M. (1993) Critical performance evaluation of catalysts and mechanistic implications for oxidative coupling of methane, *Appl. Catal. A-Gen. 104*, 11-59.
11. Kushch, L. A., Lavrushko, V. V., Misharin, Y. S., Moravsky, A. P., and Shilov, A. E. (1983) Kinetics and mechanism of methane oxidation in aqueous-solutions of platinum complexes - direct evidence for a methylplatinum intermediate, *Nouveau Journal De Chimie-New Journal of Chemistry 7*, 729-733.
12. Kao, L. C., Hutson, A. C., and Sen, A. (1991) Low-temperature, palladium(ii)-catalyzed, solution-phase oxidation of methane to a methanol derivative, *J. Am. Chem. Soc. 113*, 700-701.
13. Lin, M., and Sen, A. (1994) Direct catalytic conversion of methane to acetic-acid in an aqueous-medium, *Nature 368*, 613-615.
14. Periana, R. A., Taube, D. J., Evitt, E. R., Loffler, D. G., Wentrcek, P. R., Voss, G., and Masuda, T. (1993) A mercury-catalyzed, high-yield system for the oxidation of methane to methanol, *Science 259*, 340-343.

15. Periana, R. A., Taube, D. J., Gamble, S., Taube, H., Satoh, T., and Fujii, H. (1998) Platinum catalysts for the high-yield oxidation of methane to a methanol derivative, *Science 280*, 560-564.

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

17. Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Muller, J., and Lippard, S. J. (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.

18. Que, L., and Tolman, W. B. (2008) Biologically inspired oxidation catalysis, *Nature 455*, 333-340.

19. Friedle, S., Reisner, E., and Lippard, S. J. (2010) Current challenges of modeling diiron enzyme active sites for dioxygen activation by biomimetic synthetic complexes, *Chem. Soc. Rev. 39*, 2768-2779.

20. Ghosh, A., de Oliveira, F. T., Yano, T., Nishioka, T., Beach, E. S., Kinoshita, I., Munck, E., Ryabov, A. D., Horwitz, C. P., and Collins, T. J. (2005) Catalytically active mu-oxodiiron(IV) oxidants from iron(III) and dioxygen, *J. Am. Chem. Soc. 127*, 2505-2513.

21. Wang, D., Farquhar, E. R., Stubna, A., Munck, E., and Que, L. (2009) A diiron(IV) complex that cleaves strong C-H and O-H bonds, *Nat. Chem. 1*, 145-150.

22. Xue, G. Q., Wang, D., De Hont, R., Fiedler, A. T., Shan, X. P., Munckt, E., and Que, L. (2007) A synthetic precedent for the Fe-2(IV)(mu-O)(2) diamond core proposed for methane monooxygenase intermediate Q, *Proceedings of the National Academy of Sciences of the United States of America 104*, 20713-20718.

23. Hanson, R. S., and Hanson, T. E. (1996) Methanotrophic bacteria, *Microbiol. Rev. 60*, 439-471.

24. Dumont, M. G., and Murrell, J. C. (2005) Community-level analysis: Key genes of aerobic methane oxidation, In *Environmental Microbiology*, Elsevier Academic Press, Inc, San Diego, 413-427.

25. Hakemian, A. S., and Rosenzweig, A. C. (2007) The biochemistry of methane oxidation, *Annu. Rev. Biochem. 76*, 223-241.

26. Lieberman, R. L., and Rosenzweig, A. C. (2004) Biological methane oxidation: Regulation, biochemistry, and active site structure of particulate methane monooxygenase, *Crit. Rev. Biochem. Mol. Biol. 39*, 147-164.

27. Prior, S. D., and Dalton, H. (1985) The effect of copper ions on membrane content and methane monooxygenase activity in methanol-grown cells of *Methylococcus capsulatus* (bath), *J. Gen. Microbiol. 131*, 155-163.

28. Stanley, S. H., Prior, S. D., Leak, D. J., and Dalton, H. (1983) Copper stress underlies the fundamental change in intracellular location of methane monooxygenase in methane-oxidizing organisms - studies in batch and continuous cultures, *Biotechnol. Lett. 5*, 487-492.

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

30. Balasubramanian, R., Smith, S. M., Rawat, S., Yatsunyk, L. A., Stemmler, T. L., and Rosenzweig, A. C. (2010) Oxidation of methane by a biological dicopper centre, *Nature 465*, 115-U131.

31. Hakemian, A. S., Kondapalli, K. C., Telser, J., Hoffman, B. M., Stemmler, T. L., and Rosenzweig, A. C. (2008) The metal centers of particulate methane monooxygenase from *Methylosinus trichosporium* OB3b, *Biochemistry 47*, 6793-6801.

32. Lieberman, R. L., Shrestha, D. B., Doan, P. E., Hoffman, B. M., Stemmler, T. L., and Rosenzweig, A. C. (2003) Purified particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a dimer with both mononuclear copper and a copper-containing cluster, *Proceedings of the National Academy of Sciences of the United States of America 100*, 3820-3825.

33. Choi, D. W., Kunz, R. C., Boyd, E. S., Semrau, J. D., Antholine, W. E., Han, J. I., Zahn, J. A., Boyd, J. M., de la Mora, A. M., and DiSpirito, A. A. (2003) The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH : quinone oxido-reductase complex from *Methylococcus capsulatus* (Bath), *Journal of Bacteriology 185*, 5755-5764.

34. Zahn, J. A., and DiSpirito, A. A. (1996) Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath), *Journal of Bacteriology 178*, 1018-1029.

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

36. Basu, P., Katterle, B., Andersson, K. K., and Dalton, H. (2003) The membrane-associated form of methane mono-oxygenase from *Methylococcus capsulatus* (Bath) is a copper/iron protein, *Biochem. J. 369*, 417-427.

37. Nguyen, H. H. T., Elliott, S. J., Yip, J. H. K., and Chan, S. I. (1998) The particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a novel copper-containing three-subunit enzyme - Isolation and characterization, *Journal of Biological Chemistry 273*, 7957-7966.

38. Burrows, K. J., Cornish, A., Scott, D., and Higgins I. J. (1984) Substrate specificities of the soluble and particulate methane mono-oxygenases of *Methylosinus trichosporium* OB3b, *Microbiology*, 3327-3333.

39. Sullivan, J. P., Dickinson, D., and Chase, H. A. (1998) Methanotrophs, *Methylosinus trichosporium* OB3b, sMMO, and their application to bioremediation, *Crit. Rev. Microbiol. 24*, 335-373.

40. Elliott, S. J., Zhu, M., Tso, L., Nguyen, H. H. T., Yip, J. H. K., and Chan, S. I. (1997) Regio- and stereoselectivity of particulate methane monooxygenase from *Methylococcus capsulatus* (Bath), *J. Am. Chem. Soc. 119*, 9949-9955.

41. Wilkinson, B., Zhu, M., Priestley, N. D., Nguyen, H. H. T., Morimoto, H., Williams, P. G., Chan, S. I., and Floss, H. G. (1996) A concerted mechanism for ethane hydroxylation

by the particulate methane monooxygenase from *Methylococcus capsulatus* (Bath), *J. Am. Chem. Soc. 118*, 921-922.

42. Yu, S. S. F., Wu, L. Y., Chen, K. H. C., Luo, W. I., Huang, D. S., and Chan, S. I. (2003) The stereospecific hydroxylation of 2,2-H-2(2) butane and chiral dideuteriobutanes by the particulate methane monooxygenase from *Methylococcus capsulatus* (Bath), *Journal of Biological Chemistry 278*, 40658-40669.

43. Huang, D. S., Wu, S. H., Wang, Y. S., Yu, S. S. F., and Chan, S. I. (2002) Determination of the carbon kinetic isotope effects on propane hydroxylation mediated by the methane monooxygenases from *Methylococcus capsulatus* (Bath) by using stable carbon isotopic analysis, *Chembiochem 3*, 760-765.

44. Dalton, H. (1980) Oxidation of hydrocarbons by methane monooxygenases from a variety of microbes, In *Advances in Applied Microbiology* (Perlman, D., Ed.), Academic Press, 71-87.

45. Nakajima, T., Uchiyama, H., Yagi, O., and Nakahara, T. (1992) Purification and properties of a soluble methane monooxygenase from *Methylocystis* sp., *Biosci. Biotechnol. Biochem. 56*, 736-740.

46. Wallar, B. J., and Lipscomb, J. D. (1996) Dioxygen activation by enzymes containing binuclear non-heme iron clusters, *Chem. Rev. 96*, 2625-2657.

47. Notomista, E., Lahm, A., Di Donato, A., and Tramontano, A. (2003) Evolution of bacterial and archaeal multicomponent monooxygenases, *J. Mol. Evol. 56*, 435-445.

48. Leahy, J. G., Batchelor, P. J., and Morcomb, S. M. (2003) Evolution of the soluble diiron monooxygenases, *Fems Microbiol. Rev. 27*, 449-479.

49. Shu, L. J., Nesheim, J. C., Kauffmann, K., Munck, E., Lipscomb, J. D., and Que, L. (1997) An (Fe2O2)-O-IV diamond core structure for the key intermediate Q of methane monooxygenase, *Science 275*, 515-518.

50. Ambundo, E. A., Friesner, R. A., and Lippard, S. J. (2002) Reactions of methane monooxygenase intermediate Q with derivatized methanes, *J. Am. Chem. Soc. 124*, 8770-8771.

51. Beauvais, L. G., and Lippard, S. J. (2005) Reactions of the peroxo intermediate of soluble methane monooxygenase hydroxylase with ethers, *J. Am. Chem. Soc. 127*, 7370-7378.

52. Brazeau, B. J., and Lipscomb, J. D. (2000) Kinetics and activation thermodynamics of methane monooxygenase compound Q formation and reaction with substrates, *Biochemistry 39*, 13503-13515.

53. Liu, K. E., Valentine, A. M., Wang, D. L., Huynh, B. H., Edmondson, D. E., Salifoglou, A., and Lippard, S. J. (1995) Kinetic and spectroscopic characterization of intermediates and component interactions in reactions of methane monooxygenase from *Methylococcus capsulatus* (Bath), *J. Am. Chem. Soc. 117*, 10174-10185.

54. Valentine, A. M., Stahl, S. S., and Lippard, S. J. (1999) Mechanistic studies of the reaction of reduced methane monooxygenase hydroxylase with dioxygen and substrates, *J. Am. Chem. Soc. 121*, 3876-3887.

55. Liu, K. E., Valentine, A. M., Qiu, D., Edmondson, D. E., Appelman, E. H., Spiro, T. G., and Lippard, S. J. (1995) Characterization of a diiron(iii) peroxo intermediate in the reaction cycle of methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath), *J. Am. Chem. Soc. 117*, 4997-4998.

56. Lee, S. Y., and Lipscomb, J. D. (1999) Oxygen activation catalyzed by methane monooxygenase hydroxylase component: Proton delivery during the O-O bond cleavage steps, *Biochemistry 38*, 4423-4432.

57. Tinberg, C. E., and Lippard, S. J. (2009) Revisiting the Mechanism of Dioxygen Activation in Soluble Methane Monooxygenase from *M. capsulatus* (Bath): Evidence for a Multi-Step, Proton-Dependent Reaction Pathway, *Biochemistry 48*, 12145-12158.

58. Gherman, B. F., Dunietz, B. D., Whittington, D. A., Lippard, S. J., and Friesner, R. A. (2001) Activation of the C-H bond of methane by intermediate Q of methane monoozygenase: A theoretical study, *J. Am. Chem. Soc. 123*, 3836-3837.

59. Musaev, D. G., Basch, H., and Morokuma, K. (2002) Theoretical study of the mechanism of alkane hydroxylation and ethylene epoxidation reactions catalyzed by diiron bis-oxo complexes. The effect of substrate molecules, *J. Am. Chem. Soc. 124*, 4135-4148.

60. Siegbahn, P. E. M. (2001) O-O bond cleavage and alkane hydroxylation in methane monooxygenase, *J. Biol. Inorg. Chem. 6*, 27-45.

61. Baik, M. H., Newcomb, M., Friesner, R. A., and Lippard, S. J. (2003) Mechanistic studies on the hydroxylation of methane by methane monooxygenase, *Chem. Rev. 103*, 2385-2419.

62. Froland, W. A., Andersson, K. K., Lee, S. K., Liu, Y., and Lipscomb, J. D. (1992) Methane monooxygenase component-b and reductase alter the regioselectivity of the hydroxylase component-catalyzed reactions - a novel role for protein-protein interactions in an oxygenase mechanism, *Journal of Biological Chemistry 267*, 17588-17597.

63. Gassner, G. T., and Lippard, S. J. (1999) Component interactions in the soluble methane monooxygenase system from *Methylococcus capsulatus* (Bath), *Biochemistry 38*, 12768-12785.

64. Liu, Y., Nesheim, J. C., Lee, S. K., and Lipscomb, J. D. (1995) Gating effects of component-b on oxygen activation by the methane monooxygenase hydroxylase component, *Journal of Biological Chemistry 270*, 24662-24665.

65. Blazyk, J. L., Gassner, G. T., and Lippard, S. J. (2005) Intermolecular electron-transfer reactions in soluble methane monooxygenase: A role for hysteresis in protein function, *J. Am. Chem. Soc. 127*, 17364-17376.

66. Sazinsky, M. H., and Lippard, S. J. (2006) Correlating structure with function in bacterial multicomponent monooxygenases and related diiron proteins, *Accounts Chem. Res. 39*, 558-566.

67. Murray, L. J., and Lippard, S. J. (2007) Substrate trafficking and dioxygen activation in bacterial multicomponent monooxygenases, *Accounts Chem. Res. 40*, 466-474.

68. Whittington, D. A., Rosenzweig, A. C., Frederick, C. A., and Lippard, S. J. (2001) Xenon and halogenated alkanes track putative substrate binding cavities in the soluble methane monooxygenase hydroxylase, *Biochemistry 40*, 3476-3482.

69. Sazinsky, M. H., and Lippard, S. J. (2005) Product bound structures of the soluble methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath): Protein motion in the alpha-subunit, *J. Am. Chem. Soc. 127*, 5814-5825.

70. Wilkins, P. C., Dalton, H., Samuel, C. J., and Green, J. (1994) Further evidence for multiple pathways in soluble methane-monooxygenase-catalyzed oxidations from the measurement of deuterium kinetic isotope effects, *Eur. J. Biochem. 226*, 555-560.

71. Rataj, M. J., Kauth, J. E., and Donnelly, M. I. (1991) Oxidation of deuterated compounds by high specific activity methane monooxygenase from *Methylosinus trichosporium* - mechanistic implications, *Journal of Biological Chemistry 266*, 18684-18690.

72. Tinberg, C. E., and Lippard, S. J. (2010) Oxidation reactions performed by soluble methane monooxygenase hydroxylase intermediates H-peroxo and Q proceed by distinct mechanisms, *Biochemistry 49*, 7902-7912.

73. Lee, S. G., Goo, J. H., Kim, H. G., Oh, J. I., Kim, Y. M., and Kim, S. W. (2004) Optimization of methanol biosynthesis from methane using *Methylosinus trichosporium* OB3b, *Biotechnol. Lett. 26*, 947-950.

74. Gunay, A., and Theopold, K. H. (2010) C-H Bond activations by metal oxo compounds, *Chem. Rev. 110*, 1060-1081.

75. Smits, T. H. M., Witholt, B., and van Beilen, J. B. (2003) Functional characterization of genes involved in alkane oxidation by *Pseudomonas aeruginosa*, *Antonie Van Leeuwenhoek 84*, 193-200.

76. Ueda, T., and Coon, M. J. (1972) Enzymatic omega-oxidation .7. Reduced diphosphopyridine nucleotide-rubredoxin reductase - properties and function as an electron carrier in omega hydroxylation, *Journal of Biological Chemistry 247*, 5010-5027.

77. Peterson, J. A., Kusunose, M., Kusunose, E., and Coon, M. J. (1967) Enzymatic omega-oxilation .2. Function of rubredoxin as electron carrier in omega-hydroxylation, *Journal of Biological Chemistry 242*, 4334-4353.

78. McKenna, E. J., and Coon, M. J. (1970) Enzymatic omega-oxidation .4. Purification and properties of omega-hydroxylase of *Pseudomonas oleovorans*, *Journal of Biological Chemistry 245*, 3882-3897.

79. Ruettinger, R. T., Griffith, G. R., and Coon, M. J. (1977) Characterization of omega-hydroxylase of *Pseudomonas oleovorans* as a nonheme iron protein, *Arch. Biochem. Biophys. 183*, 528-537.

80. Nieboer, M., Kingma, J., and Witholt, B. (1993) The alkane oxidation system of *Pseudomonas oleovorans* - induction of the alk genes in *Escherichia coli* w3110(pgec47) affects membrane biogenesis and results in overexpression of alkane hydroxylase in a distinct cytoplasmic membrane subfraction, *Mol. Microbiol. 8*, 1039-1051.

81. Shanklin, J., and Whittle, E. (2003) Evidence linking the *Pseudomonas oleovorans* alkane omega-hydroxylase, an integral membrane diiron enzyme, and the fatty acid desaturase family, *FEBS Lett. 545*, 188-192.

82. Shanklin, J., Achim, C., Schmidt, H., Fox, B. G., and Munck, E. (1997) Mossbauer studies of alkane omega-hydroxylase: Evidence for a diiron cluster in an integral-membrane enzyme, *Proceedings of the National Academy of Sciences of the United States of America 94*, 2981-2986.

83. Rozhkova-Novosad, E. A., Chae, J. C., Zylstra, G. J., Bertrand, E. M., Alexander-Ozinskas, M., Deng, D. Y., Moe, L. A., van Beilen, J. B., Danahy, M., Groves, J. T., and Austin, R. N. (2007) Profiling mechanisms of alkane hydroxylase activity *in vivo* using the diagnostic substrate norcarane, *Chem. Biol. 14*, 165-172.

84. Dawson, J. H., and Sono, M. (1987) Cytochrome P450 and chloroperoxidase - thiolate-ligated heme enzymes - spectroscopic determination of their active-site structures and mechanistic implications of thiolate ligation, *Chem. Rev. 87*, 1255-1276.

85. Morant, M., Bak, S., Moller, B. L., and Werck-Reichhart, D. (2003) Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation, *Current Opinion in Biotechnology 14*, 151-162.

86. Persans, M. W., Wang, J., and Schuler, M. A. (2001) Characterization of maize cytochrome P450 monooxygenases induced in response to safeners and bacterial pathogens, *Plant Physiol. 125*, 1126-1138.

87. Schuler, M. A., and Werck-Reichhart, D. (2003) Functional genomics of P450s, *Annu. Rev. Plant Biol. 54*, 629-667.

88. de Montellano, P. R. O. (1986) *Cytochrome P450*, 1st ed., Plenum Publishing Corp., New York.

89. Isin, E. M., and Guengerich, F. P. (2007) Complex reactions catalyzed by cytochrome P450 enzymes, *Biochim. Biophys. Acta-Gen. Subj. 1770*, 314-329.

90. Denisov, I. G., Makris, T. M., Sligar, S. G., and Schlichting, I. (2005) Structure and chemistry of cytochrome P450, *Chem. Rev. 105*, 2253-2277.

91. Poulos, T. L., and Raag, R. (1992) Cytochrome P450cam - crystallography, oxygen activation, and electron-transfer, *Faseb J. 6*, 674-679.

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

93. Nishida, C. R., and de Montellano, P. R. O. (2005) Thermophilic cytochrome P450 enzymes, *Biochem. Biophys. Res. Commun. 338*, 437-445.

94. Yano, J. K., Blasco, F., Li, H. Y., Schmid, R. D., Henne, A., and Poulos, T. L. (2003) Preliminary characterization and crystal structure of a thermostable cytochrome P450 from *Thermus thermophilus*, *Journal of Biological Chemistry 278*, 608-616.

95. Ho, W. W., Li, H., Nishida, C. R., de Montellano, P. R. O., and Poulos, T. L. (2008) Crystal structure and properties of CYP231A2 from the thermoacidophilic *Archaeon Picrophilus torridus*, *Biochemistry 47*, 2071-2079.

96. Cirino, P. C., and Arnold, F. H. (2002) Protein engineering of oxygenases for biocatalysis, *Curr. Opin. Chem. Biol. 6*, 130-135.

97. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.

98. Nazor, J., Dannenmann, S., Adjei, R. O., Fordjour, Y. B., Ghampson, I. T., Blanusa, M., Roccatano, D., and Schwaneberg, U. (2008) Laboratory evolution of P450BM3 for mediated electron transfer yielding an activity-improved and reductase-independent variant, *Protein Eng. Des. Sel. 21*, 29-35.

99. Seifert, A., Vomund, S., Grohmann, K., Kriening, S., Urlacher, V. B., Laschat, S., and Pleiss, J. (2009) Rational design of a minimal and highly enriched CYP102A1 mutant library with improved regio-, stereo- and chemoselectivity, *Chembiochem 10*, 853-861.

100. van Vugt-Lussenburg, B. M. A., Stjernschantz, E., Lastdrager, J., Oostenbrink, C., Vermeulen, N. P. E., and Commandeur, J. N. M. (2007) Identification of critical residues in novel drug metabolizing mutants of cytochrome P450BM3 using random mutagenesis, *J. Med. Chem. 50*, 455-461.

101. Hasemann, C. A., Kurumbail, R. G., Boddupalli, S. S., Peterson, J. A., and Deisenhofer, J. (1995) Structure and function of cytochromes P450 - a comparative-analysis of 3 crystal-structures, *Structure 3*, 41-62.

102. Presnell, S. R., and Cohen, F. E. (1989) Topological distribution of 4-alpha-helix bundles, *Proceedings of the National Academy of Sciences of the United States of America 86*, 6592-6596.

103. Dawson, J. H., Holm, R. H., Trudell, J. R., Barth, G., Linder, R. E., Bunnenberg, E., Djerassi, C., and Tang, S. C. (1976) Oxidized cytochrome P450 magnetic circular dichroism evidence for thiolate ligation in substrate-bound form implications for catalytic mechanism, *J. Am. Chem. Soc. 98*, 3707-3709.

104. Imai, M., Shimada, H., Watanabe, Y., Matsushimahibiya, Y., Makino, R., Koga, H., Horiuchi, T., and Ishimura, Y. (1989) Uncoupling of the cytochrome P450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine - a possible role of the hydroxy amino-acid in oxygen activation, *Proceedings of the National Academy of Sciences of the United States of America 86*, 7823-7827.

105. Kimata, Y., Shimada, H., Hirose, T., and Ishimura, Y. (1995) Role of thr-252 in cytochrome P450cam - a study with unnatural amino-acid mutagenesis, *Biochem. Biophys. Res. Commun. 208*, 96-102.

106. Porubsky, P. R., Meneely, K. M., and Scott, E. E. (2008) Structures of human cytochrome P450 2E1 insights into the binding of inhibitors and both small molecular weight and fatty acid substrates, *Journal of Biological Chemistry 283*, 33698-33707.

107. Pylypenko, O., Vitali, F., Zerbe, K., Robinson, J. A., and Schlichting, I. (2003) Crystal structure of OxyC, a cytochrome P450 implicated in an oxidative C-C coupling reaction during vancomycin biosynthesis, *Journal of Biological Chemistry 278*, 46727-46733.

108. Kelly, S. L., Lamb, D. C., and Kelly, D. E. (2006) Cytochrome P450 biodiversity and biotechnology, *Biochem. Soc. Trans. 34*, 1159-1160.

109. Pylypenko, O., and Schlichting, I. (2004) Structural aspects of ligand binding to and electron transfer in bacterial and fungal p450s, *Annu. Rev. Biochem. 73*, 991-1018.

110. Gotoh, O. (1992) Substrate recognition sites in cytochrome-P450 family-2 (cyp2) proteins inferred from comparative analyses of amino-acid and coding nucleotide-sequences, *Journal of Biological Chemistry 267*, 83-90.

111. Sligar, S. G. (1976) Coupling of spin, substrate, and redox equilibria in cytochrome P450, *Biochemistry 15*, 5399-5406.

112. Groves, J. T. (1985) Key elements of the chemistry of cytochrome P450 - the oxygen rebound mechanism, *J. Chem. Educ. 62*, 928-931.

113. Coon, M. J., Vaz, A. D. N., McGinnity, D. F., and Peng, H. M. (1998) Multiple activated oxygen species in P450 catalysis - Contributions to specificity in drug metabolism, *Drug Metab. Dispos. 26*, 1190-1193.

114. Bernhardt, R. (1996) Cytochrome P450: Structure, function, and generation of reactive oxygen species, In *Reviews of Physiology Biochemistry and Pharmacology, Vol 127*, Springer-Verlag Berlin, Berlin 33, 137-221.

115. Avila, L., Wirtz, M., Bunce, R. A., and Rivera, M. (1999) An electrochemical study of the factors responsible for modulating the reduction potential of putidaredoxin, *J. Biol. Inorg. Chem. 4*, 664-674.

116. Daff, S. N., Chapman, S. K., Turner, K. L., Holt, R. A., Govindaraj, S., Poulos, T. L., and Munro, A. W. (1997) Redox control of the catalytic cycle of flavocytochrome P-450 BM3, *Biochemistry 36*, 13816-13823.

117. Guengerich, F. P., Ballou, D. P., and Coon, M. J. (1975) Purified liver microsomal cytochrome P450 - electron-accepting properties and oxidation-reduction potential, *Journal of Biological Chemistry 250*, 7405-7414.

118. Ost, T. W. B., Clark, J., Mowat, C. G., Miles, C. S., Walkinshaw, M. D., Reid, G. A., Chapman, S. K., and Daff, S. (2003) Oxygen activation and electron transfer in flavocytochrome P450BM3, *J. Am. Chem. Soc. 125*, 15010-15020.

119. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, B. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) The catalytic pathway of cytochrome P450cam at atomic resolution, *Science 287*, 1615-1622.

120. Imai, M., Shimada, H., Watanabe, Y., Matsushimahibiya, Y., Makino, R., Koga, H., Horiuchi, T., and Ishimura, Y. (1989) Uncoupling of the cytochrome P450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine - a possible role of the hydroxy amino-acid in oxygen activation, *Proceedings of the National Academy of Sciences of the United States of America 86*, 7823-7827.

121. Kadkhodayan, S., Coulter, E. D., Maryniak, D. M., Bryson, T. A., and Dawson, J. H. (1995) Uncoupling oxygen-transfer and electron-transfer in the oxygenation of camphor analogs by cytochrome P450cam - direct observation of an intermolecular isotope effect for substrate C-H activation, *Journal of Biological Chemistry 270*, 28042-28048.

122. Koshland, D. E. (1994) The key-lock theory and the induced fit theory, *Angewandte Chemie-International Edition 33*, 2375-2378.

123. Szklarz, G. D., and Paulsen, M. D. (2002) Molecular modeling of cytochrome P450 1A1: Enzyme-substrate interactions and substrate binding affinities, *J. Biomol. Struct. Dyn. 20*, 155-162.

124. Hansch, C., and Zhang, L. T. (1993) Quantitative structure-activity-relationships of cytochrome P450, *Drug Metab. Rev. 25*, 1-48.

125. Lewis, D. F. V., Jacobs, M. N., and Dickins, M. (2004) Compound lipophilicity for substrate binding to human P450s in drug metabolism, *Drug Discov. Today 9*, 530-537.

126. Frommer, U., Ullrich, V., and Stauding.H. (1970) Hydroxylation of aliphatic compounds by liver microsomes .1. Distribution pattern of isomeric alcohols, *Hoppe-Seylers Zeitschrift Fur Physiologische Chemie 351*, 903-912.

127. de Visser, S. P., Kumar, D., Cohen, S., Shacham, R., and Shaik, S. (2004) A predictive pattern of computed barriers for C-H hydroxylation by compound I of cytochrome P450, *J. Am. Chem. Soc. 126*, 8362-8363.

128. Olsen, L., Rydberg, P., Rod, T. H., and Ryde, U. (2006) Prediction of activation energies for hydrogen abstraction by cytochrome P450, *J. Med. Chem. 49*, 6489-6499.

129. Tajima, K., Edo, T., Ishizu, K., Imaoka, S., Funae, Y., Oka, S., and Sakurai, H. (1993) Cytochrome P450-butyl peroxide complex detected by ESR, *Biochem. Biophys. Res. Commun. 191*, 157-164.

130. Blake, R. C., and Coon, M. J. (1981) On the mechanism of action of cytochrome P450 - role of peroxy spectral intermediates in substrate hydroxylation, *Journal of Biological Chemistry 256*, 5755-5763.

131. Jung, C., Schunemann, V., and Lendzian, F. (2005) Freeze-quenched iron-oxo intermediates in cytochromes P450, *Biochem. Biophys. Res. Commun. 338*, 355-364.

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

133. Spolitak, T., Dawson, J. H., and Ballou, D. P. (2005) Reaction of ferric cytochrome P450cam with peracids - Kinetic characterization of intermediates on the reaction pathway, *Journal of Biological Chemistry 280*, 20300-20309.

134. Munro, A. W., Girvan, H. M., and McLean, K. J. (2007) Variations on a (t)heme - novel mechanisms, redox partners and catalytic functions in the cytochrome P450 superfamily, *Natural Product Reports 24*, 585-609.

135. Loew, G. H., and Harris, D. L. (2000) Role of the heme active site and protein environment in structure, spectra, and function of the cytochrome p450s, *Chem. Rev. 100*, 407-419.

136. Gustafsson, J. A., Rondahl, L., and Bergman, J. (1979) Iodosylbenzene derivatives as oxygen donors in cytochrome P450 catalyzed steroid hydroxylations, *Biochemistry 18*, 865-870.

137. Heimbrook, D. C., and Sligar, S. G. (1981) Multiple mechanisms of cytochrome P450-catalyzed substrate hydroxylations, *Biochem. Biophys. Res. Commun. 99*, 530-535.

138. Macdonald, T. L., Burka, L. T., Wright, S. T., and Guengerich, F. P. (1982) Mechanisms of hydroxylation by cytochrome P450 - exchange of iron-oxygen intermediates with water, *Biochem. Biophys. Res. Commun. 104*, 620-625.

139. Song, W. J., Sun, Y. J., Choi, S. K., and Nam, W. (2006) Mechanistic insights into the reversible formation of iodosylarene-iron porphyrin complexes in the reactions of oxoiron(IV) porphyrin pi-cation radicals and iodoarenes: Equilibrium, epoxidizing intermediate, and oxygen exchange, *Chemistry-a European Journal 12*, 130-137.

140. Guengerich, F. P., Yun, C. H., and Macdonald, T. L. (1996) Evidence for a 1-electron oxidation mechanism in N-dealkylation of N,N-dialkylanilines by cytochrome P450 2B1 - Kinetic hydrogen isotope effects, linear free energy relationships, comparisons with horseradish peroxidase, and studies with oxygen surrogates, *Journal of Biological Chemistry 271*, 27321-27329.

141. Bhakta, M. N., Hollenberg, P. F., and Wimalasena, K. (2005) P-450/NADPH/O-2- and P-450/PhIO-catalyzed N-dealkylations are mechanistically distinct, *J. Am. Chem. Soc. 127*, 1376-1377.

142. Brazeau, B. J., Austin, R. N., Tarr, C., Groves, J. T., and Lipscomb, J. D. (2001) Intermediate Q from soluble methane monooxygenase hydroxylates the mechanistic substrate probe norcarane: Evidence for a stepwise reaction, *J. Am. Chem. Soc. 123*, 11831-11837.

143. Kopp, D. A., and Lippard, S. J. (2002) Soluble methane monooxygenase: activation of dioxygen and methane, *Curr. Opin. Chem. Biol. 6*, 568-576.

144. Newcomb, M., Shen, R. N., Lu, Y., Coon, M. J., Hollenberg, P. F., Kopp, D. A., and Lippard, S. J. (2002) Evaluation of norcarane as a probe for radicals in cytochome P450- and soluble methane monooxygenase-catalyzed hydroxylation reactions, *J. Am. Chem. Soc. 124*, 6879-6886.

145. Groves, J. T. (2003) The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies, *Proceedings of the National Academy of Sciences of the United States of America 100*, 3569-3574.

146. Groves, J. T., McClusky, G. A., White, R. E., and Coon, M. J. (1978) Aliphatic hydroxylation by highly purified liver microsomal cytochrome P450 - evidence for a carbon radical intermediate, *Biochem. Biophys. Res. Commun. 81*, 154-160.

147. Krauser, J. A., and Guengerich, F. P. (2005) Cytochrome P450 3A4-catalyzed testosterone 6 beta-hydroxylation stereochemistry, kinetic deuterium isotope effects, and rate-limiting steps, *Journal of Biological Chemistry 280*, 19496-19506.

148. White, R. E., Miller, J. P., Favreau, L. V., and Bhattacharyya, A. (1986) Stereochemical dynamics of aliphatic hydroxylation by cytochrome P450, *J. Am. Chem. Soc. 108*, 6024-6031.

149. Shaik, S., Kumar, D., de Visser, S. P., Altun, A., and Thiel, W. (2005) Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes, *Chem. Rev. 105*, 2279-2328.

150. de Visser, S. P., Ogliaro, F., Sharma, P. K., and Shaik, S. (2002) What factors affect the regioselectivity of oxidation by cytochrome P450? A DFT study of allylic hydroxylation and double bond epoxidation in a model reaction, *J. Am. Chem. Soc. 124*, 11809-11826.

151. Ogliaro, F., Harris, N., Cohen, S., Filatov, M., de Visser, S. P., and Shaik, S. (2000) A model "rebound" mechanism of hydroxylation by cytochrome P450: Stepwise and effectively concerted pathways, and their reactivity patterns, *J. Am. Chem. Soc. 122*, 8977-8989.

152. Basch, H., Musaev, D. G., Mogi, K., and Morokuma, K. (2001) Theoretical studies on the mechanism of the methane -> methanol conversion reaction catalyzed by methane monooxygenase: O-side vs N-side mechanisms, *J. Phys. Chem. A 105*, 3615-3622.

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

154. Gillam, E. M. (2005) Exploring the potential of xenobiotic-metabolising enzymes as biocatalysts: Evolving designer catalysts from polyfunctional cytochrome P450 enzymes, *Clin. Exp. Pharmacol. Physiol. 32*, 147-152.

155. Guengerich, F. P. (2002) Cytochrome P450 enzymes in the generation of commercial products, *Nat. Rev. Drug Discov. 1*, 359-366.

156. Duetz, W. A., van Beilen, J. B., and Witholt, B. (2001) Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis, *Current Opinion in Biotechnology 12*, 419-425.

157. Fukui, Y., Tanaka, Y., Kusumi, T., Iwashita, T., and Nomoto, K. (2003) A rationale for the shift in colour towards blue in transgenic carnation flowers expressing the flavonoid 3 ',5 '-hydroxylase gene, *Phytochemistry 63*, 15-23.

158. Holton, T. A., Brugliera, F., Lester, D. R., Tanaka, Y., Hyland, C. D., Menting, J. G. T., Lu, C. Y., Farcy, E., Stevenson, T. W., and Cornish, E. C. (1993) Cloning and expression of cytochrome P450 genes-controlling flower color, *Nature 366*, 276-279.

159. Petzoldt, K. E. A. *Process for the preparation of 11-beta-hydroxy steroids*, Schering Aktiengesellschaft, Germany.

160. Peterson, D. H., Murray, H. C., Eppstein, S. H., Reineke, L. M., Weintraub, A., Meister, P. D., and Leigh, H. M. (1952) Microbiological transformations of steroids .1. Introduction of oxygen at carbon-11 of progesterone, *J. Am. Chem. Soc. 74*, 5933-5936.

161. van Beilen, J. B., Duetz, W. A., Schmid, A., and Witholt, B. (2003) Practical issues in the application of oxygenases, *Trends Biotechnol. 21*, 170-177.

162. Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., and Keasling, J. D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast, *Nature 440*, 940-943.

163. Rabe, K. S., Gandubert, V. J., Spengler, M., Erkelenz, M., and Niemeyer, C. M. (2008) Engineering and assaying of cytochrome P450 biocatalysts, *Anal. Bioanal. Chem. 392*, 1059-1073.

164. Palmer, C. N. A., Axen, E., Hughes, V., and Wolf, C. R. (1998) The repressor protein, Bm3R1, mediates an adaptive response to toxic fatty acids in *Bacillus megaterium*, *Journal of Biological Chemistry 273*, 18109-18116.

165. Palmer, C. N. A., Causevic, M., and Wolf, C. R. (1997) Modulation of fatty acid signalling by cytochrome P-450-mediated hydroxylation, *Biochem. Soc. Trans. 25*, 1160-1165.

166. Noble, M. A., Miles, C. S., Chapman, S. K., Lysek, D. A., Mackay, A. C., Reid, G. A., Hanzlik, R. P., and Munro, A. W. (1999) Roles of key active-site residues in flavocytochrome P450 BM3, *Biochem. J. 339*, 371-379.

167. Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., and Kraut, J. (1985) The 2.6-Å crystal-structure of Pseudomonas putida cytochrome P450, *Journal of Biological Chemistry 260*, 6122-6130.

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

169. Haines, D. C., Tomchick, D. R., Machius, M., and Peterson, J. A. (2001) Pivotal role of water in the mechanism of P450BM-3, *Biochemistry 40*, 13456-13465.

170. Lee, D. S., Park, S. Y., Yamane, K., Obayashi, E., Hori, H., and Shiro, Y. (2001) Structural characterization of n-butyl-isocyanide complexes of cytochromes P450nor and P450cam, *Biochemistry 40*, 2669-2677.

171. Li, H. Y., and Poulos, T. L. (1997) The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid, *Nat. Struct. Biol. 4*, 140-146.

172. Raag, R., and Poulos, T. L. (1991) Crystal-structures of cytochrome P450cam complexed with camphane, thiocamphor, and adamantane - factors controlling P450 substrate hydroxylation, *Biochemistry 30*, 2674-2684.

173. Schlichting, I., Jung, C., and Schulze, H. (1997) Crystal structure of cytochrome P-450cam complexed with the (1S)-camphor enantiomer, *FEBS Lett. 415*, 253-257.

174. Bell, S. G., Chen, X. H., Sowden, R. J., Xu, F., Williams, J. N., Wong, L. L., and Rao, Z. H. (2003) Molecular recognition in (+)-alpha-pinene oxidation by cytochrome P450(cam), *J. Am. Chem. Soc. 125*, 705-714.

175. Bell, S. G., Harford-Cross, C. F., and Wong, L. L. (2001) Engineering the CYP101 system for in vivo oxidation of unnatural substrates, *Protein Eng. 14*, 797-802.

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

177. Carmichael, A. B., and Wong, L. L. (2001) Protein engineering of *Bacillus megaterium* CYP102 - The oxidation of polycyclic aromatic hydrocarbons, *Eur. J. Biochem. 268*, 3117-3125.

178. England, P. A., Harford-Cross, C. F., Stevenson, J. A., Rouch, D. A., and Wong, L. L. (1998) The oxidation of naphthalene and pyrene by cytochrome P450(cam), *FEBS Lett. 424*, 271-274.

179. Loida, P. J., and Sligar, S. G. (1993) Engineering cytochrome P450cam to increase the stereospecificity and coupling of aliphatic hydroxylation, *Protein Eng. 6*, 207-212.

180. Bell, S. G., Rouch, D. A., and Wong, L. L. (1997) Selective aliphatic and aromatic carbon-hydrogen bond activation catalysed by mutants of cytochrome P450(cam), *J. Mol. Catal. B-Enzym. 3*, 293-302.

181. England, P. A., Rouch, D. A., Westlake, A. C. G., Bell, S. G., Nickerson, D. P., Webberley, M., Flitsch, S. L., and Wong, L. L. (1996) Aliphatic vs aromatic C-H bond activation of phenylcyclohexane catalysed by cytochrome P450cam, *Chemical Communications*, 357-358.

182. Fowler, S. M., England, P. A., Westlake, A. C. G., Rouch, D. R., Nickerson, D. P., Blunt, C., Braybrook, D., West, S., Wong, L. L., and Flitsch, S. L. (1994) Cytochrome P450cam monooxygenase can be redesigned to catalyze the regioselective aromatic hydroxylation of diphenylmethane, *J. Chem. Soc.-Chem. Commun.*, 2761-2762.

183. Harford-Cross, C. F., Carmichael, A. B., Allan, F. K., England, P. A., Rouch, D. A., and Wong, L. L. (2000) Protein engineering of cytochrome P450(cam) (CYP101) for the oxidation of polycyclic aromatic hydrocarbons, *Protein Eng. 13*, 121-128.

184. Jones, J. P., O'Hare, E. J., and Wong, L. L. (2000) The oxidation of polychlorinated benzenes by genetically engineered cytochrome P450(cam): potential applications in bioremediation, *Chemical Communications*, 247-248.

185. Manna, S. K., and Mazumdar, S. (2010) Tuning the substrate specificity by engineering the active site of cytochrome P450cam: A rational approach, *Dalton Trans. 39*, 3115-3123.

186. Nickerson, D. P., HarfordCross, C. F., Fulcher, S. R., and Wong, L. L. (1997) The catalytic activity of cytochrome P450(cam) towards styrene oxidation is increased by site-specific mutagenesis, *FEBS Lett. 405*, 153-156.

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

188. Stevenson, J. A., Westlake, A. C. G., Whittock, C., and Wong, L. L. (1996) The catalytic oxidation of linear and branched alkanes by cytochrome P450(cam), *J. Am. Chem. Soc. 118*, 12846-12847.

189. Hirakawa, H., and Nagamune, T. (2010) Molecular assembly of P450 with ferredoxin and ferredoxin reductase by fusion to PCNA, *Chembiochem 11*, 1517-1520.

190. Sabbadin, F., Hyde, R., Robin, A., Hilgarth, E. M., Delenne, M., Flitsch, S., Turner, N., Grogan, G., and Bruce, N. C. (2010) LICRED: A versatile drop-in vector for rapid generation of redox-self-sufficient cytochrome P450s, *Chembiochem 11*, 987-994.

191. Boddupalli, S. S., Oster, T., Estabrook, R. W., and Peterson, J. A. (1992) Reconstitution of the fatty-acid hydroxylation function of cytochrome P450BM3 utilizing its individual recombinant hemoprotein and flavoprotein domains, *Journal of Biological Chemistry 267*, 10375-10380.

192. Eiben, S., Kaysser, L., Maurer, S., Kuhnel, K., Urlacher, V. B., and Schmid, R. D. (2006) Preparative use of isolated CYP102 monooxygenases - A critical appraisal, *Journal of Biotechnology 124*, 662-669.

193. Boddupalli, S. S., Pramanik, B. C., Slaughter, C. A., Estabrook, R. W., and Peterson, J. A. (1992) Fatty-acid monooxygenation by P450BM3 - product identification and proposed mechanisms for the sequential hydroxylation reactions, *Arch. Biochem. Biophys. 292*, 20-28.

194. Oliver, C. F., Modi, S., Primrose, W. U., Lian, L. Y., and Roberts, G. C. K. (1997) Engineering the substrate specificity of *Bacillus megaterium* cytochrome P-450 BM3: hydroxylation of alkyl trimethylammonium compounds, *Biochem. J. 327*, 537-544.

195. Ost, T. W. B., Miles, C. S., Murdoch, J., Cheung, Y. F., Reid, G. A., Chapman, S. K., and Munro, A. W. (2000) Rational re-design of the substrate binding site of flavocytochrome P450BM3, *FEBS Lett. 486*, 173-177.

196. Dietrich, J. A., Yoshikuni, Y., Fisher, K. J., Woolard, F. X., Ockey, D., McPhee, D. J., Renninger, N. S., Chang, M. C. Y., Baker, D., and Keasling, J. D. (2009) A Novel Semi-biosynthetic Route for Artemisinin Production Using Engineered Substrate-Promiscuous P450(BM3), *ACS Chem. Biol. 4*, 261-267.

197. Lussenburg, B. M. A., Babel, L. C., Vermeulen, N. P. E., and Commandeur, J. N. M. (2005) Evaluation of alkoxyresorufins as fluorescent substrates for cytochrome P450BM3 and site-directed mutants, *Anal. Biochem. 341*, 148-155.

198. Appel, D., Lutz-Wahl, S., Fischer, P., Schwaneberg, U., and Schmid, R. D. (2001) A P450BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes, *Journal of Biotechnology 88*, 167-171.

199. Li, Q. S., Ogawa, J., Schmid, R. D., and Shimizu, S. (2001) Engineering cytochrome P450BM-3 for oxidation of polycyclic aromatic hydrocarbons, *Applied and Environmental Microbiology 67*, 5735-5739.

200. Li, Q. S., Schwaneberg, U., Fischer, P., and Schmid, R. D. (2000) Directed evolution of the fatty-acid hydyoxylase P450BM-3 into an indole-hydroxylating catalyst, *Chemistry-a European Journal 6*, 1531-1536.

201. Urlacher, V. B., Makhsumkhanov, A., and Schmid, R. D. (2006) Biotransformation of beta-ionone by engineered cytochrome P450BM-3, *Applied Microbiology and Biotechnology 70*, 53-59.

202. Sowden, R. J., Yasmin, S., Rees, N. H., Bell, S. G., and Wong, L. L. (2005) Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450(cam) and P450(BM-3), *Org. Biomol. Chem. 3*, 57-64.

203. Graham-Lorence, S., Truan, G., Peterson, J. A., Falck, J. R., Wei, S. Z., Helvig, C., and Capdevila, J. H. (1997) An active site substitution, F87V, converts cytochrome p450 BM-3 into a regio- and stereoselective (14S,15R)-arachidonic acid epoxygenase, *Journal of Biological Chemistry 272*, 1127-1135.

204. Sulistyaningdyah, W. T., Ogawa, J., Li, Q. S., Maeda, C., Yano, Y., Schmid, R. D., and Shimizu, S. (2005) Hydroxylation activity of P450BM-3 mutant F87V towards aromatic compounds and its application to the synthesis of hydroquinone derivatives from phenolic compounds, *Applied Microbiology and Biotechnology 67*, 556-562.

205. Weber, E., Seifert, A., Antonovici, M., Geinitz, C., Pleiss, J., and Urlacher, V. B. (2011) Screening of a minimal enriched P450 BM3 mutant library for hydroxylation of cyclic and acyclic alkanes, *Chemical Communications 47*, 944-946.

206. Farinas, E. T., Schwaneberg, U., Glieder, A., and Arnold, F. H. (2001) Directed evolution of a cytochrome P450 monooxygenase for alkane oxidation, *Advanced Synthesis & Catalysis 343*, 601-606.

207. Glieder, A., Farinas, E. T., and Arnold, F. H. (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase, *Nat. Biotechnol. 20*, 1135-1139.

208. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.

209. Meinhold, P., Peters, M. W., Chen, M. M. Y., Takahashi, K., and Arnold, F. H. (2005) Direct conversion of ethane to ethanol by engineered cytochrome P450BM3, *Chembiochem 6*, 1765-1768.

210. Meinhold, P., Peters, M. W., Hartwick, A., Hernandez, A. R., and Arnold, F. H. (2006) Engineering cytochrome P450BM3 for terminal alkane hydroxylation, *Advanced Synthesis & Catalysis 348*, 763-772.

211. GrahamLorence, S., Truan, G., Peterson, J. A., Falck, J. R., Wei, S. Z., Helvig, C., and Capdevila, J. H. (1997) An active site substitution, F87V, converts cytochrome p450 BM-3 into a regio- and stereoselective (14S,15R)-arachidonic acid epoxygenase, *Journal of Biological Chemistry 272*, 1127-1135.

212. Kubo, T., Peters, M. W., Meinhold, P., and Arnold, F. H. (2006) Enantioselective epoxidation of terminal alkenes to (R)- and (S)-epoxides by engineered cytochromes P450BM-3, *Chemistry-a European Journal 12*, 1216-1220.

Chapter 2

Engineered Alkane-Hydroxylating Cytochrome P450 BM3 Exhibiting Native-like Catalytic Properties

Material from this chapter appears in: Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting native-like catalytic properties, *Angewandte Chemie–International Edition* 46, 8414 – 8418, and is reprinted by permission from Wiley-VCH.

A. Abstract

New functions have been engineered in a variety of pre-existing enzymes using directed evolution, however, examples in which the engineered variants exhibit comparable catalytic properties with the non-natural substrate as the wildtype enzyme with its preferred substrates are rare. Here, we describe the *in vitro* evolution of a proficient P450 propane monooxygenase, $P450_{PMO,}$ starting from a fatty acid hydroxylase CYP102A1 (BM3). Applying only positive selection pressure in combination with a domain engineering mutagenesis strategy, which targeted the heme and reductase domains independently and in combination, re-specialization of BM3 for the non-natural substrate propane was achieved after several rounds of directed evolution. $P450_{PMO}$ supports up to 45,800 propane turnovers with 98.2% coupling of substrate oxidation and cofactor consumption, rivaling those of natural P450s with their preferred substrates. In addition, we were able to demonstrate *in vivo* propane hydroxylation using these BM3 variants in resting *Escherichia coli* cells reaching activities up to 176 U g^{-1} cdw, which surpasses the reported activities of natural alkane hydroxylases acting on their preferred substrates.

B. Introduction

Cytochrome P450 enzymes (P450s) are exceptional oxygenating catalysts (*1 – 2*) with enormous potential in drug discovery, chemical synthesis, bioremediation, and biotechnology (*3 – 4*). Compared to their natural counterparts, however, engineered P450s often exhibit poor catalytic and cofactor coupling efficiencies (*3*). Obtaining native-like catalytic proficiencies is a mandatory first step towards utilizing the power of these versatile oxygenases in chemical synthesis.

Cytochrome P450 BM3 (BM3) isolated from *B. megaterium* catalyzes the subterminal hydroxylation of long-chain (C_{12}–C_{20}) fatty acids (*5*). Its high activity and catalytic self-sufficiency (heme and diflavin reductase domains are fused in a single polypeptide chain) (*4-6*) makes BM3 an excellent platform for biocatalysis. However, despite numerous reports of the heme domain being engineered to accept nonnative substrates, including short-chain fatty acids, aromatic compounds, alkanes, and alkenes (*7 – 15*), reports of preparative-scale applications of BM3 remain scarce (*16 – 19*).

The native BM3 function is finely regulated through conformational rearrangements in the heme and reductase domains and possibly also through hinged domain motions (*5, 20 – 21*). Hydroxylation of fatty acids occurs almost fully coupled to cofactor (NADPH) utilization, 93–96% depending on the substrate (*22 – 23*). In the presence of nonnative substrates or variants containing amino acid substitutions, the mechanisms controlling efficient catalysis in P450s are disrupted (*24 – 25*), leading to the formation of reactive oxygen species and rapid enzyme inactivation (*5*). High coupling efficiencies on substrates whose physicochemical properties are substantially different from the native substrates have not been achieved, and typical coupling

efficiencies range from less than 1% to 40% (*8, 11, 14*). Strategies for addressing this "coupling problem" are needed in order to take engineered P450s to larger-scale applications.

Selective hydroxylation of small alkanes is a long-standing problem, for which no practical catalysts are available (*26 – 28*). In an effort to produce BM3-based biocatalysts for selective hydroxylation of small alkanes, we previously engineered an enzyme variant 35E11, which accepts propane and ethane as substrates (*29*). However, despite greater than 5,000 total turnover supported *in vitro*, the utility of this catalyst remained limited because of its poor *in vivo* performance (see below), which was mostly due to low coupling efficiencies between product formation to cofactor consumption (17.4% for propane and 0.01% for ethane oxidation). The goal of our work was to engineer a BM3 variant with native-like activity and coupling efficiency toward a structurally challenging, non-native substrate, propane, and evaluate the impact of these features on performance in preparative-scale biotransformations.

C. Results and Discussion

C.1. Complete mutagenesis of BM3 through a domain engineering strategy

We used a domain-based protein-engineering strategy, in which the heme, flavin mononucleotide (FMN), and flavin adenine dinucleotide (FAD) domains of variant 35E11 were evolved separately but evaluated in the context of the holoenzymes. As a final step, beneficial mutations identified in each sub-domain were recombined (Figure 2.1). Most previous engineering efforts have focused mutagenesis to the heme (hydoxylase) domain of BM3(*7, 9, 11 – 12, 14*), although mutations in the reductase and linker regions have been shown to also affect catalytic properties, specifically coupling of co-factor consumption with substrate hydroxylation (*29 – 31*). However, no systematic engineering efforts had been undertaken to engineer the complete 1,048 amino acid holoenzyme.

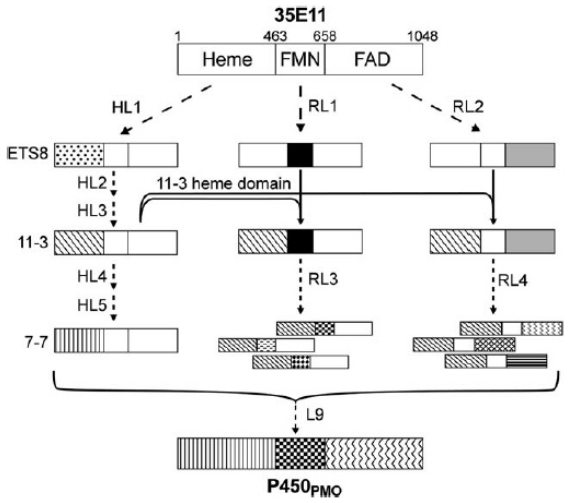


Figure 2.1: Outline of the domain engineering strategy used to improve cytochrome P450 BM3 heme and reductase domain (HL = heme domain library, RL = reductase domain libraries)

Holoenzyme libraries outlined in Figure 2.1 were created using random, saturation, and site-directed mutagenesis and screened for activity on a propane surrogate, dimethyl ether (*14*). Variants with improved dimethyl ether demethylation activity were confirmed by a rescreen,

purified, and characterized for propane hydroxylation activity using sealed head-space vials in the presence of a cofactor regeneration system. As a cumulative measure of both catalytic and coupling efficiency, improvement in total turnover number (TON), i.e., moles of propanol produced per mole of enzyme, was used as the sole selection criterion.

Measurement of the half-denaturation temperature (T_{50}), the temperature at which the enzyme retains 50% of its activity after a 15 minute incubation, of variant 35E11 heme domain demonstrated a considerable reduction in its thermostability as a consequence of the 15 accumulated mutations (T_{50}=43.4 ^{o}C vs. 55.0 ^{o}C for wild-type BM3). We therefore subjected variant 35E11 to an initial thermostabilization step (HL1), in which known stabilizing mutations from a thermostablized BM3 peroxygenase (*32*) were tested singly and in combination in the 35E11 background (see Table 2.1). Variant ETS8 (35E11-L52I-I366V), which showed the best combination of increased stability, ΔT_{50}=+5.1 ^{o}C, with only a small decrease in propane TON, $\Delta TON_{propane}$= -1,250, was selected for further directed evolution.

