LABORATORY EVOLUTION OF CYTOCHROME P450 PEROXYGENASE ACTIVITY

Thesis by

Patrick C. Cirino

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2004

(Defended June 2, 2003)

ACKNOWLEDGEMENTS

My tenure at Caltech has been a wonderful and rewarding experience. I leave this chapter of my life feeling enlightened because of what I have learned, and charmed because of the people that I have had the fortune to spend time with. I am forever indebted to these people for their contributions to my scientific and personal growth over the years.

I wish to thank my advisor Frances Arnold. I thank her for having confidence in me and for giving me the opportunity to work in her lab. I thank her for her constant encouragement, support, and guidance. She has been incredibly patient and has given me the time to grow professionally and mature personally. My respect and admiration for Frances continues to grow, and she has become a source of inspiration as I begin my own academic career. I could not have had a better advisor and I will never be able to thank her enough for all that she has given me.

I wish to acknowledge the entire Division of Chemistry and Chemical Engineering for their resources, their courses, and for providing such a supportive and inspiring work environment. In particular, the following faculty members have been extremely helpful:

I thank Dave Tirrell for all his generosity. As Division Chair, Dave has provided invaluable services which I have been able to take advantage of. I thank Dave for his supportive letters of recommendation, for giving me the opportunity to collaborate with his group, and for serving on my thesis committee. I thank Harry Gray for serving on my candidacy and thesis committees, for supplying advice and perspective throughout the years, and for amusing and stimulating conversation. I thank Mark Davis for serving on my thesis committee, for putting in a good word for me, and for good advice. I also acknowledge George Gavalas and Jack Richards for serving on my candidacy committee.

I thank Geethani Bandara, not just for keeping the lab running smoothly and helping out wherever needed, but also for helping me to keep my life in proper perspective. I want to acknowledge Donna Johnson and Kathy Bubash for all of their assistance and administrative services, and Suresh Guptha for being so helpful and for many pleasant conversations.

Zhanglin Lin was an excellent mentor who taught me most of what I know about experimental molecular biology. Zhanglin made laboratory work fun. Uli Schwaneberg and Carlos Martinez played important roles in my scientific development and particularly in my knowledge of P450s. Other Arnold lab members who have been exceptionally helpful include Volker Sieber, Holger Berk, Edgardo Farinas and Daisuke Umeno. I consider all of these people to be great friends and terrific scientists, and I look forward to future interactions with everyone.

Edgardo Farinas has become one of my closest friends over the years. Once again, I must thank Frances for putting Edgardo and I together. We were an instant team and together we explored cytochrome P450 research and established most of the lab's P450 protocols. Outside the lab, our desire to match hard work with hard play led us into memorable mischief.

I thank Yi Tang for his friendship and all his wisdom, and for helping to make our first year together tolerable. Chris Voigt has also been a good friend and a good person to talk to. Yi and Chris set the upper limit on what can be expected from our generation of scientists and I will continue to learn from both of them. I would like to acknowledge all members of the Arnold lab and Tirrell lab for their help, advice, generosity and friendship. More importantly, I thank them for putting up with me. It's no secret that I have a loud personality and a loud stereo to go with it. I only hope that my influence has rubbed off and Springsteen will continue to echo through those halls.

Isaac Carrico and Pin Wang in the Tirrell group have served as reliable sources of information and insight… these guys know everything. Denis Shcherbakov was a hard working SURF student who endured the trials of assay development with me. Anders Olson helped with screening and became a great friend. Oriana Salazar was always uplifting and a delight to work with. Mona Shahgholi provided helpful assistance with MALDI-TOF mass spectral analyses, and Nathan Dalleska and Peter Green assisted with GC procedures.

I thank my friends and family for their continued support and encouragement. My parents and sisters are the most important people in my life and they have been with me every

step of the way. There is no “significant other” to acknowledge, but my life-long friends Dan, Jeremy, Jay, and G are closer to me than brothers. I keep nothing inside; every worry, emotion and anxiety is poured onto my loved ones. There is nothing more comforting in my life than when someone knows the right thing to say to ease my mind. While this is not a complete list of the friends who have enriched my life and provided encouragement over the years, the following people (who have not already appeared above) have been particularly supportive: Tim, Ilias & Eric, Kathy & Mark, Karen (my Bruce buddy), Eimear, and Allen (Chud!).

Finally, I would like to acknowledge the organizations that have helped to fund my research: British Petroleum, the Biotechnology Research and Development Corporation (BRDC) and the NSF.

ABSTRACT

The ability of the cytochrome P450 heme monooxygenases to catalyze difficult oxidation reactions, often with high specificity and selectivity, makes them attractive for numerous bio-technological applications. However they are generally limited by low turnover rates and low stability, and their minimum requirements for catalysis include a cofactor as source of electrons (NAD(P)H), partner proteins for electron transfer, and dioxygen. Some P450s are capable of supporting low levels of peroxygenase activity, in which a peroxide is utilized to drive catalysis via a "shunt" pathway. This mechanism for substrate oxidation, although inefficient and not generally utilized in nature, simplifies P450 catalysis by eliminating the need for NAD(P)H.

Our goal was to engineer an efficient P450 peroxygenase which utilizes hydrogen peroxide (H_2O_2). Directed evolution is a powerful enzyme engineering methodology which mimics nature's algorithm for evolution. Enzyme libraries are generated via DNA mutagenesis or recombination techniques, and variants with improved function are isolated using an appropriate screen or selection. Using this strategy, in combination with site-directed mutagenesis, we have created P450 BM-3 heme domain variants with more than 100-fold improved H_2O_2-driven hydroxylation activity compared to wild-type, showing both an improved k_{cat} as well as a lower K_m for H_2O_2. Thermostability was also improved by directed evolution.

We have engineered a cell-free, biomimetic hydroxylase that requires only H_2O_2 to exploit the hydroxylating power of P450 BM-3. Peroxide-mediated inactivation as a result of heme destruction remains a major obstacle and presents an important enzyme engineering challenge. This research has broadened the potential applications of P450 biocatalysis by exploiting the versatility of heme-containing proteins.

THESIS SUMMARY

Nature provides an arsenal of biocatalysts whose capabilities we are learning to exploit and perfect through protein engineering. Oxygenases comprise several protein families that introduce one (monooxygenases) or two oxygen atoms (dioxygenases) into their substrates, typically through the activation of dioxygen (O_2) using reduction equivalents supplied from the cofactors NADH or NADPH (NAD(P)H) via electron transfer proteins (e.g., reductase). In spite of this rather complex catalytic system, these enzymes and the reactions they catalyze are highly sought-after on all levels of chemical synthesis. Among the array of transformations catalyzed by oxygenases, the ability of the heme-containing "superfamily" of enzymes termed the cytochrome P450 monooxygenases to oxidize unactivated C-H bonds is perhaps the most useful and certainly the most impressive from an applied catalysis perspective. Throughout nature these enzymes play important roles in metabolism and biosynthesis, and they are of primary importance in the design of pharmaceuticals. It is no surprise that the cytochromes P450 are the most studied of the monooxygenases and are marveled for their unmatched biocatalytic potential, diversity, and complexity.

It is often advantageous to employ *in vitro* enzyme catalysis (e.g., to avoid metabolism of desired products, for substrates that cannot permeate cell walls, or where sterile conditions are desired). *In vitro* P450 biocatalysis would require continuous regeneration of the expensive cofactor NAD(P)H from $NAD(P)^+$, adding complexity and cost to the reaction system. *In vitro* P450 catalysis is also sometimes desired in applications where it would not be feasible to use the natural cofactor to drive catalysis,

such as in the use of P450s in laundry detergents for the oxidation of surfactants and clothing stains. Many P450s are capable of using peroxides as a source of oxygen via a peroxide "shunt" pathway:

$$RH + R'OOH \xrightarrow{P450} ROH + R'OH$$

This mechanism for substrate oxidation offers the opportunity to take advantage of P450 catalysis without the need for a cofactor, and eliminates the rate-limiting electron transfer step carried out by the reductase. However, low efficiency is a major limitation as a result of this pathway not generally being utilized in nature. Our goal was to improve and harness this unnatural P450 reaction.

The heme cofactor is ubiquitous in nature and found in many enzyme families. The diverse set of functions nature has employed heme to serve is an indication of the versatility and designability of heme-containing enzymes. Studying heme enzymes and the influence of the protein scaffold on their function provides clues as to what may be possible through protein engineering of cytochrome P450. Chapter 1 of this thesis provides an introduction to heme enzymes and a detailed review of progress made in engineering select heme-containing enzymes. In particular, mutagenesis and directed evolution studies on cytochromes P450 and peroxidases are described and comparisons between these enzymes are made.

Directed evolution has proven a powerful tool for engineering enzymes to have unique or improved properties. Figure I illustrates the evolution process. Libraries of mutated genes coding for the enzyme of interest are generated. The corresponding enzyme variants are expressed and those with improvements in a desired property are selected for further rounds of evolution. A critical factor in the directed evolution

process is the choice of an appropriate screen for function (or growth selection). The assay used must facilitate rapid screening of variants (on the order of 10^3 to 10^5 per day).

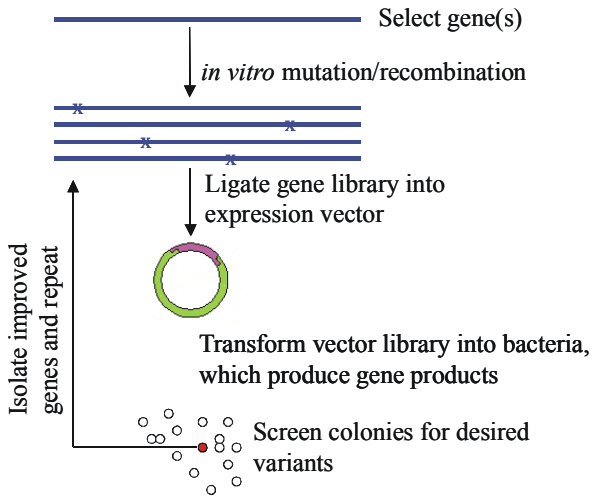


Figure I: Directed enzyme evolution.

We chose the evolutionary algorithm as the framework for improving peroxide-driven hydroxylation activity (peroxygenase activity) of cytochrome P450, using hydrogen peroxide (H_2O_2) as the oxygen donor. The P450 used in this study was a fatty acid hydroxylase from *Bacillus megaterium*, "P450 BM-3", which contains both a reductase domain and a heme domain (the catalytic domain) on a single polypeptide. This monooxygenase is one of the fastest-acting P450s known, capable of hydroxylating fatty acids (C12–C18) with rates exceeding 1000 min^{-1} when utilizing NADPH and O_2. Chapter 2 describes initial characterizations of the shunt pathway with P450 BM-3 and discusses the importance of a well-studied mutation in the active site, F87A, for supporting H_2O_2-driven catalysis. We show that although the F87A mutation causes a shift in the regiospecificity of fatty acid hydroxylation, the product distributions are essentially identical between reactions driven by NADPH and reactions driven by H_2O_2.

Whereas the wild-type enzyme is inactivated after only a few turnovers, heme domain mutant F87A ("HF87A") supports low but measurable levels of activity (~10 min^{-1} in 10 mM H_2O_2). The K_m for H_2O_2 is high (~30 mM for HF87A), meaning the peroxide concentrations required for appreciable activity result in rapid enzyme inactivation. Using directed evolution we hoped to generate variants with improved catalytic rates, reduced peroxide requirements, and enhanced peroxide stability.

Chapter 3 describes an evolved peroxygenase mutant ("21B3") which has more than 15-fold higher peroxygenase activity compared to HF87A and contains nine amino acid substitutions. Improved activity had come at the cost of stability, so we then sought to evolve the thermostability of 21B3 without sacrificing the improved peroxygenase activity. As described in Chapter 4, six additional rounds of mutagenesis and screening resulted in thermostable mutant "5H6", whose half-life at 57.5^{o}C is 50-fold longer than that of HF87A. An analysis of the point mutations accumulated throughout evolution is provided in Chapter 5.

Novel cytochrome P450 BM-3 peroxygenase variants have been generated. These variants utilize H_2O_2 to hydroxylate substrate much more efficiently than the wild-type enzyme and are more thermostable than wild-type. The evolved variants allow us to exploit this powerful and versatile hydroxylase in a cell-free reaction system without requiring a reductase or cofactor. Our "biomimetic" catalyst is easy to synthesize, requires minimal preparation (crude lysate or purified protein can be used), and is active under mild conditions.

We have made useful advances in the field of oxidative biocatalysis and in overcoming the cofactor problem for *in vitro* applications of P450s. Peroxide-mediated

inactivation remains a major obstacle, the elimination of which presents an important enzyme engineering challenge. In Chapter 6 we address this issue of enzyme inactivation and show that it is primarily the result of heme bleaching. Chapter 6 also describes a study in which all 13 methionine residues in the P450 BM-3 heme domain mutant "TH4" were replaced with the isosteric methionine analogue norleucine. This experiment was both a means of testing the functional limits of global amino acid analogue incorporation and a method to test the effects of peroxide on methionines during peroxide-driven P450 catalysis. While there was no increase in the stability of this protein to peroxide, we were surprised to find twofold increased peroxygenase activity. Finally, Chapter 7 details the experimental procedures used throughout the studies described in this thesis.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Exploring the Diversity of Heme Enzymes through Protein Engineering

Material from this chapter appears in: Cirino PC, Arnold FH: **Exploring the diversity of heme enzymes through directed evolution**. In *Directed Molecular Evolution of Proteins*. Edited by Brakmann S, Johnsson K: Wiley-VCH; 2002:215-243, and is reprinted by permission of Wiley-VCH.

Introduction

Enzymes are capable of clean, specific catalysis with high turnover rates. They have already proven useful in numerous synthetic applications, particularly for the high value and often chiral compounds demanded by the pharmaceutical, agricultural, and food industries. Redox enzymes such as peroxidases and oxygenases catalyze valuable reactions on a vast spectrum of substrates. Despite their impressive synthetic potential, these enzymes have enjoyed only limited use due to their relative complexity, instability and, in some cases, low catalytic efficiency. Demands for clean, economical oxidation processes and for increasingly complex and specific oxidation products provide a strong driving force for considering biocatalytic routes. Directed evolution may be able to eliminate some of the shortcomings of enzymes, while improving and harnessing their natural catalytic power.

Metalloporphyrins are synthesized naturally and utilized biologically as redox catalysts, and as such are essential to life. These metal complexes have different chemical functions (see [1]); nature has discovered the ability to modulate the function by incorporating them into proteins which allow for a tremendous diversity of architecture and chemical environments surrounding the prosthetic group. Within the protein framework the prosthetic group becomes a versatile tool with varying, highly specialized capabilities.

Heme serves as the active center in different families of proteins classified by structural similarity (e.g., heme-binding peroxidases, cytochromes P450, globins, catalases). Within these families the metalloporphyrin has a primary function (e.g., hydroxylation or oxygen binding), but there is also considerable functional overlap

among them. The protein regulates the function, but it is not known whether the particular folds that characterize each class are required for optimal function of that class. One could argue that nature has had a long time to optimize an enzyme's structure-function relationship. But evolution is restricted in its exploration of structure space and is contingent on previous history. P450-type hydroxylation reactions, for example, might be efficient in scaffolds very different from the one nature has adopted, but we only see the one that was discovered first. The original function of the P450 enzymes, in fact, may not even have been oxygen insertion. Some folds may be more likely to occur than others and therefore have a higher probability of being encountered, even though they may not be optimal. There are likely to be other, and perhaps even better, solutions that nature for a number of reasons may not have adopted, but that can be created in the laboratory. Directed evolution allows us to explore the interconversion of function within a structural framework and therefore address the question of how easily functions that are primarily associated with one scaffold can be grafted into another.

Evolution mainly reflects the demands of survival and reproduction. Any one function is only as good as it has to be; function must also accommodate biological needs (e.g., regulation) that may hamper its individual potential. It is clearly possible to engineer existing enzymes to improve specific functions, especially when they are removed from the context of biological compatibility. Evolution in the laboratory allows us to explore function free from biological constraints and to access and optimize functions or combinations of functions that are not biologically relevant. Such evolutionary design experiments will provide new insight into structure-function relationships while at the same time develop useful catalysts.

Heme Proteins

Heme consists of a tetrapyrrole ring system complexed with iron. The four pyrrole rings are linked by methene bridges, resulting in a highly conjugated and planar porphyrin. Shown in Figure 1.1 complexed with iron, protoporphyrin IX is one of the most common porphyrins and is the prosthetic group found in all heme enzymes discussed in this chapter. The heme iron is octahedrally coordinated by six heteroatoms. The four equatorial ligands are the porphyrin nitrogens; the remaining two axial ligands lie above and below the plane of the heme. In heme proteins that do not directly bind oxygen or hydrogen peroxide both axial coordination sites are occupied by heteroatoms from nucleophilic amino acid residues. In cytochrome *c*, for example, these atoms are a histidine imidazole nitrogen and a methionine sulfur. In heme enzymes that bind oxygen or hydrogen peroxide, only one axial site is occupied by a basic amino acid heteroatom. This 'proximal' ligand is conserved throughout each enzyme family. All heme peroxidases except for chloroperoxidase (CPO) have a histidine nitrogen as the proximal ligand. In CPO and in all P450s the proximal ligand is a cysteinate sulfur. In catalase, it is a tyrosine oxygen. The sixth coordination site, distal to the heme iron, is occupied by an oxygen atom from either O_2, H_2O or peroxide. This distal position is the catalytic center, and its coordination depends on the enzyme's status in the catalytic cycle.

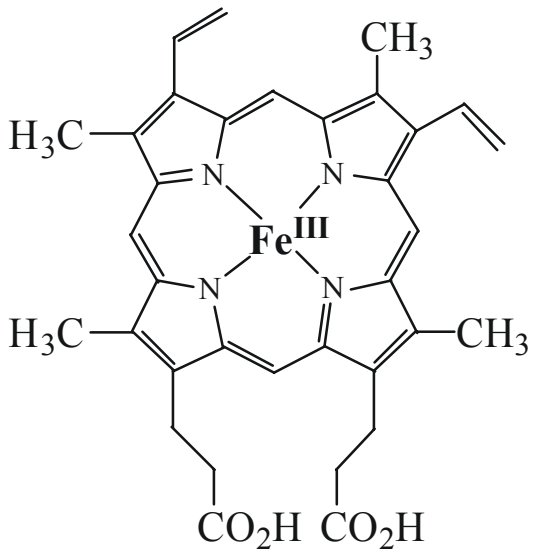


Figure 1.1. Protoporphyrin IX complexed with iron.

Heme proteins collectively have three main functions: oxygen transport, electron transfer and catalysis of redox reactions using either peroxides or oxygen plus externally supplied electrons [2]. In this chapter we will limit our discussion to "b-type" enzymes, which contain a protoporphyrin IX [3]. We further focus on proteins comprising a single heme and a single polypeptide chain. This group of hemoproteins contains many well-studied enzymes whose (relative) simplicity and varying oxidative activities make them attractive catalysts. Available crystal structures of important b-type enzymes help us understand how the different heme-protein, heme-substrate, and protein-substrate interactions modulate active-site chemistry. Additionally, protein engineering studies of these enzymes have provided useful insights into function and the catalytic potential we might be able to achieve by directed evolution.

Most of the hemoproteins of one family also intrinsically possess functionality primarily belonging to that of a different family. For example, peroxidases have activities typically associated with the cytochromes P450, and vice versa. Enzymes from

both families show catalase activity, and catalase has slight peroxidase activity. Myoglobin (Mb), whose primary function is to transport oxygen, is also capable of oxidizing substrates. Within one family there is still enormous diversity with regard to the primary reaction catalyzed, substrate specificity, catalytic rate, etc. It is the protein that controls factors such as the redox potential and stability of the oxidative iron species, the accessibility of substrates to the active site, and overall enzyme stability. With the same heme structure acting as the catalytic center for all these enzymes, it is interesting to try to understand how the protein serves to modulate heme catalysis.

Cytochromes P450

A. Introduction

A large volume of literature attests to the versatility of the cytochrome P450 monooxygenases with regard to substrate specificity, regio- and stereoselectivity and the breadth of reactions catalyzed. This family of monooxygenases was branded "P450s" because they exhibit a characteristic UV absorption maximum at 450 nm upon binding of carbon monoxide by the reduced enzyme. Entire books have been dedicated to P450 enzymes [4-6]. Many papers review what has been learned about these enzymes and their potential as catalysts [7-13]. Martinez and Stewart discuss P450 enzymes as catalysts for asymmetric olefin epoxidation [14], and Mansuy has reviewed the many reactions catalyzed by these enzymes [15]. Miles *et al*. [16] have reviewed protein engineering studies on P450s.

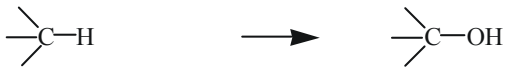
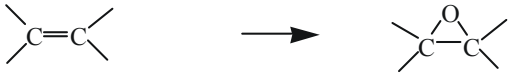
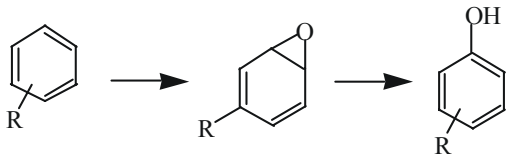
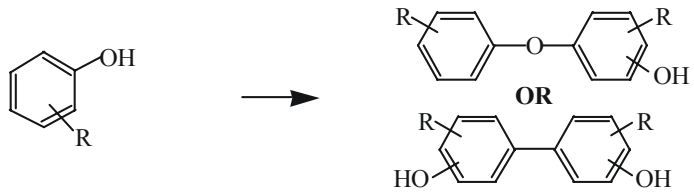
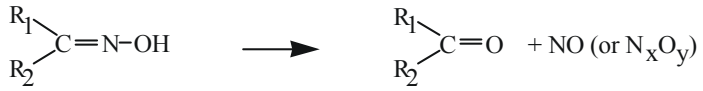
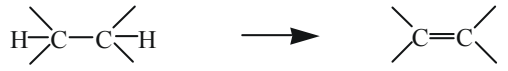
P450s are found in almost all organisms and primarily catalyze insertion of oxygen into carbon-hydrogen bonds. The general reaction equation can be written:

$$RH + NAD(P)H + H^{+} + O_2 \xrightarrow{P450} ROH + NAD(P)^{+} + H_2O.$$

Although monooxygenase-catalyzed hydroxylation and epoxidation reactions are of particular interest in chemical synthesis, P450s catalyze a number of other oxidative reactions, on substrates that range from alkanes to complex endogenous molecules such as steroids and fatty acids. Table 1.1 lists many of the oxidative reactions catalyzed by P450s. These enzymes are also known to catalyze non-oxidative dehydrase, reductase and isomerase reactions [15].

The P450s require a cofactor (NADH or NADPH) as source of reducing equivalents to reduce oxygen and use a protein electron transport system. Depending on the P450, this system is composed of either two proteins (usually a reductase and a ferredoxin protein) or a single P450 reductase flavoprotein. The best-studied of all P450 enzymes is the camphor hydroxylase ($P450_{cam}$) from the soil bacterium *Pseudomonas putida*, which uses two partner proteins (putidaredoxin reductase and putidaredoxin) for electron transfer from NADH to the heme enzyme [17]. Currently atomic structures for 11 different P450s have been determined [18-26]. While their sequence identities are quite low (typically ~20% on the amino acid level), all have similar structures.

Table 1.1. Cytochrome P450-catalyzed reactions [15].

C-H Bond hydroxylation	
Epoxidation	
	
Oxidative N-dealkylation	
Oxidative O-dealkylation	
N-Hydroxylation	
Sulfoxidation	
Peroxidase-type oxidation	
NO Synthase-type oxidation (C=N bond cleavage)	
Oxidative deformylation	
Dehydrogenation	

B. Mechanism

B.1. The Catalytic Cycle

Catalytic activity involves the generation of one or more short-lived, highly oxidizing intermediates at the heme iron and near the bound substrate. Figure 1.2 shows the catalytic reaction cycle. In the substrate-free, oxidized (ferric iron) state of the enzyme **(1)** the heme iron is in the low-spin six-coordinate form [27], with water as the sixth ligand. Binding of substrate results in dehydration of the active site so that the heme iron becomes five-coordinate (intermediate **(2)**) [7]. Additionally, the heme iron changes to predominantly high-spin and its reduction potential increases, thereby priming the enzyme for substrate turnover by allowing electron transfer to occur [7,17,28,29]. It is believed that water exclusion from the active site is important not only for the change in coordination and reduction potential, but also to improve the coupling efficiency of electron transfer (see below).

Oxygen binds to ferrous P450 after the first electron transfer, resulting in an unstable ferrous-oxy species (intermediate **(4)**) which then accepts the second electron. The electron transfer steps are believed to be rate-limiting under natural conditions [7]. The mechanism following the formation of the peroxo-iron species **(5)** involves incorporation of two protons and cleavage of the O-O bond, resulting in water formation. The two protons are pumped into the active site to the distal peroxo oxygen, with the initial formation of a hydroperoxo-iron intermediate **(6).** The two electrons required for this step come from the heme, resulting in heme oxidation to an oxy-ferryl, or iron-oxo species **(7)**. Recent studies indicate that multiple intermediates effect substrate turnover, as described below [30-32]. While the electronic structure(s) of the species performing

oxidation remains in debate, the iron-oxo species **(7)** is the most widely accepted active intermediate in P450-catalyzed reactions, and is normally depicted as (Por)+Fe^{IV}=O by analogy with heme peroxidase *Compound I* (see below).

Studies of site-directed mutants and kinetic solvent isotope effects have led to a proposed pathway for proton delivery to the heme [7,33-37]. A highly conserved threonine residue near the heme seems to play a critical role in relaying protons from the solvent to the heme. It has been proposed that intermediates **(5)**, **(6)**, and **(7)** are all active oxygenating species with varying electrophilic or nucleophilic properties, contributing to the versatility of P450 enzymes (see below) [30-32]. As shown in Figure 1.3, each intermediate, peroxo-iron, hydroperoxo-iron, or iron-oxo, is believed to catalyze a different reaction. Accessibility of protons to the heme, as controlled by the surrounding protein structure, is apparently important in governing the oxidative activity of the enzyme.

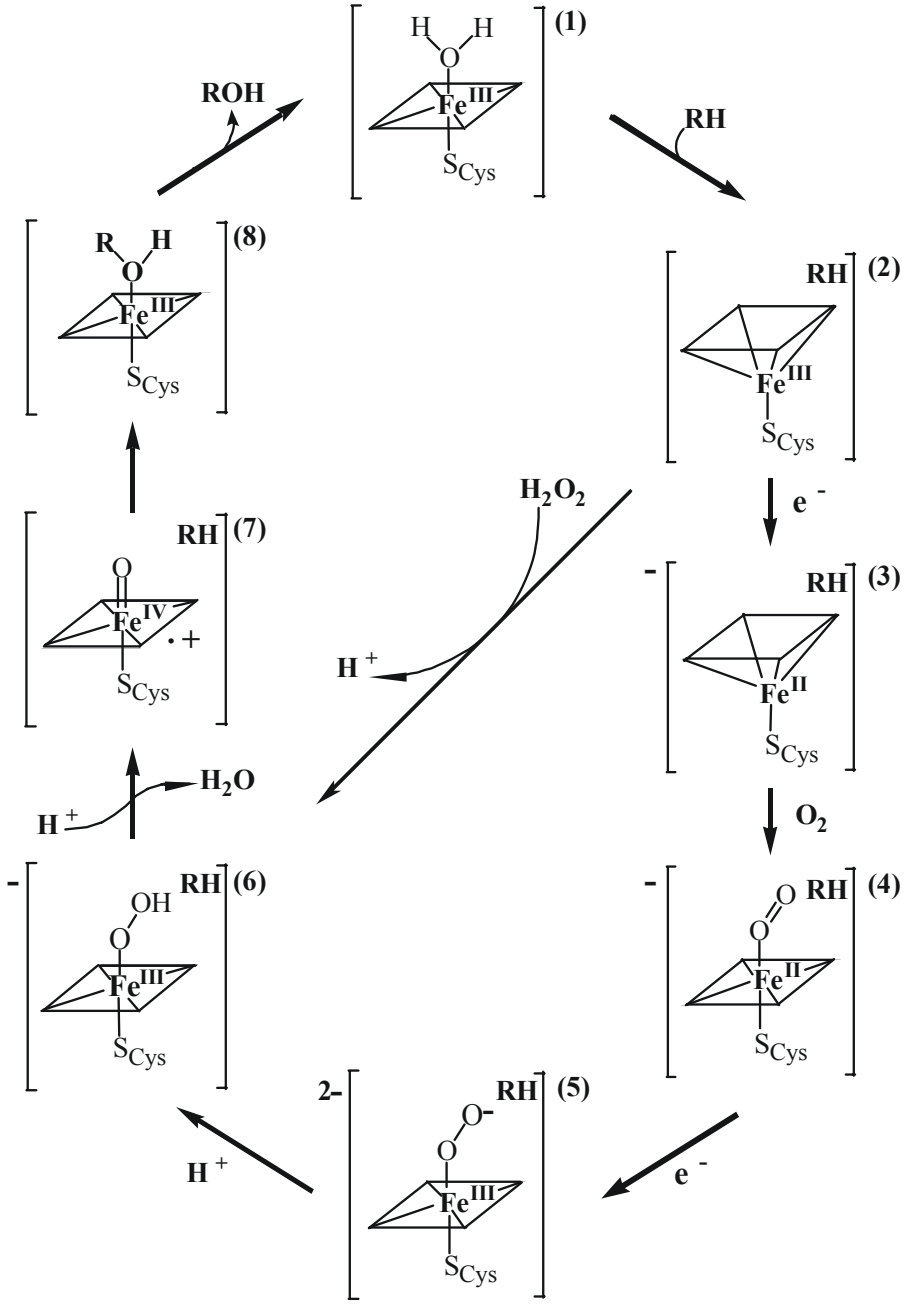


Figure 1.2. Catalytic cycle of P450 including the peroxide shunt pathway. RH is substrate, and ROH is product. The porphyrin molecule is represented as a parallelogram. The overall charge on the structures is shown to the left of each bracket. Intermediates (1), (2), (7) and (8) are neutral. Refer to text for a full description.

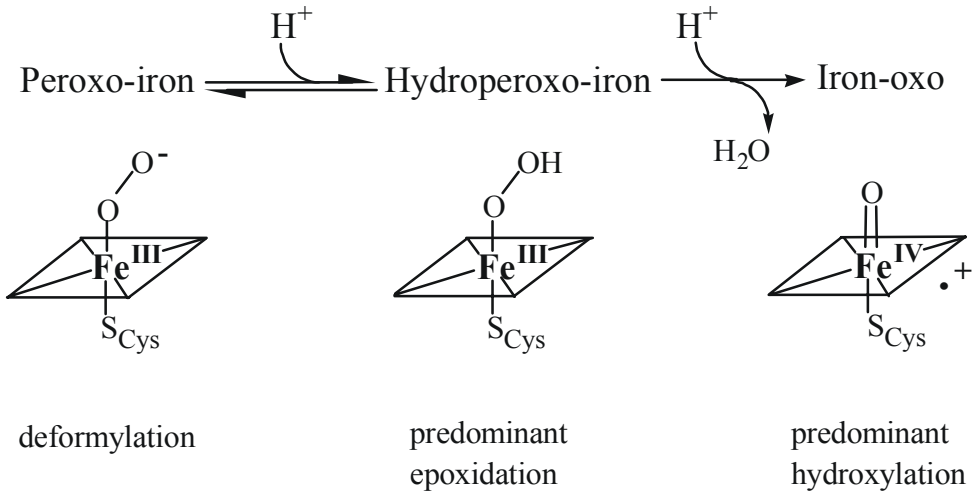


Figure 1.3. Proposed reactive intermediates in the P450 cycle and the primary reactions they catalyze (adapted from [31]). Accessibility of protons to the heme plays an important role in P450 activity.

B.2. Uncoupling

Coupling efficiency refers to the percentage of reducing equivalents from NAD(P)H that are utilized for the oxidation of substrate. The oxidation of camphor by $P450_{cam}$ occurs with 100% coupling efficiency: all electrons transferred from NADH are used in the stereospecific formation of 5-*exo*-hydroxycamphor. In contrast, the coupling efficiency of $P450_{cam}$ in styrene epoxidation is low, between 2% [38] and 7% [39]. In addition to the catalytic cycle shown in Figure 1.2, there are several pathways in which reducing equivalents transferred to the heme can be consumed and transferred away from the substrate. Autooxidation of the ferrous-oxy intermediate **(4)**, H_2O_2 formation through the decomposition of intermediate **(6)** and H_2O formation by two-electron reduction plus diprotonation of intermediate **(7)** are all uncoupling mechanisms. Uncoupling by peroxide formation competes with substrate oxidation if the O-O bond cleavage step is inhibited or if peroxide dissociation is promoted [7]. Access of water molecules to the

heme during catalysis increases polarity and could promote charge separation at the iron, thereby promoting release of hydrogen peroxide anion [17,40]. This occurs when substrates are unable to exclude water from the active site, and would result in increased uncoupling by peroxide dissociation. Reduction of intermediate **(7)** could compete with substrate oxidation if the substrate is positioned too far from the ferryl oxygen.

B.3. Peroxide Shunt Pathway

P450s are capable of utilizing an oxygen atom from peroxide to catalyze oxygen insertion without electron transport proteins or the NAD(P)H cofactor, through the "peroxide shunt" pathway:

$$RH + R'OOH \xrightarrow{P450} ROH + R'OH.$$

As illustrated in Figure 1.2, the shunt pathway bypasses a large portion of the enzyme's natural catalytic cycle, including the rate-limiting first electron transfer step (rate constant of ~15 sec^{-1} is reported for electron transfer in $P450_{cam}$ [41]). There have been many studies of the P450 peroxide shunt pathway [42-56]. Various peroxides and other oxidants (e.g., iodosobenzene, peracids and sodium periodate) will support the reaction, depending on the enzyme. This "peroxygenase" activity potentially adds to the versatility of cytochrome P450 catalysis.

Engineering cytochrome P450 to harness its exquisite catalytic power via the peroxide shunt pathway is the focus of the research described in the following chapters of this thesis. Since we are concerned with the peroxygenase reaction, which is inherent in both peroxidases and P450s, it is useful to compare structural and mechanistic features of

these two enzyme families. Understanding how peroxidases bind peroxide and which factors determine the resulting reaction specificity may provide information to help us engineer an efficient P450 peroxygenase. An introduction to peroxidases and the reactions they catalyze is therefore given below, providing a background for the section that follows in which comparisons are made between P450s and peroxidases. This comparison leads to an introduction to CPO, which shares features of peroxidases and P450s.

Peroxidases

A. Introduction

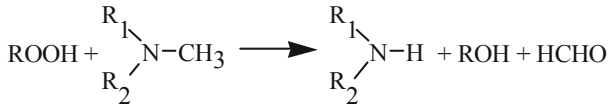
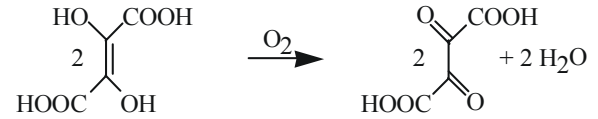
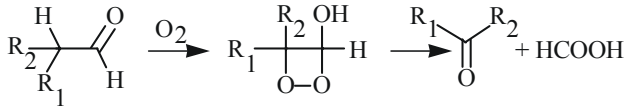
Peroxidases are ubiquitous, and many are b-type heme proteins. Several good reviews summarize years of peroxidase research and describe peroxidase applications [57-61]. Some of the reactions catalyzed by peroxidases are listed in Table 1.2 and include oxidation of aromatic and heteroatom compounds, epoxidation, enantioselective reduction of racemic hydroperoxides, free radical oligomerizations and polymerizations of electron-rich aromatics, and the oxidative degradation of lignin [57,59].

B. Mechanism

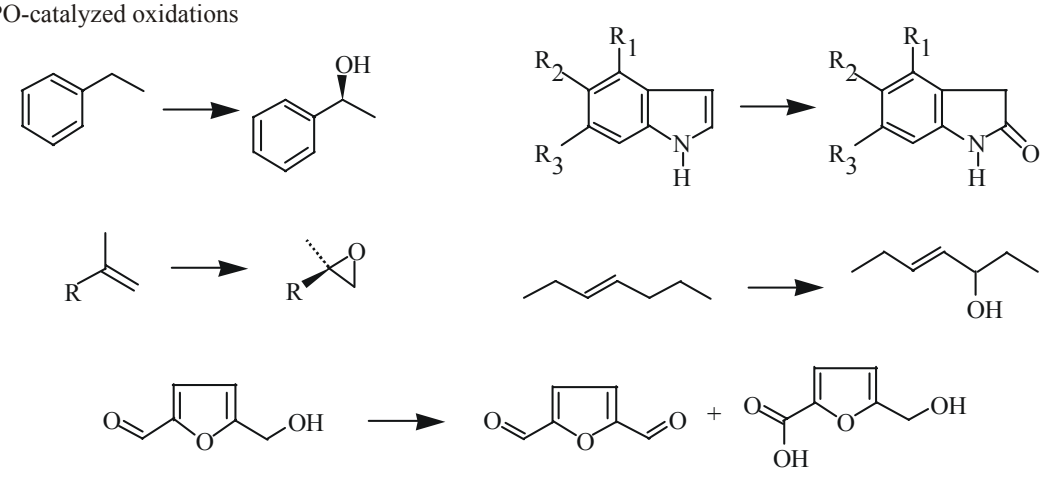
B.1. *Compound I* Formation

Most of what is understood about heme peroxidases comes from studies of the plant enzyme, horseradish peroxidase (HRP). The characteristic peroxidase activity is one-electron oxidation coupled with reduction of H_2O_2 to H_2O. Figure 1.4 shows the peroxidase catalytic cycle. The native state of the heme is the same as that for the P450s,

Table 1.2. Peroxidase-catalyzed reactions (adapted from [59] and [57]).

Reaction		Typical substrates
Electron transfer		H_2O_2, ROOH, aromatic amines
Sulfoxidation		Thioanisole, H_2O_2, ROOH
Epoxidation		Alkenes, H_2O_2
Demethylation		*N*, *N* -Dimethylaniline, ROOH
Dehydrogenation		Dihydroxyfumaric acid
α-Oxidation		Aldehydes

CPO-catalyzed oxidations



except the proximal ligand is a histidine nitrogen rather than a cysteine sulfur. In the first step of the reaction, H_2O_2 replaces H_2O at the axial position of heme Fe^{III}. The bound H_2O_2 is split heterolytically to form an iron-oxo derivative known as *Compound I*, which is formally two oxidation equivalents higher than the Fe^{III} resting state. *Compound I* is well-characterized and contains Fe^{IV}=O and a π cation radical [62].

The mechanism of O-O bond cleavage is influenced by the protein environment around the heme, and is different for peroxidases and P450s. Cleavage of the O-O bond in peroxidases to form *Compound I* is promoted by a "push-pull" mechanism [63] (Figure 1.5). Peroxidases have a catalytically critical histidine distal to the heme, along with an important cationic arginine [64]. The histidine pulls the proton from the heme-bound hydroperoxide, making the hydroperoxide a better nucleophile, while the arginine pulls on the O-O bond; together they work to cleave the O-O bond. In addition, a hydrogen-bonded carboxylate group near the proximal histidine increases electron density to the imidazole, creating an electron "push" that facilitates O-O cleavage. In hemoglobin and myoglobin, where O-O bond cleavage is not part of the primary function, there is no hydrogen-bonded carboxylate group near the proximal histidine. References [65-70] describe mutagenesis studies that have elucidated the roles of the critical distal residues in peroxidase and myoglobin. P450s generally have large substrate binding pockets, with no residues close to the heme on the distal side and therefore no "pull" effect. There, O-O cleavage is facilitated by the proton delivery system and by a very strong "push" by the cysteinate proximal ligand [2,12,71,72].

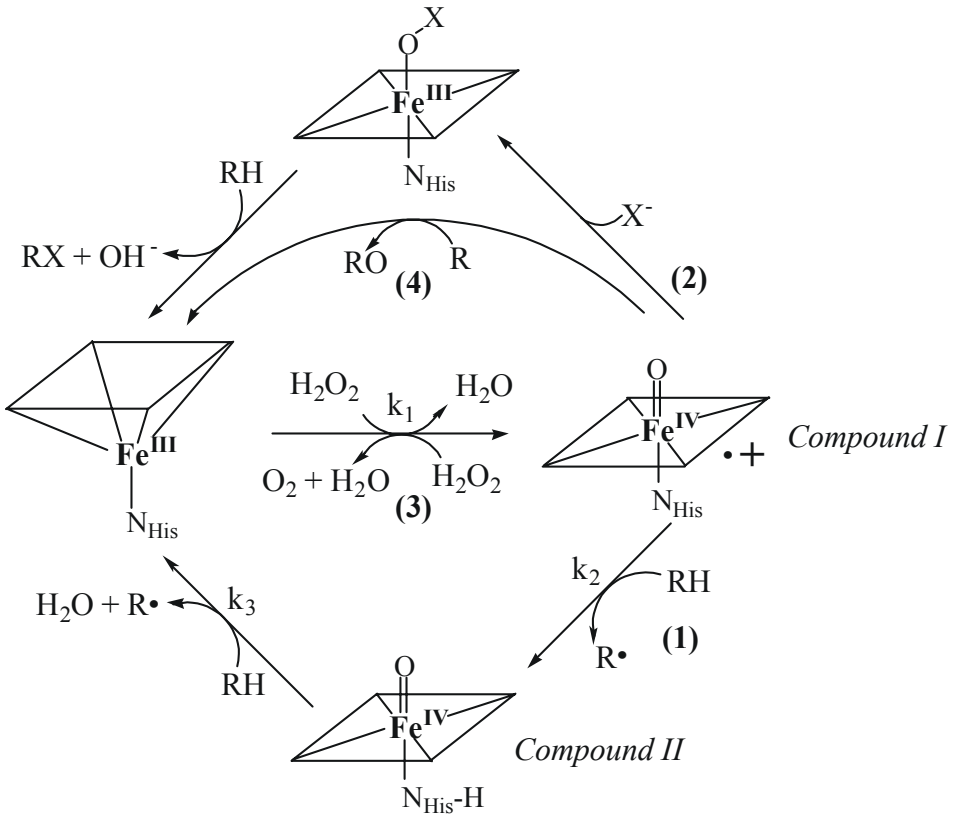


Figure 1.4. Peroxidase catalytic cycle. Four pathways are shown for the return of *Compound I* to the resting state: (1) oxidative dehydrogenation, where RH is the substrate; (2) oxidative halogenation, where X- is a halogen ion and RH is the substrate; (3) peroxide disproportionation, and (4) oxygen transfer, where R is the substrate.

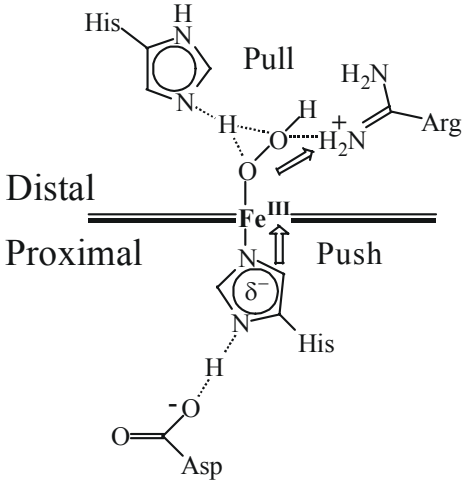


Figure 1.5. "Push-Pull" mechanism for *Compound I* formation in peroxidases (adapted from [12] and [63]). Distal residues (Arg and His) work together to cleave the O-O bond, while proximal residues (His and Asp) assist by supplying electron density.

Reduction of peroxidase *Compound I* back to the resting state can occur by one of four pathways, depending on the reaction catalyzed: oxidative dehydrogenation (pathway **(1)** in Figure 1.4), oxidative halogenation **(2)**, peroxide disproportionation **(3)** or oxygen transfer **(4)** [57].

B.2. Oxidative Dehydrogenation

Oxidative dehydrogenation is a primary biological function of peroxidases. *Compound I* is reduced in two one-electron transfers, as shown in Figure 1.4. The overall reaction is

$$2\ RH + H_2O_2 \rightarrow 2\ R\bullet + 2\ H_2O.$$

The first reduced intermediate, *Compound II*, is formed when the π cation radical is reduced and a proton is transferred to the distal base (His). *Compound II* is then reduced to the Fe^{III} resting state with the simultaneous formation of water [71]. This second electron transfer step is one to two orders of magnitude slower than *Compound I* formation and is usually rate-limiting [73]. For HRP, the rate constant for *Compound I* formation (k_1) is $2.0 \times 10^7\ M^{-1}sec^{-1}$ [74], while the rate-limiting step in phenol oxidation by HRP has a rate constant $k_3 \sim 3.0 \times 10^5\ M^{-1}sec^{-1}$ [75,76].

HRP catalyzes the oxidative dehydrogenation of a wide range of electron-rich aromatic compounds. The result of this radical formation pathway is dimerization and subsequent oligomerization of the substrates [77-79]. Peroxidases have been used to catalyze polymerizations of phenols (e.g., *p*-cresol and guaiacol) and aromatic amines (e.g., aniline, and *o*-phenyldiamine) [80,81]. *N*- and *O*-dealkylations are also useful

electron transfer reactions catalyzed by peroxidases. These reactions are used in industrial wastewater treatment and may have synthetic applications [82].

B.3. Oxidative Halogenation

Heme haloperoxidases can also use peroxide and halide ions to halogenate an activated (benzylic/allylic) carbon. The halide is first oxidized to an active halogenating intermediate (Figure 1.4, pathway **2**). Substrate is halogenated in the next step. The overall reaction is

$$RH + H_2O_2 + H^+ + X^- \rightarrow RX + 2\ H_2O.$$

B.4. Peroxide Disproportionation

The H_2O_2 disproportionation reaction decomposes H_2O_2 into water and oxygen, as shown in Figure 1.4 (pathway **3**). The overall reaction is

$$2\ H_2O_2 \rightarrow 2\ H_2O + O_2.$$

Catalase, a heme-containing enzyme with tyrosine as the proximal heme ligand, decomposes hydroperoxides and peracids by this reaction. Catalase is one of the most efficient enzymes known, with maximum turnover numbers on the order of 10^7 sec^{-1}. In the absence of an electron donor, peroxidases and particularly CPO exhibit catalase activity [59,83], as do some P450s [84].

B.5. Oxygen Transfer

The oxygen transfer reactions catalyzed by peroxidases are very interesting for synthetic applications. The overall reaction for this peroxygenase activity is

$$R + H_2O_2 \rightarrow RO + H_2O.$$

Similar to the P450s, these reactions are often stereospecific and include hetero-atom oxidations, epoxidation, and oxidation of C-H bonds in allylic/benzylic compounds, alcohols, and indole [60].

Comparison of P450s and Peroxidases

P450s and peroxidases share key elements of their mechanisms. The proximal ligands and distal and proximal protein environments influence the mechanism of O-O bond cleavage, the stability of the intermediates and the accessibility of substrates to the heme. The fact that P450s are capable of oxidizing aliphatic hydrocarbons and olefins, in addition to the activated C-H bonds oxidized by peroxidases, reflects the stronger oxidative potential of the P450 iron-oxo species [85,86]. In P450s the substrate binding pocket is relatively large and lies on the distal side of the heme, where the ferryl oxygen can be transferred to the substrate. Nature has tuned this binding pocket in the different P450s to make catalysts with a range of substrate specificities and reaction selectivities. Peroxidases, on the other hand, bind substrates near the heme edge, where electrons are transferred from the substrate to the heme center [60,87]. This ability to oxidize substrates at the heme edge accounts for the broad substrate specificity of peroxidases [88]. While peroxidases have high one-electron oxidation activities, their oxygen transfer (peroxygenase) activities are low, reflecting limited substrate access to the heme iron (with the exception of CPO). Note that some P450s catalyze the coupling of phenols that is typical of peroxidases (shown in Table 1.1) [15]:

$$2\ ArH + NAD(P)H + H^+ + O_2 \xrightarrow{P450} Ar\text{-}Ar + NAD(P)^+ + 2\ H_2O.$$

The P450 shunt pathway and peroxygenase activity of peroxidases share identical overall reaction equations. P450s generally have high K_m values for H_2O_2; values of 15 mM [42], 30 mM [89] and 250 mM [53] have been reported, while for HRP the K_m for H_2O_2 is 22 μM [90]. Reaction of H_2O_2 with the P450 heme is several orders of magnitude slower than the equivalent reaction with peroxidase. Rapid reaction of H_2O_2 with the heme Fe^{III} requires a base to assist in binding of the peroxy anion [2]. The distal histidine serves this function in peroxidases. Replacing the distal histidine with a leucine in cytochrome *c* peroxidase (CCP) reduced the rate of reaction with peroxide by five orders of magnitude [91], to a rate similar to that of a P450.

Two enzymes worthy of mention are the natural H_2O_2-utilizing peroxygenase P450s SPα from *Sphingomonas paucimobilis* (CYP152B1) [92] and BSβ from *Bacillus subtilis* (CYP152A1) [93], which share 44% amino acid sequence identity [93]. $P450_{SP\alpha}$ catalyzes H_2O_2-driven α-hydroxylation of fatty acids (carbon chain lengths between 11 and 18) with high stereoselectivity (~98%) and turnover rates greater than 1000 min^{-1} [94]. P450 BSβ produces both the β-OH and α-OH fatty acids. The active site environment in these P450s is different from other P450s and includes an Arg residue (as seen in peroxidases) [95-97], presumably to accommodate H_2O_2 binding. Only fatty acids are accepted as substrates and the carboxylate of the substrate is positioned into the active site upon substrate binding. It is believed that the carboxylate and Arg residue work together to bind peroxide and facilitate O-O bond cleavage, resulting in low K_m's for H_2O_2 (21 μM for P450 BSβ [96] and 72 μM for P450 SPα [98]) and high catalytic

efficiency. Their highly specialized catalytic mechanisms limit applications of these P450s for biotechnologically relevant transformations.

The reactions catalyzed by the cytochromes P450 are important in synthetic chemistry. The cofactor regeneration requirements of the P450s, however, severely limit their use outside of whole cells. Methods that use electrodes to drive P450 reactions are being developed, but have proven difficult [99-102]. The shunt pathway is a possible alternative to using cofactors, but this pathway is slow and the required peroxide levels are destructive to the enzyme. Directed evolution could improve the shunt pathway by increasing the stability of P450 in the presence of peroxides, increasing catalytic rates, and lowering the required peroxide concentration (reduce K_m).

P450s are generally less stable than peroxidases, although two naturally thermostable P450s (stable up to 85^{o}C) have recently been identified and characterized [23,103-107]. P450s and peroxidases are inactivated during catalysis, via heme alkylation by terminal olefins and oxidative damage by peroxides. Eukaryotic P450s are associated with cell membranes and are therefore insoluble and difficult to use outside the cell. Many of these limitations can be addressed by directed evolution.

Chloroperoxidase

CPO from the fungus *Caldariomyces fumago* catalyzes the oxidation of hydrochloric to hypochlorous acid with a turnover rate of ~10^3 sec^{-1} [58]. While functionally categorized as a haloperoxidase, CPO possesses catalytic traits characteristic of peroxidases, P450s and catalase. The CPO proximal ligand is a cysteinate sulfur as in P450s. In the distal pocket, the catalytic base used for O-O cleavage is glutamic acid rather than the typical

peroxidase histidine. Poulos and coworkers compared its crystal structure to those of peroxidases and P450s [108]. The CPO distal region is more hydrophobic than in other peroxidases, which allows it to bind substrates and promote P450-type reactions, although the presence of polar residues and restricted access to the distal face make the distal region more peroxidase-like than P450-like. The tertiary structure of CPO, however, resembles neither the P450s nor peroxidases [108], implying a separate evolutionary origin.

Several papers review the various reactions catalyzed by CPO [57-60,109,110]. Some of these are shown in Table 1.2. CPO catalyzes the sulfoxidation of thioanisole at a turnover rate of 200 sec^{-1} with >98% enantiomeric excess [110]. This is orders of magnitude higher than typical peroxidases. CPO is also known to catalyze sulfoxidations on aliphatic sulfides at a similar rate [111]. CPO is the only peroxidase known to selectively hydroxylate hydrocarbons [57] and selectively catalyze the oxidation of primary alcohols to aldehydes [112,113]. Chiral epoxides are useful in chemical synthesis because they can undergo stereospecific ring-opening to form bifunctional compounds. CPO catalyzes chiral epoxidations with high yields and high stereoselectivity [109]. Styrene epoxidation proceeds at 4.8 sec^{-1} with CPO [114], versus ~4 min^{-1} with $P450_{cam}$ [39]. Other CPO-catalyzed reactions include alkyne hydroxylation [115,116] and heteroatom dealkylation [117,118].

The rate of reaction of CPO with hydroperoxides is significantly lower than for other peroxidases, and closer to that for P450s [88]. Thus CPO is a rather slow one-electron oxidation catalyst, but it is the preferred heme enzyme for many sulfoxidation, epoxidation and hydroxylation reactions. CPO substrates, in contrast to P450 substrates,

must have electron-rich groups. One major limitation to using CPO is its instability to peroxide. Whereas its half-life at pH 5 is 40 hrs without H_2O_2, the half-life drops to 0.5 hrs in 1 mM H_2O_2 [119]. To help overcome this, methods have been developed to maintain low levels of H_2O_2, either by *in situ* generation or controlled addition [119-123].

The P450 scaffold has evolved to serve as a ubiquitous monooxygenase. Nature manipulated this structure to create an immense library of P450s with finely tuned activities and specificities. The CPO scaffold has not become a ubiquitous oxidation catalyst, but it supports high peroxygenase activity and may be superior to P450s for chemical transformations of certain substrates, particularly when it has been evolved in the laboratory for optimal performance in these applications. A better alternative may be to engineer cytochrome P450 to mimic the (peroxygenase) function of CPO while maintaining P450's substrate specificity and ability to hydroxylate unactivated C-H bonds. This is the central theme of the research described in the following chapters.

