EXPANDING THE CHEMICAL FUNCTIONALITY OF THE CELL WITH ORTHOGONAL ENZYME/COFACTOR PAIRS

Evan W. Reynolds

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill
2017

Approved by:
Eric Brustad
Dorothy Erie
Brian Kuhlman
Simon Meek
Matthew Redinbo
ABSTRACT

Evan W. Reynolds: Expanding the Chemical Functionality of the Cell with Orthogonal Enzyme/Cofactor Pairs (Under the direction of Eric M. Brustad)

Adapting enzymes for chemical synthesis is a longstanding and contemporary challenge; however, enzyme engineers have traditionally been limited to reactions found in nature. Meanwhile, synthetic reagents and catalysts provide a much broader scope of reactivity. Unnatural transition metal complexes have been incorporated into proteins, merging the unique reactivity of synthetic catalysts with the selectivity and mild reaction conditions provided by proteins. While useful, current techniques for the introduction of synthetic cofactors into proteins are labor-intensive and often performed in vitro. These limitations mitigate high-throughput methods for catalyst optimization, including directed evolution. My dissertation research introduces a novel method for expanding the chemical toolbox of enzymes through the development of engineered orthogonal enzyme/cofactor pairs. These orthogonal pairs consist of an unnatural heme derivative and mutated enzyme that interact independently of the cell’s native heme and heme enzymes, facilitating cofactor substitution in cells. We have successfully evolved a model cytochrome P450 enzyme to selectively bind the heme analog iron(III) deuteroporphyrin IX within cells, even in the presence of the native heme cofactor. This success has guided our work in exploring more complex cofactor structures, including adaptation of this system to generate an orthogonal iridium-substituted enzyme with expanded reactivity in non-natural carbenoid and nitrenoid transfer reactions. I have also synthesized unique phenyl-modified hemes that display improved activity in carbenoid transfer reactions when compared to the native cofactor. Phenyl hemes bind to cytochrome P450 enzymes, enabling the engineering of selective variants. Furthermore, we have designed a novel expanded heme-like structure that
accommodates larger metal ions that are useful for the development of highly sensitive, protein-based contrast agents for magnetic resonance imaging. This work shows that orthogonal enzyme/cofactors pairs may provide a general methodology for increasing the chemical diversity available to proteins. This strategy will increase the utility of enzymes in synthesis and beyond by introducing new activity via synthetic cofactors while maintaining the ability to optimize the protein through mutation and selection.
AKNOWLEDGEMENTS

The old proverb, “It takes a village to raise a child,” applies particularly well to my scientific journey. This dissertation, the culmination of my study of chemistry, would not have been possible without the support of my “village” of mentors, friends and family. To everyone who has impacted me along this journey, I am truly grateful.

First, I want to thank all of my mentors for their investment in my scientific development. I especially thank Dr. Eric Brustad for the opportunity to work in his lab during my doctoral studies. Eric’s drive and passion for his research initially inspired me to join his lab and has motivated me throughout the duration of my doctoral work. I am grateful for the time Eric has dedicated over these past 5 years to discussing research, editing grants and papers, giving career advice, and providing encouraging words. I could not have imagined having a better advisor and mentor for my Ph.D. study. I also want to thank my committee members, Dr. Dorothy Erie, Dr. Brian Kuhlman, Dr. Simon Meek and Dr. Matthew Redinbo, who have all provided advice that was instrumental to building these successful projects. In addition, I owe a great deal of thanks to our collaborators, Dr. Kathy Matera (Elon) and Dr. Christopher Snow (Colorado State), who have not only provided materials and expertise, but who have also patiently worked with me to broaden my skill set. I would also like to thank my undergraduate research advisor, Dr. Jim Demas, for his dedication to creating a first-class undergraduate research experience for the students in his lab. Dr. Demas provided me with innumerable opportunities to gain experience in research and in teaching and I value his friendship to this day. Finally, I would like to thank my high school chemistry teacher, Mrs. Ewart, who first inspired me to pursue the study of chemistry in college and set me on the path towards a Ph.D.

I also want to thank my friends and coworkers from my time here at UNC. I am
especially grateful for the camaraderie of current and past graduate student members of the Brustad Lab: Josh, Stef, Tim, Amy, Mack, Richard, and Adrienne. I owe a special thanks to my undergraduate colleagues, Fabien, Luke and Matt, for their hard work and dedication. I hope that you all have enjoyed my mentorship as much as I have enjoyed mentoring you and I am excited to see your accomplishments in the future. Finally, I thank my friends and colleagues in SACNAS and Alpha Chi Sigma for their friendship, encouragement, and the opportunity to spend some time working to help people other than myself.

Outside of UNC, I am grateful for the immense support I have received from my family and friends. I thank my mom and dad for, among many things, teaching me the value of hard work and instilling in me a passion for learning. I also thank my sister Stephanie, for her lifelong friendship and for giving me someone to compete against (even though Steph usually wins, I like to say I taught her everything she knows). I thank my wonderful wife Rebecca, for putting up with me despite myself, for never wavering in her love and support of me, for being someone I admire and aspire to emulate and for being my best friend. I also thank my family-in-law, Mr. and Mrs. Babski, Meredith, Emily, and Patrick, for their acceptance of me and for their help and encouragement throughout. Thank you to my friends, Sanjay, Tuong, Peter, Stephanie, Andrew, Aazim, Robin, Brad, Jessica and William, among many others, for all the memories and laughs we have shared together. I love you all.

There are countless other friends, family members, coworkers, professors, etc. whom I have not mentioned but who have all provided me with invaluable support along the way. To all of you, I say thank you for making this possible.

“Tonight we’ll stand, get off our knees
Fight for what we’ve worked for all these years
And the battle was long, it’s the fight of our lives
But we’ll stand up champions tonight.”

--Taylor Swift, “Change”
# TABLE OF CONTENTS

LIST OF TABLES.................................................................................................................. ix

LIST OF FIGURES .................................................................................................................. x

LIST OF ABBREVIATIONS........................................................................................................ xiii

CHAPTER 1: EXPANDING PROTEIN CHEMICAL FUNCTIONALITY WITH NON-NATURAL COFACTORS .................................................................................................................. 1

1.1 Introduction....................................................................................................................... 1

1.2 *In vitro* methods of non-natural cofactor introduction into proteins ............................. 3

1.3 *In vivo* methods of non-natural cofactor introduction into proteins ............................... 11

1.4 Conclusions.................................................................................................................... 13

REFERENCES...................................................................................................................... 15

CHAPTER 2: AN EVOLVED ORTHOGONAL ENZYME/COFACTOR PAIR ................................. 20

2.1 Introduction....................................................................................................................... 20

2.2 Results and Discussion .................................................................................................. 22

2.3 Conclusions.................................................................................................................... 36

2.4 Materials and Methods .................................................................................................. 38

2.5 Supplementary Information............................................................................................. 46

REFERENCES...................................................................................................................... 59

CHAPTER 3: ORTHOGONAL EXPRESSION OF AN ARTIFICIAL METALLOENZYME FOR ABIOLGICAL CATALYSIS ............................................................................................... 63

3.1 Introduction....................................................................................................................... 63

3.2 Results and Discussion .................................................................................................. 65

3.3 Conclusions and Future Work.......................................................................................... 72

3.4 Materials and Methods .................................................................................................. 74
3.5 Supplementary Information........................................................................................................81
REFERENCES..................................................................................................................................88

CHAPTER 4: TUNING HEME ENZYME CHEMISTRY USING ORTHOGONAL PORPHYRIN SCAFFOLDS..........................................................91
4.1 Introduction................................................................................................................................91
4.2 Results and Discussion .............................................................................................................93
4.3 Conclusions and Future Work.................................................................................................102
4.4 Materials and Methods ........................................................................................................104
4.5 Supplementary Information..................................................................................................117
REFERENCES..............................................................................................................................129

CHAPTER 5: GADOLINIUM-SUBSTITUTED PROTEINS AS HIGHLY SENSITIVE MRI BIOSENSORS.................................................................134
5.1 Introduction............................................................................................................................134
5.2 Results and Discussion........................................................................................................136
5.3 Conclusions and Future Work...............................................................................................141
5.4 Materials and Methods .......................................................................................................142
5.5 Supplementary Information................................................................................................148
REFERENCES..............................................................................................................................155
LIST OF TABLES

Table 2.1. Characterization of BM3h variants, including mutations, selectivity in rich media and minimal media, thermostability and yield. .................................................................28

Table 2.2. Activities and stereoselectivities of biocatalysts for the reaction of styrene with ethyl diazoacetate. ......................................................................................................................36

Table 3.1. Selectivity of BM3h variants for Ir(Me)-DPIX vs. heme in minimal media and rich media. .................................................................................................................................68

Table 4.1. Reaction optimization for cross-metathesis of metalated protoporphyrin IX dimethyl ester with styrene derivatives. .........................................................................................................................94

Table S2.1. Primers used in this study. .....................................................................................................................................................................................................................................................46

Table S2.2. Mutations present in hits from first library. ..................................................................................................................................................................................................................................47

Table S2.3. Cofactor selectivity of variants used for catalysis in minimal media and rich media.................................................................................................................................................................................................................................47

Table S2.4. Gradient used for HPLC analysis of purified proteins. ..............................................................................................................................................................................................................................47

Table S2.5. Data collection and refinement statistics for crystals. ..............................................................................................................................................................................................................................48

Table S3.1. Amino acid mutations present in BM3h variants used in this study. ...........................................................................................................................................................................................................................81

Table S3.2. UV-Vis absorbance of purified BM3h cofactor-substituted variants. ..........................................................................................................................................................................................................................82

Table S3.3. Gradient used for HPLC analysis of purified proteins. ..........................................................................................................................................................................................................................82

Table S4.1. Naming of synthetic porphyrins used in this study. ..............................................................................................................................................................................................................................117

Table S4.2. Mutations present in hits from from first library..............................................................................................................................................................................................................................117
LIST OF FIGURES

Figure 1.1. Supramolecular anchoring of transition metal complexes mediated by the biotin-(streptavidin) interaction ................................................................. 4
Figure 1.2. Artificial metalloenzymes generated via covalent anchoring ......................... 7
Figure 1.3. Dative anchoring of non-natural cofactors in hemoprotein scaffolds ................ 10
Figure 1.4. In vivo methods of non-natural cofactor introduction .................................. 13
Figure 2.1. Orthogonal enzyme/heme pairs: concept and screening ............................... 23
Figure 2.2. Engineering BM3h for selective recognition of Fe-DPIX ............................. 26
Figure 2.3. Selectivity of BM3h variants for Fe-DPIX .................................................. 29
Figure 2.4. Molecular determinants of Fe-DPIX selectivity .......................................... 32
Figure 3.1. Generation of an orthogonal metal-substituted enzyme ................................. 65
Figure 3.2. HPLC assay for metalloporphyrin quantitation ......................................... 68
Figure 3.3. Cyclopropanation of 1-octene with EDA catalyzed by WIVS-FM* T268A C400G/Ir(Me)-DPIX ................................................................. 70
Figure 3.4. Aziridination of 1-octene with tosyl azide catalyzed by WIVS-FM* T268A C400G/Ir(Me)-DPIX ................................................................. 71
Figure 3.5. Additional olefinic substrates to be tested in WIVS-FM* T268A C400G/Ir(Me)-DPIX-catalyzed cyclopropanation and aziridination ................................. 72
Figure 4.1. Aryl-substituted heme derivatives used in this study ..................................... 93
Figure 4.2. Absorbance spectra of phenyl heme derivatives ....................................... 96
Figure 4.3. Carbenoid-mediated olefin cyclopropanation catalyzed by phenyl heme derivatives ........................................................................................................ 98
Figure 4.4. Incorporation of MPH into hemoproteins ................................................. 99
Figure 4.6. Library design for improving MPH selectivity ............................................. 101
Figure 5.1. Design of a Gd-texaphyrin cofactor ......................................................... 136
Figure 5.2. Synthesis of Gd-Tex dimethyl ester ......................................................... 138
Figure 5.3. Absorbance spectra of Gd-Tex dimethyl ester and the base hydrolysis product ................................................................. 139
Figure 5.4. HPLC analysis of lipase-catalyzed hydrolysis of Gd-Tex dimethyl ester..........................................................140