Table 2.1: Thermostablized variants of 35E11

Variant	Mutations					T_{50}[a] (^{o}C)	ΔT_{50} (^{o}C)	ΔPropane TON[b]
	L52	L234	V340	I366	E442			
35E11	-	-	-	-	-	43.4	n/a	n/a
ETS1	I	-	-	-	-	44.5	1.1	-510
ETS3	-	I	-	-	-	43.2	-0.3	-1,450
ETS4	-	-	M	-	-	46.0	2.6	-1,290
ETS5	-	-	-	V	-	47.1	3.7	-850
ETS6	-	-	-	-	K	45.0	1.6	+170
ETS8	I	-	-	V	-	48.5	5.1	-1,250
ETS9	I	-	-	-	K	46.8	3.4	-2,950
ETS10	-	-	M	-	K	44.2	0.8	-3,180
ETS11	-	-	-	V	K	46.6	3.2	-1750

[a] T_{50} calculated based on a two-state denaturation model using the percentage of 450 nm CO-binding peak of P450 variants remaining after 15-min incubations at varying temperatures

[b] TON determined as nmol product/nmol enzyme. Propane reactions contained 25 – 100 nM protein, potassium phosphate buffer saturated with propane, and an NADPH regeneration system containing 100 μM $NADP^{+}$, 2 U/mL isocitrate dehydrogenase, and 10 mM isocitrate. Errors are at most 10%.

Using ETS8 as parent, heme-domain random mutagenesis libraries were generated by error-prone PCR (HL2). Variant 19A12, ETS8-L188P, was identified from this library with more than a twofold increase in propane TON (Table 2.2). Using 19A12 as the parent, a pool of active-site libraries (HL3) were constructed in which 17 positions along the substrate channel and near the active site (74, 75, 82, 87, 88, 181, 184, 188, 260, 264, 265, 268, 328, 401, 437, and 438) were subjected individually to saturation mutagenesis. From these site-saturation libraries, further improvements in propane-hydroxylating activity were achieved in multiple variants, including 11-3 (19A12-A74S) which supported 13,200 propane TONs. Recombination of the beneficial mutations identified in these active-site variants (HL4, HL5) led to variant 1-3 (19A12-A74S-V184A) and variant 7-7 (19A12-A74E-S82G), supporting 19,200 and 20,500 propane TONs, respectively.

In parallel to the mutagenesis efforts targeting the BM3 heme domain, two libraries were constructed in which random mutations were targeted to the FMN and FAD binding domains of 35E11, RL1 and RL2, respectively. Screening of more than 5,000 members from each library for dimethyl ether demethylation led to the identification of eight beneficial mutations (G443D, V445M, T480M, T515M, P654Q, T664M, D698G, and E1037G). Of these eight mutations, G443D and V445M are actually located in the C-terminus β4 sheet of the heme domain, as RL1 included not only FMN domain but also the last 32 amino acids of the heme domain due to library construction. These eight positions were further optimized through saturation mutagenesis in a holoenzyme construct having the 11-3 heme domain (RL3, RL4). By swapping the heme domains, we aimed to remove mutations whose beneficial effect is solely dependent on the presence of the 35E11 heme domain. With the 11-3 heme domain, improved variants were

found to contain G443A, V445R, P654K, T664G, D698G, and E1037G mutations and supported propane TON between 16,000 and 20,000.

In the final step, a library containing the beneficial reductase domain mutations was fused to the heme domain of variant 7-7 (L9). The most active variant isolated from this library, $P450_{PMO}R2$, supported more than 45,800 turnovers and produced 2- and 1-propanol in a 9:1 ratio. As we expected, the increase in productivity strongly correlated with the increase in coupling efficiency, which in the best variant $P450_{PMO}R2$, 98.2% reaches levels comparable to those measured for wild-type BM3 in the hydroxylation of fatty acids, 88% for myristate, 93% for palmitate, and 95% for laurate (*22 – 23*).

Table 2.2: *In vitro* propane oxidation activities of most representative P450 BM3 variants[a]

Variant	Library	Mutations versus 35E11[b]		Rate[c] (min^{-1})	Coupling[d] (%)	Propane TON
		Heme domain	Reductase domain			
35E11	-	-	-	210	17.4	5,650
19A12	HL2	L52I, L188P, I366V	-	420	44.2	10,550
11-3	HL3	L52I, A74S, L188P, I366V	-	390	55.3	13,200
1-3	HL4	L52I, A74S, V184A, L188P, I366V	-	320	72.1	19,200
7-7	HL5	L52I, A74E, S82G, L188P, I366V	-	150	90.9	20,500
$P450_{PMO}R1$	L9	L52I, A74E, S82G, L188P, I366V, G443A	P654K, E1027G	455	94.4	35,600
$P450_{PMO}R2$	L9	L52I, A74E, S82G, L188P, I366V, G443A	D698G	370	98.2	45,800

[a] Mean values from at least three replicates ± 10 % error
[b] Mutations in 35E11 are R47C, V78F, A82S, K94I, P142S, T175I, A184V, F205C, S226R, H236Q, E252G, R255S, A290V, A328F, E464G, I710T.
[c] Over the first 20 s
[d] Ratio between propanol formation rate and NADPH oxidation rate in propane-saturated buffer

The sequence of mutational events leading to $P450_{PMO}$ generation reveals a continuous rearrangement of substrate channel and active-site residues (Table 2.2). The mutation of L188P

resulted in the single largest increase in propane hydroxylation activity, 2.4-fold relative to its parent, ETS8. Leu 188 is a helix capping residue located at the C-terminus of the F helix, which along with the G-helix forms a lid that undergoes a conformational change during catalysis. In a hinged motion, these two helices move from an "open" state in the absence of a substrates to a "closed" confirmation when substrate is bound (*33*). Accurately assessing the effect of this mutation, which removes the interstrand hydrogen bond provided by the Leu188 amid NH group, is difficult in the absence of an X-ray crystal structure. However, one likely outcome of this mutation could result in an enzyme resting state with the F and B' helix being in closer proximity, mimicking the substrate-bound confirmation. In the subsequent rounds of active site optimization, two different active site configurations, 19A12-A74S-V184A and 19A12-A74E-A82G were found to supported nearly the same number of propane TONs. However, the coupling of substrate oxidation with cofactor consumption differed significantly between these two variants, 72.1% vs. 90.9%. This 18.8% difference in coupling resulted in only a 6.7% difference in propane TON, which suggests that the benefit of higher coupling efficiency for improving enzyme activity diminishes at higher coupling.

Interestingly, the activity-enhancing substitutions in the reductase domain are clustered in the same region in the FAD domain (T664G, D698G, E1037G) and nearby linker to the FMN domain (P654K) (see Figure 2.2). Perturbation of electrostatic charge distribution appears to be a prevailing trend, suggesting a more important role of these forces in BM3 function than previously proposed (*34*). In contrast, no beneficial mutations were identified in FMN domain. This may reflect its higher sensitivity to mutagenesis, as judged by the significantly lower fraction of functional variants in the FMN libraries compared to the FAD libraries. In addition,

chemical and thermal denaturation studies have shown that, among the three cofactors, FMN is the most weakly bound to the enzyme (*35*).

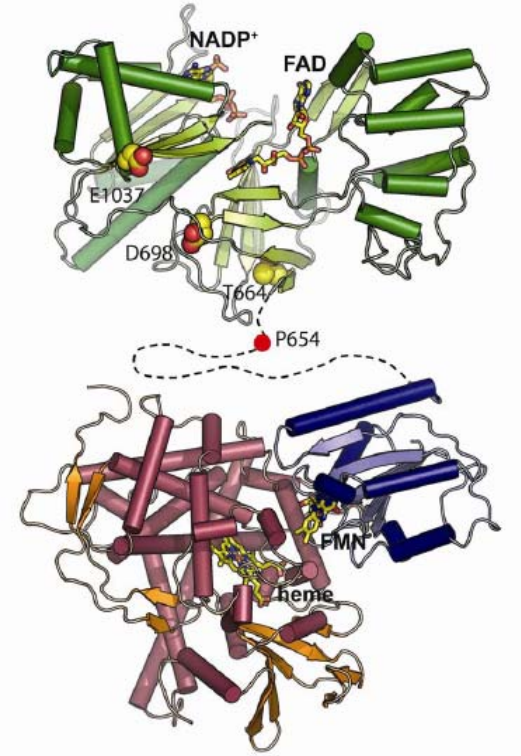


Figure 2.2: Map of the activity-enhancing reductase domain mutations on a homology model of P450BM3 FAD-binding domain prepared on the basis of the rat cytochrome P450 reductase structure (PDB: 1AMO (*36*)). Structural similarity between the two is supported by a preview of the solved but not yet published structure of P450BM3 FAD-binding domain (*6*). Heme domain and FMN domain are represented as in PDB: 1BVY (*37*). A 30 -residue linker connects the C-terminus of the FMN-binding with the N-terminus of the FAD-binding domain (dotted line).

C.2. Whole-cell bioconversion of alkane by BM3 variant with resting cells

A common strategy to reduce the prohibitive costs of NADPH-driven biotransformations is the use of cofactor regeneration systems (*16, 38 – 39*). For bulk chemical transformations such as alkane hydroxylation, these *in vitro* approaches are not economically viable (*40*). The propane-hydroxylating P450 variants were therefore evaluated in whole-cell biotransformations using resting *E. coli* cells. The expression levels of these variants in minimal medium were initially less than 0.5% of total cell mass. After optimization of growth and expression condition,

we were able to achieve expression of soluble P450s at 6–11% of total cell mass. The whole-cell biotransformations were conducted in 100 -mL fermenters using cell suspensions in nitrogen-free M9 minimal medium supplemented with glucose. The cell culture was continuously aerated with a 1:1 propane/air mixture to supply the substrate and oxidant. Under these conditions, cell densities of 0.5 – 0.9 g cdw L^{-1} were used to avoid oxygen-transfer limitations. Activities of 80–120 U g^{-1}cdw (where 1 U = 1 μmol propanol min^{-1}) were measured for $P450_{PMO}$-R1 and-R2 using various *E. coli* strains (Figure 2.3).

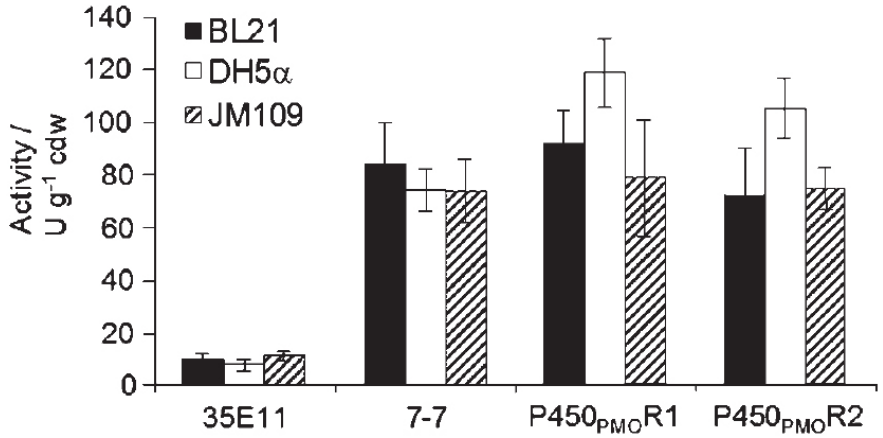


Figure 2.3: Whole-cell biotransformations of propane. Initial activities of selected P450 BM3 variants in different *E. coli* strains using air/propane (1:1) feed (pH 7.2, 25°C) measured after 1 h

The experiment was repeated in a larger fermenter (0.3 L, pH and dissolved oxygen control) with a suspension of P450-expressing DH5α cells, as DH5α was previously shown to be the most productive strain. The cell cultures were fed with a 1:1 mixture of pure oxygen and propane, and propanol formation was monitored for up to 9 h (Figure 2.4(a)). Under these conditions, very high activities (up to 176 U g^{-1} cdw) were obtained (Table 2.3). In comparison, the maximal activities of 30 U g^{-1}cdw on *n*-nonene were reported for the natural AlkB alkane hydroxylase system in both homologous (*P. oleovorans*) and heterologous strains (*E. coli*) (*41*).

Table 2.3: *In vivo* propane oxidation activities of P450BM3 variants[a]

Variant	Oxidant (propane/ oxidant ratio)	Activity [b] (U g^{-1}cdw) 0.5 h	3 h	Productivity [b,c] (mmol propanol g^{-1} P450 h^{-1})
35E11	air (1:1)	9	2	12
19A12	air (1:1)	41	9	44
7-7	air (1:1)	74	n.d.	88
$P450_{PMO}$R1	air (1:1)	118	73	119
$P450_{PMO}$R2	air (1:1)	104	68	106
$P450_{PMO}$R1	O_2 (1:1)	176	63	96
$P450_{PMO}$R2	O_2 (1:1)	119	39	94

[a] Mean values from two biological replicates ±15% error. n.d. = not determined
[b] At 0.5–0.9 g cdw L^{-1} cell density
[c] Calculated from the first hour of biotransformation

At higher cell densities (ca. 4 g cdw L^{-1}), propanol accumulated to a concentration of more than 15 mM over 4 hours (Figure 2.4 (b)). The improved coupling efficiencies resulted in considerably extended periods of whole-cell activity, 6 vs. 0.5 h (Figure 2.4 (a – b)) comparing $P450_{PMO}$ vs. 35E11. To investigate the possible causes for the decrease in productivity over time, we monitored the biocatalyst concentration over the course of the biotransformation (Figure 2.4 (c)). At the end of the experiment, approximately 52% of the initial $P450_{PMO}$R2 was still correctly folded. Control experiments using $P450_{PMO}$R2-expressing cells and propanol concentrations up to 30 mM showed no product inhibition or over-oxidation, suggesting that host related factors, rather than biocatalyst-dependent factors, are limiting. Indeed, 40–60% of the initially measured activity could be restored by resuspending cells from the plateau phase (i.e., after 4–6 h reaction) in fresh medium. In addition, the rate of biocatalyst inactivation could be reduced by varying the relative concentration of oxygen in the gas feed, with more extended whole-cell activity periods obtained at a propane/oxygen ratio of 4:1 compared to 1:1 (Table 2.3). Optimization of this parameter as well as the availability of more robust host strains (*40*) is expected to further enhance the whole-cell productivity of this engineered BM3 variant.

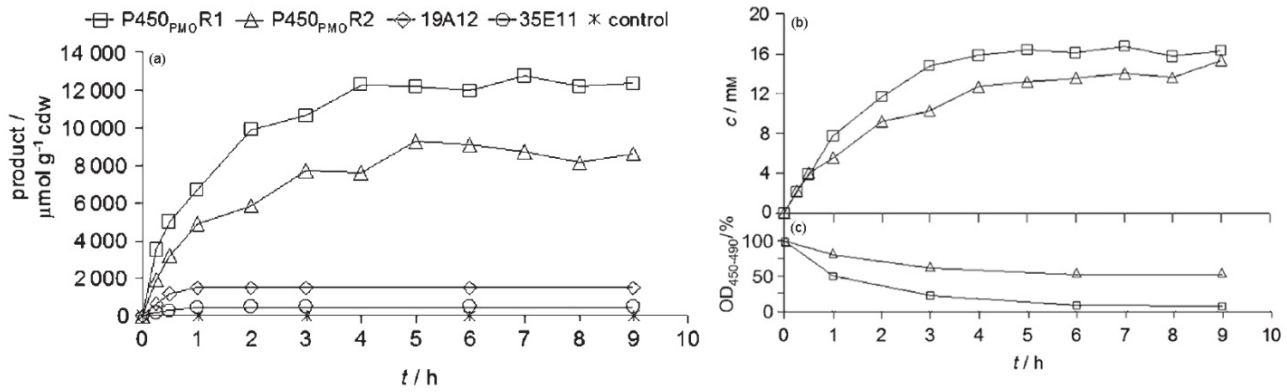


Figure 2.4: (a) Time course of propane biotransformation using recombinant DH5α cells using oxygen/propane (1:1) feed (pH 7.2, 25°C). Product amount is given per gram cell dry weight to facilitate comparison among variants. Control: no propane in the gas feed. (b) Concentration of propanol during biotransformation of propane with DH5α cells expressing P450PMOR1 (Δ) and P450PMOR2 (□) at medium cell density (4g cdw L^{-1}). (b) Relative P450 concentration as determined from CO-binding difference spectra on cell lysate; OD = optical density

D. Conclusion

Overall, a domain-based directed evolution strategy has enabled us to engineer a finely-tuned, multicofactor, multidomain enzyme to exhibit native-like catalytic properties on a substrate significantly different from the native substrate. With this approach, we could use relatively small and targeted libraries to identify beneficial mutations throughout the enzyme, which were recombined to yield the most efficiently engineered P450 reported to date. This strategy should prove useful for engineering other enzymes with multiple, interacting functional domains. With high activity and coupling efficiency for propane oxidation, $P450_{PMO}$s could be used in whole-cell biohydroxylation of propane at room temperature and pressure with air as oxidant. Total activities and product formation rates exceeding those obtained with naturally occurring alkane monooxygenases on their native substrates (*41 – 45*) were achieved in this first report of whole-cell bioconversion of propane to propanol in *E. coli.* These results open the door to considering P450-based oxidations of short-chain alkanes, with promise for green conversion of gaseous hydrocarbons into liquid fuels and chemicals.

E. References

1. Denisov, I. G., Makris, T. M., Sligar, S. G., and Schlichting, I. (2005) Structure and chemistry of cytochrome P450, *Chem. Rev. 105*, 2253-2277.
2. Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) Heme-containing oxygenases, *Chem. Rev. 96*, 2841-2887.
3. Bernhardt, R. (2006) Cytochromes P450 as versatile biocatalysts, pp 128-145, Elsevier Science Bv.
4. Urlacher, V. B., and Eiben, S. (2006) Cytochrome P450 monooxygenases: perspectives for synthetic application, *Trends Biotechnol. 24*, 324-330.
5. Munro, A. W., Leys, D. G., McLean, K. J., Marshall, K. R., Ost, T. W. B., Daff, S., Miles, C. S., Chapman, S. K., Lysek, D. A., Moser, C. C., Page, C. C., and Dutton, P. L. (2002) P450BM3: the very model of a modern flavocytochrome, *Trends Biochem.Sci. 27*, 250-257.
6. Warman, A. J., Roitel, O., Neeli, R., Girvan, H. M., Seward, H. E., Murray, S. A., McLean, K. J., Joyce, M. G., Toogood, H., Holt, R. A., Leys, D., Scrutton, N. S., and Munro, A. W. (2005) Flavocytochrome P450BM3: an update on structure and mechanism of a biotechnologically important enzyme, *Biochem. Soc. Trans. 33*, 747-753.
7. Appel, D., Lutz-Wahl, S., Fischer, P., Schwaneberg, U., and Schmid, R. D. (2001) A P450BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes, *Journal of Biotechnology 88*, 167-171.
8. Carmichael, A. B., and Wong, L. L. (2001) Protein engineering of Bacillus megaterium CYP102 - The oxidation of polycyclic aromatic hydrocarbons, *Eur. J. Biochem. 268*, 3117-3125.
9. Glieder, A., Farinas, E. T., and Arnold, F. H. (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase, *Nat. Biotechnol. 20*, 1135-1139.
10. Kubo, T., Peters, M. W., Meinhold, P., and Arnold, F. H. (2006) Enantioselective epoxidation of terminal alkenes to (R)- and (S)-epoxides by engineered cytochromes P450BM-3, *Chemistry-a European Journal 12*, 1216-1220.
11. Li, Q. S., Ogawa, J., Schmid, R. D., and Shimizu, S. (2001) Engineering cytochrome P450BM-3 for oxidation of polycyclic aromatic hydrocarbons, *Applied and Environmental Microbiology 67*, 5735-5739.
12. Li, Q. S., Ogawa, J., Schmid, R. D., and Shimizu, S. (2001) Residue size at position 87 of cytochrome P450BM-3 determines its stereoselectivity in propylbenzene and 3-chlorostyrene oxidation, *FEBS Lett. 508*, 249-252.
13. Ost, T. W. B., Miles, C. S., Murdoch, J., Cheung, Y. F., Reid, G. A., Chapman, S. K., and Munro, A. W. (2000) Rational re-design of the substrate binding site of flavocytochrome P450BM3, *FEBS Lett. 486*, 173-177.

14. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.

15. Sulistyaningdyah, W. T., Ogawa, J., Li, Q. S., Maeda, C., Yano, Y., Schmid, R. D., and Shimizu, S. (2005) Hydroxylation activity of P450BM-3 mutant F87V towards aromatic compounds and its application to the synthesis of hydroquinone derivatives from phenolic compounds, *Applied Microbiology and Biotechnology 67*, 556-562.

16. Falck, J. R., Reddy, Y. K., Haines, D. C., Reddy, K. M., Krishna, U. M., Graham, S., Murry, B., and Peterson, J. A. (2001) Practical, enantiospecific syntheses of 14,15-EET and leukotoxin B (vernolic acid), *Tetrahedron Lett. 42*, 4131-4133.

17. Maurer, S. C., Kuhnel, K., Kaysser, L. A., Eiben, S., Schmid, R. D., and Urlacher, V. B. (2005) Catalytic hydroxylation in biphasic systems using CYP102A1 mutants, *Advanced Synthesis & Catalysis 347*, 1090-1098.

18. Schneider, S., Wubbolts, M. G., Oesterhelt, G., Sanglard, D., and Witholt, B. (1999) Controlled regioselectivity of fatty acid oxidation by whole cells producing cytochrome P450(BM-3) monooxygenase under varied dissolved oxygen concentrations, *Biotechnology and Bioengineering 64*, 333-341.

19. Sowden, R. J., Yasmin, S., Rees, N. H., Bell, S. G., and Wong, L. L. (2005) Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450(cam) and P450(BM-3), *Org. Biomol. Chem. 3*, 57-64.

20. Haines, D. C., Tomchick, D. R., Machius, M., and Peterson, J. A. (2001) Pivotal role of water in the mechanism of P450BM-3, *Biochemistry 40*, 13456-13465.

21. Murataliev, M. B., and Feyereisen, R. (1996) Functional interactions in cytochrome P450BM3. Fatty acid substrate binding alters electron-transfer properties of the flavoprotein domain, *Biochemistry 35*, 15029-15037.

22. Cryle, M. J., Espinoza, R. D., Smith, S. J., Matovic, N. J., and De Voss, J. J. (2006) Are branched chain fatty acids the natural substrates for P450(BM3)?, *Chemical Communications*, 2353-2355.

23. Noble, M. A., Miles, C. S., Chapman, S. K., Lysek, D. A., Mackay, A. C., Reid, G. A., Hanzlik, R. P., and Munro, A. W. (1999) Roles of key active-site residues in flavocytochrome P450 BM3, *Biochem. J. 339*, 371-379.

24. Kadkhodayan, S., Coulter, E. D., Maryniak, D. M., Bryson, T. A., and Dawson, J. H. (1995) Uncoupling oxygen-transfer and electron-transfer in the oxygenation of camphor analogs by cytochrome P450cam - direct observation of an intermolecular isotope effect for substrate C-H activation, *Journal of Biological Chemistry 270*, 28042-28048.

25. Loida, P. J., and Sligar, S. G. (1993) Molecular recognition in cytochrome-p-450 - mechanism for the control of uncoupling reactions, *Biochemistry 32*, 11530-11538.

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

27. Shul'pin, G. B., Suss-Fink, G., and Shul'pina, L. S. (2001) Oxidations by the system "hydrogen peroxide-manganese(IV) complex-carboxylic acid" Part 3. Oxygenation of ethane, higher alkanes, alcohols, olefins and sulfides, *J. Mol. Catal. A-Chem. 170*, 17-34.

28. Yamanaka, I., Hasegawa, S., and Otsuka, K. (2002) Partial oxidation of light alkanes by reductive activated oxygen over the (Pd-black + VO(acac)(2)/VGCF) cathode of H-2-O-2 cell system at 298 K, *Appl. Catal. A-Gen. 226*, 305-315.

29. Meinhold, P., Peters, M. W., Chen, M. M. Y., Takahashi, K., and Arnold, F. H. (2005) Direct conversion of ethane to ethanol by engineered cytochrome P450BM3, *Chembiochem 6*, 1765-1768.

30. Govindaraj, S., and Poulos, T. L. (1995) Role of the linker region connecting the reductase and heme domains in cytochrome P450(BM-3), *Biochemistry 34*, 11221-11226.

31. Roitel, O., Scrutton, N. S., and Munro, A. W. (2003) Electron transfer in flavocytochrome P450BM3: Kinetics of flavin reduction and oxidation, the role of cysteine 999, and relationships with mammalian cytochrome P450 reductase, *Biochemistry 42*, 10809-10821.

32. Salazar, O., Cirino, P. C., and Arnold, F. H. (2003) Thermostabilization of a cytochrome P450 peroxygenase, *Chembiochem 4*, 891-893.

33. Arnold, G. E., and Ornstein, R. L. (1997) Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: Cytochrome P450BM-3, *Biophys. J. 73*, 1147-1159.

34. Davydov, D. R., Kariakin, A. A., Petushkova, N. A., and Peterson, J. A. (2000) Association of cytochromes P450 with their reductases: Opposite sign of the electrostatic interactions in P450BM-3 as compared with the microsomal 2B4 system, *Biochemistry 39*, 6489-6497.

35. Munro, A. W., Lindsay, J. G., Coggins, J. R., Kelly, S. M., and Price, N. C. (1996) Analysis of the structural stability of the multidomain enzyme flavocytochrome P-450 BM3, *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology 1296*, 127-137.

36. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes, *Proceedings of the National Academy of Sciences of the United States of America 94*, 8411-8416.

37. Sevrioukova, I. F., Li, H. Y., Zhang, H., Peterson, J. A., and Poulos, T. L. (1999) Structure of a cytochrome P450-redox partner electron-transfer complex, *Proceedings of the National Academy of Sciences of the United States of America 96*, 1863-1868.

38. Maurer, S. C., Schulze, H., Schmid, R. D., and Urlacher, V. (2003) Immobilisation of P450BM-3 and an NADP(+) cofactor recycling system: Towards a technical application of heme-containing monooxygenases in fine chemical synthesis, *Advanced Synthesis & Catalysis 345*, 802-810.

39. Schwaneberg, U., Otey, C., Cirino, P. C., Farinas, E., and Arnold, F. H. (2001) Cost-effective whole-cell assay for laboratory evolution of hydroxylases in *Escherichia coli*, *J. Biomol. Screen 6*, 111-117.

40. Duetz, W. A., van Beilen, J. B., and Witholt, B. (2001) Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis, *Current Opinion in Biotechnology 12*, 419-425.

41. Staijen, I. E., van Beilen, J. B., and Witholt, B. (2000) Expression, stability and performance of the three-component alkane mono-oxygenase of *Pseudomonas oleovorans* in *Escherichia coli*, *Eur. J. Biochem. 267*, 1957-1965.

42. Fujii, T., Narikawa, T., Takeda, K., and Kato, J. (2004) Biotransformation of various alkanes using the *Escherichia coli* expressing an alkane hydroxylase system from Gordonia sp TF6, *Biosci. Biotechnol. Biochem. 68*, 2171-2177.

43. Furuto, T., Takeguchi, M., and Okura, I. (1999) Semicontinuous methanol biosynthesis by Methylosinus trichosporium OB3b, *J. Mol. Catal. A-Chem. 144*, 257-261.

44. Kubota, M., Nodate, M., Yasumoto-Hirose, M., Uchiyama, T., Kagami, O., Shizuri, Y., and Misawa, N. (2005) Isolation and functional analysis of cytochrome P450 CYP153A genes from various environments, *Biosci. Biotechnol. Biochem. 69*, 2421-2430.

45. Lee, S. G., Goo, J. H., Kim, H. G., Oh, J. I., Kim, Y. M., and Kim, S. W. (2004) Optimization of methanol biosynthesis from methane using *Methylosinus trichosporium* OB3b, *Biotechnol. Lett. 26*, 947-950.

Chapter 3

Active Site Engineering of P450 BM3 for Small Alkane Hydroxylation

A. Abstract

To compare the functional richness of mutagenesis libraries generated by error-prone PCR, site-saturation mutagenesis, combinatorial active site saturation with a reduced set of amino acids and structure-based computational library design, seventeen mutagenesis libraries of cytochrome P450 BM3 were designed and constructed. Each library was evaluated for the fraction of variants that had acquired activity for demethylation of dimethyl ether and selected variants were also characterized for propane and ethane hydroxylation. Among these libraries, the ones generated by combinatorial active site saturation with a reduced set of amino acids displayed both a higher fraction of functional variants and variants with higher activity than both an error-prone PCR library with a similar mutation rate (2.1 mutation/protein) and site-saturation mutagenesis libraries targeting the same three residues. The most effective library design for generating variants for both dimethyl ether demethylation and small alkane hydroxylation was the CRAM algorithm developed and described here. While none of the isolated variants of this study achieved the level of specialization for propane hydroxylation previously obtained through multiple rounds of mutagenesis and selection, the levels of activity achieved by these variants show that jumps in sequence space from a specialized enzyme to generalist variants with desired functions are possible through various semi-rational mutagenesis approaches.

B. Introduction

Over the past decades, directed protein evolution has become a versatile tool for both the engineering of protein properties to meet industrial demands (*1 – 2*) and the exploration of structure-function relationships of biocatalysts (*3 – 4*). Using iterative cycles of sequence diversification and functional selection, enzyme variants have been reported with a variety of improved protein functions such as binding, enantioselectivity, thermostability, and altered substrate specificity (*5 – 8*). Recent advances in computational modeling (*9*), combined with the increased availability of structural and sequence information, have resulted in an expansion in the number of mutagenesis techniques and methods employed in directed protein evolution, such as SCOPE (*10*), CASTing (*11*), ISM (*12*), ISOR (*13*), and other structure-based computational library designs (*14 – 16*).

These semi-rational mutagenesis approaches aim to generate functionally enriched libraries by targeting mutations to specific regions of a protein such as an enzyme's active site or a protein-ligand interface determined to be important by structural or sequence analysis (*17*). While the viability of all these methods has been demonstrated by successful examples of their implementation, there have been very few attempts at comparing them with more traditional mutagenesis techniques such as error-prone PCR (EP-PCR) (*18*) or site-saturation mutagenesis (*19*). Part of the difficulty in comparing mutagenesis approaches is the inherent stochastic element of the directed evolution experiment. Since the outcome of such experiments relies on the specific choice of variants selected as parent for the subsequent round of evolution, repeating the same directed evolution experiment can lead to different sequence solutions. Therefore, comparing only the best variants generated through different mutagenesis techniques provides only anecdotal evidence for a method's efficacy. A more informative comparison of mutagenesis

techniques would be to evaluate the range of acquired activities and fraction of functional variants generated by each method in a single round of mutagenesis and screening for a defined function, using an identical starting point.

Here, we evaluate four mutagenesis approaches, (1) random mutagenesis by EP-PCR, (2) site-saturation mutagenesis (SSM), (3) Combinatorial Active Site saturation Test with a reduced set of amino acids (reduced CASTing) (*20*), and (4) two-structure-based computational library design approaches (*16*), for their ability to generate cytochrome P450 BM3 (BM3) (*21*) variants with activity for demethylation of dimethyl ether (DME) and hydroxylation of propane and ethane. BM3 is a self-sufficient fusion protein composed of a P450 monooxygenase and an NADPH diflavin reductase that hydroxylates C_{12}-C_{20} fatty acids as its preferred substrates (*22*) and does not have any detectable activity on these three substrates. Previous efforts in our lab generated variant $P450_{PMO}$ (PMO) (*23*) through 16 rounds of mutagenesis with 23 mutations with activity on all three substrates. In addition, PMO accepts propane as its preferred alkane substrate. The evolutionary strategy of enhancing the promiscuous alkane hydroxylation activity of BM3 and subsequent variants used to obtain PMO mimicked a natural evolution pathway and demonstrated that these functions can be acquired upon iterative rounds of mutagenesis and screening. The existing evolutionary lineage from BM3 to PMO allows us to compare the variants generated by these different mutagenesis approaches to determine the degree of specialization that can be obtained through semi-rational library design.

C. Results

C.1. Library design and composition

We designed and constructed 17 mutagenesis libraries of BM3 using EP-PCR (1), SSM (10), reduced CASTing (4) and structure-based computational methods (2), with library compositions listed in Table 3.1 and as described in Chapter 8.D. The residues targeted for SSM, reduced CASTing, and computationally-guided libraries were determined using the crystal structure of the BM3 heme domain bound with N-palmitoyl glycine, PDB:1JPZ (*24*). We identified ten residues (A74, L75, V78, A82, F87, L181, A184, L188, A328, and A330) as mutagenesis targets for both SSM and structure-based computational library design. These residues fall within various substrate recognition sites (SRS) identified for class II P450s (*25*), see Figure 3.1, and were selected over adjacent candidates because their side chains are oriented directly toward the active site. In addition, six of these ten residues (A74, V78, A82, A184, L188, and A328) were previously mutated in the evolutionary path from BM3 to PMO (*23*) and thus were known sites of beneficial mutations.

For the reduced CASTing libraries, mutations were targeted to three residues, V78, A82, and A328, as they were previously found to shift BM3's substrate specificity toward smaller substrates (*26 – 27*). The allowed amino acid cassette was restricted to L, I, M, V, F, A, and W with the use of degenerate codons following the intuition that introducing amino acids with large hydrophobic side chains should improve activity for smaller substrates. Four libraries were constructed, three of which mutated two of the three residues pairwise, and one library which mutated all three targets together. This library will be referred to as the three-site reduced CASTing library.

Instead of constructing a structural model of the BM3 transition state with small alkanes, we elected to use computational tools to find sequences that would maintain the substrate-bound conformation in the absence of substrate for the structure-based computational library design. This choice was made because a structure of the BM3 substrate-enzyme complex in a reactive conformation is not available (*28 – 29*), and the small alkane substrates lack functional groups that would aid the computational design in stabilizing potential transition states. The two designed libraries, C^{orbit} and CRAM, mutated the same ten residues as the site-saturation libraries, but only allowed for two possible amino acids at each position as determined by each algorithm (see Table 3.1). The C^{orbit} algorithm, which has previously been successful in creating diverse libraries of green fluorescent protein (*16*), models protein stability as a surrogate for protein function. In this approach, the selected mutations appeared more frequently in sequences predicted to support the desired enzyme fold. The second approach, which we termed the CRAM algorithm, aggressively packed the active site by computationally determining the largest tolerated amino acid substitutions at each of the ten target positions.

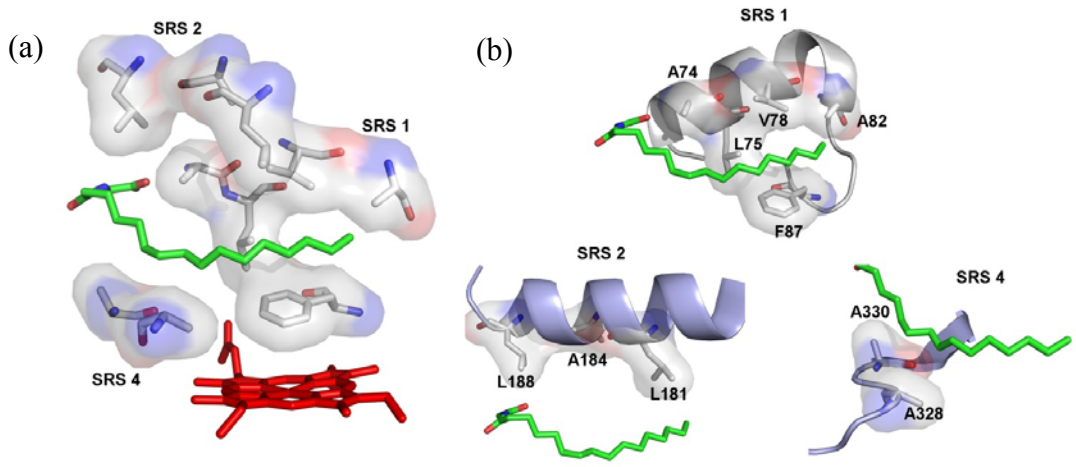


Figure 3.1: (a) The ten residues targeted for site-saturation mutagenesis and structure-based computational library design chosen based on their proximity to the bound N-palmitoyl glycine substrate in the 1JPZ structure of BM3 (*24*); (b) Close-up of SRS 1, 2, and 4 (*25*), heme shown in red. N-palmitoyl glycine is shown in green.

Table 3.1: Library designs and properties

Amino acid	EP-PCR	Site-saturation libraries					Reduced CASTing libraries				Computationally designed libraries	
		V78	A82	A328	A330	Remaining residues	V78/A82	V78/A328	A82/A328	V78/A82/ A328	CRAM	Corbit
A74	-	-	-	-	-	-	-	-	-	-	LW	AV
L75	-	-	-	-	-	-	-	-	-	-	LF	LF
V78	-	ALL[a]	-	-	-	-	LIVMFAW	LIVMFAW	-	LIVMFAW	IF	VL
A82	-	-	ALL	-	-	-	LIVMFAW	-	LIVMFAW	LIVMFAW	VL	AS
F87	-	-	-	-	-	-	-	-	-	-	FA	FA
L181	-	-	-	-	-	-	-	-	-	-	LW	LF
A184	-	-	-	-	-	-	-	-	-	-	AV	AT
L188	-	-	-	-	-	-	-	-	-	-	LW	LW
A328	-	-	-	ALL	-	-	-	LIVMFAW	LIVMFAW	LIVMFAW	VF	AF
A330	-	-	-	-	ALL	-	-	-	-	-	LW	AV
$<M_{AA}>$[b]	2.1	0.9	0.9	0.9	0.9	0.9	1.7	1.7	1.7	2.6	7.5	5
Library size (number of unique sequences)	1,166[c]	32	32	32	32	192	49	49	49	343	1,024	1,024
Clones screened	1,408	91	91	91	91	546	176	176	176	1,056	2,548	2,548
Library coverage (%)[d]	N/A	94	94	94	94	94	97	97	97	95	92	92
Fraction of folded variants[e]	0.52	0.9	0.96	0.62	0.96	0.95	0.91	0.95	0.97	0.94	0.75	0.84
Fraction of active variants[f]	0.36	0.26	0.21	0.36	0.12	0.01	0.22	0.71	0.74	0.54	0.34	0.31
Average DME activity	0.03	0.01	0.02	0.04	0.01	0.00	0.09	0.24	0.28	0.17	0.19	0.12
Standard deviation of DME activity	0.05	0.03	0.05	0.06	0.05	0.00	0.09	0.23	0.23	0.18	0.38	0.21

[a] All: 20 amino acids as encoded by the NNK codon used in the site-saturation mutagenesis. [b] $<M_{AA}>$: average mutation rate. [c] Estimation for the number of unique sequences sampled from 1,408 clones of the EP-PCR library was determined using PEDEL (*30*).[d]Library coverage was determined using GLUE (*30*). [e] Fraction of folded variants were determined by CO-binding spectroscopy, corrected for stop codon presence.
[f] Fraction of variants active for DME (of all variants) were determined in cell-free extract, corrected for background Purpald ® oxidation.

C.2. Library characterization for DME demethylation and protein folding

All 17 libraries were characterized for both DME demethylation activity and protein folding using high throughput assays; the results are summarized in Table 3.1. DME demethylation activity was quantified colorimetrically through the use of a dye, Purpald®, which reacts with the formaldehyde product of the P450 reaction to form a purple adduct with a UV/Vis peak maximum at 550 nm (*27*). The protein folding of each variant was determined by CO-binding difference spectra of the cell-free extract (*31*).

We screened 1,408 variants from a *Taq* polymerase generated EP-PCR library with an error rate of 2.1 amino acid substitutions/protein, which corresponds to sampling ~ 1,166 unique sequences, as calculated by the PEDEL algorithm for estimating the diversity of EP-PCR libraries (*30*). We found 52% of the variants to be folded (CO binding difference > 0.01) and 36% of the variants active for DME (Abs 550 > 0.13, corresponding to background Purpald® oxidation). The ten site-saturation libraries constructed with NNK codons, which encode for all 20 unique single-mutants, were screened to 94% library coverage (91 clones). Nine of the ten libraries contained a high fraction of folded variants, > 90%, with the library at A328 containing only 62% of folded variants. Variants that acquired DME activity were found in libraries targeting residues V78 (15%), A82 (9%), A328 (34%), and A330 (20%).

Given the high mutational tolerance of these active site residues for protein folding, it was unsurprising to find that the reduced CASTing libraries also contained a high fraction of folded variants, > 91%. The pairwise libraries, each having 49 unique members, were screened to 97% library coverage (176 clones), and the three-site library with 343 unique members was screened to 95% library coverage (1,056 clones). The fraction of functional variants varied from 22% for the library mutating V78/A82 to 71% and 74% for libraries mutating V78/A328 and

A82/A328, respectively. The three-site reduced CASTing library had only 54% of variants active for DME demethylation.

Finally, the two structure-based computationally designed libraries C^{orbit} and CRAM, each containing 1,024 unique members, were screened to 92% library coverage (2,548 clones), with 84% and 75% of folded variants, respectively. The fraction of variants with DME demethylation activity was similar between these two libraries, with 34% of CRAM variants and 31% of C^{orbit} variants being functional.

While all four mutagenesis strategies were able to generate variants with DME demethylation activity, the distribution of activity levels varied. The library profiles, i.e., activities of variants plotted in ranked order, for all libraries are shown in Figure 3.2. For simplicity and easier library comparisons, the variants from all ten SSM libraries were grouped and treated as single libraries for this analysis. Likewise, the three pairwise reduced CASTing libraries were also grouped.

Figure 3.2 (a) shows that both the EP-PCR library and the combined SSM libraries generated variants with DME activities up to 0.5 A_{550nm}, after correcting for background Purpald oxidation, with library averages of 0.026 $\pm$ 0.050 and 0.010 $\pm$ 0.039, respectively. These averages reflect the overall low functional richness of these libraries with a majority of both variant populations being inactive. In contrast, the variants of the reduced CASTing libraries exhibit DME activities up to 0.97 A_{550nm}, with library averages of 0.12 $\pm$ 0.19 and 0.17 $\pm$ 0.18 for the pairwise and three-site reduced CASTing libraries, respectively. A comparison of the library profiles of SSM libraries at these three residues, the pairwise, and three-site reduced CASTing libraries, Figure 3.2 (b), shows an increase in number of active variants with higher mutation rate and library size. However, the range of obtained DME activities only increased for the pairwise

reduced CASTing libraries compared to SSM libraries. The library profiles of the C^{orbit} and CRAM libraries show even higher DME activity: A_{550nm} of 2.6 was reached. However, as only ~ 30% of the variants are functional, the library averages of 0.18 ± 0.38 and 0.12 ± 0.21 are similar to the reduced CASTing libraries.

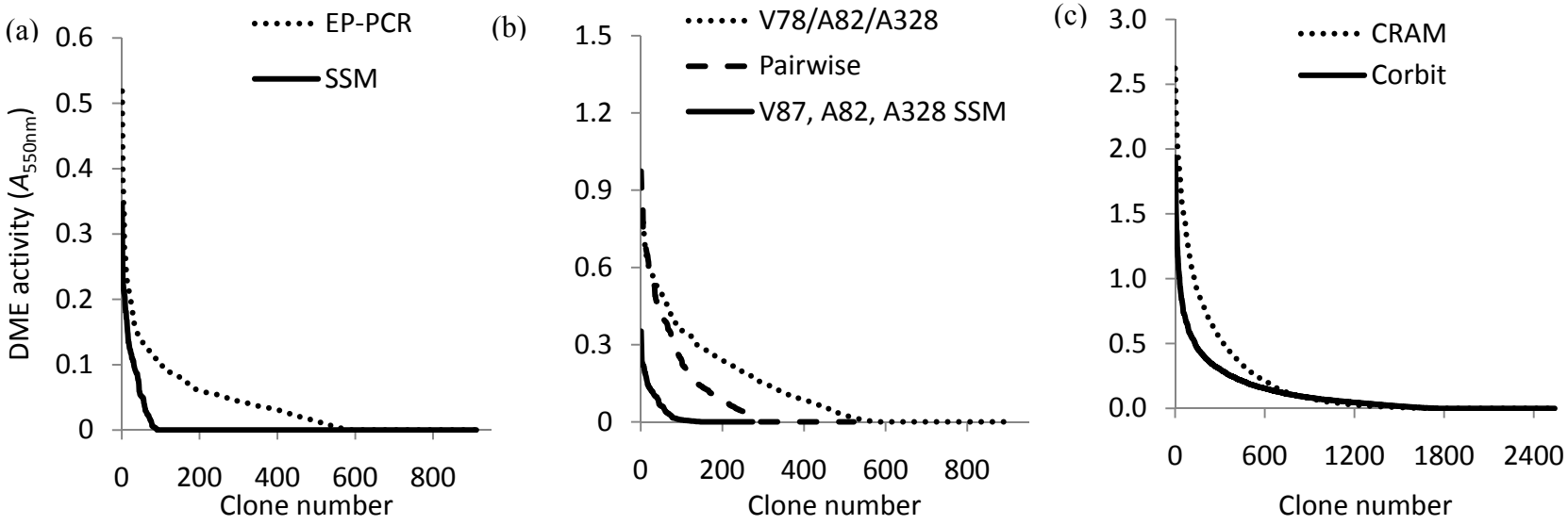


Figure 3.2: Profile of DME activity obtained by mutagenesis libraries with the activity of each variant plotted in ranked order. (a) Variants from the ten site saturation libraries and the top 910 variants of the EP-PCR library. (b) Variants from the site saturation, pairwise, and the complete reduced CASTing libraries targeting residues V78, A82, and A328. (c) Variants from C^{orbit} and CRAM libraries

C.3. Propane and ethane hydroxylation

After the DME demethylation screen was repeated for selected variants from each library in at least duplicate, top-performing variants were purified and characterized for propane and ethane hydroxylation activity (see Appendix A for complete sequence and activity information). Figure 3.3 (a) shows the histogram of propane and ethane turnover number (TON) of variants isolated from these libraries. From the ten SSM libraries, twelve variants were identified supporting propane TON ranging from 120 to 2,200. Mutations at V78 (T, C, S) and A82 (E,Q) located in the B' helix of SRS1 yielded variants with moderate propane activity, 120 to 370 TON. More active variants, > 1,000 propane TON, were obtained with mutations at residues A328 (I, P, L, V) and A330 (L, P, V), which are located in the loop between the J and K helices. The best single active-site variant, A328V, supports 2,200 propane TON with a product formation rate of 7.1 min^{-1} and 8.1% coupling of cofactor consumption.

Six variants were identified from the *Taq* EP-PCR library, supporting 130 to 3,300 propane TON. The best variant, 4F9 (F162L) supports 3,300 propane TON with a product formation rate of 19 min^{-1} and 15% coupling of cofactor consumption. The F162L mutation occurs in the linker between the E and F helices, located outside of the active site. While this residue was not mutated in variants of the PMO lineage, several residues in the adjacent F-helix were mutated, which may suggest the importance of this region for altering substrate specificity.

From the reduced CASTing libraries, nine variants were identified supporting 380 to 4,200 propane TON with only four of the nine variants containing mutations at all three targeted residues. The best variant, WT-A82L-A328V, supports 4,200 propane TON with a product formation rate of 40 min^{-1} and 44% coupling of cofactor consumption. In addition, two of these

nine variants, WT-A82L-A328L and WT-A82L-A328V, were also able to hydroxylate ethane, supporting 140 and 200 TON, respectively.

As far more variants from the CRAM and C^{orbit} libraries exhibited high DME demethylation activity compared to the other libraries, we selected the 88 most active variants from each library and screened them for propane and ethane hydroxylation directly as cell-free extracts using the assay outlined in Chapter 5.1. From this screen, 37 variants supporting at least 2,000 propane TON and 100 ethane TON as crude extracts were purified and characterized. All 37 variants were found to support at least 3,500 propane TON with 16 of the variants supporting at least 300 ethane TON as purified enzymes. A much higher number of active variants for propane and ethane hydroxylation was found in the CRAM library, 25 and 13, respectively, compared to the C^{orbit} library, which only produced twelve variants with activity on propane and three variants with ethane activity. The most active CRAM variant was E32 with mutations A74W, V78I, A82L, A184V, L188W, A328F, and A330W. E32 supported 16,800 propane TON and 1,200 ethane TON. The most active C^{orbit} variant, OD2, with mutations A74V, L181F, and A328F, supported 11,600 propane TON and 660 ethane TON. The coupling of cofactor consumption with propanol formation was also determined for a selection of these variants. Most variants exhibited coupling ranging from 36% – 52%, with the best variant, E31, having 68% coupling.

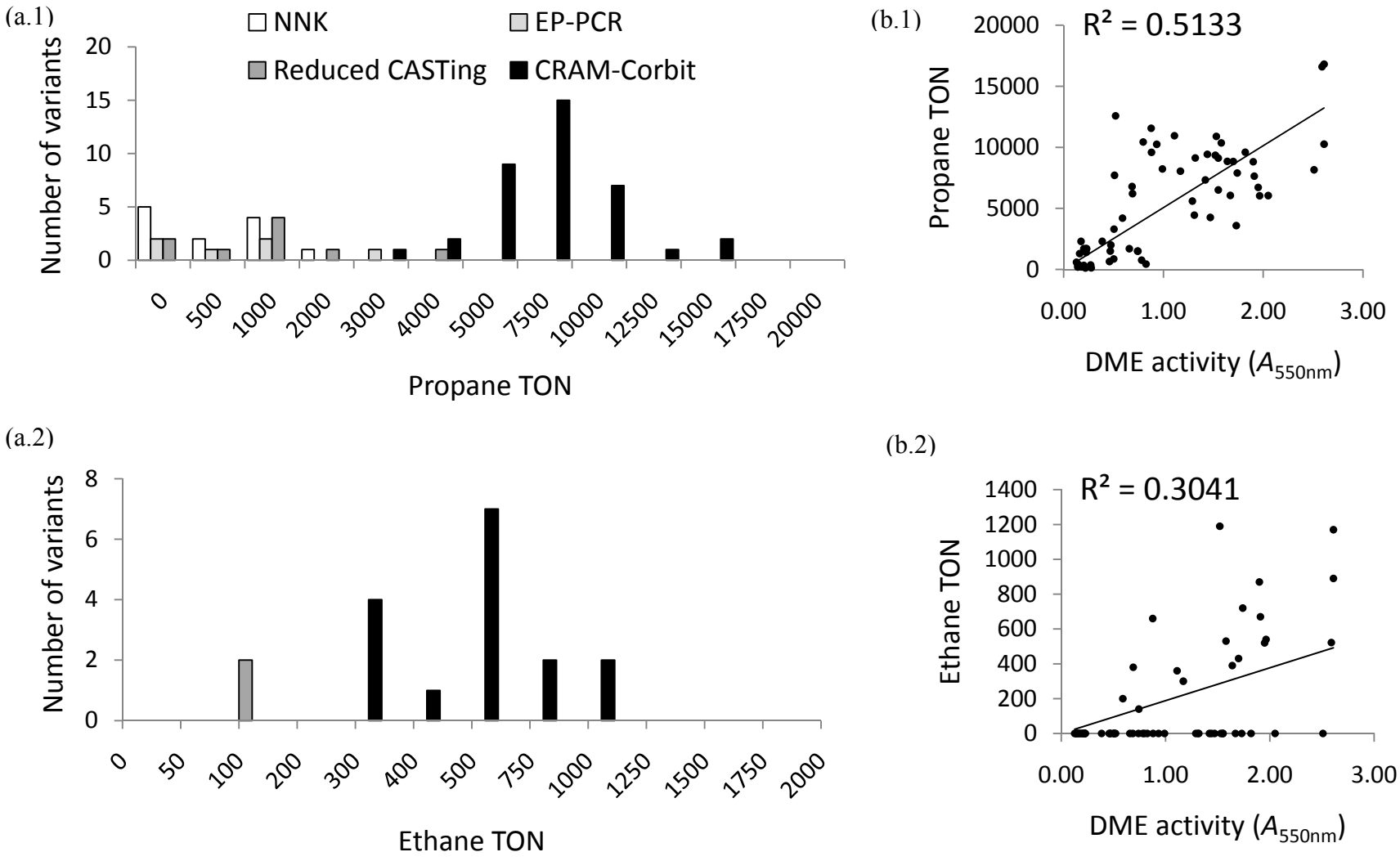
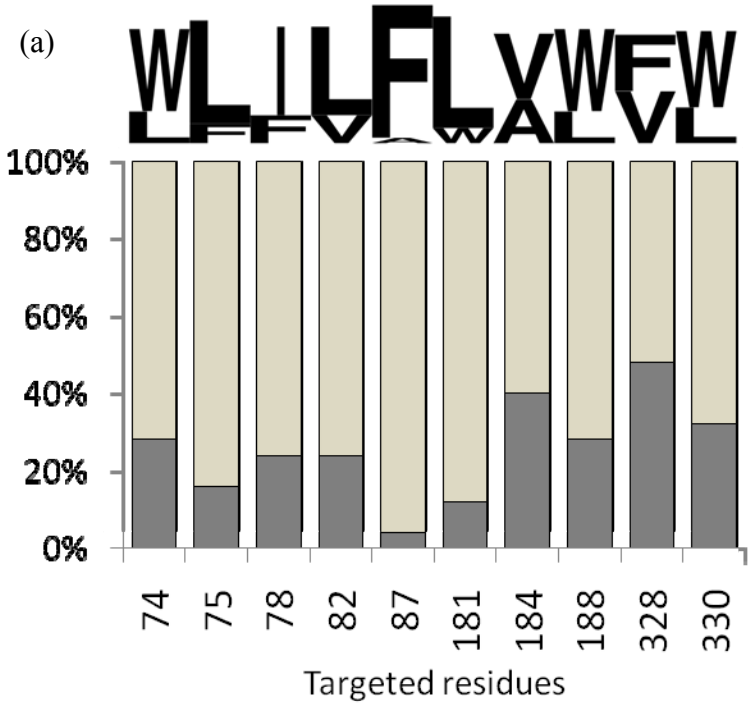


Figure 3.3: (a) Histogram of propane (a.1) and ethane (a.2) hydroxylating variants identified from various libraries. (b) Scatter plots of propane TON vs. DME activity (b.1) and ethane TON vs. DME activity (b.2) for all characterized variants

Figure 3.3 (b) shows the scatter plot of propane and ethane hydroxylation activities of all characterized variants vs. their DME demethylation activity. The data scatter of DME activity vs. propane TON, Figure 3.3 (b.1), appears to be normally distributed with a coefficient of determination (r^2) of 0.51 for linear regression of the data. In comparison, the data scatter of DME activity vs. ethane TON plot, Figure 3.3 (b.2), is not normally distributed, and a strong data bias exists representing variants with DME activity but unable to hydroxylate ethane. An r^2 of 0.30 is obtained for the linear correlation of DME activity and ethane TON. From these plots, we can conclude that both propane and ethane hydroxylation activity is positively correlated with DME demethylation since the p-values for the null hypothesis, i.e., random data scatter, are 1.5×10^{-11} and 1.26×10^{-6}, respectively. In addition, since very few variants with high DME activity were inactive for propane hydroxylation, DME demethylation is a good predictor for propane activity. However, more quantitative conclusions for the differences in predictability of DME demethylation for propane hydroxylation vs. ethane hydroxylation are difficult to determine.