Protein Engineering of P450s and Peroxidases

Heme enzymes are prime targets for biocatalyst engineering. For P450s and peroxidases, site-directed mutagenesis studies primarily focused on those enzymes for which crystal structures are available, although they can be extended to other enzymes by sequence alignment and homology modeling [16]. Mutations that change substrate specificity, reaction specificity, activity and coupling efficiency have been reported. Studies carried out on other heme enzymes show how amino acid substitutions can alter substrate specificity and explore the interconversion of function among hemoproteins, based on structure comparisons.

Whereas site-directed mutagenesis is helpful for testing structure-based mechanistic hypotheses, directed evolution methods are particularly valuable when a better catalyst is the goal. Relatively few directed evolution studies have been performed on heme enzymes, and technical hurdles of heterologous enzyme expression and screening technology remain. However, recent developments in screening methods, expression and new methods of library generation are now making it possible to rapidly isolate specialized variants. In this section we describe mutagenesis and directed evolution results that demonstrate the importance of protein structure in controlling heme activity and the ability to engineer function in P450s and peroxidases. Successes in engineering activity, substrate specificity, and reaction specificity in peroxidases not only suggest that similar protein engineering achievements can be made with P450s, but also provide clues as to how to approach such attempts.

A. P450s

A.1. $P450_{cam}$

P450s have finely tuned substrate binding pockets and active sites. Loida and Sligar investigated how mutations of active site residues at varying distances from the heme affect turnover and the partitioning between uncoupling pathways in the oxidation of ethylbenzene by $P450_{cam}$ [40,124]. Their results show that mutations can alter the coupling efficiency. As predicted, increasing the size of side chains close to the heme blocks substrate access and practically eliminates product formation. With substrates that do not fit tightly in the binding pocket, mutations further away that increased bulk help to

force the substrate closer to the heme and increase product generation. The effects of single mutations were shown to be additive.

A tyrosine is positioned directly over the heme on the distal side in the $P450_{cam}$ active site. This Tyr96 residue interacts with the natural substrate camphor through a hydrogen bond to position it for regio- and stereospecific hydroxylation [18]. Various Tyr96 substitutions create active sites of varying volume and hydrophobicity that improve the oxidation of phenyl derivatives [125,126]. For example, the oxidation of diphenylmethane, diphenylamine and 1,1-diphenylethylene by Tyr96Ala and Tyr96Gly mutants was regiospecific, with hydroxylation at the same *para* position for all three substrates, while wild-type $P450_{cam}$ showed no activity towards these substrates. Styrene oxide is produced from styrene by wild-type $P450_{cam}$ at ~4 min^{-1}. Nickerson *et al.* [39] made Tyr96Ala and Tyr96Phe mutations to increase hydrophobicity and improve binding of styrene. Styrene oxidation was improved 25-fold by the Phe mutation and 9-fold by Tyr96Ala. In addition, the coupling efficiency was increased from 7% to 32% for Tyr96Phe mutant, probably due to enhanced exclusion of water. Wong and coworkers created site-directed mutants of $P450_{cam}$ that oxidize polychlorinated benzenes with considerably enhanced activity and coupling efficiency [127]. These same mutants were also found to oxidize monoterpenes, which are of interest in fine chemical synthesis [128].

Cofactor requirements are a serious impediment to industrial use of biocatalysts such as cytochrome P450. In an effort to bypass the external cofactor and electron transfer proteins altogether, Joo *et al.* [129] evolved $P450_{cam}$ to better utilize hydrogen peroxide for hydroxylation via the shunt pathway. $P450_{cam}$ variants created by error-

prone PCR followed by DNA shuffling showed improvement in naphthalene hydroxylation utilizing H_2O_2. Additionally, mutants with different regiospecificities were identified.

A.2. P450 BM-3

Key to the success of directed evolution are rapid, functional screens for identifying improved enzymes. A clever colorimetric assay for hydroxylation activity has facilitated several recent P450 engineering studies needing high throughput activity screens. The assay, developed by Schwaneberg and coworkers [130], uses ω-*p*-nitrophenoxycarboxylic acids (pNCAs) as fatty acid surrogate substrates, which upon hydroxylation at the terminal carbon produces yellow *p*-nitrophenolate (pNP) through the formation of an unstable hemiacetal, as demonstrated in Figure 1.6 for the hydroxylation of ω-*p*-nitrophenoxydodecanoic acid (12-pNCA). Subsequent modifications to the assay have allowed its use in high throughput screening of enzyme libraries [131] and for alkane substrates [132].

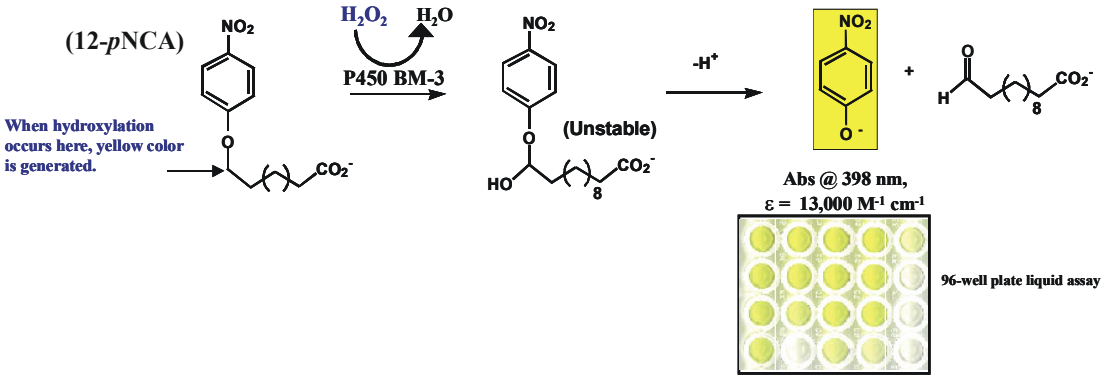


Figure 1.6. Screen for P450 fatty acid hydroxylation using a surrogate substrate.

Cytochrome P450 BM-3 (CYP102) from *Bacillus megaterium* P450 BM-3 primarily catalyzes the hydroxylation of fatty acids (~12 to 18 carbons long) at the ω-1, ω-2, and ω-3 positions, but also hydroxylates long-chain amides and alcohols and epoxidizes long chain unsaturated fatty acids [133-135]. It is called a "self-sufficient" P450 because it contains the required diflavin NADPH-P450 reductase on the same polypeptide as the P450 hemoprotein [136]. P450 BM-3 is soluble, easily expressed in recombinant *E. coli*, and highly active, with turnover rates in the thousands per minute for fatty acids.

Schwaneberg *et al.* [130] showed that P450 BM-3 hydroxylates pNCAs, with a minimum required substrate chain length of 10-11 carbons. BM-3 residue Phe87 plays an important role in determining the regioselectivity of fatty acid hydroxylation [137]. Substitution with Ala at this position is reported to shift hydroxylation towards the ω position, and the F87A mutation increases the sensitivity of the 12-pNCA assay compared to wild-type P450 BM-3 by shifting the regioselectivity for hydroxylation of this substrate to the terminal carbon [130]. Li *et al*. demonstrated the importance of residue size at position 87 in determining the stereoselectivity of oxidation of the unnatural substrates propylbenzene and 3-chlorostyrene [138].

The P450 BM-3 heme domain crystal structure shows a funnel-shaped shaft lined with hydrophobic residues for substrate binding [19,139,140]. Schmid and coworkers combined site-directed mutagenesis with site-specific saturation mutagenesis at residue positions known to affect substrate binding and screened the BM-3 mutant libraries for higher activity on pNCAs with chain lengths of 12, 10, and 8 carbons [141-144]. They found mutants that accept smaller chain substrates (a variant with five mutations that

efficiently hydroxylates 8-pNCA was obtained) [142], hydroxylate indole [141], oxidize octane and naphthalene much faster [144], and oxidize polycyclic aromatic hydrocarbons (PAHs) poorly accepted by wild-type [145].

Using site-directed mutagenesis, Carmichael and Wong engineered a BM-3 mutant capable of PAH oxidation by applying site-directed mutagenesis at residues in the active site and at the entrance of the substrate access channel [146]. Both PAH studies report orders of magnitude improvements in activity on different PAH's, although the NADPH coupling efficiencies are all extremely low. Binding and turnover of the small substrates butyrate and hexanoate were greatly increased by generating combinations of site-directed mutations that influence substrate binding [147].

In another example in which directed evolution was used to alter the substrate range of P450 BM-3, Farinas *et al.* generated mutants with improved activity for alkane hydroxylation [132] by screening mutant libraries on the surrogate substrate 8-pnpane. Two generations of laboratory evolution yielded variants with up to fivefold higher octane oxidation activity compared to wild-type P450 BM-3 and the mutants retained the very high coupling efficiency of the wild-type enzyme on its natural, fatty acid substrates. Further rounds of directed evolution have produced mutants capable of hydroxylating a variety of alkanes (C3 to C8) with turnover rates exceeding those of any known alkane hydroxylase [148].

This work suggests that P450 BM-3 can be engineered to hydroxylate a wide range of substrates. This is particularly significant in light of the superior activity of P450 BM-3 compared to most other P450s and the fact that it is easily expressed and purified from *E. coli*. The plasticity of this enzyme will certainly continue to be

exploited in future protein engineering studies and is further demonstrated by the research described in the following chapters.

A.3. Eukaryotic P450s

Although the exact mechanism for P450-catalyzed epoxidation is not known, recent studies by Coon and coworkers indicate that at least one heme intermediate is active towards epoxidizing carbon-carbon double bonds [31]. To reduce proton delivery to the active site, they replaced the conserved threonine believed to facilitate proton delivery by alanine in two P450 enzymes (P450 2B4 and P450 2E1, both from rabbit liver). They then compared the rates of product formation and the ratios of epoxidized to hydroxylated products for the wild-type enzymes and the Thr→Ala mutants for several substrates. Figure 1.3 shows the proposed heme species and the oxidations catalyzed [31]. Coon proposes that, in addition to the iron-oxo species traditionally believed to be responsible for P450-catalyzed oxidations, the hydroperoxo-iron species could also epoxidize olefins. The experiments support this: for *cis*-butene, the ratio of epoxidized to hydroxylated product increased fivefold with the Thr→Ala mutant. These results once again demonstrate how the protein structure tunes catalytic activity. It is likely that P450-catalyzed epoxidation could be greatly enhanced by directed evolution.

The ability to make functional cytochrome P450-NADPH reductase fusion proteins, particularly for mammalian P450s, would simplify the study and application of these enzymes. Such efforts have been reviewed [16]. Recently, human P450 CYP2D6 has been linked to human NADPH-cytochrome P450 oxidoreductase (CPR), which is the first report of a functionally complete human P450 fusion enzyme system [149]. Site-

directed mutagenesis at a single residue converted the human P450 redox partner NADPH CPR to a functional NADH-dependent reductase [150]. Sadeghi *et al*. report a functional fusion between the heme domain of P450 BM-3 and a flavodoxin protein from *Desulfovibrio vulgaris* [151]. Shimizu and colleagues reported that the nitric oxide synthase (NOS) reductase domain is unable to effectively substitute for that of cytochrome P450 BM-3, while the BM-3 reductase domain, in contrast, was able to support low levels of NOS activity [152].

Mammalian P450s are membrane-bound and difficult to express in recombinant organisms, and until recently [26] no crystal structures have been available. Little practical engineering has been done with mammalian P450s; most mutagenesis studies have examined structure-function relationships [153]. Sakaki and Inouye discuss practical applications of these enzymes [154]. Random mutagenesis methods have recently been applied to eukaryotic P450s resulting in P450s with altered specificities and functions [155-158].

B. Peroxidases

B.1. CiP

Cherry and coworkers evolved a fungal peroxidase (CiP) from the ink cap mushroom *Coprinus cinereus* to resist high temperature, high pH and high concentrations of peroxide [159]. The goal was to develop a peroxidase to be used as a dye-transfer inhibitor in laundry detergents, an application requiring stability under harsh conditions. Site-directed mutagenesis improved the enzyme's stability to alkali and hydrogen peroxide. These mutations were combined with other mutations discovered by random

mutagenesis (error-prone PCR) and saturation mutagenesis and screening. Combination of all the favorable mutations generated a mutant with 100-fold improved thermal stability and 2.8 times the oxidative stability of wild-type, although these improvements came at the cost of reduced overall activity. *In vivo* shuffling using a yeast homologous recombination system was then employed to improve activity. Mutants with high activity were shuffled with those showing improved thermostability but reduced activity. This generated a mutant with 174 times the thermal stability and 100 times the oxidative stability of wild-type CiP, with specific activity comparable to wild-type.

B.2. CCP

Cytochrome c peroxidase (CCP) catalyzes the oxidation of ferrocytochrome *c* (cyt *c*). Many site-directed mutagenesis studies have been performed on this enzyme in efforts to better understand and rationally engineer heme peroxidase function [160-166]. Various studies indicate that the large protein substrate cyt *c* binds to CCP in a different region than smaller substrates [166,167]. Whereas small substrates such as phenol and aniline are believed to approach the heme from its distal side and bind at the heme edge, cyt *c* lies much further from the heme, and electron transfer seems to occur from the proximal side [168]. The low activity and selectivity exhibited by CCP compared to HRP and CPO is attributed to the limited access by small substrates to the heme [169].

Iffland *et al.* [170] used directed evolution to generate CCP mutants with novel substrate specificities. They used error-prone PCR and DNA shuffling to generate mutant libraries and screened for increased activity on guaiacol by detection of the brown-colored product tetraguaiacol. After three rounds of evolution mutants were

isolated with 300-fold increased guaiacol activity and up to 1000-fold increased specificity for guaiacol relative to cyt *c*. It is interesting that all selected mutants contained the mutation Arg48His. Arg48 is a conserved distal residue that aids in *Compound I* formation and stabilization in the "push-pull" mechanism. The histidine in its place apparently reduces steric constraints for substrate access to the heme, but provides enough charge to maintain activity.

B.3. HRP

The distal residues in peroxidases impair their ability to catalyze oxygen insertion reactions. In HRP, His42 is the conserved distal residue that aids in *Compound I* formation. Site-directed mutagenesis studies have shown that this histidine can be relocated to a nearby, less obstructive position and still promote *Compound I* formation [68]. The mutation His42Ala alone reduces the rate constant of *Compound I* formation (k_1 in Figure 1.2) from ~10^7 $M^{-1}sec^{-1}$ to ~10^1 $M^{-1}sec^{-1}$, while a second mutation Phe41His brings this rate constant back up to $3x10^4$ $M^{-1}sec^{-1}$. With this double mutant, the k_{cat} for styrene oxidation increased from 10^{-6} sec^{-1} (for wild-type HRP) to $2.4x10^{-2}$ sec^{-1} (a 24,000-fold improvement), and thioanisole sulfoxidation increased from 0.05 sec^{-1} to 5.3 sec^{-1}. The peroxidase activity of this double mutant, however, was drastically reduced with respect to wild-type. In later work Savenkova *et al.* [67] used the mutation Arg38His to recover the role of the distal histidine from His42Val. For the Arg38His, His42Val double mutant, the rate of thianisole sulfoxidation was increased 680-fold, with peroxidase activity still greatly sacrificed.

HRP is widely used as a reporter enzyme in bioanalytical chemistry and diagnostics; it also has potential applications in chemical synthesis. Protein engineering of HRP has been very limited, however, due to the lack of a convenient microbial host for functional expression and mutagenesis. Expression in *E. coli*, for example, leads primarily to formation of inactive inclusion bodies [171]. HRP contains four disulfide bridges and is ~21% glycosylated by weight [171-174]. *S. cerevisiae* is known to facilitate glycosylation and disulfide formation [175], although the patterns for yeast protein modifications differ from those in fungi and plant cells [176,177]. Morawski *et al.* used directed evolution to discover mutations that would promote secretion of functional HRP in yeast [178]. Having achieved functional expression in *S. cerevisiae*, they were able to further evolve the activity and stability of the enzyme [179]. A traditional colorimetric peroxidase assay was used for screening, in which the substrate 2,2'-azino-di-(3-ethyl)benzthiazoline-6-sulfonic acid (ABTS) is oxidized in one-electron steps to generate intensely colored radical products. With three rounds of random mutagenesis and screening, the total activity of yeast culture supernatant was increased 40-fold over wild-type. The best mutants were then expressed in *Pichia pastoris*, resulting in further improvements in total activity. Additional evolution generated HRP variants with improved thermostability and greater resistance to a series of chemical denaturants, as well as higher total activity.

While peroxidases are generally more stable than P450s, both families of enzymes are generally considered to be of low stability compared to enzymes commonly used in industrial applications (e.g., hydrolases). Few naturally thermostable peroxidsases and P450s are known (refer to [180] and [179]) and little is known about the reactivities and

natural substrates of those which have been described. The work by Cherry *et al.* and Morawski *et al.* described above are demonstrations of the ability to improve the thermostability of biotechnologically important heme enzymes without sacrificing function.

My contribution to the evolution of HRP is described in more detail in Appendix A. This research was my first experience with developing and employing a screening assay for directed enzyme evolution, and helped build a foundation for my future research with cytochrome P450. Chapter 4 describes the thermostabilization of P450 BM-3 by directed evolution using a screening procedure similar to the one developed for the thermostabilization of HRP.

B.4. CPO

CPO has broad substrate specificity and does not require NAD(P)H or additional proteins for catalysis. While CPO is highly attractive for synthetic applications [58], protein engineering has been hampered by the inability to express the fungal enzyme in bacteria or yeast. Hager and coworkers, however, have used gene replacement technology to allow functional expression and production of mutants in the enzyme's natural host, *Caldariomyces fumago* [114]. In an effort to understand the importance of the proximal thiolate ligand (Cys29), Yi *et al.* [114] generated the Cys29His mutant and compared the rates of halogenation, peroxidation, epoxidation (oxygen insertion) and catalase activity for the mutant and wild-type. Surprisingly, all four rates decreased only slightly (catalase activity diminished the most – but by only 40%). Recently Conesa *et*

al. achieved functional expression of CPO in a different filamentous fungal host, *Aspergillus niger* [181].

One consequence of heme enzyme-catalyzed epoxidation of terminal olefins is the suicide inactivation of the enzyme due to *N*-alkylation of the prosthetic heme group [182-187]. Inactivation only occurs if the olefin is accepted as a substrate [187]. Neither the epoxide nor the substrate, but rather an active intermediate in the mechanism of substrate turnover, is responsible for heme alkylation [183], indicating that epoxidation and heme alkylation diverge at some point prior to epoxide formation. CPO epoxidizes olefins and is therefore subject to inactivation by primary olefins through this mechanism-based suicide reaction. Using *C. fumago* as their expression system, Rai *et al.* [188] created and screened random mutants of CPO to find ones that are more resistant to inactivation by allylbenzene. Halogenation activity was determined in a monochlorodimedone (MCD) absorbance assay, in which MCD is chlorinated to dichlorodimedone. Peroxidase activity was measured on the colorimetric assay substrate ABTS. Additionally, epoxidation activity was measured towards styrene. Three mutants that are resistant to suicide inactivation by allylbenzene (as well as 1-hexene and 1-heptene) were isolated after three and four rounds of mutagenesis. The fourth round mutant also exhibited improved styrene epoxidation activity. This same group also used directed evolution to improve *p*-nitrostyrene epoxidation activity 8-fold over wild-type [189], and indole oxidation activity was enhanced in 40% aqueous *tert*-butyl alcohol [190]. Unfortunately, using this strain to make and characterize CPO mutants is tedious and problematic [181]. The results from this work demonstrate the power of directed evolution, with a handful of

mutations, to make significant changes in a property that may at first seem a limitation inherent to the system.

Conclusions

Nature provides an arsenal of biocatalysts whose capabilities we are learning to exploit and perfect through protein engineering. The heme enzymes are a rich source of potential industrial biocatalysts and a superb testing ground for protein engineering methods. Where three-dimensional structures are known and previous studies have elucidated the roles of various residues, rational design efforts have proven quite successful. Directed evolution is an established, powerful tool for engineering proteins, particularly when little is known about structure-function relationships, and evolution approaches are already generating customized oxygenases and probing the limits of heme enzyme catalysis. P450s have been designed to oxidize novel substrates and function without requiring biological cofactors and the peroxygenase activity of CPO has been enhanced. Further investigations into mechanisms, structures and functions of heme enzymes and oxygenases remain at the forefront of enzyme research and will continue to provide clues for more fruitful engineering efforts.

Peroxidases efficiently utilize H_2O_2 to catalyze single electron oxidations, CPO efficiently utilizes H_2O_2 to catalyze select oxidations of activated carbons and heteroatoms, and P450s SPα and BSβ efficiently utilize H_2O_2 to hydroxylate fatty acids. There are limitations associated with the value, application and/or protein engineering of all of these enzymes. However, that nature has been able to evolve such enzymes is a strong indication that it should be possible to engineer a biotechnologically relevant

cytochrome P450 to function as peroxygenase, requiring only H_2O_2 to catalyze important and difficult transformations such as the regioselective hydroxylation of unactivated C-H bonds. The remainder of this thesis, along with recent publications [89,191-193], details my contribution to this effort. P450 BM-3 was first evolved to function as a peroxygenase, and then further evolved into a thermostable peroxygenase. This work attests to the "designability" of the P450 BM-3 scaffold and complements previous efforts to engineer novel substrate specificities in P450 BM-3. It is exciting to imagine future P450 BM-3 peroxygenase variants which efficiently hydroxylate alkanes or other "unnatural" substrates of interest.

References

1. Sheldon RA (Ed): *Metalloporphyrins in Catalytic Oxidations.* New York: Marcel Dekker, Inc.; 1994.
2. Erman JE, Hager LP, Sligar SG: **Cytochrome P-450 and peroxidase chemistry**. *Adv Inorg Biochem* 1994, **10**:71-118.
3. Gray HB, Ellis WR: **Electron transfer**. In *Bioinorganic Chemistry*. Edited by Bertini I, Gray HB, Lippard SJ, Valentine JS: University Science Books; 1994:315-363.
4. Phillips IR, Shephard EA (Ed): *Cytochrome P450 Protocols.* Totowa, N.J.: Humana Press; 1998.
5. Lewis DFV: *Cytochromes P450: Structure, Function, and Mechanism*. London: Taylor & Francis; 1996.
6. Ortiz de Montellano PR (Ed): *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd edn. New York and London: Plenum Press; 1995.
7. Wong LL, Westlake CG, Nickerson DP: **Protein engineering of cytochrome P450(cam)**. In *Metal Sites in Proteins and Models, Structure and Bonding*; 1997, **88**:175-207.
8. Wong LL: **Cytochrome P450 monooxygenases**. *Curr Opin Chem Biol* 1998, **2**:263-268.
9. van den Brink HM, van Gorcom RF, van den Hondel CA, Punt PJ: **Cytochrome P450 enzyme systems in fungi**. *Fungal Genet Biol* 1998, **23**:1-17.
10. Roberts GC: **The power of evolution: accessing the synthetic potential of P450s**. *Chem Biol* 1999, **6**:R269-272.
11. Kellner DG, Maves SA, Sligar SG: **Engineering cytochrome P450s for bioremediation**. *Curr Opin Biotechnol* 1997, **8**:274-278.
12. Sono M, Roach MP, Coulter ED, Dawson JH: **Heme-containing oxygenases**. *Chem Rev* 1996, **96**:2841-2887.
13. Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, et al.: **P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature**. *Pharmacogenetics* 1996, **6**:1-42.
14. Martinez CA, Stewart JD: **Cytochrome P450's: Potential catalysts for asymmetric olefin epoxidations**. *Curr Org Chem* 2000, **4**:263-282.
15. Mansuy D: **The great diversity of reactions catalyzed by cytochromes P450**. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 1998, **121**:5-14.
16. Miles CS, Ost TWB, Noble MA, Munro AW, Chapman SK: **Protein engineering of cytochromes P-450**. *Biochim Biophys Acta-Protein Struct Molec Enzym* 2000, **1543**:383-407.
17. Mueller EJ, Loida PJ, and Sligar SG: **Twenty-five years of P450cam research**. In *Cytochrome P450 Structure, Mechanism, and Biochemistry,* 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:83-124.
18. Poulos TL, Finzel BC, Howard AJ: **High-resolution crystal structure of cytochrome P450cam**. *J Mol Biol* 1987, **195**:687-700.
19. Ravichandran KG, Boddupalli SS, Hasermann CA, Peterson JA, Deisenhofer J: **Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's**. *Science* 1993, **261**:731-736.

20. Hasemann CA, Ravichandran KG, Peterson JA, Deisenhofer J: **Crystal structure and refinement of cytochrome P450terp at 2.3 Å resolution**. *J Mol Biol* 1994, **236**:1169-1185.
21. Cupp-Vickery JR, Poulos TL: **Structure of cytochrome P450eryF involved in erythromycin biosynthesis**. *Nat Struct Biol* 1995, **2**:144-153.
22. Park SY, Shimizu H, Adachi S, Nakagawa A, Tanaka I, Nakahara K, Shoun H, Obayashi E, Nakamura H, Iizuka T, et al.: **Crystal structure of nitric oxide reductase from denitrifying fungus *Fusarium oxysporum***. *Nat Struct Biol* 1997, **4**:827-832.
23. Park SY, Yamane K, Adachi S, Shiro Y, Weiss KE, Sligar SG: **Crystallization and preliminary X-ray diffraction analysis of a cytochrome P450 (CYP119) from *Sulfolobus solfataricus***. *Acta Crystallogr Sect D-Biol Crystallogr* 2000, **56**:1173-1175.
24. Ito S, Matsuoka T, Watanabe I, Kagasaki T, Serizawa N, Hata T: **Crystallization and preliminary X-ray diffraction analysis of cytochrome P450(sca-2) from *Streptomyces carbophilus* involved in production of pravastatin sodium, a tissue-selective inhibitor of HMG-CoA reductase**. *Acta Crystallogr Sect D-Biol Crystallogr* 1999, **55**:1209-1211.
25. Podust LM, Poulos TL, Waterman MR: **Crystal structure of cytochrome P450 14alpha-sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors**. *Proc Natl Acad Sci USA* 2001, **98**:3068-3073.
26. Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE: **Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity**. *Mol Cell* 2000, **5**:121-131.
27. Sligar SG, Gunsalus IC: **A thermodynamic model of regulation: modulation of redox equilibria in camphor monoxygenase**. *Proc Natl Acad Sci USA* 1976, **73**:1078-1082.
28. Fisher MT, Sligar SG: **Control of heme protein redox potential and reduction rate - linear free-energy relation between potential and ferric spin state equilibrium**. *J Am Chem Soc* 1985, **107**:5018-5019.
29. Sligar SG: **Coupling of spin, substrate, and redox equilibria in cytochrome P450**. *Biochemistry* 1976, **15**:5399-5406.
30. Toy PH, Newcomb M, Coon MJ, Vaz ADN: **Two distinct electrophilic oxidants effect hydroxylation in cytochrome P-450-catalyzed reactions**. *J Am Chem Soc* 1998, **120**:9718-9719.
31. Vaz AD, McGinnity DF, Coon MJ: **Epoxidation of olefins by cytochrome P450: evidence from site-specific mutagenesis for hydroperoxo-iron as an electrophilic oxidant**. *Proc Natl Acad Sci USA* 1998, **95**:3555-3560.
32. Vaz AD, Pernecky SJ, Raner GM, Coon MJ: **Peroxo-iron and oxenoid-iron species as alternative oxygenating agents in cytochrome P450-catalyzed reactions: switching by threonine-302 to alanine mutagenesis of cytochrome P450 2B4**. *Proc Natl Acad Sci USA* 1996, **93**:4644-4648.
33. Kimata Y, Shimada H, Hirose T, Ishimura Y: **Role of Thr-252 in cytochrome P450cam: a study with unnatural amino acid mutagenesis**. *Biochem Biophys Res Commun* 1995, **208**:96-102.

34. Aikens J, Sligar SG: **Kinetic solvent isotope effects during oxygen activation by cytochrome P-450cam**. *J Am Chem Soc* 1994, **116**:1143-1144.
35. Raag R, Martinis SA, Sligar SG, Poulos TL: **Crystal structure of the cytochrome P-450cam active site mutant Thr252Ala**. *Biochemistry* 1991, **30**:11420-11429.
36. Martinis SA, Atkins WM, Stayton PS, Sligar SG: **A conserved residue of cytochrome-P-450 is involved in heme-oxygen stability and activation**. *J Am Chem Soc* 1989, **111**:9252-9253.
37. Imai M, Shimada H, Watanabe Y, Matsushima-Hibiya Y, Makino R, Koga H, Horiuchi T, Ishimura Y: **Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation**. *Proc Natl Acad Sci USA* 1989, **86**:7823-7827.
38. Fruetel JA, Collins JR, Camper DL, Loew GH, Ortiz de Montellano PR: **Calculated and experimental absolute stereochemistry of the styrene and beta-methylstyrene epoxides formed by cytochrome P450(Cam)**. *J Am Chem Soc* 1992, **114**:6987-6993.
39. Nickerson DP, Harford-Cross CF, Fulcher SR, Wong LL: **The catalytic activity of cytochrome P450cam towards styrene oxidation is increased by site-specific mutagenesis**. *FEBS Lett* 1997, **405**:153-156.
40. Loida PJ, Sligar SG: **Molecular recognition in cytochrome P-450: mechanism for the control of uncoupling reactions**. *Biochemistry* 1993, **32**:11530-11538.
41. Hintz MJ, Mock DM, Peterson LL, Tuttle K, Peterson JA: **Equilibrium and kinetic studies of the interaction of cytochrome P-450cam and putidaredoxin**. *J Biol Chem* 1982, **257**:14324-14332.
42. Li QS, Ogawa J, Shimizu S: **Critical role of the residue size at position 87 in H2O2-dependent substrate hydroxylation activity and H2O2 inactivation of cytochrome P450BM-3**. *Biochem Biophys Res Commun* 2001, **280**:1258-1261.
43. Yu XC, Liang C, Strobel HW: **Kinetics of substrate reaction in the course of hydroperoxide-mediated inactivation of cytochrome P450 1A1**. *Biochemistry* 1996, **35**:6289-6296.
44. Anari MR, Khan S, Liu ZC, O'Brien PJ: **Cytochrome P450 peroxidase/ peroxygenase mediated xenobiotic metabolic activation and cytotoxicity in isolated hepatocytes**. *Chem Res Toxicol* 1995, **8**:997-1004.
45. Coon MJ, Blake RC, 2nd, White RE, Nordblom GD: **Assays for cytochrome P-450 peroxygenase activity**. *Methods Enzymol* 1990, **186**:273-278.
46. Estabrook RW, Martin-Wixtrom C, Saeki Y, Renneberg R, Hildebrandt A, Werringloer J: **The peroxidatic function of liver microsomal cytochrome P-450: comparison of hydrogen peroxide and NADPH-catalysed N-demethylation reactions**. *Xenobiotica* 1984, **14**:87-104.
47. McCarthy MB, White RE: **Functional differences between peroxidase compound I and the cytochrome P-450 reactive oxygen intermediate**. *J Biol Chem* 1983, **258**:9153-9158.
48. Blake RC, Coon MJ: **On the mechanism of action of cytochrome P-450. Spectral intermediates in the reaction of P-450LM2 with peroxy compounds**. *J Biol Chem* 1980, **255**:4100-4111.

49. White RE, Sligar SG, Coon MJ: **Evidence for a homolytic mechanism of peroxide oxygen-oxygen bond cleavage during substrate hydroxylation by cytochrome P-450**. *J Biol Chem* 1980, **255**:11108-11111.
50. Capdevila J, Estabrook RW, Prough RA: **Differences in the mechanism of NADPH- and cumene hydroperoxide-supported reactions of cytochrome P-450**. *Arch Biochem Biophys* 1980, **200**:186-195.
51. Rahimtula AD, O'Brien PJ: **Hydroperoxide catalyzed liver microsomal aromatic hydroxylation reactions involving cytochrome P-450**. *Biochem Biophys Res Commun* 1974, **60**:440-447.
52. Rahimtula AD, O'Brien PJ, Seifried HE, Jerina DM: **The mechanism of action of cytochrome P-450. Occurrence of the 'NIH shift' during hydroperoxide-dependent aromatic hydroxylations**. *Eur J Biochem* 1978, **89**:133-141.
53. Nordblom GD, White RE, Coon MJ: **Studies on hydroperoxide-dependent substrate hydroxylation by purified liver microsomal cytochrome P-450**. *Arch Biochem Biophys* 1976, **175**:524-533.
54. Hrycay EG, Gustafsson JA, Ingelman-Sundberg M, Ernster L: **Sodium periodate, sodium chloride, organic hydroperoxides, and H2O2 as hydroxylating agents in steroid hydroxylation reactions catalyzed by partially purified cytochrome P-450**. *Biochem Biophys Res Commun* 1975, **66**:209-216.
55. Hrycay EG, O'Brien PJ: **The peroxidase nature of cytochrome P-420 utilizing a lipid peroxide substrate**. *Arch Biochem Biophys* 1971, **147**:28-35.
56. Peterson JA, Ishimura Y, Griffin BW: ***Pseudomonas putida* cytochrome P-450: characterization of an oxygenated form of the hemoprotein**. *Arch Biochem Biophys* 1972, **149**:197-208.
57. vanDeurzen MPJ, vanRantwijk F, Sheldon RA: **Selective oxidations catalyzed by peroxidases**. *Tetrahedron* 1997, **53**:13183-13220.
58. van Rantwijk F, Sheldon RA: **Selective oxygen transfer catalysed by heme peroxidases: synthetic and mechanistic aspects**. *Curr Opin Biotechnol* 2000, **11**:554-564.
59. Adam W, Lazarus M, Saha-Moller CR, Weichold O, Hoch U, Haring D, Schreier P: **Biotransformations with peroxidases**. *Adv Biochem Eng Biotechnol* 1999, **63**:73-108.
60. Colonna S, Gaggero N, Richelmi C, Pasta P: **Recent biotechnological developments in the use of peroxidases**. *Trends Biotechnol* 1999, **17**:163-168.
61. Everse J, Everse KE, Grisham MB (Eds): *Peroxidases in Chemistry and Biology.* Boca Raton, FL: CRC Press; 1991.
62. Kyte J: *Mechanism in Protein Chemistry*. New York and London: Garland Publishing, Inc.; 1995.
63. Poulos TL: **Heme enzyme crystal structures**. *Adv Inorg Biochem* 1988, **7**:1-36.
64. Smith AT, Veitch NC: **Substrate binding and catalysis in heme peroxidases**. *Curr Opin Chem Biol* 1998, **2**:269-278.
65. Matsui T, Ozaki S, Liong E, Phillips GN, Watanabe Y: **Effects of the location of distal histidine in the reaction of myoglobin with hydrogen peroxide**. *J Biol Chem* 1999, **274**:2838-2844.

66. Ozaki S, Matsui T, Roach MP, Watanabe Y: **Rational molecular design of a catalytic site: engineering of catalytic functions to the myoglobin active site framework**. *Coord Chem Rev* 2000, **198**:39-59.
67. Savenkova MI, Kuo JM, Ortiz de Montellano PR: **Improvement of peroxygenase activity by relocation of a catalytic histidine within the active site of horseradish peroxidase**. *Biochemistry* 1998, **37**:10828-10836.
68. Savenkova MI, Newmyer SL, Ortiz de Montellano PR: **Rescue of His-42->Ala horseradish peroxidase by a Phe-41->His mutation - Engineering of a surrogate catalytic histidine**. *J Biol Chem* 1996, **271**:24598-24603.
69. Newmyer SL, Ortiz de Montellano PR: **Mutagenesis of horseradish peroxidase alters peroxidase and peroxygenase activities**. *Faseb Journal* 1995, **9**:A1353-A1353.
70. Newmyer SL, Ortiz de Montellano PR: **Horseradish peroxidase His42Ala, His42Val, and Phe41Ala mutants - histidine catalysis and control of substrate access to the heme iron**. *J Biol Chem* 1995, **270**:19430-19438.
71. Dawson JH: **Probing structure-function relations in heme-containing oxygenases and peroxidases**. *Science* 1988, **240**:433-439.
72. Marnett LJ, Weller, P., and Battista, J.R.: **Comparison of the peroxidase activity of hemeproteins and cytochrome P-450**. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 1st edn. Edited by Ortiz de Montellano PR: Plenum Press; 1986:29-77.
73. Dunford HB, Stillman JS: **On the function and mechanism of action of peroxidases.** *Coord Chem Rev*; 1976:187-125.
74. Dunford HB, Nadezhdin AD: **On the past eight years of peroxidase research**. In *Oxidases and Related Redox Systems*. Edited by King TE, Mason HS, Morrison M: Pergamon; 1982:653-670.
75. Dunford HB, Adeniran AJ: **Hammett rho sigma correlation for reactions of horseradish peroxidase compound II with phenols**. *Arch Biochem Biophys* 1986, **251**:536-542.
76. Sakurada J, Sekiguchi R, Sato K, Hosoya T: **Kinetic and molecular orbital studies on the rate of oxidation of monosubstituted phenols and anilines by horseradish peroxidase compound II**. *Biochemistry* 1990, **29**:4093-4098.
77. Pietikainen P, Adlercreutz P: **Influence of the reaction medium on the product distribution of peroxidase-catalysed oxidation of p-cresol**. *Appl Microbiol Biotechnol* 1990, **33**:455-458.
78. Dordick JS, Klibanov AM, Marletta MA: **Horseradish peroxidase catalyzed hydroxylations: mechanistic studies**. *Biochemistry* 1986, **25**:2946-2951.
79. Hewson WD, Dunford HB: **Stoichiometry of the reaction between horseradish peroxidase and p-cresol**. *J Biol Chem* 1976, **251**:6043-6052.
80. Kobayashi S, Shoda S, Uyama H: **Enzymatic polymerization and oligomerization**. *Advances in Polymer Science* 1995, **121**:1-30.
81. Dordick JS: **Enzymatic and chemoenzymatic approaches to polymer synthesis**. *Trends Biotechnol* 1992, **10**:287-293.
82. Klibanov AM, Tu TM, Scott KP: **Peroxidase-catalyzed removal of phenols from coal-conversion waste-waters**. *Science* 1983, **221**:259-260.

83. Sun W, Kadima TA, Pickard MA, Dunford HB: **Catalase activity of chloroperoxidase and its interaction with peroxidase activity**. *Biochem Cell Biol* 1994, **72**:321-331.
84. Truan G, Peterson JA: **Thr268 in substrate binding and catalysis in P450BM-3**. *Arch Biochem Biophys* 1998, **349**:53-64.
85. Macdonald TL, Gutheim WG, Martin RB, Guengerich FP: **Oxidation of substituted N,N-dimethylanilines by cytochrome P-450: estimation of the effective oxidation-reduction potential of cytochrome P-450**. *Biochemistry* 1989, **28**:2071-2077.
86. Hayashi Y, Yamazaki I: **The oxidation-reduction potentials of compound I/compound II and compound II/ferric couples of horseradish peroxidases A2 and C**. *J Biol Chem* 1979, **254**:9101-9106.
87. Casella L, Gullotti M, Ghezzi R, Poli S, Beringhelli T, Colonna S, Carrea G: **Mechanism of enantioselective oxygenation of sulfides catalyzed by chloroperoxidase and horseradish peroxidase. Spectral studies and characterization of enzyme-substrate complexes**. *Biochemistry* 1992, **31**:9451-9459.
88. Marnett LJ, Kennedy TA: **Comparison of the peroxidase activity of hemoproteins and cytochrome P450**. In *Cytochrome P450 Structure, Mechanism and Biochemistry*, 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:49-80.
89. Cirino PC, Arnold FH: **A self-sufficient peroxide-driven hydroxylation biocatalyst**. *Angew Chem-Int Edit* 2003, **In Press**.
90. Smith AT, Sanders SA, Thorneley RN, Burke JF, Bray RR: **Characterisation of a haem active-site mutant of horseradish peroxidase, Phe41Val, with altered reactivity towards hydrogen peroxide and reducing substrates**. *Eur J Biochem* 1992, **207**:507-519.
91. Erman JE, Vitello LB, Miller MA, Kraut J: **Active-site mutations in cytochrome-C peroxidase - a critical role for histidine-52 in the rate of formation of *Compound I***. *J Am Chem Soc* 1992, **114**:6592-6593.
92. Matsunaga I, Yokotani N, Gotoh O, Kusunose E, Yamada M, Ichihara K: **Molecular cloning and expression of fatty acid alpha-hydroxylase from *Sphingomonas paucimobilis***. *J Biol Chem* 1997, **272**:23592-23596.
93. Matsunaga I, Ueda A, Fujiwara N, Sumimoto T, Ichihara K: **Characterization of the ybdT gene product of *Bacillus subtilis*: novel fatty acid beta-hydroxylating cytochrome P450**. *Lipids* 1999, **34**:841-846.
94. Matsunaga I, Sumimoto T, Ueda A, Kusunose E, Ichihara K: **Fatty acid-specific, regiospecific, and stereospecific hydroxylation by cytochrome P450 (CYP152B1) from *Sphingomonas paucimobilis*: substrate structure required for alpha-hydroxylation**. *Lipids* 2000, **35**:365-371.
95. Imai Y, Matsunaga I, Kusunose E, Ichihara K: **Unique heme environment at the putative distal region of hydrogen peroxide-dependent fatty acid alpha-hydroxylase from *Sphingomonas paucimobilis* (peroxygenase P450(SP alpha))**. *J Biochem* 2000, **128**:189-194.

96. Matsunaga I, Ueda A, Sumimoto T, Ichihara K, Ayata M, Ogura H: **Site-directed mutagenesis of the putative distal helix of peroxygenase cytochrome P450**. *Arch Biochem Biophys* 2001, **394**:45-53.
97. Lee DS, Yamada A, Sugimoto H, Matsunaga I, Ogura H, Ichihara K, Adachi S, Park SY, Shiro Y: **Substrate recognition and molecular mechanism of fatty acid hydroxylation by cytochrome P450 from *Bacillus subtilis*. Crystallographic, spectroscopic, and mutational studies**. *J Biol Chem* 2003, **278**:9761-9767.
98. Matsunaga I, Yamada M, Kusunose E, Miki T, Ichihara K: **Further characterization of hydrogen peroxide dependent fatty acid alpha-hydroxylase from *Sphingomonas paucimobilis***. *J Biochem* 1998, **124**:105-110.
99. Mayhew MP, Reipa V, Holden MJ, Vilker VL: **Improving the cytochrome P450 enzyme system for electrode-driven biocatalysis of styrene epoxidation**. *Biotechnol Prog* 2000, **16**:610-616.
100. Reipa V, Mayhew MP, Vilker VL: **A direct electrode-driven P450 cycle for biocatalysis**. *Proc Natl Acad Sci USA* 1997, **94**:13554-13558.
101. Estabrook RW, Faulkner KM, Shet MS, Fisher CW: **Application of electrochemistry for P450-catalyzed reactions**. *Methods Enzymol* 1996, **272**:44-51.
102. Faulkner KM, Shet MS, Fisher CW, Estabrook RW: **Electrocatalytically driven omega-hydroxylation of fatty acids using cytochrome P450 4A1**. *Proc Natl Acad Sci USA* 1995, **92**:7705-7709.
103. Yano JK, Koo LS, Schuller DJ, Li HY, Ortiz de Montellano PR, Poulos TL: **Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus***. *J Biol Chem* 2000, **275**:31086-31092.
104. Koo LS, Tschirret-Guth RA, Straub WE, Moenne-Loccoz P, Loehr TM, Ortiz de Montellano PR: **The active site of the thermophilic CYP119 from *Sulfolobus solfataricus***. *J Biol Chem* 2000, **275**:14112-14123.
105. Chang YT, Loew G: **Homology modeling, molecular dynamics simulations, and analysis of CYP119, a P450 enzyme from extreme acidothermophilic archaeon *Sulfolobus solfataricus***. *Biochemistry* 2000, **39**:2484-2498.
106. McLean MA, Maves SA, Weiss KE, Krepich S, Sligar SG: **Characterization of a cytochrome P450 from the acidothermophilic archaea *Sulfolobus solfataricus***. *Biochem Biophys Res Commun* 1998, **252**:166-172.
107. Park SY, Yamane K, Adachi S, Shiro Y, Weiss KE, Maves SA, Sligar SG: **Thermophilic cytochrome P450 (CYP119) from *Sulfolobus solfataricus*: high resolution structure and functional properties**. *J Inorg Biochem* 2002, **91**:491-501.
108. Sundaramoorthy M, Terner J, Poulos TL: **The crystal structure of chloroperoxidase: A heme peroxidase- cytochrome P450 functional hybrid**. *Structure* 1995, **3**:1367-1377.
109. Hager LP, Lakner FJ, Basavapathruni A: **Chiral synthons via chloroperoxidase catalysis**. *J Mol Catal B-Enzym* 1998, **5**:95-101.
110. vanDeurzen MPJ, Remkes IJ, vanRantwijk F, Sheldon RA: **Chloroperoxidase catalyzed oxidations in t-butyl alcohol/water mixtures**. *J Mol Catal A -Chem* 1997, **117**:329-337.

111. Colonna S, Gaggero N, Carrea G, Pasta P: **A new enzymatic enantioselective synthesis of dialkyl sulfoxides catalysed by monooxygenases**. *Chem Commun* 1997:439-440.
112. Geigert J, Dalietos DJ, Neidleman SL, Lee TD, Wadsworth J: **Peroxide oxidation of primary alcohols to aldehydes by chloroperoxidase catalysis**. *Biochem Biophys Res Commun* 1983, **114**:1104-1108.
113. Thomas JA, Morris DR, Hager LP: **Chloroperoxidase. VII. Classical peroxidatic, catalatic, and halogenating forms of the enzyme**. *J Biol Chem* 1970, **245**:3129-3134.
114. Yi X, Mroczko M, Manoj KM, Wang X, Hager LP: **Replacement of the proximal heme thiolate ligand in chloroperoxidase with a histidine residue**. *Proc Natl Acad Sci USA* 1999, **96**:12412-12417.
115. Hu SH, Hager LP: **Highly enantioselective propargylic hydroxylations catalyzed by chloroperoxidase**. *J Am Chem Soc* 1999, **121**:872-873.
116. Hu SH, Hager LP: **Asymmetric epoxidation of functionalized cis-olefins catalyzed by chloroperoxidase**. *Tetrahedron Letters* 1999, **40**:1641-1644.
117. Kedderis GL, Hollenberg PF: **Peroxidase-catalyzed N-demethylation reactions. Substrate deuterium isotope effects**. *J Biol Chem* 1984, **259**:3663-3668.
118. Kedderis GL, Koop DR, Hollenberg PF: **N-Demethylation reactions catalyzed by chloroperoxidase**. *J Biol Chem* 1980, **255**:10174-10182.
119. vanDeurzen MPJ, Seelbach K, VanRantwijk F, Kragl U, Sheldon RA: **Chloroperoxidase: Use of a hydrogen peroxide-stat for controlling reactions and improving enzyme performance**. *Biocatal Biotransform* 1997, **15**:1-16.
120. van de Velde F, Lourenco ND, Bakker M, van Rantwijk F, Sheldon RA: **Improved operational stability of peroxidases by coimmobilization with glucose oxidase**. *Biotechnol Bioeng* 2000, **69**:286-291.
121. van de Velde F, van Rantwijk F, Sheldon RA: **Selective oxidations with molecular oxygen, catalyzed by chloroperoxidase in the presence of a reductant**. *J Mol Catal B-Enzym* 1999, **6**:453-461.
122. vanDeurzen MPJ, Groen BW, Vanrantwijk F, Sheldon RA: **A simple purification method for chloroperoxidase and its use in organic media**. *Biocatalysis* 1994, **10**:247-255.
123. Seelbach K, vanDeurzen MPJ, vanRantwijk F, Sheldon RA, Kragl U: **Improvement of the total turnover number and space-time yield for chloroperoxidase catalyzed oxidation**. *Biotechnol Bioeng* 1997, **55**:283-288.
124. Loida PJ, Sligar SG: **Engineering cytochrome P-450cam to increase the stereospecificity and coupling of aliphatic hydroxylation**. *Protein Eng* 1993, **6**:207-212.
125. Bell SG, Rouch DA, Wong LL: **Selective aliphatic and aromatic carbon-hydrogen bond activation catalysed by mutants of cytochrome P450(cam)**. *J Mol Catal B-Enzym* 1997, **3**:293-302.
126. Fowler SM, England PA, Westlake ACG, Rouch DR, Nickerson DP, Blunt C, Braybrook D, West S, Wong LL, Flitsch SL: **Cytochrome P-450(cam) monooxygenase can be redesigned to catalyze the regioselective aromatic hydroxylation of diphenylmethane**. *J Chem Soc-Chem Commun* 1994:2761-2762.

127. Jones JP, O'Hare EJ, Wong LL: **Oxidation of polychlorinated benzenes by genetically engineered CYP101 (cytochrome P450(cam))**. *Eur J Biochem* 2001, **268**:1460-1467.
128. Bell SG, Sowden RJ, Wong LL: **Engineering the haem monooxygenase cytochrome P450(cam) for monoterpene oxidation**. *Chem Commun* 2001:635-636.
129. Joo H, Lin Z, Arnold FH: **Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation**. *Nature* 1999, **399**:670-673.
130. Schwaneberg U, Schmidt-Dannert C, Schmitt J, Schmid RD: **A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A**. *Anal Biochem* 1999, **269**:359-366.
131. Schwaneberg U, Otey C, Cirino PC, Farinas E, Arnold FH: **Cost-effective whole-cell assay for laboratory evolution of hydroxylases in *Escherichia coli***. *J Biomol Screen* 2001, **6**:111-117.
132. Farinas ET, Schwaneberg U, Glieder A, Arnold FH: **Directed evolution of a cytochrome P450 monooxygenase for alkane oxidation**. *Adv Synth Catal* 2001, **343**:601-606.
133. Graham-Lorence S, Truan G, Peterson JA, Falck JR, Wei S, Helvig C, Capdevila JH: **An active site substitution, F87V, converts cytochrome P450 BM-3 into a regio- and stereoselective (14S,15R)-arachidonic acid epoxygenase**. *J Biol Chem* 1997, **272**:1127-1135.
134. Capdevila JH, Wei S, Helvig C, Falck JR, Belosludtsev Y, Truan G, Graham-Lorence SE, Peterson JA: **The highly stereoselective oxidation of polyunsaturated fatty acids by cytochrome P450BM-3**. *J Biol Chem* 1996, **271**:22663-22671.
135. Boddupalli SS, Estabrook RW, Peterson JA: **Fatty acid monooxygenation by cytochrome P450BM-3**. *J Biol Chem* 1990, **265**:4233-4239.
136. Ruettinger RT, Wen LP, Fulco AJ: **Coding nucleotide, 5' regulatory, and deduced amino acid sequences of P450BM-3, a single peptide cytochrome P-450:NADPH-P-450 reductase from *Bacillus megaterium***. *J Biol Chem* 1989, **264**:10987-10995.
137. Oliver CF, Modi S, Sutcliffe MJ, Primrose WU, Lian LY, Roberts GC: **A single mutation in cytochrome P450 BM3 changes substrate orientation in a catalytic intermediate and the regiospecificity of hydroxylation**. *Biochemistry* 1997, **36**:1567-1572.
138. Li QS, Ogawa J, Schmid RD, Shimizu S: **Residue size at position 87 of cytochrome P450BM-3 determines its stereoselectivity in propylbenzene and 3-chlorostyrene oxidation**. *Febs Letters* 2001, **508**:249-252.
139. Modi S, Sutcliffe MJ, Primrose WU, Lian LY, Roberts GC: **The catalytic mechanism of cytochrome P450 BM3 involves a 6 Å movement of the bound substrate on reduction**. *Nat Struct Biol* 1996, **3**:414-417.
140. Li H, Poulos TL: **The structure of the cytochrome P450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid**. *Nat Struct Biol* 1997, **4**:140-146.

141. Li QS, Schwaneberg U, Fischer P, Schmid RD: **Directed evolution of the fatty-acid hydroxylase P450 BM-3 into an indole-hydroxylating catalyst**. *Chem-Eur J* 2000, **6**:1531-1536.
142. Li QS, Schwaneberg U, Fischer M, Schmitt J, Pleiss J, Lutz-Wahl S, Schmid RD: **Rational evolution of a medium chain-specific cytochrome P-450 BM-3 variant**. *Biochim Biophys Acta-Protein Struct Molec Enzym* 2001, **1545**:114-121.
143. Lentz O, Qing-Shang LI, Schwaneberg U, Lutz-Wahl S, Fischer P, Schmid RD: **Modification of the fatty acid specificity of cytochrome P450BM-3 from *Bacillus megaterium* by directed evolution: A validated assay**. *J Mol Catal B-Enzym* 2001, **15**:123-133.
144. Appel D, Lutz-Wahl S, Fischer P, Schwaneberg U, Schmid RD: **A P450BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes**. *J Biotechnol* 2001, **88**:167-171.
145. Li QS, Ogawa J, Schmid RD, Shimizu S: **Engineering cytochrome P450BM-3 for oxidation of polycyclic aromatic hydrocarbons**. *Appl Environ Microbiol* 2001, **67**:5735-5739.
146. Carmichael AB, Wong LL: **Protein engineering of *Bacillus megaterium* CYP102 - The oxidation of polycyclic aromatic hydrocarbons**. *Eur J Biochem* 2001, **268**:3117-3125.
147. Ost TWB, Miles CS, Murdoch J, Cheung YF, Reid GA, Chapman SK, Munro AW: **Rational re-design of the substrate binding site of flavocytochrome P450BM3**. *Febs Letters* 2000, **486**:173-177.
148. Glieder A, Farinas ET, Arnold FH: **Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase**. *Nat Biotechnol* 2002, **20**:1135-1139.
149. Deeni YY, Paine MJI, Ayrton AD, Clarke SE, Chenery R, Wolf CR: **Expression, purification, and biochemical characterization of a human cytochrome P450CYP2D6-NADPH cytochrome P450 reductase fusion protein**. *Arch Biochem Biophys* 2001, **396**:16-24.
150. Dohr O, Paine MJI, Friedberg T, Roberts GCK, Wolf CR: **Engineering of a functional human NADH-dependent cytochrome P450 system**. *Proc Natl Acad Sci USA* 2001, **98**:81-86.
151. Sadeghi SJ, Meharenna YT, Fantuzzi A, Valetti F, Gilardi G: **Engineering artificial redox chains by molecular 'Lego'**. *Faraday Discuss* 2000:135-153.
152. Daff S, Noble MA, Craig DH, Rivers SL, Chapman SK, Munro AW, Fujiwara S, Rozhkova E, Sagami I, Shimizu T: **Control of electron transfer in neuronal NO synthase**. *Biochem Soc Trans* 2001, **29**:147-152.
153. Domanski TL, Halpert JR: **Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis**. *Curr Drug Metab* 2001, **2**:117-137.
154. Sakaki T, Inouye K: **Practical application of mammalian cytochrome P450**. *J Biosci Bioeng* 2000, **90**:583-590.
155. Nakamura K, Martin MV, Guengerich FP: **Random mutagenesis of human cytochrome P450 2A6 and screening with indole oxidation products**. *Arch Biochem Biophys* 2001, **395**:25-31.