Figure S2.1. Ferrous-CO bound absorbance spectra of BM3h with Fe-DPIX.........................49
Figure S2.2. HPLC assay for metalloporphyrin quantitation......................................................49
Figure S2.3. Heme vs. Fe-DPIX content in BM3h variants.........................................................50
Figure S2.4. Ferrous-CO bound spectra of BM3h variants expressed in rich media..........................................................51
Figure S2.5. Overall structure of BM3h variants........................................................................52
Figure S2.6. Conformation of the Fe-DPIX porphyrin ring in WIVS and WIVSFM................53
Figure S2.7. I-helix backbone rearrangements in evolved BM3h variants...............................54
Figure S2.8. Volume of the cofactor binding pocket in BM3h variants.....................................55
Figure S2.9. Determination of the ratio of cofactor-bound protein to total protein for BM3h variants..............................................................................................................56
Figure S2.10. Chiral GC analysis of cyclopropanation reactions...................................................57
Figure S2.11. 12-pNCA monooxygenation by P450BM3 variants................................................58
Figure S3.1. Absorbance spectra of Ir(Me)-DPIX-containing proteins........................................83
Figure S3.2. HPLC analysis of Ir(Me)-DPIX content in protein preparations..........................84
Figure S3.3. GC-MS analysis of cyclopropanation reaction of 1-octene with EDA catalyzed by WIVS-FM* T268A C400G/Ir(Me)-DPIX.................................................................85
Figure S3.4. 1H NMR spectrum of Ir(CO)Cl-deuteroporphyrin IX dimethyl ester......................86
Figure S3.5. 1H NMR spectrum of Ir(Me)-deuteroporphyrin IX...................................................87
Figure S4.1. Analysis of cross-metathesis reactions between protoporphyrin IX derivatives and styrene....................................................................................................................118
Figure S4.2. Incorporation of MPH and BPH into proteins in vivo.........................................119
Figure S4.3. Absorbance spectra of hits from library screening with Zn-MPH........................120
Figure S4.4. 1H NMR spectrum of MPH..................................................................................121
Figure S4.5. 1H NMR spectrum of BPH..................................................................................122
Figure S4.6. $^1$H NMR spectrum of OMe-MPH. ................................................................. 123
Figure S4.7. $^1$H NMR spectrum of CF$_3$-MPH. ................................................................. 124
Figure S4.8. $^{19}$F NMR spectrum of CF$_3$-MPH. ................................................................. 125
Figure S4.10. $^1$H NMR spectrum of Zn-MPH. ................................................................. 127
Figure S4.11. $^1$H NMR spectrum of Zn-BPH. ................................................................. 128
Figure S5.1. LC-MS analysis of Gd-Tex dimethyl ester 9. ....................................................... 148
Figure S5.2. LC-MS analysis of base-catalyzed Gd-Tex dimethyl ester hydrolysis reaction. ................................................................. 149
Figure S5.3. LC-MS analysis of lipase-catalyzed Gd-Tex dimethyl ester hydrolysis reaction. ................................................................. 150
Figure S5.4. Test incorporation of Gd-Tex into BM3h variants. .................................................. 151
Figure S5.5. $^1$H NMR of dibenzyloxytripyrane 6. ................................................................. 152
Figure S5.6. $^1$H NMR of diformyltripyrane 7. ................................................................. 153
Figure S5.7. $^1$H NMR of unmetalated texaphyrin macrocycle 8. ............................................. 154
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-pNCA</td>
<td>12-(4-nitrophenoxo)dodecanoic acid</td>
</tr>
<tr>
<td>Å</td>
<td>angstrom</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BM3h</td>
<td>heme domain of cytochrome P450&lt;sub&gt;BM3&lt;/sub&gt;</td>
</tr>
<tr>
<td>BPH</td>
<td>bisphenyl heme</td>
</tr>
<tr>
<td>CF&lt;sub&gt;3&lt;/sub&gt;-MPH</td>
<td>mono(4-trifluoromethyl)phenyl heme</td>
</tr>
<tr>
<td>CO</td>
<td>carbon monoxide</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DPIX</td>
<td>deuteroporphyrin IX</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDA</td>
<td>ethyl diazoacetate</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>e.e.</td>
<td>enantiomeric excess</td>
</tr>
<tr>
<td>F&lt;sub&gt;5&lt;/sub&gt;-MPH</td>
<td>mono(pentafluoro)phenyl heme</td>
</tr>
<tr>
<td>FM</td>
<td>BM3h G265F/F405M</td>
</tr>
<tr>
<td>GC</td>
<td>gas chromatography</td>
</tr>
<tr>
<td>Gd-Tex</td>
<td>gadolinium-texaphyrin</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrogen chloride</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>K&lt;sub&gt;A&lt;/sub&gt;</td>
<td>equilibrium association constant</td>
</tr>
<tr>
<td>k&lt;sub&gt;cat&lt;/sub&gt;</td>
<td>catalytic rate constant</td>
</tr>
<tr>
<td>K&lt;sub&gt;M&lt;/sub&gt;</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>L</td>
<td>liter</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>LmrR</td>
<td>Lactococcal multidrug resistance Regulator</td>
</tr>
<tr>
<td>m</td>
<td>meter or molality</td>
</tr>
<tr>
<td>M</td>
<td>molarity or metal</td>
</tr>
<tr>
<td>Mb</td>
<td>myoglobin</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>MPH</td>
<td>monophenyl heme</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NADPH</td>
<td>reduced nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>OMe-MPH</td>
<td>mono(4-methoxy)phenyl heme</td>
</tr>
<tr>
<td>P450&lt;sub&gt;BM3&lt;/sub&gt;</td>
<td>cytochrome P450 from <em>Bacillus megaterium</em></td>
</tr>
<tr>
<td>POP</td>
<td>prolyloligopeptidase</td>
</tr>
<tr>
<td>PPIX</td>
<td>protoporphyrin IX</td>
</tr>
<tr>
<td>r₁</td>
<td>longitudinal relaxivity</td>
</tr>
<tr>
<td>r₂</td>
<td>transverse relaxivity</td>
</tr>
<tr>
<td>RMSD</td>
<td>root-mean-square deviation</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>S-FM</td>
<td>BM3h G265F/F405M/A406S</td>
</tr>
<tr>
<td>T&lt;sub&gt;1&lt;/sub&gt;</td>
<td>longitudinal relaxation time</td>
</tr>
<tr>
<td>T&lt;sub&gt;2&lt;/sub&gt;</td>
<td>transverse relaxation time</td>
</tr>
<tr>
<td>T₅₀</td>
<td>half-denaturation temperature</td>
</tr>
<tr>
<td>TB</td>
<td>Terrific Broth</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>WIV</td>
<td>BM3h T269V/L272W/L322I</td>
</tr>
<tr>
<td>WIVS</td>
<td>BM3h T269V/L272W/L322I/A406S</td>
</tr>
<tr>
<td>WIVS-FM</td>
<td>BM3h G265F/T269V/L272W/L322I/F405M/A406S</td>
</tr>
<tr>
<td>WIVS-FM-IV</td>
<td>BM3h L52I/G265F/T269V/L272W/L322I/I366V/L405M/A406S</td>
</tr>
<tr>
<td>WIVS-FM*</td>
<td>WIVS-FM IV (see above)</td>
</tr>
</tbody>
</table>
CHAPTER 1: EXPANDING PROTEIN CHEMICAL FUNCTIONALITY WITH NON-NATURAL COFACTORS

1.1 Introduction

Nature uses proteins to achieve a large number of functions that support life. Catalytically active proteins, or enzymes, have evolved to perform under mild, aqueous conditions and catalyze reactions with high rates and selectivity. These features of enzymes have made biocatalysis an attractive option in chemical synthesis. With advances in molecular biology and protein engineering, particularly the development of techniques for systematic optimization of protein properties through iterative rounds of mutagenesis and screening (i.e., directed evolution(1)), proteins can be tailored for applications outside of their natural role. Traditionally, directed evolution efforts have focused on expanding substrate scope(2–4), altering reaction selectivity(5–7), or adapting enzymes for compatibility under harsh reaction conditions (e.g., increasing organic solvent tolerance(8) or thermostability(9)). While directed evolution is a powerful method for expanding protein function, many valuable synthetic transformations have no natural counterpart. Therefore, new strategies are needed to expand the scope of biocatalysis to non-natural reactions and thus render enzymes more broadly useful in chemical synthesis.

Synthetic chemists frequently draw inspiration from the reactive groups found in nature to design their own catalysts. Through chemical ingenuity, these synthetic catalysts have been optimized to provide access to a large number of reactions (including many that are not biologically accessible) and to operate on a broad substrate scope. Recently, enzyme engineers have sought to reverse this paradigm, looking towards synthetic chemistry to expand biocatalytic reaction space(10). Accordingly, one approach to extending the breadth of
enzymatic catalysis is to search for chemical transformations that are mechanistically similar to biological reactions. While enzymes in nature may not have been evolved for these synthetic transformations, shared mechanistic determinants often provide fortuitous, albeit weak, levels of biocatalysis for the new reaction (a concept called promiscuity). Examples of promiscuous non-natural enzymatic activities include amine-catalyzed Michael additions\(^\text{(11)}\), benzoin condensations catalyzed by thiamine diphosphate-dependent enzymes\(^\text{(12)}\), and carbenoid and nitrenoid insertion reactions catalyzed by heme enzymes\(^\text{(13)}\). Once discovered, the promiscuous enzymatic activity can be optimized by directed evolution \(^\text{(10--12)}\). Alternatively, rather than searching for promiscuous native enzymes, computational design can be used to generate pre-optimized active sites for non-natural transformations \(^\text{(14, 15)}\). These computationally designed enzymes can also be further improved through laboratory evolution \(^\text{(16--19)}\), and can possess activity in reactions beyond the original design goal \(^\text{(20, 21)}\).

However, many synthetic transformations do not share mechanistic similarities with natural reactions and thus remain inaccessible to enzymes via catalytic promiscuity or computational design strategies that rely on the reactive chemical functionality that nature provides.

To address this limitation, the introduction of non-natural cofactors into proteins has emerged as a complementary approach to broaden the scope of biocatalysis. In Nature, the chemical diversity of the amino acids is supplemented by post-translational modifications and cofactors\(^1\) to expand protein function. Researchers have mimicked this strategy to access abiological activity by generating hybrid systems in which synthetic small molecules catalysts are utilized by proteins as non-natural cofactors. Beginning with the seminal studies of Kaiser\(^\text{(22, 23)}\) and Whitesides\(^\text{(24)}\) in the late-70s, much of this work has focused on the introduction of catalytically active transition metal complexes into proteins in order to generate

---

\(^1\)Accessory molecules that bind proteins and add function are grouped into several subclasses (i.e., prosthetic groups, coenzymes and cofactors) based on binding requirements and regeneration. To facilitate writing and to be consistent with literature, we generically classify all of these accessory molecules as cofactors and non-natural cofactors.
artificial metalloenzymes. In this way, the unique reactivity of synthetic catalysts is combined with the selectivity and mild reaction conditions provided by protein cavities. In this chapter, I provide a historical perspective on the use of non-natural cofactors to expand protein chemical functionality. Recent examples of the construction of artificial metalloenzymes through supramolecular, covalent, or dative anchoring of transition metal complexes in proteins are highlighted to demonstrate the strengths and challenges of these approaches, with a distinction between \textit{in vitro} and \textit{in vivo} methods for non-natural cofactor introduction. This provides the context for my work in establishing a new and general methodology for incorporating novel metallocofactors into evolved protein scaffolds inside of cells.