By far the best source of variants with activity for propane and ethane hydroxylation is the designed library generated by the CRAM algorithm. The 25 propane hydroxylation variants with mutations at the same ten targeted residues form a concise and convenient data set for sequence analysis. The distribution of amino acids for all 25 variants, shown in Figure 3.4 (a), displays strong biases at seven of the ten targeted positions. Tryptophan appears at residue 74 and 188 in more than 72% of the sequences, likewise, strong preferences exist for L at positions 75 (84%), 82 (76%), and 181 (88%), I at 78 (76%), and F at 87 (96%). Of the remaining targeted positions, a weaker preference existed for W at 330 (68%), and V at 184 (60%), and nearly equal representations of both allowed amino acids were observed at position 328. Further sequence analysis of these variants accounting for their propane TON, Figure 3.4 (c – e), shows a fine-

tuning of amino acid preference and reduction of the sequence space with increased activity. For variants supporting less than 7,500 propane TON, a higher fraction of the less preferred amino acids are found at positions 74, 75, 78, 82, and 188. Proceeding to variants supporting higher propane TON, the occurrences of the less preferred amino acids decrease, culminating in nearly absolute preference for W at position 74, L at positions 75, 82, and 181, I at position 78, and F at position 87 for variants supporting more than 10,000 propane TONs. These results indicate that the screening process was able to find a narrow section of the total allowed sequence space containing the best solutions for propane hydroxylation. Comparing the amino acid preference of CRAM library variants, W74/**L75**/I78/L82/**F87**/**L181,** with the residues found in PMO**,** E74/**L75**/F78/G82/**F87**/**L181**, the positions with a preference for the wild-type amino acid (75, 87, 181) are not mutated in either of the CRAM variants or PMO, while mutations at the other positions (74, 78, 82) differ between the CRAM variants and PMO. The CRAM variants preferred larger hydrophobic residues at these locations, whereas PMO introduced both a charged and a hydrophilic amino acid. Although the choices of amino acids may differ, both sets of mutations may result in a similar constriction of the substrate channel. Due to both the close range of obtained activity and the low number of active variants produced by the other mutagenesis libraries, similar sequence analysis did not yield significant trends.



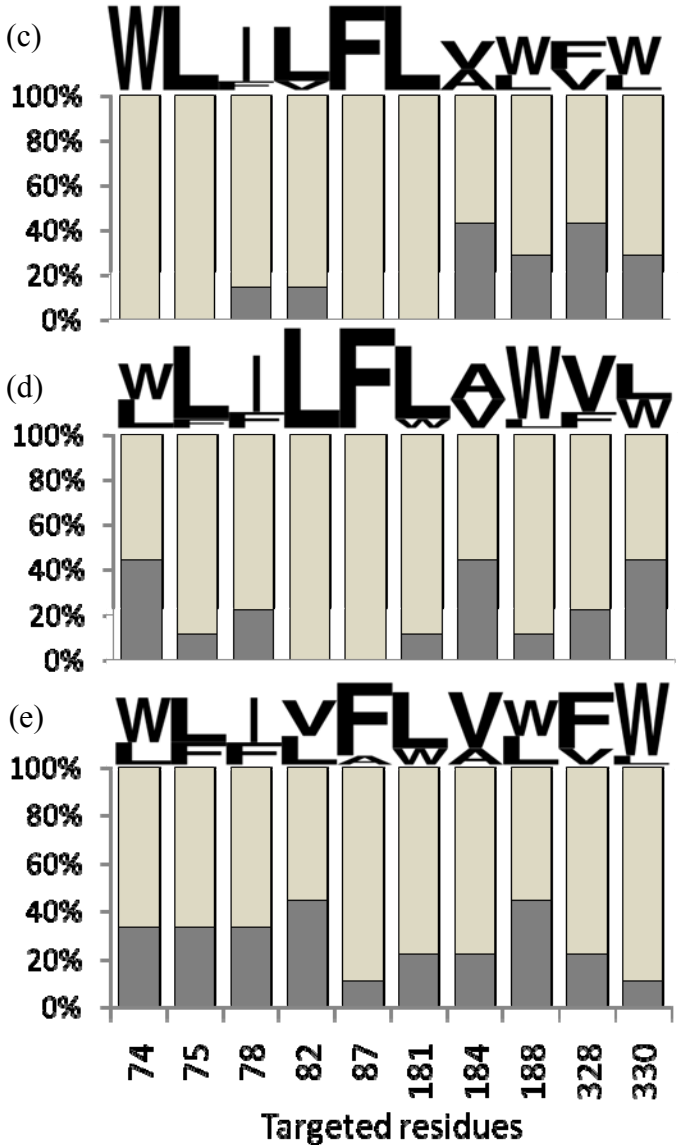


Figure 3.4: Amino acid distribution at each of the ten targeted positions of propane hydroxylating CRAM library variants: (a) all active variants, (b) identity of PMO's amino acid for these residues, (c) variants supporting less than 7,500 TON, (d) variants supporting 7,500 – 10,000 TON, (e) variants supporting > 10,000 TON, sequence logo generated by http://weblogo.berkeley.edu/

D. Discussion

Since variants with activity for DME demethylation and propane hydroxylation were identified from all the mutagenesis approaches we investigated, with as few as one mutation, it appears that finding variants with these two functions was much easier than we anticipated based on previous studies (*26, 32*). As a substrate, propane shares many similarities with hexane, the smallest known alkane hydroxylated by BM3. They are both hydrophobic with poor water solubility and possess sub-terminal alkane C-H bonds of comparable bond strength as BM3's preferred fatty acid substrates (99–100 kcal/mol). The only difference between propane and other known BM3 substrates is its smaller molecular size, which should result in a lower binding affinity. The major impact of poorly bound substrates on the P450 reaction mechanism is a weaker activation of the catalytic cycle, as the poorly bound substrates cannot displace the distal water-ligand to initiate catalysis (*33*).

For wild-type BM3, the propane binding event needs to be the sole trigger for the activation of catalysis, as the enzyme exhibits low resting state oxidase activity (~ 10 min^{-1}), which indicates that substrate-independent activation of the catalytic cycle occurs rarely. However, variants of BM3 can exhibit much higher resting state oxidase activity, thereby reducing the requirement of propane binding to induce catalysis. In fact, variants generated in the $P450_{PMO}$ lineage and many of the variants found in this study have substrate-free cofactor consumption up to an order of magnitude higher than that of the wild-type enzyme. The existence of this alternative pathway for propane hydroxylation activity, which can be achieved without appreciable propane-induced activation of the catalytic cycle (*34*), could explain the high number of identified propane-hydroxylating variants. The accessibility of this alternative

pathway for substrate hydroxylation still requires some degree of substrate binding affinity and should diminish for substrates with lower affinity, such as ethane.

Unlike the high number of variants isolated with activity for propane hydroxylation and DME demethylation, far fewer isolated variants exhibited ethane hydroxylation activity. One possible explanation for this result is the poor correlation between DME demethylation and ethane hydroxylation, as shown in Figure 3.3 (b.2). Since DME demethylation was the criterion used to filter variants for ethane hydroxylation characterization, a poor correlation between the activities would result in the elimination of ethane-hydroxylating variants with poor DME demethylation activity. Another potential explanation for this result is that variants with ethane hydroxylation activity are simply rarer in the sequence space that we investigated than variants with activity for propane hydroxylation and DME demethylation. This possibility is quite understandable since ethane is both smaller than propane and lacks the energetically favorable sub-terminal alkane C-H bond that is common to propane and BM3's preferred fatty acid substrates. Therefore, ethane hydroxylation presents challenges to not only the activation of the P450 catalytic cycle due to its smaller size, but also the ability of the P450 to break a ~ 1 kcal/mol stronger C-H bond.

Another finding from these mutagenesis libraries is the extremely high mutational tolerance of the BM3 active site. The high fraction of tolerated mutations observed with BM3 active site residues appears to contradict the general observation that mutations in the core of a protein are on average more destabilizing than mutations of solvent-exposed residues (*35*). Residues in the packed protein core generally have more interactions with neighboring amino acids than solvent-exposed residues. As a result, they have lower site entropy, which has been hypothesized to reflect decreased tolerance for mutation (*36*). However, the BM3 active site

residues inherently have a higher degree of flexibility compared to typical core residues, since the active site of BM3 undergoes significant motion between the "closed" substrate-bound state and its "open" resting state (*37*). The active site environment between these two states also differs significantly in terms of solvent accessibility (*33*), which may allow these positions to tolerate polar or even charged amino acid substitutions. Therefore, the flexible nature of the BM3 active site, which is atypical of packed protein core structures, may be responsible for the higher mutations tolerance of these residues.

In comparing the functional richness between the libraries generated by the various mutagenesis methods we investigated, it was surprising to find that the EP-PCR library appears to generate more active variants than the combined efforts of the SSM libraries (see Figure 3.2 (a)). Comparing the functional richness of EP-PCR mutagenesis with SSM is inherently subjective since a poor choice of mutagenesis sites or selection criteria can easily skew the efficacy of the site-saturation libraries. However, since comparable screening effort was required to evaluate the ten site-saturation libraries (910 clones) and the EP-PCR library (1,408 clones), this comparison is reasonable from a practical standpoint. Although the number of clones sampled from these libraries is similar, the actual sequence diversity is quite different: 200 unique variants for the combined NNK libraries vs. 1,166 expected unique variants for the EP-PCR library. This sixfold increase in number of unique sequences could account for the higher number of DME demethylating variants identified in the EP-PCR library. The screening efficacy of the site-saturation libraries can be improved with the use of more efficient codons such as NDT[1] (*20*) or a combination of codons to better match the number of unique nucleotide

[1] NDT codon degeneracy: N = A, T, C, and G; D = A, T, and G

sequences with the number of unique amino acid substitutions of the libraries. However, even with the best possible codon selection, to generate a set of site-saturation libraries with the same number of unique sequences as the EP-PCR library generated in this study would require nearly 60 site-saturation libraries. Clearly, EP-PCR is better than SSM at generating sequence diversity quickly and cheaply, at the cost of not controlling or knowing the sites of mutagenesis. Based on this inherent trade-off, EP-PCR should be superior to SSM in generating diversity for functions that are affected by mutations across protein structure, such as thermostability, solubility, or modulating promiscuous activity. Conversely, for functions that require specific changes in the enzyme's active site, such as altering regio- and enantio-selectivity, the easily generated sequence diversity of the EP-PCR library is wasted, since the majority of the mutations is not created at the necessary locations. Therefore, for these functions, SSM at active site residues has the potential to be more effective.

Comparing functional richness of the site-saturation libraries at V78, A82, and A328 with the reduced CASTing libraries mutating the same residues pairwise, Fig 3.2 (b), shows that the reduced-CASTing libraries are far better in both the range of activities obtained and the number of active variants. In terms of the sequence diversity in this comparison, the pairwise reduced CASTing libraries combine to have 147 unique variants, whereas the three SSM libraries combine to have only 60 possible variants, which should account for some of the differences in the observed functional richness. However, since the mutations found to support DME demethylation and propane hydroxylation at V78 (C, T, S) and A82 (E, Q) were not included in the allowed amino acids of the reduced CASTing libraries (L, I, V, F, M, A, and W), the pairwise reduced CASTing libraries found active variants that would not have been found through recombination of the beneficial point mutations. One obvious question is whether the

variants found from the reduced CASTing library are better than those that could have been isolated from either the recombination of the beneficial point mutants identified from site-saturation libraries or iterative rounds of SSM. We cannot answer this question directly, as neither approach was attempted. However, such comparisons would only be anecdotal and cannot determine which method is superior. Ultimately, all three approaches are flawed, as each makes a fallible assumption about the interaction of mutations. In the reduced CASTing library, the initial reduction of the allowed amino acids assumes that better solutions do not exist in the excluded amino acids. Likewise, the strategies of recombining beneficial single mutations or iterative site-saturation assumes that synergistic effects between mutations are minimal, and the best combination of mutations contains mutations found to be beneficial individually. How these assumptions accurately reflect the interaction of mutations for a particular protein will ultimately determine the efficacy of their application.

While the pairwise reduced CASTing libraries displayed both a higher range of obtained activities and a higher number of active variants than SSM libraries mutating the same residues, the range of DME demethylation activity obtained by the three-site reduced CASTing library did not increase compared to that of the pairwise reduced CASTing libraries. In addition, the fraction of functional variants for the three-site reduced CASTing library, 54%, was lower than the pairwise reduced CASTing library involving A328, 71% and 74%. This indicates that with the expanded sequence space, 343 vs. 49 library members, the larger library had more unique active sequences, but a larger fraction of the sequence space was occupied by inactive variants. One explanation for this result is that mutations at V78 and A82, which are located in close proximity, reduce the volume of the active site in the same region. Therefore, adding an additional mutation to the existing mutations at V78 and A328 or A82 and A328, which are

already sufficient for function, has only neutral or deleterious effects. Beyond this structure based argument, the reduced benefit of increasing the mutation rate of the CASTing library may be an inherent dilution effect analogous to that observed for random mutagenesis, where increasing the mutation rate beyond 1 – 2 amino acid substitutions results in lower quality libraries (*38*).

The single point mutations at V78, A82, A328, and A330 that resulted in variants with DME demethylation activity and propane hydroxylation activity generally introduced amino acids with bulkier side chains into the active site, which follows the intuition that reducing active site volume would promote activity for a smaller substrate. Since none of the PMO mutations, V78F, A82G, or A328F, were found at these positions, they are not beneficial individually. The presence of proline mutations at A328 and A330 suggests that altering the orientation of the loop containing these residues can result in improved activity in addition to mutations that simply reducing the active site volume. The best mutation, A328V, has been reported to affect fatty acid binding and cause a shift in regioselectivity of the hydroxylation reaction (*39*). A crystal structure of WT-A328V has been solved with N-palmitoyl glycine bound in the active site (PDB: 1ZOA) (*39*). A structure alignment of WT-A328V and wild-type BM3 (Figure 3.5) shows little structural deviation, with an overall RMDS of the α-carbons of only 0.23 Å. The largest deviation between the two structures occurs near the 13^{o} kink in the I-helix, the proposed site of oxygen binding (*40*). However, similar deviations can be observed between different structures in wild-type BM3 in this region, which could reflect the general flexibility rather than the result of this mutation. The methyl group of the valine side chain induces a slight twist in the bound substrate, N-palmitoyl glycine, at the C-5 carbon; otherwise the active site packing is identical.

The lack of gross changes in the active site packing between these two structures illustrates that propane hydroxylation activity is obtainable without significant structural deviations.

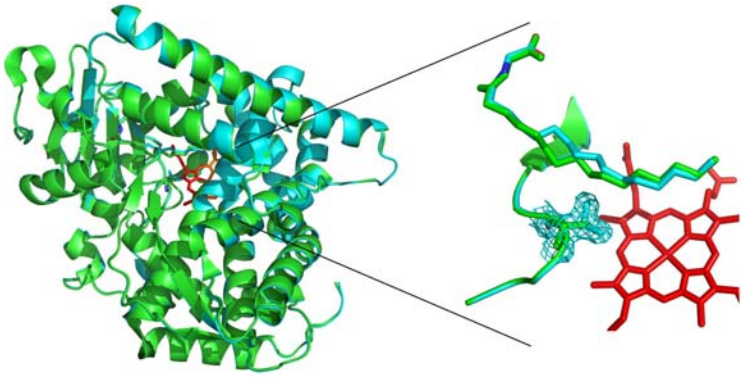


Figure 3.5: Structural alignment of BM3 (1JPZ (*24*)), shown in green, with BM3-A328V (1ZOA (*39*)), shown in cyan, heme shown in red. Close-up of the active site showing a shift of the bound N-palmitoyl glycine due to the presence of V328 side chain electron density

Of the seven mutations found through random mutagenesis, only I260V, located on the I-helix, is near the active site. The mutations F162L and I153V found in the more active variants are clustered in the region between the E and F helices. The remaining mutations E4D, T235M, D232V and Q359R, are at surface-exposed residues, which are typical of mutations found using random mutagenesis (*41*), whose effects are difficult to rationalize.

Across all nine variants isolated from the reduced CASTing libraries, V78L was the only mutation found at position 78, whereas more substitutions were beneficial at A82 (L, W, M, and V) and A328 (L, V, and F). Since the allowed amino acid set excluded glycine, only two PMO mutations, V78F and A328F, could have been found by these variants. Of these two mutations, only A328F was found in the isolated variants. Since this mutation was not found to be beneficial individually, its presence in PMO and these reduced CASTing variants suggests there are synergistic effects between this residue and neighboring amino acids. The effect of the V78L mutation on propane TON was also showed a dependence on the surrounding mutations: when V78L was introduced to WT-A328L or WT-A82L-A328L, the resulting variants lost 48% of the

parental propane activity. However, when V78L was mutated in WT-A82W-A328F, the resulting variant had sixfold improved propane activity. This illustrates the ruggedness of active site landscape, in which the effects of mutations are highly dependent on the identity of neighboring amino acids. While the effects of an amino acid substitution are heavily dependent on existing mutations, different amino acid substitutions at a given residue can result in nearly identical activities, for example, mutating A82 to either M and V produced nearly equal increases in propane TON, 3.8-and 3.3-fold in the same parental background of WT-V78L-A328L. These results suggest that reducing the allowed set of amino acids is a good trade-off for mutating more residues simultaneously, as different amino acid substitutions can achieve similar effects, a result which supports the structure-based computationally designed approach we pursued.

The most effective libraries we generated in this investigation for acquiring DME demethylation and propane hydroxylation activity are the CRAM and C^{orbit} structure-based computationally designed libraries. These two libraries mutated all ten targeted active site residues allowing for two possible amino acids at each position. By redesigning the active site in such a global fashion, we obtained variants with propane and ethane activity rivaling those achieved by variants of the $P450_{PMO}$ lineage (*34*). The best variant from the CRAM library, E32, supported 16,800 propane TON and 1,200 ethane TON, which are ~ 50% of PMO's activity on these substrates. The best variant from the C^{orbit} library, OD2, supported 11,600 propane TON and 660 ethane TON, which are 34% and 27% of PMO's activity on these substrates. As a comparison, these activity levels were obtained by variants of $P450_{PMO}$ lineage after 10 – 12 rounds of mutagenesis and screening.

Of the 37 variants we isolated from these two designed libraries, all allowed amino acids were found at least once. While all mutations were represented, there are clear biases in amino acid preference (see Figure 3.4). The consensus sequence of the active CRAM variants, W74, L75, I78, L82, F87, L181, V184, W188, F328, and W330, is actually the sequence of variant E32, the most active variant, which suggests that the screening process was able to find an optimal solution within the allowed sequence space (2^{10}). Of the six positions–75, 78, 87, 181, 184, and 328–where the PMO residue was an allowed choice by the library design, five positions converged on the PMO amino acid. This convergence on mutations found in PMO is not surprising since the PMO active site is a good solution for the selected activities. However, assigning significance to these amino acid preferences in the context of the total possible sequence space of these ten residues (10^{20}) is problematic, as each mutation was compared against only one other amino acid within a limited set of surrounding mutations. Therefore, the best solution obtained from the designed libraries is certainly not the optimal solution for the total sequence space. All we can ascertain from the CRAM library results is that, within the chosen subsection of the sequence space, multiple solutions for propane and ethane hydroxylation exist, and a locally optimal solution is obtainable.

These designed libraries also demonstrate that jumps in sequence space from BM3 to variants with moderate propane hydroxylation activity (~ 10,000 TON) are achievable. None of the obtained variants, however, reached the level of specialization that was previously obtained with $P450_{PMO}$, in either propane TON or coupling of cofactor consumption. In the evolution of BM3 to PMO, the specialization for propane hydroxylation did not occur evenly through the 16 rounds of mutagenesis. In fact, the variants of the lineage can be categorized into three distinct groups by their substrate specificity for linear alkanes as (1) preferring longer chain alkanes, (2)

having equal preference for alkanes of chain lengths C_3 – C_{10}, and (3) preferring shorter chains alkanes with the length of propane (*34*). These three groups of variants represent a transition from a specialized fatty acid hydroxylase to generalist P450s with broad alkane substrate acceptance followed by second transition to a specialized propane monooxygenase. This last transition occurs in the final four rounds of mutagenesis where the largest improvements in cofactor coupling (44% to 93%) and propane TON (10,550 to 33,400) occur. In addition, mutations acquired in these final rounds of mutagenesis are located not only in the P450 heme domain but also in the reductase domain. This suggests that mutations outside the active site or even the heme domain in general, may be necessary for functional optimization. The range of obtained propane TON (3,500 – 16,800) and coupling of co-factor consumption (36% – 68%) for variants identified by the designed libraries correspond to those values of the generalist intermediates found preceding the propane specialization phase of the PMO evolution. This suggests semi-rational library design can be an effective strategy to move away from a specialized enzyme toward generalist variants, but functional specialization still requires optimization through several rounds of random mutagenesis and screening.

E. References

1. Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W., and Hughes, G. J. (2010) Biocatalytic asymmetric synthesis of chiral amines from ketones applied to Sitagliptin manufacture, *Science 329*, 305-309.

2. Wohlgemuth, R. (2010) Biocatalysis - key to sustainable industrial chemistry, *Current Opinion in Biotechnology 21*, 713-724.

3. Fox, R. J., Davis, S. C., Mundorff, E. C., Newman, L. M., Gavrilovic, V., Ma, S. K., Chung, L. M., Ching, C., Tam, S., Muley, S., Grate, J., Gruber, J., Whitman, J. C., Sheldon, R. A., and Huisman, G. W. (2007) Improving catalytic function by ProSAR-driven enzyme evolution, *Nat. Biotechnol. 25*, 338-344.

4. Urban, P., Truan, G., and Pornpon, D. (2008) High-throughput enzymology and combinatorial mutagenesis for mining cytochrome P450 functions, *Expert Opin. Drug Metab. Toxicol. 4*, 733-747.

5. Kaur, J., and Sharma, R. (2006) Directed evolution: An approach to engineer enzymes, *Crit. Rev. Biotechnol. 26*, 165-199.

6. Reetz, M. T. (2011) Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions, *Angewandte Chemie-International Edition 50*, 138-174.

7. Tracewell, C. A., and Arnold, F. H. (2009) Directed enzyme evolution: climbing fitness peaks one amino acid at a time, *Curr. Opin. Chem. Biol. 13*, 3-9.

8. Turner, N. J. (2009) Directed evolution drives the next generation of biocatalysts, *Nature Chemical Biology 5*, 568-574.

9. Damborsky, J., and Brezovsky, J. (2009) Computational tools for designing and engineering biocatalysts, *Curr. Opin. Chem. Biol. 13*, 26-34.

10. O'Maille, P. E., Bakhtina, M., and Tsai, M. D. (2002) Structure-based combinatorial protein engineering (SCOPE), *Journal of Molecular Biology 321*, 677-691.

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

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

13. Herman, A., and Tawfik, D. S. (2007) Incorporating synthetic oligonucleotides via gene reassembly (ISOR): a versatile tool forgenerating targeted libraries, *Protein Eng. Des. Sel. 20*, 219-226.

14. Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Rothlisberger, D., Zanghellini, A., Gallaher, J. L., Betker, J. L., Tanaka, F., Barbas, C. F., Hilvert, D., Houk, K. N., Stoddard, B. L., and Baker, D. (2008) De novo computational design of retro-aldol enzymes, *Science 319*, 1387-1391.

15. Siegel, J. B., Zanghellini, A., Lovick, H. M., Kiss, G., Lambert, A. R., Clair, J. L. S., Gallaher, J. L., Hilvert, D., Gelb, M. H., Stoddard, B. L., Houk, K. N., Michael, F. E., and Baker, D. (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction, *Science 329*, 309-313.

16. Treynor, T. P., Vizcarra, C. L., Nedelcu, D., and Mayo, S. L. (2007) Computationally designed libraries of fluorescent proteins evaluated by preservation and diversity of function, *Proceedings of the National Academy of Sciences of the United States of America 104*, 48-53.

17. Bershtein, S., and Tawfik, D. S. (2008) Advances in laboratory evolution of enzymes, *Curr. Opin. Chem. Biol. 12*, 151-158.

18. Cadwell, R. C., and Joyce, G. F. (1994) Mutagenic PCR *PCR-Methods Appl. 3*, S136-S140.

19. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagensis without phenotypic selection, *Proceedings of the National Academy of Sciences of the United States of America 82*, 488-492.

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

21. Fulco, A. J. (1991) P450BM-3 and other inducible bacterial P450 cytochromes - biochemistry and regulation, *Annu. Rev. Pharmacol. Toxicol. 31*, 177-203.

22. Ost, T. W. B., Miles, C. S., Murdoch, J., Cheung, Y. F., Reid, G. A., Chapman, S. K., and Munro, A. W. (2000) Rational re-design of the substrate binding site of flavocytochrome P450BM3, *FEBS Lett. 486*, 173-177.

23. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.

24. Haines, D. C., Tomchick, D. R., Machius, M., and Peterson, J. A. (2001) Pivotal role of water in the mechanism of P450BM-3, *Biochemistry 40*, 13456-13465.

25. Pylypenko, O., and Schlichting, I. (2004) Structural aspects of ligand binding to and electron transfer in bacterial and fungal p450s, *Annu. Rev. Biochem. 73*, 991-1018.

26. Glieder, A., Farinas, E. T., and Arnold, F. H. (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase, *Nat. Biotechnol. 20*, 1135-1139.

27. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.

28. Jovanovic, T., Farid, R., Friesner, R. A., and McDermott, A. E. (2005) Thermal equilibrium of high- and low-spin forms of cytochrome P450BM-3: Repositioning of the substrate?, *J. Am. Chem. Soc. 127*, 13548-13552.

29. Modi, S., Sutcliffe, M. J., Primrose, W. U., Lian, L. Y., and Roberts, G. C. K. (1996) The catalytic mechanism of cytochrome P450 BM3 involves a 6 angstrom movement of the bound substrate on reduction, *Nat. Struct. Biol. 3*, 414-417.

30. Patrick, W. M., Firth, A. E., and Blackburn, J. M. (2003) User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries, *Protein Eng. 16*, 451-457.

31. Otey, C., and Joern, J. M. (2003) In *Methods in Molecular Biology* (Arnold, F. H., and Georgiou, G., Eds.), Humana Press, Totowa, 141-148.

32. Farinas, E. T., Schwaneberg, U., Glieder, A., and Arnold, F. H. (2001) Directed evolution of a cytochrome P450 monooxygenase for alkane oxidation, *Advanced Synthesis & Catalysis 343*, 601-606.

33. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, B. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) The catalytic pathway of cytochrome P450cam at atomic resolution, *Science 287*, 1615-1622.

34. Fasan, R., Meharenna, Y. T., Snow, C. D., Poulos, T. L., and Arnold, F. H. (2008) Evolutionary History of a Specialized P450 Propane Monooxygenase, *Journal of Molecular Biology 383*, 1069-1080.

35. Reidhaarolson, J. F., and Sauer, R. T. (1988) Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences, *Science 241*, 53-57.

36. Voigt, C. A., Mayo, S. L., Arnold, F. H., and Wang, Z. G. (2001) Computational method to reduce the search space for directed protein evolution, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3778-3783.

37. Arnold, G. E., and Ornstein, R. L. (1997) Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: Cytochrome P450BM-3, *Biophys. J. 73*, 1147-1159.

38. Drummond, D. A., Iverson, B. L., Georgiou, G., and Arnold, F. H. (2005) Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins, *Journal of Molecular Biology 350*, 806-816.

39. Hegde, A., Chen, B., Haines, D.C., Bondlela, M., Mullin, D., Graham, S.E., Tomchick, D.R., Machius, M., Peterson, J.A. Active Site Mutations of P450BM-3 that Dramatically Affect Substrate Binding and Product Formation, (*To be published).*

40. Ost, T. W. B., Clark, J., Mowat, C. G., Miles, C. S., Walkinshaw, M. D., Reid, G. A., Chapman, S. K., and Daff, S. (2003) Oxygen activation and electron transfer in flavocytochrome P450BM3, *J. Am. Chem. Soc. 125*, 15010-15020.

41. Arnold, F. H. (1998) Design by directed evolution, *Accounts Chem. Res. 31*, 125-131.

Chapter 4

In Vivo Evolution of Butane Oxidation by AlkB and CYP153A6 Terminal Alkane Hydroxylases

Material from this chapter appears in: Koch, D. J., Chen, M. M., van Beilen, J. B., and Arnold, F. H. (2009) *In Vivo* Evolution of Butane Oxidation by Terminal Alkane Hydroxylases AlkB and CYP153A6, *Applied and Environmental Microbiology* 75, 337 – 344, and is reprinted by permission of the American Society of Microbiology.

A. Abstract

Enzymes of the AlkB and CYP153 families catalyze the first step in the metabolism of medium chain-length alkanes, selective oxidation of alkanes to the 1-alkanols, and enable their host organisms to utilize alkanes as carbon sources. Small gaseous alkanes, however, are converted to alkanols by members of the evolutionarily unrelated methane monooxygenase (MMO) family. Propane and butane can be oxidized by CYP enzymes engineered in the laboratory, but these produce predominantly the 2-alkanols. Here we report the *in vivo* directed evolution of two medium chain-length terminal alkane hydroxylases, the integral-membrane di-iron enzyme AlkB from *Pseudomonas putida* GPo1 and the class I soluble CYP153A6 from *Mycobacterium* sp. HXN-1500, for enhanced activity on small alkanes. We established a *P. putida* evolution system that enables selection for terminal alkane hydroxylase activity and used it to select propane- and butane-oxidizing enzymes based on enhanced growth complementation of an adapted *P. putida* GPo12 (pGEc47ΔB) strain. The resulting enzymes exhibited higher rates of 1-butanol production from butane and maintained their preference for terminal hydroxylation. This *in vivo* evolution system could be generally useful for directed evolution of enzymes that hydroxylate small alkanes.

B. Introduction

Microbial utilization and degradation of alkanes was discovered almost a century ago (*1*). Since then, several enzyme families capable of hydroxylating alkanes to alkanols, the first step in alkane degradation, have been identified and categorized based on their preferred substrates (*2*). The soluble and particulate methane monooxygenases (sMMO and pMMO) and the related propane monooxygenase and butane monooxygenase (BMO) are specialized on gaseous small-chain alkanes (C_1 to C_4), while medium-chain (C_5 to C_{16}) alkane hydroxylation seems to be the domain of the CYP153 and AlkB enzyme families.

Conversion of C_1 to C_4 alkanes to alkanols is of particular interest for producing liquid fuels or chemical precursors from natural gas. The MMO-like enzymes that catalyze this reaction in nature, however, exhibit limited stability or poor heterologous expression (*2*) and have not been suitable for use in a recombinant host that can be engineered to optimize substrate or cofactor delivery. Alkane monooxygenases often co-metabolize a wider range of alkanes than those which support growth (*3*). We wished to determine whether it is possible to engineer a medium-chain alkane monooxygenase to hydroxylate small alkanes, thereby circumventing difficulties associated with engineering MMO-like enzymes as well as investigating the fundamental question of whether enzymes unrelated to MMO can support growth on small alkanes.

The most intensively studied medium-chain alkane hydroxylases are the AlkB enzymes (*4 – 6*), especially AlkB from *Pseudomonas putida* GPo1 (*7 – 10*). While most members of the AlkB family act on C_{10} or longer alkanes, some accept alkanes as small as C_5 (*2*). A recent study (*3*) indicated that AlkB from *P. putida* GPo1 may also be involved in propane and butane

assimilation. AlkB selectively oxidizes at the terminal carbon to produce 1-alkanols. No systematic protein engineering studies have been conducted on this diiron integral membrane enzyme, although selection and site-directed mutagenesis efforts identified one amino acid residue that sterically determines long-chain alkane degradation (*9*).

The most recent addition to the biological alkane-hydroxylating repertoire is the CYP153 family of heme-containing cytochrome P450 monooxygenases. Although their activity was detected as early as 1981 (*11*), the first CYP153 was characterized only in 2001 (*12*). Additional CYP153 enzymes were identified and studied more recently (*13 – 15*). These soluble, class I-type three-component P450 enzymes and the AlkB enzymes are the main actors in medium-chain-length alkane hydroxylation by the cultivated bacteria analyzed to date (*15*). CYP153 monooxygenases have been the subject of biochemical studies (*12 – 13, 16*), and their substrate range has been explored (*14, 17*). Known substrates include C_5-C_{11} alkanes. The best characterized member, CYP153A6, hydroxylates its preferred substrate octane predominantly (> 95%) at the terminal position (*13*).

Recent studies have shown that high activities on small alkanes can be obtained by engineering other bacterial P450 enzymes such as P450cam (CYP101, camphor hydroxylase) and P450 BM3 (CYP102A1, a fatty acid hydroxylase) (*18 – 19*). The resulting enzymes, however, hydroxylate propane and higher alkanes predominantly at the more energetically favorable subterminal position; highly selective terminal hydroxylation is difficult to achieve by engineering a subterminal hydroxylase (*20*). We wished to determine whether a small-alkane terminal hydroxylase could be obtained by directed evolution of a longer-chain alkane hydroxylase that exhibits this desirable regioselectivity. For this study, we chose to engineer

AlkB from *P. putida* GPo1 and CYP153A6 from *Mycobacterium* sp. HXN-1500 (*13, 21*) to enhance their activity on butane. Because terminal alkane hydroxylation is the first step of alkane metabolism, we reasoned that it should be possible to establish an *in vivo* evolution system that uses growth on small alkanes to select for enzyme variants exhibiting the desired activities.

The recombinant host *Pseudomonas putida* GPo12 (pGEc47ΔB) was engineered specifically for complementation studies with terminal alkane hydroxylases and was used previously to characterize members of the AlkB and CYP153 families (*15, 22*). This strain is a derivative of the natural isolate *P. putida* GPo1 lacking its endogenous OCT-plasmid (octane assimilation) (*23*), but containing cosmid pGEc47ΔB, which carries all genes comprising the *alk* machinery necessary for alkane utilization, with the exception of a deleted *alkB* gene (*24*). We show that this host can be complemented by a plasmid-encoded library of alkane hydroxylases and that growth of the mixed culture on butane leads to enrichment of novel butane-oxidizing terminal hydroxylases.

C. Results

C.1. P. putida GPo12 (pGEc47ΔB) growth on short-chain 1-alkanols

P. putida GPo12 (pGEc47ΔB) was shown in previous studies to grow on medium chain-length alkanols like 1-octanol and on the corresponding alkane only when complemented by a terminal alkane hydroxylase (*15, 22*). To determine whether this strain could be used to select or screen for improved terminal alkane hydroxylation activity, we tested its ability to grow with the primary and secondary C_1-C_8 alkanols as sole carbon sources (Figure 4.1). No growth was observed on any of the secondary alcohols or on methanol during the 18 -day period. Ethanol, 1-propanol and 1-butanol supported relatively strong growth, comparable to that of the positive control grown on glucose. Slow growth was observed on 1-pentanol, 1-hexanol, and 1-octanol.

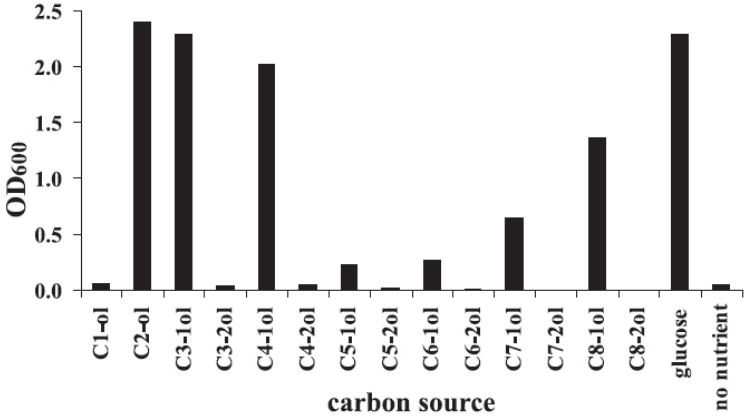


Figure 4.1: Growth of *P. putida* GPo12(pGEc47ΔB) with primary and secondary linear short and medium chain-length alkanols. The OD_{600} of the cultures was measured after 18 days of growth in liquid M9 minimal medium with 0.5% (vol/vol) primary (C_X-1ol; X = number of carbon atoms) or secondary (C_X-2ol) alcohols as carbon source dissolved in a 5% (vol/vol) organic layer of heptamethylnonane. Alcohols smaller than five carbon atoms were added without organic solvent. Cultures with 0.5% (wt/vol) glucose or no added nutrient served as controls.

The results indicated that the terminal hydroxylation products of all the short-chain *n*-alkanes except methane are readily utilized as carbon sources, while subterminal oxidation products (the *sec*-alkanols) are not. Thus, this strain should be suitable for growth-based

screening and selection for terminal hydroxylation of long, medium, and short chain-length alkanes.

C.2. Creation of gene libraries through random plasmid mutation

It was not efficient to use error-prone PCR to randomly mutate the target genes, as cloning of PCR products into the pCom vector yielded fewer than 2,000 transformants, much less than the hundreds of millions of mutants that can be used for evolution with a growth selection. Mutant libraries were therefore constructed by complementing *P. putida* GPo12 (pGEc47ΔB) strains with randomly mutated plasmids encoding AlkB or CYP153A6. The drawbacks of including mutations that affect the vector (antibiotic resistance and origin of replication) rather than just the inserted genes were compensated by the large library size and ease of library construction. Mutator strains were used to generate the plasmid libraries, and to increase diversity both available mutator strains were used, *Escherichia coli* XL1Red (Stratagene) and *E. coli* JS200 pEP Pol I (*25*). *E. coli* XL1Red has deficiencies in the DNA repair mechanism that lead to a 5,000-fold increase in the general mutation rate (*26*) (Stratagene manual). *E. coli* JS200 pEP Pol I expresses an engineered mutator DNA polymerase I, which mainly amplifies plasmid DNA with lower reliability, thus introducing mutations in the plasmid DNA (*25*). The nucleotide mutation level in XL1Red after two weeks of continuous culturing was approximately 0.1/kb, while four rounds of mutation in JS200 pEP Pol I yielded up to 0.4/kb. Cultures of both mutator strains were combined, and the mutated plasmids were transformed into *P. putida* GPo12 (pGEc47ΔB) through triparental mating with the helper strain *E. coli* CC118(pRK600) (*15*). The growth selection was performed by culturing the resulting

strain library in minimal medium with an alkane as sole carbon source for up to three weeks, as described in Chapter 8.E.6.

C.3. P. putida growth on butane through complementation with AlkB and CYP153A6 variants

A mixed culture containing alkane hydroxylase variants will become enriched in strains best adapted to use alkanes as their sole carbon source. However, not only can beneficial mutations in the hydroxylase gene lead to improved growth, but adaptations of the host and vector will do so as well. For *P. putida* GPo12 (pGEc47ΔB) complemented by CYP153 genes, it had been observed that the host had to be adapted through prolonged cultivation on alkanes to obtain significant growth on these substrates (*15*) without any mutations occurring in the *cyp153* genes themselves. To test whether host adaptation was also occurring in our experiments, 21 single colonies obtained from the first round of enrichment cultures were compared to the parent strain by plate growth tests (data not shown). Solid media growth tests were chosen over liquid media due to the growth instability of liquid minimal media cultures of *P. putida* GPo12 (pGEc47ΔB), which often showed different growth rates between replicates or, occasionally, failure to grow at all. The host, vector, and operons of the best mutants were analyzed individually by comparing them in growth tests to their wild-type counterparts. To identify and analyze potentially adapted hosts, the adapted recombinant strain was cured of the plasmid and transformed with the appropriate wild-type plasmid. Adapted vectors were isolated from the strains and the *alkB* gene or *cyp153A6 fdrA6 fdxA6* operon was replaced by the wild-type sequence by cloning, before being mated into the wild-type host. Furthermore, potentially improved hydroxylase genes were recloned into a wild-type vector and transferred into a wild-type host.

Comparison of all the resulting strains in growth assays led to the identification of several improved hosts and vectors. Strain Pcyp1, an adapted *P. putida* GPo12 (pGEc47ΔB) strain, showed faster growth on pentane than its parent when complemented by the wild-type pCom8_cyp153A6 plasmid. Further adaptation of Pcyp1 led to Pcyp2, which again grew faster on pentane. Similarly, strain Palk1 showed improved growth on propane and butane when transformed with wild-type pCom10_alkB. For the CYP153A6 system, adapted plasmid pCom8* enabled faster growth of *P. putida* GPo12 (pGEc47ΔB) on pentane, even when it contained the wild-type operon. Sequencing showed no mutation of the CYP operon in pCom8*_cyp153A6 or in the sequence 500 nucleotides up- and downstream from the operon. For the AlkB system, no improved vectors were obtained. The adapted host and vector components were specific for the particular system used, i.e., strain Palk1 did not show improved growth on short chain-length alkanes compared to the wild-type host when complemented with CYP153A6. Likewise, Pcyp1, Pcyp2, and pCom8* were only adapted for their specific systems. The nature of the mutations and how they benefit growth on short-chain alkanes is unknown.

In addition to creating these adapted hosts and plasmids, the first rounds of *in vivo* directed evolution generated enzyme mutants AlkB-BMO1 (butane monooxygenase) and CYP153A6-BMO1, both of which conferred improved growth on butane. Sequencing revealed a single nucleotide mutation in each: mutation of the codon CTA to GTA led to single amino acid substitutions L132V in AlkB-BMO1 and, through a GCA to GTA change, the substitution A94V in CYP153A6-BMO1 (A97V in the published sequence reported in reference 13). All the adapted components were combined and evaluated in plate growth tests (Table 4.1, Figure 4.2).

Table 4.1: Relative growth of adapted *P. putida* GPo12 (pGEc47ΔB) strains, expressing CYP153A6 and AlkB variants, on minimal media plates with alkanes as the sole carbon source

Complementing alkane monooxygenase	Days required for growth to a full lawn with selected carbon source				
	Ethane	Propane	Butane	Pentane	Octane
AlkB wild-type	NG[a]	5	5	3	2
AlkB-BMO1	NG	5	3	6	7
AlkB-BMO2	NG	5	2	4	8
CYP153A6 wild-type	NG	NG	NG	2	2
CYP153A6-BMO1	NG	NG	5	1.5	5

[a]NG: No growth detected during 3 -week observation

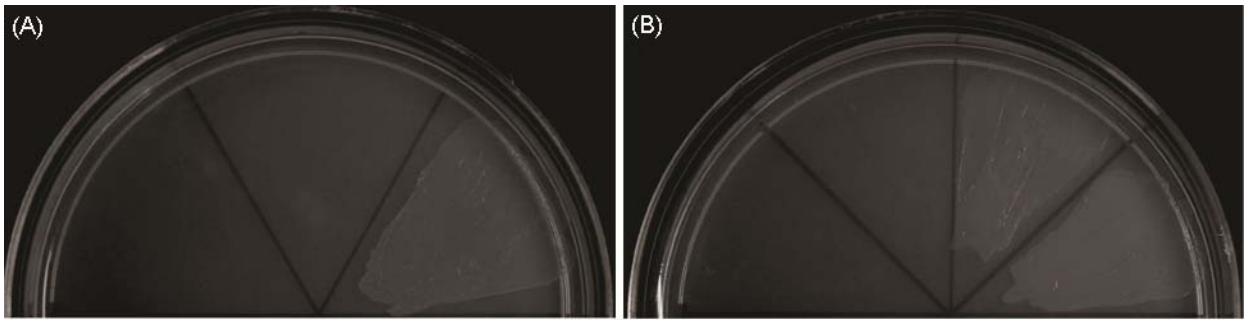


Figure 4.2: Growth of *P. putida* GPo12 (pGEc47ΔB) strains on alkanes. (A) Strains of adapted *P. putida* GPo12 (pGEc47ΔB) complemented by an empty plasmid, or expressing CYP153A6 wild type or CYP153A6-BMO1 (left, middle, and right section) were grown for 5 days with butane as sole carbon source. (B) Growth of adapted *P. putida* GPo12 (pGEc47ΔB) strains complemented by an empty plasmid, expressing the AlkB wild type, or its mutants BMO1 and BMO2, on butane after 2 days

Palk1 expressing AlkB-BMO1 showed a significant increase in rate of growth on butane compared to Palk1 expressing wild-type AlkB. In contrast, growth rates on pentane and octane were reduced. No significant growth improvement on propane was observed, and neither enzyme supported growth on ethane. These results suggest that the L132V mutation in AlkB-BMO1 specifically improves activity toward butane. A similar result was found for Pcyp2 (pCom8*_cyp153A6-BMO1), which grew more slowly on octane than Pcyp2 (pCom8*_cyp153A6) but faster on pentane. The CYP153A6-BMO1 variant also supported growth on butane, which wild-type CYP153A6 did not. Thus the A94V mutation appears to improve activity on the smaller alkanes.

Plasmids pCom8*_cyp153A6-BMO1 and pCom10_alkB-BMO1 were subjected to a second round of mutagenesis and mated into Pcyp2 and Palk1, respectively. A further-improved AlkB monooxygenase, AlkB-BMO2, was obtained after enrichment and screening. Sequencing revealed a total of three nucleotide mutations, all resulting in amino acid substitutions in AlkB-BMO2: V129M (GTG to ATG), L132V (CTA to GTA), and I233V (ATC to GTC). To ensure comparison in identical genetic backgrounds, the mutated gene was recloned into a wild-type pCom10 vector, mated into fresh Palk1, and compared in growth tests to Palk1 expressing AlkB wild-type and AlkB-BMO1 (Table 4.1, Figure 4.2). Compared to its parent AlkB-BMO1, AlkB-BMO2 performed even better in growth complementation studies with butane. Growth on pentane and octane was also improved, but was still inferior to that obtained with the wild-type enzyme. Thus, the mutations V129M and I233V improved the overall activity of AlkB-BMO2 compared to its parent, AlkB-BMO1. Enrichment and screening yielded no additional improvement in the CYP153A6 system.

C.4. Whole-cell butane bioconversions of AlkB-BMO1, -BMO2, and CYP153A6-BMO1

In order to quantify the effects of the mutations on enzyme performance, whole-cell bioconversions were performed using growth-arrested *E. coli* BL21(DE3) cells containing the CYP153A6 variants expressed from pCom8*. Although pCom plasmids are not efficient expression platforms, we nonetheless observed functional CYP153A6 expression in *E. coli*, as indicated by the CO difference spectral peak at 450 nm (Figure 4.3). In contrast, no 450 nm signal was observed for cells harboring the empty pCom8* plasmid. Cytochrome P450s are notoriously difficult to express, and it has been reported that the CO binding activity of CYP153A6 is lost shortly after cell disruption at room temperature, even when the enzyme is

isolated from its native host (*21*). However, we found CYP153A6 to express well in *E. coli* DH5α and showed a stable CO difference spectrum for hours at room temperature, if protease inhibitor was added before cell disruption. Using *E. coli* BL21(DE3) cells, which are deficient in the Lon and OmpT proteases, eliminated the need to add protease inhibitor for stable CYP153A6 expression. Cell extract from these cultures, expressing CYP153A6 or CYP153A6-BMO1, retained full CO-binding capacity for 24 h when stored at 25 °C. At 45 °C, CO-binding capacity decreased with time, showing a half-life of 638 ± 68 min for cell extract containing CYP153A6-BMO1 and 367 ± 58 min with CYP153A6.

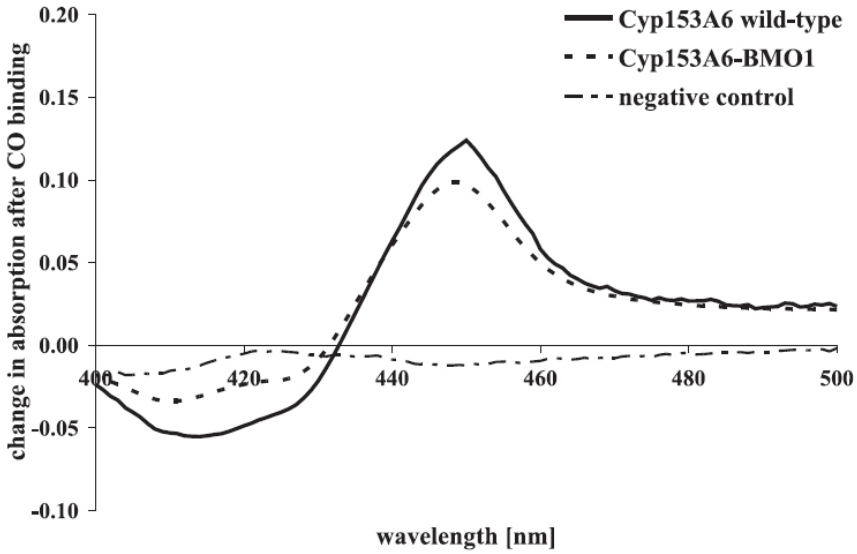


Figure 4.3: CO difference spectra of lysed *E. coli* BL21(DE3) cell suspensions. *E. coli* BL21(DE3) cultures expressing the CYP153A6 variant were concentrated 5:1 in phosphate buffer, and UV-VIS spectra were obtained from the cell-free extracts after CO saturation. The peaks correspond to 0.21 mM and 0.17 mM folded P450 for the CYP153A6 and CYP153A6-BMO1 samples, respectively. Cells carrying the empty vector treated in the same way served as the negative control.

The typical concentration of folded CYP153A6 in the cell suspensions used for bioconversions was 0.1 – 0.2 μM, with CYP153A6-BMO1 usually expressing ~ 20% less than its parent, CYP153A6. For both enzymes, we observed no significant decrease in apparent P450

concentration after the 60 -min bioconversion reactions. Bioconversion studies showed significantly altered activity and selectivity for CYP153A6-BMO1 that closely followed its growth complementation performance (Figure 4.4 (A)). Control bioconversions performed under the same conditions but using transformants of empty pCom8 vector did not produce alkanols. Butane bioconversions yielded an average total of 393 μM 1-butanol in the aqueous phase after 60 minutes with CYP153A6-BMO1, versus 277 μM with wild-type CYP153A6. From the concentration of product formed divided by the concentration of folded P450, the average turnover rate of CYP153A6-BMO1 in 1-butanol production was 49 min^{-1}, a 75% increase compared to 28 min^{-1} for CYP153A6. Interestingly, the selectivity for terminal hydroxylation also increased with the mutant, from 78% to 89% of total alkanol product. The A94V mutation also improved activity and selectivity for conversion of pentane to 1-pentanol, but had the opposite effect with propane and octane.

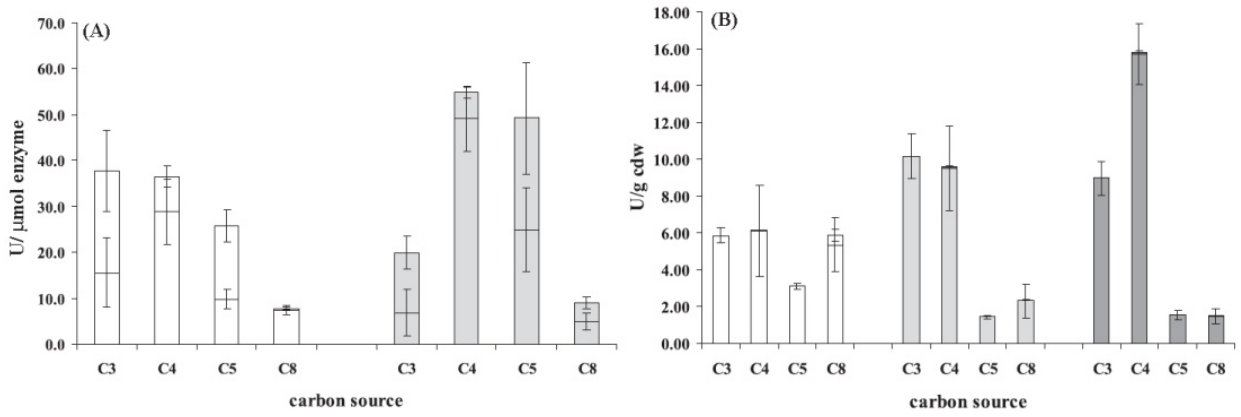


Figure 4.4: Whole-cell bioconversions were carried out with at least two replicates using resting *E. coli* BL21(DE3) cells expressing CYP153A6 (A) and AlkB (B) variants. Relative enzymatic activities in the aqueous cell suspension were calculated from the alcohol product formed per minute and the enzyme concentration (A) or total cell dry weight (B). The lower graphs represent relative activities for 1-alkanol, and the upper part shows 2-alkanol production (if detected). Solely added carbon sources were propane (C3), butane (C4), pentane (C5), and octane (C8). White graphs depict wild-type activity, light gray shows the BMO1 activity, and dark gray shows the BMO2 variant activities

Since *E. coli* cells expressing only AlkB showed no product formation, bioconversions were performed using cells transformed with pCom8_alkBFG. Minak-Bernero et al. (*27*) demonstrated that the AlkBFG system including the monooxygenase AlkB and the non-essential and essential rubredoxins AlkF and AlkG is functional in *E. coli*, without the need for the rubredoxin reductase AlkT. Results of bioconversions using AlkB, AlkB-BMO1, and AlkB-BMO2 (Figure 4.4 (B)) showed that the activity of the mutants was greater than that of wild type on butane, the substrate used for *in vivo* evolution, as well as on propane. As with CYP153A6, performance on pentane and octane decreased. The evolved mutants also showed increased activity on propane, the only change in activity that was not reflected in a similar observable change in growth complementation performance. This could reflect an activity that is too low or sub-optimal growth conditions during enrichment, such as insufficient substrate concentration or toxicity of the substrate.

In 30 minutes, butane bioconversions utilizing AlkB, AlkB-BMO1, or AlkB-BMO2 produced on average 630 μM, 1030 μM, and 1580 μM 1-butanol in the aqueous phase, respectively. The dry cell weights of the cell suspension used for the bioconversions ranged from 3.0 to 4.0 g/L. The activities (in μmol 1-butanol min^{-1} g [cell dry weight] $^{-1}$) thus increased from 6.1 to 9.5 and 15.7 units, respectively. Bioconversions with AlkB variants were highly regioselective, producing no detectable 2-alkanol from propane and pentane and very little 2-alkanol from butane and octane.

D. Discussion

D.1. Advantages and drawbacks of in vivo *evolution of terminal alkane hydroxylases*

An *in vivo* directed evolution system with selection for terminal alkane hydroxylase activity has been developed and applied to engineering enzymes from the AlkB and CYP153 families. The goal of this work was to increase hydroxylase activity on short-chain alkanes while maintaining these enzymes' remarkable preference for the thermodynamically disfavored terminal position, thereby taking a first step toward engineering small-alkane hydroxylases that can be expressed in a recombinant bacterial host amenable to further engineering.

Evolution *in vivo* using a growth selection enables searches through libraries as many as 10^8 mutants in a simple flask culture while exerting selection pressure for many useful enzyme characteristics simultaneously (e.g., substrate specificity, regioselectivity, specific activity, coupling of cofactor utilization to product formation, and expression level). In this particular case, the *Pseudomonas* host used for the directed evolution only utilizes primary alcohols, allowing selection directly for terminal hydroxylation. This is the only method we know of for efficient, high throughput selection or screening of this important activity. Previous efforts to engineer P450 BM3 or P450cam for alkane hydroxylation resulted in enzyme variants mainly targeting energetically favored subterminal positions (*18 – 19*). The present system promises to be useful for the directed evolution of a range of alkane hydroxylases, since *P. putida* GPo12 (pGEc47ΔB) can be complemented for growth on alkanes by many different AlkB and CYP153 genes (*15, 22*). Furthermore, growth tests showed that the host can utilize a range of 1-alkanols down to ethanol, potentially enabling enrichment for activity on alkanes as small as ethane.