156. Urban P, Jobert AS, Laine R, Pompon D: **Cytochrome P450 (CYP) mutants and substrate-specificity alterations: segment-directed mutagenesis applied to human CYP1A1**. *Biochem Soc Trans* 2001, **29**:128-135.
157. Yun CH, Miller GP, Guengerich FP: **Rate-determining steps in phenacetin oxidations by human cytochrome P450 1A2 and selected mutants**. *Biochemistry* 2000, **39**:11319-11329.
158. Abecassis V, Pompon D, Truan G: **High efficiency family shuffling based on multi-step PCR and in vivo DNA recombination in yeast: statistical and functional analysis of a combinatorial library between human cytochrome P450 1A1 and 1A2**. *Nucleic Acids Res* 2000, **28**:E88.
159. Cherry JR, Lamsa MH, Schneider P, Vind J, Svendsen A, Jones A, Pedersen AH: **Directed evolution of a fungal peroxidase**. *Nat Biotechnol* 1999, **17**:379-384.
160. Bhaskar B, Bonagura CA, Jamal J, Poulos TL: **Loop stability in the engineered potassium binding site of cytochrome c peroxidase**. *Tetrahedron* 2000, **56**:9471-9475.
161. Hirst J, Goodin DB: **Unusual oxidative chemistry of N(omega)-hydroxyarginine and N-hydroxyguanidine catalyzed at an engineered cavity in a heme peroxidase**. *J Biol Chem* 2000, **275**:8582-8591.
162. Bonagura CA, Bhaskar B, Sundaramoorthy M, Poulos TL: **Conversion of an engineered potassium-binding site into a calcium-selective site in cytochrome c peroxidase**. *J Biol Chem* 1999, **274**:37827-37833.
163. Wilcox SK, Putnam CD, Sastry M, Blankenship J, Chazin WJ, McRee DE, Goodin DB: **Rational design of a functional metalloenzyme: introduction of a site for manganese binding and oxidation into a heme peroxidase**. *Biochemistry* 1998, **37**:16853-16862.
164. Yeung BK, Wang X, Sigman JA, Petillo PA, Lu Y: **Construction and characterization of a manganese-binding site in cytochrome c peroxidase: towards a novel manganese peroxidase**. *Chem Biol* 1997, **4**:215-221.
165. Bonagura CA, Sundaramoorthy M, Pappa HS, Patterson WR, Poulos TL: **An engineered cation site in cytochrome c peroxidase alters the reactivity of the redox active tryptophan**. *Biochemistry* 1996, **35**:6107-6115.
166. Wilcox SK, Jensen GM, Fitzgerald MM, McRee DE, Goodin DB: **Altering substrate specificity at the heme edge of cytochrome c peroxidase**. *Biochemistry* 1996, **35**:4858-4866.
167. DePillis GD, Sishta BP, Mauk AG, Ortiz de Montellano PR: **Small substrates and cytochrome c are oxidized at different sites of cytochrome c peroxidase**. *J Biol Chem* 1991, **266**:19334-19341.
168. Pelletier H, Kraut J: **Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c**. *Science* 1992, **258**:1748-1755.
169. Casella L, Monzani E, Gullotti M, Santelli E, Poli S, Beringhelli T: **The oxidation of phenolic compounds catalysed by cytochrome C peroxidase. Comparison with other peroxidases**. *Gazz Chim Ital* 1996, **126**:121-125.
170. Iffland A, Tafelmeyer P, Saudan C, Johnsson K: **Directed molecular evolution of cytochrome c peroxidase**. *Biochemistry* 2000, **39**:10790-10798.

171. Smith AT, Santama N, Dacey S, Edwards M, Bray RC, Thorneley RN, Burke JF: **Expression of a synthetic gene for horseradish peroxidase C in *Escherichia coli* and folding and activation of the recombinant enzyme with Ca2+ and heme**. *J Biol Chem* 1990, **265**:13335-13343.
172. Gray JS, Yang BY, Montgomery R: **Heterogeneity of glycans at each N-glycosylation site of horseradish peroxidase**. *Carbohydr Res* 1998, **311**:61-69.
173. Dunford HB: **Horseradish Peroxidase: Structure and Kinetic Properties**. In *Peroxidases in Chemistry and Biology*. Edited by Everse J, Everse KE, Grisham MB: CRC Press; 1991:1-24.
174. Welinder KG: **Amino acid sequence studies of horseradish peroxidase. Amino and carboxyl termini, cyanogen bromide and tryptic fragments, the complete sequence, and some structural characteristics of horseradish peroxidase C**. *Eur J Biochem* 1979, **96**:483-502.
175. Romanos MA, Scorer CA, Clare JJ: **Foreign gene expression in yeast - a review**. *Yeast* 1992, **8**:423-488.
176. Parekh R, Forrester K, Wittrup D: **Multicopy overexpression of bovine pancreatic trypsin inhibitor saturates the protein-folding and secretory capacity of *Saccharomyces cerevisiae***. *Protein Expr Purif* 1995, **6**:537-545.
177. Sudbery PE: **The expression of recombinant proteins in yeasts**. *Curr Opin Biotechnol* 1996, **7**:517-524.
178. Morawski B, Lin Z, Cirino P, Joo H, Bandara G, Arnold FH: **Functional expression of horseradish peroxidase in *Saccharomyces cerevisiae* and *Pichia pastoris***. *Protein Eng* 2000, **13**:377-384.
179. Morawski B, Quan S, Arnold FH: **Functional expression and stabilization of horseradish peroxidase by directed evolution in *Saccharomyces cerevisiae***. *Biotechnol Bioeng* 2001, **76**:99-107.
180. Yano JK, Blasco F, Li H, Schmid RD, Henne A, Poulos TL: **Preliminary characterization and crystal structure of a thermostable cytochrome P450 from *Thermus thermophilus***. *J Biol Chem* 2003, **278**:608-616.
181. Conesa A, van de Velde F, van Rantwijk F, Sheldon RA, van den Hondel C, Punt PJ: **Expression of the *Caldariomyces fumago* chloroperoxidase in *Aspergillus niger* and characterization of the recombinant enzyme**. *J Biol Chem* 2001, **276**:17635-17640.
182. Groves JT, Fish KM, Avaria-Neisser GE, Imachi M, Kuczkowski RL: **A unique deuterium/proton exchange during cytochrome P-450 mediated expoxidation of propene and butene**. *Prog Clin Biol Res* 1988, **274**:509-524.
183. Groves JT, Han YZ: **Models and mechanisms of cytochrome P450 action**. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:3-48.
184. Tian ZQ, Richards JL, Traylor TG: **Formation of both primary and secondary N-alkylhemins during hemin-catalyzed epoxidation of terminal alkenes**. *J Am Chem Soc* 1995, **117**:21-29.
185. Ortiz de Montellano PR, Mangold BL, Wheeler C, Kunze KL, Reich NO: **Stereochemistry of cytochrome P-450-catalyzed epoxidation and prosthetic heme alkylation**. *J Biol Chem* 1983, **258**:4208-4213.

186. Kunze KL, Mangold BL, Wheeler C, Beilan HS, Ortiz de Montellano PR: **The cytochrome P-450 active site. Regiospecificity of prosthetic heme alkylation by olefins and acetylenes**. *J Biol Chem* 1983, **258**:4202-4207.
187. Ortiz de Montellano PR, Kunze KL, Mico BA: **Destruction of cytochrome P-450 by olefins: N-alkylation of prosthetic heme**. *Mol Pharmacol* 1980, **18**:602-605.
188. Rai GP, Zong Q, Hager LP: **Isolation of directed evolution mutants of chloroperoxidase resistant to suicide inactivation by primary olefins**. *Isr J Chem* 2000, **40**:63-70.
189. Rai GP, Sakai S, Florez AM, Mogollon L, Hager LP: **Directed evolution of chloroperoxidase for improved epoxidation and chlorination catalysis**. *Adv Synth Catal* 2001, **343**:638-645.
190. van de Velde F, Bakker M, van Rantwijk F, Rai GP, Hager LP, Sheldon RA: **Engineering chloroperoxidase for activity and stability**. *J Mol Catal B-Enzym* 2001, **11**:765-769.
191. Salazar O, Cirino PC, Arnold FH: **Thermostabilization of a cytochrome P450 peroxygenase**. 2003, **Submitted**.
192. Cirino PC, Arnold FH: **Regioselectivity and activity of cytochrome P450BM-3 and mutant F87A in reactions driven by hydrogen peroxide**. *Adv Synth Catal* 2002, **344**:932-937.
193. Cirino PC, Tang Y, Takahashi K, Tirrell DA, Arnold FH: **Global incorporation of norleucine in place of methionine in cytochrome P450 BM-3 heme domain increases peroxygenase activity**. *Biotechnol Bioeng* 2003, **In Press**.

CHAPTER 2

Regioselectivity and Activity of Cytochrome P450 BM-3 and Mutant F87A in Reactions Driven by Hydrogen Peroxide

Material from this chapter appears in: Cirino PC, Arnold FH: **Regioselectivity and activity of cytochrome P450BM-3 and mutant F87A in reactions driven by hydrogen peroxide**. *Adv Synth Catal* 2002, **344**:932-937*,* and is reprinted by permission of Wiley-VCH.

Summary

Cytochrome P450 monooxygenase BM-3 utilizes NADPH and dioxygen to hydroxylate fatty acids at subterminal positions. The enzyme is also capable of functioning as a peroxygenase in the same reaction, by utilizing hydrogen peroxide in place of the reductase domain, cofactor and oxygen. As a starting point for developing a practically useful hydroxylation biocatalyst, we compare the activity and regioselectivity of wild-type P450 BM-3 and its F87A mutant on various fatty acids. Neither enzyme catalyzes terminal hydroxylation under any of the conditions studied. While significantly enhancing peroxygenase activity, the F87A mutation also shifts hydroxylation further away from the terminal position. The H_2O_2-driven reactions with either the full-length BM-3 enzyme or the heme domain are slow, but yield product distributions very similar to those generated when using NADPH and O_2.

Introduction

The cytochromes P450 perform key functions in nature by catalyzing monooxygenations on a wide range of substrates. Many of these reactions are difficult to perform by synthetic methods. There is interest in using P450s for biocatalytic applications [1-4] and there are notable successes [1,5]. In many cases it is desirable to use enzymes removed from whole cells or in purified form (e.g., to avoid metabolism of desired products, for substrates that cannot permeate cell walls, or where sterile conditions are desired). *In vitro* P450 biocatalysis would require continuous regeneration of the expensive cofactor NAD(P)H from $NAD(P)^+$, adding complexity and cost to the reaction system. Several years ago it was proposed that P450s' ability to utilize peroxides in place of dioxygen and NAD(P)H to drive hydroxylation via the peroxide "shunt" pathway could provide a way to confront the cofactor regeneration problem [6]. However, the P450 peroxygenase reaction is usually very inefficient and the enzyme is rapidly inactivated by the peroxide.

We are focusing on evolving P450 BM-3 from *Bacillus megaterium* (CYP102; EC 1.14.14.1), into a practical biocatalyst by improving its shunt activity. As discussed in the previous chapter, P450 BM-3 is a water-soluble, catalytically self-sufficient P450 containing a heme (monooxygenase/hydroxylase) domain and a reductase domain [7-10]. The total length of the enzyme is 1048 amino acids. The heme domain is generally considered to end at position 472 and it is followed by a short linker before the reductase domain begins. Nucleotide and amino acid sequences for P450 BM-3 are provided in Appendix B. P450 BM-3 hydroxylates fatty acids with a chain length between C12 and C18 at subterminal positions [10,11], and also hydroxylates the corresponding fatty acid amides and alcohols and forms epoxides from unsaturated fatty acids [10,12-14]. A

peroxide-driven fatty acid hydroxylase is potentially valuable in chemical synthesis for catalyzing controlled oxidations on specific unactivated C-H bonds. One potential application of interest for a fatty acid hydroxylase driven by H_2O_2 is as an additive in laundry detergents to hydroxylate (and increase the solubility of) strong surfactants and clothing stains.

An often effective strategy for rapid screening of enzyme activity is to modify the natural substrate such that a chromophore is released upon catalysis. Our assay for screening P450 BM-3 peroxygenase activity uses the fatty acid surrogate substrate 12-*para*-nitrophenoxycarboxylic acid (12-pNCA) which, upon hydroxylation at C-12 of the fatty acid, releases colorful *p*-nitrophenolate (pNP), as described in the previous chapter [15]. Below we report the activity of P450 BM-3 on 12-pNCA under peroxide-driven catalysis and the effect of active site mutation F87A on this activity. Following those initial studies, we describe peroxygenase reactions with various fatty acid substrates and compare the activities and regioselectivities with those resulting from the natural, NADPH-driven pathway.

Results

A. Mutation F87A and its Effect on 12-pNCA Activity

Active site mutation F87A alters regioselectivity and results in enhanced sensitivity in the 12-pNCA assay [15]. Under NADPH-driven catalysis, hydroxylation of 12-pNCA by the F87A mutant results in nearly complete conversion to pNP (i.e., 100%

selective toward the terminal carbon), whereas the wild-type enzyme is only about 33% selective for the terminal carbon.

In initial characterizations of the peroxygenase activity of P450 BM-3, the activities of full-length (heme domain plus reductase domain) wild-type P450 BM-3 ("BWT"), full-length F87A mutant ("BF87A"), wild-type heme domain ("HWT") and heme domain mutant F87A ("HF87A") on 12-pNCA were compared. Cell lysates from cultures in which these various P450s were expressed showed that the F87A mutation was critical for supporting H_2O_2-driven activity on 12-pNCA, and the heme domain alone is sufficient for supporting this reaction. Whereas no pNP formation could be detected with BWT and HWT over a range of H_2O_2 concentrations (50 μM – 100 mM), HF87A and BF87A both supported the peroxygenase reaction with H_2O_2 (~50 min^{-1} in 10 mM H_2O_2 and ~90 min^{-1} in 50 mM H_2O_2). The $K_{m,app}$ of H_2O_2 for HF87A and BF87A are similar (~30 mM). The enzyme is very short-lived in the presence of peroxide: in 50 mM H_2O_2 most activity is lost after ~5 minutes. Mutants F87G and F87S also generated pNP from 12-pNCA in the presence of H_2O_2, although at rates slower than mutant F87A.

Control experiments verify that peroxygenase reactions require functional enzyme. Free hemin chloride and heat-inactivated P450 with lauric acid in the presence of H_2O_2 produced no oxidized fatty acid products. Increasing levels of enzyme denaturation due to increasing concentrations guanidinium chloride results in decreasing peroxygenase activity. Finally, a control mutant in which the heme's fifth ligand (Cys 400) was removed (C400G) has no peroxygenase activity.

Before evolving the H_2O_2-driven activity of HF87A on 12-pNCA, we wanted to understand how the F87A mutation affects not only overall activity, but also the

regioselectivity of hydroxylation on fatty acid substrates. We also wanted to know whether the peroxide-driven reaction gives the same product distribution as the reaction using cofactor and oxygen. Finally, it was important to determine whether using the hydroxylase domain alone changes the activity or distribution of hydroxylated products. Below we report regioselectivities and reaction rates for these various permutations of wild-type BM-3 and its F87A mutant acting on myristic (C14), lauric (C12), and capric (C10) acid.

B. Regioselectivities and Reaction Rates on Fatty Acids

Hydroxylated isomers were identified by GC/MS and quantified using a flame ionization detector, as described in Chapter 7 (experimental section). Rates of the NADPH-driven reactions were determined by measuring NADPH consumption and NADPH coupling efficiency, while H_2O_2-driven reaction rates were estimated by quantifying the products from one-minute reactions. Table 2.1 summarizes the results from these experiments, and representative GC traces of the derivatized products from reactions with lauric acid are shown in Figure 2.1. Figure 2.2 shows representative MS fragmentation patterns for the derivatized ω–3 and ω–5 hydroxylated lauric acid products (9- and 7-hydroxylauric acid) corresponding to the peaks at 10.95 min. and 10.55 min., respectively, from Figure 2.1. All product chromatographic resolutions and fragmentation patterns appeared as expected and are in good agreement with previous reports [16-18].

The product distributions for NADPH-driven reactions of wild-type with myristic acid and lauric acid agree with what has been reported [10]; the products of capric acid

hydroxylation have not been reported previously. The rates of NADPH-consumption for wild-type with myristic and lauric acid are also similar to literature values[20], while no rates for capric acid have been previously reported. No significant products other than those reported in Table 2.1 were detected, including ω-hydroxylation products. NADPH consumption by wild-type BM-3 with myristic acid has been indirectly shown to be tightly coupled to product formation [11]. The value we report for the coupling efficiency of wild-type with lauric acid (84%) is similar to that reported by Cowart *et al.* [21]. Coupling efficiency for wild-type with capric acid is very low (32%).

The F87A mutation reduces the reaction rate in the NADPH-driven reaction and also reduces the coupling efficiency. Similar results have been reported for the F87V mutant [21]. F87A also broadens regiospecificity and shifts hydroxylation away from the terminal position, in contrast to what has been previously reported [22]. The primary position of hydroxylation with myristic acid is shifted from ω–1 to ω–5. These results agree with a similar study on the F87V mutant [13], which suggested that replacement of Phe87 with a smaller residue allows for deeper penetration of long-chain substrates into the active site.

Table 2.1. Fatty acid hydroxylation activity of wild-type BM-3 and mutant HF87A under the natural pathway (NADPH+O_2) and the peroxide shunt pathway. BWT: full-length wild-type BM-3. HWT: heme domain wild-type. BF87A: full-length F87A mutant. HF87A: heme domain F87A. ND: could not be determined. All standard deviations were less than 10% of the average, except for the standard deviations reported below.

			Distribution of Hydroxylated Products (%)						
Fatty Acid	**P450**	**Reaction**	**ω−6**	**ω−5**	**ω−4**	**ω−3**	**ω−2**	**ω−1**	**Initial Rate (NADPH Coupling)[a]**
Myristic	**BWT**	**NADPH+O_2**			**1**	**24**	**26**	**48**	**1700[b]**
(C14)	**HWT**	**H_2O_2**				**33**	**29**	**38**	**ND[c]**
	BF87A	**NADPH+O_2**	**2**	**57**	**16**	**7**	**8**	**11**	**830±100 (ND)**
	BF87A	**H_2O_2**	**5±2**	**56**	**14**	**5**	**6**	**15**	**ND**
	HF87A	**H_2O_2**	**4±1**	**61**	**13**	**5**	**6±1**	**12**	**ND**
Lauric	**BWT**	**NADPH+O_2**				**37**	**27**	**36**	**1200 (84%)**
(C12)	**BWT**	**H_2O_2**				**42**	**28**	**30±4**	**ND[c]**
	HWT	**H_2O_2**				**46**	**27**	**28±3**	**ND[c]**
	BF87A	**NADPH+O_2**		**24**	**19**	**40**	**10**	**6**	**350±50 (67%)**
	BF87A	**H_2O_2**		**36**	**20**	**29**	**8**	**7±1**	**11**
	HF87A	**H_2O_2**		**39**	**19**	**30**	**7**	**5**	**10**
Capric	**BWT**	**NADPH+O_2**				**13**	**19**	**68**	**150 (32%)**
(C10)	**BWT**	**H_2O_2**				**21±3**	**23**	**56**	**ND[c]**
	HWT	**H_2O_2**				**18**	**23**	**59**	**ND[c]**
	BF87A	**NADPH+O_2**				**39**	**23**	**38**	**24 (18%)**
	BF87A	**H_2O_2**		**3±2**	**5±1**	**48**	**17**	**28**	**34**
	HF87A	**H_2O_2**		**3±1**	**4±1**	**53**	**17**	**23**	**46±8**

[a] Rates are reported as mole(mole P450)$^{-1}$(min)$^{-1}$. Rates for the NADPH-driven reactions are the initial rates of NADPH consumption, with coupling efficiencies reported in parentheses. Coupling efficiencies were calculated as the total moles of hydroxylated product formed per 100 moles NADPH consumed. Initial rates of H_2O_2-driven reactions were estimated by measuring the amount of product formed in one minute. [b] NADPH consumption in this reaction is reported to be tightly coupled to product formation [11,19]. [c] WT was estimated to be capable of only ~2-3 turnovers (before complete inactivation) in the peroxide-driven hydroxylation reaction.

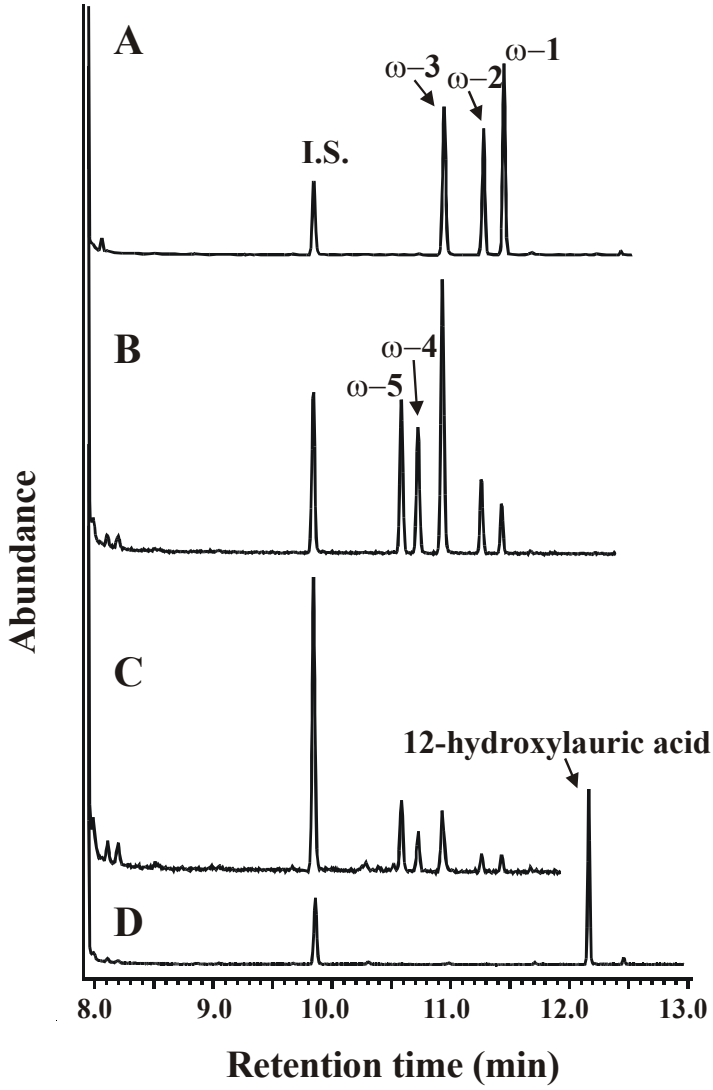


Figure 2.1. Representative chromatogram of the TMS-derivatized hydroxylated products from reactions of P450 BM-3 with lauric acid. **A**) Reaction of wild-type BM-3 with NADPH + O_2. **B**) Reaction of mutant F87A with NADPH + O_2. **C**) Reaction of mutant F87A with H_2O_2. **D**) 12-Hydroxylauric acid standard. This sample was prepared in the same manner as the reaction samples, except the enzyme was first inactivated. I.S. = internal standard (10-hydroxycapric acid). For quantitative comparisons, all peak areas were normalized relative to the internal standard peak area. Samples were prepared as described in Chapter 7.

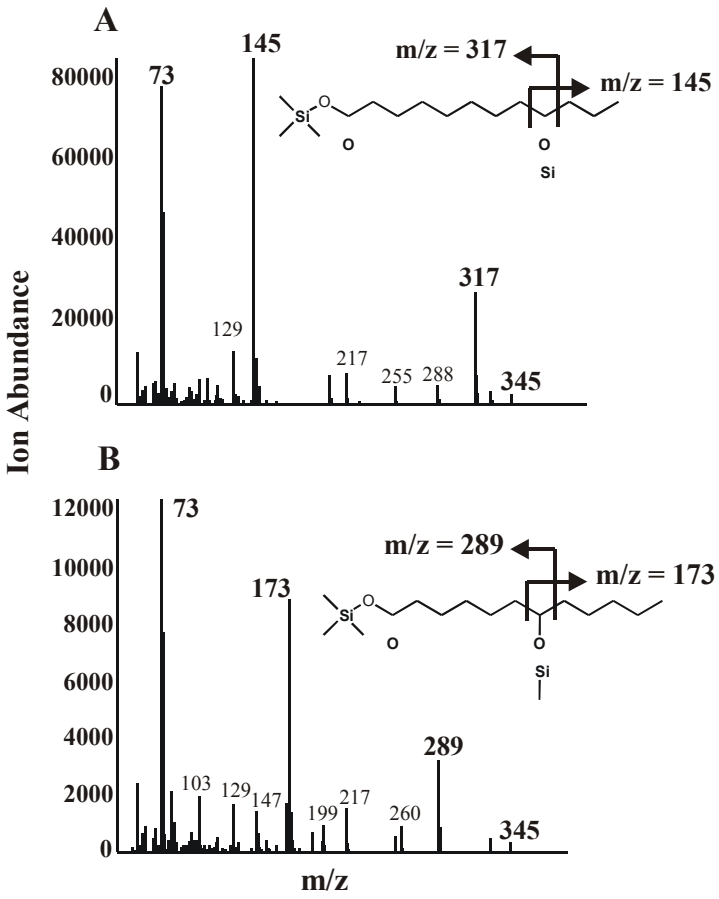


Figure 2.2. Representative mass spectrometry fragmentation patterns for the TMS-derivatized hydroxylated lauric acid peaks in Figure 2.1. **A**) MS fragments corresponding to the GC peak at 10.95 minutes, clearly identifying derivatized 9-hydroxylauric acid (ω–3 hydroxylated product). **B**) MS fragments corresponding to the GC peak at 10.55 minutes, clearly identifying derivatized 7-hydroxylauric acid (ω–5 hydroxylated product).

Reactions of P450s with various peroxides have been studied in detail (refer to [23]). Significant differences between products of the NADPH and peroxide pathways have been reported for liver microsomal P450s when organic hydroperoxides were used [24-26]. In contrast, reactions of microsomal P450s using H_2O_2 gave products identical to or very similar to those found from the NADPH reactions [27-30]. We have found that the product distributions for the monooxygenase and H_2O_2-driven peroxygenase reactions for purified P450 BM-3 are also very similar. Furthermore, removal of the reductase

domain has no significant effect on shunt pathway activity or products. Thus it is likely that the shunt pathway goes through the same active oxidative intermediate(s). The small shift in product distribution for the peroxygenase reaction could reflect slight variations in substrate placement due to the presence of H_2O_2 or the lack of dependence on electron and proton transfer to the active site for formation of the reactive intermediate(s), two processes which are likely to coincide with conformational fluctuations that affect regioselectivity.

Wild-type (both full-length BM-3 and heme domain) is capable of no more than a few turnovers before it is inactivated, and inactivation is greatly accelerated by the presence of substrate (refer to Chapter 6). The F87A mutant is more active than wild-type with H_2O_2, although inactivation still appears to be largely due to a competing, turnover-dependent reaction. Similar observations have been made [31]. While we have not elucidated the mechanism of inactivation for HWT or HF87A, we know that it is primarily a result of heme degradation, as discussed in Chapter 6.

The full-length F87A mutant is reported to have an apparent K_m for H_2O_2 of ~18 mM [31], in good agreement with our findings. Higher concentrations of H_2O_2 result in higher reaction rates, but the enzyme is also inactivated much faster. Even in 10 mM H_2O_2 the enzyme is quickly inactivated: total turnovers for the F87A mutant are approximately 70 with lauric acid and 200 with capric acid. Shunt pathway reaction rates with myristic acid could not be quantified with high accuracy because there was no good authentic standard available for calibration. However, product peak areas show that the F87A mutant made 10-15 times more hydroxylated product than wild-type in one minute, implying a rate of ~20 min^{-1}.

Discussion

There appears to be a correlation between NADPH uncoupling and shunt pathway activity. The rate of the F87A peroxygenase reaction with capric acid in 10 mM H_2O_2 is in fact higher than the highly uncoupled F87A NADPH-driven reaction, and about the same as the wild-type NADPH-driven reaction (also highly uncoupled). This correlation between peroxygenase activity and uncoupling might be explained by increased hydrophilicity in the active site. The F87A mutation is likely to create space that permits more water molecules in the active site pocket. Poorer substrates are less effective at excluding water from the active site upon binding. Water remaining after substrate binding could promote H_2O_2 access to the active site. Alternatively, the presence of water could facilitate the dissociation of hydrogen peroxide anion during NADPH-driven catalysis [32].

We are primarily interested in evolving the ability of P450 BM-3 to utilize H_2O_2 to drive the shunt pathway, both because H_2O_2 is the most practical peroxide in terms of cost and safety, and because of the potential applications of a H_2O_2-driven hydroxylase in laundry detergents. However, we also examined peroxygenase activity with several other peroxides (cumene hydroperoxide, *t*-butylhydroperoxide, peroxyacetic acid, and *m*-chloroperoxybenzoic acid) over a range of peroxide concentrations (25 μM to 50 mM). Only H_2O_2 allowed for detectable activity with the 12-pNCA assay (which is necessary for screening). This may be a result of steric interference between the peroxide and the *p*-nitrophenolate group of the substrate (so that no reaction occurs), or because hydroxylation of 12-pNCA occurs at a position other than the terminal carbon. Further

investigation revealed that both cumene hydroperoxide and t-butylhydroperoxide drive hydroxylation of fatty acids by HF87A, but with altered regioselectivity compared to reactions driven by H_2O_2. Refer to Appendix C for data and a description of these experiments. Interestingly, HWT also supported the reaction with these organic peroxides, although total turnovers were still significantly lower than those of HF87A. As discussed above, previous studies have shown that organic peroxides utilized in the shunt pathway alter P450 regioselectivity, while H_2O_2 does not.

Conclusions

We have quantified and compared the product distributions from NADPH- and H_2O_2-driven hydroxylation reactions of purified P450 BM-3 with fatty acids. That the product profiles from these different pathways are so similar suggests that we can harness peroxygenase activity without sacrificing P450 reaction specificity. Using the heme domain alone further simplifies this catalyst. However, wild-type P450 BM-3 is very slow and subject to rapid inactivation when it uses H_2O_2. Mutating Phe87 to a smaller residue increases this activity and provides a much improved starting point for further directed evolution to create a practical peroxide-driven hydroxylation catalyst.

References

1. Sakaki T, Inouye K: **Practical application of mammalian cytochrome P450**. *J Biosci Bioeng* 2000, **90**:583-590.
2. Miles CS, Ost TWB, Noble MA, Munro AW, Chapman SK: **Protein engineering of cytochromes P-450**. *Biochim Biophys Acta-Protein Struct Molec Enzym* 2000, **1543**:383-407.
3. Holland HL, Weber HK: **Enzymatic hydroxylation reactions**. *Curr Opin Biotechnol* 2000, **11**:547-553.
4. Hara M: **Application of P450s for biosensing: combination of biotechnology and electrochemistry**. *Mater Sci Eng C-Biomimetic Supramol Syst* 2000, **12**:103-109.
5. Guengerich FP: **Cytochrome P450 enzymes in the generation of commercial products**. *Nat Rev Drug Discov* 2002, **1**:359-366.
6. Joo H, Lin Z, Arnold FH: **Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation**. *Nature* 1999, **399**:670-673.
7. Li HY, Darwish K, Poulos TL: **Characterization of recombinant *Bacillus megaterium* cytochrome P-450 BM- 3 and its two functional domains**. *J Biol Chem* 1991, **266**:11909-11914.
8. Ruettinger RT, Wen LP, Fulco AJ: **Coding nucleotide, 5' regulatory, and deduced amino acid sequences of P- 450BM-3, a single peptide cytochrome P-450:NADPH-P-450 reductase from *Bacillus megaterium***. *J Biol Chem* 1989, **264**:10987-10995.
9. Miura Y, Fulco AJ: **(Omega -2) hydroxylation of fatty acids by a soluble system from *Bacillus megaterium***. *J Biol Chem* 1974, **249**:1880-1888.
10. Miura Y, Fulco AJ: **Omega-1, omega-2 and omega-3 hydroxylation of long-chain fatty acids, amides and alcohols by a soluble enzyme system from *Bacillus megaterium***. *Biochim Biophys Acta* 1975, **388**:305-317.
11. Boddupalli SS, Estabrook RW, Peterson JA: **Fatty acid monooxygenation by cytochrome P-450BM-3**. *J Biol Chem* 1990, **265**:4233-4239.
12. Capdevila JH, Wei S, Helvig C, Falck JR, Belosludtsev Y, Truan G, Graham-Lorence SE, Peterson JA: **The highly stereoselective oxidation of polyunsaturated fatty acids by cytochrome P450BM-3**. *J Biol Chem* 1996, **271**:22663-22671.
13. Graham-Lorence S, Truan G, Peterson JA, Falck JR, Wei S, Helvig C, Capdevila JH: **An active site substitution, F87V, converts cytochrome P450 BM-3 into a regio- and stereoselective (14S,15R)-arachidonic acid epoxygenase**. *J Biol Chem* 1997, **272**:1127-1135.
14. Ruettinger RT, Fulco AJ: **Epoxidation of unsaturated fatty acids by a soluble cytochrome P-450-dependent system from *Bacillus megaterium***. *J Biol Chem* 1981, **256**:5728-5734.
15. Schwaneberg U, Schmidt-Dannert C, Schmitt J, Schmid RD: **A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A**. *Anal Biochem* 1999, **269**:359-366.
16. Fang X, Halpert JR: **Dithionite-supported hydroxylation of palmitic acid by cytochrome P450BM-3**. *Drug Metab Dispos* 1996, **24**:1282-1285.

17. Sevrioukova I, Truan G, Peterson JA: **Reconstitution of the fatty acid hydroxylase activity of cytochrome P450BM-3 utilizing its functional domains**. *Arch Biochem Biophys* 1997, **340**:231-238.
18. Lentz O, Qing-Shang LI, Schwaneberg U, Lutz-Wahl S, Fischer P, Schmid RD: **Modification of the fatty acid specificity of cytochrome P450BM-3 from *Bacillus megaterium* by directed evolution: A validated assay**. *J Mol Catal B-Enzym* 2001, **15**:123-133.
19. Black SD, Linger MH, Freck LC, Kazemi S, Galbraith JA: **Affinity isolation and characterization of cytochrome-P450-102 (Bm-3) from barbiturate-induced *Bacillus megaterium***. *Arch Biochem Biophys* 1994, **310**:126-133.
20. Yeom H, Sligar SG: **Oxygen activation by cytochrome P450BM-3: effects of mutating an active site acidic residue**. *Arch Biochem Biophys* 1997, **337**:209-216.
21. Cowart LA, Falck JR, Capdevila JH: **Structural determinants of active site binding affinity and metabolism by cytochrome P450BM-3**. *Arch Biochem Biophys* 2001, **387**:117-124.
22. Oliver CF, Modi S, Sutcliffe MJ, Primrose WU, Lian LY, Roberts GC: **A single mutation in cytochrome P450 BM3 changes substrate orientation in a catalytic intermediate and the regiospecificity of hydroxylation**. *Biochemistry* 1997, **36**:1567-1572.
23. Ortiz de Montellano PR: **Oxygen activation and reactivity**. In *Cytochrome P450: Structure, Mechanism, and Biochemistry,* 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:245-303.
24. Ellin A, Orrenius S: **Hydroperoxide-supported cytochrome P-450-linked fatty acid hydroxylation in liver microsomes**. *FEBS Lett* 1975, **50**:378-381.
25. Capdevila J, Estabrook RW, Prough RA: **Differences in the mechanism of NADPH- and cumene hydroperoxide- supported reactions of cytochrome P-450**. *Arch Biochem Biophys* 1980, **200**:186-195.
26. Hlavica P, Golly I, Mietaschk J: **Comparative studies on the cumene hydroperoxide- and NADPH-supported N-oxidation of 4-chloroaniline by cytochrome P450**. *Biochem J* 1983, **212**:539-547.
27. Renneberg R, Scheller FW, Ruckpaul K, Pirrwitz J, Mohr P: **NADPH and H2O2-dependent reactions of cytochrome P-450LM compared with peroxidase catalysis**. *FEBS Lett* 1978, **96**:349-353.
28. Renneberg R: **Hydrogen peroxide-supported oxidation of benzo(a)pyrene by rat liver microsomal fractions**. *Biochem Pharmacol* 1981, **30**:843-848.
29. Estabrook RW, Martin-Wixtrom C, Saeki Y, Renneberg R, Hildebrandt A, Werringloer J: **The peroxidatic function of liver microsomal cytochrome P-450: comparison of hydrogen peroxide and NADPH-catalysed N-demethylation reactions**. *Xenobiotica* 1984, **14**:87-104.
30. Holm KA, Engell RJ, Kupfer D: **Regioselectivity of hydroxylation of prostaglandins by liver microsomes supported by NADPH versus H2O2 in methylcholanthrene-treated and control rats: formation of novel prostaglandin metabolites**. *Arch Biochem Biophys* 1985, **1985**:477-489.

31. Li QS, Ogawa J, Shimizu S: **Critical role of the residue size at position 87 in H2O2- dependent substrate hydroxylation activity and H2O2 inactivation of cytochrome P450BM-3**. *Biochem Biophys Res Commun* 2001, **280**:1258-1261.
32. Mueller EJ, Loida, P.J., and Sligar, S.G.: **Twenty-five years of P450cam research**. In *Cytochrome P450 Structure, Mechanism, and Biochemistry*, 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:83-124.

CHAPTER 3

A Self-Sufficient Peroxide-Driven Hydroxylation Biocatalyst

Material from this chapter appears in: Cirino PC, Arnold FH: **A self-sufficient peroxide driven hydroxylation biocatalyst**. *Angew Chem-Int Edit* 2003, **In press**, and is reprinted by permission of Wiley-VCH.

Summary

We have converted the heme domain of cytochrome P450 BM-3 into a highly active peroxide-driven hydroxylation catalyst. This enzyme does not require an NADPH cofactor and can therefore be used in cell-free biosynthesis applications. The P450 enzyme also no longer requires its reductase domain, which makes this simplified, 'biomimetic' catalyst amenable to further optimization, for example, to improve stability or alter its substrate range.

Introduction

Controlled, selective oxidations of unactivated C-H bonds are among the most desired transformations in catalysis (for reviews of current technologies, see [1-7]). Limitations to chemical oxidation catalysts include low turnover numbers, low or no regioselectivity, poor reaction specificity, their use of environmentally harmful components (e.g., heavy metals or halogens), or their requirements for harsh, expensive reaction conditions. Many oxygenase enzymes catalyze the insertion of oxygen into unactivated C-H bonds [8-12] and are superior to synthetic catalysts with regard to selectivity and turnover rate [13]. Enzymes have found numerous industrial applications [14,15], and the use of enzymes in industrial biocatalysis is expected to grow significantly [16]. One limitation to the application of oxygenases is their requirement for an expensive cofactor such as NAD(P)H, which precludes their use *in vitro* without addition of a coupled, cofactor-regenerating system [17-19]. The need for additional proteins to transfer electrons from the cofactor to the oxygenase presents a further complication. Here we describe a self-sufficient P450 BM-3 variant which utilizes hydrogen peroxide (H_2O_2) to catalyze hydroxylation and epoxidation at high rates.

The cytochromes P450 are heme-containing oxygenases which collectively catalyze a variety of oxidations on a diverse array of substrates [12,20]. In principle P450 peroxygenase activity offers an opportunity to employ cell-free P450 catalysis without requiring NAD(P)H regeneration, additional proteins, or dioxygen, and eliminates rate-limiting electron transfer steps [21]. In practice, this nonnatural pathway is too inefficient for any practical application. We have focused on improving the peroxygenase activity of P450 BM-3 as a step towards engineering a “biomimetic”

hydroxylation catalyst. The activities of P450 BM-3 and its corresponding F87A mutant in reactions driven by H_2O_2 are described in Chapter 2. Whereas the wild-type enzyme is inactivated after only a few turnovers, the F87A mutant supports low levels of peroxygenase activity (~70 total turnovers).

P450 BM-3 heme domain supports peroxygenase activity and is more thermostable than the full-length P450 BM-3 protein (refer to Chapter 4). Removal of the reductase domain also results in approximately fourfold increased molar expression (~70 mg/l in shake flasks). We used heme domain mutant F87A (HF87A) as the starting point (parent) to increase catalyst performance by sequential rounds of random mutagenesis and screening for H_2O_2-driven hydroxylation of 12-pNCA [22].

Results

Critical to the success of directed evolution is a sensitive screening assay. Screening conditions must be developed to accurately reflect the function of interest to be evolved. High-throughput cell growth and enzyme expression were optimized to avoid false positives. Cell lysates were used to assay mutant activity, so lysis conditions were also optimized. *E. coli* naturally produces catalase and the presence of catalase in the lysate was problematic in the development of a screening assay for shunt pathway activity. Bubbles were formed from the catalase reaction, and H_2O_2 concentrations were rapidly reduced. Therefore a catalase-deficient *E. coli* strain was used, in which the genes that code for catalase were knocked out of the host genome [23]. Use of this strain prevented bubble formation, and allowed for steady concentrations of H_2O_2 to be maintained, resulting in a sensitive screening system.

Figure 3.1 shows the time-course of *p*-nitrophenolate (pNP) formation in the reaction of HF87A with 12-pNCA with 1mM and 50 mM H_2O_2. The apparent K_m of H_2O_2 for HF87A is ~30 mM. With high concentrations of H_2O_2 the initial rate is high but the enzyme is essentially inactive within 5 minutes, while in low concentrations of H_2O_2 (1 mM) initial rates are much lower and the enzyme is more stable in the presence of peroxide. DMSO (6.3% final concentration) was used in the final screening conditions to improve the solubility of the 12-pNCA substrate. This concentration of DMSO did not have a noticeable effect on the activity of HF87A from cell lysates in 50 mM H_2O_2.

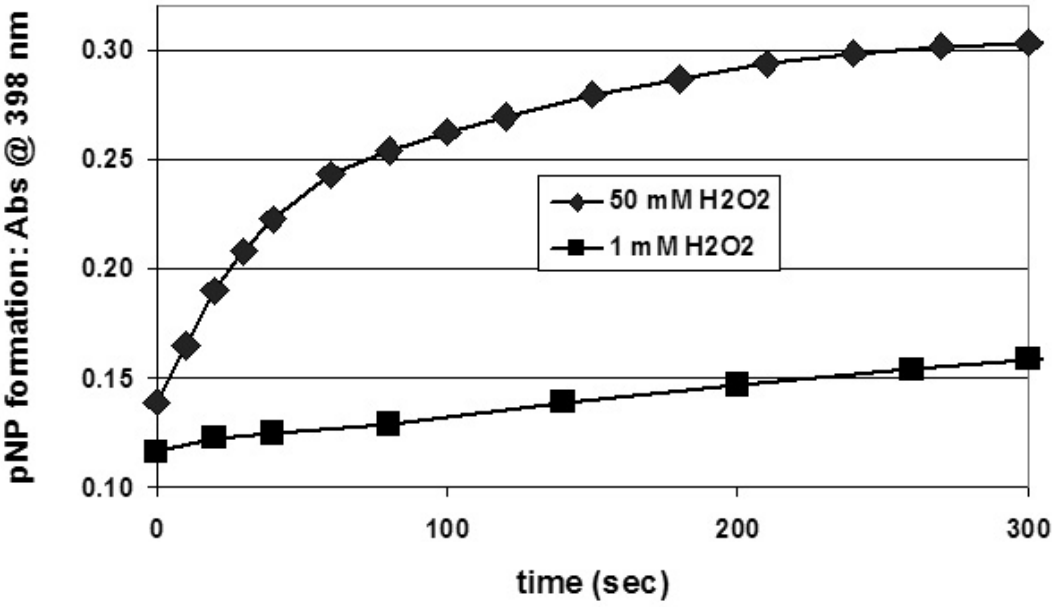


Figure 3.1. Time-course of pNP formation for the reaction of HF87A with 12-pNCA in 1 mM and 50 mM H_2O_2.

Figure 3.2 depicts the approximate activities of mutants selected throughout the evolution of peroxygenase activity, relative to HF87A. Activities in Figure 3.2 were measured as the initial rates of pNP formation (from 12-pNCA) using cell lysates in 1 mM H_2O_2. Chapter 5 lists the amino acid substitutions found from sequencing selected

mutants through different stages of evolution. In the first three rounds of evolution, mutant libraries were screened for activity in 1 mM H_2O_2 to identify variants with reduced peroxide requirements (lower K_m) and 50 mM H_2O_2 to isolate variants with increased peroxide stability. Initial rates of pNP formation as well as end-points (total turnovers; TON) were used as selection criteria. Mutants "1F8" and "2E10" from the first generation, created by error-prone PCR, were chosen as the parents for two separate second generations, also created by error-prone PCR. A total of seven improved mutants were then recombined using StEP recombination [24], in order to combine the properties of increased activity in 1 mM and 50 mM H_2O_2. One resulting variant, "step B3" was ~5 times more active than HF87A in 1 mM H_2O_2 (measured by initial rate) and ~7 times more active than HF87A in 50 mM H_2O_2 (measured by end-point). Although screening with 50 mM H_2O_2 identified variants with increased initial activity at this concentration, none were more stable to peroxide (i.e., the half-life of the enzyme in 50 mM H_2O_2 was not improved). Therefore we continued screening only in 1 mM H_2O_2, using "step B3" as the parent. Two more rounds of error-prone PCR and screening resulted in mutant "21B3". Note from Figure 3.2 that the evolutionary pathway does not appear to be approaching saturation of peroxygenase activity (a local optimum has not been reached). Thus, further rounds of mutagenesis and screening should result in further improvements in peroxygenase activity.

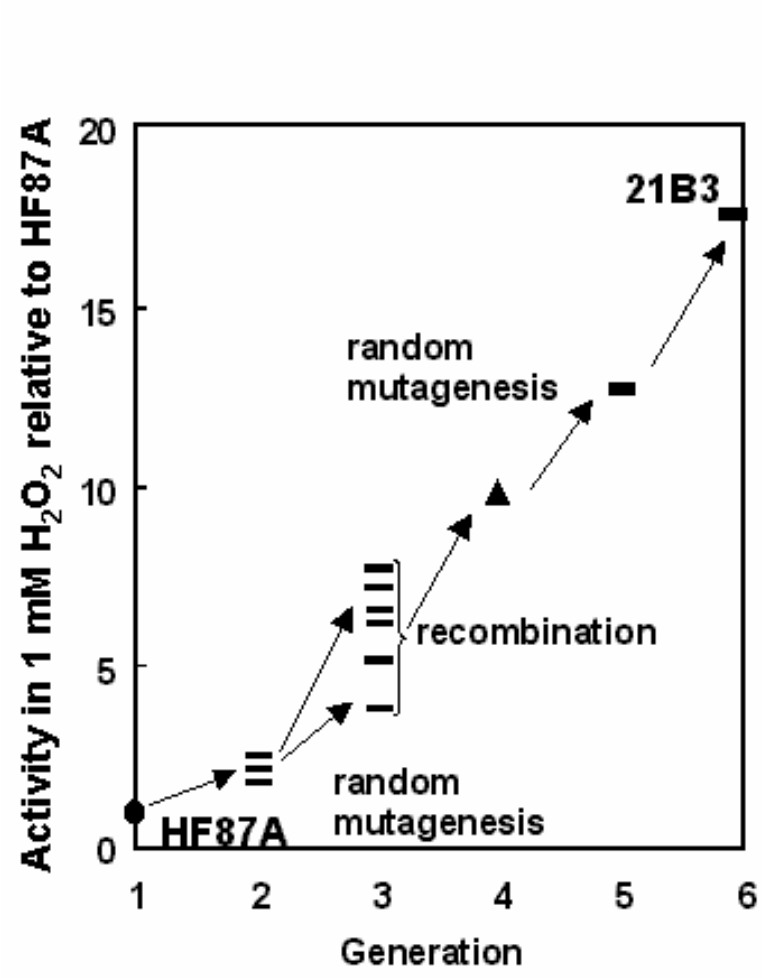


Figure 3.2. Evolution of peroxygenase activity leading to variant 21B3.

Fifth-generation mutant 21B3 is ~20-fold more active than HF87A in 10 mM H_2O_2 using 12-pNCA as substrate. Figure 3.3 shows the activities of HF87A and 21B3 at different peroxide concentrations. Activities are reported as initial rates of pNP formation at room temperature. Both HF87A and 21B3 are ~100% selective for the C-12 position of 12-pNCA, resulting in nearly complete conversion of hydroxylated products to pNP. The peroxygenase reaction follows apparent Michaelis-Menten kinetics. Whereas the HF87A apparent K_m for H_2O_2 is ~30 mM, in 21B3 the apparent K_m is reduced to ~8 mM. Total turnover numbers (TON) of ~1,000 are achieved.

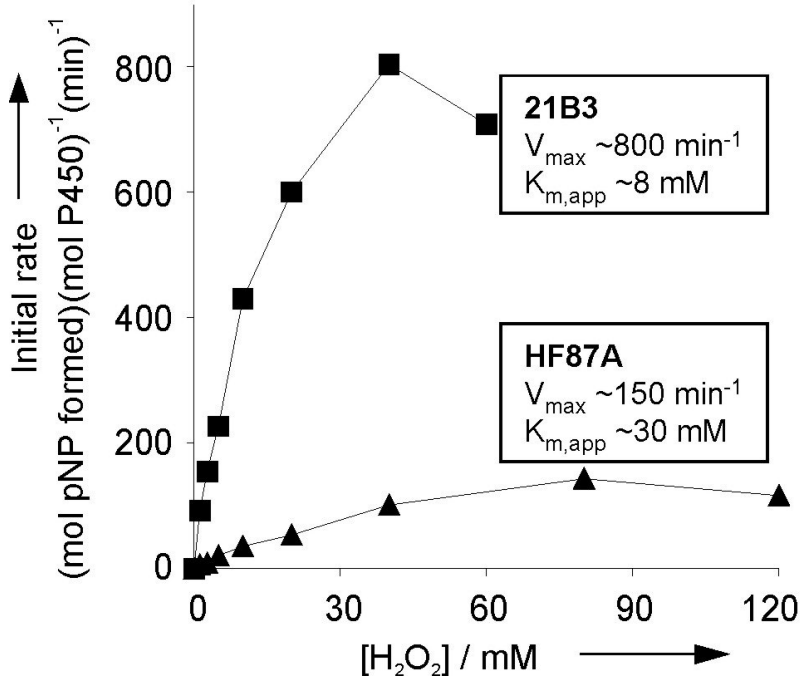


Figure 3.3. Initial rates of peroxide-driven 12-pNCA hydroxylation by P450 BM-3 variants HF87A (▲) and 21B3 (■) in different H_2O_2 concentrations. Reactions were performed at room temperature and contained 300 μM 12-pNCA, 6% DMSO, and purified enzyme (0.5 μM HF87A or 0.2 μM 21B3) in 100 mM Tris-HCl, pH 8.2.

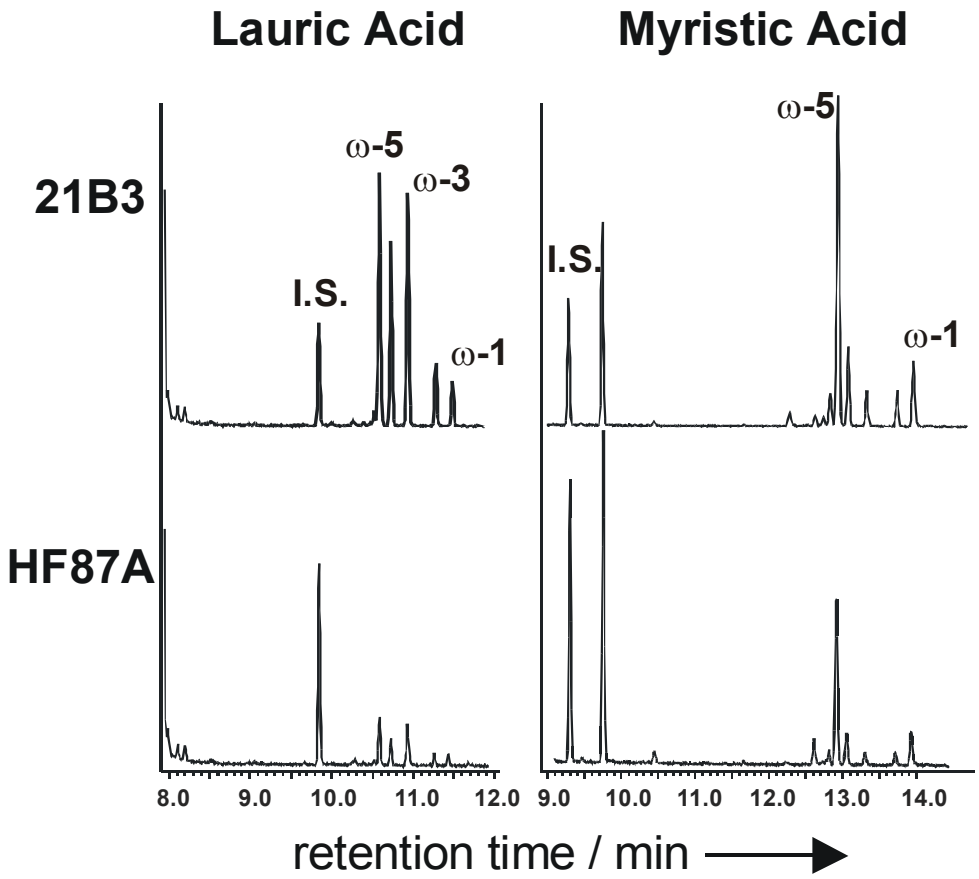


Figure 3.4. Chromatograms of the TMS-derivatized hydroxylated products from H_2O_2-driven reactions of P450 BM-3 variants HF87A and 21B3 with lauric acid and myristic acid. Reactions contained purified enzyme (1-4 μM) and 1-2 mM substrate in 500 μl 100 mM Tris-HCl, pH 8.2. Reactions were initiated by the addition of 5 mM or 10 mM H_2O_2, run at room temperature, and stopped by the addition of 7.5 μl 6 M HCl, as described in Chapter 7. I.S. is the internal standard (10-hydroxycapric acid), which was added to all samples in an equal amount (30 nmoles) at the end of each reaction so that products could be quantified relative to the area of the I.S. peak.