\subsection*{1.2 \textit{In vitro} methods of non-natural cofactor introduction into proteins}

\subsubsection*{1.2.1 Supramolecular anchoring}

A common strategy for generating artificial metalloenzymes is to exploit a protein’s natural ligand-binding ability to immobilize a transition metal complex within the protein’s ligand-binding cavity. While in some cases a protein can bind synthetic transition metal complexes directly\cite{25,26}, the supramolecular anchoring of transition metal complexes is most frequently mediated by the noncovalent association of the ligand biotin with avidin or streptavidin (Figure 1.1A)\cite{27}. The high affinity of (strept)avidin for biotin ($K_A \approx 10^{14} \text{ M}^{-1}$)\cite{27}, results in essentially irreversible binding of biotin-tethered transition metal complexes within the protein.

Wilson and Whitesides were the first to use this approach in 1978, synthesizing a biotinylated diphosphinerhodium(I) moiety which was immobilized in avidin\cite{24}. The authors found that the avidin-Rh complex catalyzed the hydrogenation of 2-acetamidoacrylic acid with 41\% e.e. in favor of the $S$ enantiomer, whereas the metal catalyst alone showed lower activity and provided no enantioselectivity\cite{24}. Although the artificial metalloenzyme displayed only modest selectivity, this early example established the utility of the biotin-(strept)avidin system for carrying out non-natural transition metal catalysis within a protein scaffold.
In 2003, more than two decades after the first study by Whitesides, Ward and coworkers revisited the same hydrogenation in the context of the protein streptavidin rather than avidin, which they hypothesized would provide a more favorable environment for enantioselective catalysis with the cationic metal complex due to streptavidin’s lower pI and deeper binding pocket (28). Remarkably, the simple change in host protein led to formation of the R enantiomer with 92% e.e., which was improved to 96% e.e. by introducing a single amino acid mutation, S112G, in streptavidin (28). This work demonstrated the potential to achieve highly selective catalysis with artificial metalloenzymes and opened the door to the study of other reactions.

Figure 1.1. Supramolecular anchoring of transition metal complexes mediated by the biotin-(streptavidin) interaction A) The strong binding affinity of streptavidin for biotin permits immobilization of biotinylated metal complexes (MLn) within the protein cavity. B) Examples of metal complexes that have been immobilized in streptavidin and the reactions they catalyze. C) Select results in non-natural reactions catalyzed by artificial metalloenzymes generated via supramolecular anchoring.

Since then, numerous biotinylated transition metal complexes have been anchored into (strept)avidin to create artificial metalloenzymes for diverse synthetic transformations (26, 27, 29). These include palladium-catalyzed allylic alkylation (30) and Suzuki-Miyaura cross-
coupling (31); iridium, rhodium and ruthenium-catalyzed transfer hydrogenation of ketones and imines (32, 33); ruthenium-catalyzed metathesis (34, 35); and rhodium-catalyzed benzannulation (Figure 1.1B) (36). This body of work has firmly established the biotin-streptavidin technology as a powerful and general approach for generating non-natural biocatalysts.

A key theme that has emerged through these studies is the importance of exploring mutational landscapes of the host protein to optimize catalysis. For example, it was found that the introduction of a glutamate or aspartate residue was pivotal for facilitating rhodium-catalyzed benzannulation in streptavidin, while the regio- and enantioselectivity of the reaction was further tuned by additional mutations in the protein cavity (Figure 1.1C) (36). In another notable example, single mutations to streptavidin were found to have pronounced effects on the selectivity of iridium-catalyzed asymmetric transfer hydrogenation of imine 3 (33). The selectivity in the reaction improved from 57% e.e. (\(R\)) with the wild type protein to 96% e.e. (\(R\)) with the S112A variant, while mutation of the same residue to lysine gave 78% e.e. for the opposite enantiomer (Figure 1.1C) (33). These and other examples have illustrated that the genetic optimization of artificial metalloenzymes is often desirable, if not necessary, to achieve meaningful results.

Unfortunately, genetic optimization of these hybrid biocatalysts via directed evolution has proven challenging. The transition metal catalysts are often inhibited by cellular components (e.g. glutathione). Consequently, (strept)avidin variants must often be purified prior to screening (37). Furthermore, there are technical challenges associated with obtaining sufficient amounts of the protein on a small scale and developing a high throughput screen for the reaction of interest (38). For these reasons, researchers have largely relied on rational mutagenesis, computational design, or small site saturation libraries to limit the number of variants that need to be screened (29, 39, 40). While significant achievements have been made utilizing these techniques, presumably much more could be gained from larger scale directed evolution experiments and the ability to construct these artificial metalloenzymes in cells.
progress toward this goal, Ward and coworkers have recently developed a strategy to permit the periplasmic assembly and directed evolution of an artificial metalloenzyme based on the biotin-streptavidin technology(35). This effort will be discussed more thoroughly in section 1.3 below.

1.2.2 Covalent anchoring

Direct conjugation of catalytically active transition metal complexes to protein scaffolds is another common technique to form artificial metalloenzymes. This approach necessitates the presence of uniquely reactive groups on both the protein and non-natural cofactor. Most frequently, cysteine residues in the protein are targeted for covalent linkage of non-natural cofactors, although modification at other naturally occurring amino acids has also been described(26). Kaiser pioneered this approach in 1977, alkylating the active site cysteine in the enzyme papain with brominated flavin derivatives(23). In particular, a semi-synthetic enzyme created from an α-bromoacetyl-derivatized flavin catalyzed the oxidation of dihydronicotinamides with an approximately 50-fold rate enhancement over the free cofactor-catalyzed reaction and with $k_{cat}/K_M$ values comparable to natural flavoenzymes(23). Although this example did not involve transition metal catalysis, it delineated a general methodology that could be applied towards covalent modification of proteins with non-natural metallocofactors.

An advantage of covalent labeling over the supramolecular anchoring approach is the flexibility in choosing a protein host. However, care must be taken in choosing a scaffold that possesses a large enough cavity to accommodate the non-natural cofactor while still inducing selective catalysis. For example, Reetz and coworkers covalently modified papain with Mn-salen and Rh-dipyridine complexes, however, these artificial metalloenzymes showed less than 10% e.e. in model epoxidation and hydrogenation reactions, respectively(41). In contrast, the Roelfes group showed that covalent modification of the transcription repressor protein LmrR with a Cu(II)-phenanthroline complex led to 97% e.e. in a benchmark Diels-Alder reaction and up to 84% e.e. in the hydration of azachalcones (Figure 1.2A)(42, 43). Although wild type protein scaffolds may not initially provide substantial activity or selectivity, mutagenesis of the
protein can improve these properties, as was seen with artificial metalloenzymes generated through supramolecular anchoring(26).

Figure 1.2. Artificial metalloenzymes generated via covalent anchoring. A) Covalent modification of an engineering cysteine residue in the protein LmrR (green) with a Cu-phenanthroline complex led to an active and selective catalyst for Diels-Alder reactions. B) Unnatural amino acid mutagenesis to introduce an azide-functionalized phenylalanine residue to the protein POP allowed for immobilization of a dirhodium catalyst for olefin cyclopropanation.

The choice of reactive handle to modify proteins is also important. Using natural amino acids to covalently anchor a cofactor can sometimes present challenges. These include the requirement for removal of any additional reactive residues prior to labeling, low efficiency depending on the reactivity of the targeted residue, and the inability to perform the conjugation in cells or in lysate due to the potential for high background labeling of other cellular components. To overcome some of these challenges, Lewis and coworkers have used unnatural amino acids to covalently link non-natural cofactors to proteins(44). Taking advantage of biorthogonal strain-promoted azide-alkyne cycloaddition chemistry, they have successfully created hybrid biocatalysts in a variety of protein scaffolds for reactions such as photocatalytic
sulfoxidation(45), manganese-catalyzed benzylic oxidation and olefin epoxidation(46), and rhodium-catalyzed cyclopropanation(44, 47). Notably, covalent anchoring of the dirhodium catalyst 4 in prolyloligopeptidase from *Pyrococcus furiosus* (POP) led to cyclopropanation of styrene with methyl phenylidazooacetate in up to 92% e.e. after engineering a histidine and two phenylalanine residues in the pocket *(Figure 1.2B)*(47).

As with supramolecular anchoring, covalent immobilization of transition metal complexes into proteins is low throughput. Inactivation of metal complexes in cellular extracts along with the potential for background labeling of natural amino acids has required the use of purified proteins, though the use of unnatural amino acids and biorthogonal chemistries could avoid the latter issue. Also, long reaction times and multiple steps are often required for conjugation and removal of any remaining free cofactor. These factors make covalent methods of artificial metalloenzyme construction labor-intensive, preventing large-scale directed evolution efforts and transfer of this technique *in vivo*.

### 1.2.3 Dative anchoring

Nearly half of all proteins in nature require metallocofactors to function(48). Researchers can exploit the native metallocofactor-binding motifs in these proteins to introduce non-natural cofactors. By removing the native cofactor and replacing it with a non-natural derivative, new functionality can be introduced into the protein.

This strategy has found broad usage especially with respect to proteins that bind metal ions. For example, Kazlauskas has replaced the active site zinc of carbonic anhydrase with rhodium to create an artificial metalloenzyme for regioselective alkene hydroformylation(49). Recently, a bipyridyl unnatural amino acid has also been used to datively anchor a copper ion in the protein LmrR, imparting activity in Friedel-Crafts alkylation reactions with up to 83% e.e(50). Alternatively, computational design can be used to craft *de novo* metal-binding proteins. Proteins have been designed to bind zinc ions, providing hydrolytic activity(51, 52), and diiron enzymes have been designed for oxidation reactions(53, 54). While promising, dative anchoring
of metal ions can be complicated by non-specific metal binding to proteins or heavy metal ion toxicity, limiting use of these systems \textit{in vivo}.

Another area of research that is especially relevant to the work described in this thesis is the dative anchoring of non-natural cofactors in hemoprotein scaffolds. Beginning in the late-50s, heme analogs with modifications to the porphyrin ring or metal center have been incorporated into hemoglobin and myoglobin in order to gain insights into the gas-binding properties of these proteins(55–58). In addition, efforts have been made to perform catalysis with cofactor-substituted heme enzymes(59). For example, the groups of Watanabe and Lu have explored the properties of myoglobin containing Mn- and Cr-salen complexes in sulfoxidation reactions(60, 61). Hayashi has also shown that myoglobin can be reconstituted with a Mn-porphycene complex to introduce C-H bond hydroxylation activity into the protein(62).

More recently, Hartwig and coworkers have shown that artificial heme enzymes can be engineered for abiological reactions(63–65). In this work, the authors screened a panel of myoglobin variants with different axial coordination environments and a variety of metal-substituted heme derivatives for activity in C-H insertion reactions (Figure 1.3). Whereas the native iron protein showed no activity in this reaction, an iridium protoporphyrin IX complex with a methyl axial ligand (Ir(Me)-PPIX) was found to catalyze the reaction with good conversion(63). This reaction was also explored in the context of the thermostable cytochrome P450 enzyme, CYP119, reconstituted with Ir(Me)-PPIX(64). Directed evolution of Ir(Me)-CYP119 afforded an artificial metalloenzyme that catalyzes abiological C-H insertion reactions with high selectivity and with kinetic parameters on par with natural enzymes (Figure 1.3)(64). In addition, the authors showed that the evolved Ir(Me)-CYP119 variant could be used in large scale reactions and recycled multiple times without suffering major losses in activity or selectivity(64). These remarkable results have demonstrated that it is possible to overcome practical problems
associated with artificial metalloenzymes, such as low activity or loss of function on preparative scales, to make these systems more broadly useful for chemical synthesis.

**Figure 1.3.** Dative anchoring of non-natural cofactors in hemoprotein scaffolds. This strategy typically involves *in vitro* removal of heme from the protein followed by reconstitution with a non-natural derivative. The generation of Ir(Me)-proteins for abiological C-H insertion reactions demonstrates the utility of this approach for introducing novel activity to the protein and the ability to optimize this activity to levels on par with native enzymes through directed evolution of these hybrid biocatalysts.