This powerful growth selection system could handle a mutation rate perhaps ten times greater than the 0.1 to 0.4 nucleotide changes per kb obtained using the mutator strains. The *P. putida* GPo12 organism and the components used in this study, however, are somewhat difficult to handle with regards to growth stability, expression, mutation rates, cloning, and transformation efficiency. Furthermore, *in vivo* evolution also has limits for identifying moderately improved mutants, since enrichment to pure culture may require weeks or months of continuous culturing, depending on the actual doubling time on a given substrate and the level of improvement over the parent. This makes the stepwise accumulation of small improvements more challenging.

D.2. Selection for growth on propane and butane improves hydroxylase function

Our results prove that propane and butane are substrates of AlkB from *P. putida* GPo1, verifying earlier assumptions (*3*). No growth complementation was observed on ethane, and the *E. coli* system utilized in the bioconversion reactions could not be used to determine whether ethane is converted to ethanol by AlkB variants, because the glycerol added to allow NADH regeneration was fermented to ethanol by the *E. coli* BL21(DE3) host. Quantification of AlkB activity on ethane would require purification and reconstitution of AlkB and its electron transport system or genetic modification of the strain used for whole-cell bioconversions.

Second-generation variant AlkB-BMO2 supported significantly better growth and butane bioconversion, more than double that of wild-type AlkB. Although specific activity (mol product/mol enzyme/minute) could not be deteremined for the AlkB variant, absolute 1-butanol production rates in bioconversion reactions were significantly higher (up to 1.6 mM in 30 min) than that of the evolved CYP153A6-BMO1 system (0.4 mM in 60 min). The higher product

formation with the AlkB system was surprising, in view of the fact that it relied on a non-native *E. coli* host reductase, while the CYP153A6 system contained the complete native electron transfer chain composed by its ferredoxin and ferredoxin reductase. In earlier studies, AlkB activity in *E. coli* cells increased from 3 U to 20 U, when the specific reductase, encoded by *alkT*, was present (*6*). However, the significantly higher *E. coli* whole-cell bioconversion rates with AlkB variants did not translate into better growth complementation of adapted *P. putida* GPo12 (pGEc47ΔB) strains. It is known that AlkB expresses two-to tenfold better in *E. coli* than *P. putida.* On the other hand, AlkB has a five- to sixfold-higher specific activity in *P. putida* (*8*). Thus, a possible explanation for the growth complementation is that CYP153 variants express better or have a higher specific activity in *P. putida* compared to AlkB variants. Futhermore, the twice-adapted host Pcyp2 might contain chromosomal improvements that are able to compensate for the lower apparent butane oxidation activity of CYP153A6-BMO1 and thus achieve the same growth rate on butane as Palk1 expressing AlkB.

Although no crystal structures are available for any enzyme of the AlkB family, a topology model for AlkB (*9, 28*) has been published. Mapping of the evolved AlkB variant mutations onto that model (Figure 4.5 (A)) shows that two of the three mutations generated in AlkB-BMO2, V129M and L132V, are close to the histidine-containing sequence motif A ($H_{138}EXXHK_{143}$), which is one of four highly conserved histidine-containing motifs in AlkB believed to coordinate the iron ions and presumed to be part of or close to the active site (*9, 29*). Thus, a direct effect of V129M and L132V on the active site of AlkB seems possible. The third amino acid substitution, I233V, is located close to the periplasm in the sixth predicted transmembrane domain, distant from the active site.

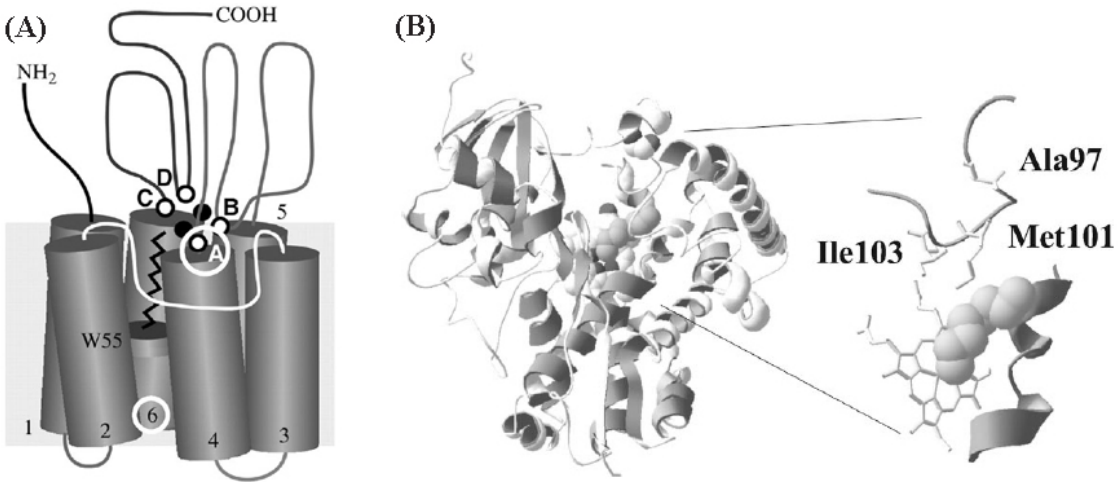


Figure 4.5: Models for (A) AlkB and (B) CYP153A6. (A) Topology model of AlkB, taken from van Beilen et al. (34). Variants AlkB-BMO1 and AlkB-BMO2 with increased activity on butane carry L132V and V129M+L132V+I233V mutations (marked by white circles), respectively. Amino acids 129 and 132 are close to the first of four histidine-containing sequence motifs (A – D) that contribute to coordinating the Fe ions. Position 233 is located in the periplasmic end of the sixth transmembrane domain (numbers 1 – 6). (B) CYP153A6 structural model with bound substrate octane (spheric object), modified from Funhoff et al. (8). CYP153A6-BMO1 carries a mutation corresponding to A97V in the model.

Evolved CYP153A6-BMO1 is the first CYP153 shown to hydroxylate butane and enable its host to grow on this short chain-length alkane. The gaseous alkanes are usually processed by distinct, specialized enzymes, like the di-iron methane and butane monooxygenases (*2*) that are unrelated to the cytochrome P450s. Complementation of *P. putida* GPo12 (pGEc47ΔB) with various CYP153 wild-type enzymes resulted in growth on alkanes ranging from pentane to decane (*15*), but never on butane. And, until now, biotransformations using CYP153 enzymes have only demonstrated activity for hexane and longer alkanes (*13 – 14, 17*). Here, we demonstrate that CYP153A6 wild type and CYP153A6–BMO1 act on butane and even propane, but only the BMO1 variant was able to complement growth of Pcyp2 on butane. The activity for 1-butanol formation of 49 min^{-1} (0.8s^{-1}) measured in CYP153A6-BMO1 assay is still less than that reported for the methane monooxygenase from *Methylosinus trichosporium* OB3b (8.8 s^{-1}) (*30*), although these values were generated under very different reaction conditions. Not only

does CYP153-BMO1 exhibit increased total activity on butane compared to its wild-type parent, but its preference for terminal hydroxylation also increased, resulting in only 11% 2-butanol formation compared to 22% for the wild-type enzyme. CYP153A6-BMO1 was converted into a genuine butane monooxygenase by *in vivo*-directed evolution.

CYP153A6-BMO1 contains a single amino acid substitution, which corresponds to A97V in the published structure model (*13*). This mutation, which has a slightly negative effect on expression level (the CYP153A6-BMO1 variant usually showed ~ 80% of the CO-binding activity of the wild type) stabilized CYP153A6 and nearly doubled its half-life at 45 °C. However, this effect is unlikely to explain the higher observed bioconversion rates, since both enzymes were stable at 25 °C, and no loss of folded P450 was observed at the end of the bioconversions. These findings argue for a direct positive effect of the A97V substitution on the butane hydroxylation activity. Based on a proposed model of CYP153 (*13*), however, A97 is predicted to be located in a loop distant from the active site, pointing toward the protein surface and not the substrate channel (Figure 4.5 (B)). Why the activity of CYP153A6 and CYP153A6-BMO1 on propane, as observed in the *E. coli* bioconversion experiments, is not accompanied by growth complementation of Pcyp2 on propane remains unknown. A minimum activity required for growth may not have yet been reached. It is also possible that the substrate is not available at a sufficient concentration or is toxic.

The appearance of improved variants of AlkB and CYP153A6 demonstrated that *in vivo* evolution in adapted *P. putida* GPo12 (pGEc47ΔB) can be applied to very different monooxygenase enzymes. Despite the higher absolute activity and better regioselectivity of AlkB variants in the bioconversion experiments, the CYP153A enzymes nonetheless offer some

important advantages for further engineering. CYP153 enzymes are soluble, and accurate determination of functional expression and concentration is easier due to observable CO binding. Furthermore, CYP153 proteins have been shown to function in whole-cell bioconversions as single-component, self-sufficient fusion proteins with the P450RhF reductase domain (*17*, *31*). Further directed evolution of terminal alkane hydroxylases like AlkB or CYP153A6 should help us to better understand and utilize their remarkable catalytic activities.

E. References

1. Sohngen, N. L. (1913) Benzin, Petroleum, Paraffinol, und Pasaffin als Kohlenstoff- und Energiequelle fur Mikroben. Zentralbl. Bacterol. Parasitenk, 595-609.
2. van Beilen, J. B., and Funhoff, E. G. (2007) Alkane hydroxylases involved in microbial alkane degradation, *Applied Microbiology and Biotechnology 74*, 13-21.
3. Johnson, E. L., and Hyman, M. R. (2006) Propane and n-butane oxidation by *Pseudomonas putida* GPo1, *Applied and Environmental Microbiology 72*, 950-952.
4. Baptist, J. N., Gholson, R. K., and Coon. M. J. (1963) Hydrocarbon oxidation by a bacterial enzyme system. I. Products of octane oxidation, *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology 73*, 1-6.
5. Nieder, M., and Shapiro, J. (1975) Physiological function of *Pseudomonas putida* ppg6 (*Pseudomonas oleovorans*) alkane hydroxylase - monoterminal oxidation of alkanes and fatty-acids, *Journal of Bacteriology 122*, 93-98.
6. van Beilen, J. B. (1994) Alkane oxidation by *Pseudomonas oleovorans*: genes and protein., University of Groningen, Groningen, The Netherlands.
7. Kok, M., Oldenhuis, R., Vanderlinden, M. P. G., Raatjes, P., Kingma, J., Vanlelyveld, P. H., and Witholt, B. (1989) The *Pseudomonas oleovorans* alkane hydroxylase gene - sequence and expression, *Journal of Biological Chemistry 264*, 5435-5441.
8. Staijen, I. E., van Beilen, J. B., and Witholt, B. (2000) Expression, stability and performance of the three-component alkane mono-oxygenase of *Pseudomonas oleovorans* in *Escherichia coli*, *Eur. J. Biochem. 267*, 1957-1965.
9. van Beilen, J. B., Smits, T. H. M., Roos, F. F., Brunner, T., Balada, S. B., Rothlisberger, M., and Witholt, B. (2005) Identification of an amino acid position that determines the substrate range of integral membrane alkane hydroxylases, *Journal of Bacteriology 187*, 85-91.
10. Vanbeilen, J. B., Kingma, J., and Witholt, B. (1994) Substrate-specificity of the alkane hydroxylase system of *Pseudomonas-oleovorans* GPo1, *Enzyme and Microbial Technology 16*, 904-911.
11. Asperger, O., Naumann, A., and Kleber, H. P. (1981) Occurrence of cytochrome P450 in acinetobacter strains after growth on normal-hexadecane, *FEMS Microbiol. Lett. 11*, 309-312.
12. Maier, T., Forster, H. H., Asperger, O., and Hahn, U. (2001) Molecular characterization of the 56-kDa CYP153 from *Acinetobacter* sp EB104, *Biochem. Biophys. Res. Commun. 286*, 652-658.
13. Funhoff, E. G., Bauer, U., Garcia-Rubio, I., Witholt, B., and van Beilen, J. B. (2006) CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation, *Journal of Bacteriology 188*, 5220-5227.

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

15. van Beilen, J. B., Funhoff, E. G., van Loon, A., Just, A., Kaysser, L., Bouza, M., Holtackers, R., Rothlisberger, M., Li, Z., and 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*, 59-65.

16. Muller, R., Asperger, O., and Kleber, H. P. (1989) Purification of cytochrome P450 from n-hexadecane-grown *Acinetobacter calcoaceticus*, *Biomedica Biochimica Acta 48*, 243-254.

17. Kubota, M., Nodate, M., Yasumoto-Hirose, M., Uchiyama, T., Kagami, O., Shizuri, Y., and Misawa, N. (2005) Isolation and functional analysis of cytochrome p450 CYP153A genes from various environments, *Biosci. Biotechnol. Biochem. 69*, 2421-2430.

18. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.

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

20. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.

21. van Beilen, J. B., Holtackers, R., Luscher, D., Bauer, U., Witholt, B., and 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*, 1737-1744.

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

23. Chakraba.Am, Chou, G., and Gunsalus, I. C. (1973) Genetic regulation of octane dissimilation plasmid in pseudomonas, *Proceedings of the National Academy of Sciences of the United States of America 70*, 1137-1140.

24. van Beilen, J. B., Penninga, D., and Witholt, B. (1992) Topology of the membrane-bound alkane hydroxylase of *pseudomonas-oleovorans*, *Journal of Biological Chemistry 267*, 9194-9201.

25. Camps, M., Naukkarinen, J., Johnson, B. P., and Loeb, L. A. (2003) Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I, *Proceedings of the National Academy of Sciences of the United States of America 100*, 9727-9732.

26. Glickman, B. W., and Radman, M. (1980) *Escherichia coli.* mutator mutants deficient in methylation-instructed DNA mismatch correction, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 77*, 1063-1067.

27. Minak-Bernero, V., Bare, R. E., Haith, C. E., and Grossman, M. J. (2004) Detection of alkanes, alcohols, and aldehydes using bioluminescence, *Biotechnology and Bioengineering 87*, 170-177.

28. van Beilen, J. B., Penninga, D., and Witholt, B. (1992) Topology of the membrane-bound alkane hydroxylase of *Pseudomonas Oleovorans*, *Journal of Biological Chemistry 267*, 9194-9201.

29. Shanklin, J., and Whittle, E. (2003) Evidence linking the *Pseudomonas oleovorans* alkane omega-hydroxylase, an integral membrane diiron enzyme, and the fatty acid desaturase family, *FEBS Lett. 545*, 188-192.

30. Duetz, W. A., van Beilen, J. B., and Witholt, B. (2001) Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis, *Current Opinion in Biotechnology 12*, 419-425.

31. Nodate, M., Kubota, M., and 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*, 455-462.

Chapter 5

Directed Evolution of P450 BM3 for Ethane Hydroxylation

A. Abstract

In continuing directed evolution of P450 BM3 for small alkane hydroxylation, we developed a high-throughput screen to directly assay for P450 alkane hydroxylation activity. With the use of a pressurizable 96-well reactor, the P450 alkane hydroxylation reaction was conducted in high throughput and the alcohol product was quantified spectroscopically by a coupled enzyme assay utilizing 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) as a horseradish peroxidase (E.C. 1.11.1.7) substrate to detect hydrogen peroxide generated by alcohol oxidase (E.C. 1.1.3.13) oxidation of the product alcohol. Applying this screen to 370 P450 variants generated in our laboratory, we identified variant E31 (WT-A74L-V78I-A82L-A184V-L188W-A328F-A330W) as the best candidate for further engineering. Through subsequent rounds of site-saturation and random mutagenesis, multiple variants supporting 1,700 – 4,000 ethane TONs were identified. Recombination of the identified mutations generated variant E31-D140E-L215P-T436R supporting 5,800 ethane TON. However, none of the BM3 variants were able to produce ethanol or methanol in whole-cell alkane bioconversions using growth-arrested *E. coli* BL21 (DE3) cells. In contrast, CYP153A6, a natural terminal alkane hydroxylase, was able to produce ethanol in whole-cell alkane bioconversions. The inability of BM3 variants to produce ethanol *in vivo* reflects their poor affinity for ethane and indicates they still lag behind a natural P450 alkane hydroxylase in terminal hydroxylation of small alkanes.

B. Introduction

Selective hydroxylation of small alkanes is a long-standing problem for which few practical catalysts are available (*1 – 3*). The lack of such catalysts that can convert gaseous alkanes into transportable liquid commodities has been a barrier to broader utilization of these resources (*4*). While this transformation has been achieved only by a limited set of transition-metal – based catalyst systems (*5 – 7*), a variety of alkane hydroxylases found in alkanotrophic microorganisms support this reaction at ambient conditions using oxygen as the oxidant (*8*). Unfortunately, since most of these hydroxylases function as a part of a larger enzyme complex and are membrane associated, their potential for industrial applications is limited. For these reasons, we have been engineering well-expressed, soluble, bacterial cytochrome P450 monooxygenases for small gaseous alkane hydroxylation.

Our previous protein engineering efforts have been focused on shifting the substrate specificity of both a self-sufficient P450 fatty acid hydroxylase from *Bacillus megaterium* CYP102A1 (BM3) (*9*) and a natural P450 medium-chain alkane hydroxylase from *Mycobacterium* sp. HXN-1500 CYP153A6 (A6) (*10*) to accept smaller alkane substrates. Using a variety of mutagenesis techniques and screening for activity on surrogate substrates such as dimethyl ether (DME) (*11*), BM3 variants with propane and ethane hydroxylation activity have been generated (*12 – 14*). Similarly, a variant with improved butane hydroxylation activity was obtained from A6 using random mutagenesis and a growth-based selection (*15*). While rapid improvements for the activities directly under selection pressure, DME demethylation and terminal butane hydroxylation, were observed in the isolated variants, activity improvements for other substrates varied. In the case of BM3 derived variants, the correlation of DME

demethylation activity with propane hydroxylation activity as measured by turnover number (TON) (r = 0.74) was much better than that with ethane hydroxylation activity (r = 0.52). Because of this difference in the correlation of activities, only relatively small improvements for ethane hydroxylation were obtained, while a complete shift in substrate specificity from fatty acids to propane was achieved through the application of the DME demethylation screen. Similarly, the growth-based selection used in the *in vivo* directed evolution of A6 also showed a poor predictability for substrates not under selection (*15*). For example, the isolated variant CYP153A6-BMO1 with improved activity for terminal hydroxylation of butane actually displayed diminished activity for propane hydroxylation.

The importance of the screen or selection to apply the desired selection pressure in directed evolution experiments has been well documented (*16 – 18*). Therefore, to continue protein engineering of P450s for hydroxylation of even smaller substrates ethane and methane a more suitable screen or selection is necessary. In this chapter, we describe the development of a high-throughput screen for terminal alkane hydroxylation activity using a pressurizable 96-well reactor to conduct the P450 reaction in high throughput and a coupled enzyme assay for colorimetric quantification of the alcohol product. Applying this screen to 370 BM3-derived variants generated in our laboratory, including variants from the alkane hydroxylation lineage (*11 – 14*), SCHEMA-guided chimeragenesis (*19 – 21*), and variants evolved for various drug compounds (*22 – 23*), we identified variant E31 (WT-A74L-V78I-A82L-A184V-L188W-A328F-A330W) as the best candidate for further engineering. Subsequent rounds of random and site saturation mutagenesis with screening for ethane hydroxylation generated variants with 1.5 to 3.3-fold improved activity, demonstrating the efficacy of the screen.

C. Results and Discussion

C.1. High-throughput ethane hydroxylation assay development

To apply selection pressure for ethane and methane hydroxylation in our laboratory evolution of P450s, a variety of small molecules was evaluated as surrogate substrates. High throughput screens were developed based on P450 oxidation of dichloromethane, chloromethane, and methanol (see Appendix C). In each instance, activity for these compounds was not a better predictor for ethane hydroxylation activity than DME demethylation. Ultimately, we pursued screening for ethane and methane hydroxylation directly using a pressurizable 96-well reactor system from Symyx (Santa Clara, CA) to conduct the P450 alkane reaction in high throughput. To quantify the alcohol product colorimetrically, the P450 reaction was first coupled to an alcohol oxidase (AO) reaction that converts the alcohol product to one equivalent of hydrogen peroxide and one equivalent of an aldehyde. In a second reaction, horseradish peroxidase (HRP) oxidizes 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) to its green radical cation using the hydrogen peroxide generated by the AO reaction. Thus, quantification of the alcohol product can be achieved spectroscopically with UV/Vis absorbance at 420 nm (*24*). The substrate scope of this assay is limited by the substrate affinity of AO, which accepts terminal linear alcohols with methanol as its preferred substrate (*25*).

Figure 5.1 summarizes the reactions of the screen and illustrates the sensitivity of the screen with both ethanol and methanol standards. Following the screening conditions detailed in Chapter 8.E.4, the spectroscopic (A_{420nm}) responses to methanol and ethanol are 0.011 A_{420nm}/μM and 0.005 A_{420nm} /μM and remain linear up to 200 and 600 μM of analyte, respectively. The

difference in alcohol sensitivity reflects the twofold higher substrate affinity (K_M) of AO for methanol compared to ethanol (*26*).

(a)

$$\text{Ethane} + \text{NADPH} + O_2 + H^+ \xrightarrow{\text{P450}} \text{Ethanol} + \text{NADP}^- + H_2O$$

$$\text{Ethanol} + O_2 \xrightarrow{\text{AO}} H_2O_2 + \text{CHO}$$

$$H_2O_2 + 2\text{ABTS} + 2H^+ \xrightarrow{\text{HRP}} 2\text{ABTS}^{+\cdot} + 2H_2O$$

(b)

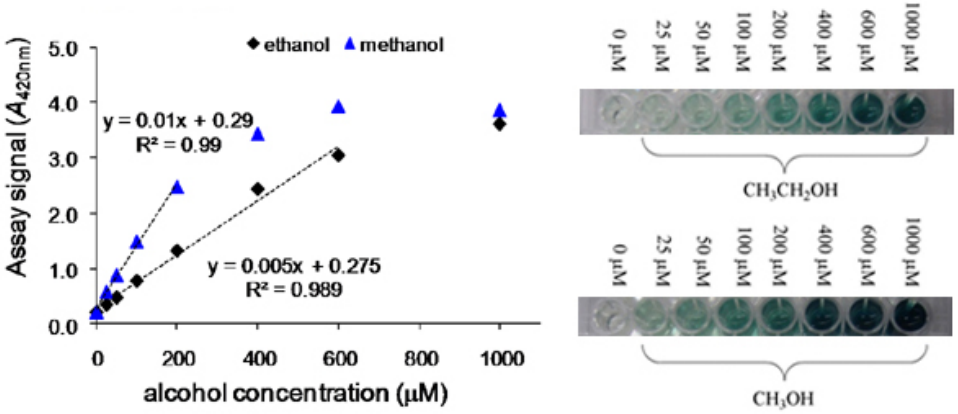


Figure 5.1: The high throughput alkane hydroxylation assay: (a) The coupled enzyme reaction scheme of the assay. (b) Assay sensitivity with methanol and ethanol standards

To validate the assay's ability to accurately quantify P450 ethane hydroxylation reactions, we selected twelve BM3 variants with various ethane activities and tested each variant for ethane hydroxylation as both cell-free extract and purified enzyme. To generate even more variety in ethanol yields, each variant was assayed at three different enzyme concentrations. The resulting ethanol product was quantified by both GC-FID as detailed in Chapter 8.H.2 and the coupled enzymatic assay. Ethanol quantifications were in good agreement between the two methods, with an R^2 of 0.96 based on linear regression of the data, see Figure 5.2. The slope of the regression, 1.20, indicates that the ABTS assay underestimates the ethanol yield in the reactions by 20%. Other than this systematic underestimation of the ethanol yield, the high

throughput colorimetric assay is as precise as GC-FID in evaluating the ethanol yield of P450 reactions.

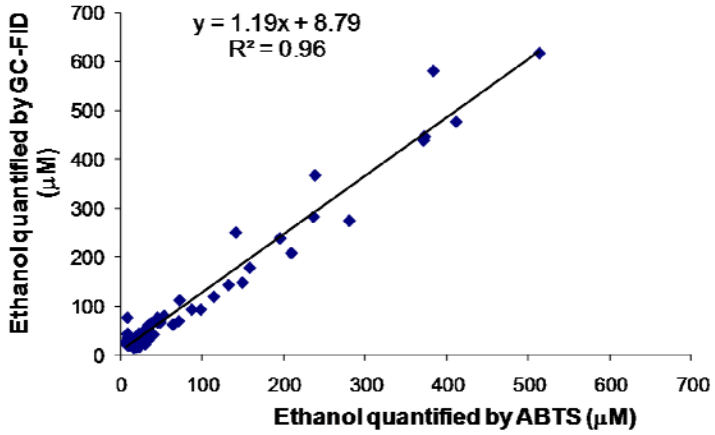


Figure 5.2: Comparison of ethanol quantification by GC-FID with enzymatic colorimetric assay

C.2. BM3 variants ethane hydroxylation evaluation

We applied this screen to evaluate BM3 variants that have accumulated in our lab from various projects for ethane and terminal propane hydroxylation activity. Previously, each variant needed to be purified and evaluated individually using GC-FID, which severely limited the number of variants that could be characterized. With the colorimetric assay, hundreds of variants can be evaluated in parallel as cell-free extracts. To this end, we selected 194 BM3 variants from the $P450_{PMO}$ alkane hydroxylation lineage (*11 – 14*), including a thermostabilized derivative AB2 ($P450_{PMO}$-C47R-I94K), thermostable chimeras of BM3 with its homologues (*20*), BM3-derived chimeras displaying a wide range of substrate activities (*19, 21*), and variants generated during evolution of BM3 for terminal alkane hydroxylation (*27*). In addition, 88 variants from each of the CRAM and C^{orbit} libraries (Chapter 3) were also selected based on their DME demethylation activities.

A total of 370 variants were tested as cell-free extracts in at least duplicate for both ethane hydroxylation and terminal propane hydroxylation. None of the selected variants supported significant 1-propanol production in the assay (> 50 μM product), while 26 variants were identified to be active for ethane hydroxylation, producing at least 45 μM ethanol (see Table 5.1). None of the 55 chimeras of BM3 were found to hydroxylate ethane, which was not surprising, since these enzymes were never selected for this activity. Of the 26 active variants, six were from the PMO lineage: PMO, 7-7, 1-3, 53-5H, 35E11, and AB2, supporting ethane TON ranging from 440 to 1,260 in cell-free extracts. Two ethane-hydroxylating variants isolated from the reduced CASTing library detailed in Chapter 3, WT-A82L-A328V and WT-A82L-A328L, were also identified by the screen.

In addition to these known ethane-hydroxylating variants, WT-V78F-A82S-A328F and WT-V78T-A82G-A328L, which were generated by grafting only the active site mutations of variants 53-5H (*14*) and 77-H9 (*14*) onto wild-type BM3, were also found to support 180 and 210 ethane TONs, respectively. Although 53-5H shows a threefold increase in ethane TON as compared to WT-V78F-A82S-A328F, both variants produce the same amount of ethanol, since the improvement in activity was offset by a decrease in protein expression. The remaining 16 variants were identified from the CRAM (13) and C^{orbit} (3) libraries as detailed in Chapter 3.

Table 5.1: Ethane hydroxylating BM3 variants identified by the high-throughput screen

Variant source	Variant	Assay signal [a] (A_{420nm})	Ethanol product [a] (μM)	Ethane TON [a]	Standard error
PMO lineage	1-3	0.64	81	1,070	150
PMO lineage	7-7	1.31	229	1,260	50
PMO lineage	AB2	0.93	154	1,110	180
PMO lineage	PMO	1.00	155	940	60
PMO lineage	53-5H	0.46	50	530	20
PMO lineage	35-E11	0.41	45	440	70
Reduced CASTing Library	WT-A82L-A328V	0.76	109	390	50
Reduced CASTing Library	WT- A82L-A328L	0.72	73	220	20
WT-779H	WT-V78T-A82G-A328L	0.58	52	210	60
WT-53-5H	WT-V78F-A82L-A328V	0.47	45	180	40
CRAM library	CB8	0.52	49	300	20
CRAM library	CD7	0.55	54	410	60
CRAM library	CG2	0.74	92	610	100
CRAM library	CE4	0.58	61	630	20
CRAM library	CA3	0.60	64	670	100
CRAM library	CA4	0.55	54	760	110
CRAM library	E78	0.84	111	930	40
CRAM library	E41	0.73	90	960	140
CRAM library	E30	0.86	114	980	90
CRAM library	E66	0.94	131	1,020	70
CRAM library	CF2	1.27	195	1,070	140
CRAM library	E32	1.00	142	1,350	20
CRAM library	E31	1.28	198	1,860	30
Corbit library	OD7	0.67	77	670	70
Corbit library	OD2	0.55	55	500	80
Corbit library	OE5	0.45	45	250	40

[a] Ethane reactions contained ca. 100 nM protein, alkane saturated potassium phosphate buffer, and an NADPH regeneration system containing 100 μM $NADP^+$, 2 U/mL isocitrate dehydrogenase, and 10 mM isocitrate (see Chapter 8.E.4 for experimental details). TON determined as nmol product/nmol enzyme.

From these 26 variants, we selected variant E31 (WT-A74L-V78I-A82L-A184V-L188W-A328F-A330W) as the parent for further directed evolution for ethane hydroxylation. It not only displayed the highest ethane TON in the cell-free extract assay, but also had a higher thermostability than wild-type BM3 with a half-denaturation temperature (T_{50}), the temperature at which the enzyme retains 50% of its activity after a 15 minute incubation, of 56.2 ^{o}C compared to 54.5 ^{o}C for wild-type BM3. In comparison, the thermostabilities of other candidates from the PMO lineage with comparable A_{420nm} such as 7-7 (T_{50} = 44.2 ^{o}C) or AB2 (T_{50} = 48.9 ^{o}C) were much lower.

To further validate the high-throughput ethane hydroxylation screen, we applied the assay across a 96-well plate containing *E. coli* cultures expressing variant E31. Using 20 μL of cell-free extract in the P450 reaction, which corresponds to ca. 120 nM of enzyme, an average TON of 1,830 was obtained with a coefficient of variance (CV) of 33% (see Figure 5.3 (a)). Increasing the enzyme loading of the reaction by using 40 μL of cell-free extract marginally increased the amount of ethanol product (from 230 μM to 310 μM) and resulted in 1,230 TON (CV of 24%). This decrease in TON with increased enzyme loading indicates that the P450 reaction with 40 μL cell-free extract is no longer limited by enzyme activity but rather by the depletion of a reactant, most likely oxygen (*11*). Therefore, the improvement in CV is merely an artifact of the reaction reaching a saturating limit rather than a systematic improvement of assay precision. Based on these results, 20 μL of cell-free extract were used in P450 reactions for the screening of E31 and subsequent mutant libraries.

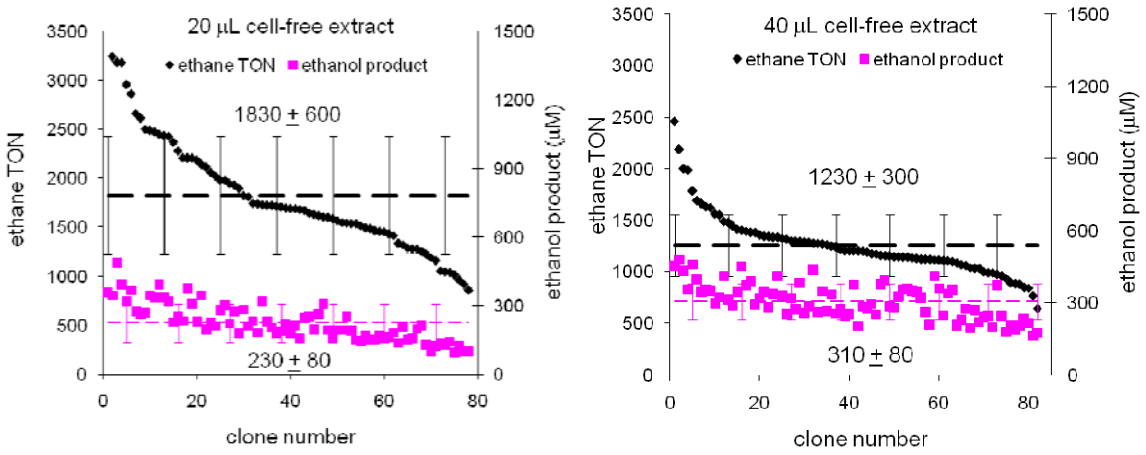


Figure 5.3: Ethane hydroxylation validation with monoclonal E31 plate: (a) 20 µL cell-free extract in 500 µL reaction (ca. 120 nM P450), (b) 40 µL cell-free extract in 500 µL reaction (ca. 240 nM P450)

C.3. Random mutagenesis of variant E31 for ethane hydroxylation

Following validation of the high-throughput ethane hydroxylation screen with the monoclonal 96-well plate of variant E31, both random mutagenesis and active site saturation mutagenesis were pursued to improve ethane hydroxylation activity. For random mutagenesis, an error-prone PCR library using the gene encoding variant E31 as template was constructed with a commercial mutagenic polymerase, Mutazyme II ®. Following the supplied protocols, a library with an average nucleotide substitution rate of 3.7/protein was obtained using 100 ng of template DNA. From screening 2,640 variants of this library for ethane hydroxylation activity, we obtained two variants, 24F8 (E31-E140D, L215P, P454S) and 22H11 (E31-D222E, A289E) with 1.6- and 1.5-fold increased ethane TON, respectively.

Surprisingly, the ethane TON of the purified enzymes was found to be significantly lower than the observed TON in the cell-free extract screening (see Table 5.2). This decrease in enzyme activity with purification contradicts our previous experience, in which propane TON is generally higher with purified enzymes compared to cell-free extracts (*11*). The most likely culprit for this difference is the reduced accuracy of P450 protein concentration determination in

cell-free extract as compared to purified enzymes. To expedite the CO-binding assay on a 96-well plate scale, the reducing agent, sodium hydrosulfite, is added to the enzyme solution before exposure to carbon monoxide (*28*). Since sodium hydrosulfite is unstable and reacts with oxygen to generate superoxide and hydrogen peroxide, which are deleterious to both the P450 and its heme prosthetic group (*29*), the inevitable protein/heme degradation may result in underestimation of the actual amount of P450 used in the screening reactions.

Table 5.2: Ethane TON of select variants as cell-free extract and purified enzyme

Variant	Ethane TON			
	Cell-free extract [a]	CV (%) [b]	Purified enzyme [a]	CV (%) [b]
E31	1,830	32.8	1,200	15.4
22H11	2,200	27.3	1,800	14.9
24F8	3,400	24.2	1,900	13.2

[a] TON determined as nmol product/nmol enzyme. Ethane reactions contained ~ 100 nM protein, alkane saturated potassium phosphate buffer, and an NADPH regeneration system containing 100 μM $NADP^+$, 2 U/mL isocitrate dehydrogenase, and 10 mM isocitrate (see Chapter 8.E.4 for experimental details).
[b] CV determined as the ratio of the standard error over the mean determined from four replicate reactions.

All five mutations found in these two variants occur at surface-exposed residues outside of the active site. Mutations L215P and D222E are located in the G helix, adjacent to residues that are part of substrate recognition site three of type II P450s (*30*). However, the side chains of these two residues are oriented away from the active site and do not appear to interact with the residues known to alter P450 substrate specificity. While the L215P mutation should disrupt the packing of this helix, rationalizing of how such an effect would lead to improved activity is difficult.

Due to the high mutation rate of the error-prone PCR library, these variants are carrying multiple mutations, which are unlikely to all be beneficial. To eliminate potential neutral and deleterious mutations, a recombination library allowing all five mutations and the corresponding wild-type amino acid at each position were constructed using splicing by overlap extension

polymerase chain reaction (SOE-PCR) (*31*). From screening 90 clones of this library with 32 unique members (94% completeness), we found two variants with improved ethane hydroxylation activity compared to variant 24F8. These two variants, RA1 (E31-D140E-L215P-D222E) and RD2 (E31-D140E-L215P), supported 2,100 and 2,200 ethane TON, respectively. Variant RD2 is variant 24F8 without the P454S mutation, which appears to be deleterious for ethane hydroxylation activity. Variant RA1 recombined the two mutations of RD2 with D222E from 22H11, this latter mutation also appears to be deleterious since the activity of RA1 is diminished compared to RD2. The thermostabilities of variants 24F8, RA1, and RD2 were determined to select a parent for the next library. Surprisingly, the T_{50}s of all three variants, 49 – 51 ^{o}C, were significantly lower than that of the parent, variant E31 (56 ^{o}C). This large decrease in thermostability was reminiscent of the introduction of the L188P mutation in the PMO evolutionary lineage (ΔT_{50} = -3 ^{o}C), which also replaced a leucine in a α-helix with a proline.

Without significant differences in thermostability between these variants, we constructed a second error-prone PCR library with the gene encoding RD2 as the template using *Taq* polymerase (*32*). From screening 2,640 members of this library with an average nucleotide substitution rate of 4.3/protein, variant 20D4 (RD2-D432G) was found with a 1.8-fold improvement in ethane hydroxylation activity. The D432G mutation occurs in the β4 sheet, close to G443A mutation identified in the PMO evolutionary lineage (*33*). This mutation would disrupt existing hydrogen bonds of D432 with Y429, E430, and E442, which could destabilize the folding of the beta-sheet. The importance of this beta-sheet structure for P450 function lies in E435, located at the bend of the sheet, which has been shown to participate in the proton

transport chain (*34*). In these two rounds of random mutagenesis, an overall 3.3-fold increase in ethane TON was achieved from variant E31 (1,200 TON) to variant 20D4 (4,000 TON).

C.4. Active site site-saturation mutagenesis of variant E31 for ethane hydroxylation

In addition to the random mutagenesis, ten active site residues of variant E31 were also targeted for site-saturation mutagenesis. Because variant E31 was isolated from the CRAM computationally designed library, which mutated ten active site residues allowing two amino acids at each position, it already contained seven active site mutations: A74L, V78I, A82L, A184V, L188W, A328F, and A330W. Instead of mutating many of these previously targeted residues again, we selected the targets for mutagenesis accounting for the sequence consensus of the ethane hydroxylation variants of the CRAM library. The ethane-hydroxylating CRAM variants displayed significant amino acid preference (> 70% representation) for F at position 87, L at position 75, W at position 188, and W at position 330. Therefore, these four positions were not mutated. Thus the ten residues targeted for saturation mutagenesis were composed of the six remaining targets from the computationally-guided libraries, A74, L75, V78, A82, A184, and A328, and residues in regions of the active site not mutated in E31, A263, I264, T436, and L437. Residues A263 and I264 are located in the I helix which constitutes the surface of the active site opposite to the B' helix that contains residues 74, 75, 78, and 82. T436 and L437 are in the β4 sheet, which contains mutations previously found to be beneficial for small alkane hydroxylation in the PMO evolutionary lineage (33).

These ten site-saturation libraries were constructed with SOE-PCR and screened to 94% completeness for ethane hydroxylation. Differences between the fraction of folded variants of these libraries and equivalent libraries constructed with wild-type BM3 revealed a decreased

mutational tolerance at these residues of E31 compared to wild-type BM3. Of the six libraries constructed with both enzymes (A74, L75, V78, A184, and A328), the fraction of folded variants decreased from an average of 0.95 with wild-type BM3 to 0.67 with variant E31. The single largest decrease occurred at position A82, for which the wild-type BM3 tolerated all possible mutations, but the equivalent E31 library contained only 36% of folded variants. Since both enzymes have similar thermostabilities, this decrease in mutational tolerance indicates that the active site of E31 is much smaller or more rigid than wild-type BM3, such that many more mutations result in steric clashes with neighboring residues.

From these site-saturation libraries, we identified four variants with mutations L74R, I78W, T436L, and T436R, supporting 1,700 – 2,600 ethane TON. Of the libraries targeting the six residues that were previously subjected to mutagenesis, only two (L74R and I78W) of the 120 total possible mutations were found to be beneficial. The lack of beneficial mutations at residues L82, V184, and F328 implies that these amino acids are the best solutions for ethane hydroxylation activity in the context of the remaining E31 mutations. Two of the four beneficial mutations introduced a positively charged arginine into the active site. The introduction of amino acids with charged side-chains into the active site also occurred in the PMO lineage and has been hypothesized to separate the active site volume into distinct pockets (*33*). The other two beneficial mutations, I78W and T436L, introduced bulkier amino acids, continuing the trend of volume-reducing mutations.

C.5. Recombination of active site mutations with mutations from error-prone PCR libraries

Using variant RD2 as the parent, the beneficial mutations for ethane hydroxylation identified from site-saturation mutagenesis and the second round of random mutagenesis were

recombined. Variant RD2 was selected as the parent, since its mutations have already been shown to be beneficial through a prior round of recombination. In addition, its mutations E140D and L215P are distant from the active site, which should reduce the probability of the mutations interacting during recombination. In contrast, the mutation identified in the second round of random mutagenesis, D432G, lies in close proximity to active site residue T436. By including this mutation in the recombination library, this residue was allowed to revert to the wild-type amino acid, which would avoid potential conflicts with mutations at T436. Starting with RD2, a recombination library allowing mutations L74R, I78W, T436L, T436R, D432G, and the corresponding wild-type amino acid at each position was constructed using SOE-PCR and screened for ethane hydroxylation to 94% completeness. From this library, we identified variant RD2-T436R (E31-D140E-L215P-T436R), supporting 5,800 ethane TON, as the most active variant.

C.6. Whole-cell alkane bioconversions

Having increased the ethane TON nearly fivefold from E31 to E31-D140E-L215P-T436R, we next tested these variants for their ability to produce methanol in *in vitro* methane hydroxylation reactions. None of variants generated from E31 were able to produce detectable amounts of methanol (> 2 μM) in these reactions. From our previous work with $P450_{PMO}$ (Chapter 2) and CYP153A6 (Chapter 4), we demonstrated that the alcohol yield of P450 whole-cell bioconversions supported with a continuous supply of alkane and oxygen can reach up to 15 mM of alcohols (*12*). Since the product yields of these whole-cell bioconversions are generally much higher than those obtained in *in vitro* P450 reactions (0.5 – 2 mM), we assayed several of

the ethane hydroxylation variants in whole-cell bioconversions for small alkane hydroxylation, including methane.

Using growth and expression conditions outlined in Chapter 8.K.1, propane bioconversions were first conducted with selected variants with PMO and CYP153A6 as controls to verify their viability in whole-cell bioconversions and to monitor the background ethanol fermentation under reaction conditions (see Figure 5.4). All six variants were able to produce propanol with activities ranging from 19 to 120 U/μmol P450, where 1 U = 1 μmol product/min. While we did not explicitly evolve variants RD2, 20D4, and RD2-T436R for improved propane hydroxylation activity, their *in vivo* propane hydroxylation activities are correlated with their *in vitro* ethane hydroxylation activities. The best ethane-hydroxylating variant RD2-T436R was nearly sixfold more productive than PMO for *in vivo* propane hydroxylation. GC-FID analysis of these propane bioconversions showed only an average of 52 $\pm$ 10 μM of ethanol was produced, which indicates minimal background ethanol fermentation.

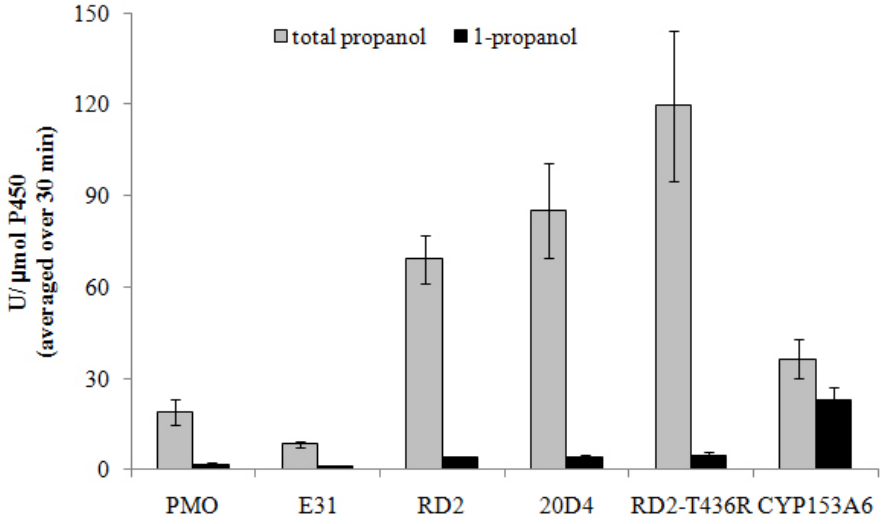


Figure 5.4: Whole-cell propane bioconversion of select P450 variants, following protocols outlined in Chapter 8.K.1

When the alkane source of the bioconversions was switched from propane to ethane with all other conditions unchanged, none of the cells containing BM3-derived variants were found to produce ethanol above background levels. The only P450 with *in vivo* ethane hydroxylation activity was CYP153A6, with a yield of 340 $\pm$ 80 μM ethanol after 30 minutes, corresponding to an activity of 8.7 U/μmol P450. None of the P450s were able to produce detectable amounts of methanol in methane bioconversions.

The lack of *in vivo* ethane hydroxylation activity of BM3 variants with *in vitro* ethane hydroxylation activity is perplexing considering that the same variants were able to produce propanol under identical conditions. This discrepancy may be due to the presence of other P450 substrates such as indole (*35 – 36*) or endogenous fatty acids that compete with the alkane during the whole-cell reactions but are absent during the *in vitro* reactions. The presence of the indole reaction is readily apparent by the visible formation of indigo over the course of the whole-cell reactions and the presence of other endogenous substrates is suggested by an increased rate of cofactor consumption when purified enzymes are assayed in the presence of cell lysate (data not shown).

D. Conclusion and Future Directions

Using a high throughput screen for P450 ethane hydroxylation, we have found variants with improved *in vitro* ethane hydroxylation through both random and site-saturation mutagenesis. However, the inability of even the most active variant, RD2-T436R, to hydroxylate ethane in whole-cell bioconversions reflects a poor affinity for ethane and highlights the gap between our current BM3 variants and a natural P450 terminal alkane hydroxylase, CYP153A6.

Although the lack of *in vivo* ethane hydroxylation is a discouraging outcome in measuring the progress of BM3 evolution, it does not eliminate the possibility that a BM3-derived variant can support *in vivo* ethane or methane hydroxylation. Iterative rounds of random and target mutagenesis that were applied in this chapter can be continued to further improve the *in vitro* ethane hydroxylation activity. However, CYP153A6, which already hydroxylates ethane as a natural, promiscuous function and supports ethane whole-cell bioconversion may be a better starting point for the engineering of a P450-based methane monooxygenase.

The drawbacks of engineering CYP153A6 are (1) it is a type I P450 with its reductase components expressed as separate enzymes, and (2) no crystal structure has been solved for any of the CYP153 family members. For these reasons, we initially pursued the *in vivo* selection-based evolution of CYP153A6 as described in Chapter 4. However, it is clear that the growth-based selection is inefficient at applying selection pressure to improve enzyme activity, as plasmid and strain adaptations were obtained with equal frequency as enzyme mutations. To engineer CYP153A6 for ethane and methane hydroxylation activity, a suitable high throughput screen is necessary. One option is to apply the ethane/methane hydroxylation screen described in this chapter and supply the reductase components through either co-expression or addition as purified enzymes. Another applicable screen is the dehalogenation of iodomethane, which releases formaldehyde that can be quantified colorimetrically by Purpald®.

The lack of a CYP153A6 crystal structure is perhaps a bigger barrier for protein engineering, since the application of the many targeted mutagenesis techniques described in Chapter 3 would not possible. While a homology model of CYP153A6 based on the CYP101 structure is available (*10*), its accuracy is questionable since the bound substrate, octane, is

largely solvent-exposed in the model. Therefore, solving the crystal structure of CYP153A6 or obtaining a suitable homology model should be pursued with a high priority. Until such structural information becomes available, random mutagenesis is the only reasonable approach.

E. References

1. Labinger, J. A., and Bercaw, J. E. (2002) Understanding and exploiting C-H bond activation, *Nature 417*, 507-514.
2. Shul'pin, G. B., Suss-Fink, G., and Shul'pina, L. S. (2001) Oxidations by the system "hydrogen peroxide-manganese(IV) complex-carboxylic acid" Part 3. Oxygenation of ethane, higher alkanes, alcohols, olefins and sulfides, *J. Mol. Catal. A-Chem. 170*, 17-34.
3. Yamanaka, I., Hasegawa, S., and Otsuka, K. (2002) Partial oxidation of light alkanes by reductive activated oxygen over the (Pd-black + VO(acac)(2)/VGCF) cathode of H-2-O-2 cell system at 298 K, *Appl. Catal. A-Gen. 226*, 305-315.
4. Arakawa, H., Aresta, M., Armor, J. N., Barteau, M. A., Beckman, E. J., Bell, A. T., Bercaw, J. E., Creutz, C., Dinjus, E., Dixon, D. A., Domen, K., DuBois, D. L., Eckert, J., Fujita, E., Gibson, D. H., Goddard, W. A., Goodman, D. W., Keller, J., Kubas, G. J., Kung, H. H., Lyons, J. E., Manzer, L. E., Marks, T. J., Morokuma, K., Nicholas, K. M., Periana, R., Que, L., Rostrup-Nielson, J., Sachtler, W. M. H., Schmidt, L. D., Sen, A., Somorjai, G. A., Stair, P. C., Stults, B. R., and Tumas, W. (2001) Catalysis research of relevance to carbon management: Progress, challenges, and opportunities, *Chem. Rev. 101*, 953-996.
5. Mizuno, N., Ishige, H., Seki, Y., Misono, M., Suh, D. J., Han, W., and Kudo, T. (1997) Low-temperature oxygenation of methane into formic acid with molecular oxygen in the presence of hydrogen catalysed by Pd0.08Cs2.5H1.34PVMo11O40, *Chemical Communications*, 1295-1296.
6. Periana, R. A., Taube, D. J., Gamble, S., Taube, H., Satoh, T., and Fujii, H. (1998) Platinum catalysts for the high-yield oxidation of methane to a methanol derivative, *Science 280*, 560-564.
7. Ushikubo, T. N., H.; Koyasu, Y.; Wajiki, S. . (1995) (Pat., U. S., Ed.).
8. van Beilen, J. B., and Funhoff, E. G. (2007) Alkane hydroxylases involved in microbial alkane degradation, *Applied Microbiology and Biotechnology 74*, 13-21.
9. Fulco, A. J. (1991) P450BM-3 and other inducible bacterial P450 cytochromes - biochemistry and regulation, *Annu. Rev. Pharmacol. Toxicol. 31*, 177-203.
10. Funhoff, E. G., Bauer, U., Garcia-Rubio, I., Witholt, B., and van Beilen, J. B. (2006) CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation, *Journal of Bacteriology 188*, 5220-5227.
11. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.
12. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.

13. Glieder, A., Farinas, E. T., and Arnold, F. H. (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase, *Nat. Biotechnol. 20*, 1135-1139.

14. Meinhold, P., Peters, M. W., Chen, M. M. Y., Takahashi, K., and Arnold, F. H. (2005) Direct conversion of ethane to ethanol by engineered cytochrome P450BM3, *Chembiochem 6*, 1765-1768.

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

16. Arnold, F. H. (1998) Design by directed evolution, *Accounts Chem. Res. 31*, 125-131.

17. You, L., and Arnold, F. H. (1996) Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide, *Protein Eng. 9*, 77-83.

18. Zhao, H. M., and Arnold, F. H. (1997) Combinatorial protein design: Strategies for screening protein libraries, *Curr. Opin. Struct. Biol. 7*, 480-485.

19. Landwehr, M., Carbone, M., Otey, C. R., Li, Y. G., and Arnold, F. H. (2007) Diversification of catalytic function in a synthetic family of chimeric cytochrome P450s, *Chem. Biol. 14*, 269-278.

20. Li, Y. G., Drummond, D. A., Sawayama, A. M., Snow, C. D., Bloom, J. D., and Arnold, F. H. (2007) A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments, *Nat. Biotechnol. 25*, 1051-1056.

21. Otey, C. R., Landwehr, M., Endelman, J. B., Hiraga, K., Bloom, J. D., and Arnold, F. H. (2006) Structure-guided recombination creates an artificial family of cytochromes P450, *PLoS. Biol. 4*, 789-798.

22. Landwehr, M., Hochrein, L., Otey, C. R., Kasrayan, A., Backvall, J. E., and Arnold, F. H. (2006) Enantioselective alpha-hydroxylation of 2-arylacetic acid derivatives and buspirone catalyzed by engineered cytochrome P450BM-3, *J. Am. Chem. Soc. 128*, 6058-6059.

23. Otey, C. R., Bandara, G., Lalonde, J., Takahashi, K., and Arnold, F. H. (2006) Preparation of human metabolites of propranolol using laboratory-evolved bacterial cytochromes P450, *Biotechnology and Bioengineering 93*, 494-499.

24. Childs, R. E., and Bardsley, W. G. (1975) Steady-state kinetics of peroxidase with 2,2'-azio-di-(3-ethylbenzthiazoline-6-sulphonic acid) as chromogen, *Biochem. J. 145*, 93-103.

25. Patel, R. N., Hou, C. T., Laskin, A. I., and Derelanko, P. (1981) Microbial oxidation of methanol - properties of crystallized alcohol oxidase from a yeast, *Pichia Sp.*, *Arch. Biochem. Biophys. 210*, 481-488.

26. Kato, N., Omori, Y., Tani, Y., and Ogata, K. (1976) Alcohol oxidases of *Kloeckera Sp.* and *Hansenula-polymorpha* - catalytic properties and subunit structures, *Eur. J. Biochem. 64*, 341-350.

27. Meinhold, P., Peters, M. W., Hartwick, A., Hernandez, A. R., and Arnold, F. H. (2006) Engineering cytochrome P450BM3 for terminal alkane hydroxylation, *Advanced Synthesis & Catalysis 348*, 763-772.

28. Otey, C., (Ed.) (2003) *High-throughput carbon monoxide binding assay for cytochrome P450*, Vol. 230, Humana Press Inc., Totowa, NJ.

29. Guengerich, F. P. (1978) Destruction of heme and hemoproteins mediated by liver microsomal reduced nicotinamide adenine-dinucleotide phosphate-cytochrome p-450 reductase, *Biochemistry 17*, 3633-3639.

30. Pylypenko, O., and Schlichting, I. (2004) Structural aspects of ligand binding to and electron transfer in bacterial and fungal P450s, *Annu. Rev. Biochem. 73*, 991-1018.

31. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagensis without phenotypic selection, *Proceedings of the National Academy of Sciences of the United States of America 82*, 488-492.