In addition to knowing how the accumulating mutations affect overall activity, we wanted to know how they affect specificity of the enzyme and, in particular, the regioselectivity of hydroxylation. Improved peroxygenase activity also extends to fatty acid substrates, with product distributions that are the same as those from HF87A, as demonstrated in Figure 3.4 for reactions of HF87A and 21B3 with lauric (dodecanoic) acid and myristic (tetradecanoic) acid. No products other than those shown in Figure 3.4

were identified from these reactions. Using 10 mM H_2O_2, the initial rate of lauric acid hydroxylation, estimated from the amount of product formed in the first minute of reaction (refer to methods in Chapter 7), is ~50 min^{-1} for 21B3. The TON for lauric acid in 5 mM H_2O_2 is ~280. Similar results are observed with capric (decanoic) acid.

Oxidation of styrene to styrene oxide is also significantly improved with 21B3. (Relaxed substrate specificity was expected since specificity was not a selection criterion during screening.) The initial rate of styrene oxide production by 21B3 is ~55 min^{-1} in 10 mM H_2O_2 (estimated from the amount of product formed in the first minute of reaction), with a TON of ~240 in 5 mM H_2O_2. In contrast, the initial rate for HF87A in 10 mM H_2O_2 is ~2.4 min^{-1}, with a TON of only ~10. For comparison, the NADPH-driven styrene oxidation activity of wild-type BM-3 (for which styrene is a poor substrate) is ~30 min^{-1}. Table 3.1 summarizes the activities of 21B3 compared to HF87A for the various substrates tested.

Note from Table 3.1 that although DMSO did not appear to influence the activity of HF87A from cell lysates using 50 mM H_2O_2, 6% DMSO does inhibit the activity of (purified) HF87A in 10 mM H_2O_2. This can be explained by the fact that DMSO acts as a competitive inhibitor of the enzyme, and H_2O_2 concentrations significantly above the K_m (e.g., 50 mM) will not show reduced initial rates. Interestingly, because enzymes were evolved in the presence of 6.3% DMSO, variant 21B3 is not inhibited by this concentration of DMSO in reactions with 12-pNCA (although inhibition does occur at much higher DMSO concentrations (refer to Chapter 6)). DMSO was not used for determining activities on styrene or lauric acid.

Table 3.1. Summary of peroxygenase activities of P450 BM-3 variants 21B3 and HF87A.

Substrate	Measurement	21B3	HF87A
12-pNCA	Initial rate[a]	430	23[d]
12-pNCA	TON[b]	980	90
Lauric Acid	Initial rate	50	10
Lauric Acid	TON	280	70
Styrene[c]	Initial rate	54	2.4
Styrene[c]	TON	240	10

[a] Initial rates were determined in 10 mM H_2O_2 and are reported as mol product(mol P450)$^{-1}$(min)$^{-1}$. [b] Total enzyme turnovers. TON values were determined in 5 mM H_2O_2. [c] Styrene oxide is the only product. [d] HF87A is nearly twofold faster in 1% DMSO (~40 min^{-1}) compared to in 6% DMSO.

The major limitation to this system is rapid enzyme inactivation as a result of peroxide-mediated heme degradation, indicated by the decrease in the heme absorbance peak in the presence of H_2O_2 (refer to Chapter 6). Figure 3.5 shows the time-courses of pNP formation from reactions of 21B3 on 12-pNCA in 1 mM, 5 mM and 10 mM H_2O_2. The concentration of 12-pNCA was not limiting in these reactions, and pNP concentrations generated were not inhibiting. In 10 mM H_2O_2 the enzyme is essentially inactive within 5 minutes. Inactivation appears to be primarily turnover-dependent (similar TONs are reached from reactions in 1 mM, 5 mM and 10 mM H_2O_2), and the ratio of peroxygenase turnovers to heme-degrading turnovers increased with each generation of evolved enzymes.

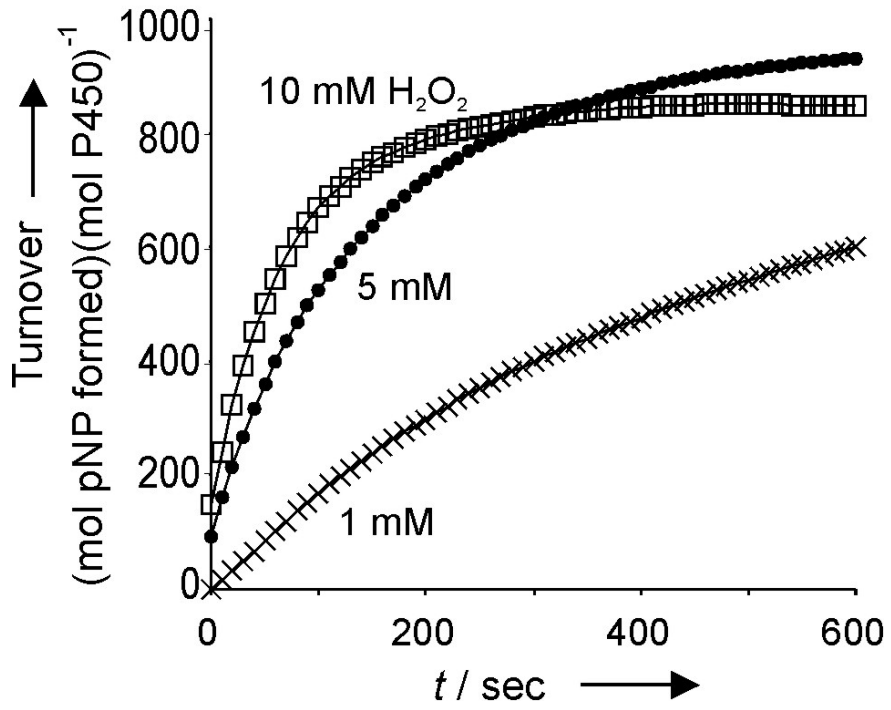


Figure 3.5. Time-courses of pNP formation for reactions catalyzed by P450 BM-3 variant 21B3 with 12-pNCA in 1, 5, and 10 mM H_2O_2. Increasing peroxide concentration increases peroxygenase activity as well as the rate of enzyme inactivation. Reactions were performed at room temperature and contained 250 μM 12-pNCA, 6.3% DMSO, and either 0.94 μM F87A or 0.15 μM 21B3 in 100 mM Tris-HCl, pH 8.2.

Sequence analysis of 21B3 reveals thirteen nucleotide substitutions compared to the HF87A sequence, nine of which result in amino acid substitutions (refer to Chapter 5). The P450 BM-3 heme domain crystal structure [25] shows that the mutations are dispersed throughout the protein scaffold. No mutations appear in the active site, substrate binding channel, or F and G helices of the heme domain. There are also no mutations in any β-sheets of the β-sheet-rich region of the structure. Four mutations lie on the protein surface. Further information on amino acid substitutions in various evolved P450 BM-3 mutants can be found in Chapter 5.

Discussion

Heme serves as the active center in different families of proteins classified by structural similarity (e.g., peroxidases, cytochromes P450, globins, catalases). Within these families the metalloporphyrin performs a primary function (e.g., hydroxylation or oxygen binding), but there is also considerable functional overlap among them. The protein regulates the function, but a given protein structure may not necessarily be optimal for the resulting function and that same protein can be evolved to perform new functions.

As discussed in Chapter 1, previous studies have been successful in converting heme enzyme function via rational design of the active site (e.g., [26-32]). Ortiz de Montellano and coworkers improved the peroxygenase activity of horseradish peroxidase (HRP) by the introduction of two active site mutations [31,32], and Watanabe and coworkers improved the peroxidase [30] and peroxygenase [29] activities of myoglobin (Mb) by the introduction of one and two mutations, respectively. Directed evolution has also been used to increase the peroxidase activity of Mb [33] and catalase [34,35].

Directed evolution allows us to explore the interconversion of function within a structural framework and understand how easily functions that are primarily associated with one scaffold can be grafted into another. We sought to modify the P450 scaffold to support efficient peroxygenase catalysis. We chose cytochrome P450 BM-3 as our target protein because it is a fast enzyme with interesting chemistry and it is functional in *E. coli*, making DNA and enzyme manipulations amenable to directed evolution protocols.

A. Comparison to Natural Peroxygenases

As discussed in Chapter 1, the existence of natural H_2O_2-driven hydroxylases such as CPO and P450s SPα and BSβ lend to the feasibility of engineering an NADPH-utilizing P450 into a functional peroxygenase. CPO has received much attention for its ability to catalyze H_2O_2-driven, enantioselective epoxidation of alkenes [37]. This enzyme has been reported to achieve total turnovers exceeding $8.5x10^5$ in the oxidation of indole using H_2O_2 concentrations maintained below 50 μM via controlled addition [38]. Interestingly, in spite of its natural ability to efficiently utilize H_2O_2, CPO is still unstable in H_2O_2 concentrations above ~100 μM, and its half-life in 1 mM H_2O_2 is only 0.5 hours [39]. This suggests that it may be very difficult or impossible to improve the stability of heme-containing enzymes to high H_2O_2 concentrations while still maintaining function (i.e., still allowing H_2O_2 to gain access to the heme). A more probable outcome from improving peroxygenase activity through directed evolution is a reduction in the peroxide requirement for activity (reduced K_m), as was found for 21B3. Significantly greater reductions in the K_m are clearly required before our evolved mutants can be driven without suffering from H_2O_2-mediated inactivation.

CPO peroxygenase reactions are limited to epoxidation, heteroatom oxidation and activated (electron-rich) C-H bond hydroxylation, and not toward the unactivated C-H bond hydroxylation reactions which are the hallmark of P450 catalysis (Toy *et al*. report unactivated C-H bond oxidation of the radical probe substrate (*trans*-2-phenyl-1-methylcyclopropane) by CPO with a TON of 20, and only 31% of product was alcohol [40]). Furthermore, protein engineering on CPO has been extremely limited, due to its lack of functional expression in bacteria or yeast (see Chapter 1). The peroxygenase

P450s SPα (CYP152B1) and BSβ (CYP152A1) have unique active sites that include an Arg residue. Only fatty acids are accepted as substrates and the carboxylate of the substrate is positioned into the active site upon substrate binding. Their highly specialized catalytic mechanisms will potentially limit applications of these P450s (see Chapter 1).

Efficient H_2O_2-driven hydroxylation of hydrophobic substrates such as fatty acids may require a fine balance between active site polarity and hydrophobicity. Peroxidases have polar residues distal to the heme which efficiently bind and cleave H_2O_2 to generate Compound I. P450s have buried, hydrophobic active sites designed for highly specific oxidations with minimal waste of reducing equivalents. The F87A mutation is a key active site modification which promotes peroxygenase activity with H_2O_2, probably by allowing H_2O to remain in the active site after substrate binding, thereby increasing the polarity of the active site so that H_2O_2 can bind and be cleaved. This is supported by the fact that hydrophobic organic peroxides (e.g., t-butyl hydroperoxide and cumene hydroperoxide) can drive the peroxygenase reaction with wild-type P450 BM-3 (see Appendix C). That the F87A mutant is still a faster peroxygenase compared to wild-type with these organic peroxides may be due to steric factors.

B. Additional Mutagenesis Attempts

B.1. Saturation Mutagenesis of Active Site Residues

Based on structural features of natural peroxygenases, it is feasible to imagine that one or more active site amino acid substitutions could facilitate peroxide binding and/or

O-O bond cleavage. Saturation mutagenesis was performed at P450 BM-3 position F87 and at various positions in the F87A mutant active site in hopes of identifying such mutations. T268 in BM-3 corresponds to a conserved residue in the P450 active site which plays an important role in the natural P450 cycle [42-44]. Truan and Peterson showed that cumene hydroperoxide-driven hydroxylation of palmitic acid was unaffected by the T268A mutation (these rates were very slow: ~0.2 nmol/nmol P450/min) [45], indicating that the Thr residue is not essential to peroxide-driven catalysis. I saturated active site residue positions T268, T268 + T269, A264, and A328 with all possible amino acids and screened the libraries using the 12-pNCA assay. Note that this search strategy is limited in that it prevents the identification of improved variants with altered regioselectivity. Exhaustive screening revealed no improved mutants.

B.2. Random Oligonucleotide Cassette Mutagenesis in the Active Site

Random oligonucleotide cassette mutagenesis is a technique used to generate libraries in which a specific region of the protein is targeted for a controlled rate of random mutagenesis [46,47]. This method can be used to find multiple, perhaps interactive, mutations which improve function. I assembled mutant libraries using semi-random oligonucleotides that code for residues 264 through 268, corresponding to the portion of the I helix lying distal to the heme. These oligonucleotides were synthesized such that only a percentage of the bases (20% and 40%) were random. For the 80% wild-type oligonucleotides, this corresponds to ~34% of the library containing double mutations and ~32% containing triple mutations. For the 60% wild-type oligonucleotides, this corresponds to ~12% of the library containing double mutations,

~31% containing triple mutations, and ~37% containing quadruple mutations. A total of ~6000 mutants were screened from these libraries (generated with the F87A mutation) using the 12-pNCA assay, and no improved variants were isolated. Loeb and coworkers have had success with this mutagenesis technique by coupling function to growth (i.e., selection) [48], allowing for much higher fractions of the libraries generated to be searched ($>10^6$ variants). Our screening capabilities do not allow for such sufficient library sampling.

B.3. Engineering the Heme Proximal Region

The environment surrounding the proximal heme ligand (cysteine-thiolate) influences the electronic character of this ligand (60) and consequently the reduction potential of the heme iron and the electron "push" for O-O bond cleavage [49] (refer to Chapter 1). P450 crystal structures show that the Cys ligand sits in a pocket in which the Cys ligand sulfur is near three peptide backbone NH groups [50] which generate a positive charge environment believed to stabilize the thiolate [50]. The residue directly C-terminal to the cysteine is often an Ile, Leu, or Val, but is a Pro in the peroxide-utilizing P450 SPα and CPO (as well as in many plant P450s). A proline in this position eliminates one NH group and imposes new structural constraints, and the functional consequences are not clear. In BM-3 the proximal cysteine ligand is residue 400, and the adjacent residue at 401 is an Ile. I was curious whether the I401P mutation would improve peroxygenase activity. Unfortunately, peroxygenase activity of the heme domain double mutant F87A/I401P was slightly lower than that of F87A.

Although my semi-rational active site design efforts were not successful in identifying improved peroxygenase variants, the screening method was limited. Active site mutations are likely to shift regioselectivity, thereby preventing the 12-pNCA assay from reporting activity. Revisiting these same experiments with a more robust peroxygenase assay, such as screening for peroxide depletion, may identify improved mutants. Whereas an assay based on measuring peroxide consumption was not initially sensitive enough for screening due to low activity and a high peroxide requirement, variant 21B3 may be sufficiently active for developing such an assay, which will allow screening with any substrate of interest.

Conclusions

Natural peroxygenases have polar residues in the active site which play critical catalytic roles. However, hydrophobic pockets are necessary for binding typical P450 substrates. A functional H_2O_2-driven P450 requires that both H_2O_2 and substrate bind efficiently, which suggests a delicate balance between polarity and hydrophobicity is required. The subtle protein modifications required to solve this problem make our system a good target for engineering by directed evolution. The hallmark of directed evolution is the ability to accumulate mutations throughout the protein scaffold with influences on function which could not be predicted. Indeed, it has proven difficult to engineer the P450 by more rational approaches due to our limited understanding of structure-function relationships in proteins and limited screening capabilities (i.e., our inability to exhaustively search for improvements resulting from combinations of active site mutations generated by saturation or random cassette mutagenesis).

Cytochrome P450 BM-3 is a versatile hydroxylase whose "designability" has been demonstrated [51] and whose biotechnological relevance has been established [52]. Directed evolution allows us to engineer into this enzyme functions not required or permitted in its natural biological context. For example, a P450 BM-3 variant which efficiently hydroxylates alkanes was recently described [53]. An efficient H_2O_2-driven P450 BM-3 variant allows us to exploit this powerful and versatile hydroxylase in a cell-free reaction system that requires neither NADPH nor reductase. This catalyst is easy to synthesize (high expression levels are achieved in *E. coli*), requires minimal preparation (crude lysate or purified protein can be used), and is active under mild conditions (i.e., 25^{o}C, 1 atm). The self-sufficiency of the H_2O_2-driven heme domain encourages further protein engineering to explore potential synthetic applications of this catalytic system.

References

1. Labinger JA, Bercaw JE: **Understanding and exploiting C-H bond activation**. *Nature* 2002, **417**:507-514.
2. Crabtree RH: **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* 2001:2437-2450.
3. Ishii Y, Sakaguchi S, Iwahama T: **Innovation of hydrocarbon oxidation with molecular oxygen and related reactions**. *Adv Synth Catal* 2001, **343**:393-427.
4. Costas M, Chen K, Que L: **Biomimetic nonheme iron catalysts for alkane hydroxylation**. *Coord Chem Rev* 2000, **200**:517-544.
5. Maravin GB, Avdeev MV, Bagrii EI: **Oxidative functionalization of saturated hydrocarbons on porphyrin metal complex catalysts (a review)**. *Petrol Chem* 2000, **40**:1-18.
6. Stahl SS, Labinger JA, Bercaw JE: **Homogeneous oxidation of alkanes by electrophilic late transition metals**. *Angew Chem-Int Edit* 1998, **37**:2181-2192.
7. Shilov AE, Shul'pin GB: **Activation of C-H bonds by metal complexes**. *Chem Rev* 1997, **97**:2879-2932.
8. Kopp DA, Lippard SJ: **Soluble methane monooxygenase: activation of dioxygen and methane**. *Curr Opin Chem Biol* 2002, **6**:568-576.
9. Smits TH, Balada SB, Witholt B, van Beilen JB: **Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria**. *J Bacteriol* 2002, **184**:1733-1742.
10. Marin MM, Smits TH, van Beilen JB, Rojo F: **The alkane hydroxylase gene of *Burkholderia cepacia* RR10 is under catabolite repression control**. *J Bacteriol* 2001, **183**:4202-4209.
11. Ratajczak A, Geissdorfer W, Hillen W: **Alkane hydroxylase from *Acinetobacter sp.* strain ADP1 is encoded by alkM and belongs to a new family of bacterial integral-membrane hydrocarbon hydroxylases**. *Appl Environ Microbiol* 1998, **64**:1175-1179.
12. Ortiz de Montellano PR (Ed): *Cytochrome P450: Structure, Mechanism, and Biochemistry,* 2nd edn. New York and London: Plenum Press; 1995.
13. Holland HL, Weber HK: **Enzymatic hydroxylation reactions**. *Curr Opin Biotechnol* 2000, **11**:547-553.
14. Kirk O, Borchert TV, Fuglsang CC: **Industrial enzyme applications**. *Curr Opin Biotechnol* 2002, **13**:345-351.
15. Liese A, Seelbach K, Wandrey C: *Industrial Biotransformations*. Weinheim: Wiley-VCH; 2000.
16. Schmid A, Dordick JS, Hauer B, Kiener A, Wubbolts M, Witholt B: **Industrial biocatalysis today and tomorrow**. *Nature* 2001, **409**:258-268.
17. Li Z, van Beilen JB, Duetz WA, Schmid A, de Raadt A, Griengl H, Witholt B: **Oxidative biotransformations using oxygenases**. *Curr Opin Chem Biol* 2002, **6**:136-144.
18. Duetz WA, van Beilen JB, Witholt B: **Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis**. *Curr Opin Biotechnol* 2001, **12**:419-425.

19. Koeller KM, Wong CH: **Enzymes for chemical synthesis**. *Nature* 2001, **409**:232-240.
20. Mansuy D: **The great diversity of reactions catalyzed by cytochromes P450**. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 1998, **121**:5-14.
21. Joo H, Lin Z, Arnold FH: **Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation**. *Nature* 1999, **399**:670-673.
22. Schwaneberg U, Schmidt-Dannert C, Schmitt J, Schmid RD: **A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A**. *Anal Biochem* 1999, **269**:359-366.
23. Nakagawa S, Ishino S, Teshiba S: **Construction of catalase deficient *Escherichia coli* strains for the production of uricase**. *Biosci Biotechnol Biochem* 1996, **60**:415-420.
24. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH: **Molecular evolution by staggered extension process (StEP) in vitro recombination**. *Nat Biotechnol* 1998, **16**:258-261.
25. Ravichandran KG, Boddupalli SS, Hasermann CA, Peterson JA, Deisenhofer J: **Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's**. *Science* 1993, **261**:731-736.
26. Ozaki S, Ortiz de Montellano PR: **Molecular engineering of horseradish peroxidase - thioether sulfoxidation and styrene epoxidation by Phe-41 leucine and threonine mutants**. *J Am Chem Soc* 1995, **117**:7056-7064.
27. Ozaki S, Matsui T, Watanabe Y: **Conversion of myoglobin into a highly stereospecific peroxygenase by the L29H/H64L mutation**. *J Am Chem Soc* 1996, **118**:9784-9785.
28. Ozaki S, Matsui T, Watanabe Y: **Conversion of myoglobin into a peroxygenase: A catalytic intermediated of sulfoxidation and epoxidation by the F43H/H64L mutant**. *J Am Chem Soc* 1997, **119**:6666-6667.
29. Ozaki S, Yang HJ, Matsui T, Goto Y, Watanabe Y: **Asymmetric oxidation catalyzed by myoglobin mutants**. *Tetrahedron-Asymmetry* 1999, **10**:183-192.
30. Ozaki S, Hara I, Matsui T, Watanabe Y: **Molecular engineering of myoglobin: The improvement of oxidation activity by replacing Phe-43 with tryptophan**. *Biochemistry* 2001, **40**:1044-1052.
31. Savenkova MI, Newmyer SL, Ortiz de Montellano PR: **Rescue of His42->Ala horseradish peroxidase by a Phe41->His mutation - Engineering of a surrogate catalytic histidine**. *J Biol Chem* 1996, **271**:24598-24603.
32. Savenkova MI, Kuo JM, Ortiz de Montellano PR: **Improvement of peroxygenase activity by relocation of a catalytic histidine within the active site of horseradish peroxidase**. *Biochemistry* 1998, **37**:10828-10836.
33. Wan L, Twitchett MB, Eltis LD, Mauk AG, Smith M: **In vitro evolution of horse heart myoglobin to increase peroxidase activity**. *Proc Natl Acad Sci USA* 1998, **95**:12825-12831.
34. Matsuura T, Yomo T, Trakulnaleamsai S, Ohashi Y, Yamamoto K, Urabe I: **Nonadditivity of mutational effects on the properties of catalase I and its application to efficient directed evolution**. *Protein Eng* 1998, **11**:789-795.
35. Trakulnaleamsai S, Yomo T, Yoshikawa M, Aihara S, Urabe I: **Experimental sketch of landscapes in protein-sequence space**. *J Ferment Bioeng* 1995, **79**:107-118.

36. Capdevila JH, Wei S, Helvig C, Falck JR, Belosludtsev Y, Truan G, Graham-Lorence SE, Peterson JA: **The highly stereoselective oxidation of polyunsaturated fatty acids by cytochrome P450BM-3**. *J Biol Chem* 1996, **271**:22663-22671.
37. van Rantwijk F, Sheldon RA: **Selective oxygen transfer catalysed by heme peroxidases: synthetic and mechanistic aspects**. *Curr Opin Biotechnol* 2000, **11**:554-564.
38. Seelbach K, vanDeurzen MPJ, vanRantwijk F, Sheldon RA, Kragl U: **Improvement of the total turnover number and space-time yield for chloroperoxidase catalyzed oxidation**. *Biotechnol Bioeng* 1997, **55**:283-288.
39. vanDeurzen MPJ, Seelbach K, VanRantwijk F, Kragl U, Sheldon RA: **Chloroperoxidase: Use of a hydrogen peroxide-stat for controlling reactions and improving enzyme performance**. *Biocatal Biotransform* 1997, **15**:1-16.
40. Toy PH, Newcomb M, Hager LP: **Hypersensitive radical probe studies of chloroperoxidase-catalyzed hydroxylation reactions**. *Chem Res Toxicol* 1998, **11**:816-823.
41. Sundaramoorthy M, Terner J, Poulos TL: **The crystal structure of chloroperoxidase: A heme peroxidase-cytochrome P450 functional hybrid**. *Structure* 1995, **3**:1367-1377.
42. Martinis SA, Atkins WM, Stayton PS, Sligar SG: **A conserved residue of cytochrome P450 is involved in heme-oxygen stability and activation**. *J Am Chem Soc* 1989, **111**:9252-9253.
43. Kimata Y, Shimada H, Hirose T, Ishimura Y: **Role of Thr-252 in cytochrome P450cam: a study with unnatural amino acid mutagenesis**. *Biochem Biophys Res Commun* 1995, **208**:96-102.
44. Imai M, Shimada H, Watanabe Y, Matsushima-Hibiya Y, Makino R, Koga H, Horiuchi T, Ishimura Y: **Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation**. *Proc Natl Acad Sci USA* 1989, **86**:7823-7827.
45. Truan G, Peterson JA: **Thr268 in substrate binding and catalysis in P450BM-3**. *Arch Biochem Biophys* 1998, **349**:53-64.
46. Dube DK, Loeb LA: **Mutants generated by the insertion of random oligonucleotides into the active site of the beta-lactamase gene**. *Biochemistry* 1989, **28**:5703-5707.
47. Black ME, Loeb LA: **Random sequence mutagenesis for the generation of active enzymes**. *Methods Mol Biol* 1996, **57**:335-349.
48. Skandalis A, Encell LP, Loeb LA: **Creating novel enzymes by applied molecular evolution**. *Chem Biol* 1997, **4**:889-898.
49. Yoshioka S, Takahashi S, Hori H, Ishimori K, Morishima I: **Proximal cysteine residue is essential for the enzymatic activities of cytochrome P450(cam)**. *Eur J Biochem* 2001, **268**:252-259.
50. Poulos TL, Cupp-Vickery JR, Li H: **Structural studies in prokaryotic cytochromes P450**. In *Cytochrome P450: Structure, Mechanism, and Biochemistry,* 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:125-150.
51. Cirino PC, Arnold FH: **Protein engineering of oxygenases for biocatalysis**. *Curr Opin Chem Biol* 2002, **6**:130-135.

52. Guengerich FP: **Cytochrome P450 enzymes in the generation of commercial products**. *Nat Rev Drug Discov* 2002, **1**:359-366.
53. Glieder A, Farinas ET, Arnold FH: **Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase**. *Nat Biotechnol* 2002, **20**:1135-1139.

CHAPTER 4

Thermostabilization of a Cytochrome P450 Peroxygenase

Material from this chapter appears in: Salazar O, Cirino PC, Arnold FH: **Thermostabilization of a cytochrome P450 peroxygenase**. Submitted to *ChemBioChem*, and is reprinted by permission of Wiley-VCH.

Summary

Chapter 3 describes a laboratory-evolved variant of the P450 BM-3 heme domain which functions as an H_2O_2-driven hydroxylase ("peroxygenase") and does not require NADPH, O_2, or the reductase [1]. This variant, which we named 21B3, allows us to carry out cytochrome P450-catalyzed biotransformations under highly simplified reaction conditions: only the heme domain and hydrogen peroxide are needed for substrate (fatty acid) hydroxylation. Because its heme domain alone is competent for catalysis, P450 BM-3 peroxygenase 21B3 offers a unique opportunity to create a thermostable, functional cytochrome P450. Here we report further directed evolution of the 21B3 peroxygenase, resulting in an enzyme which is significantly more thermostable than wild-type cytochrome P450 BM-3 and retains much of the peroxygenase activity of 21B3.

Introduction

Enzymes are often poorly stable under conditions encountered during production, storage or use. Improving enzymes for practical applications has been a major goal of protein engineering [2-5], and improving resistance to thermal denaturation ("thermostability") has been a major focus [6-8]. Improved thermostability often correlates with longer shelf-life, longer life-time during use (even at low temperatures), and a higher temperature optimum for activity [9,10].

There have been no reports of stabilizing the relatively unstable cytochrome P450 enzymes by protein engineering, however, primarily because the P450s comprise multiple subunits and contain thermolabile cofactors. P450 BM-3 heme domain containing the single amino acid substitution F87A (mutant HF87A) is significantly more active than wild-type heme domain (HWT) in reactions driven by H_2O_2 [11,12]. Variant 21B3 is much more active than HF87A in 10 mM H_2O_2, but is also less thermostable than HWT and HF87A. We therefore sought to improve the thermostability of 21B3 while maintaining its improved peroxygenase activity. While characteristic structural features of thermophilic enzymes relative to their mesophilic homologs have been identified (e.g., increased hydrophobicity, shortened loops, smaller cavities, and increased hydrogen bonding and salt bridges) [13-15], there are no good rules of thumb for introducing thermostabilizing mutations into a protein. Random mutagenesis and directed evolution on the other hand have proven effective for increasing enzyme thermostability [8,16,17].

Results

Four cycles of random mutagenesis and screening for retention of peroxygenase activity after heat treatment (refer to Methods section) yielded thermostable peroxygenase variant "TH4". Figure 4.1 shows the peroxygenase activity and stability of mutant TH4 compared to HF87A and 21B3. Thermostability in Figure 4.1 is represented as the temperature at which, after a 10-minute incubation, 50% of the initial room-temperature peroxygenase activity is retained. Figure 4.2 shows the initial rates of these three mutants as a function of H_2O_2 concentration. TH4 has a lower $K_{m,app}$ for H_2O_2 compared to both HF87A and 21B3. Thus, we were able to regain (and exceed) the thermostability of HF87A from mutant 21B3 without sacrificing the improved peroxygenase activity. TH4 was then used as the parent of another random mutagenesis library. Since we were unable to isolate from this library variants which were both more stable and more active, we chose to recombine the genes of mutants which were either more active or more thermostable using DNA shuffling [18]. Screening this recombinant library resulted in thermostable variant 5H6.

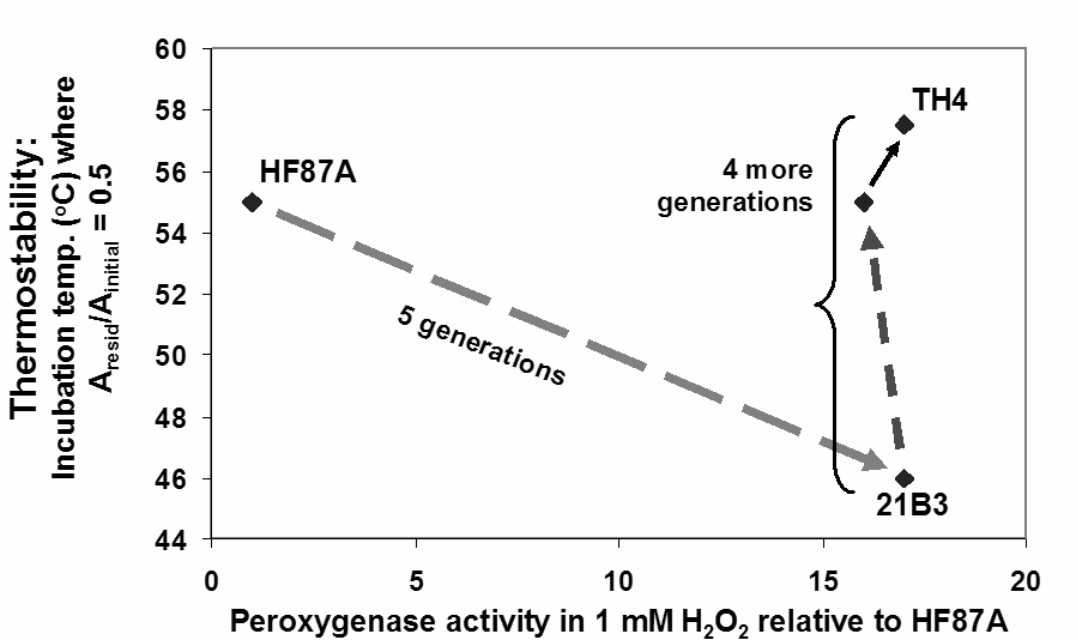


Figure 4.1. Thermostability and activities of HF87A, 21B3, and TH4.

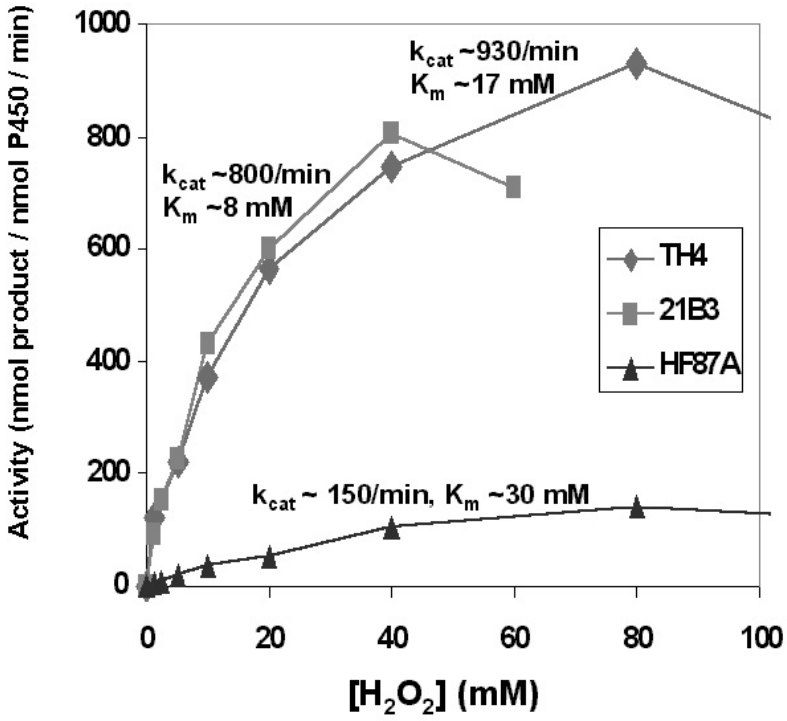


Figure 4.2. Peroxygenase activities (initial rates) of HF87A, 21B3, and TH4 in different concentrations of H_2O_2.

To characterize thermostability, we measured the fraction of folded heme domain remaining after heat-treatment, which we determined from the fraction of the ferrous heme-CO complex that retained the 450 nm absorbance peak characteristic of properly folded P450. Figure 4.3 shows the percentage of properly folded heme domain protein remaining after 10-minute incubations at different, elevated temperatures. To allow comparison to the wild-type full-length enzyme (BWT), whose stability is limited by the stability of the reductase domain and therefore cannot be determined from the CO-binding measurement, we determined the residual (NADPH-driven) activity of BWT following 10-minute incubations at the same temperatures. Data in Figure 4.3 were fit to a two-state model, and the resulting calculated temperatures corresponding to half denaturation for the 10-minute heat incubations (T_{50}) are listed in Table 4.1. These values are in good agreement with the midpoints of the melting curves of Figure 4.3. According to this measure of stability, variant 5H6 ($T_{50} = 61^{o}C$) is much more thermostable than the natural catalytic system, BWT ($T_{50} = 43^{o}C$). It is also significantly more thermostable than HF87A and 21B3.

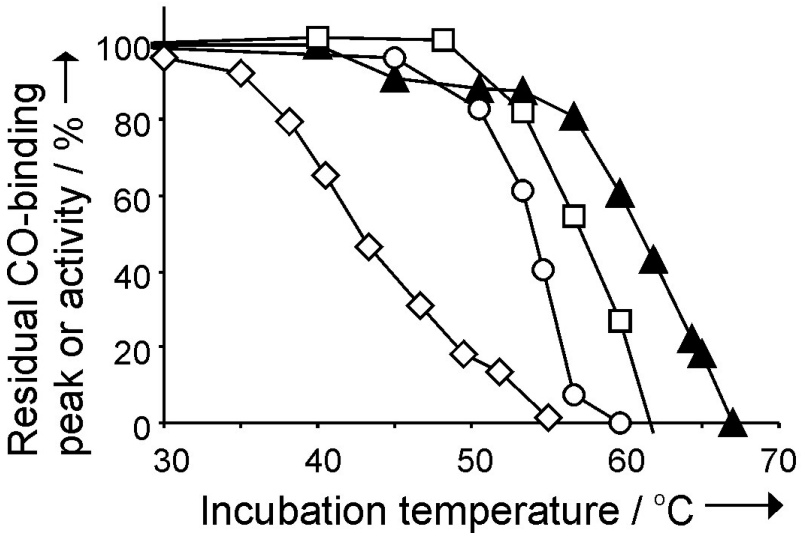


Figure 4.3. Percentage of 450 nm CO-binding peak of cytochrome P450 BM-3 heme domain HWT (□), HF87A (○) and 5H6 (▲) remaining after 10-minute incubation at the indicated temperatures. For the holoenzyme BWT (◇), the percentage of initial NADPH-driven activity remaining after 10-minute incubations is shown.

Table 4.1. Thermostability and activity parameters for evolved and parental P450s. BWT = full-length, wild-type P450 BM-3; HWT = wild-type P450 BM-3 heme domain; HF87A = P450 BM-3 heme domain containing mutation F87A; 21B3 & 5H6 = evolved heme domain peroxygenase variants.

Mutant	T_{50} for 10-minute incubations[a] (°C)	$t_{1/2}$ at 57.5°C[b] (minutes)	Peroxygenase Activity[c] (minute^{-1})
BWT	43	0.46	<5
HWT	57	n.d.	<5
HF87A	54	2.3	23
21B3	46	n.d.	430
5H6	61	115	220

[a] Calculated from the data in Figure 1, fit to two-state denaturation equation. [b] Calculated from the data in Figure 2, fit to a first-order exponential decay equation. [c] Reported as initial rates at room temperature on 12-pNCA in 10 mM H_2O_2 and 6% DMSO. n.d.: not determined.

Peroxygenase activities were measured at room temperature, using a colorimetric assay with 12-*p*-nitrophenoxycarboxylic acid (12-pNCA) as substrate (see [19] and Chapter 1). 5H6 retains ~50% of the high activity of 21B3 and is almost ten times as active as HF87A (Table 4.1). Another useful measure of enzyme stability comes from the rate of inactivation at high temperature. Figure 4.4 shows the percentage of activity that remains for the different variants upon heating at 57.5^{o}C. The activities decay exponentially with time (first-order), and the half-life ($t_{1/2}$) of each catalytic system is listed in Table 4.1. HF87A (which is less thermostable than HWT) is significantly more resistant to inactivation at 57.5^{o}C compared to BWT. (The half-life of HF87A is also higher than that of BWT at room temperature (data not shown).) The half-life of 5H6 at 57.5^{o}C is 50 times longer than that of HF87A and 250 times longer than BWT. The fraction of peroxygenase activity remaining after heat treatment correlated with the fraction of remaining CO-binding peak for HF87A and 5H6. (Residual activity of HWT could not be correlated to the remaining CO-binding peak because HWT has essentially no peroxygenase activity.)

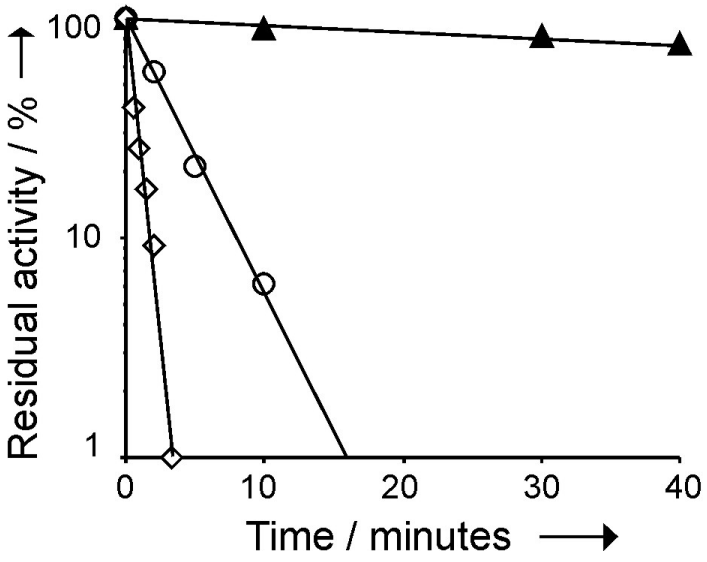


Figure 4.4. Heat-inactivation of cytochrome P450 BM-3 holoenzyme BWT (◇) and peroxygenase mutants HF87A (○) and 5H6 (▲), calculated as the percentage of activity remaining after incubation at 57.5°C for the indicated periods of time. Peroxygenase activity was measured for HF87A and 5H6, while NADPH-driven activity was measured for BWT.

Figures 4.3 and 4.4 show that the peroxide-driven catalytic system, requiring only the heme domain, is significantly more thermostable than the natural catalytic system (BWT). This may be in part due to a greater instability of the reductase domain compared to the heme domain, or a greater instability of one or more protein components involved in the electron transfer process used by the NADPH pathway compared to the heme domain. Regardless, this is strong evidence that it is easier to engineer stability in the heme domain alone than in the full length P450 BM-3 enzyme.

Thermostable peroxygenase 5H6 contains eight new amino acid substitutions compared to 21B3. Refer to Chapter 5 for a description of these mutations. Enzyme thermostabilization often leads to a shift in the activity-temperature profile to higher temperatures, reflecting the higher stability of the folded protein [9]. Measurements of peroxygenase activity at different temperatures, however, showed no significant increase

in the optimum temperature for activity for 5H6 compared to HF87A (both were 25-30^{o}C).

As described in Chapter 1, two thermostable cytochrome P450s (CYP119 and CYP175A1) from thermophilic organisms have recently been identified [20,21] and their (heme domain) crystal structures determined [22-24]. CYP119 exhibits a melting temperature of ~91^{o}C. Aromatic stacking, salt-link networks and shortened loops are believed to help stabilize these enzymes. Unfortunately, the functions of these P450s are not known, and reported activities are extremely low (e.g., 0.35 min^{-1} in the NADH-driven hydroxylation of lauric acid [25]).

Directed evolution allows us to explore "unnatural" functions and properties of P450s. In the first example of engineering a P450 to resist thermal denaturation, we have used directed evolution to improve the thermostability of BM-3 peroxygenase variant 21B3. These results reinforce the biotechnological relevance of cytochrome P450 BM-3 and the catalytic potential of our biomimetic hydroxylase.

References

1. Cirino PC, Arnold FH: **A self-sufficient peroxide-driven hydroxylation biocatalyst**. *Angew Chem-Int Edit* 2003, **In Press**.
2. Arnold FH: **Combinatorial and computational challenges for biocatalyst design**. *Nature* 2001, **409**:253-257.
3. Kirk O, Borchert TV, Fuglsang CC: **Industrial enzyme applications**. *Curr Opin Biotechnol* 2002, **13**:345-351.
4. Brannigan JA, Wilkinson AJ: **Protein engineering 20 years on**. *Nat Rev Mol Cell Biol* 2002, **3**:964-970.
5. Dalboge H, Borchert TV: **Protein engineering of enzymes**. *Biochim Biophys Acta* 2000, **1543**:203-455.
6. Lehmann M, Pasamontes L, Lassen SF, Wyss M: **The consensus concept for thermostability engineering of proteins**. *Biochim Biophys Acta* 2000, **1543**:408-415.
7. Lehmann M, Wyss M: **Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution**. *Curr Opin Biotechnol* 2001, **12**:371-375.
8. Wintrode PL, Arnold FH: **Temperature adaptation of enzymes: lessons from laboratory evolution**. In *Evolutionary Protein Design*. Edited by Arnold FH: Academic Press; 2001:161-225.
9. Daniel RM, Danson MJ, Eisenthal R: **The temperature optima of enzymes: a new perspective on an old phenomenon**. *Trends Biochem Sci* 2001, **26**:223-225.
10. Martinek K, Mozhaev VV: **Need for thermostability, its benefits and main strategies for thermostabilization**. In *Thermostability of Enzymes*. Edited by Gupta MN: Springer-Verlag; 1993.
11. Li QS, Ogawa J, Shimizu S: **Critical role of the residue size at position 87 in H2O2-dependent substrate hydroxylation activity and H2O2 inactivation of cytochrome P450BM-3**. *Biochem Biophys Res Commun* 2001, **280**:1258-1261.
12. Cirino PC, Arnold FH: **Regioselectivity and activity of cytochrome P450BM-3 and mutant F87A in reactions driven by hydrogen peroxide**. *Adv Synth Catal* 2002, **344**:932-937.
13. Russell RJ, Taylor GL: **Engineering thermostability: lessons from thermophilic proteins**. *Curr Opin Biotechnol* 1995, **6**:370-374.
14. Perl D, Schmid FX: **Some like it hot: The molecular determinants of protein thermostability**. *Chembiochem* 2002, **3**:39-44.
15. Jaenicke R, Bohm G: **The stability of proteins in extreme environments**. *Curr Opin Struct Biol* 1998, **8**:738-748.
16. Miyazaki K, Wintrode PL, Grayling RA, Rubingh DN, Arnold FH: **Directed evolution study of temperature adaptation in a psychrophilic enzyme**. *J Mol Biol* 2000, **297**:1015-1026.
17. Gershenson A, Arnold FH: **Enzyme stabilization by directed evolution**. *Genet Eng (N Y)* 2000, **22**:55-76.
18. Stemmer WP: **Rapid evolution of a protein in vitro by DNA shuffling**. *Nature* 1994, **370**:389-391.

19. Schwaneberg U, Schmidt-Dannert C, Schmitt J, Schmid RD: **A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A**. *Anal Biochem* 1999, **269**:359-366.
20. Chang YT, Loew G: **Homology modeling, molecular dynamics simulations, and analysis of CYP119, a P450 enzyme from extreme acidothermophilic archaeon *Sulfolobus solfataricus***. *Biochemistry* 2000, **39**:2484-2498.
21. McLean MA, Maves SA, Weiss KE, Krepich S, Sligar SG: **Characterization of a cytochrome P450 from the acidothermophilic archaea *Sulfolobus solfataricus***. *Biochem Biophys Res Commun* 1998, **252**:166-172.
22. Yano JK, Koo LS, Schuller DJ, Li H, Ortiz de Montellano PR, Poulos TL: **Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus***. *J Biol Chem* 2000, **275**:31086-31092.
23. Park SY, Yamane K, Adachi S, Shiro Y, Weiss KE, Maves SA, Sligar SG: **Thermophilic cytochrome P450 (CYP119) from *Sulfolobus solfataricus*: high resolution structure and functional properties**. *J Inorg Biochem* 2002, **91**:491-501.
24. Park SY, Yamane K, Adachi S, Shiro Y, Weiss KE, Sligar SG: **Crystallization and preliminary X-ray diffraction analysis of a cytochrome P450 (CYP119) from *Sulfolobus solfataricus***. *Acta Crystallogr D Biol Crystallogr* 2000, **56 (Pt 9)**:1173-1175.
25. Puchkaev AV, Koo LS, Ortiz de Montellano PR: **Aromatic stacking as a determinant of the thermal stability of CYP119 from *Sulfolobus solfataricus***. *Arch Biochem Biophys* 2003, **409**:52-58.

CHAPTER 5

Description and Analysis of Mutations Acquired During Directed Evolution

Mutations Revealed by Sequencing Select Mutants

All mutations described in this chapter are in addition to those in variant HF87A. Nucleotide and amino acid sequences of P450 BM-3 are available in Appendix B. Table 5.1 lists the amino acid substitutions found from sequencing various mutants selected throughout the evolution of peroxygenase activity and thermostability in P450 BM-3 heme domain mutant F87A (HF87A). The parent used to generate each mutant is included in parentheses. The first round of evolution (using HF87A as the parent) resulted in (more active) mutants "2H1", "1F8", and "2E10". Two separate second generations resulted in several more mutants with increased peroxygenase activity, including "2E10-1", "2E10-2", "2E10-3", and "2E10-4" (from parent "2E10") and "1F8-1" and "1F8-2" (from parent "1F8"). Recombination was then performed using HF87A, 2H1, "1F8-1", "1F8-2", "2E10-1", "2E10-2", "2E10-3", and "2E10-4" as the parent genes. Mutants "step B3" and "step B6" were found from screening this recombinant library for mutants with improved activity in both 1 mM and 50 mM H_2O_2. Table 5.2 lists the nucleotide and amino acid substitutions in mutants "step B3" and "step B6" (in addition to those found in HF87A).

Starting with "step B3" as the parent, two further rounds of evolution (using error-prone PCR) resulted in peroxygenase variant "21B3" (described in Chapter 3). Nucleotide substitutions in "21B3" are listed in Table 5.3. As described in Chapter 4, four rounds of directed evolution to improve thermostability using error-prone PCR (starting with "21B3" as the parent) resulted in variant "TH4", whose nucleotide substitutions are listed in Table 5.4. Two final rounds of evolution (first using error-

prone PCR and then recombination, as described in Chapter 4) resulted in mutant "5H6", whose nucleotide substitutions are listed in Table 5.5.

Table 5.1. Amino acid substitutions (in addition to F87A) found from sequencing selected mutants throughout evolution. The parent of each mutant is indicated in parenthesis. "(rec)" indicates that that mutant is from a recombinant library with multiple parents.

Mutant (Parent)	Amino Acid Mutations in Addition to HF87A
2H1 (HF87A)	K434E
1F8 (HF87A)	K9I, H100R
2E10 (HF87A)	K113E, K434E
2E10-1 (2E10)	K113E, D217V, and K434E
2E10-3 (2E10)	E93G, K113E, N186S, and K434E
2E10-4 (2E10)	K113E, M237L, and K434E
step B3 (rec)	H100R, M145V, S274T, and K434E
step B6 (rec)	H100R, M145V, M237L, and K434E
21B3 (11C3)	I58V, H100R, F107L, A135S, M145V, N239H, S274T, K434E, and V446I
TH3 (TH2)	I58V, H100R, F107L, A135S, M145V, N239H, S274T, L324I, I366V, K434E, E442K, and V446I
TH4 (TH3)	I58V, H100R, F107L, A135S, M145A, N239H, S274T, L324I, I366V, K434E, E442K, and V446I
5H6 (rec)	L52I, F87A, H100R, F107L, A135S, A184V, N239H, S274T, L324I, V340M, I366V, K434E, E442K, and V446I (plus one H deleted in 6-H tag)

Table 5.2. Substitutions in "step B3" and "step B6" P450 BM-3 variants (in addition to those in HF87A). 'X' indicates the substitution was present; '-' indicates the substitution was not present.

Base Substitution	Amino Acid Substitution	step B3	step B6
A299G	H100R	X	X
A433G	M145V	X	X
A709T	M237L	-	X
T820A	S274T	X	-
T1188A	(SYNONYMOUS)	X	X
A1300G	K434E	X	X

Table 5.3. Mutations in peroxygenase variant 21B3, in addition to those in HF87A.

Base Change	Amino Acid Change
A172G	I58V
A195T	(SYNONYMOUS)
A299G	H100R
C321A	F107L
G403T	A135S
A433G	M145V
A684G	(SYNONYMOUS)
A715C	N239H
T810C	(SYNONYMOUS)
T820A	S274T
T1188A	(SYNONYMOUS)
A130OG	K434E
G1336A	V446I

Table 5.4. Mutations in thermostable peroxygenase variant TH4, in addition to those in HF87A. (Percentage values represent the changes in codon usage by *E. coli*)

Base Change(s)	**Amino Acid Change**
A172G	I58V
A195T	SYNONYMOUS (S);14% to 15%
A299G	H100R
C321A	F107L
G403T	A135S
A433G + T434C	M145A
A684G	SYNONYMOUS (E); 67% to 33%
A715C	N239H
T810C	SYNONYMOUS (S); 16% to 26%
T820A	S274T
T970A	L324I
A1096G	I366V
T1188A	SYNONYMOUS (G); 33% to 13%
A1300G	K434E
T1309C	SYNONYMOUS (L); 14% to 4%
G1324A	E442K
G1336A	V446I

Table 5.5. Mutations in thermostable peroxygenase variant 5H6, in addition to those in HF87A. (Percentage values represent the changes in codon usage by *E. coli*)

Base Change(s)	**Amino Acid Change**
T154A	L52I
A172G	I58V
A195T	SYNONYMOUS (S);14% to 15%
A299G	H100R
C318G	S106R
C321A	F107L
G403T	A135S
C489T	SYNONYMOUS (N); 52% to 48%
C551T	A184V
A684G	SYNONYMOUS (E); 67% to 33%
A715C	N239H
T810C	SYNONYMOUS (S); 16% to 26%
T820A	S274T
T970A	L324I
G1018A	V340M
A1096G	I366V
T1188A	SYNONYMOUS (G); 33% to 13%
A1300G	K434E
T1309C	SYNONYMOUS (L); 14% to 4%
G1324A	E442K
G1336A	V446I
CAT (1405, 1406, 1407)	DELETED

Location of Mutations in the Heme Domain Crystal Structure

Figure 5.1 depicts the structural layout of the P450 BM-3 heme domain crystal structure [1]. Helices are represented by black bars and strands of β-sheets are shown as open rectangles. The structural elements are grouped into the α-helical-rich domain and the β-sheet-rich domain. The heme is shown by the oval at the NH_2-terminal end of the L-helix. Figures 5.2 and 5.3 show two alternate views of the P450 BM-3 heme domain crystal structure [2], represented as ribbon drawings. Secondary structural elements shown in Figure 5.1 are labeled in Figures 5.2 and 5.3. Mutations acquired during

evolution of peroxygenase activity and which appear in mutant "21B3" are shown as black balls in Figures 5.2 and 5.3. Mutations acquired during further evolution of thermostability and which appear in variant "5H6" are shown as gray balls in Figures 5.2 and 5.3. Finally, Table 5.6 describes where the various mutations listed above appear within the secondary structure elements of the heme domain. Table 5.6 also indicates whether the residue position is exposed to the protein surface or is buried.