<table>
<thead>
<tr>
<th>Protein (Reaction)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$K_M$ (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>evolved Ir(Me)-Myoglobin (carbene insertion)</td>
<td>0.73</td>
<td>1.1</td>
</tr>
<tr>
<td>evolved Ir(Me)-CYP119 (carbene insertion)</td>
<td>45.8</td>
<td>0.17</td>
</tr>
<tr>
<td>native heme P450s (oxygen insertion)</td>
<td>312 (median)</td>
<td>0.13 (median)</td>
</tr>
</tbody>
</table>

In the majority of studies with artificial heme enzymes, researchers have relied on *in vitro* reconstitution protocols to introduce the non-natural cofactor. Most frequently, this involves partially denaturing the protein in acidic solution to release the native heme cofactor, followed by extraction of the heme into organic solvent and then reconstitution of the apoprotein with the non-natural cofactor (**Figure 1.3**) (66). This process is labor-intensive and often low yielding,
limiting throughput. Some researchers have made efforts to reduce the number of steps in this process by isolating the apoprotein directly through expression in minimal media lacking iron to reduce cellular heme production (63, 67). This approach was used by Hartwig to enable medium throughput generation of artificial metalloenzyme variants for their directed evolution studies (63). However, this requires a stable apoprotein that can be expressed in good yield, and small amounts of heme contamination identified in these preparations could cause problems for identifying and optimizing new activity if the native iron-heme protein can also catalyze the reaction of interest. Recent efforts discussed below have sought to directly produce cofactor-substituted heme enzymes in vivo, which could further facilitate directed evolution studies and promote non-natural biocatalysis within cells (68, 69).

1.3 In vivo methods of non-natural cofactor introduction into proteins

The successes discussed above highlight the potential for non-natural cofactors to expand protein chemical functionality. However, most common methods for generating artificial metalloenzymes require significant and time-consuming in vitro manipulations. This has made genetic optimization of these systems difficult and has prevented their widespread application to chemical synthesis. For these reasons, several groups have recently turned their attention to developing new methods for in cell construction of artificial metalloenzymes. These methods can reduce the labor-cost associated with non-natural cofactor introduction, facilitate directed evolution to create efficient and selective biocatalysts for abiological reactions, and open the door to integration of artificial metalloenzymes into synthetic metabolic pathways within cells.

Ward and coworkers have recently adapted the biotin-streptavidin approach to construct artificial metalloenzymes for metathesis in cells (35). A key realization in this work was that although the artificial enzyme is strongly inhibited by glutathione, it is not inhibited by the oxidized form of glutathione, glutathione disulfide, which is the primary form of glutathione in the periplasm of E. coli. Developing a strategy based off this information, it was found that a biotinylated ruthenium complex could effectively cross the outer membrane of E. coli and
associate with streptavidin localized to the periplasm to yield an active catalyst for metathesis (Figure 1.4A)(35). The ability to perform the periplasmic assembly of the catalyst-protein complex accelerated the work flow for directed evolution to improve activity. Using a model ring-closing metathesis reaction that produces a fluorescent umbelliferone molecule, 5, the authors screened over 3,000 variants to identify a streptavidin variant that catalyzed the reaction with 5-fold enhancement in cell-specific activity and which outperformed traditional metathesis catalysts in reactions with several other substrates(35). This represents the first example of an artificial metalloenzyme used for abiological catalysis in cells. Moving forward, it will be interesting to see if this methodology can be extended to other metal catalysts and if the artificial metalloenzyme can be integrated into pathways with naturally occurring enzymes, as has been demonstrated several times with biosynthetic cascade reactions in vitro(70).

Cofactor-substitution of hemoproteins has also been demonstrated in cells. In order to accomplish this, special conditions must be used to avoid contamination of the protein with native heme. In one strategy, Marletta made use of E. coli strain RP523, a heme auxotroph strain with an unidentified mutation that renders the membrane permeable to heme(68). Culturing the cells under anaerobic conditions allows for growth without heme and non-natural cofactors can be supplemented into the media and passed into the cell for incorporation into a protein as it is expressed. Using this method, high purity cofactor-substituted hemoproteins have been obtained for use as phosphorescent oxygen sensors and as biosensors in magnetic resonance imaging (MRI)(71–73). A challenge of this approach is the necessity of anaerobic conditions for cell growth, which requires specialized fermentation equipment and significantly limits yield and throughput. An alternative approach to achieving in vivo cofactor substitution of hemoproteins involves exploitation of the E. coli heme transport protein ChuA (Figure 1.4B). Growth of cells in minimal media lacking iron reduces levels of heme biosynthesis, and overexpression of ChuA permits cellular uptake of non-natural heme analogs which can then be incorporated into proteins(69). This method has been used to create a Mn-
P450 MRI biosensor and Mn- and Co-myoglobin variants for C-H amination reactions (69, 74).

This system has the advantage of not requiring anaerobic conditions, however the requirement for minimal media reduces cell growth and yield, and small but significant levels of heme contamination are still observed (69).

**Figure 1.4. In vivo methods of non-natural cofactor introduction.** A) The supramolecular anchoring of a ruthenium methasis catalysis in streptavidin localized to the periplasm permits catalysis of ring-closing metathesis reactions in vivo. B) Overexpression of the heme transport protein ChuA (orange) allows for cellular uptake of non-natural heme derivatives and incorporation into hemoproteins (purple) as they express to generate artificial hemoproteins.

1.4 Conclusions

Non-natural cofactor introduction into proteins has proven to be broadly useful for a variety of applications. Despite these successes, most common methods for generating these semi-synthetic proteins are limited to *in vitro* immobilization of non-natural cofactors. This work
is labor-intensive and mitigates optimization through directed evolution. Efforts to achieve \textit{in vivo} cofactor substitution have met with some success, though these techniques require specialized expression conditions that makes them less generalizable and limit yield and throughput.

In the work described in this dissertation, I have sought to overcome the challenges associated with current strategies for non-natural cofactor introduction through the development of orthogonal enzyme/cofactor pairs. An orthogonal enzyme/cofactor pair consists of a non-natural cofactor and mutated enzyme that interact self-sufficiently without cross-talk with native cofactors or enzymes. This permits the construction of artificial metalloenzymes within cells under native growth conditions, thus improving the efficiency in generating these systems, facilitating directed evolution and allowing integration of non-natural biocatalysts into metabolic pathways.

In particular, I have applied this methodology to cofactor-substitution of heme enzymes, due to the significant interest this has received within the artificial metalloenzyme community and the diversity of porphyrin structures and reactivities that have been developed synthetically. Chapter 2 details the development of the first example of an orthogonal enzyme/cofactor pair for generating an artificial metalloenzyme \textit{in vivo}, accomplished by evolving a model Cytochrome P450 enzyme from \textit{Bacillus megaterium} (P450\textsubscript{BM3}) to selectively recognize the non-proteinogenic cofactor, iron(III) deuteroporphyrin IX (Fe-DPIX). In Chapter 3, this orthogonal enzyme was used to selectively incorporate a metal-substituted cofactor, Ir(Me)-DPIX, to improve the protein’s activity in non-natural reactions. The work presented in Chapter 4 shows that modifications to the porphyrin scaffold can also tune catalysis, exemplified by the development of novel aryl-substituted heme analogs. Finally, in Chapter 5, the potential to generate metalloproteins for use as highly sensitive MRI biosensors is explored through the design of a gadolinium-texaphyrin cofactor. Taken as a whole, this work has established orthogonal enzyme/cofactor pairs as promising tools to expand the chemical functionality of the cell.
REFERENCES


CHAPTER 2: AN EVOLVED ORTHOGONAL ENZYME/COFACTOR PAIR

2.1 Introduction

The introduction of non-natural metallocofactors into proteins is a powerful means of expanding protein chemical functionality and has been utilized to probe protein structure and function(1–3), to create novel biosensors(4–6) and to promote non-natural biocatalysis(7–9). Several methods have been developed for the in vitro immobilization of artificial metallocofactors within proteins(7), including the supramolecular anchoring of biotinylated metal complexes into streptavidin(10, 11) and the covalent anchoring of metal complexes via strain-promoted azide-alkyne cycloadditions(12). Others have sought to repurpose existing cofactor sites within proteins. In particular, hemoproteins have been the target of a variety of cofactor substitution protocols(13–15, 3). Such substitution is commonly accomplished via acid denaturation and subsequent organic extraction to remove heme, followed by reconstitution of the protein in the presence of a non-natural metal cofactor (e.g. metallo-salens and -porphyrins, etc.)(13, 14, 3). These in vitro methods generate new metal centers that can be further tuned through mutagenesis(16–18); however, they are labor-intensive and often result in low reconstituted protein yields.

To overcome challenges associated with in vitro cofactor replacement, two complementary strategies have been developed to permit the cellular uptake and incorporation of artificial heme-like cofactors into proteins. Overexpression of a native heme-transport channel, ChuA, has enabled the import of metal-swapped protoporphyrin IX derivatives in

---

Escherichia coli(4, 19). This approach, however, is limited by competition between the synthetic cofactor and native heme, which is the preferred prosthetic group of wild type hemoproteins. Accordingly, metal-substitution via ChuA is carried out in iron-deficient minimal media to minimize heme biosynthesis, and even in these cases a small amount of heme contamination (~2-5%) is observed(4). Heme-substituted proteins have been produced with high purity (>99%) using a specialized strain of E. coli that combines a knockout in heme biosynthesis and an uncharacterized mutation that promotes membrane permeability to synthetic hemes, permitting incorporation of a variety of heme derivatives with modifications to both the metal center and/or the porphyrin scaffold(20, 5, 21, 6). Nevertheless, this method necessitates anaerobic conditions for growth, which limits yield and throughput(20).

We introduce a new approach for the construction of non-natural metallocofactor proteins via the evolution of orthogonal enzyme/heme pairs (Figure 2.1A). Orthogonal enzyme/heme pairs are comprised of a non-natural heme-like cofactor and a complementary heme-binding protein that selectively interact independent of the cell’s native heme machinery. This strategy of generating orthogonal enzyme/cofactor pairs has been applied to other systems, including nicotinamide adenine dinucleotide-(22, 23) and S-adenosyl methionine (SAM)-utilizing enzymes(24).

Here, we demonstrate the in vivo, selective incorporation of a non-proteinogenic metalloporphyrin, iron deuteroporphyrin IX (Fe-DPIX), in place of heme in a model bacterial cytochrome P450 from Bacillus megaterium (P450BM3)(25). Through directed evolution(26) of the P450BM3 heme-binding pocket, a P450-variant was isolated that selectively binds Fe-DPIX with > 99% purity in minimal media, and 95% enrichment of Fe-DPIX in nutrient-rich environments under which cells produce endogenous heme. Crystal structures of variants along the evolutionary trajectory reveal an iterative reduction in size of the cofactor-binding pocket, resulting in exclusion of heme from the enzyme while accommodating Fe-DPIX. Furthermore, the orthogonal enzyme/cofactor pair shows non-natural catalytic activity in carbenoid-mediated
cyclopropanation, highlighting the potential of this system to expand biocatalytic reaction space.

We anticipate that orthogonal enzyme/cofactor pairs may provide a general strategy for increasing cofactor diversity and expanding protein chemical functionality in cells.

2.2 Results and Discussion

2.2.1 BM3h/Fe-DPIX as a model orthogonal enzyme/cofactor pair

To demonstrate the feasibility of generating orthogonal enzyme/cofactor pairs, cytochrome P450BM3 was chosen as a target scaffold. P450BM3 is a soluble, self-sufficient (i.e. consists of fused heme and P450-reductase domains) monoxygenase that has been extensively studied(25). P450BM3 has been engineered for a number of applications including altered substrate scope(27), thermostability(28), and tolerance to organic solvent(29), attesting to the flexibility of this scaffold to accommodate mutations. Furthermore, the isolated heme domain of P450BM3 (BM3h) has been the subject of several cofactor substitution studies(4, 30, 31).