32. Cadwell, R. C., and Joyce, G. F. (1994) Mutagenic PCR *PCR-Methods Appl. 3*, S136-S140.

33. Fasan, R., Meharenna, Y. T., Snow, C. D., Poulos, T. L., and Arnold, F. H. (2008) Evolutionary history of a specialized P450 propane monooxygenase, *Journal of Molecular Biology 383*, 1069-1080.

34. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, B. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) The catalytic pathway of cytochrome P450cam at atomic resolution, *Science 287*, 1615-1622.

35. Li, Q. S., Schwaneberg, U., Fischer, P., and Schmid, R. D. (2000) Directed evolution of the fatty-acid hydyoxylase P450BM-3 into an indole-hydroxylating catalyst, *Chemistry-a European Journal 6*, 1531-1536.

36. Whitehouse, C. J. C., Bell, S. G., Tufton, H. G., Kenny, R. J. P., Ogilvie, L. C. I., and Wong, L. L. (2008) Evolved CYP102A1 (P450(BM3)) variants oxidise a range of non-natural substrates and offer new selectivity options, *Chemical Communications*, 966-968.

Chapter 6

P450 Alkane Hydroxylation Using Terminal Oxidants

A. Abstract

We investigated five P450s for the ability of their active heme-ferryl oxidant to hydroxylate the 104.9 kcal/mol C-H bond of methane within their native active site configuration, using terminal oxidants to circumvent the substrate binding activation of dioxygen in the P450 catalytic cycle. We found that the soluble cytochrome P450 from *Mycobacterium* sp. HXN-1500, CYP153A6, hydroxylates methane with 0.05 turnovers using iodosylbenzene as the oxidant. The methanol product of the iodosylbenzene reaction was validated by isotope labeling using $^{13}CH_4$ and $H_2{}^{18}O$. Attempts to demonstrate this activity under turnover conditions ($NADH/O_2$) in reactions utilizing reconstituted reductase proteins were unsuccessful. We attribute this to the low methane binding affinity of CYP153A6, which does not exhibit a spin-shift in the presence of methane. In contrast, CYP153A6 was found to support both ethane hydroxylation and iodomethane dehalogenation with product formation rates of 61 min^{-1} and 58 min^{-1}, respectively.

B. Introduction

Selective methane conversion to a liquid fuel such as methanol remains one of the great challenges in hydrocarbon chemistry (*1*). Thus far, only a few catalyst systems utilizing transition metals, such as platinum or gold dissolved in specialized environments such as concentrated acid solvents, are capable of this transformation (*2 – 3*). For a variety of reasons including catalyst cost, reaction medium toxicity, catalyst poisoning by oxidized products, and low activity, none of these catalysts are practical (*4*). Methane monooxygenases (MMOs) found in methanotrophic bacteria (*5*) appear to be ideal catalysts for this reaction, as they convert methane to methanol at rates up to 220 min^{-1} using oxygen at ambient conditions (*6*). However, despite decades of research, these enzymes have yet to be functionally expressed in heterologous hosts (*7 – 8*). For these reasons, we and others have been engineering another class of enzymes, cytochrome P450s (P450s), which shares a similar C-H bond activation mechanism as MMOs and utilizes similar high-valent iron oxygen species as the active oxidant, for small alkane hydroxylation with the ultimate goal of achieving methane oxidation (*9 – 13*). In contrast to MMOs, which are only found in methanotropic bacteria, P450s are ubiquitous across all kingdoms of life. Of the more than 11,500 known P450s (data source: http://drnelson.utmem.edu/CytochromeP450.html as published in August 2009), none has been shown to naturally hydroxylate methane.

Using both directed evolution and rational design, propane and ethane hydroxylation activities have been successfully engineered with two different P450s, CYP102A1 (BM3) (*14*) and CYP101 ($P450_{cam}$) (*13*). However, activity for methane remained elusive. Methane presents several challenges as a P450 substrate: its small molecular size presents challenges for both (1)

the initiation of the P450 catalytic cycle, which normally occurs upon the displacement of a distal water ligand as the result of substrate binding, and (2) the formation of the active radical $[(Porp)^{+\cdot}Fe^{IV}=O]^{+}$, known as Compound I (CMP I), which requires substrate-induced water expulsion from the active site. These challenges are faced for any poorly fitting non-natural substrate, albeit magnified by the small size and apolar nature of methane. The strength of the methane C-H bond, however, presents a unique challenge to the oxidizing potential of CMP I.

From density functional theory (DFT) calculations, the transition state for the H-atom abstraction presents the largest barrier in the oxygen insertion reaction of CMP I (*15*). For methane, this barrier height is 26.7 kcal/mol, which is significantly higher than the ca. 19 kcal/mol barrier for substrates such as camphor or propane's secondary C-H bond, or even the 21.6 – 21.8 kcal/mol barriers for ethane and the terminal propane C-H bond (*16 – 19*). As a comparison, this transition state barrier has been calculated to be as low as 13.8 kcal/mol for sMMO acting on methane (*20*).

The complete absence of methane oxidation activity in numerous BM3 variants evolved for propane and ethane hydroxylation activity led us to question if the P450 CMP I can overcome the 26.7 kcal/mol barrier needed to abstract the hydrogen of the methane C-H bond. In this chapter, we separated the substrate binding problem presented by the small size of methane from the challenge of the higher activation barrier presented by the methane C-H bond by assaying the reactivity of CMP I directly through terminal oxidant-supported P450 reactions. Using iodosylbenzene (PhIO), 3-chloroperoxybenzoic acid (MCPBA), and hydrogen peroxide (H_2O_2), the ability of the CMP I of five P450s, BM3, $P450_{PMO}$ (PMO) (*9*), $P450_{cam}$, CYP153A6(A6), and

CYP153A6 BMO-1(*21*), to hydroxylate alkanes ranging from methane to octane was determined.

BM3 and $P450_{cam}$ were selected because they have been engineered to hydroxylate alkanes as small as ethane (*9, 13*). PMO is a laboratory-evolved BM3 variant, which exhibits wild-type like coupling and catalytic efficiency for propane as its preferred substrate. A6, a natural P450 alkane hydroxylase which prefers medium-chain-length alkanes, and its laboratory-evolved variant CYP153A6 BMO-1 with improved butane hydroxylation activity (*21*) were chosen for their ability to hydroxylate medium-chain-length alkanes at the terminal position. In PhIO-supported reactions, CMP I is formed directly, while a ferric hydroperoxo complex (CMP 0) is formed in reactions with peroxides. The generation of CMP I from this complex requires protonation at the distal oxygen followed by heterolytic O-O bond cleavage (Figure 6.1).

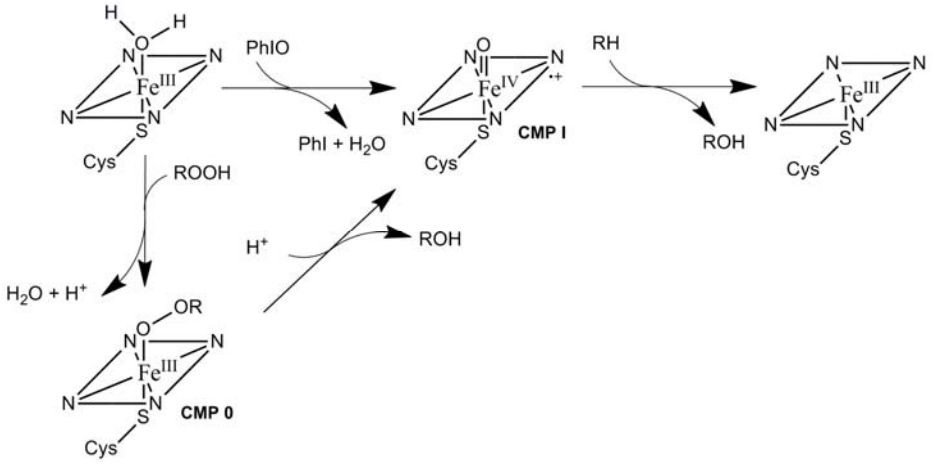


Figure 6.1: Reaction scheme for PhIO and peroxide-supported alkane hydroxylation

From these terminal oxidant-supported P450 reactions, we found A6 and A6 BMO-1, to be able to break the methane C-H bond supporting 0.05 and 0.02 turnovers, respectively, using PhIO as the oxidant. This demonstrates both the feasibility of P450 methane oxidation and the use of terminal oxidant-supported P450 reactions as an assay to investigate the compatibility of

P450 active sites for small alkane oxidation. By chemically generating the active radical, we eliminated the requirement for substrate binding to initiate P450 catalysis, which enabled us to determine the innate substrate range of each P450 active site.

Recently, P450 methane hydroxylation has also been demonstrated with wild-type BM3 through the utilization of perfluoro carboxylic acid additives (*22*). This "chemical tuning" approach apparently resolves the aforementioned challenges of methane as a P450 substrate through the generation of a catalytically active enzyme complex with reduced active site volume using an inert molecule as an external trigger to initiate catalysis. This study also demonstrates that the barrier for P450 methane oxidation is poor activation of the P450 catalytic cycle due to low methane binding affinity rather than the strength of the methane C-H bond.

C. Results and Discussion

C.1. P450s alkane hydroxylation using terminal oxidants

Given the inherent destructive nature of the terminal-oxidant supported reaction for both the enzyme and heme prosthetic group and the likely low efficiency of the oxidation reaction with small alkane substrates, we elected to use a much larger quantity of enzyme (100 μM) in these reactions than previous studies (1 – 2 μM) (*23 – 24*). Only the hydroxylase domain of each P450 was used since terminal oxidant-supported reactions do not require heme iron reduction. Each enzyme was purified through a three-step purification and lyophilized prior to use (see Chapter 8.C for details). The reactions were initiated by the addition of oxidants to pre-incubated mixtures of enzymes and substrates, following previously established protocols (*23 – 25*).

The results of the alkane hydroxylation reactions supported by terminal oxidants are summarized in Table 6.1. In general, PhIO-supported reactions resulted in higher product yields than peroxide-supported reactions with a few exceptions where yields were similar. This difference highlights uncoupling at CMP 0 as one of the difficulties in the hydroxylation of non-natural substrates. With preferred substrates, substrate binding expels water from the active site such that the protonation of CMP 0 only occurs through the proton transport chain, resulting in CMP I formation (*26*). In contrast, the binding of poorly fitting, non-natural substrates, such as small alkanes, does not fully expel water from the active site, and protonation can occur at the proximal oxygen, resulting in unproductive release of peroxide (*27*).

Table 6.1: Alkane hydroxylation by P450s utilizing terminal oxidants

	product formed, μmol (μmol P450)$^{-1}$ [a]										
Substrate	BM3			PMO	$P450_{cam}$			A6			A6 BMO-1
	PhIO	MCPBA	H_2O_2	PhIO	PhIO	MCPBA	H_2O_2	PhIO	MCPBA	H_2O_2	PhIO
Methane	-[b]	-	-	-	-	-	-	**0.05 (0.02)**	-	-	**0.02 (0.01)**
Ethane	-	-	-	-	-	-	-	2.5	0.34	0.23	2.1
Propane	1.0	0.35	0.30	0.77	0.83	0.78	0.96	3.9	1.6	0.41	3.0
Hexane	1.7	0.35	0.31	0.76	0.53	0.38	0.06	0.48	0.57	0.24	n.d.[c]
Octane	1.4	0.12	0.20	0.29	0.31	0.22	0.07 (0.02)	0.51	0.15	0.30	n.d.

[a] Alkanes (2.5 mM, or saturated at 20 psi) were incubated with P450 (100 μM) and terminal oxidant (5 mM) at 25 ^{o}C for 10 min. The data represent the averages of at least two experiments and do not correct for P450 destruction; standard errors are within 20% of the reported average with exceptions given in parentheses.
[b] Dash indicates a lack of detectable amounts of product.
[c] n.d.–not determined

Wild-type BM3 and $P450_{cam}$ were found to hydroxylate alkanes as small as propane using all three oxidants producing only sub-terminal alcohols. This indicates that alkanes as small as propane can be properly oriented near CMP I in the native active sites for oxidation to occur. In the case of BM3, under turnover conditions ($NADPH/O_2$), alkane hydroxylation has been observed only for hexane and nothing smaller. The lack of activity for the smaller alkanes under turnover conditions is solely due to poor substrate binding, which results in insufficient activation of the catalytic cycle and uncoupling at CMP 0. PMO exhibits the same substrate range as BM3 with similar TON for propane using terminal oxidants. This indicates that the laboratory evolution from BM3-to PMO-enabled propane binding to both activate the catalytic cycle and generate CMP I efficiently without changes in H-atom abstraction reaction, i.e., the reaction between CMP I of BM3 and PMO with propane, remain the same. The readiness of BM3's CMP I to react with propane may also explain the ease with which propane hydroxylation activity was obtained from BM3 through various mutations (Chapter 3).

A6 was found to hydroxylate all alkane substrates, even methane, with PhIO as the oxidant. This demonstrates that direct methane-to-methanol conversion by a P450 heme porphyrin catalyst at ambient conditions is possible and does not necessarily require the use of additional effectors to alter the active site geometry. With PhIO as the oxidant, A6 is able to hydroxylate methane with 0.05 TON. This low TON shows that although methane can be oxidized by A6, it is a poor substrate with minimal reactivity even in the presence of a pre-generated CMP I. The A6 methane TON is 50-fold lower compared to A6 ethane TON, which may reflect poor binding of methane in the active site compared to ethane, since the generation of CMP I is substrate-independent. The higher methane C-H bond strength (104.9 vs. 101.0 kcal/mol) may also be responsible for the decrease in TON. A6 methane reactions with MCPBA

and H_2O_2 did not yield detectable methanol product. Considering the low yield of the methane reaction with PhIO and the general trend of peroxide reactions being less efficient, the absence of methanol product in these reactions could be due to the limited ability to detect the product (2.0 μM, corresponding to a signal-to-noise ratio of 3). Surprisingly, reactions with the preferred substrates of A6, hexane and octane, yielded far less product as compared to reactions with ethane and propane. This may be the result of competition between these preferred substrates and PhIO for active site access. Finally, CYP153A6 BMO-1 also exhibited methane oxidation with PhIO, but with only 0.02 TON. Its propane TON also decreased compared to A6 from 3.9 to 3.0, which reflects the diminished activity observed under turnover conditions. Since CYP153A6 BMO-1 was only selected for growth complementation on butane, this loss of activity for smaller alkanes is a consequence of natural drift as these activities were not under direct selection (*28*).

In PhIO-supported reactions of A6 and PMO with hexane and octane, slightly different regioselectivities were observed compared to those obtained under turnover conditions, whereas reactions with BM3 and $P450_{cam}$ displayed similar regioselectivities to those observed under turnover conditions. For A6, 1-hexanol and 1-octanol were obtained with > 95% selectivity under turnover conditions, but in PhIO-mediated reactions, significant sub-terminal products were generated: 34% for hexane and 46% octane. In contrast, PMO produces 2-hexanol and 2-octanol with > 90% selectivity in $NADH/O_2$ supported reactions, but in PhIO-mediated reactions 24% of 1-hexanol and 30% of 1-octanol were produced. Regioselectivity differences between PhIO-supported and $NADPH/O_2$ supported reactions have been reported for several substrates (*29*) and suggest there are differences in active site packing between PhIO- and $NADH/O_2$-supported reactions. These differences could be the result of lingering iodoarene in the active site or the different order in which substrate binding and CMP I formation occurs. Under turnover

conditions, substrate binding precedes CMP I formation, which may allow the substrate to orient in the preferred conformation, whereas in PhIO-supported reactions, the CMP I species is generated in the absence of the substrate. In A6 PhIO-supported reactions with propane, 1-propanol was produced with a similar regioselectivity (35%) as the NADH/O2 supported reaction (41%). Propane hydroxylation may be less sensitive to differences in active site packing between PhIO- and NADH/O_2-supported reactions since it already occurs with low regioselectivity.

^{13}C and ^{18}O labeling experiments were conducted to verify the carbon and oxygen sources of the methanol product generated in PhIO reactions with A6. Reactions with ^{13}C methane produced an m/z 33 ion peak unique to ^{13}C-methanol, which corresponds to a +1 m/z shift of the major ^{12}C-methanol ion of m/z 32 (see Appendix D). This result confirms the carbon source of the methanol product to be the supplied methane gas. Quantification against authentic ^{13}C-methanol standards showed A6 produced 0.035 $\pm$ 0.009 TON with ^{13}C-methane. To confirm the oxygen source of the methanol product, we took advantage of the fact that CMP I generated with PhIO undergoes oxygen exchange with solvent water (*30*), such that solvent oxygen incorporation is a hallmark of the reaction mediated by PhIO. Reactions in the presence of 50% $H_2{}^{18}O$ also produced an m/z 33 ions peak corresponding to a +2 m/z shift of the m/z 31 ion of ^{12}C-methanol, unique to ^{18}O-methanol. Quantification for ^{18}O incorporation was not possible due to the low yield and the presence of a mixture of ^{16}O- and ^{18}O-methanol products. As a general comparison for the PhIO reaction with A6, ethane reactions in the presence of 50% $H_2{}^{18}O$ resulted in a 50% decrease in ^{16}O ethanol, which suggest nearly quantitative ^{18}O incorporation, which is in good accordance with literature values (*25*). These results confirm the obtained methanol product is generated through a PhIO-mediated P450 reaction with methane.

C.2. A6 alkane hydroxylation under turnover conditions

Convinced that the A6 CMP I can hydroxylate methane, we investigated A6 for oxidation of methane and other alkanes under turnover conditions utilizing reconstituted A6 reductase proteins. For the *in vitro* A6 alkane hydroxylation reactions, its reductase components, ferredoxin reductase (fdrA6) and ferredoxin (fdxA6) were purified following literature protocols established for the purification of putidaredoxin and putidaredoxin reductase of $P450_{cam}$ (*31*), which share 44% and 43% sequence identity with the ferredoxin and ferredoxin reductase of A6. The isolated proteins were quantified using known extinction coefficients for their FAD and [Fe_2-S_2] cofactors (*31*) and characterized for electron transfer activity using cytochrome c reduction (*32*).

An optimum ratio of reductase components of 1:1:10 for A6:fdrA6:fdxA6 was determined from octane hydroxylation reactions and used for subsequent experiments (see Appendix D). At this ratio of reductase components, the substrate-free cofactor consumption rate was only 3 – 5 min^{-1}. In the presence of octane, the rate of cofactor consumption reaches ca. 80 min^{-1}, which is significantly lower than that of typical bacterial P450s, which often reaches 1,000s min^{-1} (*33*), but is consistent with the rates of other natural alkane hydroxylases acting on their preferred substrates: sMMO ca. 220 min^{-1} on methane (*6*), AlkB ca. 150 min^{-1} on octane (*34*), and butane monooxygenase ca. 36 min^{-1} on butane (*35*). The rate of octanol formation was found to be 75 min^{-1}, which is slightly higher than the reported value for the same activity determined with A6 in *P. putida* GPo12 cell extracts (*36*). Reactions with ethane at 20 -psi headspace pressure resulted in ethanol formation rates of 32 min^{-1}, but reactions with both ^{12}C- and ^{13}C-methane did not produce detectable amounts of methanol.

The absence of A6 methane hydroxylation activity under turnover conditions can be rationalized by a lack of CMP I formation during the reaction, analogous to the lack of wild-type BM3 propane hydroxylation under turnover conditions. P450 catalysis is initiated by substrate binding, which displaces the distal water ligand inducing a spin-shift from the low spin resting state to a catalytically active high spin state (*37*). For A6, this spin-shift is indicated by UV/Vis difference spectra (*38*) and is observed for alkanes as small as ethane, but is absent for methane (Figure 6.2). This lack of spin-shift in the presence of methane demonstrates the absence of CMP I formation as a barrier for catalysis under turnover conditions.

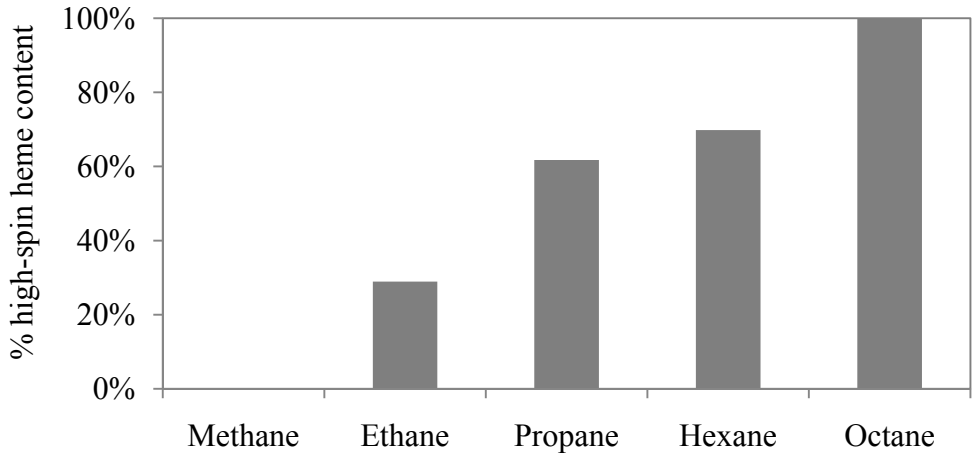


Figure 6.2: Alkane induced spin-shift of A6 as determined at saturation by absorbance difference between A_{392} and A_{418}. The percentage of high-spin content was determined relative to the spin-shift induced by the preferred substrate octane. For gaseous substrates, the spin-shift was determined with 40 -psi head-space pressure. For liquid alkanes, the spin-shift was determined with 1 mM substrate in a 1% ethanol solution.

To gain more insight into the differences in A6-catalyzed oxidation of a preferred substrate vs. a smaller, non-natural substrate, the kinetic parameters (K_M and k_{cat}) and the kinetic isotope effect (KIE) were determined for the oxidation of hexane, octane, and the dehalogenation of iodomethane (Table 6.2). Attempts to characterize ethane hydroxylation kinetics were unsuccessful, as saturating kinetics were not observed over the pressure range investigated (see Appendix D). Therefore, iodomethane was chosen as a surrogate for the small gaseous alkanes,

because it possesses both a molecular size and a C-H bond strength (102.9 kcal/mol) intermediate of those of methane (104.9 kcal/mol) and ethane (101.0 kcal/mol). The liquid form of iodomethane offers the additional benefit that saturating kinetics can be observed. From the preferred substrate octane to iodomethane, a 50-fold increase in K_M from 0.32 mM to 17.7 mM was observed. Surprisingly, there was only a small difference in the k_{cat} values for these two substrates, 75 min^{-1} for octane vs. 58 min^{-1} for iodomethane. The overall 70-fold decrease in catalytic efficiency from $3.9x10^3$ $M^{-1}s^{-1}$ for octane to 55 $M^{-1}s^{-1}$ for iodomethane is thus largely due to the higher K_M.

The kinetic isotope effect (KIE) of a P450 reaction, determined by comparing the reaction rate of deuterated and non-deuterated substrates, can indicate if the C-H bond activation is the rate-limiting step in catalysis. Since a C-D bond has lower vibrational frequency compared to a C-H bond, it has a lower zero point energy. As a result of this lower ground state energy, the deuterated substrate has a higher activation barrier for reaction, which would produce a slower reaction rate if breaking of the C-H bond is rate-limiting. For the preferred A6 substrates hexane and octane, KIEs of near unity were observed. This indicates the reaction of CMP I with these substrates is not rate-limiting under turnover conditions, which is expected, as the second electron transfer step is generally rate-limiting for P450s acting on their preferred substrates (*39*). In contrast, a KIE of 5.8 was observed for iodomethane dehalogenation, which demonstrates that the H-atom abstraction reaction has become rate-limiting. A KIE of 5.8 falls within the classical limit and indicates an absence of hydrogen atom tunneling, which has been suggested to occur during sMMO oxidation of methane (*40*). As a comparison, a similar KIE of 6.4 has been observed for $P450_{cam}$ hydroxylation of (1R)-5,5-difluorocamphor (*41*). By blocking the preferred hydroxylation site, this study showed that the C-H bond activation step can become rate-limiting

when the oxidation occurs under unfavorable geometries. The higher mobility of iodomethane within the A6 active site may have a similar effect such that the substrate C-H bond is not properly oriented near CMP I for reaction.

Table 6.2: A6 alkane hydroxylation under turnover conditions

	k_{cat} (min^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})	Coupling (%)	KIE (k_H/k_D)
Methane [a]	0	n.d.[b]	n.d.	n.d.	n.d.
Iodomethane [c]	58 (5.1)	17.7(1.4)	55	42	5.8
Ethane [a]	61 (8.3)	n.d.	n.d.	74	n.d.
Hexane [c]	98 (7.0)	0.78 (0.04)	$2.1x10^3$	96	1.0
Octane [c]	75 (7.2)	0.32 (0.02)	$3.9x10^3$	98	1.0

[a] Reactions contained 0.5 μM A6, 0.5 μM FdrA6, 5 μM FdxA6,1 mM NADH, in 0.1 M phosphate buffer, pH 8.0 under alkane atmosphere with head-space pressure ranging from 20 – 60 psi. Ethane k_{cat} and coupling were determined at 40 psi head-space pressure, corresponding to the maximum rate of ethanol formation observed. [b] n.d.–not determined. [c]Reactions contained 0.5 μM A6, 0.5 μM FdrA6, 5 μM FdxA6,1 mM NADH, the substrate in 2% ethanol and 0.1 M phosphate buffer, pH 8.0. The data represent the averages of three replicates; values in parentheses are the standard errors.

In conclusion, we have demonstrated the use of terminal oxidants for evaluating the innate substrate specificity of P450s, independent of the requirement for substrate binding to initiate catalysis. Using this assay, we were able to show that CMP I of A6 can support methane oxidation, just as the CMP I of BM3 is poised for propane oxidation, despite the fact that neither activity is observed under turnover conditions. This result confirms that the methane C-H bond of 104.9 kcal/mol can be oxidized by a P450 and suggests that A6 can be a good starting point for the engineering of a P450 methane monooxygenase.

D. References

1. Labinger, J. A., and Bercaw, J. E. (2002) Understanding and exploiting C-H bond activation, *Nature 417*, 507-514.
2. Periana, R. A., Taube, D. J., Gamble, S., Taube, H., Satoh, T., and Fujii, H. (1998) Platinum catalysts for the high-yield oxidation of methane to a methanol derivative, *Science 280*, 560-564.
3. Shilov, A. E., and Shteinman, A. A. (1977) Activation of saturated-hydrocarbons by metal-complexes in solution, *Coord. Chem. Rev. 24*, 97-143.
4. Labinger, J. A. (2004) Selective alkane oxidation: hot and cold approaches to a hot problem, *J. Mol. Catal. A-Chem. 220*, 27-35.
5. Hanson, R. S., and Hanson, T. E. (1996) Methanotrophic bacteria, *Microbiol. Rev. 60*, 439-+.
6. Fox, B. G., Froland, W. A., Jollie, D. R., and Lipscomb, J. D. (1990) Methane monooxygenase from *Methylosinus trichosporium* OB3B, *Method Enzymol. 188*, 191-202.
7. Kopp, D. A., and Lippard, S. J. (2002) Soluble methane monooxygenase: activation of dioxygen and methane, *Curr. Opin. Chem. Biol. 6*, 568-576.
8. Lieberman, R. L., and Rosenzweig, A. C. (2004) Biological methane oxidation: Regulation, biochemistry, and active site structure of particulate methane monooxygenase, *Crit. Rev. Biochem. Mol. Biol. 39*, 147-164.
9. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.
10. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.
11. Bell, S. G., Stevenson, J. A., Boyd, H. D., Campbell, S., Riddle, A. D., Orton, E. L., and Wong, L. L. (2002) Butane and propane oxidation by engineered cytochrome P450(cam), *Chemical Communications*, 490-491.
12. Stevenson, J. A., Westlake, A. C. G., Whittock, C., and Wong, L. L. (1996) The catalytic oxidation of linear and branched alkanes by cytochrome P450(cam), *J. Am. Chem. Soc. 118*, 12846-12847.
13. Xu, F., Bell, S. G., Lednik, J., Insley, A., Rao, Z. H., and Wong, L. L. (2005) The heme monooxygenase cytochrome P450(cam) can be engineered to oxidize ethane to ethanol, *Angewandte Chemie-International Edition 44*, 4029-4032.
14. Meinhold, P., Peters, M. W., Chen, M. M. Y., Takahashi, K., and Arnold, F. H. (2005) Direct conversion of ethane to ethanol by engineered cytochrome P450BM3, *Chembiochem 6*, 1765-1768.

15. Ensing, B., Buda, F., Gribnau, M. C. M., and Baerends, E. J. (2004) Methane-to-methanol oxidation by the hydrated iron(IV) oxo species in aqueous solution: A combined DFT and car-parrinello molecular dynamics study, *J. Am. Chem. Soc. 126*, 4355-4365.

16. Shaik, S., Kumar, D., de Visser, S. P., Altun, A., and Thiel, W. (2005) Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes, *Chem. Rev. 105*, 2279-2328.

17. de Visser, S. P., Kumar, D., Cohen, S., Shacham, R., and Shaik, S. (2004) A predictive pattern of computed barriers for C-H hydroxylation by compound I of cytochrome P450, *J. Am. Chem. Soc. 126*, 8362-8363.

18. de Visser, S. P., Ogliaro, F., Sharma, P. K., and Shaik, S. (2002) What factors affect the regioselectivity of oxidation by cytochrome P450? A DFT study of allylic hydroxylation and double bond epoxidation in a model reaction, *J. Am. Chem. Soc. 124*, 11809-11826.

19. Ogliaro, F., Harris, N., Cohen, S., Filatov, M., de Visser, S. P., and Shaik, S. (2000) A model "rebound" mechanism of hydroxylation by cytochrome P450: Stepwise and effectively concerted pathways, and their reactivity patterns, *J. Am. Chem. Soc. 122*, 8977-8989.

20. Siegbahn, P. E. M. (2001) O-O bond cleavage and alkane hydroxylation in methane monooxygenase, *J. Biol. Inorg. Chem. 6*, 27-45.

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

22. Zilly, F. E., Acevedo, J. P., Augustyniak, W., Deege, A., Häusig, U. W., and Reetz, M. T. (2011) Tuning a P450 Enzyme for Methane Oxidation, *Angewandte Chemie 123*, 2772-2776.

23. Anzenbacher, P., Niwa, T., Tolbert, L. M., Sirimanne, S. R., and Guengerich, F. P. (1996) Oxidation of 9-alkylanthracenes by P450 2B1, horseradish peroxidase, and iron tetraphenylporphine iodosylbenzene systems: Anaerobic and aerobic mechanisms, *Biochemistry 35*, 2512-2520.

24. Guengerich, F. P., Yun, C. H., and Macdonald, T. L. (1996) Evidence for a 1-electron oxidation mechanism in N-dealkylation of N,N-dialkylanilines by cytochrome P450 2B1 - Kinetic hydrogen isotope effects, linear free energy relationships, comparisons with horseradish peroxidase, and studies with oxygen surrogates, *Journal of Biological Chemistry 271*, 27321-27329.

25. Heimbrook, D. C., and Sligar, S. G. (1981) Multiple mechanisms of cytochrome P450-catalyzed substrate hydroxylations, *Biochem. Biophys. Res. Commun. 99*, 530-535.

26. Imai, M., Shimada, H., Watanabe, Y., Matsushimahibiya, Y., Makino, R., Koga, H., Horiuchi, T., and Ishimura, Y. (1989) Uncoupling of the cytochrome P450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine - a

possible role of the hydroxy amino-acid in oxygen activation, *Proceedings of the National Academy of Sciences of the United States of America 86*, 7823-7827.

27. Davydov, R., Perera, R., Jin, S. X., Yang, T. C., Bryson, T. A., Sono, M., Dawson, J. H., and Hoffman, B. M. (2005) Substrate modulation of the properties and reactivity of the oxy-ferrous and hydroperoxo-ferric intermediates of cytochrome P450cam as shown by cryoreduction-EPR/ENDOR spectroscopy, *J. Am. Chem. Soc. 127*, 1403-1413.

28. Arnold, F. H., Wintrode, P. L., Miyazaki, K., and Gershenson, A. (2001) How enzymes adapt: lessons from directed evolution, *Trends Biochem.Sci. 26*, 100-106.

29. Gustafsson, J. A., Rondahl, L., and Bergman, J. (1979) Iodosylbenzene derivatives as oxygen donors in cytochrome P450 catalyzed steroid hydroxylations, *Biochemistry 18*, 865-870.

30. Nam, W. W., and Valentine, J. S. (1993) Reevaluation of the significance of O-18 incorporation in metal complex-catalyzed oxygenation reactions carried out in the presence of (H2O)-O-18), *J. Am. Chem. Soc. 115*, 1772-1778.

31. Gusalus, I. C., and Wagner, G. C. (1978) *Methods Enzymol 52*, 166-188.

32. Roome, P. W., Philley, J. C., and Peterson, J. A. (1983) Purification and properties of putidaredoxin reductase, *Journal of Biological Chemistry 258*, 2593-2598.

33. Ost, T. W. B., Miles, C. S., Murdoch, J., Cheung, Y. F., Reid, G. A., Chapman, S. K., and Munro, A. W. (2000) Rational re-design of the substrate binding site of flavocytochrome P450BM3, *FEBS Lett. 486*, 173-177.

34. Shanklin, J., Achim, C., Schmidt, H., Fox, B. G., and Munck, E. (1997) Mossbauer studies of alkane omega-hydroxylase: Evidence for a diiron cluster in an integral-membrane enzyme, *Proceedings of the National Academy of Sciences of the United States of America 94*, 2981-2986.

35. Cooley, R. B., Dubbel, B. L., Sayavedra-Soto, L. A., Bottomley, P. J., and Arp, D. J. (2009) Kinetic characterization of the soluble butane monooxygenase from *Thauera butanivorans*, formerly *'Pseudomonas butanovora*', *Microbiology-(UK) 155*, 2086-2096.

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

37. Ost, T. W. B., Clark, J., Mowat, C. G., Miles, C. S., Walkinshaw, M. D., Reid, G. A., Chapman, S. K., and Daff, S. (2003) Oxygen activation and electron transfer in flavocytochrome P450BM3, *J. Am. Chem. Soc. 125*, 15010-15020.

38. de Montellano, P. R. O. (1986) *Cytochrome P450*, 1st ed., Plenum Publishing Corp., New York.

39. Gelb, M. H., Heimbrook, D. C., Malkonen, P., and Sligar, S. G. (1982) Stereochemistry and deuterium-isotope effects in camphor hydroxylation by the cytochrome P450cam monooxygenase system, *Biochemistry 21*, 370-377.

40. Baik, M. H., Newcomb, M., Friesner, R. A., and Lippard, S. J. (2003) Mechanistic studies on the hydroxylation of methane by methane monooxygenase, *Chem. Rev. 103*, 2385-2419.

41. Kadkhodayan, S., Coulter, E. D., Maryniak, D. M., Bryson, T. A., and Dawson, J. H. (1995) Uncoupling oxygen-transfer and electron-transfer in the oxygenation of camphor analogs by cytochrome P450cam - direct observation of an intermolecular isotope effect for substrate C-H activation, *Journal of Biological Chemistry 270*, 28042-28048.

Chapter 7

Panel of Cytochrome P450 BM3 Variants to Produce Drug Metabolites and Diversify Lead Compounds

Material from this chapter appears in: Sawayama, A. M., Chen, M. M. Y., Kulanthaivel, P., Kuo, M. S., Hemmerle, H., and Arnold, F. H. (2009) A Panel of Cytochrome P450 BM3 Variants to Produce Drug Metabolites and Diversify Lead Compounds, *Chemistry–A European Journal* 15, 11723-11729, and is reprinted by permission from Wiley-VCH.

A. Abstract

Herein we demonstrate that a small panel of variants derived from cytochrome P450 BM3 from *Bacillus megaterium* covers the breadth of reactivity of human P450s by producing 12 of 13 mammalian metabolites for two marketed drugs, verapamil and astemizole, and one research compound. The most active enzymes support preparation of individual metabolites for preclinical bioactivity and toxicology evaluations. Underscoring their potential utility in drug lead diversification, engineered P450 BM3 variants also produce novel metabolites by catalyzing reactions at carbon centers beyond those targeted by animal and human P450s. Production of a specific metabolite can be improved by directed evolution of the enzyme catalyst. Some variants are more active on the more hydrophobic parent drug than on its metabolites, which limits production of multiple-hydroxylated species, a preference that appears to depend on the evolutionary history of the P450 variant.

B. Introduction

Selective C-H oxidation represents one of the great challenges for which synthetic chemists find only substrate-specific solutions (*1 – 4*). Breakthroughs in selective C-H oxidation methodology could benefit drug discovery, among other fields, by enabling rapid and parallel analog construction for a specific molecular scaffold. Such improvements in efficiency would greatly increase the number and variety of compounds that could be produced and raise the likelihood of identifying effective therapeutic agents.

Many synthetic strategies for C-H oxidation rely on a reactive intermediate that plays upon subtle differences in C-H bond strength (1–5 kcal mol^{-1}) to achieve regioselectivity (*5*). Owing to the large number of C-H bonds in most bioactive chemicals, identifying a reagent that can react at one C-H bond in preference to all others can be difficult, or even impossible. Nature solves the selectivity problem by incorporating discrete molecular recognition elements into enzyme catalysts so that they can use specific enzyme–substrate interactions to impart reactivity to a specific C-H bond.

Cytochromes P450 (CYPs) are a large superfamily of heme-containing C-H oxidation enzymes. In humans, CYPs play key roles in drug metabolism and clearance. An ability to predict how potential pharmaceuticals will be metabolized will better equip us to identify derivatives with improved biological activity, solubility, toxicity, stability, or bioavailability (*6*). In fact, FDA guidelines indicate that uniquely human metabolites and metabolites present at disproportionately higher levels in humans as compared to the animal species used during standard toxicology testing may require safety assessment before beginning large-scale clinical trials (*7*).

Preparation of these metabolites at sufficient scale for evaluation is not trivial and often requires a *de novo* synthesis for each metabolite. Biosynthetic methods that employ purified human CYPs or crude liver microsomes are not much better, because human CYPs are poorly stable, membrane-bound, multi-component protein systems that generally exhibit low reaction rates. As an alternative to using human CYPs as biocatalysts for metabolite production, we and others have focused on soluble, bacterial P450 BM3 (also known as CYP102A1) (*8*) as a C-H oxidation platform (*9 – 11*). Derived from *Bacillus megaterium*, P450 BM3 has properties that greatly facilitate its engineering and use in synthesis: it can be expressed at high levels in *E. coli* (~ 12% dry cell mass), and, unlike nearly all other CYPs, its hydroxylase, reductase, and electron-transfer domains are all in one contiguous polypeptide chain. This last feature might contribute to its relatively high activity (> 1,000 turnovers per min) on its preferred fatty acid substrates (*8*). Like most CYPs implicated in anabolic pathways, P450 BM3 is substrate specific, and hydroxylates a C_{20} fatty acid over a C_{12} fatty acid with more than 200-fold higher efficiency (*12*).

We chose three structurally diverse drug compounds with known patterns of mammalian CYP-dependent clearance to evaluate whether P450 BM3 variants can catalyze similar C-H oxidations. Verapamil is a calcium channel blocker used in the treatment of hypertension and arrhythmia (*13*). Astemizole is a potent H_1-histamine receptor antagonist used for treatment of common sinus allergy symptoms (*14*). The third compound, LY294002, is an antiproliferative agent that inhibits phosphatidylinositol 3-kinase, a target with potential for treatment of malignancies (*15*). We report here that a small collection of P450 BM3 variants can produce nearly all the known human (or rat) metabolites for each of the three drugs. Within each set of active enzyme variants, we identified several that produce selected metabolites in yields and

activities suitable for preparative scale synthesis. We also identified variants that generate metabolites not produced by rat liver microsome controls or known to be human metabolites; the results demonstrate the ability of this C-H oxidation platform to target a range of carbon centers.

C. Results

C.1. Selection of P450 variants

From our extensive collections we selected 120 P450 BM3 variants that had previously demonstrated activity on substrates not in the wild-type enzyme's repertoire. These variants were constructed by using a variety of commonly implemented genetic diversification techniques, including error-prone PCR and targeted mutagenesis of active site residues. We also included variants derived from structure-guided recombination of BM3 and its homologs, CYP102A2 and CYP102A3 (*16*). These enzymes had been selected based on their activities toward a variety of substrates, including straight-chain alkanes, cycloalkanes, alkyl ethers (*17 – 19*), aromatic compounds (*20*), derivatives of fatty acids (*21*), and drug compounds (*10, 22*).

We also distinguished between enzyme variants that function as monooxygenases and use NADPH to reduce the catalytic Fe^{III} center and peroxygenases that contain only the heme domain and use H_2O_2 in place of O_2 and NADPH. Each has its own catalytic and operational advantages and disadvantages that have been described elsewhere (*21*). Although we had evidence that the presence of a reductase domain impacts activity favorably (*20*), we were unsure what effect, if any, this domain would have on regioselectivity. The variants, their sequences relative to wild type, and the criteria for their prior selection from mutant libraries are described in Appendix E.

C.2. Reactions with human and bacterial CYPs

Verapamil and astemizole undergo extensive biotransformation by the major human CYPs to produce demethylated and dealkylated products (*23*) by hydroxylation of activated carbon centers adjacent to heteroatoms (summarized in Tables 7.1 (a), 7.2 (a)). Experiments

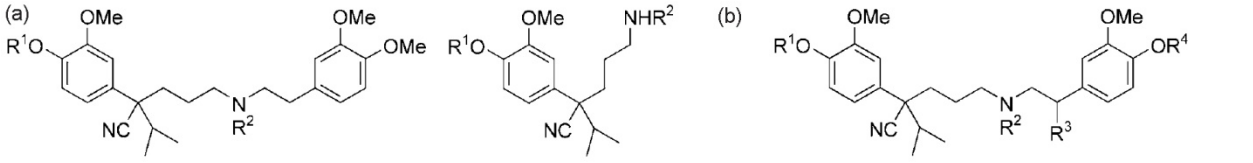
conducted with rat liver microsomes *in vitro* indicate that LY294002 metabolism is confined to single- and double- alkyl hydroxylations of the morpholine ring (Table 7.3), a pattern also observed in other drugs possessing this moiety. Reactions of the 120 P450 BM3 variants with each of the three drugs were assessed by HPLC. Metabolite products were evident in the HPLC traces for a significant number of variants (43 exhibited some activity on verapamil, 42 on astemizole, and 18 on LY294002); the metabolites and their distributions were subsequently characterized by using LCMS and MS/MS. Metabolite structure assignments were aided by comparison of MS and MS/MS spectra with spectra of known metabolites obtained in our laboratory using rat microsomal systems.

C.3. Verapamil

Table 7.1 summarizes the performance of selected enzymes on verapamil (all 43 active enzymes and their product distributions are listed in Appendix E). For example, entry 4 displays the activity of chimera 22313333 on verapamil. It converts 34% of the starting material into an assortment of products, the distribution of which is 41% **2**, 15% **5**, 20% **6**, 9% **7**, and 15% **10**. Variants 2C11 and 9C1 furnished the widest array of products, with six and eight individual metabolites, respectively. In particular, 9C1 produced several metabolites that had undergone two hydroxylation events (**1**, **4**, and **8**). Variant 22313333 (entry 4) was the only enzyme capable of producing double hydroxylation product **5**, while 32313233 (entry 6) best produced norverapamil (3). Interestingly, the addition of a reductase as a fusion to the 32313233 heme domain (entry 7) rendered the enzyme unable to produce new metabolite **7**. Variant 7-11D (entry 13) differs by no more than two mutations from the variants in each of entries 9–12, but its product distribution more closely parallels chimera 21312332. The 9-10A family of enzymes

was the most active, with conversions exceeding 30% (entries 9–12), and produced metabolite **7** with excellent selectivity.

Table 7.1: Conversion of verapamil to: (a) its most abundant[a] human metabolites, and (b) new metabolites by different P450 BM3 variants



	M_w	R^1	R^2
Verapamil	454	Me	Me
1	426	H	H
2	440	H	Me
3	440	Me	H

	M_w	R^1	R^2
4	276	H	Me
5	276	Me	H
6	290	Me	Me

	M_w	R^1	R^2	R^3	R^4	other
7	440	Me	Me	H	H	
8	456	Me	Me	OH	H	
9	470	Me	Me	OH	Me	
10	470	Me	Me	H	Me	OH

	Variant[b]	% Conversion	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	*2C11*	25		8	28			20	8		8	28
2	*9C1*[c]	31	3		29	3		6	13	3		12
3	*D6H10*	78			26	8		31				24
4	**22313333**	34		41			15	20	9			15
5	**22313231**	43		46				33	5			16
6	**32313233**	24			50			21	13			17
7	**32313233-R1**	25			40			44				16
8	**21312332**	6			17			33	17			33
9	9-10A A78F	30		10					83			
10	9-10A A82L	51						2	94			
11	9-10A F87L	49						6	94			
12	12-10C	34			6			6	79		3	
13	7-11D	21			28			29	14			29

[a] Defined as > 1% abundance following oral ^{14}C verapamil administration in humans (*23*).

[b] Variants shown in italics were selected for activity on propranolol (*22*), variants in bold are chimeras (*16*), variants in normal type were selected for activity one alkanes (*18 – 19*) chimeras were written according to fragment composition: 32313233-R1, for example, represents a protein that inherits the first fragment from parent CYP102A3, the second from CYP102A2, the third from CYP102A4 and so on. R1 connotes a fusion to the reductase domain from parent A1. Chimera fusions were used as monooxygenases; chimera heme domains were used as peroxygenases.

[c] Because not all P450 BM3 oxidation products could be identified, product distribution totals can be less than 100%.

The enzymes that best mimicked human CYP reactivity on verapamil were derived from two variant families. Three had been isolated by directed evolution for activity on propranolol (*22*), while the remaining four were chimeras (*16, 24*). The BM3 variants that best produced new

metabolites were derived from the alkane-hydroxylating 9-10A variant (*18 – 19*). These enzymes catalyzed a new regioselective demethylation reaction at position R^4 and a new benzylic oxidation at position R^3. Even without optimization of the reaction conditions, some enzymes were highly active on verapamil; D6H10 (entry 3), for example, transformed verapamil into metabolites at 78% conversion and a total turnover number (TON) greater than 1500.

The enzymes possess contrasting degrees of regioselectivity. Some were unselective and produced a spectrum of metabolites (e.g., propranolol-evolved enzymes 2C11, 9C1, D6H10). By contrast, chimera 22313231 produced metabolite **2**, chimera 32313233-R1 made dealkylated compound **6** (entries 5 and 7), and many variants produced new metabolites **7** and **10** (entries 8–13) with sufficient selectivity ($\geq$ 30%) for larger scale production without further optimization. To demonstrate that useful quantities of metabolites can be produced, we used purified 9-10A F87L to produce metabolite **7** (9.4 mg) from verapamil (25 mg) in 39% yield (1560 TON, 0.025 mol% catalyst).

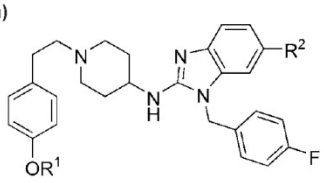
C.4. Astemizole

Activity of the enzyme panel on astemizole resulted in the seven metabolites described in Table 7.2 (all 42 active enzymes and their product distributions are listed in Appendix E). Variants DE10, 21313311, 22313333, and 32313233 (entries 1–3 and 9) produced dealkylated metabolite **14** in preference to other compounds. The identity of the residue at position 78 had a strong impact on the product distribution within the 9-10A backbone; mutations A78F, A78T, and A78S (entries 4, 5, and 12) produced metabolites **11**, **13,** and **15** as the most abundant respective products. Furthermore, 41-5B (entry 8), a 9-10A family member that contains the A78F mutation (in addition to A82G and A328V) also favors demethylated **11**. Variants 9-10A

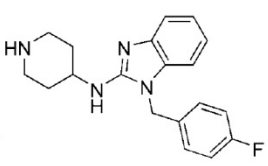
A82S, and 9-10A F87L both produced metabolite **13** in good yield and selectivity (entries 6 and 7). Two chimeric monooxygenases, 32313233-R1 and 32312333-R1 (entries 10 and 11), were effective in generating new aromatic hydroxylation product **15**.

Table 7.2: Conversion of astemizole to: (a) its most abundant[a] human metabolites, and (b) new metabolites by different P450 BM3 variants

(a)

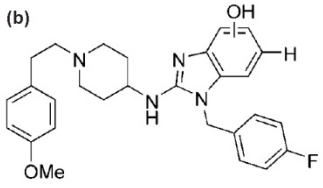


	M_w	R^1	R^2
Astemizole	458	Me	H
11	444	H	H
12	460	H	OH
13	474	Me	OH

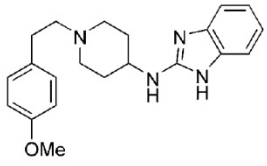


	M_w
14	324

(b)



	M_w	other
15	474	
16	490	OH



	M_w
17	350

	Variant[b]	% Conversion	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	*DE10*	9				56			
2	**21313311**	10	20			40			40
3	**22313333**	9	22			56			22
4	9-10A A78F	21	48		38			14	
5	9-10A A78T	27	37	7	44			4	
6	9-10A A82S	30	20	7	70			3	
7	9-10A F87L	36		3	88			6	
8	41-5B	5	100						
9	**32313233**	11				45	27		27
10	**32313233-R1**	16				13	69		
11	**32312333-R1**	78				4	67		9
12	9-10A A78S	15	13	7			80		
13	9-10A A82L	49	24				61	4	
14	9-10A A82I	32	31	6	63				
15	9-10A F87A	31			23	16	45		

[a] Defined as > 1% abundance following oral ^{14}C verapamil administration in humans (*25*)
[b] See Table 7.1 for further explanation of variant nomenclature.

The 9-10A-derived monooxygenases were the most adept at producing human metabolites of astemizole (entries 4–8). However, these enzymes were unable to dealkylate astemizole and produce metabolite **14**. Oxidation of this more hindered C-H bond was best accomplished by the propranolol active variants and chimeric peroxygenases (entries 1, 2, and 3, respectively). A new benzylic site not hydroxylated by human CYPs was targeted by six

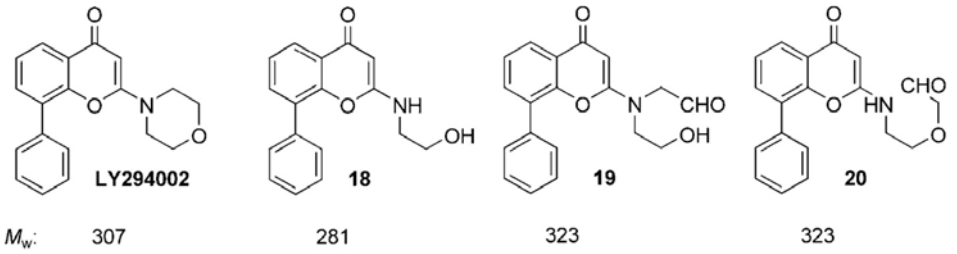
chimeric BM3 enzymes to produce **17**. A second benzimidazole site was hydroxylated (as demonstrated by metabolites **15** and **16**), which is different from that observed in human metabolites **12** and **13**. Discrimination of these sterically and electronically identical sp^2 C-H bonds is virtually impossible by using traditional transition metal catalysts and emphasizes the power of molecular recognition as a regiocontrolling element.

Small changes in the substrate channel can affect the regioselectivity of aromatic C-H oxidation. For instance, the presence and position of a single methyl group in Thr78 vs. Ser78 (entry 5 vs. 12) and Ile82 vs. Leu82 (entry 14 vs. 13) were sufficient to bias oxidation at C6 instead of at another position on the benzimidazole ring. Product **16** has undergone two oxidations. Although this metabolite was observed in reactions with ten variants, it was produced in greater than 10% abundance by only one variant (entry 4). New metabolite **15** is a candidate for scale-up and could be produced at high conversion (> 25% selectivity) by 15 variants of different lineages (Table E.4 in the Appendix E). Even without optimization of reaction conditions, several enzymes showed good activity towards astemizole. Variant 32312333- R1, for example, converted 78% astemizole into metabolites (70 TON).

C.5. LY294002

The most abundant metabolites of LY294002 produced by rat liver microsomes are detailed in Table 7.3. Both single hydroxylation products **19** and **20** were identified in BM3-variant reactions with very good regioselectivity. Aminoalcohol **18** is the sole metabolite of all three drugs not observed in the reactions. Derivative **18** requires two oxidations and might not appear simply due to the low conversion in these 96-well plate reactions.

Table 7.3: Conversion of LY294002 by different P450 BM3 variants into the most abundant[a] metabolites produced by rat liver microsomes



	Variant[b]	% Conversion	**18**	**19**	**20**
1	9-10A L75I	10		50	
2	9-10A A78S	7		71	
3	9-10A F87V	6			100
4	9-10A T260S	11		9	
5	68-8F	9		11	

[a] Define as > 1% abundance following in vitro reaction of LY294002 with rat liver microsomes.
[b] Variants in normal type were selected for activity on alkanes (*18 – 19*).

The P450s catalyzing the single hydroxylations were monooxygenases derived from 9-10A (Table 7.3). Although several variants could produce metabolite **19**, only 9-10A F87V oxidized the other position of the morpholine ring and made derivative **20** (entry 3). The remaining twelve metabolite-producing variants were also monooxygenases of both 9-10A and chimeric origin, and produced a metabolite with Mw=238; this indicates morpholine loss and addition of water to the bis-aryl backbone. In all cases, LY294002 conversion was very low (< 15%) and correlated with the low rat liver microsome conversion for LY294002 (40%) relative to microsome activity on verapamil (85%) and astemizole (85%). LY294002 better evades both mammalian and bacterial P450-catalyzed C-H oxidation than verapamil and astemizole under the same conditions.

C.6. Activities on singly-hydroxylated metabolites

Of the twelve mammalian metabolites produced by the P450 BM3 variants, eight were made with sufficiently high selectivity to enable preparative scale production without further

optimization (**2**, **3**, **6**, **11**, **13**, **14**, **19**, and **20**). Of the seven new metabolites to which structures could be assigned, four were made with sufficiently high regioselectivity to enable preparative scale production (**7**, **10**, **15**, and **17**). All eleven of the highly produced metabolites arose from single oxidations. Seven of the eight remaining metabolites are the products of two or more oxidations and are present as minor components in mixtures with the singly-hydroxylated products. Of the 69 individual reactions that underwent appreciable conversion ($\geq$ 10% drug consumed) only five generated products of multiple C-H oxidations in $\geq$ 10% abundance.

The remainder contained single-oxidation products, including mixtures of singly-oxidized species. The lack of bishydroxylated products could reflect a regiochemical effect in which the second C-H site is less reactive. In this case, one of the two possible single-hydroxylation products would be favored. Sixteen variants produced only one single-hydroxylation product; reaction mixtures from ten additional variants contained two (metabolites were present in 1:1 to 3:1 ratios in 6/11 reactions). It is also possible that the relative lack of bis-hydroxylation products reflects discrimination on the part of the enzymes in which the BM3 binds and hydroxylates the parent drug in preference to its metabolites. Because BM3-catalyzed C-H oxidation increases polarity by either unmasking heteroatoms or formally substituting hydrogen with a hydroxyl group, production of polyhydroxylated metabolites could be disfavored, for example, if the enzyme prefers a more hydrophobic substrate.