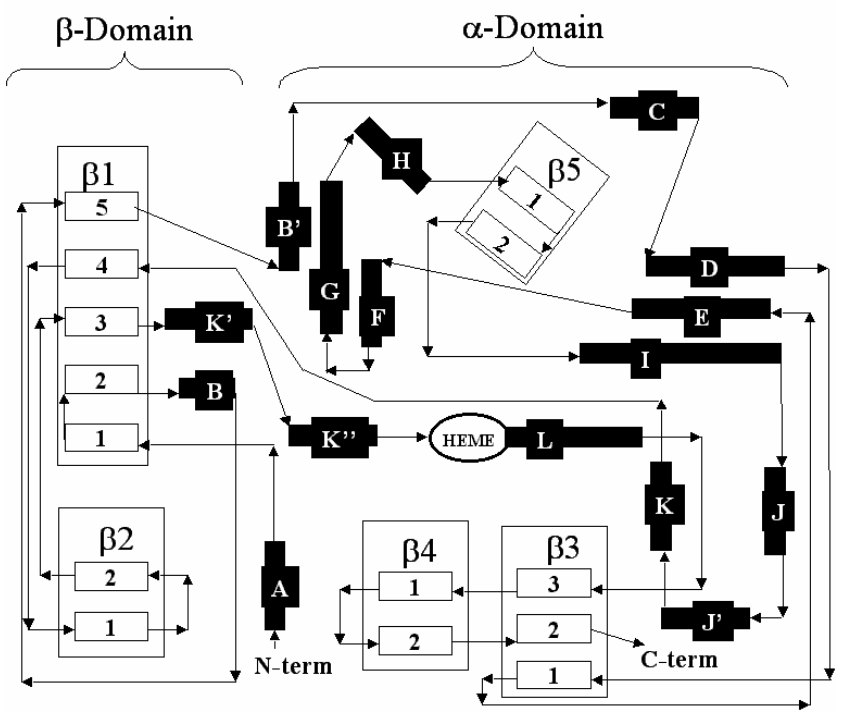


Figure 5.1. Topology drawing of P450 BM-3. Helices are represented by black bars and strands of β-sheets are shown with open rectangles. The structural elements are grouped into the α-helical-rich domain and the β-sheet-rich domain. The heme is shown by the oval at the NH_2-terminal end of the L-helix.

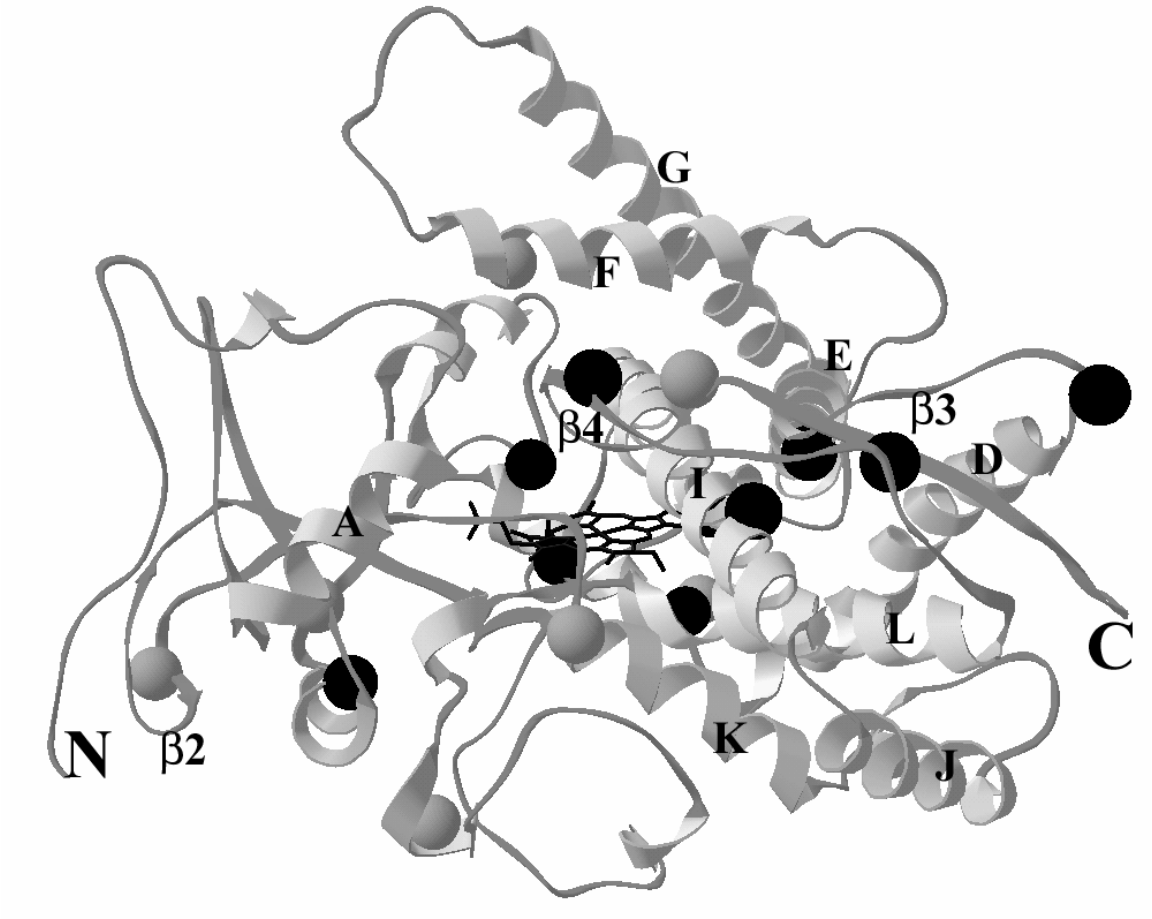


Figure 5.2. Ribbon drawing of the wild-type cytochrome P450 BM-3 heme domain with secondary structure elements labeled. The β-domain appears to the left of the figure. Mutations acquired during evolution of peroxygenase activity and which appear in mutant "21B3" are shown as black balls. Mutations acquired through further evolution of thermostability and which appear in variant "5H6" are shown as gray balls. The atomic coordinates of P450 BM-3 [2] were used to create this image with the Swiss PDB Viewer.

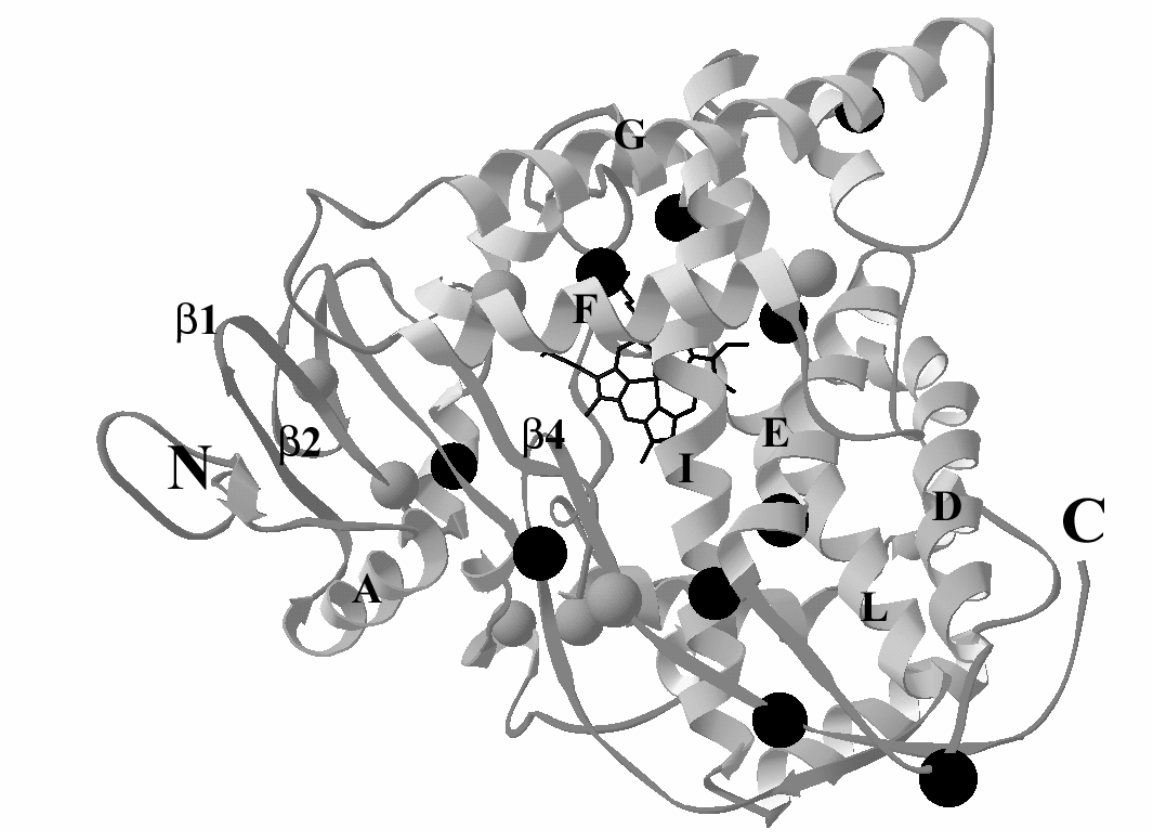


Figure 5.3. Alternate view (looking at Figure 5.2 from the top) of the wild-type cytochrome P450 BM-3 heme domain, with secondary structure elements labeled. Mutations acquired through evolution of peroxygenase activity and which appear in mutant "21B3" are shown as black balls. Mutations acquired through further evolution of thermostability and which appear in variant "5H6" are shown as gray balls. The atomic coordinates of P450 BM-3 [2] were used to create this image with the Swiss PDB Viewer.

Table 5.6. Locations of select amino acid substitutions (in addition to F87A)

Mutation	Position in Heme Domain Structure
L52I	β-sheet 1-2 (buried)
I58V	helix B (buried)
H100R	helix C (surface)
S106R	loop between helices C & D (surface)
F107L	loop between helices C & D (buried)
A135S	loop between helices D & E (surface)
M145V/A	helix E (buried)
A184V	helix F (buried)
N239H	end of helix H (surface)
S274T	helix I (buried)
L324I	end of helix K (surface)
V340M	β-sheet 2-1 (surface)
I366V	helix K' (surface)
K434E	β-sheet 4-1 (surface)
E442K	end of β-sheet 4-2 (surface)
V446I	β-sheet 3-2 (buried)

Analysis of Mutations

A. Mutations in "21B3"

The mutations in "21B3" are dispersed throughout the protein scaffold. No mutations appear in the active site or the substrate binding channel. Additionally, there are no mutations between amino acid positions 146 and 238, corresponding to the F and G helices of the heme domain, which undergo major conformational changes to effect catalysis upon substrate binding [3]. Interestingly, in a similar study in which BM-3 was evolved to hydroxylate alkanes (under the natural NADPH pathway), five out of 11 mutations occurred in the F and G helices [4]. There are also no mutations in any β-

sheets of the β-sheet-rich region of the structure, and four mutations in 21B3 lie on the protein surface.

Mutations H100R and F107L are less than 5 Å from the heme edge on the proximal side. Sequence alignments with other P450s indicate that H100 corresponds to a conserved residue position where the residue is typically an Arg [5]. These residues are believed to participate in hydrogen bonding to one of the heme propionate groups. In BM-3, H100 lies too far from the heme to directly hydrogen bond with the heme propionate. Instead, H100 is in hydrogen bonding distance of a crystallographically ordered water molecule, which in turn hydrogen bonds with a heme propionate.[1] It is possible that His and Arg are interchangeable in this role, or that the H100R mutation replaces the water molecule and directly interacts with the heme. The crystal structure of the heme-FMN domain interface indicates that H100 also forms a salt bridge in this complex [6], a function no longer necessary in our system.

B. Thermostabilizing Mutations

The only difference between the mutations in "TH4" and the mutations in the mutant from the previous generation (mutant "TH3", which was the parent used to generate the library that resulted in TH4) is that (previously occurring) mutation M145V was further changed to M145A. Thus, throughout the course of evolving shunt pathway activity and stability, a single codon was mutated on two separate occasions, resulting in an amino acid (Ala) that could not be reached by a single base mutation. This mutation later reverted to Met upon back-crossing with HF87A during DNA shuffling, resulting in

mutant “5H6”. M145 and mutations M145V and M145A probably all have very similar effects on activity and stability (i.e., are nearly neutral).

Thermostable peroxygenase mutant “5H6” contains eight new amino acid substitutions compared to “21B3” and 15 substitutions compared to HF87A. “5H6” also contains a deletion resulting in the removal of one His residue from the 6-His sequence included at the C-terminus. Five of the amino acid substitutions in “5H6” are conservative with regard to hydrophobicity and size: L52I, A184V, L324I, V340M and I366V. Ser 106 was converted to a positively charged Arg residue (S106R) and a negatively-charged Glu residue was converted to a positively-charged Lys (E442K). It is difficult to rationalize how these mutations increase the enzyme’s stability. According to the heme domain crystal structure, substitutions S106R, L324I, V340M, I366V, and E442K are located on the protein surface; the others are buried.

As shown in Figure 5.4, four stabilizing mutations are close to positions where mutations that improved peroxygenase activity accumulated in earlier experiments: L52I (in β-sheet 1-2) is adjacent to I58V (helix B) from 21B3, S106R (in a loop connecting helices C and D) lies next to mutation F107L from 21B3, E442K (in β-sheet 4-2) lies adjacent to K434E (in β-sheet 4-1) from 21B3, and the reversion to M145 (helix E) is adjacent to S274T (helix I) from 21B3. The new stabilizing mutations may serve to alleviate structural perturbations introduced by the original (activity) mutations.

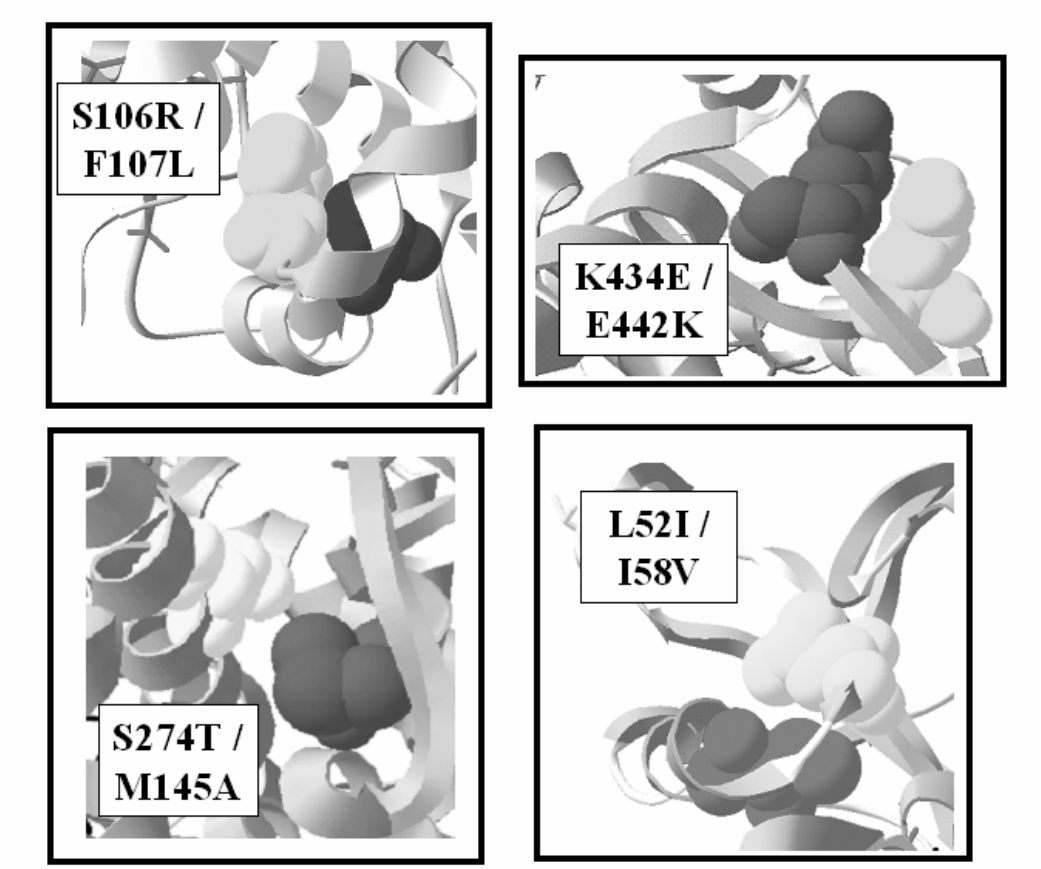


Figure 5.4. Four residue positions where mutations acquired during evolution of thermostability (L52, S106, E442, and M145) lie adjacent to positions (in the heme domain structure) where mutations were previously acquired during evolution of peroxygenase activity.

C. Additional Studies on Various Mutations

Mutation K434E appeared in two separately evolved mutants ("2H1" and "2E10"), indicating this mutation or other mutations at this position might be especially effective in improving peroxygenase activity. Saturation mutagenesis at position 434 in mutants "step B3" and "step B6" did not identify other mutations with higher activity.

The purpose of recombining sequences from several different mutants is to explore the additive effects of beneficial point mutations and to remove deleterious mutations. Therefore the amino acid substitutions found in mutants "step B3", "step B6",

and “5H6” are considered to be particularly important for improving peroxygenase activity and/or stability since they were selected by screening recombinant products. While our StEP library successfully generated improved mutants as a result of multiple crossovers, I was concerned that the StEP PCR reaction was not able to generate all combinations of crossovers since many mutations from the parents are close to one another in sequence. To test combinations which may not have been synthesized in the StEP PCR reaction, all combinations of E93E/G, K113K/E, N186N/S, and M237M/L were created in mutant “step B3” using two different PCR methods: assembly PCR by overlap extension [7] and a recently described “ligation during amplification” (LDA) method [8]. Figure 5.5 shows the activities of the resulting mutants. Assembly PCR resulted in mutant “B3 ASS”, which has ~2.6-fold higher activity compared to “step B3” in 1 mM H_2O_2 (see Figure 5.5). In addition to the mutations found in “step B3”, this mutant contained mutations K113E, N186S, and M237L. “B3 ASS” also contained two new amino acid substitutions (H236R and E372D) as a result of using HotStar DNA Taq polymerase (Qiagen), which apparently has extremely low fidelity. The LDA method resulted in mutant “B3 LDA”, which has ~2.8-fold higher activity compared to “step B3” in 1 mM H_2O_2. This mutant was not sequenced.

Note from Table 5.2 that there is only one mutation different between “step B3” and “step B6” (“step B3” has S274T, while “step B6” has M237L). Mutation M237L was added directly to mutant “step B3”. This mutation increases peroxygenase activity (see Figure 5.5) but also reduces the enzyme’s structural stability, as indicated by the appearance of a 420 nm absorbance peak in the CO-binding difference spectrum of the reduced heme [9] in all mutants containing this mutation.

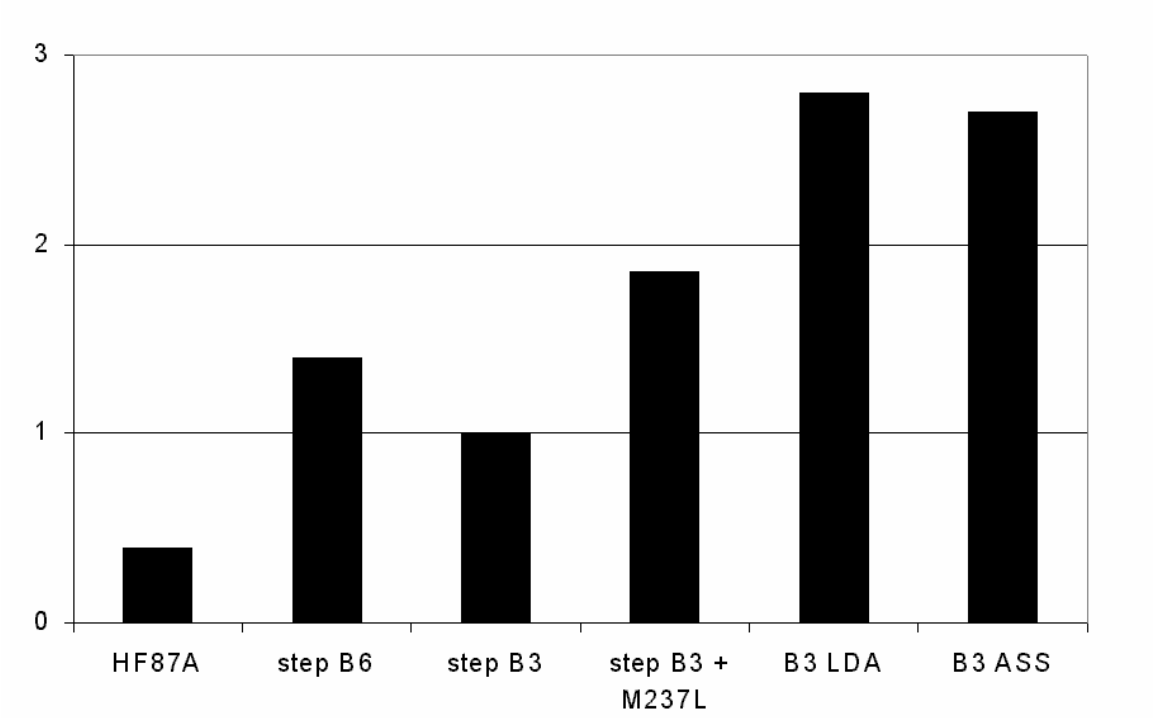


Figure 5.5. Peroxygenase activities of various mutants in 1 mM H_2O_2 relative to the activity of mutant "step B3".

References

1. Peterson JA, Graham-Lorence S: **Bacterial P450s: structural similarities and functional differences**. In *Cytochrome P450 Structure, Mechanism, and Biochemistry*, 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:151-180.
2. Li H, Poulos TL: **Modeling protein-substrate interactions in the heme domain of cytochrome P450BM-3**. *Acta Crystallogr.* 1994, **D51**:21-32.
3. Li HY, Poulos TL: **Fatty acid metabolism, conformational change, and electron transfer in cytochrome P-450(BM-3)**. *BBA-Mol Cell Biol Lipids* 1999, **1441**:141-149.
4. Glieder A, Farinas ET, Arnold FH: **Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase**. *Nat Biotechnol* 2002, **20**:1135-1139.
5. Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, et al.: **P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature**. *Pharmacogenetics* 1996, **6**:1-42.
6. Sevrioukova IF, Li HY, Zhang H, Peterson JA, Poulos TL: **Structure of a cytochrome P450-redox partner electron-transfer complex**. *Proc Natl Acad Sci USA* 1999, **96**:1863-1868.
7. McPherson MJ, Moller SG: **PCR Mutagenesis**. In *PCR*. Edited by McPherson MJ, Moller SG: BIOS Scientific Publishers Limited; 2000:143-182.
8. Sawano A, Miyawaki A: **Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis**. *Nucleic Acids Res* 2000, **28**:E78.
9. Martinis SA, Blanke SR, Hager LP, Sligar SG, Hoa GHB, Rux JJ, Dawson JH: **Probing ng the heme iron coordination structure of pressure-induced cytochrome P420(cam)**. *Biochemistry* 1996, **35**:14530-14536.

CHAPTER 6

Peroxide-Mediated Inactivation of Cytochrome P450 BM-3 Heme Domain and the Effect of Replacing Methionine with Norleucine on Peroxygenase Activity

Material from this chapter appears in: Cirino PC, Tang Y, Takahashi K, Tirrell DA, Arnold FH: **Global incorporation of norleucine in place of methionine in cytochrome P450 BM-3 heme domain increases peroxygenase activity**. *Biotechnol Bioeng* 2003, **In Press**, and is reprinted by permission of John Wiley & Sons, Inc.

Summary

Our laboratory-evolved P450 BM-3 heme domain peroxygenase mutants are subject to rapid H_2O_2-mediated enzyme inactivation. Inactivation corresponds to heme degradation and appears to be primarily turnover-dependent. As a result of using DMSO in the screening conditions for evolution, maximum activity and TON for the evolved mutants is achieved in the presence of DMSO, and DMSO apparently protects the heme during catalysis. It is not clear whether protein oxidation or modification also plays a role in enzyme inactivation. We therefore replaced all 13 methionine residues in heme domain mutant TH4 with the isosteric methionine analog norleucine as an efficient way to test whether inactivation during peroxide-driven P450 catalysis involves methionine oxidation. Additionally, this experiment provided a means of testing the functional limits of globally incorporating into an enzyme an unnatural amino acid in place of its natural analog. While there was no increase in the stability of TH4 under standard reaction conditions (in 10 mM H_2O_2), complete substitution with norleucine resulted in nearly twofold increased peroxygenase activity. Thermostability was significantly reduced. The fact that the enzyme can tolerate such extensive amino acid replacement suggests that we can engineer enzymes with unique chemical properties via incorporation of unnatural amino acids while retaining or improving catalytic properties. This system additionally provides a platform for directing enzyme evolution using an extended set of protein building blocks.

Introduction

The cytochrome P450 monooxygenases are among the most widely studied enzymes. Their physiological roles are important in drug metabolism and drug design, and their catalytic capabilities are increasingly valued in chemical synthesis and bioremediation (see [1-4]). Cytochrome P450 BM-3 is one of the fastest fatty acid hydroxylases known, preferring straight-chain substrates with 12 to 18 carbons [5,6].

As described in the preceding chapters, directed evolution has been used to isolate mutants of P450 BM-3 with higher k_{cat} values for the peroxygenase reaction and lower K_m values for H_2O_2 binding, eliminating the need for the reductase domain, NADPH and O_2 to drive catalysis. After several rounds of mutagenesis and screening for peroxygenase activity, we generated a highly active, thermostable BM-3 heme domain variant ("TH4"; refer to Chapter 4) with up to 20-fold improved peroxygenase activity compared to HF87A. Using increased TON as the selection criterion we hoped to identify mutants which were more stable to peroxide (higher TONs would be reached). We found that mutants with higher TONs had higher initial peroxygenase rates but were no more stable to peroxide. TH4 is still rapidly inactivated by H_2O_2: in 10 mM H_2O_2, the enzyme is essentially inactive after 10 minutes (similar to mutant 21B3, as shown in Figure 3.5). Experiments are described below in which methionines in TH4 are replaced with norleucine. The primary goal of this experiment was to test whether enzyme inactivation is also partly due to protein oxidation by H_2O_2. This experiment additionally allowed us to probe the limits of globally incorporating an unnatural amino acid into an enzyme. Following a discussion of the results from those experiments, we characterize and analyze H_2O_2-mediated inactivation.

Replacement of Methionine with Norleucine in TH4

A. Introduction

Although H_2O_2-mediated inactivation of TH4 is primarily the result of heme degradation, we nonetheless wanted to test whether protein modification played a role in inactivating the enzyme. Methionine residues are prime targets for oxidation by peroxide [7-9], and methionine oxidation often results in a decrease or loss of biological function [9-13]. Both buried and surface methionines can be oxidized [14], and the rates of peroxide-mediated oxidation of the methionines within a single protein can vary greatly [15]. BM-3 mutant TH4 contains 13 methionines, and it would be tedious to introduce unnatural methionine analogs or other natural amino acids at all 13 positions in a site-directed fashion. We decided therefore to test whether global incorporation of a methionine analog might enhance stability to peroxide.

Incorporating unnatural amino acids into proteins is an established technique which has been useful for studying protein biosynthesis, probing protein structure and function, and adding unique functionality to proteins (see [16-18]). One simple method useful for incorporating some unnatural amino acids in place of their natural homologs is to exploit the promiscuity of the aminoacyl-tRNA synthetases. Methionine and norleucine (2-aminohexanoic acid) are isosteric amino acids, with norleucine the more hydrophobic [19]. Norleucine can substitute for methionine in the acylation of $tRNA^{Met}$ [20], resulting in proteins containing norleucine in place of methionine. The replacement of methionine with norleucine often has little influence on protein structure or function

[21-25], except in cases where methionines play key functional roles [26]. Gilles *et al*. incorporated norleucine into six methionine positions in adenylate kinase, resulting in 16-20% of the expressed kinase molecules resistant to CNBr cleavage and having improved stability in H_2O_2 [22]. Here we report the replacement of 13 methionines with norleucine in P450 BM-3 heme domain mutant TH4 and the effects of this global replacement on enzyme activity and stability.

B. Results

Table 6.1 lists the three culture conditions used during induced expression of TH4, following the shift of cultures (OD_{600} of ~0.8) to methionine-free M9AA medium. Expression resulted in cultures containing TH4 with only methionine incorporated (TH4(Met)), TH4 with both methionine and norleucine incorporated (TH4(Mix)), and TH4 with norleucine incorporated almost exclusively (TH4(Nor)). Figure 6.1 shows a PAGE gel of purified enzymes from cultures with different concentrations of norleucine. Table 6.1 lists approximate P450 expression levels for each culture. As expected, expression levels and cell growth rates were significantly reduced under methionine-restricted conditions. The expression level in the control experiment indicates that a small percentage of TH4 methionine positions will still contain methionine in TH4(Nor). Co-expression of the methionyl-tRNA synthetase has been shown to increase expression of proteins incorporating norleucine [27]. In our case, however, co-expression of the synthetase had little effect on expression of TH4(Nor) or cell growth (not shown). The reduced expression probably reflects norleucine's toxicity to cells [28] as well as the slow kinetics of norleucine incorporation [29].

Table 6.1. Methionine / norleucine concentrations used to supplement cultures following medium shift to methionine-free M9AA, and the resulting cytochrome P450 BM-3 mutant TH4 expression levels.

Cultures	TH4(Met)	TH4(Mix)	TH4(Nor)	Control
[L-Methionine] (mg/L)	100	10	-	-
[L-Norleucine] (mg/L)	-	300	300	-
P450 expression[a] (mg/L)	34	21	10	0.4

[a]P450 expression varies considerably between cultures. Here we report representative values, determined from the CO-binding difference spectrum of the clarified lysate, prior to purification.

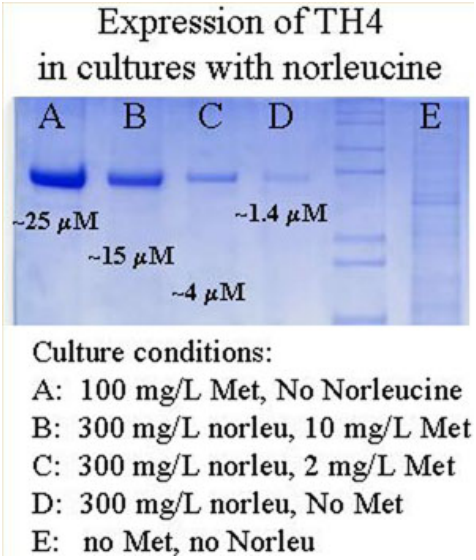


Figure 6.1. PAGE gel of purified protein recovered after expression of TH4 in cultures with different concentrations of methionine and norleucine. The corresponding protein concentrations are indicated.

Figure 6.2 shows mass spectra of fragments from tryptic digests of the three purified TH4 samples. Virtually complete substitution of methionine (MW 131.2) with norleucine (MW 113.2) in TH4(Nor) is verified by the shift of 18 mass units in three fragments that each contain a single methionine. Low-abundance, non-shifted peaks confirm that a small percentage of TH4 methionine positions still contain methionine in

TH4(Nor). Partial incorporation was achieved in the mixed culture TH4(Mix), as indicated by the presence of both sets of fragments. These three TH4 samples displayed typical P450 spectra (450 nm CO-binding peak and 417 nm low-spin absorbance peak), indicating that their overall folds and heme environments are similar to those of wild-type BM-3. A small shoulder occasionally appeared at 420 nm in the CO-binding spectrum of TH4(Nor). This is indicative of a P450 whose active site is disrupted [30] and suggests that stability has been reduced as a result of norleucine substitution.

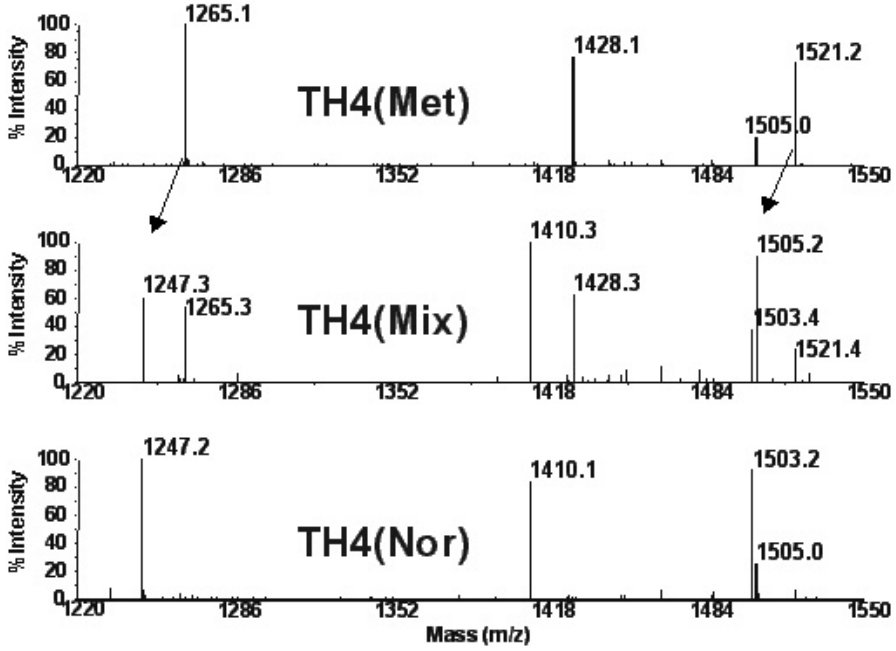


Figure 6.2. MALDI-TOF mass spectra of tryptic peptides derived from TH4 expression in medium supplemented with methionine only (TH4(Met)), methionine plus norleucine (TH4(Mix)) and norleucine only (TH4(Nor)). Peaks at m/z = 1265, 1428, and 1521 arise from peptides that each contain a single methionine. A shift of 18 mass units indicates replacement of methionine with norleucine.

Specific peroxygenase activities of the three purified TH4 samples were measured using 10 mM H_2O_2 and 12-*p*-nitrophenoxycarboxylic acid (12-pNCA) as substrate. TH4

is ~100% selective towards hydroxylating the C-12 carbon of 12-pNCA (adjacent to the pNP group), resulting in complete conversion to pNP under substrate-limiting conditions. Product formation was measured by monitoring pNP absorption at 398 nm. Figure 6.3 shows the time course of pNP formation for the three TH4 samples. The initial rate of product formation increases with the level of norleucine incorporation, and complete incorporation of norleucine results in almost twofold improved activity over TH4(Met). In repeated experiments with enzyme from duplicate cultures and varying reaction conditions, TH4(Nor) was consistently almost twice as active as TH4(Met). Under substrate-limiting conditions, the same amount of pNP was formed with TH4(Met) as with TH4(Nor), eliminating the possibility that increased activity is due to altered regioselectivity. The increased peroxygenase activity of TH4(Nor) was verified using styrene as substrate and quantifying styrene oxide by GC/FID (not shown). TH4(Nor) is more active than TH4(Met) in 2% DMSO as well as 6% DMSO (not shown).

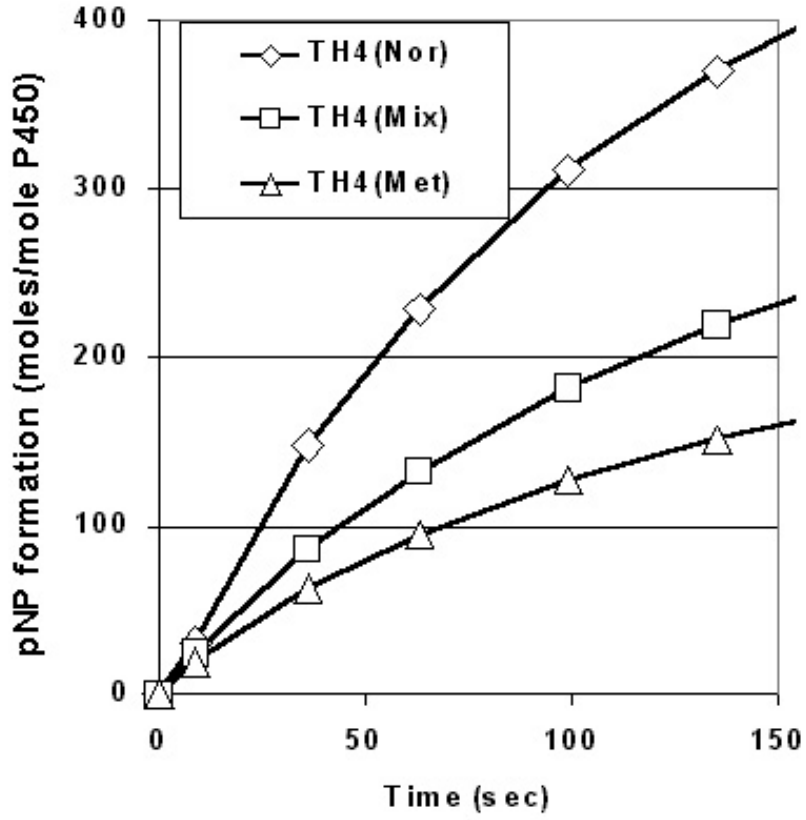


Figure 6.3. Time course of product formation during the peroxygenase reactions of TH4(Met), TH4(Mix), and TH4(Nor) with 12-pNCA in 10 mM H_2O_2. Absorbance at 398 nm was monitored to measure formation of product (pNP). Data points represent averages from three experiments; standard deviations were less than 10%.

To estimate the relative thermostabilities of the TH4 samples, each was heated to 57.5^{o}C for ten minutes, cooled, and re-assayed for activity at room temperature. Figure 6.4 shows the total turnovers achieved by each enzyme sample before and after heat treatment. Thermostability decreases with increased norleucine incorporation such that TH4(Nor) is almost completely inactivated by the heat treatment. Norleucine incorporation therefore decreases stability but increases peroxygenase activity. Since the majority of amino acid substitutions that can be introduced into an enzyme will have neutral to deleterious effects on both activity and stability, the observed improvement in

activity is unexpected. We are unable to offer an explanation for how norleucine increases peroxygenase activity.

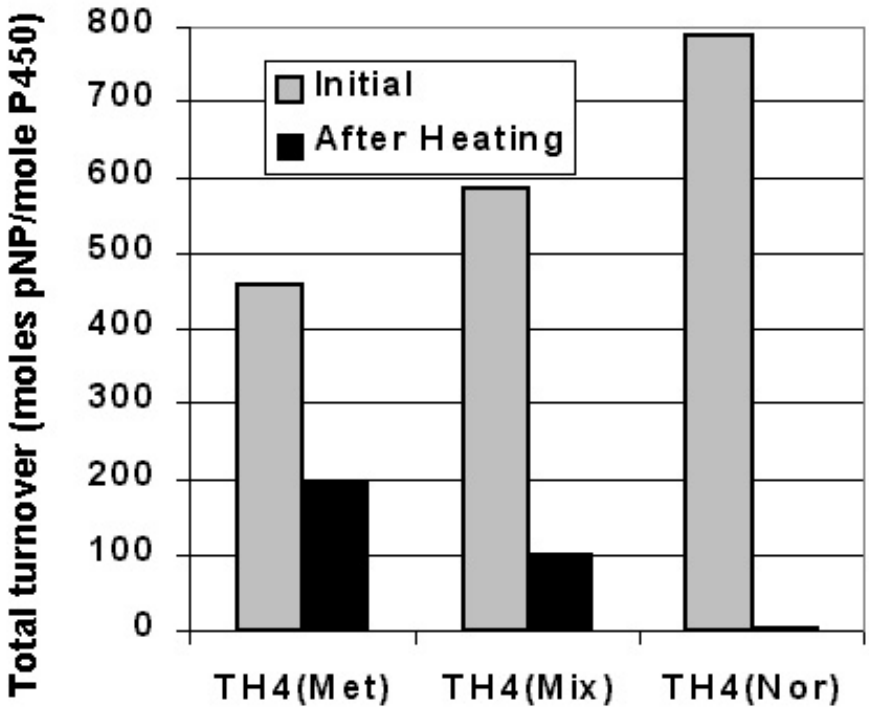


Figure 6.4. Total peroxide-driven turnovers achieved by TH4(Met), TH4(Mix), and TH4(Nor) before and after heat treatment. Reactions were performed at room temperature with 12-pNCA as substrate. 10 mM H_2O_2 was added to start each reaction. Heated samples were incubated at 57.5°C for ten minutes and then cooled before measuring activity.

While initial activity is increased with norleucine incorporation, the time courses of product formation are very similar for the three mutants, indicating that enzyme inactivation is not affected by norleucine incorporation and TH4 methionines are not oxidized under the peroxide concentrations used in the peroxygenase reaction (≤ 10 mM) during these reaction times. Mass spectral analysis of fragments from trypsin, endoproteinase Lys-C, and chymotrypsin digests of peroxide-treated TH4 have verified that 8 of the 13 methionine residues have not been oxidized (refer to discussion below

and Appendix D for details of the MALDI-TOF MS results). The remaining five residues were not represented by the detectable fragments so oxidation at these positions could not be checked.

H_2O_2-Mediated Inactivation of P450 BM-3

A. Effect of F87A on Stability to Peroxide

Figure 6.5 shows the fraction of initial activity of full-length wild-type P450 BM-3 (BWT) and the F87A mutant (BF87A) remaining after incubation in 10 mM H_2O_2. Activities were measured as the rate of NADPH consumption in the presence of myristic acid. Figure 6.6 shows the rate of heme bleaching for BWT and BF87A in the presence of 1 mM H_2O_2 and myristic acid, measured by monitoring the decrease in the 392 nm heme absorbance peak (heme Soret band). The wild-type enzyme is inactivated much more rapidly than the F87A mutant, and inactivation coincides with heme degradation. This helps to explain why peroxygenase activity is much higher with the F87A mutant. One explanation for how the F87A mutation prolongs the life of the heme in the presence of H_2O_2 could be that the mutation allows water molecules to remain in the active site after substrate binding. These water molecules may help to stabilize radical intermediate(s) formed at the heme which will otherwise destroy the heme (see below).

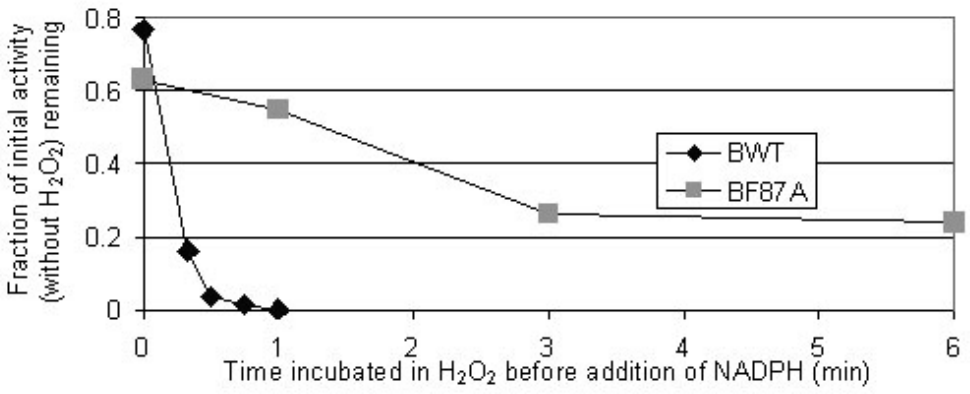


Figure 6.5. NADPH consumption rates for BWT and BF87A in the presence of myristic acid following incubations in 10 mM H_2O_2.

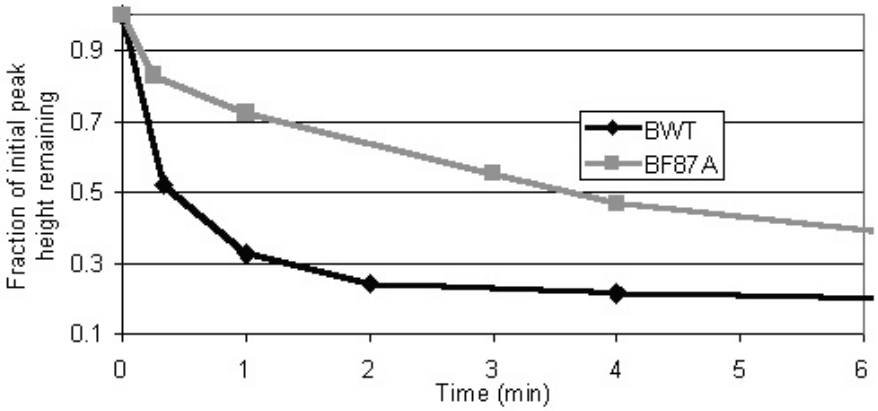


Figure 6.6. Heme bleaching: decay of the heme absorbance peak (392 nm) for BWT and BF87A in the presence of 1 mM H_2O_2 and myristic acid.

B. Effect of DMSO and Substrate on Stability to Peroxide

Recall that mutant libraries were screened for activity on 12-pNCA in the presence of 6.3% DMSO. As mentioned in Chapter 3, this concentration of DMSO inhibits peroxygenase activity of HF87A at low H_2O_2 concentrations (1-10 mM), while inhibition is almost insignificant using 50 mM H_2O_2, implying DMSO competitively inhibits HF87A (V_{max} is not changed). Figure 6.7 shows the effect of DMSO on the activity (initial rate and TON) of HF87A and TH4. Increasing DMSO concentrations

increasingly inhibit HF87A activity (initial rates), although similar TONs are reached in 1.6%, 5%, and 10% DMSO. Conversely, inhibition of the evolved mutant by DMSO is only significant at much higher DMSO concentrations (15%), presumably a result of screening for increased TONs and initial rates in 6.3% DMSO. Additionally, increasing DMSO concentrations actually prolong the life of TH4 and allows for higher TONs. Similar results are seen with mutant 21B3, as shown in Figure 6.8.

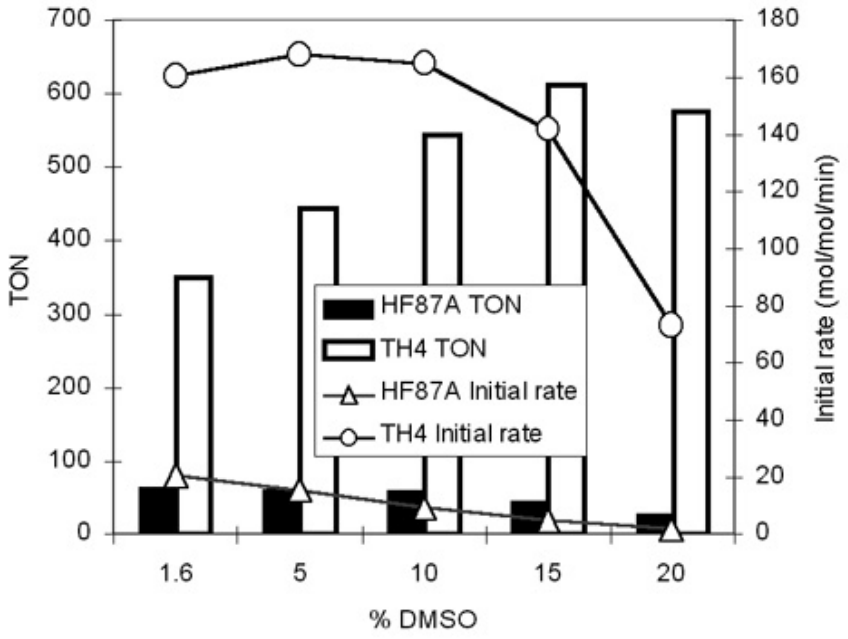


Figure 6.7. Effect of DMSO concentration on peroxygenase activity (initial rate and TON) of HF87A and TH4 using 12-pNCA as substrate.

Figure 6.9 ((a), (b), and (c)) shows the effects of DMSO (10% by volume) and substrate (1 mM myristate) on heme decay in the presence of 5 mM H_2O_2 for the wild-type BM-3 heme domain (HWT) and mutants HF87A and TH4. The change in representative spectra over time in 5 mM H_2O_2 and in the presence and absence of substrate are shown in Figure 6.10. In the absence of substrate the 418 nm absorbance peak was measured to monitor the heme, while the 392 nm peak was used when myristate was present.

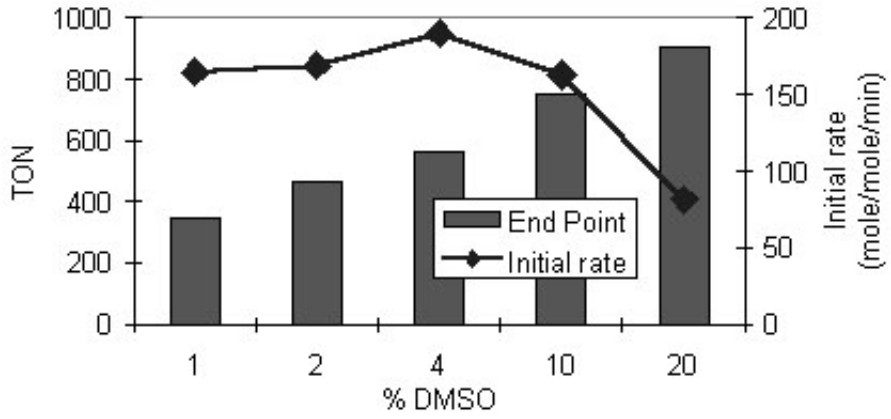
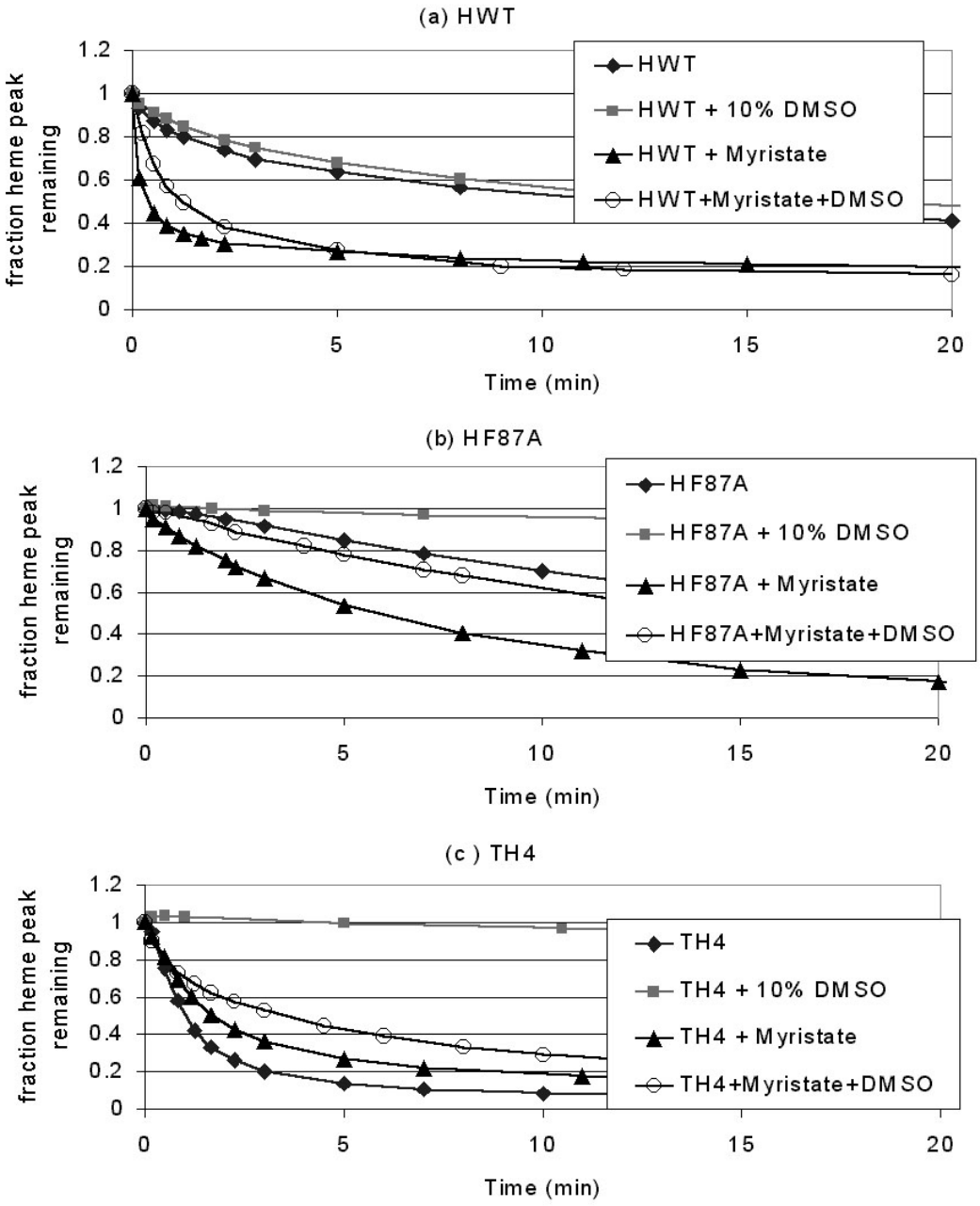


Figure 6.8. Effect of DMSO concentration on peroxygenase activity (initial rate and TON) of 21B3 using 12-pNCA as substrate.

Heme degradation is accelerated by the presence of substrate for HWT and HF87A. This is indicative of a turnover-dependent inactivation process. DMSO significantly inhibits heme degradation for HF87A and TH4, but not HWT. Perhaps DMSO cannot gain access to the active site in HWT, while the F87A mutation may create enough space to accommodate DMSO. That DMSO inhibits both HF87A activity (Figure 6.7) and heme degradation (Figure 6.9 (b)) further implies heme degradation is largely turnover-dependent. Also, the fact that similar TONs are reached by HF87A in the presence of 1.6%, 5% and 10% DMSO, even though the extent of inhibition varies significantly over this range of DMSO concentrations, indicates inactivation is largely mechanism-based. Similar results are seen when 1-butanol is added to reactions with HF87A (that is, the reaction is inhibited but similar TONs are reached before inactivation).



Figures 6.9. Effect of DMSO (10% by volume) and substrate (1 mM myristate) on heme bleaching in HWT, HF87A, and TH4 in the presence of 5 mM H_2O_2. Heme was monitored by either the 418 nm (without substrate) or 392 nm (with substrate) absorbance peaks.