We initiated cofactor replacement efforts by focusing on a commercially available metalloporphyrin, Fe-DPIX, which differs from native heme by the deletion of the two exocyclic vinyl moieties (Figure 2.1B). We hypothesized that the remaining heme-like features of Fe-DPIX, including iron-metal center, aromatic tetrapyrrole scaffold, and carboxylate side chains, would promote promiscuous uptake by ChuA and provide conserved motifs to facilitate binding within the BM3h heme pocket. Concomitantly, the “holes” created by deletion of the vinyl groups would serve as discriminating handles to accommodate mutations that enrich binding of Fe-DPIX in lieu of heme (Figure 2.1B). Similar “bump-hole” strategies have been applied to a variety of different systems, including design of ATP mimetics that selectively interact with engineered protein kinases(32), and SAM analogs for engineered histone methyltransferases(24), among others(33, 34).

2.2.2 In vivo incorporation of Fe-DPIX into wild type BM3h

To test for cellular uptake of Fe-DPIX and its accommodation within BM3h, ChuA and BM3h were co-expressed in E. coli grown in iron-deficient minimal media (to restrict cellular
Figure 2.1. Orthogonal enzyme/heme pairs: concept and screening. A) Heme transport channel, ChuA (orange), enables cellular import of non-natural heme derivatives (purple) that selectively bind to orthogonal protein scaffolds (purple) without cross-talk with endogenous heme and hemoproteins (black and blue, respectively). B) Replacing the vinyl groups of heme (blue) with hydrogens (purple) in Fe-DPIX creates a cavity that may be filled by amino acid mutations (green semi-circles) located within adjacent hydrophobic pockets-1 and -2. C) Absorbance spectra of purified ferrous CO-bound BM3h in complex with Fe-DPIX (purple) or heme (blue). D) Absorbance spectrum (solid black line) of ferrous CO-bound BM3h expressed in rich media supplemented with Fe-DPIX and heme. The relative contribution of the Fe-DPIX- and heme-bound enzyme to the composite spectrum have been approximated by a Gaussian fit centered at 434 nm (purple dotted line) and 450 nm (blue dotted line), respectively.
heme production) that was supplemented with 10 μM Fe-DPIX. Reduced CO-bound BM3h/Fe-DPIX shows a distinctive Soret band at 434 nm (Figure 2.1C) that is blue-shifted from native BM3h, which displays a Soret peak at 450 nm characteristic of cytochrome P450s(35). The hypsochromic shift observed for the Fe-DPIX Soret band is consistent with decreased conjugation around the porphyrin ring resulting from removal of the vinyl groups. BM3h/Fe-DPIX production was dependent on ChuA overexpression (Figure S2.1). This observation suggests that ChuA can provide broadened cellular access to metallocofactors that contain at least small modifications to the protoporphyrin IX scaffold.

Fe-DPIX and heme content in wild type BM3h preparations were quantified using a customized HPLC assay (Figure S2.2, Table S2.4). Under reverse-phase conditions, BM3h denatures, releasing both cofactors and allowing for separation and comparison against calibration curves constructed for each cofactor.

Using this assay, the yield of Fe-DPIX-containing protein expressed in minimal media was determined to be ~30 mg/L. Even under these stringent conditions, small levels of heme contamination (~2%) were observed (yields and mol% Fe-DPIX are reported in Table 2.1). In contrast, BM3h expressed in rich media (i.e. in the presence of endogenous heme) supplemented with ectopic Fe-DPIX yields a 62.7:37.3 mixture of Fe-DPIX and heme-containing protein, respectively (Table 2.1).

2.2.3 An absorbance-based assay for cofactor enrichment

The 16 nm Soret band shift observed in BM3h/Fe-DPIX provides a convenient spectroscopic handle to distinguish Fe-DPIX and heme content in complex mixtures, which can be used to screen for mutations that lead to changes in cofactor selectivity. When expressed in rich media (Terrific Broth, TB) supplemented with 10 μM heme and 10 μM Fe-DPIX, a mixture of Fe-DPIX- and heme-bound BM3h is observed. A Matlab script was developed to fit CO-bound absorbance spectra as Gaussian functions centered at 434 and 450 nm (Figure 2.1D) to
approximate the ratio of incorporation of Fe-DPIX with respect to heme for a particular variant. This assay is readily adaptable to high throughput analysis of BM3h-hemocompositions in lysate and on a multiwell format.

2.2.4 Evolution of an Fe-DPIX-specific protein.

Analysis of BM3h crystal structures (36, 37) reveals two distinct binding pockets, which we denote as pocket-1 and pocket-2, that accommodate the vinyl groups of heme (Figure 2.2A). These pockets are composed primarily of hydrophobic amino acids that provide a complementary surface to accommodate the porphyrin scaffold. We hypothesized that mutations that fill the cavities imparted by removal of the heme vinyl groups may provide productive binding interactions for the selective discrimination of Fe-DPIX while concomitantly introducing steric clashes that prevent heme recognition (Figure 2.1B).

We initiated our screening efforts by investigating the ability of mutations to replace the pocket-1 vinyl moiety. Since protein/Fe-DPIX interactions in this pocket are likely to involve hydrophobic and van der Waals contacts, we created a limited-diversity library randomizing active site residues to ten hydrophobic or aromatic amino acids – Gly, Ala, Val, Leu, Ile, Met, His, Phe, Tyr, and Trp – at each position, allowing us to restrict library size and in parallel maximize the number of residues available for mutation. Similar limited-diversity strategies have been used to improve the likelihood of identifying hits while reducing screening effort (38).

Mutations were focused at residues T269, L272, and L322 (Figure 2.2A, left) providing a library diversity of $10^3$ possible variants. From this library, 1,000 transformants were screened in rich media, as described above. Three variants were isolated that showed improvements in Fe-DPIX binding (Figure 2.2B and Table S2.2). Notably, all three variants converged on similar mutations including a Trp and Ile mutation at positions 272 and 322, respectively, and a small hydrophobic amino acid at position 269. The best variant, BM3h-T269V/L272W/L322I (denoted WIV), was chosen for further optimization. A CO-bound absorbance spectrum showing the increase in Fe-DPIX incorporation by WIV with respect to wild type BM3h is shown in Figure
Figure 2.2. Engineering BM3h for selective recognition of Fe-DPIX. A) Crystal structure of wild type BM3h (green, PDB: 2IJ2(36)) highlighting side chains targeted for mutation (blue) as well as the heme cofactor (yellow). B) Absorbance spectra for ferrous-CO bound BM3h variants along the evolutionary trajectory: Wild type (red), WIV (green), WIVS (blue), and WIVS-FM (purple).

2.2B. To achieve further improvements in pocket-1 Fe-DPIX selectivity, we next constructed a site-saturation library (all 20 possible amino acids substitutions) focused at position A406 using WIV as a template. An additional mutation WIV-A406S (WIVS) was identified that improved Fe-DPIX incorporation when compared to the WIV parent (Figure 2.2B). Introduction of the A406S mutation resulted in a slight blue-shift of the Soret band of the heme- and Fe-DPIX-bound protein, to 444 nm and 430 nm, respectively.
Efforts to further enhance Fe-DPIX selectivity through mutations in or around pocket-1 were unsuccessful. Accordingly, we focused subsequent libraries to pocket-2 using WIVS as a parent template. We again chose three residues – I153, G265, and F405 for randomization, each sampling the same ten amino acids as our initial library. Screening 1,000 transformants from this library yielded a single variant bearing two additional mutations G265F and F405M (WIVS-FM). Remarkably, this variant showed no detectable heme peak in the CO-bound absorbance spectrum under screening conditions (Figure 2.2B).

2.2.5 WIVS-FM is highly selective for Fe-DPIX

To more thoroughly investigate cofactor selectivity, evolved BM3h variants were expressed in the presence of 10 μM Fe-DPIX in either minimal media (low cellular heme concentrations) or rich media (TB, i.e. in the presence of endogenous heme), and the Fe-DPIX and heme content of purified BM3h preparations was quantified via HPLC (Table 2.1, Figure S2.2). Minimal media expressions provided BM3h variants that were highly enriched (> 98 mol %) for Fe-DPIX (Table 2.1 and Fig. S2.3A), and the most selective variant, WIVS-FM, showed no detectable heme binding under these conditions. On the other hand, when expressed in rich media, wild type BM3h showed considerable contamination by native heme (Fe-DPIX:heme = 62.7 mol% : 37.3 mol%; Figure 2.3A and B, Table 2.1 and Figures S2.2A and S2.3B).

Enhancements in Fe-DPIX selectivity as measured by Soret band absorbance (Figure 2.2B) were corroborated by HPLC quantitation. WIV and WIVS, which contain mutations isolated to pocket-1, showed small but iterative improvements in Fe-DPIX selectivity (71.4 mol% and 74.3 mol%, respectively; Figure 2.3B, Table 2.1 and Figures S2.2A and S2.3B). Selectivity was significantly improved in WIVS-FM, which in rich media provided protein purities comparable to those produced under conditions that restrict heme biosynthesis (Fe-DPIX:heme = 95.0 mol% : 5.0 mol%; Figure 2.3A and B, Table 2.1 and Figures S2.2A and S2.3B). The ratio of cofactor-bound- to apo-WIVS-FM was found to be consistent with levels of heme incorporation in wild type BM3h (Figure S2.9). Furthermore, expression of WIVS-FM in rich
media and in the absence of Fe-DPIX produced only small quantities of heme-bound WIVS-FM (<0.5 mg/L; Figure 2.3A), the majority of which possessed a CO-bound Soret peak at 420 nm, indicative of a misligated heme state (often called P420, Figure S2.4A)(35). These data suggest that WIVS-FM effectively restricts heme from the active site, while accommodating Fe-DPIX in an orthogonal fashion.

Table 2.1. Characterization of BM3h variants, including mutations, selectivity in rich media and minimal media, thermostability and yield. *

<table>
<thead>
<tr>
<th>Variant</th>
<th>Mutations</th>
<th>mol %a (TB media)</th>
<th>mol %a (M9 media)</th>
<th>T50 (°C)</th>
<th>M9 yield (mg/L)d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fe-DPIX heme</td>
<td>error*</td>
<td>Fe-DPIX heme</td>
<td>error*</td>
</tr>
<tr>
<td>WT</td>
<td>none</td>
<td>62.7</td>
<td>37.3</td>
<td>4.5</td>
<td>98.8</td>
</tr>
<tr>
<td>WIV</td>
<td>T269V, L272W, L322I</td>
<td>71.4</td>
<td>28.6</td>
<td>9.5</td>
<td>99.5</td>
</tr>
<tr>
<td>WIVS</td>
<td>T269V, L272W, L322I, A406S</td>
<td>74.3</td>
<td>25.7</td>
<td>3.8</td>
<td>98.3</td>
</tr>
<tr>
<td>WIVS-FM</td>
<td>G265F, T269V, L272W, L322I, F405M, A406S</td>
<td>95.0b</td>
<td>5.0b</td>
<td>1.8b</td>
<td>&gt;99.9c</td>
</tr>
<tr>
<td>FM</td>
<td>G265F, F405M</td>
<td>77.4</td>
<td>22.6</td>
<td>6.0</td>
<td>94.5</td>
</tr>
<tr>
<td>S-FM</td>
<td>G265F, F405M, A406S</td>
<td>89.5</td>
<td>10.5</td>
<td>4.0</td>
<td>99.7</td>
</tr>
<tr>
<td>WIV-FM</td>
<td>G265F, T269V, L272W, L322I, F405M</td>
<td>89.6</td>
<td>10.4</td>
<td>2.3</td>
<td>99.9</td>
</tr>
<tr>
<td>WIVS-FM-IV</td>
<td>L52I, G265F, T269V, L272W, L322I, I366V, F405M, A406S</td>
<td>84.6</td>
<td>15.4</td>
<td>2.0</td>
<td>97.6</td>
</tr>
</tbody>
</table>

*values reported are the average and standard deviations from three independent experiments unless otherwise noted
a determined by HPLC analysis of purified proteins from expressions in the denoted media
b, c average and standard deviation from b eight and c seven independent experiments
d yields were determined via HPLC analysis or from ferrous CO-bound difference spectra, using ε430nm=130 mM⁻¹cm⁻¹

2.2.6 Mutations in pocket-1 and pocket-2 provide additive improvements in Fe-DPIX selectivity

While we pursued an evolutionary trajectory that initially targeted mutations to pocket-1 (followed by pocket-2 in later rounds), efforts to affect changes in Fe-DPIX selectivity could feasibly use either cavity as a starting point. The most significant improvement in Fe-DPIX selectivity was achieved upon the addition of two mutations to pocket-2 (i.e. G265F and F405M in WIVS-FM). This observation led us to question whether improvements that result from these
**Figure 2.3. Selectivity of BM3h variants for Fe-DPIX.** A) Cell pellets of wild type BM3h and WIVS-FM after expressions in rich media in the presence or absence of Fe-DPIX. Red cell pellets are indicative of high levels of properly-folded BM3h. Yields were determined by HPLC analysis. B) Improvements in Fe-DPIX selectivity between variants along the original evolutionary trajectory (black bars and lines) and via alternative evolutionary paths (gray bars and dashed lines).

mutations were dependent on amino acid substitutions identified in prior rounds of selection, or whether pocket-2 mutations could function in isolation.