To assess enzyme activity on the metabolites produced by a single oxidation, all 120 enzymes were incubated with purified demethylated metabolites norverapamil (**3**) and desmethylastemizole (*26*) (**11**). Product distributions and extents of reaction were determined by HPLC. In every case in which an enzyme was active on both parent drug and metabolite,

regioselectivities were unchanged; this indicates that the metabolite and parent drug bind in similar orientations. Comparison of each enzyme's activity on the parent drug and the associated metabolite showed that enzymes from the alkane-evolved 9-10A lineage tended to be more active on astemizole and verapamil than on their metabolites (true of 49 out of 55 active enzymes; Table 7.4).

Table 7.4: Preference for more hydrophobic parent drug over its metabolite depends on evolutionary history of the variant.

P450 Variant family	Number of P450s more active on parent drug	Number of P450s more active on metabolite
Alkane-selected lineage	48/55	6/55
Propranolol-selected and chimera lineages	18/30	12/30

Because neither the N-methyl group of verapamil nor the O-methyl group of astemizole is the preferred site for C-H oxidation by the alkane-evolved variants, this bias could reflect a preference for the more hydrophobic substrate. In contrast, when this pairwise reaction matrix is analyzed across the propranolol-evolved and chimeric P450 BM3 lineages, there is no statistical difference from random at the 95% confidence interval ($0.43 \leq p = 0.60 \leq 0.77$).

C.7. Directed evolution can improve metabolite production

None of the 120 members of this catalyst panel had been selected for activity on any of the three drugs. Thus, any initial activity represents a promiscuous activity; such side activities are often easy to improve by directed evolution (*27 – 28*). Of the 103 enzymes that reacted with these substrates, 9-10A F87L possessed the best combination of activity and regioselectivity (Table 7.2, entry 7); its reaction with astemizole, which produces metabolite **13** with 88% selectivity. We screened 2,000 variants made by error-prone PCR of the 9-10A F87L gene using a colorimetric screen for products of aromatic hydroxylation (*29*) and identified three new

sequences with improved metabolite production (Table 7.5). The new variants improved the conversion to 51–52% while preserving high regioselectivity (~ 80% for metabolite **13**).

Table 7.5: Production of astemizole metabolites by 9-10A F87L variants

	Variant	% Conversion	% Selectivity **13**
1	9-10A F87L	34	88
2	E4D D68G	51	80
3	C205R D338G	51	80
4	E4D H92Q	52	75

D. Discussion

A small, 120-member panel of P450 BM3 variants captured nearly all of the mammalian P450 scope of reactivity by producing 12 of 13 known metabolites. In their ability to mimic human CYPs, the P450 BM3 variants demonstrated considerable versatility, activating C-H bonds of varying strength (90–105 kcal mol^{-1}) and steric encumbrance (sp^2 vs. sp^3 and 1^o vs. 2^o carbon centers). We were able to assign structures to seven new metabolites (Table 7.1 (b), 7.2 (b)), all of which have undergone oxidation at new carbon centers.

Cytochrome P450 enzymes are versatile catalysts, the biological activities of which are encoded in a wide range of primary sequences. The 26 CYPs, for which the crystal structures have been solved, possess as little as 15% sequence identity (*30*), but share a highly conserved fold. Emphasizing the versatility of the P450 fold, P450cam and P450cin catalyze the oxidation of isosteres camphor and cineole. Although the substrates are nearly identical, the two enzymes differ not only in sequence (27% identity), but also in the structure of their active sites, most notably a complete lack of the B' helix in P450cin (*31 – 32*). Whereas large changes in sequence are possible, only small perturbations are necessary to produce significant changes in function.

For example, CYP2A4 and CYP2A5 differ by only eleven amino acids, yet catalyze hydroxylations on structurally dissimilar coumarin and testosterone substrates (*33*) .

Only 57 human enzymes (*34*) are responsible for known CYP-dependent drug metabolism, and a single enzyme, CYP3A4, accounts for > 50% of the burden for xenobiotic CYP-mediated clearance (*35*). Because CYPs can be broadly or narrowly specific, we and others have speculated that it should be possible to take advantage of the high native activity of P450 BM3 and use mutation to either relax or shift its substrate specificity in order to generate useful C-H oxidation catalysts (*11, 36 – 39*). For example, BM3 variant 9-10A, which is 13 mutations away from the wild type, exhibits broad activity across short- and medium-chain alkanes—activity that is low or completely absent in its wild-type parent (*18 – 19*). Furthermore, 9-10A could be respecialized to hydroxylate propane, preferring it over alkanes that differ by a single methylene group (*40*). We also showed previously that a variant of 9-10A was able to hydroxylate drug-like compounds efficiently and selectively (*10*). Here, we wanted to determine whether variants of 9-10A and other BM3-derived enzymes could cover or even exceed the broad substrate range of mammalian CYPs. Within any catalyst panel, both extremes of regioselectivity can be useful: enzymes that already possess the desired selectivity can be used to produce individual metabolites, whereas less selective enzymes can be used to survey metabolite possibilities. Both can serve as starting sequences for directed evolution to enhance activity or tune selectivity. That a systematic and general evolutionary algorithm can be used for catalyst improvement is a particularly appealing aspect of DNA-encoded reagents. Complementary optimization studies used in traditional synthesis methods usually rely on chemical intuition to improve a catalyst and require a good understanding of the catalytic mechanism.

When we examined conversion of the parent drug versus its demethylated metabolite, we noted that one enzyme family consistently converted more of the parent drug than the demethylated and more polar metabolite. These enzymes tend to catalyze single hydroxylations, while enzymes evolved for activity on propranolol and the chimeric variants often catalyze bis-hydroxylations. Structure–activity relationships of this type should help in the future to select enzymes that are most likely to react with as-of-yet-untested substrates and could also help predict product profiles.

E. Conclusion

This panel should enable rapid identification and production of relevant quantities of the human metabolites of drug candidates for pharmacological and toxicological evaluations in preclinical species (*41 – 42*). Although we have highlighted the potential of these enzymes to accelerate preparation of metabolites for pharmacological and toxicological testing, this enzyme panel is likely also to be useful further upstream in the drug-development process as general reagents for lead diversification. Reagents that rely on molecular recognition will always be restricted in their scope of use. However, because the functionality of small molecules is not evenly distributed across all possible molecular architectures (*43*), it should be worthwhile to engineer P450-derived reagents that are active on privileged scaffolds that reside in these densely functional regions of structure space. The plurality of C-H sites targeted by this small P450 BM3 variant set—including and extending human P450 metabolism—augurs well for the development of a truly general panel of C-H oxidation catalysts.

F. References

1. Bergman, R. G. (2007) Organometallic chemistry - C-H activation, *Nature 446*, 391-393.
2. Crabtree, R. H. (2001) Alkane C-H activation and functionalization with homogeneous transition metal catalysts: a century of progress-a new millennium in prospect, *J. Chem. Soc.-Dalton Trans.*, 2437-2450.
3. Godula, K., and Sames, D. (2006) C-H bond functionalization in complex organic synthesis, *Science 312*, 67-72.
4. Labinger, J. A., and Bercaw, J. E. (2002) Understanding and exploiting C-H bond activation, *Nature 417*, 507-514.
5. Chen, M. S., and White, M. C. (2007) A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis, *Science 318*, 783-787.
6. Johnson, M. D., Zuo, H., Lee, K. H., Trebley, J. P., Rae, J. M., Weatherman, R. V., Desta, Z., Flockhart, D. A., and Skaar, T. C. (2004) Pharmacological characterization of 4-hydroxy-N-desmethyl tamoxifen, a novel active metabolite of tamoxifen, *Breast Cancer Res. Treat. 85*, 151-159.
7. FDA. *FDA Guidance of Industry Safety Testing of Drug Metabolites Homepage*: http://www.questpharm.com/FDA.html.
8. Narhi, L. O., and Fulco, A. J. (1986) Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P450 monooxygenase induced by barbiturates in *Bacillus megaterium*, *Journal of Biological Chemistry 261*, 7160-7169.
9. Budde, M., Morr, M., Schmid, R. D., and Urlacher, V. B. (2006) Selective hydroxylation of highly branched fatty acids and their derivatives by CYP102A1 from *Bacillus megaterium*, *Chembiochem 7*, 789-794.
10. Landwehr, M., Hochrein, L., Otey, C. R., Kasrayan, A., Backvall, J. E., and Arnold, F. H. (2006) Enantioselective alpha-hydroxylation of 2-arylacetic acid derivatives and buspirone catalyzed by engineered cytochrome P450BM-3, *J. Am. Chem. Soc. 128*, 6058-6059.
11. van Vugt-Lussenburg, B. M. A., Stjernschantz, E., Lastdrager, J., Oostenbrink, C., Vermeulen, N. P. E., and Commandeur, J. N. M. (2007) Identification of critical residues in novel drug metabolizing mutants of cytochrome P450BM3 using random mutagenesis, *J. Med. Chem. 50*, 455-461.
12. Ost, T. W. B., Miles, C. S., Munro, A. W., Murdoch, J., Reid, G. A., and Chapman, S. K. (2001) Phenylalanine 393 exerts thermodynamic control over the heme of flavocytochrome P450BM3, *Biochemistry 40*, 13421-13429.
13. McTavish, D., and Sorkin, E. M. (1989) Verapamil - an updated review of its pharmacodynamic and pharmacokinetic properties, and therapeutic use in hypertension, *Drugs 38*, 19-76.

14. Richards, D. M., Brogden, R. N., Heel, R. C., Speight, T. M., and Avery, G. S. (1984) Astemizole - a review of its pharmacodynamic properties and therapeutic efficacy, *Drugs 28*, 38-61.

15. Vlahos, C. J., Matter, W. F., Hui, K. Y., and Brown, R. F. (1994) A specific inhibitor of phosphatidylnositol 3-kinase, 2-(4-morpholinyl)-8-phenyl-4h-1-benzopyran-4-one (LY294002), *Journal of Biological Chemistry 269*, 5241-5248.

16. Otey, C. R., Landwehr, M., Endelman, J. B., Hiraga, K., Bloom, J. D., and Arnold, F. H. (2006) Structure-guided recombination creates an artificial family of cytochromes P450, *PLoS. Biol. 4*, 789-798.

17. Glieder, A., Farinas, E. T., and Arnold, F. H. (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase, *Nat. Biotechnol. 20*, 1135-1139.

18. Meinhold, P., Peters, M. W., Hartwick, A., Hernandez, A. R., and Arnold, F. H. (2006) Engineering cytochrome P450BM3 for terminal alkane hydroxylation, *Advanced Synthesis & Catalysis 348*, 763-772.

19. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.

20. Landwehr, M., Carbone, M., Otey, C. R., Li, Y. G., and Arnold, F. H. (2007) Diversification of catalytic function in a synthetic family of chimeric cytochrome P450s, *Chem. Biol. 14*, 269-278.

21. Cirino, P. C., and Arnold, F. H. (2003) A self-sufficient peroxide-driven hydroxylation biocatalyst, *Angewandte Chemie-International Edition 42*, 3299-3301.

22. Otey, C. R., Bandara, G., Lalonde, J., Takahashi, K., and Arnold, F. H. (2006) Preparation of human metabolites of propranolol using laboratory-evolved bacterial cytochromes P450, *Biotechnology and Bioengineering 93*, 494-499.

23. Eichelbaum, M., Ende, M., Remberg, G., Schomerus, M., and Dengler, H. J. (1979) Metabolism of dl-verapamil c-14 in man, *Drug Metab. Dispos. 7*, 145-148.

24. Li, Y. G., Drummond, D. A., Sawayama, A. M., Snow, C. D., Bloom, J. D., and Arnold, F. H. (2007) A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments, *Nat. Biotechnol. 25*, 1051-1056.

25. Meuldermans, W., Hendrickx, J., Lauwers, W., Hurkmans, R., Swysen, E., and Heykants, J. (1986) Excretion and biotransformation of astemizole in rats, guinea-pigs, dogs, and man, *Drug Dev. Res. 8*, 37-51.

26. Janssens, F., Torremans, J., Janssen, M., Stokbroekx, R. A., Luyckx, M., and Janssen, P. A. J. (1985) New antihistaminic n-heterocyclic 4-piperidinamines .2. Synthesis and antihistaminic activity of 1- (4-fluorophenyl)methyl -n-(4-piperidinyl)-1h-benzimidazol-2-amines, *J. Med. Chem. 28*, 1934-1943.

27. Joo, H., Lin, Z. L., and Arnold, F. H. (1999) Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation, *Nature 399*, 670-673.

28. Moore, J. C., and Arnold, F. H. (1996) Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents, *Nat. Biotechnol. 14*, 458-467.

29. Otey, C., and Joern, J. M. (2003) In *Methods in Molecular Biology* (Arnold, F. H., and Georgiou, G., Eds.), Humana Press, Totowa, 141-148.

30. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Clustal-W - improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673-4680.

31. Meharenna, Y. T., Li, H. Y., Hawkes, D. B., Pearson, A. G., De Voss, J., and Poulos, T. L. (2004) Crystal structure of P450cin in a complex with its substrate, 1,8-cineole, a close structural homologue to D-camphor, the substrate for P450cam, *Biochemistry 43*, 9487-9494.

32. Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., and Kraut, J. (1985) The 2.6-Å crystal-structure of *Pseudomonas putida* cytochrome P450, *Journal of Biological Chemistry 260*, 6122-6130.

33. Lindberg, R. L. P., and Negishi, M. (1989) Alteration of mouse cytochrome P450coh substrate-specificity by mutation of a single amino-acid residue, *Nature 339*, 632-634.

34. Lewis, D. F. V. (2004) 57 varieties: the human cytochromes P450, *Pharmacogenomics 5*, 305-318.

35. Guengrich, F. P. (1999) Cytochrome P-450 3A4: Regulation and role in drug metabolism, *Annu. Rev. Pharmacol. Toxicol. 39*, 1-17.

36. Appel, D., Lutz-Wahl, S., Fischer, P., Schwaneberg, U., and Schmid, R. D. (2001) A P450BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes, *Journal of Biotechnology 88*, 167-171.

37. Kubo, T., Peters, M. W., Meinhold, P., and Arnold, F. H. (2006) Enantioselective epoxidation of terminal alkenes to (R)- and (S)-epoxides by engineered cytochromes P450BM-3, *Chemistry-a European Journal 12*, 1216-1220.

38. Ost, T. W. B., Miles, C. S., Murdoch, J., Cheung, Y. F., Reid, G. A., Chapman, S. K., and Munro, A. W. (2000) Rational re-design of the substrate binding site of flavocytochrome P450BM3, *FEBS Lett. 486*, 173-177.

39. Sulistyaningdyah, W. T., Ogawa, J., Li, Q. S., Maeda, C., Yano, Y., Schmid, R. D., and Shimizu, S. (2005) Hydroxylation activity of P450BM-3 mutant F87V towards aromatic compounds and its application to the synthesis of hydroquinone derivatives from phenolic compounds, *Applied Microbiology and Biotechnology 67*, 556-562.

40. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.

41. Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E.,

Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D., and Dumas, B. (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast, *Nat. Biotechnol. 21*, 143-149.

42. van Beilen, J. B., Holtackers, R., Luscher, D., Bauer, U., Witholt, B., and 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*, 1737-1744.

43. Bemis, G. W., and Murcko, M. A. (1996) The properties of known drugs .1. Molecular frameworks, *J. Med. Chem. 39*, 2887-2893.

Chapter 8

Materials and Methods

A. Reagents

All chemicals, including liquid alkanes and product standards, were purchased from Sigma-Aldrich (St. Louis, MO). Solvents (hexanes, chloroform, methylene chloride) were purchased from EMD (Gibbstown, NJ). Ethane (99.95% or 99.99%), propane (99.5 %), dimethyl ether (99.8%), and methane (99.5%) were purchased from Special Gas Services (Warren, NJ) and Airgas (El Monte, CA). Methane (99.0%), ethane (99.99%), propane (99.975), butane (99%), and gaseous halomethanes, chloromethane (99.8%), and bromomethane (99.5%) were purchased from Sigma-Aldrich. NAD^+, DL-isocitric acid trisodium salt, and recombinant isocitrate dehydrogenase used as an NADPH regeneration system in experiments described in Chapters 2 – 4 were purchased from Sigma-Aldrich. $NADP^+$ and NADPH were also purchased from Codexis, Inc. (Redwood City, CA). Recombinant isocitrate dehydrogenase was also purchased from OYC, Inc. (Azusawa Japan). Alcohol oxidase, horse radish peroxidase, and ABTS used in ethanol and methanol detection assays in experiments described in Chapters 3 and 5, were purchased from Sigma-Aldrich. Restriction enzymes (*Bam*HI, *Sac*I, *Eco*RI, *Nde*I, *Xho*I, *Kpn*I) and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA), Taq DNA polymerase from Roche (Indianapolis, IN), Pfu turbo DNA polymerase from Stratagene (La Jolla, CA), and Phusion DNA polymerase from Finnzymes (Woburn, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from MP Biomedicals, Inc. (Aurora, OH).

B. Expression of P450s

B.1. BM3 holoenzymes (Chapters 2, 3, 5, and 7)

The P450 BM3 gene was cloned downstream of a double *tac* promoter of the expression vector pCwori using the flanking restriction sites *BamH*I and *EcoR*I. Silent mutations were used

to introduce a *Sac*I site 130 bp upstream of the end of the heme domain and a *Kpn*I site at 81 bp downstream of the end of the heme domain to facilitate cloning of heme libraries. *Escherichia coli* Dh5α cells transformed with the plasmids containing P450 BM3 and its derived variants were used for the expression on both 50 – 500 mL scales as well as in 96-well plates. For protein production, terrific broth (TB) medium containing 100 μg/ml ampicillin was inoculated with an overnight culture at 1:500 dilution for shake flask cultures and 1:20 dilution for 96-well plates. The cultures were initially incubated at 37 ^{o}C and 250 rpm for 4.5 hours, and after cooling to 25 ^{o}C, expression was induced by the addition of a mixture of IPTG (1 mM) and δ-aminolevulinic acid hydrochloride (δ-ALA; 0.5 mM). The cultures were harvested by centrifugation at 4 ^{o}C and 5,000 rpm for 10 min 20 to 24 hours after induction and stored at -20 ^{o}C. The protein yields varied from about 10 to 100 mg/L depending on the variant identities.

B.2. Expression of P450$_{cam}$, CYP153A6, fdrA6, fdxA6, and BM3 heme domain (Chapter 6)

The genes of P450$_{cam,}$ CYP153A6, CYP153A6 ferredoxin, and CYP153A6 ferredoxin reductase were cloned into the pET-22b(+) expression vector downstream of a T7 promoter using flanking restriction sites *Nde*I and *Xho*I. A his$_6$ affinity tag along with the amino acid pair of leucine and glutamic acid corresponding to the *Xho*I site were also introduced to the C-terminus of the P450cam and CYP153A6 constructs. The heme domain of P450 BM3 and its derived variants (amino acids 1-463) were cloned into a modified pCwori vector with flanking restriction sites *BamH*I and *Xho*I. The same his$_6$-tag with the leucine and glutamic acid appendage was also introduced to the C-terminus. *E. coli* BL21(DE3) cells transformed with these plasmids were used for expression in 50 – 500 mL scales. For expression, supplemented TB medium (500 mL, 100 μg/ml ampicillin, trace metal mix: 50 μM $FeCl_3$, 20 μM $CaCl_2$, 10

μM $MnSO_4$, 10 μM $ZnSO_4$, 2 μM $CoSO_4$, 2 μM $CuCl_2$, 2 μM $NiCl_2$, 2 μM $NaMoO_4$, and 2 μM H_3BO_3) was inoculated with TB overnight culture to an initial optical density at 600 nm (OD_{600}) of 0.5. After 3.5 hours of incubation at 37 ^oC and 250 rpm, the cultures were cooled to 25 ^oC, and expression was induced with the addition of IPTG (1 mM). The cells were harvested by centrifugation at 4 ^oC, 5,000 rpm for 10 mins, 20 to 24 hours after induction and stored at -20 ^oC. For the expression of fdrA6 and fdxA6, the cultures were cooled to 30 ^oC for induction and allowed to express for 24 to 30 hours at 30 ^oC.

C. Purification of P450s and Reductase Components

C.1. Single-step purification of P450 BM3 (Chapters 2, 3, 5,and 7)

A single-step purification protocol was developed based on a published procedure (*1*). During the purification, all samples were kept on ice or at 4 ^oC. Cell pellets of BM3 and its derived variants were resuspended in Tris-HCl (25 mM, pH 8.0, 0.5 ml/gram cell weight) and lysed by sonication (2 x 30 s, output control 5, 50% duty cycle; Sonicator, Heat Systems–Ultrasonic, Inc.). The lysate was centrifuged at 20,000 rpm for 30 min, and the supernantent was filtered with a 22 μM filter (Millipore, Billerica, MA) before loading on a pre-equilibrated Toyopearl® Super Q-650M column for ion exchange chromatography. After washing with five column volumes of 25 mM Tris-HCl and five column volumes of 0.15 M NaCl in 25 mM Tris-HCl, the protein was eluted with 0.34 M NaCl in 25 mM Tris-HCl. The P450 fractions were collected and concentrated using a 30 kDa molecular weight cut-off centrifugal (mwco) filter (Millipore, Billerica, MA). The protein solution was then buffer-exchanged with potassium phosphate buffer (100 mM, pH 8.0) by either using PD-10 desalting columns (GE heathcare, Piscataway, NJ), or through three dilution and concentration cycles in Amicon spin columns. The

P450 enzyme concentrations were quantified by CO-binding difference spectra as described (*2*), using 91 mM^{-1} cm^{-1} as the extinction coefficient. The protein solutions were flash-frozen with dry ice and stored at -80 ^{o}C.

C.2. Three-step purification of P450cam, CYP153A6, and BM3 heme domains (Chapter 6)

To obtain highly pure protein preparations for terminal oxidant reactions described in Chapter 6, a three-step purification was employed. During the purification, all samples were kept on ice or at 4 ^{o}C. Cell pellets were resuspended in Ni-NTA buffer A, (25 mM Tris-HCl, 100 mM NaCl, 30 mM imidazole, pH 8.0, 0.5 ml/gcw) and lysed by sonication (2 x 30 s, output control 5, 50% duty cycle; Sonicator, Heat Systems–Ultrasonic, Inc.). The lysate was centrifuged at 20,000 rpm for 30 min, and the supernantent was filtered with a 22 μM filter before loading on a pre-equilibrated Ni-NTA column.

After washing with five column volumes of Ni-NTA buffer A, the P450 was eluted with buffer B (25 mM Tris-HCl, 100 mM NaCl, 100 mM imidazole, pH 8.0). After buffer-exchange with 25 mM Tris-HCl, the desalted protein solution was loaded onto a pre-equilibrated Q sepharose^{TM} column (HiTrap^{TM} Q HP, GE healthcare, Piscataway, NJ). A wash step with five column volumes of 25 mM Tris-HCl was followed by elution of the P450 with a linear gradient of NaCl from 0 to 500 mM over 10 column volumes.

The P450 fractions were collected and concentrated using a 30 kDa mwco centrifugal filter. Following buffer exchange with 20 mM HEPES buffer, the protein solution was loaded onto a pre-equilibrated HiPrep 16/60 Sephacryl S-100 HR column (GE healthcare, Piscataway, NJ). The protein was eluted with a flow rate of 0.5 ml/min; purified fractions were concentrated to 2 mL using a 30 kDa mwco centrifugal filter. To remove potential alcohol contaminants, the purified

protein was flash-frozen with dry ice and lyophilized overnight on a bench-top lyophilizer (Millrock Technology, Kingston, NY). P450 concentration was determined in triplicate after resuspension in potassium phosphate buffer (100 mM pH 8.0) by CO-difference spectroscopy.

C.3. Two-step purification of fdrA6 and fdxA6 (Chapter 6)

To obtain the A6 reductase components for NADH/O_2-supported alkane reactions described in Chapter 6, a two-step purification was employed. During the purification, all samples were kept on ice or at 4 ^{o}C. Cell pellets were resuspended in 25 mM Tris-HCl at 0.5 ml/gcw and lysed by sonication (2 x 30 s, output control 5, 50% duty cycle; Sonicator, Heat Systems–Ultrasonic, Inc.). The lysate was centrifuged at 20,000 rpm for 30 min, and the supernantent was filtered with a 22 μM filter before loading on a pre-equilibrated Q sepharose^{TM} column (HiTrap^{TM} Q HP, GE healthcare, Piscataway, NJ). A wash step with five column volumes of 25 mM Tris-HCl was followed by elution of the P450 with a linear gradient of NaCl from 0 to 500 mM over 10 column volumes.

The protein fractions were collected and concentrated using a 10 and 30 kDa mwco centrifugal filter for fdxA6 and fdrA6, respectively. Following buffer exchange with 20 mM HEPES buffer, the protein solution was loaded onto a pre-equilibrated HiPrep 16/60 Sephacryl S-100 HR column (GE healthcare, Piscataway, NJ). The proteins were eluted with a flow rate of 0.5 ml/min; purified fractions were concentrated and flash-frozen with dry ice before lyophilization with a bench-top lyophilizer (Millrock Technology, Kingston, NY).

D. Mutagenesis and Library Construction of P450 BM3 Variants

D.1. Thermostabilization of variant 35E11 (Chapter 2)

In our effort to thermostabilize variant 35E11, we grafted known stabilizing mutations of the P450 heme domain (*3*). The following mutations L52I, S106R, M145A, L324I, V340M, I366V, and E442K in conjunction were found to stabilize a P450 BM3-derived peroxygenase by 18 ^{o}C. Each of these mutations was first individually introduced into the pCwori_35E11 plasmid by site-directed overlap extension PCR (SOE-PCR) (*4*). BamHI_Fwd and SacI_rev served as flanking primers. For all primer sequences see section D.10. For each mutation, two separate PCRs were performed with *Pfu* turbo (Stratagene), each using a perfectly complementary flanking primer and a mutagenic primer. The two resulting overlapping fragments containing the nucleotide substitutions were then annealed in a second PCR to amplify the complete mutated gene. The amplified gene was then digested with *BamH*I and *Sac*I restriction enzymes and ligated with T4 DNA ligase into pCwori_WT, previously digested with *BamH*I and *Sac*I to remove the wild-type gene. The ligation mixtures were used to transform electro-competent *E. coli* DH5α cells. The transformed cells were plated onto Luria-Bertani (LB) (*5*) agar plates containing ampicillin to form single colonies. Colonies from each of the seven variants were expressed and purified as described above. Mutations that resulted in an improved T_{50} (see Section J) and less than 20% loss of propane turnover number (TON) (see Section F.1) were recombined in a pair-wise fashion using SOE-PCR. The best variant identified in this thermostabilization effort was ETS8 with mutation L52I and I366V.

D.2. Random mutagenesis of ETS8 (Chapter 2)

A random mutagenesis library of the heme domain of variant ETS8 (aa 1-433) was created by error-prone PCR based on published protocols (*6*). Four different concentrations of $MnCl_2$ (50, 100, 200, 300 μM) were used to induce mutations with *Taq* DNA polymerase (Roche) in the mutagenic PCR containing approximately 50 ng of plasmid DNA as template, and BamHI_Fwd and SacI_rev serving as end primers. The PCR product was then digested with *BamH*I and *Sac*I restriction enzymes and ligated with T4 DNA ligase into pCwori_WT, previously digested with *BamH*I and *Sac*I to remove the wild-type gene. The ligation mixtures were then electoporated into electro-competent *E. coli* DH5α cells. The transformed cells were plated onto LBamp agar plates to form single colonies. For each Mn^{2+} concentration, a small library of 88 mutants in a 96-well plate (four wells containing parent ETS8, four blank wells) was picked and screened for both folding by CO-binding difference spectroscopy (E.2) as well as activity toward dimethyl ether as described below (E.3). The library prepared with a Mn^{2+} concentration of 200 μM yielded ~ 50% inactive mutants, and thus it was selected for screening of 2,500 members for activity towards dimethyl ether. Variant 19A12 was selected from the DME screening with roughly twice the parental activity.

D.3. Site-saturation mutagenesis and recombination of variant 19A12 (Chapter 2)

Using pCwori_19A12 as template, 17 residues along the active site channel, identified from the P450 BM3 crystal structure 1JPZ, were selected for saturation mutagenesis. NNK libraries were constructed individually at positions 74, 75, 78, 82, 87, 88, 181, 184, 188, 260, 264, 265, 268, 328, 401, 437, and 438 by SOE-PCR. The mutagenic primer for each library contained an NNK (N = A, T, G or C, K = T or G) codon at the mutation site, which encoded all 20 possible amino acids and only one stop codon out of the possible 32 codons. For each library, two separate PCRs were performed, each using a perfectly complementary end primer and a

mutagenic primer. The two resulting overlapping fragments were then annealed in a second PCR to amplify the complete mutated gene. For the libraries at residue 437 and 438, BamHI_Fwd and EcoRI_rev were used as the end primers, and digestion with *BamH*I and *EcoR*I was employed. For all other libraries BamHI_Fwd and SacI_Rev were used. End primers and the subsequent cloning into pre-digested pCwori_WT occurred as previously described (D.1).

For each library, 88 mutants in a 96-well plate (four wells containing parent 19A12, four blank wells) were picked and screened for folding and activity toward dimethyl ether. The five beneficial mutations A74S, A74Q, V184S, V184T, and V184A were recombined using degenerate primers or a mixture of primers. A second recombination library was also constructed that allowed for 74NNK, 82S/G, and 184NNK. Screening these two libraries produced variants 1-3 (19A12-A74S, V184A) and 7-7 (19A12-A74E, S82G) with comparable propane TTN.

D.4. Reductase libraries construction (Chapter 2)

The FMN and FAD sub-domains of the BM3 reductase were mutated separately by error-prone PCR with *Taq* Polymerase (Roche). Using pCwori_35E11 as template, the primer pairs FMN_for/ FMN_rev and FAD_for/FAD_rev were used in the PCR relying on $MnCl_2$ (100, 200, 300 μM) as the mutation inducer. The amplified genes were cloned into pre-digested pCwori_35E11 using restriction sites *Sac*I and *Nsi*I for FMN libraries and *Nsi*I and *Eco*RI for the FAD libraries. Screening for improved DME activity yielded beneficial mutations at positions 443, 445, 515, 580, 654, 664, 698, and 1037.

Using pCwori_11-3 as template, site-saturation libraries were constructed at all eight positions identified from the 35E11 random reductase libraries using NNK primers and SOE PCR as described (D.1). FMN_for and EcoRI_rev end primers were used to amplify the entire

reductase. Following *Sac*I and *EcoR*I digest, the mutagenized genes were ligated into a pre-digested pCwori_11-3 and electoporated into *E. coli* Dh5α. Screening of 88 colonies from each library identified beneficial mutations G443A, P654K, T664G, D698G, and E1037G. A recombination library was constructed using degenerate or mixed primers allowing for both the wild-type amino acid and mutated amino acid at each position. This reductase library was fused to pCwori_7-7 through the *Sac*I and *EcoR*I restriction sites. Screening of the resulting library yielded $P450_{PMO}$ R1 (7-7-G443A, P654K, and E1037G) and $P450_{PMO}$ R2 (7-7-G443A, D698G).

D.5. Targeted mutagenesis of the P450 BM3 active site (Chapter 3)

Ten residues along the active site channel, identified from the P450 BM3 crystal structure 1JPZ, were selected for saturation mutagenesis. Using pCwori_WT as template, NNK libraries were constructed individually at positions 74, 75, 78, 82, 87, 181, 184, 188, 328, and 330 by SOE-PCR as described (D.1). At least 91 colonies of each library were screened for folding by CO-binding difference spectroscopy and activity for dimethyl ether.

Combinatorial active site saturation test with a reduced set of allowed amino acid (reduced CASTing) libraries targeting two and three sites simultaneously were constructed at positions 78, 82, and 328 with degenerate codons that encoded for L, I, F, V, A, M, and W. To obtain the desired amino acid restriction at each site, three primers containing the codons GYN (encodes A and V), WTS (encodes L, I, M and F), and TGG (encodes W) were mixed to a 2:4:1 ratio to achieve equal representation of each amino acid. A modified SOE-PCR strategy was used for two-site focused libraries, where the target residues were more than 20 nucleotides apart, A328+V78 and A328+A82. Three overlapping fragments were first generated using the primer pairs, BamHI_Fwd/Site1_rev, Site1_fwd/Site2_rev, and Site2_fwd/SacI_Rev in three

separate reactions. A second PCR annealed these three fragments and amplified the full length gene with the end-primers BamHI_Fwd and SacI_Rev. The two-site focused library with target residues less than 20 nucleotides apart, V78+A82, was constructed with a sequential SOE-PCR strategy. Standard SOE-PCR was used to introduce the first mutagenic codon into the full-length gene, which then served as template for a second round of SOE-PCR. The three-site focus library, A328+V78+A82, was constructed similarly by sequential SOE-PCRs.

D.6. Random mutagenesis of P450 BM3 (Chapter 3)

A random mutagenesis library of the heme domain of P450 BM3 (aa 1-433) was created by error-prone PCR as described (D.2). Four different concentrations of $MnCl_2$ (50, 100, 150, and 200 μM) were used to induce mutations with *Taq* DNA polymerase (Roche). From screening 88 colonies of each of the four libraries for both folding and activity toward dimethyl ether, the library prepared with a 150 μM Mn^{2+} concentration, which contained ~ 50% inactive mutants, was selected for full screening. Ten variants were randomly selected for sequencing, which revealed an average error rate of 2.1 amino acids/protein. In the subsequent full screening, 1,408 members of this library were assayed for activity towards dimethyl ether.

D.7. Construction of CRAM- and C^{orbit}- designed libraries (Chapter 3)

The CRAM and C^{orbit} libraries were constructed to allow for two possible amino acids at ten selected residues. The allowed amino acids were determined for each design following the algorithms detailed in Appendix B. The selected amino acids and the degenerate codons or mixed codons that were used in library construction are summarized below in Table 8.1.

Table 8.1 CRAM and C^{orbit} designs; allowed amino acids and codon usage. (K = T or G, S = C or G, W = A or T, Y = C or T, R = A or G)

	CRAM Design		C^{orbit} Design	
Target Residue	Allowed Amino Acid	Codon	Allowed Amino Acid	Codon
A74	L/W	TKG	A/V	GYC
L75	L/F	TTS	L/F	TTS
V78	F/I	WTT	L/V	STG
A82	L/V	STG	A/S	KCG
F87	F/A	TTC/GCA	F/A	TTC/GCA
L181	L/W	TKG	L/F	TTS
A184	A/V	GYC	A/T	RCG
L188	L/W	TKG	L/W	TKG
A328	F/V	KTC	A/F	TTC/GCA
A330	L/W	TKG	A/V	GYC

For both libraries, the ten target positions were clustered into three different nodes for ease of primer design and PCR. The first node included residues 74, 75, 78, 82, and 87. A single mutagenic primer was able to span this entire node and introduce the desired codon mixture to all five positions. For the remaining targeted positions, residues 181, 184, and 188 were grouped into a second node, and residues 328 and 330 were grouped into a third node. Four overlapping fragments were first generated using BamHI_Fwd and SacI_Rev in conjunction with forward and reverse mutagenic primers at each node. These four fragments were annealed in a second PCR to amplify the full-length gene. Digestion with *BamH*I and *Sac*I was followed by ligation with T4 DNA ligase into a pre-digested pCwori_WT. The ligation mixture was then electroporated into electro-competent *E. coli* Dh5α cells and plated on LBamp. Single colonies

were picked into 96-well plates and screened for protein folding and activity for dimethyl ether as described above.

D.8. Random mutagenesis of AlkB and Cyp153A6 (Chapter 4)

Mutagenesis of pCom plasmids was performed in *E. coli* XL1-Red strains according to the manufacturer's manual (Stratagene) and in *E. coli* JS200 (pEPPol I) as described previously (*7*). Mutated *alkB* genes or the CYP153A6 gene along with the *fdrA6* and *fdxA6* operons were cloned into the original pCom10 or into pCom8* plasmids as *Eco*RI-*Hind*III or *Kpn*I-digested fragments, respectively. Plasmid pCom8_alkBFG was constructed by amplifying the *alkBFG* operon from plasmid pblaP4_*alkJBFG-luxAB* (*8*) using the primers alkBFG_1 and alkBFG_2 in a standard PCR. The resulting fragment was cloned into the pCom8 vector using the *Nde*I and *Xma*I restriction sites introduced by the primers. The resulting plasmid was digested with *Spe*I, and the *alkB*-containing 3.7 -kb fragment was replaced with the appropriate fragment from *Spe*I-digested pCom10_alkB, pCom10_alkB-BMO1, and pCom10_alkB-BMO2 by cloning, resulting in constructs pCom8_alkBFG, pCom8_alkB-BMO1_alkFG, and pCom8_alkBBMO2_alkFG.

D.9. Site-saturation and random mutagenesis of variant E31 (Chapter 5)

Ten residues along the active site channel, identified from the P450 BM3 crystal structure 1JPZ, were selected for saturation mutagenesis. Using pCwori_E31 identified from the CRAM and C^{orbit} screening as template, NNK libraries were constructed individually at positions 74, 75, 78, 82, 184, 263, 264, 328, 436, and 437 by SOE-PCR as described (D.1). At least 88 colonies of each library were screened for protein folding by CO-binding difference spectroscopy and activity for ethane hydroxylation (E.4).

The first-generation random mutagenesis library targeting the heme domain of variant E31 (aa 1-490) was created by error-prone PCR using the Genemorph II kit (Stratagene, La Jolla, CA) according to the manufacturer's protocol. Approximately 100 ng of plasmid DNA were used as template along with primers Heme_fwd_1 and KpnI_rev in the PCR. Sequencing of ten randomly selected variants revealed an average nucleotide substitution rate of 3.7/protein was obtained for this library. After screening 3,000 members of this library, we obtained variants 24F8 and 22F11 with improved ethane hydroxylation activity with three and two mutations, respectively. SOE-PCR was used to generate a recombination library with these five mutations. From screening 90 colonies of this library for improved ethane hydroxylation activity, we obtained variant RD2 with only two of the five mutations.

The second-generation random mutagenesis library targeting the heme domain was constructed using pCwori_RD2 as the template at three different concentrations of $MnCl_2$ (175, 200, and 300 μM) to induce mutations with *Taq* DNA polymerase (Roche). Sequencing of ten randomly selected variants revealed that an average nucleotide substitution rate of 4.3/protein was obtained for this library. A pre-screen of 88 colonies from each library led us to choose the 175 μM Mn^{2+} library for screening. After screening 3,000 members of this library for improved ethane hydroxylation activity, we obtained variants 20D4.

D.10. Primer list

Table 8.2 Primers used in P450 library construction

#	Primer name	SEQUENCE
1	L52I_for	5'-CGCGCTACATATCAAGTCAGC-3'
2	L52I_rev	5'-GCTGACTTGATATGTAGCGCG-3'
3	M145A_for	5'-GTATCGGAAGACGCGACACGTTTAACG-3'
4	M145A_rev	5'-GTATCGGAAGACGCGACACGTTTAACG-3'
5	V340M_for	5'-GAAGATACGATGCTTGGAGGAG-3'

6	V340M_rev	5’-CTCCTCCAAGCATCGTATCTTC-3’
7	I366V_for	5’-CGTGATAAAACAGTTTGGGGAGACG-3’
8	I366V_rev	5’-CGTCTCCCCAAACTGTTTTATCACG-3’
9	E442K_for	5’-CGTTAAAACCTAAAGGCTTTGTGG-3’
10	E442K_rev	5’-CCACAAAGCCTTTAGGTTTTAACG-3’
11	L324I_for	5’-CGAAGCGCTGCGCATCTGGCCAACTT-3’
12	L324I_rev	5’-AAGTTGGCCAGATGCGCAGCGCTTCG-3’
13	S106R_for	5’-CTTACTTCCAAGGTTCAGTCAGCAGG-3’
14	S106R_rev	5’-CCTGCTGACTGAACCTTGGAAGTAAG-3’
15	BamHI_fwd	5’-CACAGGAAACAGGATCCATCGATGCTTAGG-3’
16	SacI_rev	5’-CTAGGTGAAGGAATACCGCCAAGCGGA-3’
17	L437NNK_for	5’-CGATATTAAAGAAACTNNKACGTTAAAACC-3’
18	L437NNK_rev	5’-GGTTTTAACGTMNNAGTTTCTTTAATATCG-3’
19	T438NNK_for	5’-CGATATTAAAGAAACTTTANNKTTAAAACC-3’
20	T438NNK_rev	5’-GGTTTTAAMNNTAAAGTTTCTTTAATATCG-3’
21	EcoRI_Rev	5’-CCGGGCTCAGATCTGCTCATGTTTGACAGC-3’
22	L181NNK_for	5’-GGTCCGTGCANNKGATGAAGTAATG-3’
23	L181NNK_rev	5’-CATTACTTCATCMNNTGCACGGACC-3’
24	A82NNK_for	5’-CGTGATTTTNNKGGAGACGGGTTA-3’
25	A82NNK_rev	5’-TAACCCGTCTCCMNNAAAATCACG-3’
26	A74NNK_for	5’-AACTTAAGTCAANNKCTTAAATTC-3’
27	A74NNK_rev	5’-GAATTTAAGMNNTTGACTTAAGTT-3’
28	L75NNK_for	5’-GTCAAGCGNNKAAATTCTTTCGTG-3’
29	L75NNK_rev	5’-CACGAAAGAATTTMNNCGCTTGAC-3’
30	V78NNK_for	5’-GTCAAGCGCTTAAATTCNNKCGTGATTTT-3’
31	V78NNK_rev	5’-AAAATCACGMNNGAATTTAAGCGCTTGAC-3’
32	A328NNK_for	5’-GGCCAACTNNKCCTGCGTTTTCC-3’
33	A328NNK_rev	5’-GGAAAACGCAGGMNNAGTTGGCC-3’
34	A184NNK_for	5’-GCACTGGATGAANNKATGAACAAG-3’
35	A184NNK_rev	5’-CTTGTTCATMNNTTCATCCAGTGC-3’
36	L188NNK_for	5’-GAACAAGNNKCAGCGAGCAAATCC-3’
37	L188NNK_rev	5’-GGATTTGCTCGCTGMNNCTTGTTC-3’
38	I401NNKfwd	5’-GCGTGCGTGTNNKGGTCAGCAG-3’
39	I401NNKrev	5’-CTGCTGACCMNNACACGCACGC-3’
40	T268NNKfwd	5’-GCGGGACACGAANNKACAAGTGGTC-3’
41	T268NNKrev	5’-GACCACTTGTMNNTTCGTGTCCCGC-3’
42	G265NNKfwd	5’-CATTCTTAATTGCGNNKCACGAAACAACAAGTG-3’
43	G265NNKrev	5’-CACTTGTTGTTTCGTGMNNCGCAATTAAGAATG-3’
44	A264NNKfwd	5’-CATTCTTAATTNNKGGACACGAAACAACAAGTG-3’
45	A264NNKrev	5’-CACTTGTTGTTTCGTGTCCMNNAATTAAGAATG-3’

46	T260NNKfwd	5’-CAAATTATTNNKTTCTTAATTGCGGGAC-3’
47	T260NNKrev	5’-GTCCCGCAATTAAGAAMNNAATAATTTG-3’
48	L75NNKfwd	5’-GTCAAGCGNNKAAATTTGTACG-3’
49	L75NNKrev	5’-GTCCCGCAATTAAGAAMNNAATAATTTG-3’
50	T88NNKfwd	5’-GACGGGTTATTTNNKAGCTGGACGCATG-3’
51	T88NNKrev	5’-GTCCCGCAATTAAGAAMNNAATAATTTG-3’
52	F87NNKfwd	5’-GACGGGTTANNKACAAGCTGG-3’
53	F87NNKrev	5’-CCAGCTTGTMNNTAACCCGTC-3’
54	A82G_for	5’-CGTGATTTTGGTGGAGACGGGTTA-3’
55	A82G_rev	5’-TAACCCGTCTCCACCAAAATCACG-3’
56	FMN_for	5’-GCTGGTACTTGGTATGATGCT-3’
57	FMN_rev	5’-CCAGACGGATTTGCTGTGAT-3’
58	FAD_for	5’-CGTGTAACAGCAAGGTTCGG-3’
59	FAD_rev	5’-CTGCTCATGTTTGACAGCTTATC-3’
60	G443NNK_for	5’-CGTTAAAACCTGAANNKTTTGTGG-3’
61	G443NNK_rev	5’-CCACAAAMNNTTCAGGTTTTAACG-3’
62	V445NNK_for	5’-CCTGAAGGCTTTNNKGTAAAAGCA-3’
63	V445NNK_rev	5’-TGCTTTTACMNNAAAGCCTTCAGG-3’
64	T480NNK_for	5’-CGCTCATAATNNKCCGCTGCTTGTG-3’
65	T480NNK_rev	5’-CACAAGCAGCGGMNNATTATGAGCG-3’
66	T515NNK_for	5’-CCGCAGGTCGCANNKCTTGATTCAC-3’
67	T515NNK_rev	5’-GTGAATCAAGMNNTGCGACCTGCGG-3’
68	P654NNK_for	5’-GCGGATATGNNKCTTGCGAAAATG-3’
69	P654NNK_rev	5’-CATTTTCGCAAGMNNCATATCCGC-3’
70	T664NNK_for	5’-GGTGCGTTTTCANNKAACGTCGTAGCA-3’
71	T664NNK_rev	5’-TGCTACGACGTTMNNTGAAAACGCACC-3’
72	D698NNK_for	5’-CAAGAAGGANNKCATTTAGGTG-3’
73	D698NNK_rev	5’-CACCTAAATGMNNTCCTTCTTG-3’
74	E1037NNK_for	5’-CAGCAGCTAGAANNKAAAGGCCG-3’
75	E1037NNK_rev	5’-CGGCCTTTMNNTTCTAGCTGCTG-3’
76	BMfor_1504	5’-GCAGATATTGCAATGAGCAAAGG-3’
77	BMrev1504	5’-CCTTTGCTCATTGCAATATCTGC-3’
78	BMfor2315	5’-CGGTCTGCCCGCCGCATAAAG-3’
79	BMrev2315	5’-CTTTATGCGGCGGGCAGACCG-3’
80	Heme fwd 1	5'-CAGGAAACAGGATCAGCTTACTCCCC-3'
82	74NNK fwd	5'-GATAAAAACTTAAGTCAANNKCTTAAATTTGTACGTG-3'
83	74NNK rev	5'-CACGTACAAATTTAAGMNNTTGACTTAAGTTTTTATC-3'
84	75NNK fwd	5'-GATAAAAACTTAAGTCAAGCGNNKAAATTTGTACGTG-3'
85	75NNK rev	5'-CACGTACAAATTTMNNCGCTTGACTTAAGTTTTTATC-3'
86	78NNK fwd	5'-GCGCTTAAATTTNNKCGTGATTTTGCAGG-3'

87	78NNK rev	5'-CCTGCAAAATCACGMNNAAATTTAAGCGC-3'
88	81NNK fwd	5'-GTACGTGATNNKGCAGGAGACGGG-3'
89	81NNK rev	5'-CCCGTCTCCTGCMNNATCACGTAC-3'
90	82NNK fwd	5'-GTACGTGATTTTNNKGGAGACGGG-3'
91	82NNK rev	5'-CCCGTCTCCMNNAAAATCACGTAC-3'
92	87NNK fwd	5'-GACGGGTTANNKACAAGCTGGACGC-3'
93	87NNK rev	5'-GCGTCCAGCTTGTMNNTAACCCGTC-3'
94	88NNK fwd	5'-GACGGGTTATTTNNKAGCTGGACGC-3'
95	88NNK rev	5'-GCGTCCAGCTMNNAAATAACCCGTC-3'
96	181NNK fwd	5'-GTCCGTGCANNKGATGAAGCAATGAAC-3'
97	181NNK rev	5'-GTTCATTGCTTCATCMNNTGCACGGAC-3'
98	184NNK fwd	5'-GCACTGGATGAANNKATGAACAAGCTG-3'
99	184NNK rev	5'-CAGCTTGTTCATMNNTTCATCCAGTGC-3'
100	188NNK fwd	5'-CAATGAACAAGNNKCAGCGAGCAAATCC-3'
101	188NNK rev	5'-GGATTTGCTCGCTGMNNCTTGTTCATTG-3'
102	263NNK fwd	5'-AATTATTACATTCTTANNKGCGGGACACGAAACAAC-3'
103	263NNK rev	5'-GTTGTTTCGTGTCCCGCMNNTAAGAATGTAATAATT-3'
104	264NNK fwd	5'-AATTATTACATTCTTAATTNNKGGACACGAAACAAC-3'
105	264NNK rev	5'-GTTGTTTCGTGTCCMNNAATTAAGAATGTAATAATT-3'
106	328NNK fwd	5'-GGCCAACTNNKCCTGCGTTTTCCC-3'
107	328NNK rev	5'-GGGAAAACGCAGGMNNAGTTGGCC-3'
108	330NNK fwd	5'-GCCAACTGCTCCTNNKTTTTCCCTATATG-3'
109	330NNK rev	5'-CATATAGGGAAAAMNNAGGAGCAGTTGGC-3'
110	V78VA fwd	5'-GCGCTTAAATTTGYNCGTGATTTTGCAGG-3'
111	V78VA rev	5'-CCTGCAAAATCACGNRCAAATTTAAGCGC-3'
112	V78LIMF fwd	5'-GCGCTTAAATTTWTSCGTGATTTTGCAGG-3'
113	V78LIMF rev	5'-CCTGCAAAATCACGSAWAAATTTAAGCGC-3'
114	V78W fwd	5'-GCGCTTAAATTTTGGCGTGATTTTGCAGG-3'
115	V78W rev	5'-CCTGCAAAATCACGCCAAAATTTAAGCGC-3'
116	A82VA fwd	5'-GTACGTGATTTTGYNGGAGACGGG-3'
117	A82VA rev	5'-CCCGTCTCCNRCAAAATCACGTAC-3'
118	A82LIMF fwd	5'-GTACGTGATTTTWTSGGAGACGGG-3'
119	A82LIMF rev	5'-CCCGTCTCCSAWAAAATCACGTAC-3'
120	A82W fwd	5'-GTACGTGATTTTTGGGGAGACGGG-3'
121	A82W rev	5'-CCCGTCTCCCCAAAAATCACGTAC-3'
122	A328VA fwd	5'-GGCCAACTGYNCCTGCGTTTTCCC-3'
123	A328VA rev	5'-GGGAAAACGCAGGNRCAGTTGGCC-3'
124	A328LIMF fwd	5'-GGCCAACTWTSCCTGCGTTTTCCC-3'
125	A328LIMF rev	5'-GGGAAAACGCAGGSAWAGTTGGCC-3'
126	A328W fwd	5'-GGCCAACTTGGCCTGCGTTTTCCC-3'

127	A328W rev	5'-GGGAAAACGCAGGCCAAGTTGGCC-3'
128	V78A82MNYY fwd	5'-GCGCTTAAATTTATGCGTGATTTTNYYGGAGACGGG-3'
129	V78A82MNYY rev	5'-CCCGTCTCCRRNAAAATCACGCATAAATTTAAGCGC-3'
130	V78A82NYYM fwd	5'-GCGCTTAAATTTNYYCGTGATTTTATGGGAGACGGG-3'
131	V78A82NYYM rev	5'-CCCGTCTCCCATAAAATCACGRRNAAATTTAAGCGC-3'
132	V78A82KGGNYY fwd	5'-GCGCTTAAATTTKGGCGTGATTTTNYYGGAGACGGG-3'
133	V78A82KGGNYY rev	5'-CCCGTCTCCRRNAAAATCACGCCMAAATTTAAGCGC-3'
134	V78A82NYYKGG fwd	5'-GCGCTTAAATTTNYYCGTGATTTTKGGGGAGACGGG-3'
135	V78A82NYYKGG rev	5'-CCCGTCTCCCCMAAAATCACGRRNAAATTTAAGCGC-3'
136	V78A82KGG2 fwd	5'-GCGCTTAAATTTKGGCGTGATTTTKGGGGAGACGGG-3'
137	V78A82KGG2 rev	5'-CCCGTCTCCCCMAAAATCACGCCMAAATTTAAGCGC-3'
138	V78A82NYY2 fwd	5'-GCGCTTAAATTTNYYCGTGATTTTNYYGGAGACGGG-3'
139	V78A82NYY2 rev	5'-CCCGTCTCCRRNAAAATCACGRRNAAATTTAAGCGC-3'
140	V78A82M2 fwd	5'-GCGCTTAAATTTATGCGTGATTTTATGGGAGACGGG-3'
141	V78A82M2 rev	5'-CCCGTCTCCCATAAAATCACGCATAAATTTAAGCGC-3'
142	F87T88MNYY fwd	5'-GACGGGTTAATGNYYAGCTGGACGC-3'
143	F87T88MNYY rev	5'-GCGTCCAGCTRRNCATTAACCCGTC-3'
144	F87T88NYYM fwd	5'-GACGGGTTANYYATGAGCTGGACGC-3'
145	F87T88NYYM rev	5'-GCGTCCAGCTCATRRNTAACCCGTC-3'
146	F87T88KGGNYY fwd	5'-GACGGGTTAKGGNYYAGCTGGACGC-3'
147	F87T88KGGNYY rev	5'-GCGTCCAGCTRRNCCMTAACCCGTC-3'
148	F87T88NYYKGG fwd	5'-GACGGGTTANYYKGGAGCTGGACGC-3'
149	F87T88NYYKGG rev	5'-GCGTCCAGCTCCMRRNTAACCCGTC-3'
150	F87T88KGG2 fwd	5'-GACGGGTTAKGGKGGAGCTGGACGC-3'
151	F87T88KGG2 rev	5'-GCGTCCAGCTCCMCCMTAACCCGTC-3'
152	F87T88NYY2 fwd	5'-GACGGGTTANYYNYYAGCTGGACGC-3'
153	F87T88NYY2 rev	5'-GCGTCCAGCTRRNRRNTAACCCGTC-3'
154	F87T88M2 fwd	5'-GACGGGTTAATGATGAGCTGGACGC-3'
155	F87T88M2 rev	5'-GCGTCCAGCTCATCATTAACCCGTC-3'
156	A328KGG fwd	5'-GGCCAACTKGGCCTGCGTTTTCCC-3'
157	A328KGG rev	5'-GGGAAAACGCAGGCCMAGTTGGCC-3'
158	A328NYY fwd	5'-GGCCAACTNYYCCTGCGTTTTCCC-3'
159	A328NYY rev	5'-GGGAAAACGCAGGRRNAGTTGGCC-3'
160	N1F fwd	5'-CTTAAGTCAATKGTTSAAATTTWTTCGTGATTTTS TGGGAGACGGGTTATTCACAAGCTGG-3'
161	N1F rev	5'-CCAGCTTGTGAATAACCCGTCTCCCASAAAATCACG AAWAAATTTSAACMATTGACTTAAG-3'
162	N1A fwd	5'-CTTAAGTCAATKGTTSAAATTTWTTCGTGATTTTSTGG GAGACGGGTTAGCAACAAGCTGG-3'
163	N1A rev	5'-CCAGCTTGTTGCTAACCCGTCTCCCASAAAATCACGA AWAAATTTSAACMATTGACTTAAG-3'