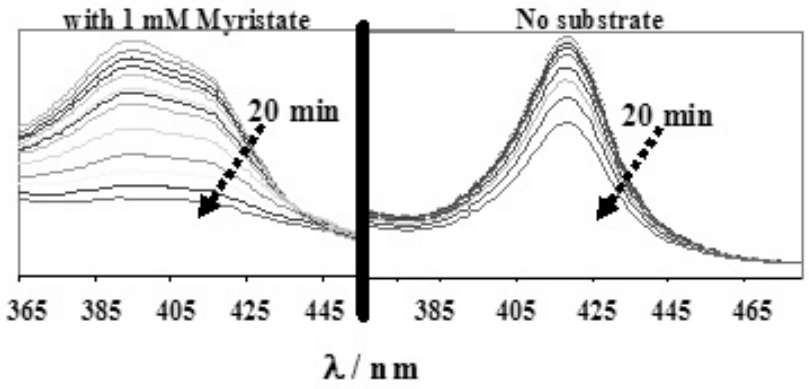


Figure 6.10. Change in HF87A heme absorbance spectra over 20 minutes in 5 mM H_2O_2 and the presence or absence of substrate (myristate).

Note that TH4 heme bleaching is much faster than that of HF87A when substrate is not present. In the absence of other substrates, H_2O_2 serves as a (poor) substrate, in the catalase reaction. The faster inactivation of TH4 in the absence of substrate and DMSO may partly be explained by the fact that TH4 has significantly higher catalase activity compared to HF87A (not shown). DMSO increases the TON of TH4 and therefore may be involved in protecting the heme. TH4 may have evolved to accommodate DMSO in the active site, where it can scavenge proposed radical heme intermediate species (see below). Thus, one possible method of improving TONs is by adding reagents which will protect the heme and prolong enzyme life by quenching otherwise suicidal heme intermediates. Other known radical scavengers (e.g., isopropanol, glutathione, phenol) were tested in reactions with TH4, but none resulted in TONs higher than those with DMSO. This also supports the conclusion that radicals, if formed, are localized at the heme and are not released from the active site.

C. Discussion

Mechanism-based ("suicide") inactivation has been studied in the H_2O_2-driven peroxidase reaction ([32] and references therein). In that case, *Compound* II (see Chapter 1) reacts with a second H_2O_2 molecule, resulting in the formation of a destructive radical species called *Compound* III. In the peroxygenase reaction, it is possible that a similar radical species is formed by the attack of a second H_2O_2 molecule by the P450 iron-oxo (*Compound* I-like) species through a catalase-type reaction. The resulting intermediate can either dissociate into H_2O and O_2 (the catalase reaction) or form a destructive radical species similar to *Compound* III. This proposed mechanism for heme destruction explains why inactivation is turnover-dependent: it requires the formation of the iron-oxo intermediate. Two H_2O_2 molecules are required to generate the suicide intermediate. Low concentrations of H_2O_2 relative to substrate should significantly reduce the rate of this "competing" reaction. Peroxidases show that low H_2O_2 concentrations (<100 μM) are significantly less destructive to the heme (competing suicide reactions occur at much slower rates). The high total turnovers achieved by CPO are only possible when low concentrations of H_2O_2 are maintained (< 50 μM) [33]. The reported half-lives of CPO in high concentrations of H_2O_2 (e.g., 1 min. in 30 mM H_2O_2 [32] and 30 min. in 1 mM H_2O_2 [34]) are very similar to the half-lives we find with P450 BM-3. Current efforts are therefore aimed at greatly lowering the peroxide requirement for the peroxygenase reaction using more stringent screening conditions.

Inactive heme intermediates have been spectroscopically observed in the presence of organic peroxides ([35] and references therein), but we were unable to detect such intermediates or other modified heme species by monitoring absorbance and difference

spectra during the P450 reaction/inactivation with H_2O_2. That free radicals (•OH) are not formed is supported by the fact that glutathione and other radical scavengers did not improve TONs, and mass spectra of heme domain peptide fragments generated after reactions with H_2O_2 do not identify any protein modification/oxidation (see below and Appendix D).

Heme modification and degradation due to the formation of suicide intermediates upon reaction with peroxides is well studied (refer to [32,35-37] and references therein). The observed heme bleaching is the result of a variety of reactions whereby the heme is degraded to smaller heme fragments [38] and possibly bound to apoprotein [39-41]. Correia and coworkers have shown that the L helix in the heme proximal region of cytochromes P450 2B1 [41] and 3A4 [40] is subject to attack by alkylating heme fragments after treatment with cumene hydroperoxide. Substrates have also been shown to form adducts with the heme and portions of the highly conserved distal I helix [39]. An analysis of MALDI-TOF mass spectra of peptide fragments from trypsin, endoproteinase Lys-C, and chymotrypsin digests of peroxide-treated BM-3 heme domain proteins (HWT, HF87A and TH4) is provided in Appendix D. While treatment with 10 mM H_2O_2 does not identify any peptide modifications, only ~66% of the total heme domain is represented by the identifiable peptide fragments. Unfortunately, fragments representing the majority of the L and I helices, which respectively lie proximal and distal to the heme, could not be detected from control (without peroxide) or peroxide-treated samples. Unidentifiable peaks occasionally appeared in the peroxide-treated samples and did not appear in control samples, although these peaks did not coincide with the disappearance of peaks in the control (without peroxide) experiments (for

example, fragment with m/z = 2583 from the endoproteinase Lys-C digest of H_2O_2-treated TH4, in Appendix D). It is therefore possible that the heme becomes alkylated to the I and/or L helices, or that protein from yet-unidentified fragments is modified/oxidized. The MALDI-TOF data do show that neither the distal region containing Phe 87 (peptide fragment consisting of residues 77 to 94, with m/z = 2113 for HWT and m/z = 2037 for HF87A, from endoproteinase Lys-C digests) nor the peptide fragment containing the heme proximal ligand Cys 400 (TH4 peptide fragment consisting of residues 394 to 405, with m/z = 1278, from a chymotrypsin digest) are modified during reaction with peroxide.

Conclusions

Peroxygenase activity of mutant TH4 is retained, and even improved, after global replacement of methionines with their oxidation-resistant, more hydrophobic analog norleucine. These results are particularly encouraging in light of recent successes in the engineering of expression systems to utilize unnatural protein building blocks [18,42]. Novel properties can be engineered into an enzyme via unnatural amino acid incorporation without sacrificing the enzyme's natural catalytic functions, provided the substitutions are not overly disruptive. Directed evolution experiments will soon be extended to evolving enzymes expressed in the presence of unnatural amino acids, where one probable goal will be to improve stability or function compromised as a result of analog incorporation. The system described here provides a framework for such evolution experiments.

Peroxide-mediated enzyme inactivation is the major limitation to our peroxygenase enzymes. That the H_2O_2-stability of the norleucine-containing TH4 mutant was no greater than that of the methionine-containing enzyme indicates that methionine oxidation does not play a role in enzyme inactivation, and further suggests that protein oxidation in general is not a problem. Thus, heme degradation is likely to be the only cause of peroxide-mediated inactivation, although we have not been able to verify that protein modifications do not occur. As a result of using DMSO in the screening conditions for evolution, maximum activity and TON for the evolved mutants are achieved in the presence of DMSO. DMSO apparently protects the heme during catalysis, suggesting a possible method of improving TON through engineering a reaction system in which destructive heme intermediates are quenched by a suitable reagent. The formation of suicide heme intermediates in peroxidases requires the reaction of an activated heme complex with a second H_2O_2 molecule, and a similar mechanism is likely to occur in P450s. Thus, as is the case with peroxidases, low H_2O_2 concentrations (50 μM) should greatly reduce the frequency of suicide reactions. Significantly improved TONs should therefore be achievable by reducing the K_m for H_2O_2 in the P450 peroxygenase reaction.

References

1. Guengerich FP: **Cytochrome P450 enzymes in the generation of commercial products**. *Nat Rev Drug Discov* 2002, **1**:359-366.
2. McGinnity DF, Riley RJ: **Predicting drug pharmacokinetics in humans from in vitro metabolism studies**. *Biochem Soc Trans* 2001, **29**:135-139.
3. Ortiz de Montellano PR (Ed): *Cytochrome P450: Structure, Mechanism, and Biochemistry,* 2nd edn. New York and London: Plenum Press; 1995.
4. Wong LL: **Cytochrome P450 monooxygenases**. *Curr Opin Chem Biol* 1998, **2**:263-268.
5. Miura Y, Fulco AJ: **Omega-1, omega-2 and omega-3 hydroxylation of long-chain fatty acids, amides and alcohols by a soluble enzyme system from *Bacillus megaterium***. *Biochim Biophys Acta* 1975, **388**:305-317.
6. Ruettinger RT, Wen LP, Fulco AJ: **Coding nucleotide, 5' regulatory, and deduced amino acid sequences of P-450BM-3, a single peptide cytochrome P-450:NADPH-P-450 reductase from *Bacillus megaterium***. *J Biol Chem* 1989, **264**:10987-10995.
7. Caldwell KA, Tappel AL: **Reactions of seleno- and sulfoamino acids with hydroperoxides**. *Biochemistry* 1964, **3**:1643-&.
8. Schechter Y, Burstein Y, Patchornik A: **Proceedings: Selective oxidation of methionine residues in proteins**. *Isr J Med Sci* 1975, **11**:1171.
9. Stauffer CE, Etson D: **Effect of subtilisin activity of oxidizing a methionine residue**. *J Biol Chem* 1969, **244**:5333-&.
10. Vogt W: **Oxidation of methionyl residues in proteins: tools, targets, and reversal**. *Free Radic Biol Med* 1995, **18**:93-105.
11. Brot N, Weissbach H: **Biochemistry of methionine sulfoxide residues in proteins**. *Biofactors* 1991, **3**:91-96.
12. Estell DA, Graycar TP, Wells JA: **Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation**. *J Biol Chem* 1985, **260**:6518-6521.
13. Kim YH, Berry AH, Spencer DS, Stites WE: **Comparing the effect on protein stability of methionine oxidation versus mutagenesis: steps toward engineering oxidative resistance in proteins**. *Protein Eng* 2001, **14**:343-347.
14. Keck RG: **The use of t-butyl hydroperoxide as a probe for methionine oxidation in proteins**. *Anal Biochem* 1996, **236**:56-62.
15. Lu HS, Fausset PR, Narhi LO, Horan T, Shinagawa K, Shimamoto G, Boone TC: **Chemical modification and site-directed mutagenesis of methionine residues in recombinant human granulocyte colony- stimulating factor: Effect on stability and biological activity**. *Arch Biochem Biophys* 1999, **362**:1-11.
16. Cornish VW, Mendel D, Schultz PG: **Probing protein-structure and function with an expanded genetic-code**. *Angew Chem-Int Edit* 1995, **34**:621-633.
17. Dougherty DA: **Unnatural amino acids as probes of protein structure and function**. *Curr Opin Chem Biol* 2000, **4**:645-652.
18. Wang L, Schultz PG: **Expanding the genetic code**. *Chem Commun (Camb)* 2002:1-11.

19. Schenck HL, Dado GP, Gellman SH: **Redox-triggered secondary structure changes in the aggregated states of a designed methionine-rich peptide**. *J Am Chem Soc* 1996, **118**:12487-12494.
20. Trupin J, Dickerman H, Nirenberg M, Weissbach H: **Formylation of amino acid analogues of methionine sRNA**. *Biochem Biophys Res Commun* 1966, **24**:50-55.
21. Anfinsen CB, Corley LG: **An active variant of staphylococcal nuclease containing norleucine in place of methionine**. *J Biol Chem* 1969, **244**:5149-&.
22. Gilles AM, Marliere P, Rose T, Sarfati R, Longin R, Meier A, Fermandjian S, Monnot M, Cohen GN, Barzu O: **Conservative replacement of methionine by norleucine in *Escherichia coli* adenylate kinase**. *J Biol Chem* 1988, **263**:8204-8209.
23. Naider F, Bohak Z, Yariv J: **Reversible alkylation of a methionyl residue near the active site of beta-galactosidase**. *Biochemistry* 1972, **11**:3202-3208.
24. Randhawa ZI, Witkowska HE, Cone J, Wilkins JA, Hughes P, Yamanishi K, Yasuda S, Masui Y, Arthur P, Kletke C, et al.: **Incorporation of norleucine at methionine positions in recombinant human macrophage colony stimulating factor (M-CSF, 4-153) expressed in *Escherichia coli*: structural analysis**. *Biochemistry* 1994, **33**:4352-4362.
25. Rocchi R, Scatturin A, Moroder L, Marchiori F, Tamburro AM, Scoffone E: **Synthesis of peptide analogs of the N-terminal eicosapeptide sequence of ribonuclease A. XI. Synthesis and conformational studies of [Orn-10, Nle-13]-S-peptide**. *J Am Chem Soc* 1969, **91**:492-496.
26. Yuan T, Vogel HJ: **Substitution of the methionine residues of calmodulin with the unnatural amino acid analogs ethionine and norleucine: biochemical and spectroscopic studies**. *Protein Sci* 1999, **8**:113-121.
27. Kiick KL, Tirrell DA: **Protein engineering by in vivo incorporation of non-natural amino acids: Control of incorporation of methionine analogues by methionyl-tRNA synthetase**. *Tetrahedron* 2000, **56**:9487-9493.
28. Bogosian G, Violand BN, Dorward-King EJ, Workman WE, Jung PE, Kane JF: **Biosynthesis and incorporation into protein of norleucine by *Escherichia coli***. *J Biol Chem* 1989, **264**:531-539.
29. van Hest JCM, Kiick KL, Tirrell DA: **Efficient incorporation of unsaturated methionine analogues into proteins in vivo**. *J Am Chem Soc* 2000, **122**:1282-1288.
30. Martinis SA, Blanke SR, Hager LP, Sligar SG, Hoa GHB, Rux JJ, Dawson JH: **Probing the heme iron coordination structure of pressure- induced cytochrome P420(cam)**. *Biochemistry* 1996, **35**:14530-14536.
31. Schwaneberg U, Schmidt-Dannert C, Schmitt J, Schmid RD: **A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A**. *Anal Biochem* 1999, **269**:359-366.
32. Valderrama B, Ayala M, Vazquez-Duhalt R: **Suicide inactivation of peroxidases and the challenge of engineering more robust enzymes**. *Chem Biol* 2002, **9**:555-565.
33. Seelbach K, vanDeurzen MPJ, vanRantwijk F, Sheldon RA, Kragl U: **Improvement of the total turnover number and space-time yield for chloroperoxidase catalyzed oxidation**. *Biotechnol Bioeng* 1997, **55**:283-288.

34. vanDeurzen MPJ, Seelbach K, VanRantwijk F, Kragl U, Sheldon RA: **Chloroperoxidase: Use of a hydrogen peroxide-stat for controlling reactions and improving enzyme performance**. *Biocatal Biotransform* 1997, **15**:1-16.
35. Ortiz de Montellano PR: **Oxygen activation and reactivity**. In *Cytochrome P450: Structure, Mechanism, and Biochemistry,* 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:245-303.
36. Ortiz de Montellano PR, Correia MA: **Inhibition of cytochrome P450 enzymes**. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 2nd edn. Edited by Ortiz de Montellano PR: Plenum Press; 1995:305-364.
37. Karuzina, II, Archakov AI: **The oxidative inactivation of cytochrome P450 in monooxygenase reactions**. *Free Radic Biol Med* 1994, **16**:73-97.
38. Schaefer WH, Harris TM, Guengerich FP: **Characterization of the enzymatic and nonenzymatic peroxidative degradation of iron porphyrins and cytochrome P-450 heme**. *Biochemistry* 1985, **24**:3254-3263.
39. He K, Falick AM, Chen B, Nilsson F, Correia MA: **Identification of the heme adduct and an active site peptide modified during mechanism-based inactivation of rat liver cytochrome P450 2B1 by secobarbital**. *Chem Res Toxicol* 1996, **9**:614-622.
40. He K, Bornheim LM, Falick AM, Maltby D, Yin H, Correia MA: **Identification of the heme-modified peptides from cumene hydroperoxide-inactivated cytochrome P450 3A4**. *Biochemistry* 1998, **37**:17448-17457.
41. Yao K, Falick AM, Patel N, Correia MA: **Cumene hydroperoxide-mediated inactivation of cytochrome P450 2B1. Identification of an active site heme-modified peptide**. *J Biol Chem* 1993, **268**:59-65.
42. Datta D, Wang P, Carrico IS, Mayo SL, Tirrell DA: **A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins**. *J Am Chem Soc* 2002, **124**:5652-5653.

CHAPTER 7

Materials and Experimental Procedures

General Remarks

All chemical reagents were procured from Aldrich, Sigma, or Fluka. H_2O_2 was purchased as a 35 wt% solution (Aldrich). NADPH (tetrasodium salt) was purchased from BioCatalytics, Inc. (Pasadena, CA). Restriction enzymes were purchased from New England Biolabs and Roche. Deep-well plates (96 wells, 1 ml volume per well) for growing mutant libraries were purchased from Becton Dickinson. Flat bottom 96-well microplates (300 μl per well) for screening mutant library activities were purchased from Rainin. Trypsin and chymotrypsin were purchased from Promega. Endoproteinase LysC was from Sigma. Fatty acids were added to reaction mixtures as 25-50 mM stock solutions of the sodium salts, which were prepared fresh each day in 50 mM Na_2CO_3 and stored at 80°C. H_2O_2 and NADPH stock solutions were prepared fresh each day in 100 mM Tris-HCl, pH 8.2. Buffers and media were prepared according to standard protocols [1].

The 12-pNCA substrate was synthesized as described [2], except hydrolysis of the ester was not performed enzymatically. Instead, the ester was dissolved into THF, mixed with an equal volume of an aqueous solution of 1 M KOH, and allowed to reflux with stirring for 6 hours. Following hydrolysis, the aqueous layer was separated, and 12-pNCA was crystallized by addition of H_2SO_4 and then washed. The THF layer was saved for a second round of hydrolysis. The final product was verified by ^1H NMR, recorded on a Bruker DPX 300 spectrometer.

Purified enzyme samples were stored at –80°C until use, at which time they were thawed at room temperature and then kept on ice. Concentrations of properly folded

P450 enzyme were determining from the 450 nm CO-binding difference spectra of the reduced heme, as described [3].

Enzyme Expression

All enzymes were expressed in catalase-deficient *E. coli* [4] using the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible pCWori(+) vector [5], which is under the control of the double Ptac promoter. The P450 BM-3 heme domain (and mutants thereof) consisted of the first 463 amino acids of P450 BM-3 followed by a 6-His sequence at the C-terminus, which had no significant influence on activity. Expression was accomplished by growth in terrific broth (TB) supplemented with 0.5 mM thiamine and trace elements [6]. Cultures were shaken at 150 rpm and 35ºC, induced after 9 hours by the addition of 1 mM δ-aminolevulinic acid and 0.5 mM IPTG, and grown for an additional 22 hours at 100 rpm and 35ºC.

Protein expression for norleucine incorporation experiments (Chapter 6) was performed using the medium shift method as described [7], using the *E. coli* methionine auxotroph CAG18491 (*λ⁻, rph-1, metEo-30769::Tn10*). Briefly, cells were transformed with the pCWori(+)(TH-4) plasmid and grown on LB-agar plates containing 100 mg/l ampicillin. A single, freshly transformed colony was used to inoculate 5 ml M9AA medium supplemented with 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4 wt% glucose, 1 mM thiamine, and 100 mg/l ampicillin. This 5 ml culture was grown overnight and used to inoculate a 500 ml culture containing the same ingredients. A medium shift was performed when this culture reached an OD_{600} of ~0.8. Cells were sedimented and resuspended in a 0.9% NaCl solution for washing, and this was repeated three times.

Cells were then resuspended in 500 ml of the M9AA medium described above, without methionine. The culture was supplemented with 1 mM δ-aminolevulinic acid (a heme precursor) plus trace elements [6] and grown for 10 min. at 37°C to deplete any remaining methionine. IPTG (0.5 mM) was then added to induce protein expression, and the culture was divided into four 125 ml cultures. One was grown without methionine or norleucine, as a negative control. The positive control (TH-4(Met)) was grown with 100 mg/l L-methionine added. TH-4(Mix) was grown with 300 mg/l L-norleucine plus 10 mg/l L-methionine, and TH-4(Nor) was grown with 300 mg/l L-norleucine only. These cultures were grown at 30°C for 15-20 hours.

Protein Purification

Purification of full-length enzymes (BWT and BF87A) was performed essentially as described [8] using an Äkta explorer system (Pharmacia Biotech) and SuperQ-650M column packing (Toyopearl). Purification of the heme domain enzymes took advantage of the 6-His tag using the QIAexpressionist kit (Qiagen) for purification under native conditions. Briefly, cultures were grown as described above, and after centrifugation and resuspension in lysis buffer (10 mM imidazole, 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl) the cells were lysed by sonication. Cell lysates were centrifuged, filtered, and loaded onto a Qiagen Ni-NTA column. The column was washed with wash buffer (20 mM imidazole, 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl), and the bound P450 was eluted with elution buffer (200 mM imidazole, 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl). Aliquots of the purified proteins were placed into liquid nitrogen and stored at -80 ºC. When used, the frozen aliquots were rapidly thawed and buffer-exchanged with 100 mM

Tris-HCl, pH 8.2 using a PD-10 Desalting column (Amersham Pharmacia Biotech). P450 enzyme concentrations were quantified by CO-binding difference spectra of the reduced heme as described [3], using an extinction coefficient of 91 $mM^{-1}cm^{-1}$ for the 450 nm minus 490 nm peak. P450 total expression yields were estimated from CO-binding difference spectra of the clarified lysates.

Determination of Rates and Product Distributions with Fatty Acids

Reactions contained purified P450 (1 μM for the NADPH-driven reactions, 4 μM for the H_2O_2-driven reactions) and 1-2 mM substrate in 500 μl 100 mM Tris-HCl, pH 8.2. Initial rate reactions were run for one minute. Total turnover number reactions were run for 30 min. Reactions were run at room temperature and initiated by the addition of 5 mM H_2O_2 (total turnover determinations) 10 mM H_2O_2 (initial rate measurements), or 2 mM NADPH and stopped by the addition of 7.5 μl 6 M HCl. At the end of each reaction an internal standard was added prior to extraction. For reactions with myristic and lauric acid, 30 nmoles of 10-hydroxycapric acid was used as the internal standard. For reactions with capric acid, 30 nmoles of 12-hydroxylauric acid was added as the internal standard.

Reactions were extracted twice with 1 ml ethyl acetate. The ethyl acetate layer was dried with sodium sulfate and then evaporated to dryness in a vacuum centrifuge. The resulting product residue was dissolved in 100 μl of a 1:1 pyridine:BSTFA (bis-(trimethylsilyl-trifluoroacetamide)) mixture containing 1% trimethylchlorosilane (TMCS). This mixture was heated at 80ºC for 30 minutes to allow for complete

derivatization of the acid and alcohol groups to their respective trimethylsilyl (TMS) esters and ethers.

Reaction products were identified by GC/MS using a Hewlett Packard 5890 Series II gas chromatograph coupled with a Hewlett Packard 5989A mass spectrometer. MS fragmentation patterns clearly identified the hydroxylated isomers. Quantification of the reaction products was accomplished using a Hewlett Packard 5890 Series II Plus gas chromatograph equipped with a flame ionization detector (FID). The GCs were fitted with an HP-5 column. The peak areas from each GC trace were normalized by dividing by the peak area of the internal standard. Authentic standards for each hydroxylated isomer of the fatty acids were not available, so standard curves were generated using the available ω-hydroxylated standards (12-hydroxylauric acid and 10-hydroxycapric acid). Authentic standard samples were prepared in the same fashion as the reaction samples, except the standards were added to the reaction mixture and the enzyme was inactivated by the addition of HCl before adding NADPH or H_2O_2. It was assumed that the FID response is the same for all regioisomers of a given hydroxylated fatty acid. The standard curves were linear in the range of product detection, with R-squared values of 0.99 for the linear regression fits. A standard curve for hydroxylated myristic acid was prepared using β-hydroxymyristic acid, but there was significant loss of this standard during sample work-up compared to the reaction products and the resulting calibration curve was not used for quantifying hydroxylated products.

Determination of Peroxygenase Activity on Styrene

Reactions using styrene as substrate were prepared in a similar fashion: Styrene reaction mixtures contained 5 mM H_2O_2 and 2% DMSO (total turnover reactions) or 10 mM H_2O_2 and 6% DMSO (initial rate reactions). Reactions were stopped by the addition of 1 ml pentane followed by vigorous shaking. 200 nmoles of 3-chlorostyrene oxide was added as the internal standard. The products were extracted twice with 1 ml pentane. The pentane layer was evaporated down to ~200 μl and injected directly onto the GC without derivatization. Peroxygenase reactions with fatty acids as well as styrene achieved maximum rates near pH 8.0, and Tris buffer supported higher rates than phosphate or MOPS.

Determination of NADPH-driven Reaction Rates (Chapters 2 and 4)

Rates of NADPH-driven reactions were determined from initial rates of NADPH consumption. Reactions contained purified P450 (1 μM for reactions with myristic and lauric acid and 3 μM for reactions with capric acid) and 0.5-1 mM substrate in 100 mM Tris-HCl, pH 8.2 in a quartz cuvette. Reactions were initiated by addition of 0.2 mM NADPH and the initial rate of NADPH consumption was determined by measuring the decrease in absorbance at 340 nm, using an extinction coefficient of 6.2 $mM^{-1}cm^{-1}$ [9]. The background level of NADPH consumption (i.e., with no substrate present) was determined for wild-type (~10 min^{-1}) and the F87A mutant (~9 min^{-1}) and subtracted from the NADPH consumption rates.

Determination of Coupling Efficiency (Chapter 2)

Coupling efficiencies for NADPH-driven reactions with lauric acid and capric acid were determined by measuring the moles of product formed per 100 moles of NADPH consumed. Reactions (500 μl) containing 1-2 μM P450 and 2 mM substrate in 100 mM Tris-HCl (pH 8.2) were initiated by addition of 100 μM (50 nmoles) NADPH and allowed to react to NADPH depletion. Products were then quantified from these reactions as described above.

Mutant Library Synthesis

Error-prone PCR libraries were generated using standard protocols [10]. Specifically, 100 μl reactions contained 7 mM Mg^{2+}, 0.2 mM dNTPs plus excess concentrations of dCTP and either dTTP or dATP (0.8 mM each), 20 fmole template DNA (as plasmid), 30 pmole of each outside primer, 1X Taq buffer (Roche) and 1 μl (1 Unit) Taq DNA polymerase (Roche). Due to the high Mg^{2+} concentration and excess of two dNTPs, no Mn^{2+} was necessary to generate mutant libraries with a suitable fitness landscape (30% to 40% of variants had <5% of parent activity). The temperature cycle used was: 94°C for 1 min followed by 29 cycles of 94°C for 1 min then 55°C for 1 min then 72°C for 1:40.

StEP recombination was performed similar to the protocol described [11] using HotStar Taq DNA polymerase (Qiagen). A 50 μl StEP reaction contained ~0.12 pmole total template DNA (comprised of approximately equal concentrations of seven mutant genes), 0.2 mM dNTPs, 5 pmole outside primers, 1X HotStar buffer (containing 15 mM Mg^{2+}), and 2.5 U HotStar Taq DNA polymerase. The temperature protocol was as

follows: (hot start) 95°C for 3 min, followed by 100 cycles of 94°C for 30 sec and 58°C for 8 sec.

Two libraries during evolution of thermostability were prepared with the GeneMorph™ PCR Mutagenesis Kit (Stratagene). In the final generation leading to mutant 5H6, a recombinant library was prepared by DNA shuffling [12] using Pfu Ultra DNA Polymerase (Stratagene). For all PCR manipulations on the entire P450 BM-3 heme domain gene the forward primer sequence was:
5'-ACAGGATCCATCGATGCTTAGGAGGTCATATG-3'
and the reverse primer sequence was: 5'-GCTCATGTTTGACAGCTTATCATCG-3'.

The heme domain gene was cloned into the pCWori vector using the unique restriction sites BamHI upstream of the gene and EcoRI downstream. All site-directed mutants, saturation mutagenesis libraries, and random oligonucleotide mutagenesis libraries were generated by standard SOEing PCR procedures [13,14] using oligonucleotides ordered from the Caltech Biopolymer Synthesis Facility and either HotStar Taq DNA polymerase (Qiagen), Pfu DNA polymerase (Stratagene), or Taq DNA polymerase (Roche).

Screening Procedure

Colonies from transformations of mutant libraries were picked into 96-well deep-well plates containing LB medium (300 μl) and ampicillin (100 μg/ml). Plates were incubated at 30 ºC, 270 rpm, and 80% relative humidity in a Kühner ISF-1-W shaker. After 24 hours, 20 μl of culture liquid from each well was used to inoculate 300 μl of TB medium containing ampicillin (100 μg/ml), thiamine (0.5 mM), and trace elements [6] contained

in 96-deep-well plates. Handling of 96 cultures at once was accomplished using a Beckman Multimek 96-channel pipetting robot. The LB pre-cultures were stored at 4°C as master plates. The TB cultures were grown at 30ºC, 270 rpm with shaking for approximately three hours before being induced for protein expression by the addition of δ-aminolevulinic acid (1 mM) and IPTG (0.5 mM). Cultures were then grown for an additional 18 hours.

After cell growth the plates were centrifuged and cell pellets were frozen at -20ºC before lysing. Lysis was accomplished by resuspending the cell pellets in 300-700 µl Tris-HCl buffer (100 mM, pH 8.2) containing lysozyme (0.5-1 mg/ml) and deoxyribonuclease I (1.5-4 Units/ml). The pellets were resuspended and lysed by mixing for approximately 15 minutes before centrifugation. An appropriate volume (10-50 µl) of the clarified cell lysates containing soluble heme domain mutants were used in the activity and thermostability assay, as described below.

The most active and/or thermostable mutants in a generation were streaked onto agar plates to obtain single colonies, which were then picked for rescreening. Rescreening was performed as described above, except 10 ml TB cultures were grown instead of deep-well plate cultures. Cell pellets were resuspended in 1 ml Tris-HCl (100 mM, pH 8.2) and lysed by sonication. P450 concentrations were quantified from the clarified lysates by CO-binding difference spectra [3]. Specific activities, total enzyme turnover values, and/or thermostabilities were determined to verify that the selected mutants were improved with respect to the parent enzyme.

Approximately 1000 to 3000 mutants were screened per generation. The screening procedure was developed and optimized to ensure that the coefficient of variation in activity measurements were less than ~15%.

Peroxygenase Activity Screening Assay

P450 activity measurements using 12-pNCA were performed by monitoring the formation of *p*-nitrophenolate (pNP) (398 nm) at room temperature using a 96-well plate spectrophotometer (SPECTRAmax, Molecular Devices). Activity (hydroxylation of 12-pNCA) requires folded enzyme displaying the characteristic 450 nm CO-binding absorbance peak. Enzyme samples containing either partially or fully denatured protein (as a result of guanidinium chloride treatment or heat treatment) are either less active than the native enzyme (for partially denatured samples) or not active at all (for fully denatured samples). Control reactions containing H_2O_2 and 12-pNCA with or without hemin chloride show no background activity.

A typical reaction well contained 140 µl 100 mM Tris-HCl buffer pH 8.2, 10 µl stock solution of substrate (4 mM 12-pNCA) in DMSO, and purified enzyme (or clarified lysates when screening). Reactions were initiated by the addition of 10 µl H_2O_2 stock solution. Final concentrations were 250 µM 12-pNCA, 6.3% DMSO, 1-50 mM H_2O_2, and 0.1-1.0 µM enzyme. DMSO (6% final concentration) was used to solubilize the 12-pNCA. The 398 nm absorbance reading for each well was blanked before addition of H_2O_2 so that end point turnovers could be calculated. Rates of peroxygenase activity were calculated as the rate of pNP formation (or the increase in absorbance at 398 nm over time). The value for (extinction coefficient)x(path length) for pNP under the exact

conditions used in the spectrophotometer assay was calculated from a standard curve generated with known concentrations of pNP. This factor was used to quantify turnover of substrate.

Data for accurate determination of 12-pNCA turnover rates with purified enzyme were collected using a BioSpec-1601 spectrophotometer (Shimadzu), where absorbance changes could be registered every 0.1 seconds.. These reactions contained 250 μM 12-pNCA, 1-6% DMSO, 1-10 mM H_2O_2 and 0.1-1.0 μM P450 in 100 mM Tris-HCl, pH 8.2. Reactions were carried out in cells with a 1 cm path length, and pNP formation was measured by monitoring absorbance at 398 nm. The extinction coefficient of pNP under the exact conditions used in the spectrophotometer assay was calculated from a standard curve generated with known concentrations of pNP. Initial rates were calculated in the first few seconds of reaction, where product formation was linear with time.

Library Screening for Thermostability (Chapter 3)

Screening was performed as described above, except cell lysates were additionally subjected to a heat inactivation step and screened for residual activity (see also [15]). Clarified lysates were transferred to 96-well PCR plates (GeneMate) and heated to an appropriate temperature (48ºC - 57.5ºC) in a PTC200 thermocycler (MJ Research) for 10-15 minutes, rapidly cooled to 4ºC, and then brought to room temperature. The residual activities of these heat-treated lysates were then measured in the same manner as the initial activities. Clones showing a higher fraction of activity remaining after heat treatment and high initial activity were characterized further.

T_{50} Determination (Chapter 4)

Purified enzyme samples (~20 μM) in Tris-HCl buffer (100 mM, pH 8.2) were incubated for 10 minutes at different temperatures. Samples were then cooled on ice, and the concentration of properly folded heme domain (diluted 8x) was estimated from the 450 nm CO-binding difference spectra and compared to the CO-binding peak prior to heat treatment. Residual NADPH-consumption activity was measured for BWT.

$t_{1/2}$ Determination (Chapter 4)

Concentrated purified enzyme (70 μM) was added to pre-heated (57.5°C) Tris-HCl buffer (100 mM, pH 8.2) and incubated at 57.5°C. Samples were removed at time intervals, quenched by dilution into cold buffer, brought to room temperature, and assayed for residual activity.

Preparation of Samples for Matrix Assisted Laser Desorption Ionization – Time-Of-Flight Mass Spectra (MALDI-TOF MS) Analyses (Chapter 6)

To determine whether heme domain proteins (HWT, HF87A, and TH4) were modified during reactions in H_2O_2, enzymes were first incubated in H_2O_2 with myristic or lauric acid. H_2O_2 was not added to control samples so comparisons could be made with unmodified protein. Reactions (100 μl) contained 4 μM enzyme, 10 mM H_2O_2, and 1 mM substrate. Following 20-minute incubations in H_2O_2, catalase (100 ng) was added to each reaction to remove all H_2O_2 (catalase was also added to the control reactions). This amount of catalase was low enough to not interfere with the mass spectra of the P450

fragments. Urea was then added to each sample to a final concentration of 2 M, and the samples were heated to 60^{o}C for 20 min and then cooled.

Tryptic and chymotryptic digests (50 μl) contained 1-5 μM purified P450 and 8 ng/μl trypsin or chymotrypsin in 50 mM NH_4HCO_3. Following overnight digestion at room temperature, digests were quenched by addition of trifluoroacetic acid (TFA) (final pH ~2) and cleaned using the $ZipTip_{C18}$ procedure (Millipore). Peptide fragments were eluted in 2 μL (50-80% acetonitrile/1% TFA/water; elutions were performed in gradients with increasing acetonitrile concentration). The peptide samples were mixed (1:3 ratio) with an α-cyanohydroxycinnaminic acid solution (10 mg/ml in 50% acetonitrile / 0.05% TFA / water) to form the MALDI matrix. Mass spectra were recorded on an Applied Biosystems Voyager DE-PRO MALDI-TOF mass spectrometer operated in reflected ion mode.

Digests with endoproteinase Lys-C were treated similar to those with trypsin, except 0.2 μg Lys-C was first added and samples were incubated for 2 hours at 37^{o}C, and then 0.2 μg more of Lys-C was added and the samples were digested overnight at room temperature.

References

1. Sambrook J, Russell DW: *Molecular Cloning: A Laboratory Manual, 3rd ed.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
2. Schwaneberg U, Schmidt-Dannert C, Schmitt J, Schmid RD: **A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A**. *Anal Biochem* 1999, **269**:359-366.
3. Omura T, Sato RJ: **The carbon monoxide-binding pigment of liver microsomes**. *J. Biol. Chem.* 1964, **239**:2370-2378.
4. Nakagawa S, Ishino S, Teshiba S: **Construction of catalase deficient *Escherichia coli* strains for the production of uricase**. *Biosci Biotechnol Biochem* 1996, **60**:415-420.
5. Barnes HJ, Arlotto MP, Waterman MR: **Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli***. *Proc Natl Acad Sci USA* 1991, **88**:5597-5601.
6. Joo H, Arisawa A, Lin Z, Arnold FH: **A high-throughput digital imaging screen for the discovery and directed evolution of oxygenases**. *Chem Biol* 1999, **6**:699-706.
7. van Hest JCM, Kiick KL, Tirrell DA: **Efficient incorporation of unsaturated methionine analogues into proteins in vivo**. *J Am Chem Soc* 2000, **122**:1282-1288.
8. Schwaneberg U, Sprauer A, Schmidt-Dannert C, Schmid RD: **P450 monooxygenase in biotechnology. I. Single-step, large-scale purification method for cytochrome P450 BM-3 by anion-exchange chromatography**. *J Chromatogr A* 1999, **848**:149-159.
9. Yeom H, Sligar SG: **Oxygen activation by cytochrome P450BM-3: effects of mutating an active site acidic residue**. *Arch Biochem Biophys* 1997, **337**:209-216.
10. Cirino PC, Mayer KM, Umeno D: **Generating mutant libraries using error-prone PCR**. In *Directed Evolution Library Creation: Methods and Protocols*. Edited by Arnold FH, Georgiou G: Humana Press; 2003:3-9.
11. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH: **Molecular evolution by staggered extension process (StEP) in vitro recombination**. *Nat Biotechnol* 1998, **16**:258-261.
12. Stemmer WP: **Rapid evolution of a protein in vitro by DNA shuffling**. *Nature* 1994, **370**:389-391.
13. Horton RM, Ho SN, Pullen JK, Hunt HD, Cai Z, Pease LR: **Gene splicing by overlap extension**. *Methods Enzymol* 1993, **217**:270-279.
14. McPherson MJ, Moller SG: **PCR mutagenesis**. In *PCR*. Edited by McPherson MJ, Moller SG: BIOS Scientific Publishers Limited; 2000:143-182. Beynon R, Howe C (Series Editor): The Basics
15. Cirino PC, Georgescu R: **Screening for thermostability**. In *Directed Enzyme Evolution: Screening and Selection Methods*. Edited by Arnold FH, Georgiou G: Humana Press; 2003.

Appendix A

Directed Evolution of Horseradish Peroxidase

Introduction

H_2O_2 damages the heme and is known to oxidize protein residues. The half-life of horseradish peroxidase (HRP) in 1 mM H_2O_2 is 33 minutes [1]. Figure A.1 shows the catalytic cycle of HRP and mechanisms for the H_2O_2-mediated formation of destructive radical species as well as the inactive heme species E_{670}. In the study described here, directed evolution was used to improve the stability of HRP to H_2O_2 and high temperature. Wild-type HRP contains four disulfide bonds and is highly glycosylated, and the inability to functionally express this enzyme in *E. coli* has been a major roadblock in the development of improved enzyme variants [2]. Expression of HRP in *E. coli* results in high levels of the polypeptide chain in inclusion bodies, which is probably a consequence of *E. coli*'s lack of glycosylation function and poor ability to support disulfide formation. We therefore sought to achieve functional expression of HRP (and mutants thereof) in *S. cerevisiae* (yeast), which facilitates glycosylation and disulfide formation [3]. One round of random mutagenesis followed by screening for improved total activity resulted in an HRP "expression variant" containing mutation L37I. Starting with this L37I mutant as the parent, I used the same yeast expression system to evolve resistance to H_2O_2 inactivation, increased thermostability and total activity.

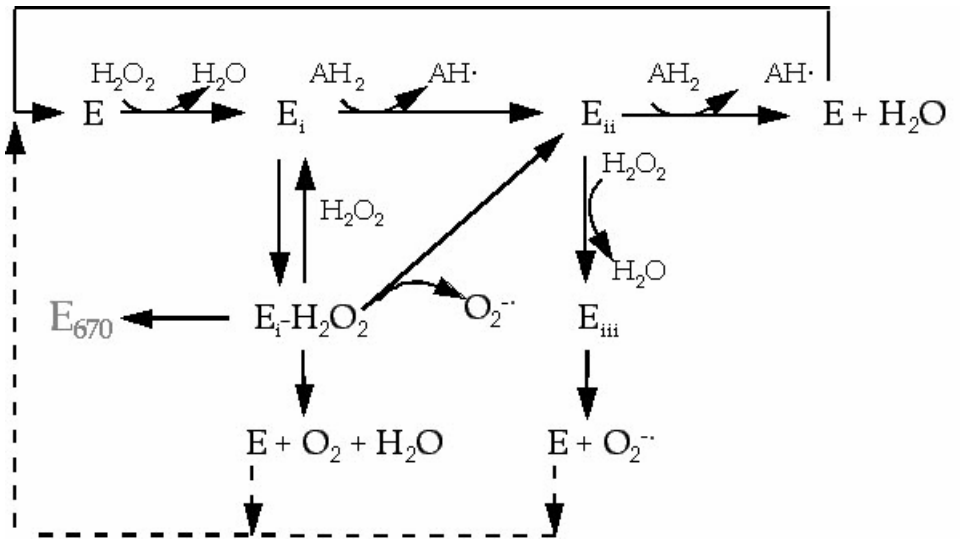


Figure A.1: Catalytic cycle and inactivation pathways of HRP in the presence of aromatic substrate (AH_2) and H_2O_2, showing the H_2O_2-mediated formation of inactive enzyme E_{670}.

Results

Wild-type enzyme studies were first carried out to determine appropriate screening conditions. HRP activity was measured using the conventional spectrophotometric activity assay with ABTS (0.5 mM). It was determined that an incubation temperature of 73^{o}C for ten minutes was suitable for screening thermostability (wild-type retains about 25% activity). Calcium chloride (2 mM) is required for optimal resistance to thermal deactivation. To determine H_2O_2-resistance assay conditions, a range of H_2O_2 concentrations (up to 50 mM) were tested in the presence of 0.5 mM ABTS. Additionally, HRP was pre-incubated in various concentrations of H_2O_2 for various time intervals prior to measuring activity. The optimum H_2O_2 concentration for HRP activity was found to be between 0.5 mM and 5 mM H_2O_2. A 30 minute incubation in 25 mM H_2O_2 prior to measuring activity in 25 mM H_2O_2 was chosen for screening H_2O_2-resistant HRP mutants.

After screening one generation of random mutants in *S. cerevisiae*, the fraction of initial activity (measured at 25^{o}C) to activity (also at 25^{o}C) following a 10 min incubation at 73^{o}C was improved nearly fourfold (from 23% to 80%), as shown in Figure A.2. Figure A.2 shows that $T_{1/2}$, the transition midpoint of the inactivation curve, for the most thermostable mutant is more than 6^{o}C higher than that of the parent. As shown in Figure A.3, the residual activity following incubation in 25 mM H_2O_2 was improved from 42% to 60% for another mutant. Finally, the total initial activity of one mutant lysate was improved more than fivefold, as shown in Figure A.4. Sequencing revealed single point mutations in the improved mutants, and these mutations are given for each mutant in Figures A.2 through A.4. Two further rounds of directed evolution were carried out, resulting in a 40-fold increase in total HRP activity in the *S. cerevisiae* culture supernatant compared with wild-type (260 units/l/OD_{600} as measured on ABTS), as described in Chapter 1 and in Morawski *et al*. [2].

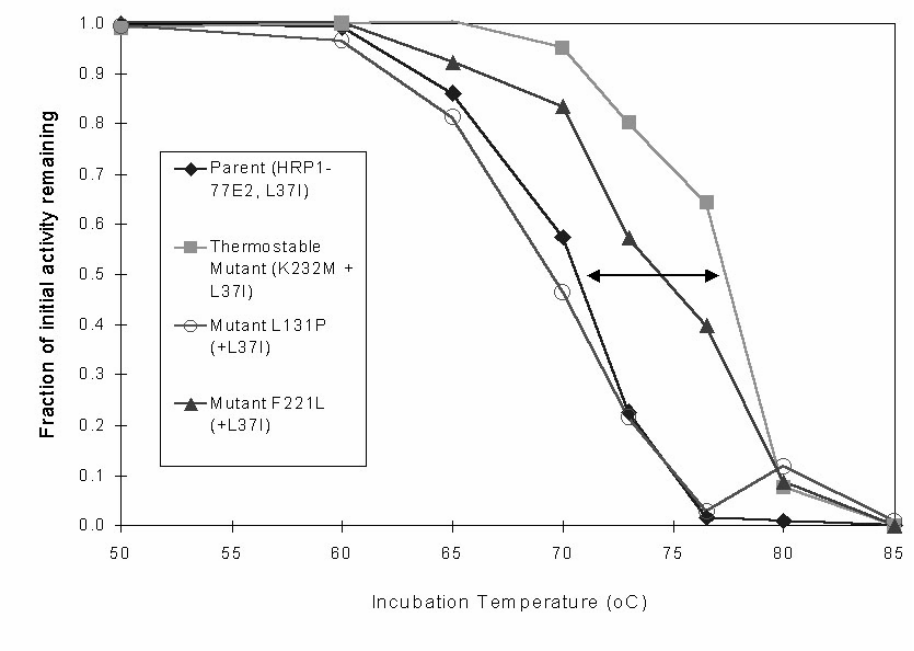


Figure A.2. Thermostability of HRP mutants: residual activities after 10 minute incubations at the indicated temperatures.

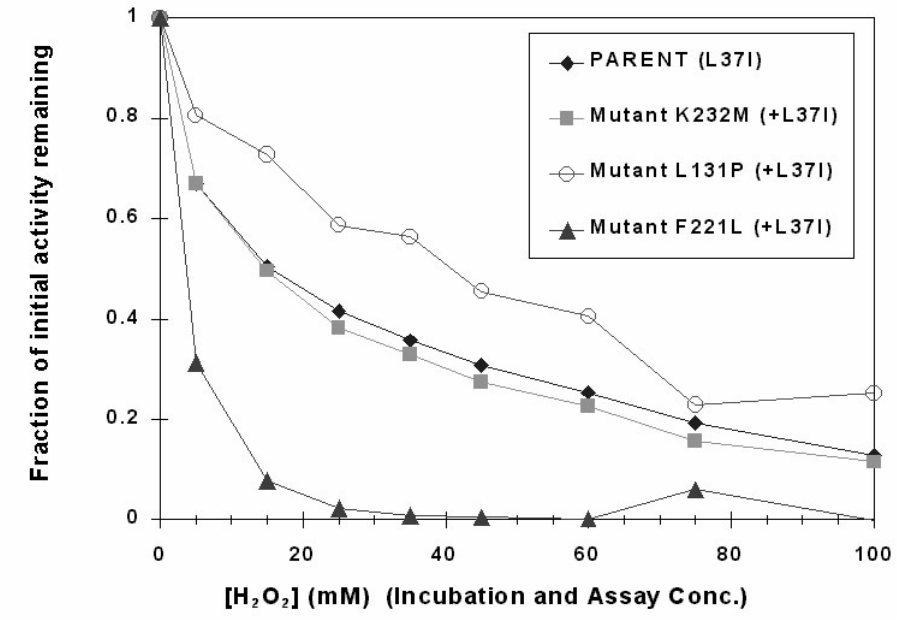


Figure A.3. H_2O_2-stability of HRP mutants: residual activities of mutants following 30 minute incubations in the indicated H_2O_2 concentrations.

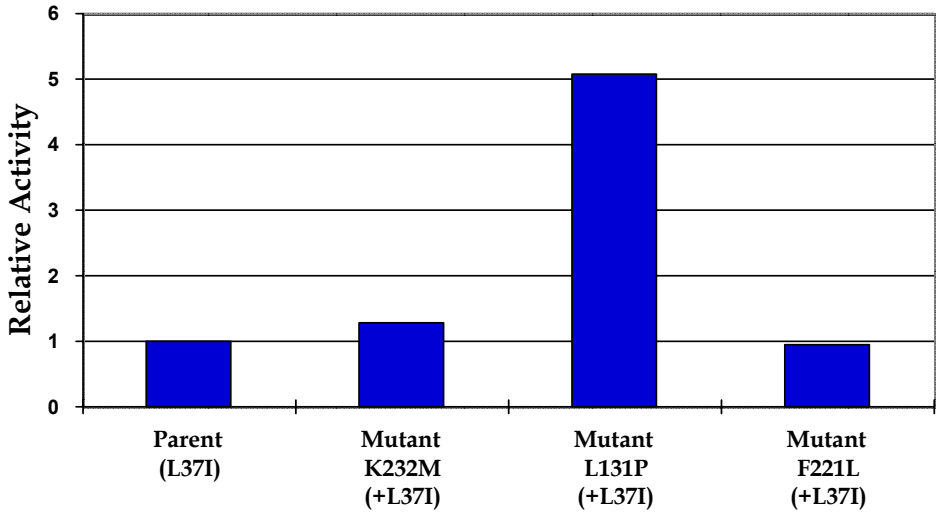


Figure A.4. Total activities of HRP mutant lysates relative to the parent.

Methods

The mutant library was generated using error-prone PCR with 0.1 mM Mn^{2+}, resulting in an error rate of approximately 1-2 mutations per gene (HRP is a 900 bp gene). The PCR products were purified with a Promega Wizard PCR kit and then digested with Sac I and Bam HI. The digestion products were then gel-purified with a QIAEX II gel extraction kit, and the HRP fragments were ligated into the plasmid pYEX-S1, shown in Figure A.5. The vector carries the yeast selectable nutrition markers *leu*2-d and URA3, as well as the *E. coli* ampicillin resistance gene. The plasmids were first transformed into *E. coli* strain HB101 to allow for plasmid propagation. Following plasmid recovery from *E. coli*, the plasmids were then transformed into *S. cerevisiae* strain BJ5464 for protein expression and secretion. Transformed yeast cells were plated on YNB selective medium supplemented with leucine, histidine, adenine, and tryptophan. Colonies were picked and

grown in 96-well microplates in YEPD medium at 30°C in an air-circulating incubator for 2 days plus 16 hours, and then screened.

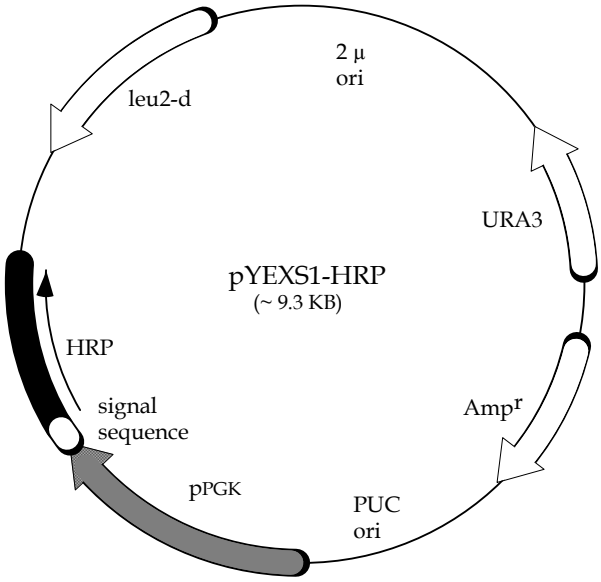


Figure A.5. Plasmid map of pYEXS1-HRP, used for propagation of the HRP gene in *E. coli* and expression of HRP in yeast. In this construct, expression is under the control of the constitutive phosphoglycerate kinase promoter.

Screening HRP Thermostability and H_2O_2-Stability

Yeast cultures expressing the HRP mutant library in 96-well microplates were centrifuged to pellet the cells. For initial activity measurement, clarified lysates (10 μl) were transferred to standard 96-well microplates. The plates were incubated at 25°C for 20 min and then 150 μl of ABTS assay solution (ABTS / H_2O_2 (0.5 mM) @ 25°C) was added to each plate and activity was measured in a plate reader as the rate of increase of 405 nm absorbance. For measurement of residual activity after heat treatment, clarified lysates (40 μl) were first added to 96-well PCR plates (GeneMate) containing 5 μl per

well of 20 mM $CaCl_2$ in H_2O. These plates were heated to 73°C for 10 min, rapidly cooled to 4°C for 10 min, and finally brought to 25°C for 5 min before residual activity was measured (using 10 μl of this lysate at 25°C, as in the initial activity measurement). Finally, for measurement of residual activity after incubation in H_2O_2, clarified lysates (30 μl) were transferred to standard 96-well microplates containing 30 μl of 50 mM H_2O_2 in YEPD. This mixture (20 μl) was assayed immediately for initial activity and the remainder was incubated for 30 min before residual activity was measured.

References

1. Villegas JA, Mauk AG, Vazquez-Duhalt R: **A cytochrome c variant resistant to heme degradation by hydrogen peroxide**. *Chem Biol* 2000, **7**:237-244.
2. Morawski B, Lin Z, Cirino P, Joo H, Bandara G, Arnold FH: **Functional expression of horseradish peroxidase in *Saccharomyces cerevisiae* and *Pichia pastoris***. *Protein Eng* 2000, **13**:377-384.
3. Romanos MA, Scorer CA, Clare JJ: **Foreign gene-expression in yeast - a review**. *Yeast* 1992, **8**:423-488.

APPENDIX B

Sequences of P450 BM-3 and pCWori(+) Vector

General Information

The heme domain includes the first 463 amino acids. There was no difference in activity in heme domain mutant F87A containing the first 463 amino acids compared to F87A mutant containing the first 470 amino acids. A 6-His tag was cloned onto the C-terminus of the heme domain.

All mutants containing mutation F87A (ttt to GCC) also contain a mutation at the third nucleotide of the Leu 86 codon (nucleotide 258: ttA to ttG). Therefore the total change in nucleotide sequence at amino acid positions 86 - 87 are:

256 - TTA TTT…

to 256 - TTG GCC…

Figure B.1 lists the nucleotide sequence of full-length, wild-type cytochrome P450 BM-3.