To examine the role of mutational directionality in orthogonal enzyme/cofactor pair evolution, we constructed three additional BM3h variants sampling alternative evolutionary trajectories (Figure 2.3B and Figure S2.4B). Variant BM3h-G265F/F405M (denoted FM), which contains only pocket-2 mutations, increased the mol % incorporation of Fe-DPIX to 77.4%, as compared to 62.7% in the wild type scaffold (Table 2.1). The degree of enrichment afforded by these two mutations alone (+ 14.7%) is similar in magnitude to that observed upon addition of the same mutations to WIVS (+ 20.7%) (Figure 2.3B). Grafting the G265F/F405M mutations onto the WIV scaffold (lacking the A406S mutation) to produce WIV-FM (BM3h-T269V/L272W/L322I/G265F/F405M) produced similar gains in Fe-DPIX selectivity (+ 18.1 mol%) (Figure 2.3B). These results show that pocket-1 and pocket-2 mutations, in isolation, are sufficient to provide significant improvements in Fe-DPIX binding and provide additive enhancements when combined. The functional independence of both pockets may permit the
evolution of protein scaffolds that selectively recognize more complicated cofactor structures including heme analogs baring asymmetric-substitution patterns in place of the vinyl moieties.

In contrast to other mutations, A406S showed greater context-dependent effects with respect to the parent scaffold. While the introduction of A406S into WIV (to produce WIVS) provided only a modest improvement in Fe-DPIX binding (2.9 mol %), this mutation provided a 12.1 mol % improvement in binding in the context of the FM scaffold (S-FM) (Figure 2.3B). While our library designs include this residue in pocket-1, it is situated on an active site helix that occupies a position at the interface of both cavities. These results suggest that mutations at A406 may facilitate pocket-2 remodeling in a manner independent of pocket-1.

2.2.7 Fe-DPIX selectivity increases as result of decreased stability of the heme-bound enzyme

Increased Fe-DPIX selectivity in BM3h variants may reflect improved enzyme/cofactor interactions with the Fe-DPIX scaffold, decreased affinity/stability with respect to native heme, or a combination of both. To investigate the extent to which these effects contribute to observed increases in Fe-DPIX binding, we measured the stability of purified heme- and Fe-DPIX-bound variants (Table 2.1) as a function of $T_{50}$ (the temperature at which half the protein remains folded after a 10 minute incubation). Native BM3h/heme was more stable than BM3h/Fe-DPIX by $\sim 5 \degree C$ ($T_{50, \text{heme}} = 55.5 \degree C$, $T_{50, \text{Fe-DPIX}} = 50.4 \degree C$; Table 2.1), consistent with the fact that the wild type pocket is highly optimized for heme. Mutations that confer increases in Fe-DPIX selectivity resulted in lower stability of the heme-bound protein. While significant decreases in stability relative to wild type were observed for heme-bound WIV ($T_{50, \text{heme}} = 52.7 \degree C$; Table 2.1) and WIVS ($T_{50, \text{heme}} = 49.7 \degree C$; Table 2.1), these variants maintained similar stabilities in the presence of Fe-DPIX ($T_{50, \text{Fe-DPIX}} \approx 50 \degree C$; Table 2.1). WIVS-FM/heme expressions resulted in a misfolded P420 state (Figure S2.4A), which prevented thermostability measurements. This result suggests that heme-bound WIVS-FM is further destabilized relative to other variants.

The decrease in stability observed in WIVS-FM/heme was accompanied by a slight loss in stability of WIVS-FM/Fe-DPIX ($T_{50, \text{Fe-DPIX}} = 47.1 \degree C$; Table 2.1). In attempt to improve the
stability of the WIVS-FM/Fe-DPIX pair, we introduced two mutations (L52I and I366V, to produce WIVS-FM-IV), which have previously been shown to increase the thermostability of BM3h variants (39). WIVS-FM-IV/Fe-DPIX showed increased yield and stability ($T_{50,Fe-DPIX} = 51.4 \degree C$; Table 2.1) compared to WIVS-FM/Fe-DPIX. However, these gains were accompanied by a concomitant increase in stability and yield of the heme-bound protein, resulting in a loss of Fe-DPIX selectivity (Table 2.1, Figures S2.3A, B, S2.4C). These stability trends suggest that engineered mutations provide enrichment in Fe-DPIX primarily through destabilization of the heme-bound protein.

2.2.8 Structural determinants of Fe-DPIX selectivity in evolved BM3h variants

To elucidate the molecular basis for Fe-DPIX selectivity, we determined the crystal structures of WIVS and WIVS-FM to 2.34 Å and 2.16 Å, respectively (see Table S2.4 for statistics on data collection and refinement). Each structure contains eight molecules in the unit cell. Overall, both structures closely resemble wild type BM3h (Figure S2.5), with the most significant changes occurring to side chain positions in the heme-binding pocket (36, 37). WIVS and WIVS-FM aligned to the substrate-free structure of wild type BM3h (PDB: 2IJ2, molecule A) with average rsmd values of 0.88 Å and 0.92 Å, respectively, and to the substrate-bound structure (PDB: 1JPZ, molecule A) with rmsd values of 0.40 Å and 0.43 Å, respectively.

Active site mutations of WIVS-FM appear to fill the cavities imparted by deletion of the heme-vinyl groups. The indole side chain of L272W occupies the space previously filled by the vinyl group in pocket-1 (Figure 2.4A). The side chain of L322I, which packs against L272W, may orient the tryptophan side chain to more favorably exclude heme. The T269V mutation overlays closely with the isosteric wild type side chain, and while this mutation may not contribute significantly to steric occlusion of native heme, it may increase hydrophobicity in pocket-1. In addition, this mutation interrupts a cognate hydrogen bond between the backbone carbonyl of residue 265 with the T269 side chain (Figure S2.7A) (37), which may lead to
Figure 2.4. Molecular determinants of Fe-DPIX selectivity. A) Pocket-1 and B) pocket-2 of wild type BM3h (green, 2IJ2) overlaid against WIVS-FM side chains (cyan). The Fe-DPIX cofactor of WIVS-FM is shown in yellow. C) Cartoon overlay of the I-helix of WIVS-FM (cyan) with that of substrate-free wild type BM3h (green, 2IJ2) and substrate-bound BM3h (purple, PDB: 1JPZ(37)). D) The I-helix of WIVS-FM (cyan) overlaid with the substrate-free wild type structure (green, 2IJ2), shown in stick representation. Only the side chain of residue G265F is shown for clarity. E) The cofactor-binding pocket (gray surface) of the wild type enzyme (green, 2IJ2) compared to WIVS-FM (cyan, cofactor in yellow). F) Quantified cavity volumes along the evolutionary trajectory. Error bars represent standard deviations. n=8 molecules. *p < 0.01, **p < 0.0001 ***p < 0.0000001.
observed backbone conformational changes in the I-helix, discussed below. The A406S mutation appears to improve packing with the porphyrin ring. In pocket-2, the G265F mutation fills the vinyl cavity of pocket-2 (similar to L272W in pocket-1) while F405M appears to create a flexible packing interface to accommodate the G265F mutation (Figure 2.4B).

Several unexpected features were observed in crystal structures of both WIVS and WIVS-FM. For example, the conformation of the Fe-DPIX cofactor is rotated 180° about the plane of the porphyrin with respect to the preferred conformation typically observed in heme-bound BM3h (see heme/Fe-DPIX methyl groups in Figure 2.4E and Figure S2.6A, B, and E). Modeling suggests that, in the “un-flipped” conformation, steric clashes would occur between the Fe-DPIX exocyclic-methyl groups and the side chains of A406S and G265F; however, these inhibitory interactions are alleviated in the ring-flipped conformation (Figure S2.6C and D). This alternate “ring-flipped” porphyrin conformation is observed for heme in high resolution structures of BM3h and is estimated to make up less than 20% of the population in the wild type enzyme(37). The ability of the cofactor to occupy different conformations may allow some variants to maintain heme binding by circumventing steric clashes between the heme vinyl moieties and the introduced amino acid mutations (e.g. clashes between the heme vinyl group in pocket-1 and the L272W side chain are avoided in the “ring-flipped” conformation) (Figure S2.6E).

The structures of WIVS and WIVS-FM also show unexpected rearrangements in the backbone conformation of the BM3h I-helix (a long, conserved helix that spans the length of the P450 active site and is important for catalysis). Substrate-free structures of wild type BM3h show a significant kink in the middle of the I-helix due to an interruption of the i-i+4 hydrogen bonding pattern by a conserved water molecule, which inserts itself between the carbonyl of I263 and the amide nitrogen of E267(36). In substrate-bound structures, this water molecule is displaced resulting in a straightening of the I-helix(36). In both WIVS and WIVS-FM, which were crystallized in the absence of substrate, we do not observe the typical water molecule between
I263 and E267 and the I-helix adopts an uninked helical backbone conformation that more closely resembles the substrate-bound form of the protein (Figure 2.4C and Figure S2.7A and B). Notably, the I-helix becomes progressively less distorted along the evolutionary trajectory.

Previous structural studies of P450<sub>BM3</sub> variants have noted that mutations in the I-helix can lock the enzyme in the substrate-bound conformation even in the absence of substrate (37, 40, 41). Unbending of the I-helix observed in WIVS may result from the T269V mutation, which interrupts a cognate backbone carbonyl-side chain hydrogen bond between G265 and the side chain of T269 (Figure S2.7A). Further straightening of the helix is observed in WIVS-FM with the addition of the G265F and F405M mutation (Figure S2.7A and S2.7C). Notably, G265F, resides in the I-helix; the removal of a helix-breaking glycine residue at this location may contribute significantly to increased helicity. As a result of this backbone conformational change, the Cα atom of G265F (in WIVS-FM) shifts 1.6 Å with respect to the wild type structure (Figure 2.4D). This movement helps prevent steric clashes between the phenylalanine side chain and the Fe-DPIX porphyrin ring (Figure 2.4D). This observation suggests that while the I-helix is well situated to interact with native heme, backbone conformational rearrangements may be important to accommodate mutations that enrich for artificial porphyrin binding.