164	N2 fwd	5'-GTCCGTGCATKGGATGAAGYCATGAACAAGTKGCAGCGAGCAAATC-3'
165	N2 rev	5'-GATTTGCTCGCTGCMACTTGTTCATGRCTTCATCCMATGCACGGAC-3'
166	N3 fwd	5'-GCTTATGGCCAACTKTCCCTTKGTTTTCCC-3'
167	N3 rev	5'-GGGAAAACMAAGGGAMAGTTGGCCATAAGC-3'
168	O N1F fwd	5'-CTTAAGTCAAGYCTTSAAATTTSTGCGTGATTTTKCGGGA GACGGGTTATTCACAAGCTGG-3'
169	O N1F rev	5'-CCAGCTTGTGAATAACCCGTCTCCCGMAAAATCACGCASAA ATTTSAAGRCTTGACTTAAG-3'
170	O N1A fwd	5'-CTTAAGTCAAGYCTTSAAATTTSTGCGTGATTTTKCGGGAG ACGGGTTAGCAACAAGCTGG-3'
171	O N1A rev	5'-CCAGCTTGTTGCTAACCCGTCTCCCGMAAAATCACGCASA AATTTSAAGRCTTGACTTAAG-3'
172	O N2 fwd	5'-GTCCGTGCATTSGATGAARCGATGAACAAGTKGCAGCGAGCAAATC-3'
173	O N2 rev	5'-GATTTGCTCGCTGCMACTTGTTCATCGYTTCATCSAATGCACGGAC-3'
174	O N3F fwd	5'-GCTTATGGCCAACTTTCCCTGYCTTTTCCC-3'
175	O N3F rev	5'-GGGAAAAGRCAGGGAAAGTTGGCCATAAGC-3'
176	O N3A fwd	5'-GCTTATGGCCAACTGCACCTGYCTTTTCCC-3'
177	O N3A rev	5'-GGGAAAAGRCAGGTGCAGTTGGCCATAAGC-3'
178	N1 NNK fwd	5'-CTTAAGTCAANNKNNKAAATTTNNKCGTGATTTTNNKG GAGACGGGTTANNKACAAGCTGG-3'
179	N1 NNK rev	5'-CCAGCTTGTMNNTAACCCGTCTCCMNNAAAATCACGM NNAAATTTMNNMNNTTGACTTAAG-3'
180	N2 NNK fwd	5'-GTCCGTGCANNKGATGAANNKATGAACAAGNNK CAGCGAGCAAATC-3'
181	N2 NNK rev	5'-GATTTGCTCGCTGMNNCTTGTTCATMNNTTCATC MNNTGCACGGAC-3'
182	N3 NNK fwd	5'-GCTTATGGCCAACTNNKCCTNNKTTTTCCC-3'
183	N3 NNK rev	5'-GGGAAAAMNNAGGMNNAGTTGGCCATAAGC-3'
184	alkBFG_1	5'-GATCTACATATGCTTGAGAAACACAGAGTTCTGGATTC-3'
185	alkBFG_2	5'-GATCTACCCGGGTCACTTTTCCTCGTAGAGCACATAGTC-3'
186	E31-74NNK fwd	5'-GATAAAAACTTAAGTCAANNKCTGAAATTTGTACGTG-3'
187	E31-74NNK rev	5'-CACGTACAAATTTCAGMNNTTGACTTAAGTTTTTATC-3'
188	E31-75NNK fwd	5'-GATAAAAACTTAAGTCAACTGNNKAAATTTGTACGTG-3'
189	E31-75NNK rev	5'-CACGTACAAATTTMNNCAGTTGACTTAAGTTTTTATC-3'
190	E31-78NNK fwd	5'-CTGCTGAAATTTNNKCGTGATTTTTTGGG-3'
191	E31-78NNK rev	5'-CCCAAAAAATCACGMNNAAATTTCAGCAG-3'
192	E31-82NNK fwd	5'-ATTCGTGATTTTNNKGGAGACGGG-3'
193	E31-82NNK rev	5'-CCCGTCTCCMNNAAAATCACGAAT-3'
194	E31-184NNK fwd	5'-GCACTGGATGAANNKATGAACAAGTGG-3'
195	E31-184NNK rev	5'-CCACTTGTTCATMNNTTCATCCAGTGC-3'

196	E31-263NNK fwd	5'-AATTATTACATTCTTANNKGCGGGACACGAAACAAC-3'
197	E31-263NNK rev	5'-GTTGTTTCGTGTCCCGCMNNTAAGAATGTAATAATT-3'
198	E31-264NNK fwd	5'-AATTATTACATTCTTAATTNNKGGACACGAAACAAC-3'
199	E31-264NNK rev	5'-GTTGTTTCGTGTCCMNNAATTAAGAATGTAATAATT-3'
200	E31-328NNK fwd	5'-GGCCAACTNNKCCTTGGTTTTCCC-3'
201	E31-328NNK rev	5'-GGGAAAACCAAGGMNNAGTTGGCC-3'
202	E31-436NNK fwd	5'-CGATATTAAAGAANNKTTAACGTTAAAACC-3'
203	E31-436NNK rev	5'-GGTTTTAACGTTAAMNNTTCTTTAATATCG-3'
204	E31-437NNK fwd	5'-CGATATTAAAGAAACTNNKACGTTAAAACC-3'
205	E31-437NNK rev	5'-GGTTTTAACGTMNNAGTTTCTTTAATATCG-3'
206	E31-140 fwd	5'-GATGAGCATATTGAWGTACCGGAAGAC-3'
207	E31-140 rev	5'-GTCTTCCGGTACWTCAATATGCTCATC-3'
208	E31-215 fwd	5'-GGTGATGAACGACCYAGTAGATAAAATTATTGCAG-3'
209	E31-215 rev	5'-CTGCAATAATTTTATCTACTRGGTCGTTCATCACC-3'
210	E31-454 fwd	5'-CGAAAAAAATTYCGCTTGGCGGTATTCC-3'
211	E31-454 rev	5'-GGAATACCGCCAAGCGRAATTTTTTTCG-3'
212	E31-222 fwd	5'-GATAAAATTATTGCAGAKCGCAAAGCAAGCGG-3'
213	E31-222 rev	5'-CCGCTTGCTTTGCGMTCTGCAATAATTTTATC-3'
214	E31-289 fwd	5'-GTATTACAAAAAGMAGCAGAAGAAGCAGC-3'
215	E31-289 rev	5'-GCTGCTTCTTCTGCTKCTTTTTGTAATAC-3'
216	RD2-432 fwd	5'-CTACGAGCTCGRTATTAAAGAAAC-3'
217	RD2-432 rev	5'-GTTTCTTTAATAYCGAGCTCGTAG-3'
218	RD2-74 fwd	5'-GATAAAAACTTAAGTCAACKGCTGAAATTTGTACGTG-3'
219	RD2-74 rev	5'-CACGTACAAATTTCAGCMGTTGACTTAAGTTTTTATC-3'
220	RD2-78I fwd	5'-CTGCTGAAATTTATTCGTGATTTTTTGGG-3'
221	RD2-78I rev	5'-CCCAAAAAATCACGAATAAATTTCAGCAG-3'
222	RD2-78W fwd	5'-CTGCTGAAATTTTGGCGTGATTTTTTGGG-3'
223	RD2-78W rev	5'-CCCAAAAAATCACGCCAAAATTTCAGCAG-3'
224	RD2-436 fwd	5'-CGATATTAAAGAAASATTAACGTTAAAACC-3'
225	RD2-436 rev	5'-GGTTTTAACGTTAATSTTTCTTTAATATCG-3'

E. High-Throughput Screening of P450 BM3 Libraries

E.1. Lysate preparation of high-throughput screening (Chapters 2, 3, 5, and 7)

96-well plates with 1-mL wells containing LB medium (300 μL, supplemented with 100 mg/mL ampicillin) were inoculated with single colonies either by hand or using a Qpix colony

picking robot (Genetix, Beaverton, OR). Two to four wells were also inoculated with the parent by hand, and one to four wells was left uninoculated as sterility and negative control for high-throughput screening. After overnight incubation (at least 12 hrs) at 37 ^{o}C, 250 rpm, and 80% relative humidity in a Kühner shaker (Kühner, Birsfeden, Switzerland), 50 µL of the LB culture was used to inoculate 96-well plates with 2-mL wells containing TB medium (600 – 800 µL, supplemented with 100 mg/mL ampicillin). The TBamp expression culture was incubated at 37 ^{o}C, 250 rpm, and 80% relative humidity for four hours before the temperature was reduced to 25 ^{o}C. After thirty minutes at the reduced temperature, protein expression was induced with 0.5 mM IPTG and 0.5 mM δ-ALA. After 20 – 24 hours of expression at 25 ^{o}C, the cultures were harvested by centrifugation at 5,000 rpm and 4 ^{o}C, and stored at -20 ^{o}C.

The frozen cell pellets were resuspended in 0.1 M phosphate buffer, pH 8.0, containing 10 mM $MgCl_2$, 0.5 mg/ml lysozyme (Sigma), and 2.0 U/mL DNaseI. After lysis at 37 ^{o}C for 60 minutes, the lysates were centrifuged at 5,000 rpm and 4 ^{o}C for 10 min, and the supernatants were transferred to 96-well microtiter plates for high-throughput assays.

E.2. High-throughput P450 protein folding assay (Chapters 2 – 7)

The CO-binding difference spectroscopy has been well established for determination of folded P450 concentrations. Following a modified protocol (*9*) for the CO-binding assay in 96-well microtiter plates, 40 – 100 µL of 0.4 M sodium hydrosulfite were added to 100 – 160 µL of lysate in each well of a 96-well microtiter plate. The UV/Vis absorbance from 400 to 500 nm was taken using a microtiter plate reader before exposure to carbon monoxide. The microtiter plates were placed in a vacuum chamber, which was evacuated to 20 mmHg. Carbon monoxide was then used to fill the chamber to atmosphere pressure. After a 15 minute incubation period,

the 400 – 500 nm spectrum was obtained and subtracted from the pre-incubation spectrum. The extinction coefficient of 91 mM cm^{-1} was used to obtain the P450 concentration of each well.

E.3. High-throughput demethylation assay using Purpald (Chapters 2, 3 and 7)

For quantifying the demethylation reaction of P450 with substrates such as dimethyl ether, purpald was used to detect the resulting aldehydes (*10*). In a 96-well microtiter plate, 120 µL of 0.1 M, pH 8.0 phosphate buffer pre-saturated with dimethyl ether were added to 30 µL of cell lysate, and the reaction was initiated with 50 µL of 4 mM NADPH. The reaction was allowed to proceed for 15 – 30 minutes, after which 50 µL of 168 mM purpald in 2 M NaOH were added to quench the reaction and develop the purple adduct with the released aldehydes. After 15 minutes, the purple color was quantified by absorbance at 550 nm using a microtiter plate reader.

E.4. High-throughput ethane hydroxylation assay (Chapters 3 and 5)

The ethanol product from the ethane hydroxylation reaction was quantified with a coupled enzyme assay utilizing alcohol oxidase and horseradish peroxidase. The alcohol oxidase converted the ethanol product from the P450 reaction into one molecule of acetaldehyde and one molecule of hydrogen peroxide. The horseradish peroxidase was then able to oxidize ABTS into a green soluble end-product using hydrogen peroxide.

To minimize ethanol fermentation from intact cells, the lysate was first filtered through a 0.2 micron filter plate (Pall, Ann Arbor, MI). In a 96-well Symx reactor (Symx, Miami, FL), 20 µL of lysate were added to each well along with 440 µL of 0.1 M phosphate buffer, pH 8.0, pre-saturated with ethane, and the reaction was initiated with 40 µL of an NADPH regeneration system consisting of 20 mM DL-isocitric acid trisodium salt, 400 µM $NADP^+$, and 0.5 U/mL

isocitrate dehydrogenase. The reactor was then pressurized to 30 psi with ethane and allowed to react for 24 – 48 hours at 4 ^{o}C. After the reactor was vented to atmospheric pressure, the reaction mixture was transferred to a 2-mL 96-well plate. To each well, 15 μL of 5 M HCl were added to oxidize the remaining NADPH to $NADP^{+}$ to prevent the NADPH from reducing the oxidized ABTS product. The 96-well plate was then shaken to mix at 150 rpm, 25 ^{o}C for 10 minutes. After centrifugation to pellet the denatured protein, 150 μL of the reaction mixture were transferred to a 96-well microtiter plate. Another 3 μL of 5 M HCl were added followed by 75 μL of 1 M phosphate buffer, pH 8.0, to raise the solution to a neutral pH. In the final color development setup, 50 μl of ABTS (1.3 mg/mL), 10 μL of HRP (3 mg/ml), and 10 μL of AOX (0.5 U/ml) were added in this sequence, and after 10 minutes of incubation, the green color was quantified at 420 nm.

E.5. Determination of NAD(P)H consumption rates (Chapter 2 – 7)

The P450 enzymes were purified and quantified by CO-binding as described above. Initial rates of NADPH consumption were determined by monitoring the absorption decrease at 340 nm measured with a 1 cm pathlength cuvette on a Cary 100 UV/VIS spectrophotometer (Varian, Walnut Creek, CA) or with a 96-well microtiter plate (pathlength 0.48 cm for 150 μL sample volume) on a Tecan plate reader (Tecan, Durham, NC). For liquid substrates such as alkanes, substrate stock solutions in ethanol (50-fold) were added to a protein solution (50 to 200 nM) in 0.1 M, pH 8.0 phosphate buffer and incubated for at least 30 seconds. The reaction was initiated by the addition of NAD(P)H varying from 200 μM to 1 mM depending on the rate of consumption and instrument range. For gaseous substrates, such as propane or ethane, the protein

solution was made in pre-saturated buffer. The consumption rates were determined using an extinction coefficient of ε_{340nm} = 6,210 M^{-1} cm^{-1}. Each rate was determined at least in triplicate.

E.6. Growth selection of AlkB and CYP153 libraries (Chapter 4)

Luria-Bertani broth and modified M9 medium with 1.5% yeast extract (*11*) supplemented with appropriate antibiotics, or E2 (*12*) and M9 (*13*) minimal media supplemented with carbon sources, were used for growth. All cultures were grown aerobically at 30 °C (*P. putida*) or 37 °C (*Escherichia coli*). Antibiotic concentrations were 100 μg/ml ampicillin, 15 μg/ml tetracycline, 20 μg/ml chloramphenicol, and 10 μg/ml gentamicin for *E. coli* cultures and 50 μg/ml for *Pseudomonas*. Bacterial strains were grown on solid E2 minimal medium with liquid alkanes provided through the gas phase as described previously (*14*). Solid minimal medium growth tests on gaseous alkanes were conducted in gas-tight plastic containers (GasPak 150 large anaerobic vented system; VWR), pressurized at 20 psi for 20 s (ethane and propane) or 10 psi for 6 s (butane). Liquid minimal medium cultures growing on alkanes were shaken in custom-made gas-tight flasks with 1% liquid alkane (pentane and octane) in a reservoir or in gas-tight serum bottles (Alltech), pressurized with gaseous alkanes as described above. For growth tests of *P. putida* GPo12(pGEc47ΔB) on alkanols, cells from an LB preculture were washed three times with M9 medium and used to inoculate the 5-mL M9 main cultures in 14-mL tubes to an optical density (OD_{600}) of 0.1, then grown at 30 °C with continuous shaking. Libraries of *P. putida* GPo12(pGEc47ΔB) strains expressing AlkB or CYP153A6 variants were precultured on E2 minimal medium plates with antibiotics and 0.2% (wt/vol) citrate as carbon source and then enriched for improved strains through continuous growth in liquid E2 minimal medium with small-chain-length alkanes as the sole carbon source

F. P450 Reactions (Chapters 2 – 7)

F.1. Gaseous alkane hydroxylation with P450 BM3 variants (Chapter 2 and 3)

Gaseous alkane reactions were carried out in 10 or 20 mL crimp-top head-space vials (Agilent, Santa Clara, CA). The 1.0 – 5.0 mL reaction mixture contained 20 – 200 μM P450 in 0.1 M phosphate buffer, pH 8.0, saturated with alkane and oxygen. The buffer was saturated with the gaseous alkane through vigorous bubbling for at least 10 minutes. After the addition of the P450 mixture, the crimp-top vial was sealed and stirred moderately at 25 ^{o}C. The reaction was then initiated by the addition of 0.5 mL of a NADPH regeneration system consisting of 10 mg/mL isocitrate, 0.2 mg/mL $NADP^{+}$, and 10 U/mL of isocitrate dehydrogenase. For determination of the TON, the reaction was carried out at 4 ^{o}C over 24 hrs. For initial rate of formation determination, the reaction was initiated by the addition of 500 μM NADPH and was quenched with 200 μL of concentrated H_2SO_4 after an appropriate time depending on the enzyme activity (20 second to 10 minutes) at 25 ^{o}C.

The alcohol product was initially analyzed as alkyl nitrites using a published method (*15*). To a 5.0 mL reaction mixture, 0.3 g of sodium nitrite and 2 mL of hexanes were added along with 1-butanol as an internal standard, and the mixture was cooled on ice. The reaction was initiated with the addition of 0.2 mL concentrated sulfuric acid and quenched by dilution with 5 mL of deionized water after 15 minutes. The reaction mixture was further washed with 20 mL of water in a separatory funnel. The organic phase was collected and analyzed with GC-ECD (see H.1). Alternatively, the alcohol products were also analyzed by GC-FID using Supelco SPB-1 column and 1-pentanol as an internal standard (see H.2). All measurements were carried out in triplicate.

F.2. Alkane hydroxylation with CYP153A6 (Chapter 6)

For alkane reactions with CYP153A6, 0.15 – 0.3 mL reaction mixtures contained 0.5 – 2.0 μM A6 and 0.5 – 2.0 μM FDR and 5.0 – 20.0 μM FDX at a 1:1:10 ratio in 0.1 M phosphate buffer, pH 8.0. In the case of liquid alkanes, a stock solution in ethanol was used to add 25 μM to 10 mM alkane in a final solution containing 2% ethanol. For gaseous alkanes, the reaction was carried out in crimp-top head-space vials pressurized to 20 psi with the alkane. The reaction was initiated with the addition of 1 – 2 mM NADH. For determination of the initial rate, the reactions were quenched with 20 μL of 3.0 M HCl after 5 minutes.

Reactions with liquid alkanes were extracted with 100 – 150 μL of chloroform after the addition of 2-nonanol as an internal standard. The mixture was then vortexed for 30 seconds before centrifuging at 14,000x*g* for 2 minutes. The bottom organic phase was collected and analyzed by GC-FID, see H.2 for details. Reactions with gaseous alkanes were quenched with 20 μL of 3.0 M HCl and neutralized with 75 μL of 1.0 M phosphate buffer, pH 8.0. This acidification and neutralization sequence resulted in the desired precipitation of the enzymes in solution. After centrifuging at 14,000x*g* for 2 minutes to pellet the denatured protein, the clarified solution was analyzed by GC-MS (see H.3 for details).

F.3. Alkane hydroxylation with terminal oxidants (Chapter 6)

For alkane reactions with P450s using terminal oxidants, 0.3 mL reaction mixture contained 50 – 250 μM of highly pure enzyme, as described in C.2, in 0.1 M phosphate buffer, pH 8.0. The liquid alkanes were added from an ethanol stock solution, to yield a final solution containing 2.5 mM alkane and 2% ethanol. For gasesous alkanes, lyophilized protein was added directly to a 10 mL crimp-top headspace vial. The vial was then sealed and flushed with the

alkane substrate for 2 minutes before the addition of 0.27 mL of pre-saturated 0.1 M phosphate buffer, pH 8.0. The reaction was then pressurized to 20 psi with the gaseous alkane. For reactions with PhIO, the reaction was initiated with the addition of 0.03 mL of 5 mM PhIO, solubilized in deionized water by sonication. For reactions with MCPBA, the reaction was initiated with the addition of 5 mM MCPBA in an ethanol or isopropanol stock solution. Reactions with H_2O_2 were similarly initiated with the addition of 5 mM H_2O_2. All reactions were allowed to proceed for 10 minutes at 25 ^{o}C. The work-up and analysis of the alcohol products followed the same protocol as described in Section F.2.

F.4. Iodomethane reaction with CYP153A6 (Chapter 6)

For iodomethane reactions with CYP153A6, 0.15 – 0.3 mL reaction mixtures contained 0.5 μM A6 and 0.5 μM FDR and 5.0 μM FDX at a 1:1:10 ratio in 0.1 M phosphate buffer, pH 8.0. Iodomethane in an ethanol stock solution was used to add 25 μM to 10 mM alkane in a final solution containing 2% ethanol. The reaction was initiated with the addition of 1 mM NADH and quenched after 5 minutes with either 168 mM purpald in 2 M NaOH or 20 μL of 3.0 M HCl. Reactions quenched with purpald were quantified by absorbance at 550 nm after 10 minutes of color development. For reactions quenched with acid, a neutralization step followed with 75 μL of 1.0 M phosphate buffer, pH 8.0, to precipitate the proteins in solution. After centrifuging at 14,000x*g* for 2 minutes, the clarified solution was analyzed by GC-MS (see H.3 for details).

F.5. Metabolite production with P450s (Chapter 7)

Reactions of verapamil, astemizole, and LY294002 with a panel of 120 P450s were conducted in 96-well microtiter plates. For P450 holoenzymes, 60 μL of lysate, generated as described in E.1, were added to 110 μL of 0.1 M phosphate buffer, pH 8.0, and 10 μL of

substrate (5 mM). The reaction was initiated with the addition of 20 μL NADPH (20 mM). For P450 peroxygenases, 50 μL lysate were added along with 100 μL of 0.1 M EPPS buffer, pH 8.2, and 10 μL of substrate (20 mM). The reaction was initiated with the addition of 40 μL of H_2O_2 (5 mM). The plates were shaken briefly after the addition of NADPH or H_2O_2, and incubated for two hours at 25 ^oC. After this time, 200 μL of acetonitrile were added to quench the reaction. The resulting mixture was centrifuged, and the supernatant was used for subsequent HPLC and LC/MS analysis.

G. Absorption Spectra of P450s (Chapter 6)

The absorption spectra over the range from 350 to 500 nm were taken using a 5 μM solution of purified P450 in 0.1 M, pH 8.0, phosphate buffer with 1 cm pathlength cuvette on a Cary 100 UV/VIS spectrophotometer. Liquid substrate in an ethanol stock solution was used to add 10 mM substrate in a final solution containing 2% ethanol. For gaseous substrates, the protocol for alkane hydroxylation with terminal oxidants (see F.3) was used to expose the enzymes to substrates. The differences in absorption spectra at A_{418nm} corresponding to a low-spin ferric state and A_{390nm} corresponding to a high-spin state were used to determine the extent of the induced spin-shift.

H. Gas Chromatography

All peak assignments in our gas chromatographs were determined by comparison to authentic standards. Analyte peaks were quantified using calibrations curves containing at least four standard concentrations. Unless stated otherwise, every analysis was done with a 1 μL sample injection.

H.1. GC-ECD analysis of alkane reactions

Analysis of the ethanol and propanol products as their alkyl nitrite derivatives followed published procedures (*15*), using a Hewlett-Packard 5890 Series II Plus gas chromatograph (Agilant, Santa Clara, CA) with a HP-1 column (30 m length, 0.32 mm ID, 0.25 μm film thickness) connected to an electron capture detector (ECD). A typical temperature program used was: 60 °C injector, 150 °C detector, 35 °C oven for 3 min, 10 °C/min gradient to 60 °C, 25 °C/min gradient to 200 °C, and then 200 °C for 5 min.

H.2. GC-FID analysis of alkane reactions

Analysis of hydroxylation products of small alkanes was performed on a Hewlett Packard 5890 Series II Plus gas chromatograph with a flame ionization detector (FID) and fitted with a HP-7673 autosampler system. Direct analysis of ethane, propane, butane, and pentane, hydroxylation products was performed on a Suplerco SPB-1 column, (60 m length, 0.32 mm ID, 0.25 μm film thickness). A typical temperature program used for separating the alcohol products was 250 °C injector, 300 °C detector, 50 °C oven for 3 minutes, then 10 °C/minute gradient to 200 °C, 25 °C/minute gradient to 250 °C, then 250 °C for 3 minutes.

Analysis of hydroxylation products of hexane and octane was similarly performed as described above using a DB-624 column, (30 m length, 0.53 mm ID, 3.0 μm film thickness) connected to a FID detector. A typical temperature program used for separating the alcohol products was 250 °C injector, 300 °C detector, 40 °C oven for 5 minutes, then 10 °C/minute gradient to 260 °C, then 260 °C for 3 minutes.

H.3. GC-MSD analysis

Analysis of hydroxylation products of methane, ethane, propane, and iodomethane was performed on Hewlett Packard 6890 Series II Plus gas chromatograph with a 5975C mass selective detector (MSD). Direct analysis of the alcohol products in buffer were performed using a DB-624 column, (30 m length, 0.53 mm ID, 3.0 μm film thickness) with a 0.5 μL sample injection. A typical temperature program used for separating the alcohol products consisted of 250 °C injector, 300 °C detector, 45 °C oven for 5 minutes, then 10 °C/minute gradient to 150 °C, then 25 ^{o}C/minute gradient to 250 °C, and a hold for 3 minutes at 250^{o}C.

I. HPLC/ LCMS

Analysis of the reactions of drug compounds verapamil, astemizole, and lead compound LY294002 with a panel of P450s was performed using a Supelco Discovery C18 column (2.1 x 150 mm, 3 μm) on a Waters 2690 Separation module in conjunction with a Waters 996 PDA detector. Clarified supernantant (25 μL) was analyzed with 0.2% formic acid (solvent A) and acetonitrile (solvent B) at the following conditions: 0 – 3 min, A/B 90:10; 3 – 25min, linear gradient to A/B 30:70; 25 – 30 min, linear gradient to A/B 10:90. LCMS and MS/MS spectra were obtained with a ThermoFinnigan LCQ classic (San Jose, CA), using identical conditions to the HPLC method detailed above for the LC conditions. The MS was operated in positive ESI mode, and the MS/MS spectra were acquired for the most abundant ions.

J. Half-Denaturation Temperature (T_{50}) Determination

Samples of purified enzyme (3 – 5 μM) were incubated for 15 minutes at various temperatures from 25 ^{o}C at the bench top to 65 ^{o}C using a PCR thermocycler or a thermoblock. After centrifugation at 14,000xg to pellet the denatured protein fraction, 160 μL of the

supernatant were mixed with 40 μL of 0.4 M sodium hydrosulfite one a microtiterplate for CO-difference spectroscopy, see E.2 for details. The T_{50} values were determined by data fitting to a 2-state denaturation model using Sigmaplot (Systat, San Jose, CA).

K. Whole-Cell Bioconversion with P450 BM3 Variants (Chapters 2 and 5)

Propane bioconversions were carried out at 80 mL and 300 mL scale using temperature controlled 100-mL (Ochs-Labor, Bovenden, Germany) and 1-L fermenter (DasGip, Jülich, Germany), respectively. Freshly transformed *Escherichia coli* Dh5α cells were grown in modified M9 medium supplemented with 0.4% glucose, and 1.5% yeast extract. In addition to the standard salts, the modified M9 contained additional nutrients: calcium pantothenate, p-aminobenzoic acid, p-hydroxybenzoic acid, thiamine, and trace metals (CoCl2, CuSO4, MnCl2, ZnSO4). After reaching an OD_{600} of 1.2, the cells were induced with 0.25 mM IPTG and 0.25 mM δ-aminolevulininc acid, and harvested after 10 – 12 hours of expression. The cells were centrifuged at 5,000 rpm and 4 ^{o}C for 10 minutes.

The resulting pellet was resuspended in nitrogen-free modified M9 medium supplemented with 1% glucose. A gas mixture of propane and air at 1:2 ratio was bubbled through the resting cells at a flow rate of 15 L/hr. A bubbler filled with an equal volume of water as the reaction volume was connected to the reactor's gas outlet to retain any alcohol product removed by the gas flow. At defined time intervals, 1-mL samples of the cell suspension and the bubbler fraction were removed for GC analysis of the alcohol product. The cell suspension samples were centrifuged and filtered prior to GC analysis. In addition, the P450 concentration was also determined by CO-binding difference spectroscopy on cell lysate obtained by sonication of the cell suspension pellet.

L. Whole-cell bioconversion with AlkB and CYP153 (Chapter 4)

Freshly transformed *E. coli* BL21(DE3) cells expressing AlkB variants were precultured in LB medium supplemented with the appropriate antibiotic at 37 ^{o}C and shaken at 250 rpm for 24 hours. LB cultures (12 mL) in 1-L flasks were then inoculated to a staring OD_{600} of 1.0 and incubated at 37 °C, 250 rpm, for 2.5 h. These cultures were then cooled to 25 °C and induced with 0.4 mM dicyclopropylketone (Sigma-Aldrich) after 30 minutes. After 20 hours of protein expression, the cell cultures were centrifuged at 3,300x*g*, 10 min, and 25 ^{o}C. CYP153A6 variants were similarly cultured, with the exception of using modified M9 medium supplemented with 1.5% yeast extract for only 14 hours of expression.

The cell pellets were resuspended in equal volume of either 0.1 phosphate buffer, pH 7.0, for AlkB or nitrogen-free modified M9 medium for CYP153. For bioconversion of liquid alkanes, 250 μL of alkane, 1% glycerol were added to 1-mL cell suspensions in a glass vial. The reaction was capped and shaken at 200 rpm for 1 hour at 25 ^{o}C. The addition of 200 μL of 1 M HCl and 250 μL of hexanes and vigorous vortexing quenched the reaction and extracted the alcohol products to the organic phase. After centrifuging at 14,000x*g* for 5 minutes, the organic layer was collected for GC analysis.

For bioconversion of gaseous alkanes, 80 mL of the cell suspension and 15 μL of antifoam (Sigma-Aldrich) were stirred in a 100-mL fermenter at 25 ^{o}C. The gaseous alkane was bubbled through the reaction mixture with air at 10 L/hr and at a 1:3 ratio. The reaction was initiated with the addition of 1% final concentration of glycerol for AlkB and 20 mM glucose for CYP153A6. Similar to the bioconversion of P450 BM3, at defined time intervals, 1-mL samples of the cell suspension and the bubbler fraction were removed for GC analysis of the alcohol

product. The P450 concentration was also determined by CO-binding difference spectroscopy on cell lysate obtained by sonication.

M. References

1. Schwaneberg, U., Sprauer, A., Schmidt-Dannert, C., and Schmid, R. D. (1999) P450 monooxygenase in biotechnology - I. Single-step, large-scale purification method for cytochrome P450BM-3 by anion-exchange chromatography, *J. Chromatogr. A 848*, 149-159.
2. Omura, T., and Sato, R. (1964) Carbon monoxide-binding pigment of liver microsomes .i. Evidence for its hemoprotein nature, *Journal of Biological Chemistry 239*, 2370-2379.
3. Salazar, O., Cirino, P. C., and Arnold, F. H. (2003) Thermostabilization of a cytochrome P450 peroxygenase, *Chembiochem 4*, 891-893.
4. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) A general-method of in vitro preparation and specific mutagenesis of dna fragments - study of protein and dna interactions, *Nucleic Acids Res. 16*, 7351-7367.
5. Sambrook, J., Fritch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Press, Cold Spring Harbor, N. Y.
6. Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006) Protein stability promotes evolvability, *Proceedings of the National Academy of Sciences of the United States of America 103*, 5869-5874.
7. Camps, M., Naukkarinen, J., Johnson, B. P., and Loeb, L. A. (2003) Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I, *Proceedings of the National Academy of Sciences of the United States of America 100*, 9727-9732.
8. Minak-Bernero, V., Bare, R. E., Haith, C. E., and Grossman, M. J. (2004) Detection of alkanes, alcohols, and aldehydes using bioluminescence, *Biotechnology and Bioengineering 87*, 170-177.
9. Schwaneberg, U., Otey, C., Cirino, P. C., Farinas, E., and Arnold, F. H. (2001) Cost-effective whole-cell assay for laboratory evolution of hydroxylases in Escherichia coli, *J. Biomol. Screen 6*, 111-117.
10. Quesenberry, M. S., and Lee, Y. C. (1996) A rapid formaldehyde assay using purpald reagent: Application under periodation conditions, *Anal. Biochem. 234*, 50-55.
11. Fasan, R., Chen, M. M., Crook, N. C., and Arnold, F. H. (2007) Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties, *Angewandte Chemie-International Edition 46*, 8414-8418.
12. Lageveen, R. G., Huisman, G. W., Preusting, H., Ketelaar, P., Eggink, G., and Witholt, B. (1988) Formation of polyesters by Pseudomonas oleovorans - effect of substrates on formation and composition of poly-(r)-3-hydroxyalkanoates and poly-(r)-3-hydroxyalkenoates, *Applied and Environmental Microbiology 54*, 2924-2932.
13. Miller, J. H. (1972) *Experiments in molecular genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

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

15. Nguyen, H. T. H., Takenaka, N., Bandow, H., and Maeda, Y. (2001) Trace level determination of low-molecular-weight alcohols in aqueous samples based on alkyl nitrite formation and gas chromatography, *Anal. Sci. 17*, 639-643.

Appendix A

Sequence and Activities of Cytochrome P450 BM3 Variants

The gene encoding for cytochrome P450 BM3 (GenBank number J04832) was cloned in the expression vector pCWori (pBM3_WT18-6) downstream of the double *tac* promoter with flanking *BamH*I and *EcoR*I sites. This plasmid was used for all subsequent cloning procedures and protein expression.

Figure A.1 lists the nucleotide sequence of full-length, wild-type cytochrome P450 BM3.

Figure A.2 lists the amino acid sequence of full-length, wild-type cytochrome P450 BM3.

Table A.1 lists the mutations and activities of P450 BM3 variants generated by random mutagenesis using error-prone PCR, site-saturation mutagenesis, combinatorial active site saturation with a reduced set of amino acids, and structure-based computational library design, as detailed in Chapter 3.

Figure A.1: Nucleotide sequence of full-length, wild-type cytochrome P450 BM-3

```
   1 ATGACAATTA AAGAAATGCC TCAGCCAAAA ACGTTTGGAG AGCTTAAAAA
  51 TTTACCGTTA TTAAACACAG ATAAACCGGT TCAAGCTTTG ATGAAAATTG
 101 CGGATGAATT AGGAGAAATC TTTAAATTCG AGGCGCCTGG TCGTGTAACG
 151 CGCTACTTAT CAAGTCAGCG TCTAATTAAA GAAGCATGCG ATGAATCACG
 201 CTTTGATAAA AACTTAAGTC AAGCGCTTAA ATTTGTACGT GATTTTGCAG
 251 GAGACGGGTT ATTTACAAGC TGGACGCATG AAAAAAATTG GAAAAAAGCG
 301 CATAATATCT TACTTCCAAG CTTCAGTCAG CAGGCAATGA AAGGCTATCA
 351 TGCGATGATG GTCGATATCG CCGTGCAGCT TGTTCAAAAG TGGGAGCGTC
 401 TAAATGCAGA TGAGCATATT GAAGTACCGG AAGACATGAC ACGTTTAACG
 451 CTTGATACAA TTGGTCTTTG CGGCTTTAAC TATCGCTTTA ACAGCTTTTA
 501 CCGAGATCAG CCTCATCCAT TTATTACAAG TATGGTCCGT GCACTGGATG
 551 AAGCAATGAA CAAGCTGCAG CGAGCAAATC CAGACGACCC AGCTTATGAT
 601 GAAAACAAGC GCCAGTTTCA AGAAGATATC AAGGTGATGA ACGACCTAGT
 651 AGATAAAATT ATTGCAGATC GCAAAGCAAG CGGTGAACAA AGCGATGATT
 701 TATTAACGCA TATGCTAAAC GGAAAAGATC CAGAAACGGG TGAGCCGCTT
 751 GATGACGAGA ACATTCGCTA TCAAATTATT ACATTCTTAA TTGCGGGACA
 801 CGAAACAACA AGTGGTCTTT TATCATTTGC GCTGTATTTC TTAGTGAAAA
 851 ATCCACATGT ATTACAAAAA GCAGCAGAAG AAGCAGCACG AGTTCTAGTA
 901 GATCCTGTTC CAAGCTACAA ACAAGTCAAA CAGCTTAAAT ATGTCGGCAT
 951 GGTCTTAAAC GAAGCGCTGC GCTTATGGCC AACTGCTCCT GCGTTTTCCC
1001 TATATGCAAA AGAAGATACG GTGCTTGGAG GAGAATATCC TTTAGAAAAA
1051 GGCGACGAAC TAATGGTTCT GATTCCTCAG CTTCACCGTG ATAAAACAAT
1101 TTGGGGAGAC GATGTGGAAG AGTTCCGTCC AGAGCGTTTT GAAAATCCAA
1151 GTGCGATTCC GCAGCATGCG TTTAAACCGT TTGGAAACGG TCAGCGTGCG
1201 TGTATCGGTC AGCAGTTCGC TCTTCATGAA GCAACGCTGG TACTTGGTAT
1251 GATGCTAAAA CACTTTGACT TTGAAGATCA TACAAACTAC GAGCTCGATA
1301 TTAAAGAAAC TTTAACGTTA AAACCTGAAG GCTTTGTGGT AAAAGCAAAA
1351 TCGAAAAAAA TTCCGCTTGG CGGTATTCCT TCACCTAGCA CTGAACAGTC
1401 TGCTAAAAAA GTACGCAAAA AGGCAGAAAA CGCTCATAAT ACGCCGCTGC
1451 TTGTGCTATA CGGTTCAAAT ATGGGAACAG CTGAAGGAAC GGCGCGTGAT
1501 TTAGCAGATA TTGCAATGAG CAAAGGATTT GCACCGCAGG TCGCAACGCT
1551 TGATTCACAC GCCGGAAATC TTCCGCGCGA AGGAGCTGTA TTAATTGTAA
1601 CGGCGTCTTA TAACGGTCAT CCGCCTGATA ACGCAAAGCA ATTTGTCGAC
1651 TGGTTAGACC AAGCGTCTGC TGATGAAGTA AAAGGCGTTC GCTACTCCGT
1701 ATTTGGATGC GGCGATAAAA ACTGGGCTAC TACGTATCAA AAAGTGCCTG
1751 CTTTTATCGA TGAAACGCTT GCCGCTAAAG GGGCAGAAAA CATCGCTGAC
1801 CGCGGTGAAG CAGATGCAAG CGACGACTTT GAAGGCACAT ATGAAGAATG
1851 GCGTGAACAT ATGTGGAGTG ACGTAGCAGC CTACTTTAAC CTCGACATTG
1901 AAAACAGTGA AGATAATAAA TCTACTCTTT CACTTCAATT TGTCGACAGC
1951 GCCGCGGATA TGCCGCTTGC GAAAATGCAC GGTGCGTTTT CAACGAACGT
2001 CGTAGCAAGC AAAGAACTTC AACAGCCAGG CAGTGCACGA AGCACGCGAC
2051 ATCTTGAAAT TGAACTTCCA AAAGAAGCTT CTTATCAAGA AGGAGATCAT
2101 TTAGGTGTTA TTCCTCGCAA CTATGAAGGA ATAGTAAACC GTGTAACAGC
2151 AAGGTTCGGC CTAGATGCAT CACAGCAAAT CCGTCTGGAA GCAGAAGAAG
2201 AAAAATTAGC TCATTTGCCA CTCGCTAAAA CAGTATCCGT AGAAGAGCTT
2251 CTGCAATACG TGGAGCTTCA AGATCCTGTT ACGCGCACGC AGCTTCGCGC
```

```
2301 AATGGCTGCT AAAACGGTCT GCCCGCCGCA TAAAGTAGAG CTTGAAGCCT
2351 TGCTTGAAAA GCAAGCCTAC AAAGAACAAG TGCTGGCAAA ACGTTTAACA
2401 ATGCTTGAAC TGCTTGAAAA ATACCCGGCG TGTGAAATGA AATTCAGCGA
2451 ATTTATCGCC CTTCTGCCAA GCATACGCCC GCGCTATTAC TCGATTTCTT
2501 CATCACCTCG TGTCGATGAA AAACAAGCAA GCATCACGGT CAGCGTTGTC
2551 TCAGGAGAAG CGTGGAGCGG ATATGGAGAA TATAAAGGAA TTGCGTCGAA
2601 CTATCTTGCC GAGCTGCAAG AAGGAGATAC GATTACGTGC TTTATTTCCA
2651 CACCGCAGTC AGAATTTACG CTGCCAAAAG ACCCTGAAAC GCCGCTTATC
2701 ATGGTCGGAC CGGGAACAGG CGTCGCGCCG TTTAGAGGCT TTGTGCAGGC
2751 GCGCAAACAG CTAAAAGAAC AAGGACAGTC ACTTGGAGAA GCACATTTAT
2801 ACTTCGGCTG CCGTTCACCT CATGAAGACT ATCTGTATCA AGAAGAGCTT
2851 GAAAACGCCC AAAGCGAAGG CATCATTACG CTTCATACCG CTTTTTCTCG
2901 CATGCCAAAT CAGCCGAAAA CATACGTTCA GCACGTAATG GAACAAGACG
2951 GCAAGAAATT GATTGAACTT CTTGATCAAG GAGCGCACTT CTATATTTGC
3001 GGAGACGGAA GCCAAATGGC ACCTGCCGTT GAAGCAACGC TTATGAAAAG
3051 CTATGCTGAC GTTCACCAAG TGAGTGAAGC AGACGCTCGC TTATGGCTGC
3101 AGCAGCTAGA AGAAAAAGGC CGATACGCAA AAGACGTGTG GGCTGGG
```

Figure A.2: Amino acid sequence of full-length, wild-type cytochrome P450 BM-3

```
   1 MTIKEMPQPK TFGELKNLPL LNTDKPVQAL MKIADELGEI FKFEAPGRVT
  51 RYLSSQRLIK EACDESRFDK NLSQALKFVR DFAGDGLFTS WTHEKNWKKA
 101 HNILLPSFSQ QAMKGYHAMM VDIAVQLVQK WERLNADEHI EVPEDMTRLT
 151 LDTIGLCGFN YRFNSFYRDQ PHPFITSMVR ALDEAMNKLQ RANPDDPAYD
 201 ENKRQFQEDI KVMNDLVDKI IADRKASGEQ SDDLLTHMLN GKDPETGEPL
 251 DDENIRYQII TFLIAGHETT SGLLSFALYF LVKNPHVLQK AAEEAARVLV
 301 DPVPSYKQVK QLKYVGMVLN EALRLWPTAP AFSLYAKEDT VLGGEYPLEK
 351 GDELMVLIPQ LHRDKTIWGD DVEEFRPERF ENPSAIPQHA FKPFGNGQRA
 401 CIGQQFALHE ATLVLGMMLK HFDFEDHTNY ELDIKETLTL KPEGFVVKAK
 451 SKKIPLGGIP SPSTEQSAKK VRKKAENAHN TPLLVLYGSN MGTAEGTARD
 501 LADIAMSKGF APQVATLDSH AGNLPREGAV LIVTASYNGH PPDNAKQFVD
 551 WLDQASADEV KGVRYSVFGC GDKNWATTYQ KVPAFIDETL AAKGAENIAD
 601 RGEADASDDF EGTYEEWREH MWSDVAAYFN LDIENSEDNK STLSLQFVDS
 651 AADMPLAKMH GAFSTNVVAS KELQQPGSAR STRHLEIELP KEASYQEGDH
 701 LGVIPRNYEG IVNRVTARFG LDASQQIRLE AEEEKLAHLP LAKTVSVEEL
 751 LQYVELQDPV TRTQLRAMAA KTVCPPHKVE LEALLEKQAY KEQVLAKRLT
 801 MLELLEKYPA CEMKFSEFIA LLPSIRPRYY SISSSPRVDE KQASITVSVV
 851 SGEAWSGYGE YKGIASNYLA ELQEGDTITC FISTPQSEFT LPKDPETPLI
 901 MVGPGTGVAP FRGFVQARKQ LKEQGQSLGE AHLYFGCRSP HEDYLYQEEL
 951 ENAQSEGIIT LHTAFSRMPN QPKTYVQHVM EQDGKKLIEL LDQGAHFYIC
1001 GDGSQMAPAV EATLMKSYAD VHQVSEADAR LWLQQLEEKG RYAKDVWAG
```

Table A.1 Sequence and activities of BM3 variants

Variants	Mutations										Other Residues	DME activity [a] ($A_{550\ nm}$)	Propane TON [b]	Propanol formation rate [c] (min^{-1})	Coupling [d] (%)	Ethane TON [b]	T_{50} [e] (°C)
	74	75	78	82	87	181	184	188	328	330							
BM3	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	54.5
V78C	-	-	C	-	-	-	-	-	-	-	-	0.14	200	2.3	4	0	-
V78S	-	-	S	-	-	-	-	-	-	-	-	0.18	210	1.8	6	0	-
V78T	-	-	T	-	-	-	-	-	-	-	-	0.14	370	1.8	9	0	-
A82E	-	-	-	E	-	-	-	-	-	-	-	0.20	330	-	-	0	-
A82Q	-	-	-	Q	-	-	-	-	-	-	-	0.22	120	-	-	0	-
A328I	-	-	-	-	-	-	-	-	I	-	-	0.47	1,500	3.5	3	0	-
A328P	-	-	-	-	-	-	-	-	P	-	-	0.16	1,300	4.7	10	0	-
A328L	-	-	-	-	-	-	-	-	L	-	-	0.50	860	4.0	8	0	-
A328V	-	-	-	-	-	-	-	-	V	-	-	0.17	2,300	7.1	8	0	59.1
A330L	-	-	-	-	-	-	-	-	-	L	-	0.22	1,700	-	-	0	-
A330P	-	-	-	-	-	-	-	-	-	P	-	0.13	590	-	-	0	-
A330V	-	-	-	-	-	-	-	-	-	V	-	0.23	1,700	-	-	0	-
4F9	-	-	-	-	-	-	-	-	-	-	F162L	0.47	3,300	19	18	0	-
8D12	-	-	-	-	-	-	-	-	-	-	I260V	0.20	650	11	14	0	-
4B1	-	-	-	-	-	-	-	-	-	-	E4D,I153V	0.50	2,000	9	15	0	-
1G9	-	-	-	-	-	-	-	-	-	-	T235M	0.46	1,700	-	-	-	-
7F12	-	-	-	-	-	-	-	-	-	-	D232V	0.28	133	-	-	-	-
3F1	-	-	-	-	-	-	-	-	-	-	Q359R	0.27	364	-	-	-	-
2C11	-	-	L		-	-	-	-	L	-	-	0.83	440	10	31	0	61.3
1B4	-	-	-	L	-	-	-	-	L	-	-	0.74	1,500	25	16	140	59.5
4E10	-	-	-	L	-	-	-	-	V	-	-	0.59	4,200	40	44	200	55.4
1B1	-	-	-	W	-	-	-	-	F	-	-	0.15	380	8	2	0	48.4
1E5	-	-	L	W	-	-	-	-	-	-	-	0.23	1,400	27	8	0	54.5
6C8	-	-	L	L	-	-	-	-	L	-	-	0.78	760	6	16	0	52.5
4F7	-	-	L	M	-	-	-	-	L	-	-	0.66	1,700	8	38	0	63.1
7B7	-	-	L	V	-	-	-	-	L	-	-	0.74	1,500	8	32	0	61.9
6E12	-	-	L	W	-	-	-	-	F	-	-	0.39	2,300	28	7	0	50.7

Variants	Mutations										Other Residues	DME activity [a] ($A_{550\ nm}$)	Propane TON [b]	Propanol formation rate [c] (min^{-1})	Coupling [d] (%)	Ethane TON [b]	T_{50} [e] (^oC)
	74	75	78	82	87	181	184	188	328	330							
E 32	W	-	I	L	-	-	V	W	F	W	-	2.61	16,800	236	36	1,200	56.6
E 30	W	-	I	L	-	-	A	-	F	W	-	2.59	16,600	188	41	520	53.0
CC2	W	-	F	L	-	-	V	-	F	W	-	0.52	12,600	-	-	0	-
E 92	W	-	I	L	-	-	V	W	V	L	-	1.11	11,000	138	36	360	52.5
CF3	W	-	I	V	-	-	A	W	V	L	-	1.53	10,900	-	-	0	-
E 35	W	-	I	L	-	-	V	W	V	W	-	1.58	10,400	-	-	530	-
E 66	W	-	I	L	-	-	A	W	F	W	-	2.61	10,300	195	52	890	52.1
CE5	W	-	I	L	-	-	A	-	V	L	-	1.82	9,600	-	-	0	-
E 31	L	-	I	L	-	-	V	W	F	W	-	1.52	9,350	208	68	1,200	56.2
CA3	W	-	F	L	-	-	A	W	V	L	-	1.32	9,130	-	-	0	-
CC5	W	F	I	L	-	-	V	W	V	L	-	1.55	9,120	-	-	0	-
CF1	W	-	I	L	-	W	A	W	V	L	-	1.64	8,850	-	-	390	-
E 77	L	-	I	L	-	-	A	W	V	L	-	1.79	8,840	-	-	430	-
E 78	L	-	I	L	-	-	A	W	F	W	-	1.90	8,820	-	-	870	-
CB8	W	-	F	L	-	-	V	W	V	W	-	1.17	8,050	-	-	300	-
CE4	L	-	I	L	-	-	V	W	V	W	-	1.91	7,640	-	-	670	-
OB6	W	-	F	V	-	-	V	-	V	W	-	1.42	7,330	-	-	0	-
CA1	W	-	F	L	-	-	V	W	V	L	-	1.95	6,720	-	-	520	-
OD3	W	-	I	V	-	-	A	-	F	W	-	1.55	6,510	-	-	0	-
CG3	L	F	I	L	A	-	A	W	F	W	-	1.67	6,060	-	-	0	-
CD3	L	F	I	L	-	-	V	W	F	W	-	2.05	6,050	-	-	0	-
CG2	L	-	F	V	-	-	V	W	F	W	-	1.96	6,030	-	-	540	-
CE7	W	-	I	V	-	-	V	-	F	W	-	1.29	5,600	-	-	0	-
CH1	W	-	I	L	-	W	V	-	F	W	-	1.47	4,260	-	-	0	-
CA4	W	F	I	V	-	W	V	W	F	W	-	1.73	3,590	-	-	0	-
OD2	V	-	-	-	-	F	-	-	F	-	-	0.88	11,600	168	42	660	50.8
OC5	-	-	L	S	-	-	T	W	F	-	-	0.80	10,400	-	-	0	-
OD1	V	-	-	S	-	F	-	-	F	V	-	0.93	10,300	-	-	0	-
OA1	V	-	L	S	-	F	-	-	F	-	-	0.88	9,590	-	-	0	-

Variants	Mutations										Other Residues	DME activity [a] ($A_{550\ nm}$)	Propane TON [b]	Propanol formation rate [c] (min^{-1})	Coupling [d] (%)	Ethane TON [b]	T_{50} [e] (^oC)
	74	75	78	82	87	181	184	188	328	330							
OC6	V	-	-	S	-	F	-	-	F	-	-	1.44	9,430	-	-	0	-
OG9	-	-	-	S	-	-	T	W	F	V	-	0.99	8,230	-	-	0	-
OH3	-	F	L	-	-	-	-	-	-	-	-	2.51	8,160	-	-	0	-
OD7	-	F	-	S	A	-	T	-	F	V	-	1.74	7,900	86	39	720	53.1
OC9	-	-	-	S	-	-	-	W	F	-	-	0.51	7,710	-	-	0	-
OH9	-	-	-	S	-	F	T	W	-	V	-	0.69	6,790	-	-	0	-
OE5	V	F	-	-	-	-	-	W	F	V	-	0.69	6,210	131	41	380	51.9
OA2	V	-	-	-	-	F	-	W	F	-	-	1.31	4,450	-	-	0	-

[a] DME demethylation activity was determined in cell-free extract, corrected for background Purpald ® oxidation.

[b] TON determined as nmol product/nmol enzyme. Alkane reactions contained 25 – 250 nM protein, potassium phosphate buffer saturated with propane, and an NADPH regeneration system containing 100 μM $NADP^+$, 2 U/mL isocitrate dehydrogenase, and 10 mM isocitrate. Standard errors are within 15% of average.

[c] Propanol formation rates were determined over 1 – 10 minutes with 100 nM protein, propane saturated potassium phosphate buffer, and 1 mM NADPH.

[d] Coupling determined by ratio of product formation rate to NADPH consumption rate

[e] T_{50} calculated based on a two-state denaturation model using the percentage of 450 nm CO-binding peak of P450 variants remaining after 15-minute incubations at varying temperatures

Appendix B

C^{orbit} and CRAM Algorithm and Mutation Evaluations

B.1 C^{orbit} algorithm and frequency table of mutations

Under the C^{orbit} design, all amino acids except C, M, and P were allowed at the ten target residues 74, 75, 78, 82, 87, 181, 184, 188, 328, and 330, resulting in a total possible sequence space of 10^{17}. In the first step of the modeling process, the structure of the BM3 with N-palmitoyl glycine bound (pdb: 1JPZ), was subjected to 50 steps of conjugate gradient minimization using the DREIDING force field (*1*). After the minimization, the substrate was removed and the backbone of "substrate-less" structure was used for the computational analysis. In addition to mutating the ten target residues, a shell of 26 residues with side chain atoms within 4 Å of these design positions, 20, 25, 69, 71, 72, 73, 77, 81, 88, 177, 180, 185, 189, 205, 259, 260, 263, 264, 267, 268, 329, 354, 356, 436, 437, and 438, were allowed to change side chain conformation, but not amino acid identity.

The globally optimal sequence was determined using the FASTER algorithm (*2*) with a backbone- independent conformer library with binning level of 1.0 (*3*) and scoring by the energy function previously demonstrated for mutagenesis of GFP (*4*). From this optimal sequence, Monte Carlo sampling with 100 temperature cycles between 150 K and 4,000 K at 10^6 steps per cycle was carried out to sample sequences around the global optimum. The top 20,000 scoring sequences were compiled to generate a frequency table (see table B.1). From this table, the most frequent amino acid sharing a degenerate codon with the wild-type residue was selected as the allowed mutation at each position.