Figure B.2 lists the amino acid sequence of full-length, wild-type cytochrome P450 BM-3.

Figure B.3 lists the nucleotide sequence of the heme domain, wild-type cytochrome P450 BM-3 containing nucleotides coding for the C-terminal 6-His tag.

Figure B.4 lists the amino acid sequence of the heme domain, wild-type cytochrome P450 BM-3 including the 6-His tag.

Figure B.5 lists the nucleotide sequence of the pCWori(+) vector. The two Ptac promoters in this vector begin at nucleotide positions 167 and 262 (refer to plasmid map in Figure B.6).

Figure B.6 shows a map of plasmid pCWori(+)_BM-3(heme), which is the pCWori(+) vector with the P450 BM-3 heme domain cloned in.

Figure B.7 lists the nucleotide sequence of plasmid pCWori(+)_BM-3(HF87A), which contains the pCWori(+) vector with P450 BM-3 heme domain mutant F87A cloned into BamHI (5'-end, nucleotide position 334 in the pCWori+ vector sequence) and EcoRI (3'-end, nucleotide position 334 in the pCWOri vector sequence) sites. The numbering here corresponds to that of Figure B.6.

Figure B.1. Nucleotide sequence of full-length WT BM-3 (not including start codon).

```
   1 ACAATTAAAG AAATGCCTCA GCCAAAAACG TTTGGAGAGC TTAAAAATTT
  51 ACCGTTATTA AACACAGATA AACCGGTTCA AGCTTTGATG AAAATTGCGG
 101 ATGAATTAGG AGAAATCTTT AAATTCGAGG CGCCTGGTCG TGTAACGCGC
 151 TACTTATCAA GTCAGCGTCT AATTAAAGAA GCATGCGATG AATCACGCTT
 201 TGATAAAAAC TTAAGTCAAG CGCTTAAATT TGTACGTGAT TTTGCAGGAG
 251 ACGGGTTATT TACAAGCTGG ACGCATGAAA AAAATTGGAA AAAAGCGCAT
 301 AATATCTTAC TTCCAAGCTT CAGTCAGCAG GCAATGAAAG GCTATCATGC
 351 GATGATGGTC GATATCGCCG TGCAGCTTGT TCAAAAGTGG GAGCGTCTAA
 401 ATGCAGATGA GCATATTGAA GTACCGGAAG ACATGACACG TTTAACGCTT
 451 GATACAATTG GTCTTTGCGG CTTTAACTAT CGCTTTAACA GCTTTTACCG
 501 AGATCAGCCT CATCCATTTA TTACAAGTAT GGTCCGTGCA CTGGATGAAG
 551 CAATGAACAA GCTGCAGCGA GCAAATCCAG ACGACCCAGC TTATGATGAA
 601 AACAAGCGCC AGTTTCAAGA AGATATCAAG GTGATGAACG ACCTAGTAGA
 651 TAAAATTATT GCAGATCGCA AAGCAAGCGG TGAACAAAGC GATGATTTAT
 701 TAACGCATAT GCTAAACGGA AAAGATCCAG AAACGGGTGA GCCGCTTGAT
 751 GACGAGAACA TTCGCTATCA AATTATTACA TTCTTAATTG CGGGACACGA
 801 AACAACAAGT GGTCTTTTAT CATTTGCGCT GTATTTCTTA GTGAAAAATC
 851 CACATGTATT ACAAAAAGCA GCAGAAGAAG CAGCACGAGT TCTAGTAGAT
 901 CCTGTTCCAA GCTACAAACA AGTCAAACAG CTTAAATATG TCGGCATGGT
 951 CTTAAACGAA GCGCTGCGCT TATGGCCAAC TGCTCCTGCG TTTTCCCTAT
1001 ATGCAAAAGA AGATACGGTG CTTGGAGGAG AATATCCTTT AGAAAAAGGC
```

1051 GACGAACTAA TGGTTCTGAT TCCTCAGCTT CACCGTGATA AAACAATTTG
1101 GGGAGACGAT GTGGAAGAGT TCCGTCCAGA GCGTTTTGAA AATCCAAGTG
1151 CGATTCCGCA GCATGCGTTT AAACCGTTTG GAAACGGTCA GCGTGCGTGT
1201 ATCGGTCAGC AGTTCGCTCT TCATGAAGCA ACGCTGGTAC TTGGTATGAT
1251 GCTAAAACAC TTTGACTTTG AAGATCATAC AAACTACGAG CTCGATATTA
1301 AAGAAACTTT AACGTTAAAA CCTGAAGGCT TTGTGGTAAA AGCAAAATCG
1351 AAAAAAATTC CGCTTGGCGG TATTCCTTCA CCTAGCACTG AACAGTCTGC
1401 TAAAAAAGTA CGCAAAAAGG CAGAAAACGC TCATAATACG CCGCTGCTTG
1451 TGCTATACGG TTCAAATATG GGAACAGCTG AAGGAACGGC GCGTGATTTA
1501 GCAGATATTG CAATGAGCAA AGGATTTGCA CCGCAGGTCG CAACGCTTGA
1551 TTCACACGCC GGAAATCTTC CGCGCGAAGG AGCTGTATTA ATTGTAACGG
1601 CGTCTTATAA CGGTCATCCG CCTGATAACG CAAAGCAATT TGTCGACTGG
1651 TTAGACCAAG CGTCTGCTGA TGAAGTAAAA GGCGTTCGCT ACTCCGTATT
1701 TGGATGCGGC GATAAAAACT GGGCTACTAC GTATCAAAAA GTGCCTGCTT
1751 TTATCGATGA AACGCTTGCC GCTAAAGGGG CAGAAAACAT CGCTGACCGC
1801 GGTGAAGCAG ATGCAAGCGA CGACTTTGAA GGCACATATG AAGAATGGCG
1851 TGAACATATG TGGAGTGACG TAGCAGCCTA CTTTAACCTC GACATTGAAA
1901 ACAGTGAAGA TAATAAATCT ACTCTTTCAC TTCAATTTGT CGACAGCGCC
1951 GCGGATATGC CGCTTGCGAA AATGCACGGT GCGTTTTCAA CGAACGTCGT
2001 AGCAAGCAAA GAACTTCAAC AGCCAGGCAG TGCACGAAGC ACGCGACATC
2051 TTGAAATTGA ACTTCCAAAA GAAGCTTCTT ATCAAGAAGG AGATCATTTA
2101 GGTGTTATTC CTCGCAACTA TGAAGGAATA GTAAACCGTG TAACAGCAAG
2151 GTTCGGCCTA GATGCATCAC AGCAAATCCG TCTGGAAGCA GAAGAAGAAA
2201 AATTAGCTCA TTTGCCACTC GCTAAAACAG TATCCGTAGA AGAGCTTCTG
2251 CAATACGTGG AGCTTCAAGA TCCTGTTACG CGCACGCAGC TTCGCGCAAT
2301 GGCTGCTAAA ACGGTCTGCC CGCCGCATAA AGTAGAGCTT GAAGCCTTGC
2351 TTGAAAAGCA AGCCTACAAA GAACAAGTGC TGGCAAAACG TTTAACAATG
2401 CTTGAACTGC TTGAAAAATA CCCGGCGTGT GAAATGAAAT TCAGCGAATT
2451 TATCGCCCTT CTGCCAAGCA TACGCCCGCG CTATTACTCG ATTTCTTCAT
2501 CACCTCGTGT CGATGAAAAA CAAGCAAGCA TCACGGTCAG CGTTGTCTCA
2551 GGAGAAGCGT GGAGCGGATA TGGAGAATAT AAAGGAATTG CGTCGAACTA
2601 TCTTGCCGAG CTGCAAGAAG GAGATACGAT TACGTGCTTT ATTTCCACAC
2651 CGCAGTCAGA ATTTACGCTG CCAAAAGACC CTGAAACGCC GCTTATCATG
2701 GTCGGACCGG GAACAGGCGT CGCGCCGTTT AGAGGCTTTG TGCAGGCGCG
2751 CAAACAGCTA AAAGAACAAG GACAGTCACT TGGAGAAGCA CATTTATACT
2801 TCGGCTGCCG TTCACCTCAT GAAGACTATC TGTATCAAGA AGAGCTTGAA
2851 AACGCCCAAA GCGAAGGCAT CATTACGCTT CATACCGCTT TTTCTCGCAT
2901 GCCAAATCAG CCGAAAACAT ACGTTCAGCA CGTAATGGAA CAAGACGGCA
2951 AGAAATTGAT TGAACTTCTT GATCAAGGAG CGCACTTCTA TATTTGCGGA
3001 GACGGAAGCC AAATGGCACC TGCCGTTGAA GCAACGCTTA TGAAAAGCTA
3051 TGCTGACGTT CACCAAGTGA GTGAAGCAGA CGCTCGCTTA TGGCTGCAGC
3101 AGCTAGAAGA AAAAGGCCGA TACGCAAAAG ACGTGTGGGC TGGGTAA

Figure B.2. Amino acid sequence of full-length WT BM-3.

1 TIKEMPQPKT FGELKNLPLL NTDKPVQALM KIADELGEIF KFEAPGRVTR
51 YLSSQRLIKE ACDESRFDKN LSQALKFVRD FAGDGLFTSW THEKNWKKAH
101 NILLPSFSQQ AMKGYHAMMV DIAVQLVQKW ERLNADEHIE VPEDMTRLTL
151 DTIGLCGFNY RFNSFYRDQP HPFITSMVRA LDEAMNKLQR ANPDDPAYDE
201 NKRQFQEDIK VMNDLVDKII ADRKASGEQS DDLLTHMLNG KDPETGEPLD
251 DENIRYQIIT FLIAGHETTS GLLSFALYFL VKNPHVLQKA AEEAARVLVD
301 PVPSYKQVKQ LKYVGMVLNE ALRLWPTAPA FSLYAKEDTV LGGEYPLEKG
351 DELMVLIPQL HRDKTIWGDD VEEFRPERFE NPSAIPQHAF KPFGNGQRAC
401 IGQQFALHEA TLVLGMMLKH FDFEDHTNYE LDIKETLTLK PEGFVVKAKS
451 KKIPLGGIPS PSTEQSAKKV RKKAENAHNT PLLVLYGSNM GTAEGTARDL
501 ADIAMSKGFA PQVATLDSHA GNLPREGAVL IVTASYNGHP PDNAKQFVDW
551 LDQASADEVK GVRYSVFGCG DKNWATTYQK VPAFIDETLA AKGAENIADR
601 GEADASDDFE GTYEEWREHM WSDVAAYFNL DIENSEDNKS TLSLQFVDSA
651 ADMPLAKMHG AFSTNVVASK ELQQPGSARS TRHLEIELPK EASYQEGDHL
701 GVIPRNYEGI VNRVTARFGL DASQQIRLEA EEEKLAHLPL AKTVSVEELL
751 QYVELQDPVT RTQLRAMAAK TVCPPHKVEL EALLEKQAYK EQVLAKRLTM
801 LELLEKYPAC EMKFSEFIAL LPSIRPRYYS ISSSPRVDEK QASITVSVVS
851 GEAWSGYGEY KGIASNYLAE LQEGDTITCF ISTPQSEFTL PKDPETPLIM
901 VGPGTGVAPF RGFVQARKQL KEQGQSLGEA HLYFGCRSPH EDYLYQEELE
951 NAQSEGIITL HTAFSRMPNQ PKTYVQHVME QDGKKLIELL DQGAHFYICG
1001 DGSQMAPAVE ATLMKSYADV HQVSEADARL WLQQLEEKGR YAKDVWAG

Figure B.3. Nucleotide sequence of wild-type heme domain P450 BM-3 (not including start codon).

```
   1 ACAATTAAAG AAATGCCTCA GCCAAAAACG TTTGGAGAGC TTAAAAATTT
  51 ACCGTTATTA AACACAGATA AACCGGTTCA AGCTTTGATG AAAATTGCGG
 101 ATGAATTAGG AGAAATCTTT AAATTCGAGG CGCCTGGTCG TGTAACGCGC
 151 TACTTATCAA GTCAGCGTCT AATTAAAGAA GCATGCGATG AATCACGCTT
 201 TGATAAAAAC TTAAGTCAAG CGCTTAAATT TGTACGTGAT TTTGCAGGAG
 251 ACGGGTTATT TACAAGCTGG ACGCATGAAA AAAATTGGAA AAAAGCGCAT
 301 AATATCTTAC TTCCAAGCTT CAGTCAGCAG GCAATGAAAG GCTATCATGC
 351 GATGATGGTC GATATCGCCG TGCAGCTTGT TCAAAAGTGG GAGCGTCTAA
 401 ATGCAGATGA GCATATTGAA GTACCGGAAG ACATGACACG TTTAACGCTT
 451 GATACAATTG GTCTTTGCGG CTTTAACTAT CGCTTTAACA GCTTTTACCG
 501 AGATCAGCCT CATCCATTTA TTACAAGTAT GGTCCGTGCA CTGGATGAAG
 551 CAATGAACAA GCTGCAGCGA GCAAATCCAG ACGACCCAGC TTATGATGAA
 601 AACAAGCGCC AGTTTCAAGA AGATATCAAG GTGATGAACG ACCTAGTAGA
 651 TAAAATTATT GCAGATCGCA AAGCAAGCGG TGAACAAAGC GATGATTTAT
 701 TAACGCATAT GCTAAACGGA AAAGATCCAG AAACGGGTGA GCCGCTTGAT
 751 GACGAGAACA TTCGCTATCA AATTATTACA TTCTTAATTG CGGGACACGA
 801 AACAACAAGT GGTCTTTTAT CATTTGCGCT GTATTTCTTA GTGAAAAATC
 851 CACATGTATT ACAAAAAGCA GCAGAAGAAG CAGCACGAGT TCTAGTAGAT
 901 CCTGTTCCAA GCTACAAACA AGTCAAACAG CTTAAATATG TCGGCATGGT
 951 CTTAAACGAA GCGCTGCGCT TATGGCCAAC TGCTCCTGCG TTTTCCCTAT
1001 ATGCAAAAGA AGATACGGTG CTTGGAGGAG AATATCCTTT AGAAAAAGGC
1051 GACGAACTAA TGGTTCTGAT TCCTCAGCTT CACCGTGATA AAACAATTTG
1101 GGGAGACGAT GTGGAAGAGT TCCGTCCAGA GCGTTTTGAA AATCCAAGTG
1151 CGATTCCGCA GCATGCGTTT AAACCGTTTG GAAACGGTCA GCGTGCGTGT
1201 ATCGGTCAGC AGTTCGCTCT TCATGAAGCA ACGCTGGTAC TTGGTATGAT
1251 GCTAAAACAC TTTGACTTTG AAGATCATAC AAACTACGAG CTCGATATTA
1301 AAGAAACTTT AACGTTAAAA CCTGAAGGCT TTGTGGTAAA AGCAAAATCG
1351 AAAAAAATTC CGCTTGGCGG TATTCCTTCA CCTAGCACTC ATCATCATCA
1401 TCATCATTAA
```

Figure B.4. Amino acid sequence of heme domain wild-type BM-3.

1 TIKEMPQPKT FGELKNLPLL NTDKPVQALM KIADELGEIF KFEAPGRVTR
51 YLSSQRLIKE ACDESRFDKN LSQALKFVRD FAGDGLFTSW THEKNWKKAH
101 NILLPSFSQQ AMKGYHAMMV DIAVQLVQKW ERLNADEHIE VPEDMTRLTL
151 DTIGLCGFNY RFNSFYRDQP HPFITSMVRA LDEAMNKLQR ANPDDPAYDE
201 NKRQFQEDIK VMNDLVDKII ADRKASGEQS DDLLTHMLNG KDPETGEPLD
251 DENIRYQIIT FLIAGHETTS GLLSFALYFL VKNPHVLQKA AEEAARVLVD
301 PVPSYKQVKQ LKYVGMVLNE ALRLWPTAPA FSLYAKEDTV LGGEYPLEKG
351 DELMVLIPQL HRDKTIWGDD VEEFRPERFE NPSAIPQHAF KPFGNGQRAC
401 IGQQFALHEA TLVLGMMLKH FDFEDHTNYE LDIKETLTLK PEGFVVKAKS
451 KKIPLGGIPS PSTHHHHHH

Figure B.5. Nucleotide sequence of vector pCWori.

1 AGCGGGCAGT GAGCGCAACG CAATTAATGT GAGTTAGCTC ACTCATTAGG
51 CACCCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATAAT GTGTGGAATT
101 GTGAGCGGAT AACAATTTCA CACAGGAAAC AGGATCGATC CATCGATGAG
151 CTTACTCCCC ATCCCCCTGT TGACAATTAA TCATCGGCTC GTATAATGTG
201 TGGAATTGTG AGCGGATAAC AATTTCACAC AGGAAACAGG ATCAGCTTAC
251 TCCCCATCCC CCTGTTGACA ATTAATCATC GGCTCGTATA ATGTGTGGAA
301 TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGGATCCA TCGATGCTTA
351 GGAGGTCATA TGTGAATTCA TCGATGATAA GCTGTCAAAC ATGAGCAGAT
401 CTGAGCCCGC CTAATGAGCG GGCTTTTTTT TCAGATCTGC TTGAAGACGA
451 AAGGGCCTCG TGATACGCCT ATTTTTATAG GTTAATGTCA TGATAATAAT
501 GGTTTCTTAG CGTCAAAGCA ACCATAGTAC GCGCCCTGTA GCGGCGCATT
551 AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT ACACTTGCCA
601 GCGCCCTAGC GCCCGCTCCT TTCGCTTTCT TCCCTTCCTT TCTCGCCACG
651 TTCGCCGGCT TTCCCCGTCA AGCTCTAAAT CGGGGGCTCC CTTTAGGGTT
701 CCGATTTAGT GCTTTACGGC ACCTCGACCC CAAAAAACTT GATTTGGGTG
751 ATGGTTCACG TAGTGGGCCA TCGCCCTGAT AGACGGTTTT TCGCCCTTTG
801 ACGTTGGAGT CCACGTTCTT TAATAGTGGA CTCTTGTTCC AAACTGGAAC
851 AACACTCAAC CCTATCTCGG GCTATTCTTT TGATTTATAA GGGATTTTGC
901 CGATTTCGGC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC
951 GCGAATTTTA ACAAAATATT AACGTTTACA ATTTCAGGTG GCACTTTTCG
1001 GGGAAATGTG CGCGGAACCC CTATTTGTTT ATTTTTCTAA ATACATTCAA
1051 ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT TCAATAATAT
1101 TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC
1151 CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG
1201 TGAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC
1251 GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC GCCCCGAAGA
1301 ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT
1351 TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG CATACACTAT
1401 TCTCAGAATG ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC
1451 GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG
1501 ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG
1551 CTAACCGCTT TTTTGCACAA CATGGGGGAT CATGTAACTC GCCTTGATCG
1601 TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA
1651 CGATGCCTGC AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA
1701 CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA
1751 TAAAGTTGCA GGACCACTTC TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA
1801 TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA
1851 GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC
1901 GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG
1951 GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT
2001 ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT
2051 GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT
2101 CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA
2151 GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC
2201 GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC
2251 CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC TGTCCTTCTA
2301 GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC
2351 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA
2401 AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG
2451 CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG
2501 AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG
2551 CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG
2601 GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA
2651 TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT

2701 TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG
2751 GCCTTTTTAC GGTTCCTGGC CTTTTGCTGG CCTTTTGCTC ACATGTTCTT
2801 TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC GCCTTTGAGT
2851 GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG
2901 AGCGAGGAAG CGGAAGAGCG CCTGATGCGG TATTTTCTCC TTACGCATCT
2951 GTGCGGTATT TCACACCGCA TATATGGTGC ACTCTCAGTA CAATCTGCTC
3001 TGATGCCGCA TAGTTAAGCC AGTATACACT CCGCTATCGC TACGTGACTG
3051 GGTCATGGCT GCGCCCCGAC ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC
3101 GGGCTTGTCT GCTCCCGGCA TCCGCTTACA GACAAGCTGT GACCGTCTCC
3151 GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA AACGCGCGAG
3201 GCAGAACGCC ATCAAAAATA ATTCGCGTCT GGCCTTCCTG TAGCCAGCTT
3251 TCATCAACAT TAAATGTGAG CGAGTAACAA CCCGTCGGAT TCTCCGTGGG
3301 AACAAACGGC GGATTGACCG TAATGGGATA GGTTACGTTG GTGTAGATGG
3351 GCGCATCGTA ACCGTGCATC TGCCAGTTTG AGGGGACGAC GACAGTATCG
3401 GCCTCAGGAA GATCGCACTC CAGCCAGCTT TCCGGCACCG CTTCTGGTGC
3451 CGGAAACCAG GCAAAGCGCC ATTCGCCATT CAGGCTGCGC AACTGTTGGG
3501 AAGGGCGATC GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GGCGAAAGGG
3551 GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT TTTCCCAGTC
3601 ACGACGTTGT AAAACGACGG CCAGTGAATC CGTAATCATG GTCATAGCTG
3651 TTTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA ACATACGAGC
3701 CGGAAGCATA AAGTGTAAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA
3751 CATTAATTGC GTTGCGCTCA CTGCCCGCTT TCCAGTCGGG AAACCTGTCG
3801 TGCCAGCTGC ATTAATGAAT CGGCCAACGC GCGGGGAGAG GCGGTTTGCG
3851 TATTGGGCGC CAGGGTGGTT TTTCTTTTCA CCAGTGAGAC GGGCAACAGC
3901 TGATTGCCCT TCACCGCCTG GCCCTGAGAG AGTTGCAGCA AGCGGTCCAC
3951 GCTGGTTTGC CCCAGCAGGC GAAAATCCTG TTTGATGGTG GTTAACGGCG
4001 GGATATAACA TGAGCTGTCT TCGGTATCGT CGTATCCCAC TACCGAGATA
4051 TCCGCACCAA CGCGCAGCCC GGACTCGGTA ATGGCGCGCA TTGCGCCCAG
4101 CGCCATCTGA TCGTTGGCAA CCAGCATCGC AGTGGGAACG ATGCCCTCAT
4151 TCAGCATTTG CATGGTTTGT TGAAAACCGG ACATGGCACT CCAGTCGCCT
4201 TCCCGTTCCG CTATCGGCTG AATTTGATTG CGAGTGAGAT ATTTATGCCA
4251 GCCAGCCAGA CGCAGACGCG CCGAGACAGA ACTTAATGGG CCCGCTAACA
4301 GCGCGATTTG CTGGTGACCC AATGCGACCA GATGCTCCAC GCCCAGTCGC
4351 GTACCGTCTT CATGGGAGAA AATAATACTG TTGATGGGTG TCTGGTCAGA
4401 GACATCAAGA AATAACGCCG GAACATTAGT GCAGGCAGCT TCCACAGCAA
4451 TGGCATCCTG GTCATCCAGC GGATAGTTAA TGATCAGCCC ACTGACGCGT
4501 TGCGCGAGAA GATTGTGCAC CGCCGCTTTA CAGGCTTCGA CGCCGCTTCG
4551 TTCTACCATC GACACCACCA CGCTGGCACC CAGTTGATCG GCGCGAGATT
4601 TAATCGCCGC GACAATTTGC GACGGCGCGT GCAGGGCCAG ACTGGAGGTG
4651 GCAACGCCAA TCAGCAACGA CTGTTTGCCC GCCAGTTGTT GTGCCACGCG
4701 GTTGGGAATG TAATTCAGCT CCGCCATCGC CGCTTCCACT TTTTCCCGCG
4751 TTTTCGCAGA AACGTGGCTG GCCTGGTTCA CCACGCGGGA AACGGTCTGA
4801 TAAGAGACAC CGGCATACTC TGCGACATCG TATAACGTTA CTGGTTTCAC
4851 ATTCACCACC CTGAATTGAC TCTCTTCCGG GCGCTATCAT GCCATACCGC
4901 GAAAGGTTTT GCGCCATTCG ATGGTGTCCT GGCACGACAG GTTTCCCGAC
4951 TGGAA

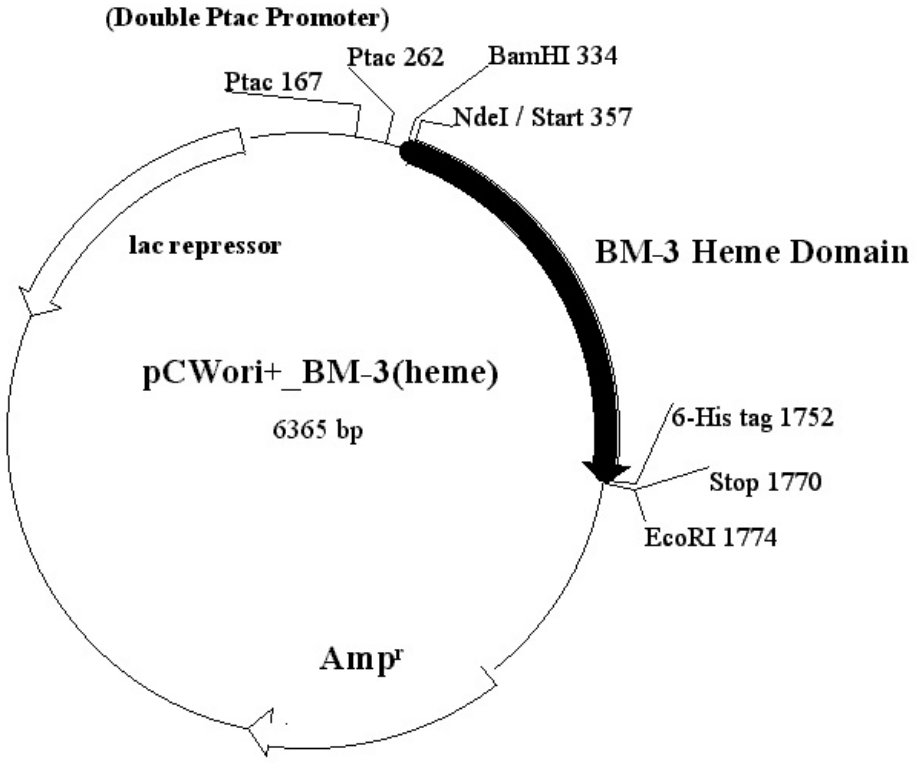


Figure B.6. Plasmid map of pCWori(+)_BM-3(Heme).

Figure B.7. Nucleotide sequence of plasmid containing vector pCWori and heme domain mutant HF87A.

1 AGCGGGCAGT GAGCGCAACG CAATTAATGT GAGTTAGCTC ACTCATTAGG
51 CACCCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATAAT GTGTGGAATT
101 GTGAGCGGAT AACAATTTCA CACAGGAAAC AGGATCGATC CATCGATGAG
151 CTTACTCCCC ATCCCCCTGT TGACAATTAA TCATCGGCTC GTATAATGTG
201 TGGAATTGTG AGCGGATAAC AATTTCACAC AGGAAACAGG ATCAGCTTAC
251 TCCCCATCCC CCTGTTGACA ATTAATCATC GGCTCGTATA ATGTGTGGAA
301 TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGGATCCA TCGATGCTTA
351 GGAGGTCATA TGACAATTAA AGAAATGCCT CAGCCAAAAA CGTTTGGAGA
401 GCTTAAAAAT TTACCGTTAT TAAACACAGA TAAACCGGTT CAAGCTTTGA
451 TGAAAATTGC GGATGAATTA GGAGAAATCT TTAAATTCGA GGCGCCTGGT
501 CGTGTAACGC GCTACTTATC AAGTCAGCGT CTAATTAAAG AAGCATGCGA
551 TGAATCACGC TTTGATAAAA ACTTAAGTCA AGCGCTTAAA TTTGTACGTG
601 ATTTTGCAGG AGACGGGTTG GCCACAAGCT GGACGCATGA AAAAAATTGG
651 AAAAAAGCGC ATAATATCTT ACTTCCAAGC TTCAGTCAGC AGGCAATGAA
701 AGGCTATCAT GCGATGATGG TCGATATCGC CGTGCAGCTT GTTCAAAAGT
751 GGGAGCGTCT AAATGCAGAT GAGCATATTG AAGTACCGGA AGACATGACA
801 CGTTTAACGC TTGATACGAT TGGTCTTTGC GGCTTTAACT ATCGCTTTAA
851 CAGCTTTTAC CGAGATCAGC CTCATCCATT TATTACAAGT ATGGTCCGTG
901 CACTGGATGA AGCAATGAAC AAGCTGCAGC GAGCAAATCC AGACGACCCA
951 GCTTATGATG AAAACAAGCG CCAGTTTCAA GAAGATATCA AGGTGATGAA
1001 CGACCTAGTA GATAAAATTA TTGCAGATCG CAAAGCAAGC GGTGAACAAA
1051 GCGATGATTT ATTAACGCAC ATGCTAAACG GAAAAGATCC AGAAACGGGT
1101 GAGCCGCTTG ATGACGAGAA CATTCGCTAT CAAATTATTA CATTCTTAAT
1151 TGCGGGACAC GAAACAACAA GTGGTCTTTT ATCATTTGCG CTGTATTTCT
1201 TAGTGAAAAA TCCACATGTA TTACAAAAAG CAGCAGAAGA AGCAGCACGA
1251 GTTCTAGTAG ATCCTGTTCC AAGCTACAAA CAAGTCAAAC AGCTTAAATA
1301 TGTCGGCATG GTCTTAAACG AAGCGCTGCG CTTATGGCCA ACTGCTCCTG
1351 CGTTTTCCCT ATATGCAAAA GAAGATACGG TGCTTGGAGG AGAATATCCT
1401 TTAGAAAAAG GCGACGAACT AATGGTTCTG ATTCCTCAGC TTCACCGTGA
1451 TAAAACAATT TGGGGAGACG ATGTGGAAGA GTTCCGTCCA GAGCGTTTTG
1501 AAAATCCAAG TGCGATTCCG CAGCATGCGT TTAAACCGTT TGGAAACGGT
1551 CAGCGTGCGT GTATCGGTCA GCAGTTCGCT CTTCATGAAG CAACGCTGGT
1601 ACTTGGTATG ATGCTAAAAC ACTTTGACTT TGAAGATCAT ACAAACTACG
1651 AGCTGGATAT TAAAGAAACT TTAACGTTAA AACCTGAAGG CTTTGTGGTA
1701 AAAGCAAAAT CGAAAAAAAT TCCGCTTGGC GGTATTCCTT CACCTAGCAC
1751 TCATCATCAT CATCATCATT AATGAATTCA TCGATGATAA GCTGTCAAAC
1801 ATGAGCAGAT CTGAGCCCGC CTAATGAGCG GGCTTTTTTT TCAGATCTGC
1851 TTGAAGACGA AAGGGCCTCG TGATACGCCT ATTTTTATAG GTTAATGTCA
1901 TGATAATAAT GGTTTCTTAG CGTCAAAGCA ACCATAGTAC GCGCCCTGTA
1951 GCGGCGCATT AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT
2001 ACACTTGCCA GCGCCCTAGC GCCCGCTCCT TTCGCTTTCT TCCCTTCCTT
2051 TCTCGCCACG TTCGCCGGCT TTCCCCGTCA AGCTCTAAAT CGGGGGCTCC
2101 CTTTAGGGTT CCGATTTAGT GCTTTACGGC ACCTCGACCC CAAAAAACTT
2151 GATTTGGGTG ATGGTTCACG TAGTGGGCCA TCGCCCTGAT AGACGGTTTT
2201 TCGCCCTTTG ACGTTGGAGT CCACGTTCTT TAATAGTGGA CTCTTGTTCC
2251 AAACTGGAAC AACACTCAAC CCTATCTCGG GCTATTCTTT TGATTTATAA
2301 GGGATTTTGC CGATTTCGGC CTATTGGTTA AAAAATGAGC TGATTTAACA
2351 AAAATTTAAC GCGAATTTTA ACAAAATATT AACGTTTACA ATTTCAGGTG
2401 GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT ATTTTTCTAA
2451 ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT
2501 TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG
2551 CCCTTATTCC CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA

2601 GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT
2651 GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC
2701 GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT
2751 GGCGCGGTAT TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG
2801 CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA GTCACAGAAA
2851 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA
2901 ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG
2951 ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT CATGTAACTC
3001 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG
3051 CGTGACACCA CGATGCCTGC AGCAATGGCA ACAACGTTGC GCAAACTATT
3101 AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA
3151 TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC CCTTCCGGCT
3201 GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG
3251 TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA
3301 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC
3351 GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT
3401 TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA
3451 GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA
3501 CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG
3551 ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA
3601 AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA
3651 ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC
3701 TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG
3751 CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC
3801 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC
3851 GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA
3901 GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA
3951 TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
4001 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA
4051 ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG
4101 CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC
4151 CAGCAACGCG GCCTTTTTAC GGTTCCTGGC CTTTTGCTGG CCTTTTGCTC
4201 ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC
4251 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG
4301 CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG CCTGATGCGG TATTTTCTCC
4351 TTACGCATCT GTGCGGTATT TCACACCGCA TATATGGTGC ACTCTCAGTA
4401 CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGTATACACT CCGCTATCGC
4451 TACGTGACTG GGTCATGGCT GCGCCCCGAC ACCCGCCAAC ACCCGCTGAC
4501 GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA TCCGCTTACA GACAAGCTGT
4551 GACCGTCTCC GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA
4601 AACGCGCGAG GCAGAACGCC ATCAAAAATA ATTCGCGTCT GGCCTTCCTG
4651 TAGCCAGCTT TCATCAACAT TAAATGTGAG CGAGTAACAA CCCGTCGGAT
4701 TCTCCGTGGG AACAAACGGC GGATTGACCG TAATGGGATA GGTTACGTTG
4751 GTGTAGATGG GCGCATCGTA ACCGTGCATC TGCCAGTTTG AGGGGACGAC
4801 GACAGTATCG GCCTCAGGAA GATCGCACTC CAGCCAGCTT TCCGGCACCG
4851 CTTCTGGTGC CGGAAACCAG GCAAAGCGCC ATTCGCCATT CAGGCTGCGC
4901 AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT TACGCCAGCT
4951 GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT
5001 TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATC CGTAATCATG
5051 GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA
5101 ACATACGAGC CGGAAGCATA AAGTGTAAAG CCTGGGGTGC CTAATGAGTG
5151 AGCTAACTCA CATTAATTGC GTTGCGCTCA CTGCCCGCTT TCCAGTCGGG
5201 AAACCTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACGC GCGGGGAGAG
5251 GCGGTTTGCG TATTGGGCGC CAGGGTGGTT TTTCTTTTCA CCAGTGAGAC
5301 GGGCAACAGC TGATTGCCCT TCACCGCCTG GCCCTGAGAG AGTTGCAGCA
5351 AGCGGTCCAC GCTGGTTTGC CCCAGCAGGC GAAAATCCTG TTTGATGGTG

5401 GTTAACGGCG GGATATAACA TGAGCTGTCT TCGGTATCGT CGTATCCCAC
5451 TACCGAGATA TCCGCACCAA CGCGCAGCCC GGACTCGGTA ATGGCGCGCA
5501 TTGCGCCCAG CGCCATCTGA TCGTTGGCAA CCAGCATCGC AGTGGGAACG
5551 ATGCCCTCAT TCAGCATTTG CATGGTTTGT TGAAAACCGG ACATGGCACT
5601 CCAGTCGCCT TCCCGTTCCG CTATCGGCTG AATTTGATTG CGAGTGAGAT
5651 ATTTATGCCA GCCAGCCAGA CGCAGACGCG CCGAGACAGA ACTTAATGGG
5701 CCCGCTAACA GCGCGATTTG CTGGTGACCC AATGCGACCA GATGCTCCAC
5751 GCCCAGTCGC GTACCGTCTT CATGGGAGAA AATAATACTG TTGATGGGTG
5801 TCTGGTCAGA GACATCAAGA AATAACGCCG GAACATTAGT GCAGGCAGCT
5851 TCCACAGCAA TGGCATCCTG GTCATCCAGC GGATAGTTAA TGATCAGCCC
5901 ACTGACGCGT TGCGCGAGAA GATTGTGCAC CGCCGCTTTA CAGGCTTCGA
5951 CGCCGCTTCG TTCTACCATC GACACCACCA CGCTGGCACC CAGTTGATCG
6001 GCGCGAGATT TAATCGCCGC GACAATTTGC GACGGCGCGT GCAGGGCCAG
6051 ACTGGAGGTG GCAACGCCAA TCAGCAACGA CTGTTTGCCC GCCAGTTGTT
6101 GTGCCACGCG GTTGGGAATG TAATTCAGCT CCGCCATCGC CGCTTCCACT
6151 TTTTCCCGCG TTTTCGCAGA AACGTGGCTG GCCTGGTTCA CCACGCGGGA
6201 AACGGTCTGA TAAGAGACAC CGGCATACTC TGCGACATCG TATAACGTTA
6251 CTGGTTTCAC ATTCACCACC CTGAATTGAC TCTCTTCCGG GCGCTATCAT
6301 GCCATACCGC GAAAGGTTTT GCGCCATTCG ATGGTGTCCT GGCACGACAG
6351 GTTTCCCGAC TGGAA

APPENDIX C

Peroxygenase Activity Tests with Organic Peroxides

General Information

Using 12-pNCA as substrate, no peroxygenase activity was detected with HF87A or evolved mutant TH4 using the following organic peroxides: cumene hydroperoxide ("CuOOH"), *t*-butyl hydroperoxide ("tBuOOH"), or *m*-chloroperoxybenzoic acid ("mCPBA"). Peroxide concentrations were varied from 50 μM to 10 mM. Activity with peroxyacetic acid ("AcOOH") could be detected but was much less than that with H_2O_2 (using equal concentration of either peroxide), and is probably due to the presence of H_2O_2 in AcOOH.

Table C.1 shows the effects of adding 40 mM H_2O_2 to mixtures containing HF87A and 12-pNCA and different concentrations of tBuOOH. The middle column ("end-point reached") represents the end-point reached with the indicated amount of peroxide, before the addition of an extra 40 mM H_2O_2.

Table C.1. Effect of adding 40 mM H_2O_2 to reactions of HF87A with 12-pNCA and tBuOOH.

first added	end-point reached:	then add 40 mM H2O2
40 mM H2O2	0.395	0.396
100 mM tBuOOH	0	-0.01
10 mM tBuOOH	0.04	0.055
1 mM tBuOOH	0.062	0.262
0.1 mM tBuOOH	0.054	0.35
0.01 mM tBuOOH	0.055	0.388

No activity is seen with tBuOOH, although the enzyme is still active with H_2O_2 in the presence of 1 mM tBuOOH or less. Higher concentrations of tBuOOH inactivate the enzyme.

The heme absorbance peak is bleached much more slowly in 1 mM tBuOOH compared to 1 mM H_2O_2 for HF87A (2 μM) in the presence of myristic acid and 6% DMSO (without substrate, the heme is not bleached at any appreciable rate with either peroxide). 1 mM CuOOH causes rapid bleaching with and without substrate, while heme bleaching is slow in 50 μM CuOOH plus myristic acid.

Product Distributions and Relative TONs from Reactions with Myristic Acid Driven by H_2O_2, CuOOH, and tBuOOH

Table C.2 lists the results from reactions of various heme domain mutants with myristic acid and the indicated peroxide. Reported are the relative amounts of total hydroxylated product formed in each reaction (TON) as well as the product distributions from each reaction (reported as the percent of total hydroxylated product). Reactions (500 μl) contained 1 μM or 4 μM P450, 2% DMSO (final concentration in all reactions before extraction) and 1 mM myristic acid. Products were extracted and derivatized before analysis by GC/MS, as described in Chapter 7. Multiple additions of 0.1 mM or 0.05 mM peroxide were also tested, and are denoted as "0.1 mM x 3" or "0.05 mM x 4". These multiple additions were done every 5 minutes.

Table C.2. Hydroxylated product distributions and relative TONs from peroxygenase reactions of heme domain enzymes HF87A, 5H6, and HWT with myristic acid. Reactions were allowed to proceed for 20 minutes. TON represents the sum of all peak areas from all hydroxylated products formed in a given reaction relative to the area of an internal standard, which was added in an equal amount to all reactions.

REACTIONS: Myristic acid substrate, 2% DMSO, 4 uM HFA or 1 uM 5H6

Product Distributions (% of total hydroxylated product)

			(retention time, min.)					
		relative	(18.09)	(18.16)	(18.25)	(18.41)	(18.64)	(18.76)
HF87A		**TON**	**w-6**	**w-5**	**w-4**	**w-3**	**w-2**	**w-1**
H202	10 mM	1.11	3.5	60.2	14.9	5.0	4.1	12.4
	1 mM	0.51	3.4	60.4	14.3	5.4	4.1	12.4
	1 mM	0.36		62.3	16.9	5.5	3.8	11.5
tBuOOH	10 mM	0.72	3.3	53.2	12.9	4.4	3.3	23.0
	1 mM	0.98	3.0	53.2	11.5	3.8	3.1	25.5
	1 mM	0.87	2.7	53.8	13.4	4.0	2.9	23.2
	0.1 mM	0.14		59.3	12.6			28.1
	0.1 mM	0.15		56.6	12.6	3.1	2.8	24.9
	0.1 mM x 3	0.34	3.1	53.9	11.2	3.9	2.8	25.2
CuOOH	1 mM	0.54	1.8	21.9	6.7	2.8	2.7	64.1
	1 mM	0.43		22.6	7.8	4.0	3.3	62.3
	0.1 mM	0.88	1.2	21.0	5.3	1.8	2.0	68.7
	0.1 mM	0.62		22.4	6.0	2.2	2.2	67.1
	0.1 mM x 3	1.84	1.5	24.1	6.3	2.0	2.3	63.9
	0.05 mM x 4	1.90	1.4	22.4	5.8	2.2	2.2	65.9
5H6								
H2O2	10 mM	8.18	5.9	58.0	19.2	6.2	4.1	6.6
	1 mM	6.55	6.0	58.1	18.2	6.3	4.2	7.1
tBuOOH	10 mM	1.45	5.5	57.8	18.4	5.3	3.1	10.0
	1 mM	1.82	6.3	55.1	13.9	4.5	2.9	17.3
CuOOH	1 mM	0.98	3.1	23.9	8.2	4.5	5.0	55.4
HWT								
tBuOOH	1 mM	0.12				32.3	24.5	43.2
CuOOH	1 mM	0.32				33.4	25.2	41.4

Different regioselectivites are observed from reactions with tButOOH and CuOOH compared to with H_2O_2, as has been noted with other P450s (see Chapter 2). tBuOOH behaves much like H_2O_2 in that TONs increase with increasing amounts of

peroxide used. This reflects the relatively high K_ms for these peroxides compared to CuOOH. In reactions with HF87A, 100 μM CuOOH gives more product than 1 mM CuOOH, indicating that the K_m for CuOOH is low enough to (significantly) drive reactions with CuOOH concentrations that are significantly less destructive to the enzyme. Adding 100 μM CuOOH three times gives approximately three times more product than adding 100 μM CuOOH only once, indicating that enzyme inactivation in 100 μM CuOOH is very slow (perhaps there is no inactivation). Adding 50 μM CuOOH four times (200 μM total) gives even more product than adding 100 μM CuOOH three times (300 μM total). This is either because 50 μM is less destructive than 100 μM or because higher percentages of CuOOH get consumed in a catalase-like reaction as concentrations increase.

The products from these reactions were not quantified because no good authentic standard is available (refer to Chapter 2). However, if we assume that the "ratio" value of "1.1" for the reaction of HF87A in 10 mM H_2O_2 represents a TON of ~100 (see Chapter 2), then the reaction of HF87A with 50 μM CuOOH x 4 reaches a TON of ~200, indicating complete utilization of CuOOH toward substrate hydroxylation. Thus, peroxygenase reactions of the P450 BM-3 heme domain with CuOOH behave much like CPO with H_2O_2: the K_m for the peroxide is low enough to drive catalysis at micromolar concentrations, and as long as peroxide concentrations are maintained at these low concentrations, enzyme inactivation is suppressed.

TONs reached with tBuOOH and CuOOH are higher with evolved mutant 5H6 compared to with HF87A, although the improvements (~2-fold with 1 mM of either peroxide) are not as great as those seen with H_2O_2 (>10-fold with 1 mM H_2O_2).

While HWT has essentially no peroxygenase activity with H_2O_2 (refer to Chapter 2), HWT supports activity with CuOOH and tBuOOH, although this activity is still lower than that of HF87A.

APPENDIX D

MALDI-TOF MS Analysis of Heme Domain Protein Fragments before and after Treatment with H_2O_2

The spreadsheet below (Figure D.1) describes which peptide fragments have been identified (from digests of HWT, HF87A, or TH4) in samples prepared from enzyme incubated in H_2O_2 plus substrate as well as in control samples not exposed to H_2O_2. Peptide fragments were detected using Matrix Assisted Laser Desorption Ionization – Time-Of-Flight Mass Spectrometry (MALDI-TOF MS). *In silico* digest results reported by the software package "Sherpa Lite 4.0" are also listed below (Tables D.1 - D.7).

Note that the majority of possible peaks were not present in the control samples or the H_2O_2-treated samples, and that overall 34% of the protein was not identified by peptide fragments. There are on obvious peptide modifications, although some "unknown" peaks did appear occasionally.

Figure D.1. MALDI-TOF MS analysis of BM-3 heme domain peptide fragments.

* **"Covered" = Protein fragment identified in samples with and without H2O2**

ENDO LYSC DIGESTS			TRYPSIN DIGESTS			CHYMOTRYP DIGESTS		
COVERED *	ENZYME	MASS	COVERED *	ENZYME	MASS	COVERED *	ENZYME	MASS
	HFA	1107	42-47	TH4	677	* - AMBIGUOUS PEAK IDENTIFICATION		
203-210	HFA	1064	283-289	TH4	834	394-405	TH4	1278
290-306	HFA	1815		TH4	1051		TH4	1255
77-94	HFA	2037		TH4	946	116-126	TH4	1228
	HFA	2080	313-323	TH4	1265	* 422-429	TH4	1041
313-336	HFA	2711	168-179	TH4	1428	182-198	TH4	1948
188-202	HFA	1746	350-362	TH4	1521	* 239-256 / 280-298	TH4	2085
350-364	HFA	1764	365-378	TH4	1735	* 131-148 / 199-215	TH4	2123
452-469	HFA	1989	363-378	TH4	1978	1-11	TH4	1320
420-434	HFA	1923	379-398	TH4	2242			
453-469	HFA	1861	420-442	TH4	2834			
435-447	HFA	1461	337-362	TH4	2951			
			310-336	TH4	3078			
77-94	WT	2113	114-129	TH4	1803			
99-113	TH4	1670	133-147	TH4	1725			
420-442	TH4	2833						
42-59	TH4	2108						
32-59	TH4	3223						

UNKNOWN PEAKS:	MASS
ENDO LYS C DIGEST OF (TH4 + H2O2)	2583
CHYMOTRYPSIN DIGEST OF (TH4 + H2O2)	1744

PEAKS THAT DISAPPEARED AFTER TREATMENT WITH H2O2:		FRAGMENT POSITION
ENDO LYS C DIGEST OF HF87A AND HWT	1200	60-69
ENDO LYS C DIGEST OF HF87A AND HWT	2121	42-59

** I HELIX: 250-282

** L HELIX: 401-430

Ambiguous peaks: 239-256 / 280-298, 131-148 / 199-215

COVERED		NOT COVERED		Contains Met
1	11			M
		11	31	M
32	59			
		59	76	
77	94			
		94	98	
99	113			M
114	129			MM
		129	132	
133	147			
		147	167	
168	179			M
		179	181	
182	198			M
188	202			
203	210			
		210	282	MM
283	289			
290	306			
		306	309	
310	336			M
337	362			
350	364			M
363	378			
365	378			
379	398			
394	405			
		405	419	MM
420	434			
420	442			
420	442			
435	447			
		447	451	
452	469			
453	469			

159 residues not covered
34% NOT COVERED

Table D.1. HWT, Endoproteinase Lys-C digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 3	360.237	360.454	<TIK> E
1..2	1 - 9	1070.579	1071.307	<TIKEMPQPK> T
2	4 - 9	728.353	728.868	K <EMPQPK> T
2..3	4 - 15	1403.712	1404.651	K <EMPQPKTFGELK> N
3	10 - 15	693.370	693.798	K <TFGELK> N
3..4	10 - 31	2469.366	2470.960	K <TFGELKNLPLLNTDKPVQALMK> I
4	16 - 31	1794.007	1795.177	K <NLPLLNTDKPVQALMK> I
4..5	16 - 41	2909.594	2911.457	K <NLPLLNTDKPVQALMKIADELGEIF K> F
5	32 - 41	1133.597	1134.295	K <IADELGEIFK> F
5..6	32 - 59	3235.772	3237.751	K <IADELGEIFKFEAPGRVTRYLSSQR LIK> E
6	42 - 59	2120.186	2121.471	K <FEAPGRVTRYLSSQRLIK> E
6..7	42 - 69	3300.667	3302.720	K <FEAPGRVTRYLSSQRLIKEACDESR FDK> N
7	60 - 69	1198.492	1199.264	K <EACDESRFDK> N
7..8	60 - 76	1952.926	1954.149	K <EACDESRFDKNLSQALK> F
8	70 - 76	772.444	772.900	K <NLSQALK> F
8..9	70 - 94	2866.440	2868.202	K <NLSQALKFVRDFAGDGLFTSWTHEK> N
9	77 - 94	2112.007	2113.317	K <FVRDFAGDGLFTSWTHEK> N
9..10	77 - 97	2540.224	2541.809	K <FVRDFAGDGLFTSWTHEKNWK> K
10	95 - 97	446.228	446.507	K <NWK> K
10..11	95 - 98	574.323	574.681	K <NWKK> A
11	98 - 98	146.106	146.189	K <K> A
11..12	98 - 113	1811.972	1813.155	K <KAHNILLPSFSQQAMK> G
12	99 - 113	1683.877	1684.981	K <AHNILLPSFSQQAMK> G
12..13	99 - 129	3467.788	3470.146	K <AHNILLPSFSQQAMKGYHAMMVDIA VQLVQK> W
13	114 - 129	1801.922	1803.180	K <GYHAMMVDIAVQLVQK> W
13..14	114 - 187	8685.205	8690.963	K <GYHAMMVDIAVQLVQKWERLNADEH IEVPEDMTRLTLDTIGLCGFNYRFN SFYRDQPHPFITSMVRALDEAMNK> L
14	130 - 187	6901.294	6905.798	K <WERLNADEHIEVPEDMTRLTLDTIG LCGFNYRFNSFYRDQPHPFITSMVR ALDEAMNK> L
14..15	130 - 202	8628.085	8633.606	K <WERLNADEHIEVPEDMTRLTLDTIG LCGFNYRFNSFYRDQPHPFITSMVR ALDEAMNKLQRANPDDPAYDENK> R
15	188 - 202	1744.802	1745.824	K <LQRANPDDPAYDENK> R
15..16	188 - 210	2789.337	2790.987	K <LQRANPDDPAYDENKRQFQEDIK> V
16	203 - 210	1062.546	1063.179	K <RQFQEDIK> V
16..17	203 - 218	1976.999	1978.258	K <RQFQEDIKVMNDLVDK> I
17	211 - 218	932.464	933.094	K <VMNDLVDK> I
17..18	211 - 224	1628.892	1629.942	K <VMNDLVDKIIADRK> A
18	219 - 224	714.439	714.864	K <IIADRK> A
18..19	219 - 241	2511.275	2512.828	K <IIADRKASGEQSDDLLTHMLNGK> D
19	225 - 241	1814.847	1815.979	K <ASGEQSDDLLTHMLNGK> D
19..20	225 - 282	6422.194	6426.178	K <ASGEQSDDLLTHMLNGKDPETGEPL DDENIRYQIITFLIAGHETTSGLLS FALYFLVK> N
20	242 - 282	4625.358	4628.214	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLSFALYFLVK> N
20..21	242 - 289	5441.819	5445.173	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLSFALYFLVKNPHVLQK> A
21	283 - 289	834.471	834.974	K <NPHVLQK> A
21..22	283 - 306	2630.418	2632.016	K <NPHVLQKAAEEAARVLVDPVPSYK>Q
22	290 - 306	1813.957	1815.057	K <AAEEAARVLVDPVPSYK> Q
22..23	290 - 309	2169.179	2170.495	K <AAEEAARVLVDPVPSYKQVK> Q
23	307 - 309	373.233	373.453	K <QVK> Q
23..24	307 - 312	742.470	742.917	K <QVKQLK> Y
24	310 - 312	387.248	387.480	K <QLK> Y
24..25	310 - 336	3078.673	3080.689	K <QLKYVGMVLNEALRLWPTAPAFSLY AK> E

25	313 - 336	2709.435	2711.224	K <YVGMVLNEALRLWPTAPAFSLYAK>E
25..26	313 - 349	4140.128	4142.787	K <YVGMVLNEALRLWPTAPAFSLYAKE DTVLGGEYPLEK> G
26	337 - 349	1448.703	1449.578	K <EDTVLGGEYPLEK> G
26..27	337 - 364	3193.633	3195.642	K <EDTVLGGEYPLEKGDELMVLIPQLH RDK> T
27	350 - 364	1762.940	1764.080	K <GDELMVLIPQLHRDK> T
27..28	350 - 419	7971.006	7976.188	K <GDELMVLIPQLHRDKTIWGDDVEEF RPERFENPSAIPQHAFKPFGNGQRA CIGQQFALHEATLVLGMMLK> H
28	365 - 419	6226.077	6230.124	K <TIWGDDVEEFRPERFENPSAIPQHA FKPFGNGQRACIGQQFALHEATLVL GMMLK> H
28..29	365 - 434	8129.915	8135.135	K <TIWGDDVEEFRPERFENPSAIPQHA FKPFGNGQRACIGQQFALHEATLVL GMMLKHFDFEDHTNYELDIK> E
29	420 - 434	1921.848	1923.026	K <HFDFEDHTNYELDIK> E
29..30	420 - 447	3363.666	3365.745	K <HFDFEDHTNYELDIKETLTLKPEGF VVK> A
30	435 - 447	1459.829	1460.734	K <ETLTLKPEGFVVK> A
30..31	435 - 449	1658.961	1659.987	K <ETLTLKPEGFVVKAK> S
31	448 - 449	217.143	217.268	K <AK> S
31..32	448 - 451	432.270	432.521	K <AKSK> K
32	450 - 451	233.138	233.268	K <SK> K
32..33	450 - 452	361.233	361.442	K <SKK> I
33	452 - 452	146.106	146.189	K <K> I
33..34	452 - 469	1988.024	1989.231	K <KIPLGGIPSPSTHHHHHH>
34	453 - 469	1859.929	1861.057	K <IPLGGIPSPSTHHHHHH>