In general, we observed a significant decrease in the size of the cofactor binding pocket of engineered BM3h variants when compared to the wild type enzyme (Figure 2.4E and Figure S2.8A). To quantify these differences, we calculated changes in solvent-accessible pocket volume in the immediate vicinity of the cofactor (Figure 2.4F and Figure S2.8B). To provide a statistical analysis, we computed active site volumes in all 8 monomers of the WIVS and WIVS-FM structures, as well as 8 molecules of substrate-free wild type BM3h from the structures 2IJ2, 1BU7, 2HPD and 2BMH (37, 42–44). The cofactor binding pocket volume shrinks substantially along the evolutionary trajectory (wild type = 212 Å³, WIVS = 181 Å³, WIVS-FM = 161 Å³; Figure 2.4F), providing a complementary binding surface for Fe-DPIX at the exclusion of the larger heme porphyrin.
2.2.9 WIVS-FM/Fe-DPIX catalyzes non-natural carbenoid-mediated olefin cyclopropanation

A major goal of this work is to enable the generation of new catalysts for reactions not found in nature. To demonstrate the potential for using orthogonal enzyme/heme pairs as catalytic scaffolds, we examined the activity of WIVS-FM/Fe-DPIX in non-natural carbenoid-mediated olefin cyclopropanation. Recent work has shown that heme enzymes can be engineered as highly active and selective catalysts for this reaction (45–48). In particular, the mutation T268A in BM3h has been shown to be highly activating towards cyclopropanation (45). Therefore, we introduced the T268A mutation into the WIVS-FM scaffold. Addition of this mutation did not affect selective Fe-DPIX recognition (Table S2.3). We then tested the *in vitro* activity of WIVS-FM T268A/Fe-DPIX in the model cyclopropanation reaction between styrene and ethyl diazoacetate (Table 2.2 and Figure S2.10). WIVS-FM T268A/Fe-DPIX shows levels of cyclopropanation activity near that of BM3h T268A/heme, although the selectivity of the T268A variant is altered in the context of WIVS-FM compared to the wild type protein. The observed changes in selectivity may be a result of altered substrate binding conformations caused by the observed backbone changes in the I-helix (described above).

Orthogonal P450-FeDPIX scaffolds were inactive in native P450-mediated monooxygenation reactions (Figure S2.11). As such, engineered FeDPIX variants provide a new enzyme scaffold that is exclusively competent for non-natural cyclopropanation catalysis, without competition with native P450 chemistry. We hypothesize that the loss of monooxygenation catalysis results from significant structural rearrangements in the I-helix, which has been shown to play a pivotal role in proton transfer events that are necessary for the P450 catalytic cycle (36).

The observation that WIVS-FM is no longer active in monooxygenation reactions led us to also investigate cyclopropanation under aerobic conditions, which are typically inhibitory towards P450-mediated carbene insertion reactions. WIVS-FM T268A/Fe-DPIX maintained a greater proportion of its activity and selectivity under aerobic conditions when compared to
Table 2.2. Activities and stereoselectivities of biocatalysts for the reaction of styrene with ethyl diazoacetate.

![Chemical reaction diagram]

<table>
<thead>
<tr>
<th>Catalyst</th>
<th>Yield</th>
<th>TTN&lt;sup&gt;a&lt;/sup&gt;</th>
<th>dr (cis:trans)</th>
<th>ee cis [%]&lt;sup&gt;b&lt;/sup&gt;</th>
<th>ee trans [%]&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>heme</td>
<td>7</td>
<td>35</td>
<td>16:84</td>
<td>0</td>
<td>-11</td>
</tr>
<tr>
<td>Fe-DPIX</td>
<td>5</td>
<td>23</td>
<td>13:87</td>
<td>3</td>
<td>-11</td>
</tr>
<tr>
<td>BM3h T268A/heme</td>
<td>57</td>
<td>284</td>
<td>1:99</td>
<td>-6</td>
<td>-97</td>
</tr>
<tr>
<td>BM3h T268A/heme aerobic</td>
<td>2</td>
<td>11</td>
<td>10:90</td>
<td>-15</td>
<td>-42</td>
</tr>
<tr>
<td>WIVS-FM T268A/Fe-DPIX</td>
<td>44</td>
<td>221</td>
<td>12:88</td>
<td>-51</td>
<td>-46</td>
</tr>
<tr>
<td>WIVS-FM T268A/Fe-DPIX aerobic</td>
<td>12</td>
<td>59</td>
<td>12:88</td>
<td>-33</td>
<td>-37</td>
</tr>
</tbody>
</table>

<sup>a</sup> TTN = total turnover number.  
<sup>b</sup> (1<sup>R</sup>, 2<sup>S</sup>) – (1<sup>S</sup>, 2<sup>R</sup>).  
<sup>c</sup> (1<sup>R</sup>, 2<sup>R</sup>) – (1<sup>S</sup>, 2<sup>S</sup>). TTN and stereoselectivities determined by chiral GC analysis.

BM3h T268A/heme, further illustrating the potential of this orthogonal system as a starting point towards generating novel biocatalysts. Notably, the activity of the enzyme/cofactor pair correlates with the activity of the free cofactor. This allows for potential orthogonal cofactors to be screened for a desired reaction activity prior to engineering the protein for selective binding. In this way, orthogonal enzyme/heme pairs allow for optimization of new chemistry via “evolution” of both the cofactor and protein.

2.3 Conclusions

We report the design and evolution of an orthogonal cytochrome P450 scaffold that shows altered specificity towards a non-natural metallocofactor, Fe-DPIX. WIVS-FM was isolated after only three rounds of directed evolution, sampling ~2,100 total transformants, suggesting that limited-diversity randomization of the heme-binding cavity can produce rapid enhancements in selectivity for non-natural cofactors. Structural investigations suggests that
isolated mutations exclude native heme through iterative reduction in size of the cofactor binding cavity, providing selective recognition of the smaller deuteroporphyrin scaffold. While WIVS-FM/Fe-DPIX no longer catalyzes native P450 monooxygenation reactions, it shows significant activity in non-natural cyclopropanation reactions, providing a new starting point for the development of non-natural biocatalysts.

We anticipate that this approach can be expanded to a range of synthetic porphyrin and non-porphyrin systems bearing alternate functionality that enables discrimination from native-heme and in parallel that tune the electronic properties of the metal center. In cases where the cofactor structure is significantly divergent from native heme, the scope of promiscuous transport by ChuA will need to be addressed. Evolution of ChuA prior to, or in parallel with, evolution the orthogonal enzyme may be required for more complex scaffolds. Additionally, the very small levels of heme contamination observed for orthogonal enzyme preparations in rich media may complicate screens for novel catalytic activity, as care will have to be taken to ensure observed activity is not due to contaminating heme enzyme. However, this may not be an issue in screening for reactions that heme itself cannot catalyze.

This work provides a strategy for reshaping heme-binding pockets to recognize designed, artificial metallocofactors. For example, the ability of evolved Fe-DPIX-binding proteins to produce non-iron proteins is currently being investigated. As was recently shown, alternative metal centers have the potential to greatly expand the reaction scope of heme enzymes(49) and thus is a promising direction for the generation of highly-active orthogonal enzyme/heme pairs in reactions that cannot be catalyzed by native heme. In addition, the design of a panel of orthogonal proteins that selectively recognize unique cofactor scaffolds will allow for the expression of tailored metal-substituted proteins in cells. These engineered metalloproteins may then be integrated into biosynthetic pathways or used as probes of protein function. Accordingly, orthogonal enzyme/cofactor pairs represent versatile tools to expand nature’s chemical repertoire.
2.4 Materials and Methods

2.4.1 Materials

Iron(III) deuteroporphyrin IX chloride and hemin chloride were obtained from Frontier Scientific. Enzymes and reagents used for cloning were obtained from New England BioLabs. Oligonucleotides were obtained from Integrated DNA Technologies. All other chemical reagents were obtained from chemical suppliers (Acros, Fisher Scientific or Sigma-Aldrich) and used without further purification.

2.4.2 Plasmids and library construction

Genes for BM3 variants were cloned, as previously described (50) into plasmid pCWori. This plasmid contains an IPTG inducible promoter for P450 expression, and a pBR322 origin and ampicillin resistance marker for plasmid replication and maintenance. Heme transport protein, ChuA, was expressed from plasmid pChuA (available from Addgene), which contains a T7 promoter for gene expression and a kanamycin resistance cassette for plasmid maintenance.

Hydrophobic pocket libraries were constructed through standard overlap extension PCR assembly of PCR fragments. Forward primers containing library mutations were designed using DC Analyzer (51) to include degenerate codons that encode only hydrophobic residues (G, W, F, L, I, M, H, A, V, and Y). Site-specific mutants and single site-saturation libraries were assembled via scarless restriction ligation of BsaI digested PCR fragments containing the desired mutation or an NNK codon (codes for all 20 amino acids and a TAG stop codon). See Table S2.1 for a full list of primers used in this study.

2.4.3 Protein expression

Introduction of Fe-DPIX to BM3h was accomplished in a manner similar to previously described methodology (4). BM3h variants in pCWori were co-transformed with pChuA into electrocompetent E. coli BL21-AI cells (Novagen). For expressions in minimal media, individual colonies were inoculated into liquid starter cultures of M9 minimal media supplemented with
0.5% glucose, 0.2% casamino acids, 340 µg/mL thiamine hydrochloride, 2 mM MgSO₄, 0.1 mM CaCl₂, 100 µg/mL kanamycin and 200 µg/mL ampicillin. For rich media expressions, colonies were inoculated into liquid starter cultures of LB media supplemented with 100 µg/mL kanamycin and 200 µg/mL ampicillin. The cultures were grown overnight at 37 °C and 225 rpm. One liter of M9 minimal media or TB supplemented with 100 µg/mL Kanamycin and 200 µg/mL ampicillin was inoculated with 5 mL of overnight starter culture (the M9 or LB culture, respectively) and incubated for 3-4 hours at 37 °C and 225 rpm to an OD₆₀₀nm of 0.6-0.7. The cultures were then induced by adding arabinose (to induce ChuA expression in BL21-AI) and Isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.2% and 1 mM, respectively. A concentrated stock solution of the Fe-DPIX in DMSO was added at the time of induction to a final concentration of 10 µM. Upon induction, the incubation temperature was reduced to 25 °C and expressions were allowed to continue for 16-20 hours. After harvesting the cells by centrifugation (4 °C, 10 min, 4,000 × g), the cell pellets were frozen at -20 °C for at least 2 hours.

2.4.4. Protein purification

Frozen cell pellets were thawed and resuspended in 25 mM Tris-HCl buffer pH 7.5, 300 mM NaCl, 25 mM imidazole (buffer A) and then lysed by sonication (8 min, 20% amplitude, pulse on 0.5s, pulse off 0.8s; Sonic Dismembrator Model 500, Fisher Scientific). Cell debris was pelleted by centrifugation for 30 min at 4 °C and 20,000 × g, and the supernatant was subjected to 6 × His-tag purification on a Ni-NTA HisTrap HP column (5 mL HisTrap HP, GE Healthcare, Piscataway, NJ) using an ATKA Purifier (GE Healthcare). The column was washed with buffer A until detector signal reached ≤ 60 mAu. Protein was then eluted with 55% buffer B (buffer A + 300 mM imidazole). The purified protein was exchanged into 0.1 M potassium phosphate buffer, pH 8.0, via dialysis or a 30 kDa molecular weight cut-off Amicon centrifugal filtration device (EMD Millipore). Protein aliquots were flash frozen on dry ice and stored at -20 °C.
For crystallization experiments, the His-tagged purified proteins were subjected to an additional round of purification by anion exchange. The 6 × His-tagged purified proteins were exchanged into 25 mM Tris-HCl pH 7.5, 25 mM NaCl, using a 30 kDa molecular weight cut-off Amicon centrifugal filter and loaded onto a Q sepharose column (5 mL HiTrap Q FF, GE Healthcare). Protein was eluted from the column using a linear gradient from 25 mM to 500 mM NaCl over 10 column volumes. The purified protein was exchanged into 25 mM Tris-HCl pH 7.5, 25 mM NaCl and concentrated to approximately 10 mg/mL using a 30 kDa molecular weight cut-off Amicon centrifugal filter. The protein was flash frozen on dry ice and stored at -20 °C.