Table B.1: Frequency table for the most stable 20,000 sequences as determined by C^{orbit}

Amino acid [a]	74	75	78	82	87	181	184	188	328	330
ALA	**0.14**	0.00	0.00	**0.27**	0.01	0.00	**0.69**	0.24	**0.01**	**0.01**
ARG	0.00	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.01	0.00
ASN	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.02	0.00	0.00
ASP	0.00	0.00	0.00	0.04	0.00	0.00	0.05	0.04	0.00	0.00
GLN	0.00	0.01	0.00	0.20	0.00	0.00	0.00	0.01	0.01	0.00
GLU	0.01	0.00	0.00	0.13	0.00	0.01	0.00	0.03	0.02	0.00
GLY	0.02	0.00	0.01	0.02	0.00	0.00	0.06	0.02	0.00	0.00
ILE	0.00	0.00	0.36	0.00	0.00	0.00	0.00	0.01	0.00	0.08
LEU	0.01	**0.84**	0.37	0.02	0.00	**0.73**	0.01	**0.44**	0.01	0.31
LYS	0.00	0.01	0.04	0.11	0.00	0.04	0.00	0.01	0.08	0.01
PHE	0.36	0.12	0.00	0.00	**0.99**	0.16	0.00	0.01	0.74	0.43
SER	0.02	0.00	0.00	0.15	0.00	0.00	0.09	0.04	0.00	0.00
THR	0.01	0.00	0.00	0.01	0.00	0.00	0.09	0.02	0.00	0.00
VAL	0.03	0.00	**0.19**	0.00	0.00	0.00	0.00	0.01	0.00	0.03
TRP	0.32	0.00	0.01	0.00	0.00	0.00	0.00	0.11	0.03	0.01
TYR	0.03	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02
His	0.06	0.01	0.01	0.00	0.00	0.06	0.00	0.01	0.08	0.10
MET	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PRO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CYS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Selected amino acid	**A**, V	**L**, F	**V**, L	**A**, S	**F**, A	**L**, F	**A**, T	**L**, W	**A**, F	**A**, V

[a] Amino acid and values shown in bold correspond to wild-type residues, values shown with green highlight correspond to selected mutations.

B.2 CRAM algorithm and energy evaluation of mutations

As a counterpoint to the minimum energy strategy, we decided to test a partially structure-guided design. The principal objective of this latter design was to maximally reduce the binding channel volume through incorporation of the largest possible amino acids. However, we recognized that incorporation of tryptophan at all positions would likely result in a largely non-functional library. Also, requiring a degenerate codon that specified exactly the wild-type amino acid and the largest permitted amino acid was overly restrictive. Therefore to aggressively reduce volume, we relaxed our constraint to always include the wild-type residue.

To prevent the incorporation of residues that would inevitably result in a steric clash, we performed side-chain optimization calculations as follows. First, we mutated the targeted ten active positions (74, 75, 78, 82, 87, 181, 184, 188, 328, and 330) to glycine. Second, we individually mutated each position to all amino acids except C, M, P, R, K, E, and D. Holding all other amino acids in the protein fixed in place, we performed a thorough optimization of the mutated side chain to determine if a van der Waal clash was avoidable. This optimization used the RosettaDesign software (*5*), the all-atom Rosetta force field (*6*), and several non-standard RosettaDesign settings to minimize the occurrence of van der Waal clashes. These include an expanded rotamer library (-ex1, –ex2aro_only, -extrachi_cutoff 18), reduced van der Waal radii (-small_radii), and a non-rotameric energy minimization of the side chain position (-mcmin_trials).

With this resulting map of unavoidable van der Waal clashes, we manually customized a degenerate codon library enriched for wild-type representation and for representation of the largest individually tolerated amino acid. We used amino acid volumes as given by Zamyatin (*7*).

Table B.2: Best case repulsive van der Waal energy as determined by ROSETTA

Amino acid	74	75	78	82	87	181	184	188	328	330
ALA	**115.73**	114.82	115.57	**115.57**	115.9	115.61	**115.9**	115.77	**115.59**	**115.81**
PHE	115.87	92.81	181.1	385.92	**97.92**	144.2	536.66	160.32	91.19	116.46
GLY	115.8	115.56	115.98	115.98	115.66	115.73	115.34	115.89	116.07	115.47
HIS	115.99	115.26	127.99	161.58	115.49	99.65	202.6	101.42	114.96	98.98
ILE	138.8	129.53	98.29	147.43	117.49	144.09	143.7	117.65	235.98	143.34
LEU	125.58	**115.89**	97.78	118.73	115.76	**114.86**	150.74	**97.69**	116.3	117.1
ASN	92.96	115.7	92.17	93.18	89.65	91.03	94.7	115.55	116.05	93.39
GLN	92.45	115.33	115.26	115.77	90.75	114.98	185.13	114.84	94.78	92.58
SER	90.66	116.3	115.58	115.8	116.21	89.83	115.83	89.67	115.87	90.58
THR	115.62	93.73	116.38	97.42	115.67	92.08	116.52	92.32	123.17	116.42
VAL	147.03	36.98	**116.05**	122.11	94.63	121.52	135.11	118.18	131.04	123.42
TRP	94.32	97.06	115.15	354.69	93.83	155.82	983.14	116.54	93.76	115.79
TYR	91.30	90.69	171.88	308.64	94.14	156.26	634.72	142.68	91.00	116.07
Selected amino acid	L, W	**L**, F	I, F	V, L	**F**, A	**L**, W	**A**, V	**L**, W	V, F	L, W

[a] Amino acid and values shown in bold correspond to wild-type residues, values shown with green highlight correspond to selected mutations.

References

1. Mayo, S. L., Olafson, B. D., and Goddard, W. A. (1990) DREIDING - a generic force-field for molecular simulations, *J. Phys. Chem. 94*, 8897-8909.

2. Desmet, J., Spriet, J., and Lasters, I. (2002) Fast and Accurate Side-Chain Topology and Energy Refinement (FASTER) as a new method for protein structure optimization, *Proteins 48*, 31-43.

3. Lassila, J. K., Privett, H. K., Allen, B. D., and Mayo, S. L. (2006) Combinatorial methods for small-molecule placement in computational enzyme design, *Proceedings of the National Academy of Sciences of the United States of America 103*, 16710-16715.

4. Hayes, R. J., Bentzien, J., Ary, M. L., Hwang, M. Y., Jacinto, J. M., Vielmetter, J., Kundu, A., and Dahiyat, B. I. (2002) Combining computational and experimental screening for rapid optimization of protein properties, *Proceedings of the National Academy of Sciences of the United States of America 99*, 15926-15931.

5. Kuhlman, B., and Baker, D. (2000) Native protein sequences are close to optimal for their structures, *Proceedings of the National Academy of Sciences of the United States of America 97*, 10383-10388.

6. Rohl, C. A., Strauss, C. E. M., Misura, K. M. S., and Baker, D. (2004) Protein Structure Prediction Using Rosetta, In *Method Enzymol.* (Ludwig, B., and Michael, L. J., Eds.), Academic Press, 66-93.

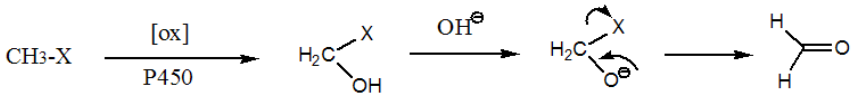
7. Zamyatnin, A. A. (1972) Protein volume in solution, *Progress in Biophysics and Molecular Biology 24*, 107-123.

Appendix C

Candidate High-Throughput Screens for Small Alkane Hydroxylation

C.1 Halomethane Dehalogenation Screen

Halomethanes (CH_3X) are an obvious surrogate for methane and ethane as they share similar molecular size and C-H bond dissociation energies. The P450 catalyzed dehalogenation of halomethanes ultimately yields formaldehyde (see equation (C.1)), which can be quantified calorimetrically with the use of Purpald®, similarly as the DME screen detailed in Chapter 8.E.3.



(C.1)

After assaying BM3 variants for CH_3I, CH_3F, CH_3Br, and CH_3Cl dehalogenation, we determined CH_3Cl to be the most amenable substrate for high-throughput screening as it had the highest activity (20 – 70 TON) and assay signal (A_{550nm} = 0.2 – 0.5). While the assay captured differences in dehalogenation activity of BM3 variants (see Figure C.1), when it was used to evaluate mutant libraries, the variants with improved dehalogenation activity did not frequently exhibit improved ethane hydroxylation activity.

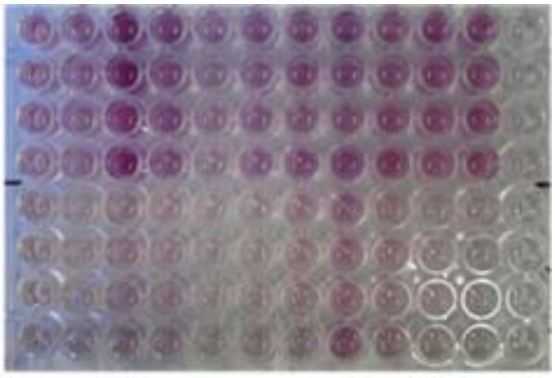
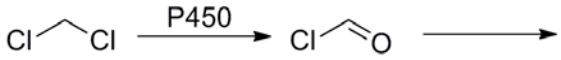


Figure C.1: Colorimetric screen for chloromethane dehalogenation; each column contains eight reactions with the same variant–the top four wells contain reactions with chloromethane, the bottom four wells are control reactions without chloromethane. Column one contains variant 35E11, columns two through eleven contain BM3 variants derived from 35E11, and column 12 contains wild-type BM3.

C.2 Dichloromethane Dehalogenation Screen

In addition to using chloromethane dehalogenation as surrogate for ethane or methane hydroxylation, we also attempted to use dehalogenation of dichloromethane as a high throughput screen. The P450 catalyzed dehalogenation of dichloromethane results in the *in situ* formation of carbon monoxide (see equation (C.2)), which when bound to P450 heme generates the characteristic 450 nm soret peak.



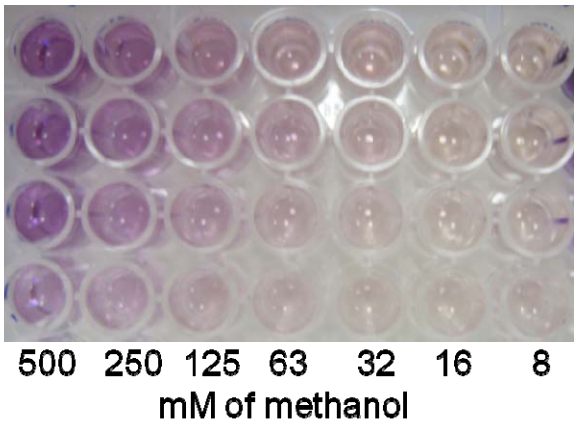
(C.2)

Despite having a weaker C-H bond as compared to chloromethane (93 kcal/mol vs. 98 kcal/mol), the activity of BM3 variants for dehalogenation of dichloromethane was equally poor, 20 – 50 TON. In order to quantify the catalytic release of CO, purified P450 heme domains are needed in excess. The final drawback of this assay is the inhibitory nature of CO of the P450 reaction.

C3. Methanol Oxidation Screen

The small size of methane and ethane significantly limits the molecules that can serve as a suitable surrogate. In addition to halomethanes, methanol is another compound with an intermediate size between methane and ethane. While methanol oxidation activity is non-ideal for a selective methane oxidizing catalyst, it is present in MMOs. Since ethanol is a known P450 substrate, pursuing methanol oxidation as a high-throughput screen was appealing. The P450 catalyzed methanol oxidation resulted in the production of formaldehyde, which can be quantified by Purpald® (see Figure C.2).

(a)



(b)

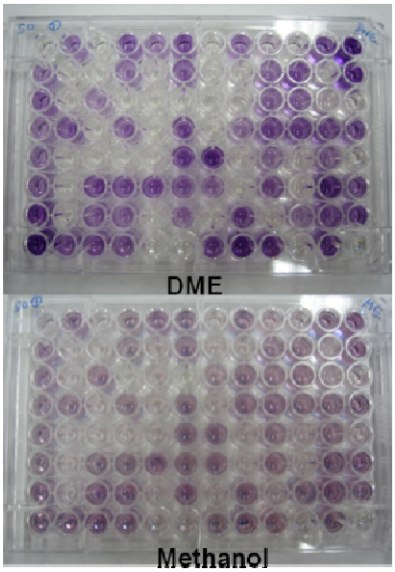


Figure C.2: High-throughput methanol oxidation screen: (a) the methanol hydroxylation activity of 7-7 AB2, a $P450_{PMO}$-derived BM3 variant, at various methanol concentrations; (b) a sample screening plate from 7-7 AB2 random heme library assayed for DME demethylation and methanol oxidation activity

Methanol oxidation activity was not observed in wild-type BM3 and most variants in the $P450_{PMO}$ lineage. $P450_{PMO}$ and its derived variants were able to oxidize methanol with detectable activity with more than 50 mM methanol present in the reaction. From screening of a random heme domain library for both DME demethylation and methanol oxidation it was evident that these two activities were well-correlated (see Figure C.2 (b)). As a result, there was not a significant difference between variants identified from these two assays.

Appendix D

Chapter 6 Supplemental Materials

D.1 GC/MS analysis of terminal oxidant-supported methane reactions

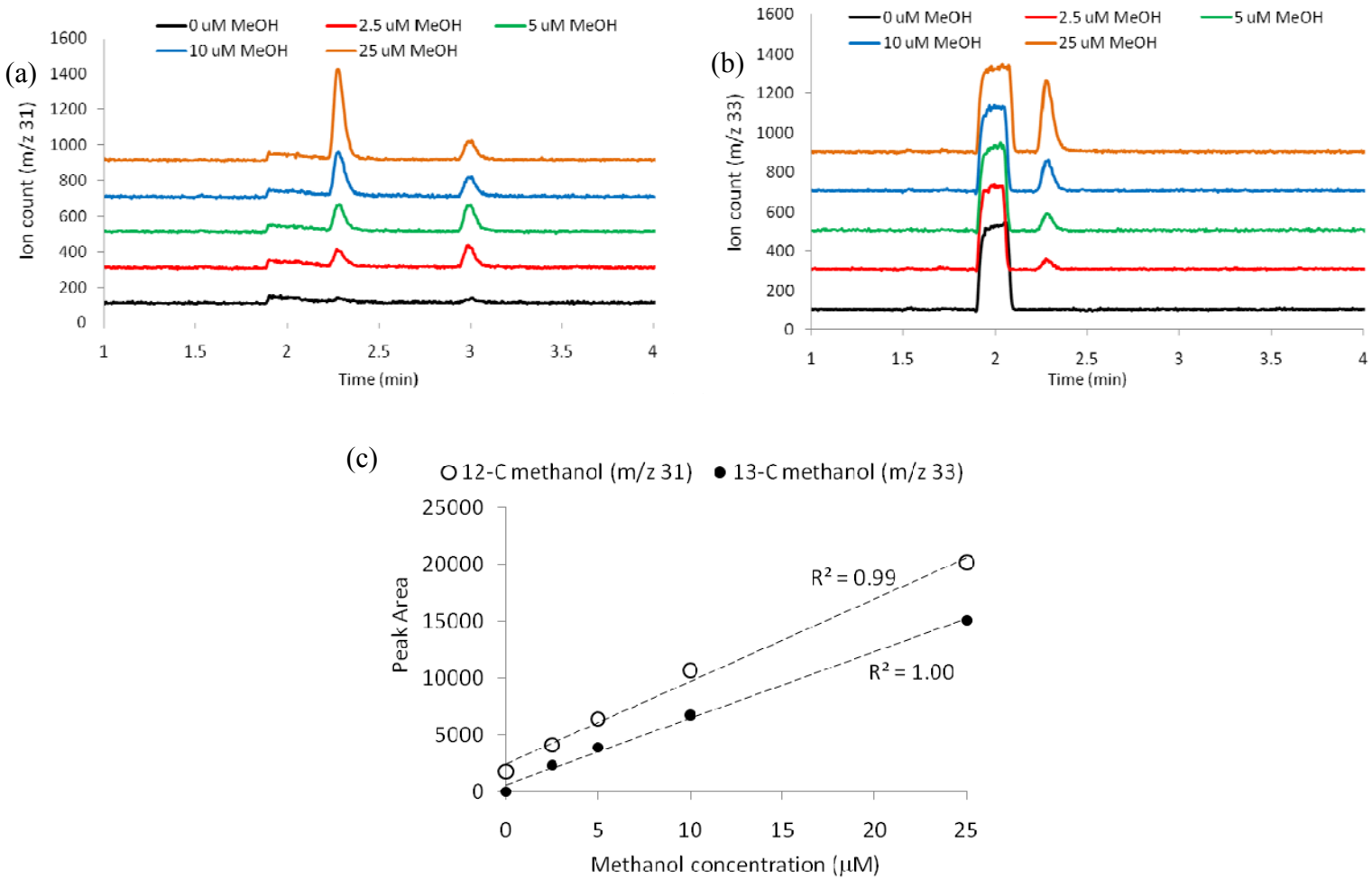


Figure D.1: GC/MS-SIM chromatogram of ^{12}C and ^{13}C methanol calibration standards. (a) ^{12}C methanol standards quantified by m/z 31 ion with peak retention at 2.26 min, shown with a vertical offset. Peak at 3.0 min corresponds to background ethanol. (b) ^{13}C methanol standards quantified by m/z 33 ion. (c) Calibration curves for ^{12}C and ^{13}C methanol used for methanol quantification

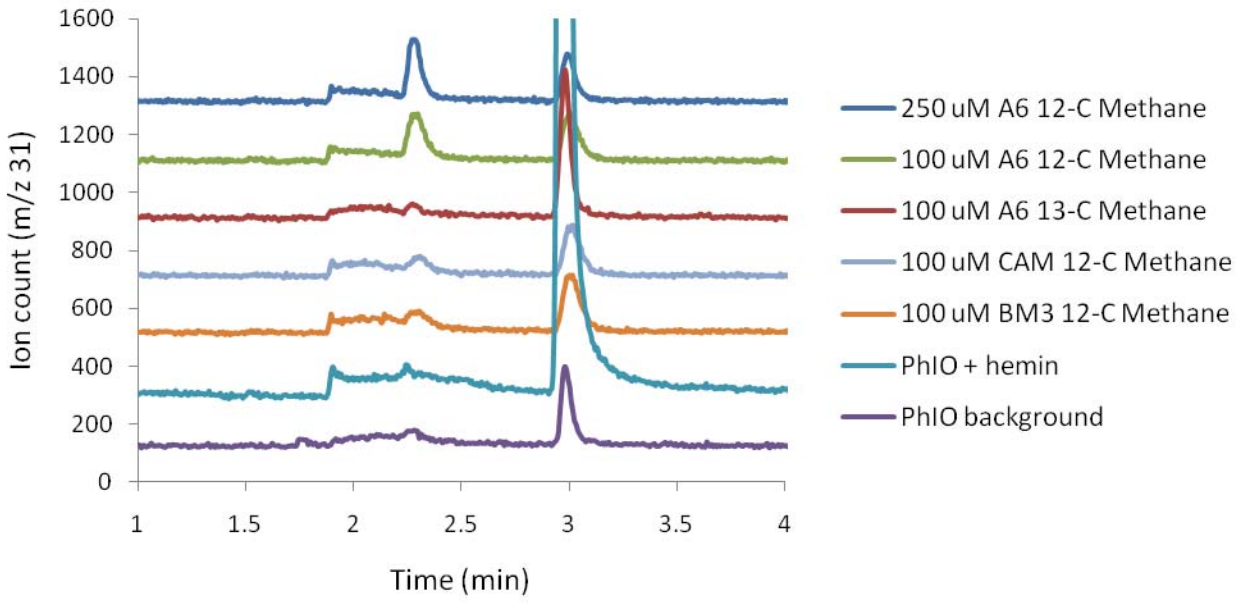


Figure D.2: GC/MS-SIM chromatogram of PhIO-supported ^{12}C-methane reactions with BM3, CAM, and A6 with controls reactions containing PhIO, PhIO, and hemin dissolved in ethanol, and PhIO-supported ^{13}C-methane reaction with A6, shown with a vertical offset. The methanol peak has a retention time of 2.26 min with background ethanol found as a peak at 3.0 min.

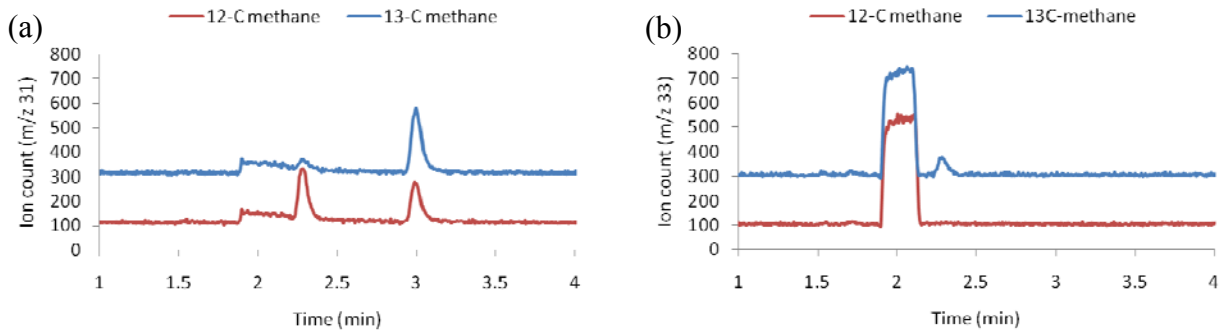


Figure D.3: GC/MS-SIM chromatogram of PhIO-supported A6 methane reactions with ^{12}C-and ^{13}C-methane, shown with a vertical offset: (a) m/z 31 ion chromatogram with methanol peak with retention of 2.26 min and background ethanol with retention of 3.0 min; (b) m/z 33 ion chromatogram with methanol peak with retention of 2.26 min

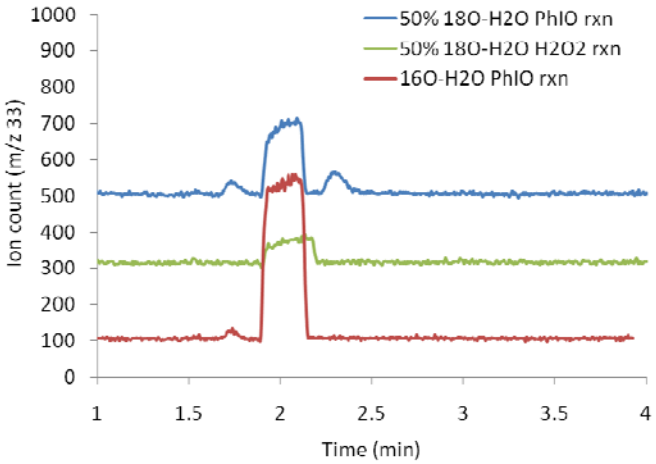


Figure D.4: GC/MS-SIM (m/z 33) chromatogram of terminal oxidant-supported A6 methane reactions with ^{16}O-and ^{18}O-water, shown with a vertical offset. Methanol peak with retention of 2.26 min

D.2 Quantification of fdrA6 and fdxA6 and determination of optimal reductase component ratios

As a type I P450, A6 differs from BM3 in that its reductase components fdrA6 and fdxA6 are separate proteins rather than being fused on a single polypeptide chain. We sub-cloned fdrA6 and fdxA6 individually into the pET-22b(+) expression vector and expressed the enzymes in *E. coli* BL21(DE3) transformed with the resulting plasmids. Following literature protocols established for the purification of putidaredoxin and putidaredoxin reductase of $P450_{cam}$ (*1*), we purified fdrA6 and fdxA6 with the characteristic UV-Vis spectra for the FAD and [Fe_2-S_2] cofactors (see Figure E.5). The concentrations of fdrA6 and fdxA6 were determined using the extinction coefficients, ε_{455nm} = 10.4 mM^{-1} cm^{-1} and ε_{412nm} = 11.0 mM^{-1} cm^{-1} (*1*), respectively.

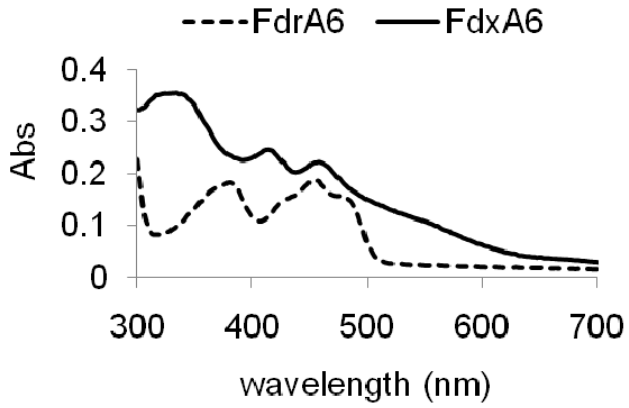


Figure D.5: UV/Vis spectrum of purified FdrA6 and FdxA6

To determine the optimum ratio of reductase components, the A6 NADH consumption rate was determined in the presence and absence of octane at various ratios of the reductase components (see Figure E.6). The substrate-free NADH consumption rate was found to be 3 – 5 min^{-1} at moderate ratios of fdxA6:A6, 2.5 – 10:1, independent of fdrA6. At higher ratios of fdxA6:A6, the substrate-free NADH consumption increases linearly with increasing fdxA6 reaching a rate of 15 min^{-1} at 80:1 fdxA6:A6. In the presence of octane, the rate of NADH consumption asymptotes above a ratio of 10:1 fdxA6:A6 and 1:1fdrA6:A6 at ca. 80 min^{-1}. Based

on these results, a ratio of 1:1:10 for A6:fdrA6:fdxA6 was chosen for all NADH/O_2-supported reactions with A6.

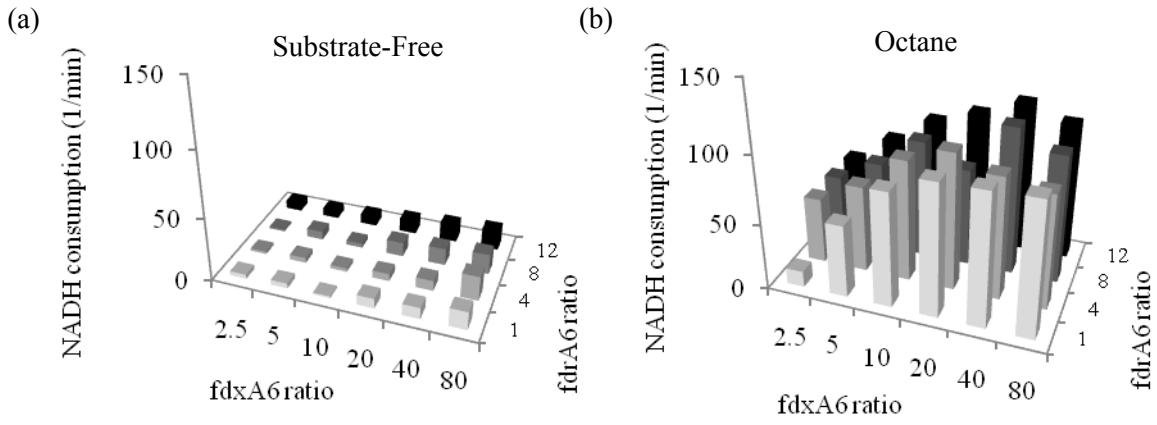


Figure D.6: (a) Substrate-free NADH consumption in the presence of various ratios of fdxA6 and fdrA6 as determined by absorption at 340 nm. (b) NADH consumption in the presence of octane with various ratios of fdxA6 and fdrA6

D.3 Determination of Michaelis-Menten kinetics parameters

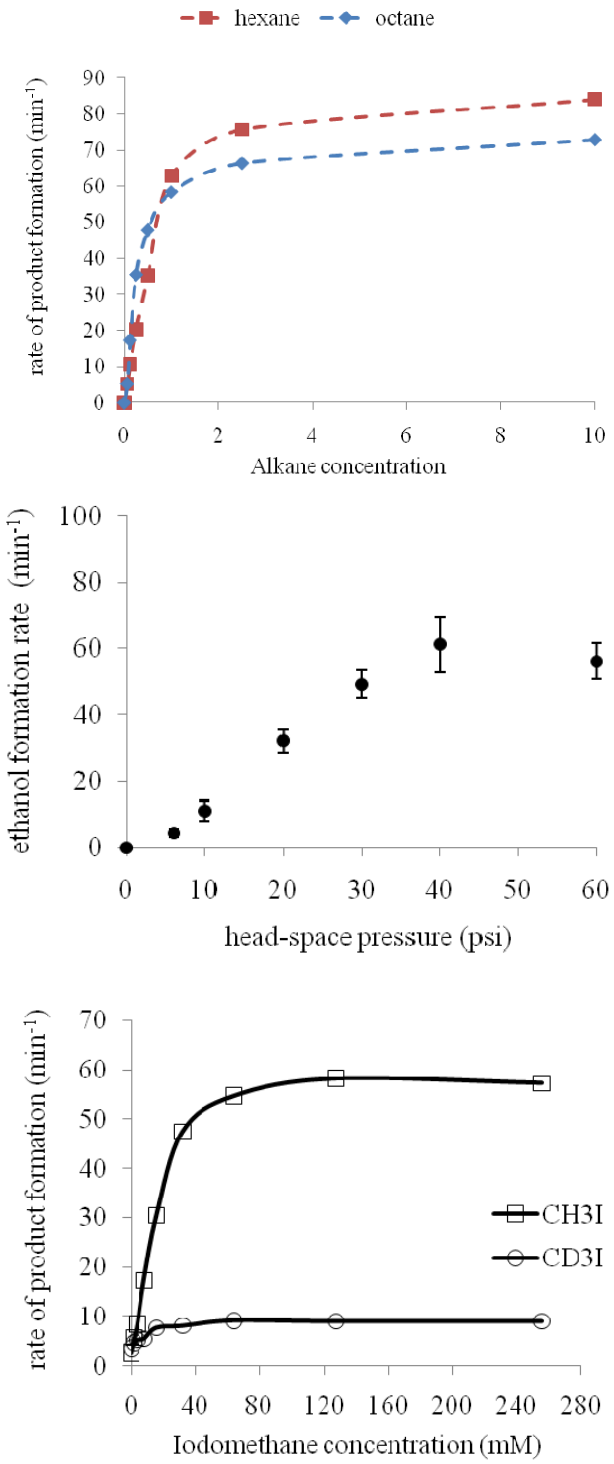


Figure D.7: Plots of initial product formation rates at various substrate concentrations. (a) Octane and hexane reactions; (b) Ethane reactions; (c) Iodomethane and d_3-Iodomethane reactions; (see Chapter 8.F for reaction details)

D.4 UV/Vis spectroscopy

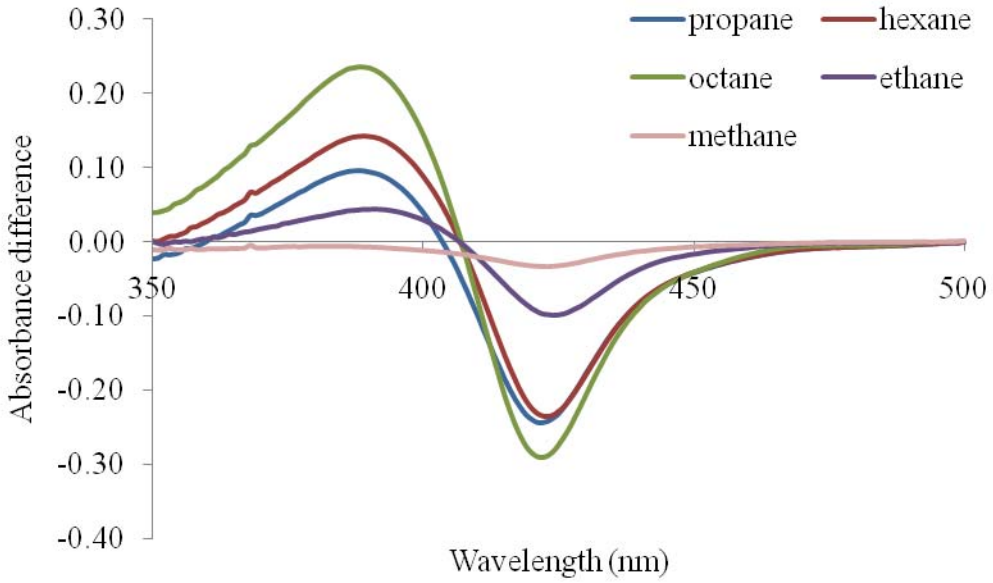


Figure D.8: Difference UV/Vis spectra of CYP153A6 in the presence of different alkanes. Hexane and octane were present at 1 mM in 1% ethanol and referenced against an A6 solution containing 1% ethanol. Gaseous alkanes were present at 60 -psi headspace pressure and referenced against an A6 solution.

Reference

1. Gusalus, I. C., and Wagner, G. C. (1978) *Methods Enzymol. 52*, 166-188.

Appendix E

Variant Selection for Production of Drug Metabolites and Diversified Lead Compounds

Table E.1: Identity of engineered P450 BM3 variant panel: enzyme family, name, sequence, number of mutations from closest wild-type parent, and required oxidant

Entry	Family	Name[a]	Sequence	Mutations vs. WT	Selection Criteria	Preferred Oxidant
1	WT	P450 BM3	WT	0	fatty acid	O_2
2	WT	CYP102A2	CYP102A2	0	fatty acid	O_2
3	WT	CYP102A3	CYP102A3	0	fatty acid	O_2
4	WT	A1	CYP102A1 Heme Domain Only	0	none	H_2O_2
5	WT	A2	CYP102A2 Heme Domain Only	0	none	H_2O_2
6	WT	WT F87A	WT F87A	1	none	O_2
7	Propranolol	9C1	**21B3** I102T A145V L324I I366V E442K	14	Propranolol	O_2
8	Propranolol	D6H10	**9C1** L75H V78E A82P	17	Propranolol	O_2
9	Propranolol	DE10	**9C1** A74V A82L A87G	16	Propranolol	H_2O_2
10	Propranolol	2C11	**DE10** K24R R47H	18	Propranolol	H_2O_2
11	Chimera	11113311	chimera	35	none	H_2O_2
12	Chimera	12112333	chimera	96	none	H_2O_2
13	Chimera	21112233	chimera	98	none	H_2O_2
14	Chimera	21112331	chimera	85	none	H_2O_2
15	Chimera	21112333	chimera	89	none	H_2O_2
16	Chimera	21113312	chimera	97	none	H_2O_2
17	Chimera	21113333	chimera	75	none	H_2O_2
18	Chimera	21212233	chimera	89	none	H_2O_2
19	Chimera	21212333	chimera	87	none	H_2O_2
20	Chimera	21311231	chimera	81	none	H_2O_2
21	Chimera	21311233	chimera	97	none	H_2O_2
22	Chimera	21311311	chimera	63	none	H_2O_2
23	Chimera	21311313	chimera	95	none	H_2O_2
24	Chimera	21311331	chimera	81	none	H_2O_2
25	Chimera	21311333	chimera	81	none	H_2O_2
26	Chimera	21312133	chimera	100	none	H_2O_2
27	Chimera	21312211	chimera	76	none	H_2O_2
28	Chimera	21312213	chimera	99	none	H_2O_2
29	Chimera	21312231	chimera	94	none	H_2O_2
30	Chimera	21312233	chimera	96	none	H_2O_2
31	Chimera	21312311	chimera	76	none	H_2O_2
32	Chimera	21312313	chimera	98	none	H_2O_2
33	Chimera	21312331	chimera	94	none	H_2O_2
34	Chimera	21312332	chimera	83	none	H_2O_2
35	Chimera	21312333	chimera	80	none	H_2O_2
36	Chimera	21313111	chimera	58	none	H_2O_2

37	Chimera	21313231	chimera	96	none	H_2O_2
38	Chimera	21313233	chimera	82	none	H_2O_2
39	Chimera	21313311	chimera	78	none	H_2O_2
40	Chimera	21313313	chimera	84	none	H_2O_2
41	Chimera	21313331	chimera	96	none	H_2O_2
42	Chimera	21313333	chimera	66	none	H_2O_2
43	Chimera	21333233	chimera	61	none	H_2O_2
44	Chimera	22112233	chimera	81	none	H_2O_2
45	Chimera	22112333	chimera	93	none	H_2O_2
46	Chimera	22212333	chimera	88	none	H_2O_2
47	Chimera	22223132	chimera	55	none	H_2O_2
48	Chimera	22311233	chimera	92	none	H_2O_2
49	Chimera	22311331	chimera	98	none	H_2O_2
50	Chimera	22311333	chimera	85	none	H_2O_2
51	Chimera	22312231	chimera	78	none	H_2O_2
52	Chimera	22312233	chimera	79	none	H_2O_2
53	Chimera	22312331	chimera	94	none	H_2O_2
54	Chimera	22312333	chimera	84	none	H_2O_2
55	Chimera	22313231	chimera	92	none	H_2O_2
56	Chimera	22313233	chimera	86	none	H_2O_2
57	Chimera	22313331	chimera	102	none	H_2O_2
58	Chimera	22313333	chimera	70	none	H_2O_2
59	Chimera	23132233	chimera	70	none	H_2O_2
60	Chimera	32312231	chimera	101	none	H_2O_2
61	Chimera	32312333	chimera	53	none	H_2O_2
62	Chimera	32313233	chimera	55	none	H_2O_2
63	Chimera	11113311-R1	chimera	35	none	O_2
64	Chimera	12112333-R1	chimera	96	none	O_2
65	Chimera	21113312-R1	chimera	97	none	O_2
66	Chimera	21113312-R2	chimera	97	none	O_2
67	Chimera	21311231-R1	chimera	81	none	O_2
68	Chimera	21311233-R1	chimera	97	none	O_2
69	Chimera	21313311-R1	chimera	78	none	O_2
70	Chimera	21333233-R2	chimera	61	none	O_2
71	Chimera	22132231-R1	chimera	77	none	O_2
72	Chimera	22223132-R1	chimera	55	none	O_2
73	Chimera	22312333-R1	chimera	84	none	O_2
74	Chimera	22313233-R1	chimera	86	none	O_2
75	Chimera	23132233-R1	chimera	70	none	O_2
76	Chimera	23132233-R2	chimera	70	none	O_2
77	Chimera	32312231-R1	chimera	101	none	O_2
78	Chimera	32312333-R1	chimera	53	none	O_2
79	Chimera	32313233-R1	chimera	55	none	O_2

80	Propane	139-3	**WT** V78A H138Y T175I V178I A184V H236Q E252G R255S A290V A295T L353V	10	octane	O_2
81	Propane	J	**139-3** Y138H I178V F205C S226R A290V	10	propane	O_2
82	Propane	9-10A	**J** R47C K94I P142S	13	propane	O_2
83	Propane	9-10A A328F	9-10A A328F	14	propane	O_2
84	Propane	9-10A A328M	9-10A A328M	14	propane	O_2
85	Propane	9-10A A328V	9-10A A328V	14	propane	O_2
86	Propane	9-10A A78F	9-10A A78F	13	propane	O_2
87	Propane	9-10A A78S	9-10A A78S	13	propane	O_2
88	Propane	9-10A A78T	9-10A A78T	13	propane	O_2
89	Propane	9-10A A82C	9-10A A82C	14	propane	O_2
90	Propane	9-10A A82F	9-10A A82F	14	propane	O_2
91	Propane	9-10A A82G	9-10A A82G	14	propane	O_2
92	Propane	9-10A A82I	9-10A A82I	14	propane	O_2
93	Propane	9-10A A82L	9-10A A82L	14	propane	O_2
94	Propane	9-10A A82S	9-10A A82S	14	propane	O_2
95	Propane	9-10A A82T	9-10A A82T	14	propane	O_2
96	Propane	9-10A F87A	9-10A F87A	14	propane	O_2
97	Propane	9-10A F87I	9-10A F87I	14	propane	O_2
98	Propane	9-10A F87L	9-10A F87L	14	propane	O_2
99	Propane	9-10A F87V	9-10A F87V	14	propane	O_2
100	Propane	9-10A L75I	9-10A L75I	14	propane	O_2
101	Propane	9-10A L75W	9-10A L75W	14	propane	O_2
102	Propane	9-10A T260L	9-10A T260L	14	propane	O_2
103	Propane	9-10A T260N	9-10A T260N	14	propane	O_2
104	Propane	9-10A T260S	9-10A T260S	14	propane	O_2
105	Propane	9-10A T88L	9-10A T88L	14	propane	O_2
106	Propane	9-10A A82L A328V	9-10A-A82L-A328V	15	propane	O_2
107	Propane	9-10A A82F A328V	9-10A A82F A328V	15	propane	O_2
108	Propane	9-10A F87V A328L	9-10A F87V A328L	15	propane	O_2
109	Propane	9-10A A82G F87V A328V	9-10A A82G F87V A328V	16	propane	O_2
110	Propane	9-10A A78T A328V	9-10A A78T A328V	14	propane	O_2
111	Propane	9-10A A78T A82G F87V A328L	9-10A A78T A82G F87V A328L	16	propane	O_2
112	Propane	9-10A A82F A328F	9-10A A82F A328F	15	propane	O_2

113	Propane	9-10A A78F A82G A328F	9-10A A78F A82G A328F	14	propane	O_2
114	Propane	9-10A A78F A82S A328L	9-10A A78F A82S A328L	14	propane	O_2
115	Propane	9-10A A78F A82G A328V	9-10A A78F A82G A328V	14	propane	O_2
116	Propane	9-10A A78T A82G F87V A328L	9-10A A78T A82G F87V A328L	16	propane	O_2
117	Propane	9-10A A82G F87L A328L	9-10A A82G F87L A328L	16	propane	O_2
118	Propane	9-10A A78F A82S A328F	9-10A A78F A82S A328F	15	propane	O_2
119	Propane	9-10A A78F A82G A328L	9-10A A78F A82G A328L	15	propane	O_2
120	Propane	9-10A A78T A82G A328L	9-10A A78T A82G A328L	15	propane	O_2

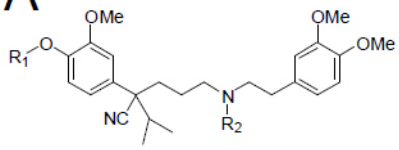
[a]Underlined variants contain wild-type sequences, variants in italics were selected for propranolol activity, variants in bold are chimeras, variants in normal type were selected for activity on alkanes. **21B3** contains the following mutations relative to wild-type: I58V, H100R, F107L, A135S, M145V, N239H, S274T, K434E and V446I. Chimeras are written according to fragment composition: 32313233-R1, for example, represents a protein which inherits the first fragment from parent CYP102A3, the second from CYP102A2, the third from CYP102A3, and so on. The specific amino acid sequence of each block is contained in Table E.2. R1 connotes the presence of the reductase domain from parent A1, indicating that this chimera is a monooxygenase.

Table E.2 Amino acid sequence of blocks 1 – 8 of the cytochrome P450 chimeras

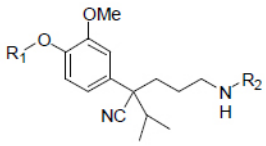
Block	Parent	Sequence
1	A1	-TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDE
1	A2	KETSPIPQPKTFGPLGNLPLIDKDKPTLSLIKLAEEQGPIFQIHTPAGTTIVVSGHELVKEVCDE
1	A3	KQASAIPQPKTYGPLKNLPHLEKEQLSQSLWRIADELGPIFRFDFPGVSSVFVSGHNLVAEVCDE
2	A1	SRFDKNLSQALKFVRDFAGDGLATSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDI
2	A2	ERFDKSIEGALEKVRAFSGDGLATSWTHEPNWRKAHNILMPTFSQRAMKDYHEKMVDI
2	A3	KRFDKNLGKGLQKVREFGGDGLATSWTHEPNWQKAHRILLPSFSQKAMKGYHSMMLDI
3	A1	AVQLVQKWERLNADEHIEVPEDMTRLTLDTIGLCGFNYRFNSFY
3	A2	AVQLIQKWARLNPNEAVDVPGDMTRLTLDTIGLCGFNYRFNSYY
3	A3	ATQLIQKWSRLNPNEEIDVADDMTRLTLDTIGLCGFNYRFNSFY
4	A1	RDQPHPFITSMVRALDEAMNKLQRANPDDPAYDENKRQFQEDIKVMNDLV
4	A2	RETPHPFINSMVRALDEAMHQMQRLDVQDKLMVRTKRQFRYDIQTMFSLV
4	A3	RDSQHPFITSMLRALKEAMNQSKRLGLQDKMMVKTKLQFQKDIEVMNSLV
5	A1	DKIIADRKASGEQ-SDDLLTHMLNGKDPETGEPLDDENIRYQIITFLIAGHET
5	A2	DSIIAERRANGDQDEKDLLARMLNVEDPETGEKLDDENIRFQIITFLIAGHET
5	A3	DRMIAERKANPDENIKDLLSLMLYAKDPVTGETLDDENIRYQIITFLIAGHET
6	A1	TSGLLSFALYFLVKNPHVLQKAAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTAA
6	A2	TSGLLSFATYFLLKHPDKLKKAYEEVDRVLTDAAPTYKQVLELTYIRMILNESLRLWPTA
6	A3	TSGLLSFAIYCLLTHPEKLKKAQEEADRVLTDDTPEYKQIQQLKYIRMVLNETLRLYPTA
7	A1	PAFSLYAKEDTVLGGEYPLEKGDELMVLIPQLHRDKTIWGDDVEEFRPERFENPSAIPQHAFKPF GNGQRACIGQQ
7	A2	PAFSLYPKEDTVIGGKFPITTNDRISVLIPQLHRDRDAWGKDAEEFRPERFEHQDQVPHHAYKPF GNGQRACIGMQ
7	A3	PAFSLYAKEDTVLGGEYPISKGQPVTVLIPKLHRDQNAWGPDAEDFRPERFEDPSSIPHHAYKPF GNGQRACIGMQ
8	A1	FALHEATLVLGMMLKHFDFEDHTNYELDIKETLTLKPEGFVVKAKSKKIPLGGIPSPST
8	A2	FALHEATLVLGMILKYFTLIDHENYELDIKQTLTLKPGDFHISVQSRHQEAIHADVQAAE
8	A3	FALQEATMVLGLVLKHFELINHTGYELKIKEALTIKPDDFKITVKPRKTAAINVQRKEQA

Table E.3 Complete list of active enzymes and their metabolite distributions with verapamil

A

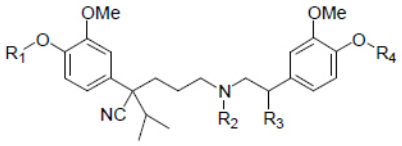


	Mw	R_1	R_2
Verapamil	454	Me	Me
1	426	H	H
2	440	H	Me
3	440	Me	H



	Mw	R_1	R_2
4	276	H	Me
5	276	Me	H
6	290	Me	Me

B



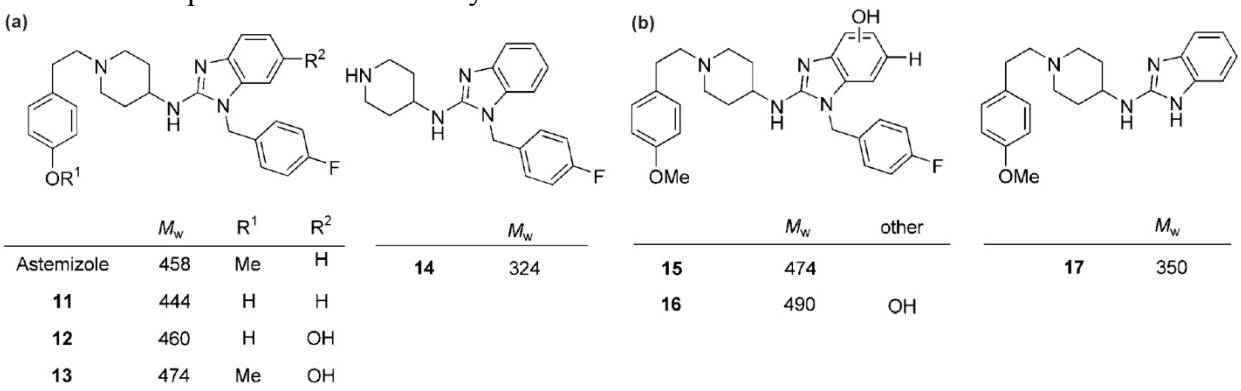
	Mw	R_1	R_2	R_3	R_4	other
7	440	Me	Me	H	H	
8	456	Me	Me	OH	H	
9	470	Me	Me	OH	Me	
10	470	Me	Me	H	Me	OH

Entry	Variant [a]	% Conversion	% 1	% 2	% 3	% 4	% 5	% 6	% 7	% 8	% 9	% 10
1	*2C11*	25		8	28			20	8		8	28
2	*9C1*	31	3		29	3		6	13	3	10	13
3	*D6H10*	78			26	8		31			3	24
4	*DE10*	20						15	35		10	15
5	**11113311**	25			48			28	8			16
6	**21113312**	17		6	12			12	41	6	6	12
7	**21113333**	5						20	40	20		20
8	**21312332**	6			17			33	17			33
9	**21313111**	6						67				33
10	**21313311**	12			33			33	8			25
11	**21313331**	5			20			20	20	20		20
12	**22313231**	43		47				33	5			16
13	**22313233**	12	8					8	17	8	17	8
14	**22313333**	34		41			15	21	9			15
15	**32313233**	24			50			21	13			17
16	**23132233-R1**	15			27			73				
17	**32312231-R1**	23			30			43				26
18	**32312333-R1**	44		7	7				36			
19	**32313233-R1**	25			40			44				16
20	139-3	14		14	14			14	57			
21	9-10A	27		7	7			7	63	7		
22	L75I	34		6	6			6	74	3		
23	A78F	30		10					83			
24	A78S	20			25			35	35			5
25	A78T	25		8	8			8	64	4		
26	A82C	16			6			6	81			
27	A82F	30			10			3	83			
28	A82G	13			31			46	23			
29	A82I	33		3	6			3	79	3	3	
30	A82L	51						2	94			
31	A82S	32		3	3			6	88			
32	A82T	28		4	7			4	86			

Entry	Variant [a]	% Conversion	% 1	% 2	% 3	% 4	% 5	% 6	% 7	% 8	% 9	% 10
33	F87A	12			8			17	58			
34	F87I	13		8				8	69			
35	F87L	49						6	94			
36	F87V	24	4	4				8	63			
37	T88L	14			21			43	21			14
38	T260S	34			12			29	38	9		12
39	A328M	36	3	3	3			6	75		6	
40	1-12G	19			16			21	53			
41	7-11D	21			29			29	14			
42	12-10C	34			6			6	79	3		29
43	41-5B	18			11			6	67			

[a]Variants in italics were selected for propranolol activity (*1*), variants in bold are chimeras (*2*), variants in normal type were selected for activity on alkanes (*3 – 4*).

Table E.4 Complete list of active enzymes and their metabolite distributions with astemizole



	M_w	R^1	R^2
Astemizole	458	Me	H
11	444	H	H
12	460	H	OH
13	474	Me	OH

	M_w
14	324

	M_w	other
15	474	
16	490	OH

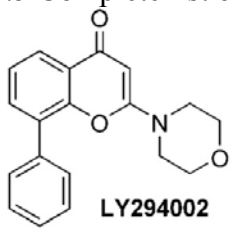
	M_w
17	350

Entry	Variant [a]	% Conversion	% 11	% 12	% 13	% 14	% 15	% 16	% 17
1	*2C11*	4				100			
2	*9C1*	2				100			
3	*DE10*	9				56			
4	**21113312**	6	33			67			
5	**21313111**	4	25			75			
6	**21313311**	10	20			40			40
7	**22313231**	9	22			44			33
8	**22313233**	7	14			57			29
9	**22313333**	9	22			56			22
10	**32312333**	3				100			
11	**32313233**	11				45	27		27
12	**21313111-R1**	11	27		27	18	27		
13	**22313233-R1**	6	33			33			33
14	**32312333-R1**	78				4	67		9
15	**32313233-R1**	16				13	69		
16	139-3	6			67		33		
17	J	23	26	4	65				

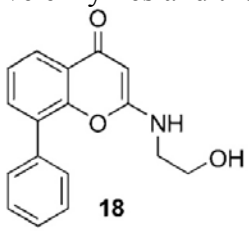
Entry	Variant [a]	% Conversion	% 11	% 12	% 13	% 14	% 15	% 16	% 17
18	9-10A	50	14	10			70	6	
19	L75I	15	33		47				
20	A78F	21	48		38			14	
21	A78S	15	13	7			80		
22	A78T	27	37	7	44			4	
23	A82C	24	33		67				
24	A82F	15	60						
25	A82G	35	29	6	63			3	
26	A82I	32	31	6	63				
27	A82L	49	24				61	4	
28	A82S	30	20	7	70			3	
29	A82T	34	24	12	59			6	
30	F87A	31			23	16	45		
31	F87I	50			48		36	4	
32	F87L	36		3	88			6	
33	F87V	26			38		31		
34	T88L	28			75		25		
35	T260L	2			100				
36	T260N	40		25	75				
37	T260S	25			32	8	44		
38	11-8E	2	100						
39	12-10C	26	38			12	46		
40	23-11B	2	100						
41	41-5B	5	100						
42	49-9B	5	60			40			

[a]Variants in italics were selected for propranolol activity (*1*), variants in bold are chimeras (*2*), variants in normal type were selected for activity on alkanes (*3 – 4*).

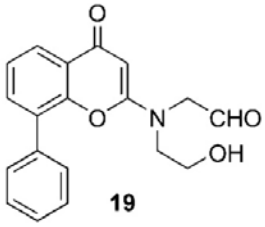
Table E.5 Complete list of active enzymes and their metabolite distributions with LY294002



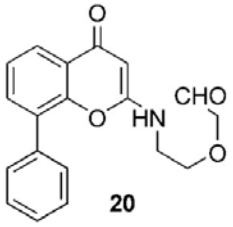
LY294002



18



19



20

M_w: 307 281 323 323

Entry	Variant [a]	% Conversion	% 18	% 19	% 20	% metabolite A [b]
1	**12112333-R1**	8				100
2	**21113312-R1**	8				100
3	**21313111-R1**	7				100
4	**21333233-R1**	4				100
5	**22132231-R1**	7				100
6	**32312231-R1**	6				100
7	**32313233-R1**	8				100
8	J	9				100
9	L75I	10		50		
10	A78S	7		71		
11	A82S	2		100		
12	F87V	6			100	
13	T260S	11		9		91
14	A328F	7				100
15	7-11D	7				100
16	29-10E	12				83
17	68-8F	9		11		89
18	77-9H	11				91

[a] Variants in italics were selected for propranolol activity (*1*), variants in bold are chimeras (*2*), variants in normal type were selected for activity on alkanes (*3 – 4*).
[b] Metabolite of unknown structure, Mw = 238.

1. Otey, C. R., Bandara, G., Lalonde, J., Takahashi, K., and Arnold, F. H. (2006) Preparation of human metabolites of propranolol using laboratory-evolved bacterial cytochromes P450, *Biotechnology and Bioengineering 93*, 494-499.
2. Otey, C. R., Landwehr, M., Endelman, J. B., Hiraga, K., Bloom, J. D., and Arnold, F. H. (2006) Structure-guided recombination creates an artificial family of cytochromes P450, *PLoS. Biol. 4*, 789-798.
3. Meinhold, P., Peters, M. W., Hartwick, A., Hernandez, A. R., and Arnold, F. H. (2006) Engineering cytochrome P450BM3 for terminal alkane hydroxylation, *Advanced Synthesis & Catalysis 348*, 763-772.
4. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc. 125*, 13442-13450.