Table D.2. HWT, Trypsin digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 3	360.237	360.454	<TIK> E
1..2	1 - 9	1070.579	1071.307	<TIKEMPQPK> T
2	4 - 9	728.353	728.868	K <EMPQPK> T
2..3	4 - 15	1403.712	1404.651	K <EMPQPKTFGELK> N
3	10 - 15	693.370	693.798	K <TFGELK> N
3..4	10 - 31	2469.366	2470.960	K <TFGELKNLPLLNTDKPVQALMK> I
4	16 - 31	1794.007	1795.177	K <NLPLLNTDKPVQALMK> I
4..5	16 - 41	2909.594	2911.457	K <NLPLLNTDKPVQALMKIADELGEIF K> F
5	32 - 41	1133.597	1134.295	K <IADELGEIFK> F
5..6	32 - 47	1790.920	1792.022	K <IADELGEIFKFEAPGR> V
6	42 - 47	675.334	675.743	K <FEAPGR> V
6..7	42 - 50	1031.551	1032.168	K <FEAPGRVTR> Y
7	48 - 50	374.228	374.441	R <VTR> Y
7..8	48 - 56	1108.599	1109.251	R <VTRYLSSQR> L
8	51 - 56	752.382	752.826	R <YLSSQR> L
8..9	51 - 59	1106.645	1107.319	R <YLSSQRLIK> E
9	57 - 59	372.274	372.508	R <LIK> E
9..10	57 - 66	1162.565	1163.318	R <LIKEACDESR> F
10	60 - 66	808.302	808.825	K <EACDESR> F
10..11	60 - 69	1198.492	1199.264	K <EACDESRFDK> N
11	67 - 69	408.201	408.455	R <FDK> N
11..12	67 - 76	1162.635	1163.340	R <FDKNLSQALK> F
12	70 - 76	772.444	772.900	K <NLSQALK> F
12..13	70 - 79	1174.682	1175.397	K <NLSQALKFVR> D
13	77 - 79	420.249	420.512	K <FVR> D
13..14	77 - 94	2112.007	2113.317	K <FVRDFAGDGLFTSWTHEK> N
14	80 - 94	1709.769	1710.821	R <DFAGDGLFTSWTHEK> N
14..15	80 - 97	2137.986	2139.312	R <DFAGDGLFTSWTHEKNWK> K
15	95 - 97	446.228	446.507	K <NWK> K
15..16	95 - 98	574.323	574.681	K <NWKK> A
16	98 - 98	146.106	146.189	K <K> A
16..17	98 - 113	1811.972	1813.155	K <KAHNILLPSFSQQAMK> G
17	99 - 113	1683.877	1684.981	K <AHNILLPSFSQQAMK> G
17..18	99 - 129	3467.788	3470.146	K <AHNILLPSFSQQAMKGYHAMMVDIA VQLVQK> W
18	114 - 129	1801.922	1803.180	K <GYHAMMVDIAVQLVQK> W
18..19	114 - 132	2273.145	2274.697	K <GYHAMMVDIAVQLVQKWER> L
19	130 - 132	489.234	489.532	K <WER> L
19..20	130 - 147	2239.033	2240.439	K <WERLNADEHIEVPEDMTR> L
20	133 - 147	1767.810	1768.923	R <LNADEHIEVPEDMTR> L
20..21	133 - 161	3334.596	3336.752	R <LNADEHIEVPEDMTRLTLDTIGLCG FNYR> F
21	148 - 161	1584.797	1585.845	R <LTLDTIGLCGFNYR> F
21..22	148 - 167	2399.173	2400.744	R <LTLDTIGLCGFNYRFNSFYR> D
22	162 - 167	832.387	832.914	R <FNSFYR> D
22..23	162 - 179	2241.079	2242.547	R <FNSFYRDQPHPFITSMVR> A
23	168 - 179	1426.703	1427.648	R <DQPHPFITSMVR> A
23..24	168 - 187	2299.109	2300.646	R <DQPHPFITSMVRALDEAMNK> L
24	180 - 187	890.417	891.013	R <ALDEAMNK> L
24..25	180 - 190	1287.661	1288.491	R <ALDEAMNKLQR> A
25	188 - 190	415.254	415.493	K <LQR> A
25..26	188 - 202	1744.802	1745.824	K <LQRANPDDPAYDENK> R
26	191 - 202	1347.558	1348.346	R <ANPDDPAYDENK> R
26..27	191 - 203	1503.659	1504.533	R <ANPDDPAYDENKR> Q
27	203 - 203	174.112	174.203	K <R> Q
27..28	203 - 210	1062.546	1063.179	K <RQFQEDIK> V
28	204 - 210	906.445	906.991	R <QFQEDIK> V
28..29	204 - 218	1820.898	1822.070	R <QFQEDIKVMNDLVDK> I
29	211 - 218	932.464	933.094	K <VMNDLVDK> I
29..30	211 - 223	1500.797	1501.768	K <VMNDLVDKIIADR> K
30	219 - 223	586.344	586.689	K <IIADR> K
30..31	219 - 224	714.439	714.864	K <IIADRK> A
31	224 - 224	146.106	146.189	R <K> A

31..32	224 - 241	1942.942	1944.154	R <KASGEQSDDLLTHMLNGK> D
32	225 - 241	1814.847	1815.979	K <ASGEQSDDLLTHMLNGK> D
32..33	225 - 255	3395.542	3397.593	K <ASGEQSDDLLTHMLNGKDPETGEPL DDENIR> Y
33	242 - 255	1598.706	1599.629	K <DPETGEPLDDENIR> Y
33..34	242 - 282	4625.358	4628.214	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLSFALYFLVK> N
34	256 - 282	3044.663	3046.601	R <YQIITFLIAGHETTSGLLSFALYFL VK> N
34..35	256 - 289	3861.123	3863.560	R <YQIITFLIAGHETTSGLLSFALYFL VKNPHVLQK> A
35	283 - 289	834.471	834.974	K <NPHVLQK> A
35..36	283 - 296	1532.806	1533.708	K <NPHVLQKAAEEAAR> V
36	290 - 296	716.345	716.749	K <AAEEAAR> V
36..37	290 - 306	1813.957	1815.057	K <AAEEAARVLVDPVPSYK> Q
37	297 - 306	1115.623	1116.323	R <VLVDPVPSYK> Q
37..38	297 - 309	1470.845	1471.761	R <VLVDPVPSYKQVK> Q
38	307 - 309	373.233	373.453	K <QVK> Q
38..39	307 - 312	742.470	742.917	K <QVKQLK> Y
39	310 - 312	387.248	387.480	K <QLK> Y
39..40	310 - 323	1632.902	1633.976	K <QLKYVGMVLNEALR> L
40	313 - 323	1263.665	1264.512	K <YVGMVLNEALR> L
40..41	313 - 336	2709.435	2711.224	K <YVGMVLNEALRLWPTAPAFSLYAK> E
41	324 - 336	1463.781	1464.728	R <LWPTAPAFSLYAK> E
41..42	324 - 349	2894.474	2896.290	R <LWPTAPAFSLYAKEDTVLGGEYPLE K> G
42	337 - 349	1448.703	1449.578	K <EDTVLGGEYPLEK> G
42..43	337 - 362	2950.511	2952.380	K <EDTVLGGEYPLEKGDELMVLIPQLH R> D
43	350 - 362	1519.818	1520.817	K <GDELMVLIPQLHR> D
43..44	350 - 364	1762.940	1764.080	K <GDELMVLIPQLHRDK> T
44	363 - 364	261.132	261.278	R <DK> T
44..45	363 - 378	1990.938	1992.133	R <DKTIWGDDVEEFRPER> F
45	365 - 378	1747.817	1748.870	K <TIWGDDVEEFRPER> F
45..46	365 - 398	3970.914	3973.337	K <TIWGDDVEEFRPERFENPSAIPQHA FKPFGNGQR> A
46	379 - 398	2241.108	2242.482	R <FENPSAIPQHAFKPFGNGQR> A
46..47	379 - 419	4496.271	4499.269	R <FENPSAIPQHAFKPFGNGQRACIGQ QFALHEATLVLGMMLK> H
47	399 - 419	2273.173	2274.802	R <ACIGQQFALHEATLVLGMMLK> H
47..48	399 - 434	4177.011	4179.812	R <ACIGQQFALHEATLVLGMMLKHFDF EDHTNYELDIK> E
48	420 - 434	1921.848	1923.026	K <HFDFEDHTNYELDIK> E
48..49	420 - 447	3363.666	3365.745	K <HFDFEDHTNYELDIKETLTLKPEGF VVK> A
49	435 - 447	1459.829	1460.734	K <ETLTLKPEGFVVK> A
49..50	435 - 449	1658.961	1659.987	K <ETLTLKPEGFVVKAK> S
50	448 - 449	217.143	217.268	K <AK> S
50..51	448 - 451	432.270	432.521	K <AKSK> K
51	450 - 451	233.138	233.268	K <SK> K
51..52	450 - 452	361.233	361.442	K <SKK> I
52	452 - 452	146.106	146.189	K <K> I
52..53	452 - 469	1988.024	1989.231	K <KIPLGGIPSPSTHHHHHH>
53	453 - 469	1859.929	1861.057	K <IPLGGIPSPSTHHHHHH>

Table D.3. HF87A, Endoproteinase Lys-C digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 3	360.237	360.454	<TIK> E
1..2	1 - 9	1070.579	1071.307	<TIKEMPQPK> T
2	4 - 9	728.353	728.868	K <EMPQPK> T
2..3	4 - 15	1403.712	1404.651	K <EMPQPKTFGELK> N
3	10 - 15	693.370	693.798	K <TFGELK> N
3..4	10 - 31	2469.366	2470.960	K <TFGELKNLPLLNTDKPVQALMK> I
4	16 - 31	1794.007	1795.177	K <NLPLLNTDKPVQALMK> I
4..5	16 - 41	2909.594	2911.457	K <NLPLLNTDKPVQALMKIADELGEIF K> F
5	32 - 41	1133.597	1134.295	K <IADELGEIFK> F
5..6	32 - 59	3235.772	3237.751	K <IADELGEIFKFEAPGRVTRYLSSQR LIK> E
6	42 - 59	2120.186	2121.471	K <FEAPGRVTRYLSSQRLIK> E
6..7	42 - 69	3300.667	3302.720	K <FEAPGRVTRYLSSQRLIKEACDESR FDK> N
7	60 - 69	1198.492	1199.264	K <EACDESRFDK> N
7..8	60 - 76	1952.926	1954.149	K <EACDESRFDKNLSQALK> F
8	70 - 76	772.444	772.900	K <NLSQALK> F
8..9	70 - 94	2790.409	2792.104	K <NLSQALKFVRDFAGDGLATSWTHEK> N
9	77 - 94	2035.975	2037.220	K <FVRDFAGDGLATSWTHEK> N
9..10	77 - 97	2464.192	2465.711	K <FVRDFAGDGLATSWTHEKNWK> K
10	95 - 97	446.228	446.507	K <NWK> K
10..11	95 - 98	574.323	574.681	K <NWKK> A
11	98 - 98	146.106	146.189	K <K> A
11..12	98 - 113	1811.972	1813.155	K <KAHNILLPSFSQQAMK> G
12	99 - 113	1683.877	1684.981	K <AHNILLPSFSQQAMK> G
12..13	99 - 129	3467.788	3470.146	K <AHNILLPSFSQQAMKGYHAMMVDIA VQLVQK> W
13	114 - 129	1801.922	1803.180	K <GYHAMMVDIAVQLVQK> W
13..14	114 - 187	8685.205	8690.963	K <GYHAMMVDIAVQLVQKWERLNADEH IEVPEDMTRLTLDTIGLCGFNYRFN SFYRDQPHPFITSMVRALDEAMNK> L
14	130 - 187	6901.294	6905.798	K <WERLNADEHIEVPEDMTRLTLDTIG LCGFNYRFNSFYRDQPHPFITSMVR ALDEAMNK> L
14..15	130 - 202	8628.085	8633.606	K <WERLNADEHIEVPEDMTRLTLDTIG LCGFNYRFNSFYRDQPHPFITSMVR ALDEAMNKLQRANPDDPAYDENK> R
15	188 - 202	1744.802	1745.824	K <LQRANPDDPAYDENK> R
15..16	188 - 210	2789.337	2790.987	K <LQRANPDDPAYDENKRQFQEDIK> V
16	203 - 210	1062.546	1063.179	K <RQFQEDIK> V
16..17	203 - 218	1976.999	1978.258	K <RQFQEDIKVMNDLVDK> I
17	211 - 218	932.464	933.094	K <VMNDLVDK> I
17..18	211 - 224	1628.892	1629.942	K <VMNDLVDKIIADRK> A
18	219 - 224	714.439	714.864	K <IIADRK> A
18..19	219 - 241	2511.275	2512.828	K <IIADRKASGEQSDDLLTHMLNGK> D
19	225 - 241	1814.847	1815.979	K <ASGEQSDDLLTHMLNGK> D
19..20	225 - 282	6422.194	6426.178	K <ASGEQSDDLLTHMLNGKDPETGEPL DDENIRYQIITFLIAGHETTSGLLS FALYFLVK> N
20	242 - 282	4625.358	4628.214	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLSFALYFLVK> N
20..21	242 - 289	5441.819	5445.173	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLSFALYFLVKNPHVLQK> A
21	283 - 289	834.471	834.974	K <NPHVLQK> A
21..22	283 - 306	2630.418	2632.016	K <NPHVLQKAAEEAARVLVDPVPSYK> Q
22	290 - 306	1813.957	1815.057	K <AAEEAARVLVDPVPSYK> Q
22..23	290 - 309	2169.179	2170.495	K <AAEEAARVLVDPVPSYKQVK> Q
23	307 - 309	373.233	373.453	K <QVK> Q
23..24	307 - 312	742.470	742.917	K <QVKQLK> Y
24	310 - 312	387.248	387.480	K <QLK> Y
24..25	310 - 336	3078.673	3080.689	K <QLKYVGMVLNEALRLWPTAPAFSLY AK> E
25	313 - 336	2709.435	2711.224	K <YVGMVLNEALRLWPTAPAFSLYAK> E

25..26	313 - 349	4140.128	4142.787	K <YVGMVLNEALRLWPTAPAFSLYAKE DTVLGGEYPLEK> G
26	337 - 349	1448.703	1449.578	K <EDTVLGGEYPLEK> G
26..27	337 - 364	3193.633	3195.642	K <EDTVLGGEYPLEKGDELMVLIPQLH RDK> T
27	350 - 364	1762.940	1764.080	K <GDELMVLIPQLHRDK> T
27..28	350 - 419	7971.006	7976.188	K <GDELMVLIPQLHRDKTIWGDDVEEF RPERFENPSAIPQHAFKPFGNGQRA CIGQQFALHEATLVLGMMLK> H
28	365 - 419	6226.077	6230.124	K <TIWGDDVEEFRPERFENPSAIPQHA FKPFGNGQRACIGQQFALHEATLVL GMMLK> H
28..29	365 - 434	8129.915	8135.135	K <TIWGDDVEEFRPERFENPSAIPQHA FKPFGNGQRACIGQQFALHEATLVL GMMLKHFDFEDHTNYELDIK> E
29	420 - 434	1921.848	1923.026	K <HFDFEDHTNYELDIK> E
29..30	420 - 447	3363.666	3365.745	K <HFDFEDHTNYELDIKETLTLKPEGF VVK> A
30	435 - 447	1459.829	1460.734	K <ETLTLKPEGFVVK> A
30..31	435 - 449	1658.961	1659.987	K <ETLTLKPEGFVVKAK> S
31	448 - 449	217.143	217.268	K <AK> S
31..32	448 - 451	432.270	432.521	K <AKSK> K
32	450 - 451	233.138	233.268	K <SK> K
32..33	450 - 452	361.233	361.442	K <SKK> I
33	452 - 452	146.106	146.189	K <K> I
33..34	452 - 469	1988.024	1989.231	K <KIPLGGIPSPSTHHHHHH>
34	453 - 469	1859.929	1861.057	K <IPLGGIPSPSTHHHHHH>

Table D.4. HF87A, Trypsin digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 3	360.237	360.454	<TIK> E
1..2	1 - 9	1070.579	1071.307	<TIKEMPQPK> T
2	4 - 9	728.353	728.868	K <EMPQPK> T
2..3	4 - 15	1403.712	1404.651	K <EMPQPKTFGELK> N
3	10 - 15	693.370	693.798	K <TFGELK> N
3..4	10 - 31	2469.366	2470.960	K <TFGELKNLPLLNTDKPVQALMK> I
4	16 - 31	1794.007	1795.177	K <NLPLLNTDKPVQALMK> I
4..5	16 - 41	2909.594	2911.457	K <NLPLLNTDKPVQALMKIADELGEIF K> F
5	32 - 41	1133.597	1134.295	K <IADELGEIFK> F
5..6	32 - 47	1790.920	1792.022	K <IADELGEIFKFEAPGR> V
6	42 - 47	675.334	675.743	K <FEAPGR> V
6..7	42 - 50	1031.551	1032.168	K <FEAPGRVTR> Y
7	48 - 50	374.228	374.441	R <VTR> Y
7..8	48 - 56	1108.599	1109.251	R <VTRYLSSQR> L
8	51 - 56	752.382	752.826	R <YLSSQR> L
8..9	51 - 59	1106.645	1107.319	R <YLSSQRLIK> E
9	57 - 59	372.274	372.508	R <LIK> E
9..10	57 - 66	1162.565	1163.318	R <LIKEACDESR> F
10	60 - 66	808.302	808.825	K <EACDESR> F
10..11	60 - 69	1198.492	1199.264	K <EACDESRFDK> N
11	67 - 69	408.201	408.455	R <FDK> N
11..12	67 - 76	1162.635	1163.340	R <FDKNLSQALK> F
12	70 - 76	772.444	772.900	K <NLSQALK> F
12..13	70 - 79	1174.682	1175.397	K <NLSQALKFVR> D
13	77 - 79	420.249	420.512	K <FVR> D
13..14	77 - 94	2035.975	2037.220	K <FVRDFAGDGLATSWTHEK> N
14	80 - 94	1633.737	1634.723	R <DFAGDGLATSWTHEK> N
14..15	80 - 97	2061.954	2063.214	R <DFAGDGLATSWTHEKNWK> K
15	95 - 97	446.228	446.507	K <NWK> K
15..16	95 - 98	574.323	574.681	K <NWKK> A
16	98 - 98	146.106	146.189	K <K> A
16..17	98 - 113	1811.972	1813.155	K <KAHNILLPSFSQQAMK> G
17	99 - 113	1683.877	1684.981	K <AHNILLPSFSQQAMK> G
17..18	99 - 129	3467.788	3470.146	K <AHNILLPSFSQQAMKGYHAMMVDIA VQLVQK> W
18	114 - 129	1801.922	1803.180	K <GYHAMMVDIAVQLVQK> W
18..19	114 - 132	2273.145	2274.697	K <GYHAMMVDIAVQLVQKWER> L
19	130 - 132	489.234	489.532	K <WER> L
19..20	130 - 147	2239.033	2240.439	K <WERLNADEHIEVPEDMTR> L
20	133 - 147	1767.810	1768.923	R <LNADEHIEVPEDMTR> L
20..21	133 - 161	3334.596	3336.752	R <LNADEHIEVPEDMTRLTLDTIGLCG FNYR> F
21	148 - 161	1584.797	1585.845	R <LTLDTIGLCGFNYR> F
21..22	148 - 167	2399.173	2400.744	R <LTLDTIGLCGFNYRFNSFYR> D
22	162 - 167	832.387	832.914	R <FNSFYR> D
22..23	162 - 179	2241.079	2242.547	R <FNSFYRDQPHPFITSMVR> A
23	168 - 179	1426.703	1427.648	R <DQPHPFITSMVR> A
23..24	168 - 187	2299.109	2300.646	R <DQPHPFITSMVRALDEAMNK> L
24	180 - 187	890.417	891.013	R <ALDEAMNK> L
24..25	180 - 190	1287.661	1288.491	R <ALDEAMNKLQR> A
25	188 - 190	415.254	415.493	K <LQR> A
25..26	188 - 202	1744.802	1745.824	K <LQRANPDDPAYDENK> R
26	191 - 202	1347.558	1348.346	R <ANPDDPAYDENK> R
26..27	191 - 203	1503.659	1504.533	R <ANPDDPAYDENKR> Q
27	203 - 203	174.112	174.203	K <R> Q
27..28	203 - 210	1062.546	1063.179	K <RQFQEDIK> V
28	204 - 210	906.445	906.991	R <QFQEDIK> V
28..29	204 - 218	1820.898	1822.070	R <QFQEDIKVMNDLVDK> I
29	211 - 218	932.464	933.094	K <VMNDLVDK> I
29..30	211 - 223	1500.797	1501.768	K <VMNDLVDKIIADR> K
30	219 - 223	586.344	586.689	K <IIADR> K
30..31	219 - 224	714.439	714.864	K <IIADRK> A
31	224 - 224	146.106	146.189	R <K> A

31..32	224 - 241	1942.942	1944.154	R <KASGEQSDDLLTHMLNGK> D
32	225 - 241	1814.847	1815.979	K <ASGEQSDDLLTHMLNGK> D
32..33	225 - 255	3395.542	3397.593	K <ASGEQSDDLLTHMLNGKDPETGEPL DDENIR> Y
33	242 - 255	1598.706	1599.629	K <DPETGEPLDDENIR> Y
33..34	242 - 282	4625.358	4628.214	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLSFALYFLVK> N
34	256 - 282	3044.663	3046.601	R <YQIITFLIAGHETTSGLLSFALYFL VK> N
34..35	256 - 289	3861.123	3863.560	R <YQIITFLIAGHETTSGLLSFALYFL VKNPHVLQK> A
35	283 - 289	834.471	834.974	K <NPHVLQK> A
35..36	283 - 296	1532.806	1533.708	K <NPHVLQKAAEEAAR> V
36	290 - 296	716.345	716.749	K <AAEEAAR> V
36..37	290 - 306	1813.957	1815.057	K <AAEEAARVLVDPVPSYK> Q
37	297 - 306	1115.623	1116.323	R <VLVDPVPSYK> Q
37..38	297 - 309	1470.845	1471.761	R <VLVDPVPSYKQVK> Q
38	307 - 309	373.233	373.453	K <QVK> Q
38..39	307 - 312	742.470	742.917	K <QVKQLK> Y
39	310 - 312	387.248	387.480	K <QLK> Y
39..40	310 - 323	1632.902	1633.976	K <QLKYVGMVLNEALR> L
40	313 - 323	1263.665	1264.512	K <YVGMVLNEALR> L
40..41	313 - 336	2709.435	2711.224	K <YVGMVLNEALRLWPTAPAFSLYAK> E
41	324 - 336	1463.781	1464.728	R <LWPTAPAFSLYAK> E
41..42	324 - 349	2894.474	2896.290	R <LWPTAPAFSLYAKEDTVLGGEYPLE K> G
42	337 - 349	1448.703	1449.578	K <EDTVLGGEYPLEK> G
42..43	337 - 362	2950.511	2952.380	K <EDTVLGGEYPLEKGDELMVLIPQLH R> D
43	350 - 362	1519.818	1520.817	K <GDELMVLIPQLHR> D
43..44	350 - 364	1762.940	1764.080	K <GDELMVLIPQLHRDK> T
44	363 - 364	261.132	261.278	R <DK> T
44..45	363 - 378	1990.938	1992.133	R <DKTIWGDDVEEFRPER> F
45	365 - 378	1747.817	1748.870	K <TIWGDDVEEFRPER> F
45..46	365 - 398	3970.914	3973.337	K <TIWGDDVEEFRPERFENPSAIPQHA FKPFGNGQR> A
46	379 - 398	2241.108	2242.482	R <FENPSAIPQHAFKPFGNGQR> A
46..47	379 - 419	4496.271	4499.269	R <FENPSAIPQHAFKPFGNGQRACIGQ QFALHEATLVLGMMLK> H
47	399 - 419	2273.173	2274.802	R <ACIGQQFALHEATLVLGMMLK> H
47..48	399 - 434	4177.011	4179.812	R <ACIGQQFALHEATLVLGMMLKHFDF EDHTNYELDIK> E
48	420 - 434	1921.848	1923.026	K <HFDFEDHTNYELDIK> E
48..49	420 - 447	3363.666	3365.745	K <HFDFEDHTNYELDIKETLTLKPEGF VVK> A
49	435 - 447	1459.829	1460.734	K <ETLTLKPEGFVVK> A
49..50	435 - 449	1658.961	1659.987	K <ETLTLKPEGFVVKAK> S
50	448 - 449	217.143	217.268	K <AK> S
50..51	448 - 451	432.270	432.521	K <AKSK> K
51	450 - 451	233.138	233.268	K <SK> K
51..52	450 - 452	361.233	361.442	K <SKK> I
52	452 - 452	146.106	146.189	K <K> I
52..53	452 - 469	1988.024	1989.231	K <KIPLGGIPSPSTHHHHHH>
53	453 - 469	1859.929	1861.057	K <IPLGGIPSPSTHHHHHH>

Table D.5. TH4, Endoproteinase Lys-C digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 3	360.237	360.454	<TIK> E
1..2	1 - 9	1070.579	1071.307	<TIKEMPQPK> T
2	4 - 9	728.353	728.868	K <EMPQPK> T
2..3	4 - 15	1403.712	1404.651	K <EMPQPKTFGELK> N
3	10 - 15	693.370	693.798	K <TFGELK> N
3..4	10 - 31	2469.366	2470.960	K <TFGELKNLPLLNTDKPVQALMK> I
4	16 - 31	1794.007	1795.177	K <NLPLLNTDKPVQALMK> I
4..5	16 - 41	2909.594	2911.457	K <NLPLLNTDKPVQALMKIADELGEIF K> F
5	32 - 41	1133.597	1134.295	K <IADELGEIFK> F
5..6	32 - 59	3221.756	3223.724	K <IADELGEIFKFEAPGRVTRYLSSQR LVK> E
6	42 - 59	2106.170	2107.445	K <FEAPGRVTRYLSSQRLVK> E
6..7	42 - 69	3286.652	3288.693	K <FEAPGRVTRYLSSQRLVKEACDESR FDK> N
7	60 - 69	1198.492	1199.264	K <EACDESRFDK> N
7..8	60 - 76	1952.926	1954.149	K <EACDESRFDKNLSQALK> F
8	70 - 76	772.444	772.900	K <NLSQALK> F
8..9	70 - 94	2790.409	2792.104	K <NLSQALKFVRDFAGDGLATSWTHEK> N
9	77 - 94	2035.975	2037.220	K <FVRDFAGDGLATSWTHEK> N
9..10	77 - 97	2464.192	2465.711	K <FVRDFAGDGLATSWTHEKNWK> K
10	95 - 97	446.228	446.507	K <NWK> K
10..11	95 - 98	574.323	574.681	K <NWKK> A
11	98 - 98	146.106	146.189	K <K> A
11..12	98 - 113	1797.029	1798.184	K <KARNILLPSLSQQAMK> G
12	99 - 113	1668.935	1670.010	K <ARNILLPSLSQQAMK> G
12..13	99 - 129	3452.846	3455.175	K <ARNILLPSLSQQAMKGYHAMMVDIA VQLVQK> W
13	114 - 129	1801.922	1803.180	K <GYHAMMVDIAVQLVQK> W
13..14	114 - 187	8641.196	8646.843	K <GYHAMMVDIAVQLVQKWERLNSDEH IEVPEDATRLTLDTIGLCGFNYRFN SFYRDQPHPFITSMVRALDEAMNK> L
14	130 - 187	6857.285	6861.678	K <WERLNSDEHIEVPEDATRLTLDTIG LCGFNYRFNSFYRDQPHPFITSMVR ALDEAMNK> L
14..15	130 - 202	8584.076	8589.486	K <WERLNSDEHIEVPEDATRLTLDTIG LCGFNYRFNSFYRDQPHPFITSMVR ALDEAMNKLQRANPDDPAYDENK> R
15	188 - 202	1744.802	1745.824	K <LQRANPDDPAYDENK> R
15..16	188 - 210	2789.337	2790.987	K <LQRANPDDPAYDENKRQFQEDIK> V
16	203 - 210	1062.546	1063.179	K <RQFQEDIK> V
16..17	203 - 218	1976.999	1978.258	K <RQFQEDIKVMNDLVDK> I
17	211 - 218	932.464	933.094	K <VMNDLVDK> I
17..18	211 - 224	1628.892	1629.942	K <VMNDLVDKIIADRK> A
18	219 - 224	714.439	714.864	K <IIADRK> A
18..19	219 - 241	2534.291	2535.865	K <IIADRKASGEQSDDLLTHMLHGK> D
19	225 - 241	1837.863	1839.017	K <ASGEQSDDLLTHMLHGK> D
19..20	225 - 282	6459.226	6463.243	K <ASGEQSDDLLTHMLHGKDPETGEPL DDENIRYQIITFLIAGHETTSGLLT FALYFLVK> N
20	242 - 282	4639.374	4642.241	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLTFALYFLVK> N
20..21	242 - 289	5455.834	5459.200	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLTFALYFLVKNPHVLQK> A
21	283 - 289	834.471	834.974	K <NPHVLQK> A
21..22	283 - 306	2630.418	2632.016	K <NPHVLQKAAEEAARVLVDPVPSYK> Q
22	290 - 306	1813.957	1815.057	K <AAEEAARVLVDPVPSYK> Q
22..23	290 - 309	2169.179	2170.495	K <AAEEAARVLVDPVPSYKQVK> Q
23	307 - 309	373.233	373.453	K <QVK> Q
23..24	307 - 312	742.470	742.917	K <QVKQLK> Y
24	310 - 312	387.248	387.480	K <QLK> Y
24..25	310 - 336	3078.673	3080.689	K <QLKYVGMVLNEALRIWPTAPAFSLY AK> E
25	313 - 336	2709.435	2711.224	K <YVGMVLNEALRIWPTAPAFSLYAK> E

25..26	313 - 349	4140.128	4142.787	K <YVGMVLNEALRIWPTAPAFSLYAKE DTVLGGEYPLEK> G
26	337 - 349	1448.703	1449.578	K <EDTVLGGEYPLEK> G
26..27	337 - 364	3193.633	3195.642	K <EDTVLGGEYPLEKGDELMVLIPQLH RDK> T
27	350 - 364	1762.940	1764.080	K <GDELMVLIPQLHRDK> T
27..28	350 - 419	7956.991	7962.162	K <GDELMVLIPQLHRDKTVWGDDVEEF RPERFENPSAIPQHAFKPFGNGQRA CIGQQFALHEATLVLGMMLK> H
28	365 - 419	6212.061	6216.097	K <TVWGDDVEEFRPERFENPSAIPQHA FKPFGNGQRACIGQQFALHEATLVL GMMLK> H
28..29	365 - 442	9027.396	9033.159	K <TVWGDDVEEFRPERFENPSAIPQHA FKPFGNGQRACIGQQFALHEATLVL GMMLKHFDFEDHTNYELDIEETLTL KPK> G
29	420 - 442	2833.345	2835.077	K <HFDFEDHTNYELDIEETLTLKPK> G
29..30	420 - 447	3377.682	3379.772	K <HFDFEDHTNYELDIEETLTLKPKGF VIK> A
30	443 - 447	562.348	562.710	K <GFVIK> A
30..31	443 - 449	761.480	761.963	K <GFVIKAK> S
31	448 - 449	217.143	217.268	K <AK> S
31..32	448 - 451	432.270	432.521	K <AKSK> K
32	450 - 451	233.138	233.268	K <SK> K
32..33	450 - 452	361.233	361.442	K <SKK> I
33	452 - 452	146.106	146.189	K <K> I
33..34	452 - 469	1988.024	1989.231	K <KIPLGGIPSPSTHHHHHH>
34	453 - 469	1859.929	1861.057	K <IPLGGIPSPSTHHHHHH>

Table D.6. TH4, Trypsin digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 3	360.237	360.454	<TIK> E
1..2	1 - 9	1070.579	1071.307	<TIKEMPQPK> T
2	4 - 9	728.353	728.868	K <EMPQPK> T
2..3	4 - 15	1403.712	1404.651	K <EMPQPKTFGELK> N
3	10 - 15	693.370	693.798	K <TFGELK> N
3..4	10 - 31	2469.366	2470.960	K <TFGELKNLPLLNTDKPVQALMK> I
4	16 - 31	1794.007	1795.177	K <NLPLLNTDKPVQALMK> I
4..5	16 - 41	2909.594	2911.457	K <NLPLLNTDKPVQALMKIADELGEIF K> F
5	32 - 41	1133.597	1134.295	K <IADELGEIFK> F
5..6	32 - 47	1790.920	1792.022	K <IADELGEIFKFEAPGR> V
6	42 - 47	675.334	675.743	K <FEAPGR> V
6..7	42 - 50	1031.551	1032.168	K <FEAPGRVTR> Y
7	48 - 50	374.228	374.441	R <VTR> Y
7..8	48 - 56	1108.599	1109.251	R <VTRYLSSQR> L
8	51 - 56	752.382	752.826	R <YLSSQR> L
8..9	51 - 59	1092.629	1093.292	R <YLSSQRLVK> E
9	57 - 59	358.258	358.482	R <LVK> E
9..10	57 - 66	1148.550	1149.291	R <LVKEACDESR> F
10	60 - 66	808.302	808.825	K <EACDESR> F
10..11	60 - 69	1198.492	1199.264	K <EACDESRFDK> N
11	67 - 69	408.201	408.455	R <FDK> N
11..12	67 - 76	1162.635	1163.340	R <FDKNLSQALK> F
12	70 - 76	772.444	772.900	K <NLSQALK> F
12..13	70 - 79	1174.682	1175.397	K <NLSQALKFVR> D
13	77 - 79	420.249	420.512	K <FVR> D
13..14	77 - 94	2035.975	2037.220	K <FVRDFAGDGLATSWTHEK> N
14	80 - 94	1633.737	1634.723	R <DFAGDGLATSWTHEK> N
14..15	80 - 97	2061.954	2063.214	R <DFAGDGLATSWTHEKNWK> K
15	95 - 97	446.228	446.507	K <NWK> K
15..16	95 - 98	574.323	574.681	K <NWKK> A
16	98 - 98	146.106	146.189	K <K> A
16..17	98 - 100	373.244	373.456	K <KAR> N
17	99 - 100	245.149	245.282	K <AR> N
17..18	99 - 113	1668.935	1670.010	K <ARNILLPSLSQQAMK> G
18	101 - 113	1441.796	1442.744	R <NILLPSLSQQAMK> G
18..19	101 - 129	3225.708	3227.909	R <NILLPSLSQQAMKGYHAMMVDIAVQ LVQK> W
19	114 - 129	1801.922	1803.180	K <GYHAMMVDIAVQLVQK> W
19..20	114 - 132	2273.145	2274.697	K <GYHAMMVDIAVQLVQKWER> L
20	130 - 132	489.234	489.532	K <WER> L
20..21	130 - 147	2195.024	2196.319	K <WERLNSDEHIEVPEDATR> L
21	133 - 147	1723.801	1724.802	R <LNSDEHIEVPEDATR> L
21..22	133 - 161	3290.588	3292.632	R <LNSDEHIEVPEDATRLTLDTIGLCG FNYR> F
22	148 - 161	1584.797	1585.845	R <LTLDTIGLCGFNYR> F
22..23	148 - 167	2399.173	2400.744	R <LTLDTIGLCGFNYRFNSFYR> D
23	162 - 167	832.387	832.914	R <FNSFYR> D
23..24	162 - 179	2241.079	2242.547	R <FNSFYRDQPHPFITSMVR> A
24	168 - 179	1426.703	1427.648	R <DQPHPFITSMVR> A
24..25	168 - 187	2299.109	2300.646	R <DQPHPFITSMVRALDEAMNK> L
25	180 - 187	890.417	891.013	R <ALDEAMNK> L
25..26	180 - 190	1287.661	1288.491	R <ALDEAMNKLQR> A
26	188 - 190	415.254	415.493	K <LQR> A
26..27	188 - 202	1744.802	1745.824	K <LQRANPDDPAYDENK> R
27	191 - 202	1347.558	1348.346	R <ANPDDPAYDENK> R
27..28	191 - 203	1503.659	1504.533	R <ANPDDPAYDENKR> Q
28	203 - 203	174.112	174.203	K <R> Q
28..29	203 - 210	1062.546	1063.179	K <RQFQEDIK> V
29	204 - 210	906.445	906.991	R <QFQEDIK> V
29..30	204 - 218	1820.898	1822.070	R <QFQEDIKVMNDLVDK> I
30	211 - 218	932.464	933.094	K <VMNDLVDK> I
30..31	211 - 223	1500.797	1501.768	K <VMNDLVDKIIADR> K
31	219 - 223	586.344	586.689	K <IIADR> K

31..32	219 - 224	714.439	714.864	K <IIADRK> A
32	224 - 224	146.106	146.189	R <K> A
32..33	224 - 241	1965.958	1967.191	R <KASGEQSDDLLTHMLHGK> D
33	225 - 241	1837.863	1839.017	K <ASGEQSDDLLTHMLHGK> D
33..34	225 - 255	3418.558	3420.630	K <ASGEQSDDLLTHMLHGKDPETGEPL DDENIR> Y
34	242 - 255	1598.706	1599.629	K <DPETGEPLDDENIR> Y
34..35	242 - 282	4639.374	4642.241	K <DPETGEPLDDENIRYQIITFLIAGH ETTSGLLTFALYFLVK> N
35	256 - 282	3058.678	3060.627	R <YQIITFLIAGHETTSGLLTFALYFL VK> N
35..36	256 - 289	3875.139	3877.586	R <YQIITFLIAGHETTSGLLTFALYFL VKNPHVLQK> A
36	283 - 289	834.471	834.974	K <NPHVLQK> A
36..37	283 - 296	1532.806	1533.708	K <NPHVLQKAAEEAAR> V
37	290 - 296	716.345	716.749	K <AAEEAAR> V
37..38	290 - 306	1813.957	1815.057	K <AAEEAARVLVDPVPSYK> Q
38	297 - 306	1115.623	1116.323	R <VLVDPVPSYK> Q
38..39	297 - 309	1470.845	1471.761	R <VLVDPVPSYKQVK> Q
39	307 - 309	373.233	373.453	K <QVK> Q
39..40	307 - 312	742.470	742.917	K <QVKQLK> Y
40	310 - 312	387.248	387.480	K <QLK> Y
40..41	310 - 323	1632.902	1633.976	K <QLKYVGMVLNEALR> I
41	313 - 323	1263.665	1264.512	K <YVGMVLNEALR> I
41..42	313 - 336	2709.435	2711.224	K <YVGMVLNEALRIWPTAPAFSLYAK> E
42	324 - 336	1463.781	1464.728	R <IWPTAPAFSLYAK> E
42..43	324 - 349	2894.474	2896.290	R <IWPTAPAFSLYAKEDTVLGGEYPLE K> G
43	337 - 349	1448.703	1449.578	K <EDTVLGGEYPLEK> G
43..44	337 - 362	2950.511	2952.380	K <EDTVLGGEYPLEKGDELMVLIPQLH R> D
44	350 - 362	1519.818	1520.817	K <GDELMVLIPQLHR> D
44..45	350 - 364	1762.940	1764.080	K <GDELMVLIPQLHRDK> T
45	363 - 364	261.132	261.278	R <DK> T
45..46	363 - 378	1976.923	1978.106	R <DKTVWGDDVEEFRPER> F
46	365 - 378	1733.801	1734.843	K <TVWGDDVEEFRPER> F
46..47	365 - 398	3956.898	3959.311	K <TVWGDDVEEFRPERFENPSAIPQHA FKPFGNGQR> A
47	379 - 398	2241.108	2242.482	R <FENPSAIPQHAFKPFGNGQR> A
47..48	379 - 419	4496.271	4499.269	R <FENPSAIPQHAFKPFGNGQRACIGQ QFALHEATLVLGMMLK> H
48	399 - 419	2273.173	2274.802	R <ACIGQQFALHEATLVLGMMLK> H
48..49	399 - 442	5088.507	5091.864	R <ACIGQQFALHEATLVLGMMLKHFDF EDHTNYELDIEETLTLKPK> G
49	420 - 442	2833.345	2835.077	K <HFDFEDHTNYELDIEETLTLKPK> G
49..50	420 - 447	3377.682	3379.772	K <HFDFEDHTNYELDIEETLTLKPKGF VIK> A
50	443 - 447	562.348	562.710	K <GFVIK> A
50..51	443 - 449	761.480	761.963	K <GFVIKAK> S
51	448 - 449	217.143	217.268	K <AK> S
51..52	448 - 451	432.270	432.521	K <AKSK> K
52	450 - 451	233.138	233.268	K <SK> K
52..53	450 - 452	361.233	361.442	K <SKK> I
53	452 - 452	146.106	146.189	K <K> I
53..54	452 - 469	1988.024	1989.231	K <KIPLGGIPSPSTHHHHHH>
54	453 - 469	1859.929	1861.057	K <IPLGGIPSPSTHHHHHH>

Table D.7. TH4, Chymotrypsin digest fragments

#	RESIDUES	MASS (Da.) MONO.	AVG.	SEQUENCE
1	1 - 11	1318.696	1319.588	<TIKEMPQPKTF> G
1..2	1 - 14	1617.844	1618.915	<TIKEMPQPKTFGEL> K
2	12 - 14	317.159	317.342	F <GEL> K
2..3	12 - 19	882.517	883.056	F <GELKNLPL> L
3	15 - 19	583.369	583.729	L <KNLPL> L
3..4	15 - 20	696.453	696.888	L <KNLPLL> N
4	20 - 20	131.095	131.175	L <L> N
4..5	20 - 29	1097.608	1098.265	L <LNTDKPVQAL> M
5	21 - 29	984.524	985.106	L <NTDKPVQAL> M
5..6	21 - 36	1784.934	1786.080	L <NTDKPVQALMKIADEL> G
6	30 - 36	818.421	818.990	L <MKIADEL> G
6..7	30 - 40	1264.637	1265.494	L <MKIADELGEIF> K
7	37 - 40	464.227	464.519	L <GEIF> K
7..8	37 - 42	739.391	739.870	L <GEIFKF> E
8	41 - 42	293.174	293.366	F <KF> E
8..9	41 - 51	1322.709	1323.518	F <KFEAPGRVTRY> L
9	43 - 51	1047.546	1048.167	F <EAPGRVTRY> L
9..10	43 - 52	1160.630	1161.327	F <EAPGRVTRYL> S
10	52 - 52	131.095	131.175	Y <L> S
10..11	52 - 57	702.402	702.809	Y <LSSQRL> V
11	53 - 57	589.318	589.650	L <SSQRL> V
11..12	53 - 67	1753.842	1754.942	L <SSQRLVKEACDESRF> D
12	58 - 67	1182.534	1183.308	L <VKEACDESRF> D
12..13	58 - 71	1652.783	1653.834	L <VKEACDESRFDKNL> S
13	68 - 71	488.259	488.541	F <DKNL> S
13..14	68 - 75	887.471	887.989	F <DKNLSQAL> K
14	72 - 75	417.222	417.463	L <SQAL> K
14..15	72 - 77	692.386	692.813	L <SQALKF> V
15	76 - 77	293.174	293.366	L <KF> V
15..16	76 - 81	810.439	810.952	L <KFVRDF> A
16	78 - 81	535.275	535.601	F <VRDF> A
16..17	78 - 86	948.466	949.032	F <VRDFAGDGL> A
17	82 - 86	431.202	431.446	F <AGDGL> A
17..18	82 - 90	876.398	876.922	F <AGDGLATSW> T
18	87 - 90	463.207	463.491	L <ATSW> T
18..19	87 - 96	1258.573	1259.344	L <ATSWTHEKNW> K
19	91 - 96	813.377	813.868	W <THEKNW> K
19..20	91 - 103	1636.916	1637.906	W <THEKNWKKARNIL> L
20	97 - 103	841.550	842.053	W <KKARNIL> L
20..21	97 - 107	1251.803	1252.567	W <KKARNILLPSL> S
21	104 - 107	428.263	428.529	L <LPSL> S
21..22	104 - 115	1321.670	1322.549	L <LPSLSQQAMKGY> H
22	108 - 115	911.417	912.035	L <SQQAMKGY> H
22..23	108 - 126	2120.022	2121.534	L <SQQAMKGYHAMMVDIAVQL> V
23	116 - 126	1226.615	1227.515	Y <HAMMVDIAVQL> V
23..24	116 - 130	1767.916	1769.166	Y <HAMMVDIAVQLVQKW> E
24	127 - 130	559.312	559.666	L <VQKW> E
24..25	127 - 133	957.540	958.129	L <VQKWERL> N
25	131 - 133	416.238	416.478	W <ERL> N
25..26	131 - 148	2122.029	2123.265	W <ERLNSDEHIEVPEDATRL> T
26	134 - 148	1723.801	1724.802	L <NSDEHIEVPEDATRL> T
26..27	134 - 150	1937.933	1939.067	L <NSDEHIEVPEDATRLTL> D
27	149 - 150	232.142	232.280	L <TL> D
27..28	149 - 155	731.406	731.845	L <TLDTIGL> C
28	151 - 155	517.275	517.580	L <DTIGL> C
28..29	151 - 158	824.374	824.953	L <DTIGLCGF> N
29	156 - 158	325.110	325.389	L <CGF> N
29..30	156 - 160	602.216	602.669	L <CGFNY> R
30	159 - 160	295.117	295.295	F <NY> R
30..31	159 - 162	598.286	598.659	F <NYRF> N
31	161 - 162	321.180	321.380	Y <RF> N
31..32	161 - 165	669.323	669.738	Y <RFNSF> Y
32	163 - 165	366.154	366.374	F <NSF> Y
32..33	163 - 166	529.217	529.550	F <NSFY> R

33	166 - 166	181.074	181.191	F <Y> R
33..34	166 - 173	1058.493	1059.150	F <YRDQPHPF> I
34	167 - 173	895.430	895.974	Y <RDQPHPF> I
34..35	167 - 181	1766.925	1768.074	Y <RDQPHPFITSMVRAL> D
35	174 - 181	889.505	890.115	F <ITSMVRAL> D
35..36	174 - 188	1690.875	1692.034	F <ITSMVRALDEAMNKL> Q
36	182 - 188	819.380	819.934	L <DEAMNKL> Q
36..37	182 - 198	1946.879	1948.101	L <DEAMNKLQRANPDDPAY> D
37	189 - 198	1145.510	1146.182	L <QRANPDDPAY> D
37..38	189 - 205	2062.946	2064.159	L <QRANPDDPAYDENKRQF> Q
38	199 - 205	935.446	935.993	Y <DENKRQF> Q
38..39	199 - 215	2121.016	2122.344	Y <DENKRQFQEDIKVMNDL> V
39	206 - 215	1203.581	1204.367	F <QEDIKVMNDL> V
39..40	206 - 233	3144.561	3146.481	F <QEDIKVMNDLVDKIIADRKASGEQS DDL> L
40	216 - 233	1958.991	1960.129	L <VDKIIADRKASGEQSDDL> L
40..41	216 - 234	2072.075	2073.289	L <VDKIIADRKASGEQSDDLL> T
41	234 - 234	131.095	131.175	L <L> T
41..42	234 - 238	613.326	613.779	L <LTHML> H
42	235 - 238	500.242	500.620	L <THML> H
42..43	235 - 249	1660.788	1661.857	L <THMLHGKDPETGEPL> D
43	239 - 249	1178.557	1179.252	L <HGKDPETGEPL> D
43..44	239 - 256	2083.945	2085.172	L <HGKDPETGEPLDDENIRY> Q
44	250 - 256	923.398	923.935	L <DDENIRY> Q
44..45	250 - 261	1525.741	1526.667	L <DDENIRYQIITF> L
45	257 - 261	620.353	620.747	Y <QIITF> L
45..46	257 - 262	733.437	733.906	Y <QIITFL> I
46	262 - 262	131.095	131.175	F <L> I
46..47	262 - 272	1097.572	1098.222	F <LIAGHETTSGL> L
47	263 - 272	984.488	985.062	L <IAGHETTSGL> L
47..48	263 - 273	1097.572	1098.222	L <IAGHETTSGLL> T
48	273 - 273	131.095	131.175	L <L> T
48..49	273 - 275	379.211	379.456	L <LTF> A
49	274 - 275	266.127	266.297	L <TF> A
49..50	274 - 277	450.248	450.535	L <TFAL> Y
50	276 - 277	202.132	202.254	F <AL> Y
50..51	276 - 278	365.195	365.430	F <ALY> F
51	278 - 278	181.074	181.191	L <Y> F
51..52	278 - 279	328.142	328.368	L <YF> L
52	279 - 279	165.079	165.192	Y <F> L
52..53	279 - 280	278.163	278.351	Y <FL> V
53	280 - 280	131.095	131.175	F <L> V
53..54	280 - 287	918.565	919.135	F <LVKNPHVL> Q
54	281 - 287	805.481	805.976	L <VKNPHVL> Q
54..55	281 - 298	1972.122	1973.307	L <VKNPHVLQKAAEEAARVL> V
55	288 - 298	1184.651	1185.346	L <QKAAEEAARVL> V
55..56	288 - 305	1942.016	1943.188	L <QKAAEEAARVLVDPVPSY> K
56	299 - 305	775.375	775.857	L <VDPVPSY> K
56..57	299 - 311	1499.835	1500.759	L <VDPVPSYKQVKQL> K
57	306 - 311	742.470	742.917	Y <KQVKQL> K
57..58	306 - 313	1033.628	1034.267	Y <KQVKQLKY> V
58	312 - 313	309.169	309.365	L <KY> V
58..59	312 - 318	808.452	809.041	L <KYVGMVL> N
59	314 - 318	517.293	517.690	Y <VGMVL> N
59..60	314 - 322	944.500	945.148	Y <VGMVLNEAL> R
60	319 - 322	445.217	445.473	L <NEAL> R
60..61	319 - 331	1484.778	1485.706	L <NEALRIWPTAPAF> S
61	323 - 331	1057.571	1058.249	L <RIWPTAPAF> S
61..62	323 - 333	1257.687	1258.486	L <RIWPTAPAFSL> Y
62	332 - 333	218.127	218.253	F <SL> Y
62..63	332 - 334	381.190	381.429	F <SLY> A
63	334 - 334	181.074	181.191	L <Y> A
63..64	334 - 341	937.476	938.046	L <YAKEDTVL> G
64	335 - 341	774.412	774.870	Y <AKEDTVL> G
64..65	335 - 347	1390.698	1391.541	Y <AKEDTVLGGEYPL> E
65	342 - 347	634.296	634.687	L <GGEYPL> E
65..66	342 - 353	1305.609	1306.392	L <GGEYPLEKGDEL> M
66	348 - 353	689.323	689.721	L <EKGDEL> M
66..67	348 - 356	1032.516	1033.211	L <EKGDELMVL> I
67	354 - 356	361.204	361.506	L <MVL> I
67..68	354 - 360	812.483	813.072	L <MVLIPQL> H

68	357 - 360	469.290	469.582	L <IPQL> H
68..69	357 - 367	1391.767	1392.624	L <IPQLHRDKTVW> G
69	361 - 367	940.488	941.058	L <HRDKTVW> G
69..70	361 - 374	1731.785	1732.827	L <HRDKTVWGDDVEEF> R
70	368 - 374	809.308	809.785	W <GDDVEEF> R
70..71	368 - 379	1494.674	1495.569	W <GDDVEEFRPERF> E
71	375 - 379	703.377	703.799	F <RPERF> E
71..72	375 - 390	1894.944	1896.096	F <RPERFENPSAIPQHAF> K
72	380 - 390	1209.578	1210.312	F <ENPSAIPQHAF> K
72..73	380 - 393	1581.794	1582.780	F <ENPSAIPQHAFKPF> G
73	391 - 393	390.227	390.483	F <KPF> G
73..74	391 - 405	1649.810	1650.882	F <KPFGNGQRACIGQQF> A
74	394 - 405	1277.594	1278.415	F <GNGQRACIGQQF> A
74..75	394 - 407	1461.715	1462.653	F <GNGQRACIGQQFAL> H
75	406 - 407	202.132	202.254	F <AL> H
75..76	406 - 412	753.402	753.854	F <ALHEATL> V
76	408 - 412	569.281	569.615	L <HEATL> V
76..77	408 - 414	781.433	781.908	L <HEATLVL> G
77	413 - 414	230.163	230.307	L <VL> G
77..78	413 - 418	662.350	662.916	L <VLGMML> K
78	415 - 418	450.197	450.624	L <GMML> K
78..79	415 - 421	862.419	863.116	L <GMMLKHF> D
79	419 - 421	430.233	430.507	L <KHF> D
79..80	419 - 423	692.328	692.772	L <KHFDF> E
80	422 - 423	280.106	280.281	F <DF> E
80..81	422 - 429	1039.388	1040.011	F <DFEDHTNY> E
81	424 - 429	777.293	777.746	F <EDHTNY> E
81..82	424 - 431	1019.420	1020.021	F <EDHTNYEL> D
82	430 - 431	260.137	260.290	Y <EL> D
82..83	430 - 437	960.465	961.034	Y <ELDIEETL> T
83	432 - 437	718.339	718.759	L <DIEETL> T
83..84	432 - 439	932.470	933.024	L <DIEETLTL> K
84	438 - 439	232.142	232.280	L <TL> K
84..85	438 - 444	789.475	789.973	L <TLKPKGF> V
85	440 - 444	575.343	575.709	L <KPKGF> V
85..86	440 - 455	1781.166	1782.290	L <KPKGFVIKAKSKKIPL> G
86	445 - 455	1223.833	1224.597	F <VIKAKSKKIPL> G
86..87	445 - 469	2742.531	2744.202	F <VIKAKSKKIPLGGIPSPSTHHHHHH>
87	456 - 469	1536.708	1537.621	L <GGIPSPSTHHHHHH>