**2.4.5 HPLC analysis of purified proteins for heme content**

Proteins for HPLC analysis were exchanged into distilled water via dialysis or a 30 kDa molecular weight cut-off Amicon centrifugal filter (EMD Millipore). Proteins were diluted to approximately 5 µM in distilled water and filtered through glass wool plugs prior to HPLC analysis on a 1260 Infinity binary liquid chromatography system (Agilent). Separations were achieved on a Restek Viva C4 column (5 µm, 2.1 x 150 mm) using the gradient described in Table S2.4. Calibration curves (Figure S2.2B) were generated for Fe-DPIX and heme using commercial standards to determine metalloporphyrin concentrations in the protein sample.

**2.4.6 Determination of the ratio of cofactor-bound protein to total protein**

WIVS-FM/Fe-DPIX and wild type BM3h/heme were expressed and purified as described above, from TB media supplemented with 10 µM Fe-DPIX or 10 µM heme, respectively. Purified BM3h variants were run on SDS-PAGE (Figure S2.9) to assess the purity of the protein preparations. Total protein content was then determined by measuring absorbance of pure protein samples at 280 nm, using calculated extinction coefficients(52) of $\epsilon_{280nm} = 49,860 \text{ M}^{-1}\text{cm}^{-1}$ and $44,170 \text{ M}^{-1}\text{cm}^{-1}$ for WIVS-FM and wild type, respectively. The amount of cofactor-bound protein was determined by HPLC analysis as described above, and the ratio of cofactor-bound to total protein was calculated.
2.4.7 Carbon monoxide binding assay

The carbon monoxide difference spectra for purified P450 enzymes, or P450s in lysate, were obtained as previously described in either multiwell plate format(47) or in cuvettes(35). The UV-vis data was obtained on a Tecan M1000 PRO UV/Vis plate reader or Cary 100 UV/Vis spectrometer (Agilent), respectively.

2.4.8 High-throughput screen for Fe-DPIX-selective variants

Variant libraries of BM3h in vector pCWori were co-transformed with the plasmid pChuA into BL21-AI electrocompetent cells and selected on LB Agar supplemented with 50 µg/mL kanamycin and 100 µg/mL ampicillin. Single colonies were inoculated into 96 well blocks containing 1 mL of LB supplemented with 100 µg/mL kanamycin and 200 µg/mL ampicillin. One to three wells of each block were used for controls harboring wild type BM3h or parent mutants from previous rounds of selection, as well as an uninoculated well to control for contamination. Blocks were sealed with M3 medical tape to permit aeration during cell growth. Cultures were grown to saturation overnight at 37 °C, 990 rpm and 80% humidity in a Multitron Pro 3 mm orbital incubation shaker (Infors HT). The following day, 100 µL of the cultures was mixed with 100 µL of 50% sterile glycerol in 96 well plates for storage at -80 °C. Another 50 µL of culture was used to inoculate 96 well blocks containing 1 mL of TB media supplemented with 100 µg/mL kanamycin and 200 µg/mL ampicillin. Cells were grown for 3 hours at 37 °C, 990 rpm and 80% humidity to an OD$_{600nm}$ of approximately 0.6-0.7, and then induced with 0.2% arabinose, 1 mM IPTG, 10 µM heme, and 10 µM Fe-DPIX. Proteins were expressed overnight at 25 °C, 990 rpm, and 40% humidity. Following overnight expression, cells were harvested at 4000 × g for 10 minutes. The supernatant was discarded and the pellets were frozen at -20 °C for at least 2 hours. Thawed pellets were resuspended in 400 µL of 0.1 M potassium phosphate buffer, pH 8.0, supplemented with 0.5 mg/mL lysozyme and a small amount (approximately 0.01 mg/mL) of DNaseI and incubated at 37 °C, 990 rpm for 1 hour. The lysate was cleared of cell debris by centrifugation at 4000 × g for 10 minutes. 200 µL of the clarified lysate was transferred
to microtiter plates containing 50 µL of 0.4 M sodium dithionite in 1 M potassium phosphate pH 8.0 for high-throughput CO binding experiments. The CO-difference spectra were measured with a Tecan M1000 PRO UV/Vis plate reader from 380 – 500 nm. The relative contribution of deuteroporphyrin-bound and heme-bound protein was estimated by fitting the data to a weighted sum of gaussians, using a non-linear least squares algorithm (53) in Matlab (Mathworks). We used the weighting of the Gaussians (ω) to calculate an approximate percent Fe-DPIX in the enzyme as:

\[
\% \text{FeDPIX} = \frac{\omega_{434\text{nm}}}{\omega_{434\text{nm}} + \omega_{450\text{nm}}}
\]

We then compared the % Fe-DPIX between variants and parent proteins to identify hits with improved Fe-DPIX selectivity.

2.4.9 Extinction coefficient determination

The carbon monoxide difference spectrum of purified WIVS-FM-Fe-DPIX was measured and compared with exact Fe-DPIX content determined via HPLC as described. From these data, the extinction coefficient for the Soret peak was determined to be \(\varepsilon_{430\text{nm}} = 130 \pm 15 \text{ mM}^{-1}\text{cm}^{-1} \) (n=3).

2.4.10 Thermostability

Proteins were diluted to 7 µM in 0.1 M potassium phosphate buffer, pH 8.0. 200 µL of diluted protein was heated over a range of temperatures (35 – 70 °C) for 10 minutes then cooled to 4 °C for 1 minute using an Eppendorf ProS thermocycler. The samples were centrifuged to remove precipitated protein and the concentration of folded P450 was determined by the CO-difference spectra. The temperature at which half of the protein was denatured was calculated by fitting the data to the equation (46):

\[
f(T) = \frac{100}{1 + e^{-a(\frac{T-1}{T_{50}})}}
\]
2.4.11 Protein crystallography

Proteins were crystallized by vapor diffusion in 24 well sitting drop plates (Hampton Research). A 1.5 µL:1 µL mixture of protein stock (10 mg/mL in 25 mM Tris pH 7.5, 25 mM NaCl) to mother liquor was combined in each drop. Optimal crystallization conditions for WIVS-FeDPIX and WIVSFM-FeDPIX were 0.25 M MgCl$_2$, 17% PEG3350. Crystals typically grew over a span of 10 – 15 days at room temperature.

2.4.12 X-ray data collection and protein structure determination

X-ray diffraction data were collected at Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (Argonne National Laboratory) using beamline 22-ID and a MAR300HS CCD detector. Data were collected at 100 K. Statistics for data collection and refinement are listed in Table S2. Diffraction data sets were integrated and scaled with the automated data processing software KYLIN(54) provided by SER-CAT. For WIVS-FM, initial phases were determined by molecular replacement against the wild-type BM3h structure deposited under PDB code 2IJ2(37) (chain A) using Phenix Phaser(55). The WIVS-FM model was used for molecular replacement of WIVS, maintaining the $R_{\text{free}}$ flags from WIVS-FM. Refinement was accomplished by iterative cycles of manual model building with Coot(56) and automated refinement using Phenix Refine(55) with noncrystallographic symmetry restraints. Model quality was assessed with the Phenix Validation tool(55). All of the protein structure figures and alignments were generated using PyMOL software (The PyMOL Molecular Graphics System, Version 1.6, Schrödinger LLC.).

2.4.13 Cofactor binding pocket volume calculations

A voxel grid approach was used to compute volume differences in the different heme pockets. Specifically, for each heme pocket model a 20 Å x 20 Å x 10 Å rectangular cuboid was constructed and aligned to the heme cofactor, centered 1 Å below the iron. Within this box, a 3-D grid is built, with grid lines every 0.5 Å. Each such grid point represents one voxel of 0.125
cubic Å. Voxels were deemed accessible to solvent if a probe sphere placed at the grid point in question did not clash with any of the surrounding protein atoms (excluding the heme cofactor). The probe sphere radius was set to 1.4 Å (to mimic water), and protein atom radii were obtained using the MSMS utility script pdb_to_xyzr (57). To ensure that the resulting collections of accessible voxels provided a fair representation of the volume in heme pocket the resulting point clouds were manually inspected in PyMOL (Figure S2.8B).

2.4.14 Monooxygenation Reactions

Monooxygenation reactions were carried out with 12-(4-nitrophenoxy)dodecanoic acid (12-pNCA) as a substrate, which was synthesized as previously described (58). WIVS-FM/Fe-DPIX variants were expressed in minimal media and purified via 6 x His Tag chromatography as described above, while BM3h/heme variants were expressed from TB media as described elsewhere (47) and purified as above. For full-length WIVS-FM/Fe-DPIX and P450<sub>BM3</sub>/heme consisting of the heme domain with the fused reductase domain, expressions were carried out under the same conditions as for the BM3h variants, and the reactions performed in clarified lysate. Reactions were conducted in microplate format in a total volume of 250 μL. Typical reactions set-up was in 0.1 M Tris-HCl pH 8.1, adding 2.5 μL of a stock solution of substrate (10–20 mM 12-pNCA in DMSO, reactions contained 1% DMSO) and enzyme. Reactions with BM3h variants were initiated by the addition of 25 μL of H₂O₂ stock solution (50 mM H₂O₂ in 0.1 M Tris pH 8.1), while reactions with full length P450<sub>BM3</sub> variants were initiated by addition of 25 μL NADPH stock solution (8.7 mM in 0.1 M Tris pH 8.1). Final concentrations of reagents were 0.1 – 3 μM enzyme, 100 – 200 μM 12-pNCA, and 5 mM H₂O₂ or 870 μM NADPH in 0.1 M Tris-HCl pH 8.1. The formation of 4-nitrophenol over time was monitored by measuring absorbance at 398 nm with a Tecan M1000 PRO UV/Vis plate reader.

2.4.15 Cyclopropanation Reactions

To prepare proteins for catalysis, WIVS-FM/Fe-DPIX variants were expressed in minimal media and purified as described above, while BM3h/heme variants were expressed from TB
media as described previously (47) and purified as above. Small scale (400 µL) anaerobic reactions were setup in 2 mL crimp vials (Agilent). Distilled water and a 10X concentrated solution of 0.1 M sodium dithionite in 1 M potassium phosphate buffer (pH 8.0) were degassed in separate sealed vials for at least 10 minutes prior to use. The headspace of a sealed crimp vial containing enzyme in 0.1 M potassium phosphate buffer (pH 8.0) (or Fe-DPIX or heme in DMSO) and a stir bar was flushed with argon, taking care to avoid bubbling of the protein solution. Using glass syringes, 40 µL of the 10X buffer/dithionite solution was added under argon, along with water to a final volume of 380 µL. 10 µL of a styrene solution (1.2 M in EtOH) was then added via syringe and allowed to mix for 30 seconds prior to adding 10 µL of ethyl diazoacetate (400 mM in EtOH). The final concentrations of the reagents were: 20 µM catalyst, 10 mM sodium dithionite, 30 mM styrene, and 10 mM ethyl diazoacetate in 0.1 M potassium phosphate (pH 8.0). The vial was removed from the argon line, the septum covered with parafilm, and the reaction was allowed to stir overnight at room temperature. Aerobic reactions were set up in a similar manner, simply open to air rather than under argon. Reactions were quenched by the addition of 30 µL of 3 M HCl. The vials were opened and 20 µL of an internal standard solution (100 mM cyclopropyl phenyl ketone in EtOH) was added. The reaction was extracted with 1 mL of ethyl acetate, transferred to a 1.7 mL microcentrifuge tube, vortexed, and centrifuged at 18,400 x g for 2 minutes. The organic layer was then removed and dried over sodium sulfate in a separate microcentrifuge tube. The organic layer was then filtered through a glass wool plug into glass vials for chiral gas chromatography analysis. GC analyses were carried out on a 7820A gas chromatograph with flame ionization detector (Agilent), using a J&W Scientific Cyclosil-B column (30 m x 0.25 mm x 0.25 µm film). Separations were achieved at constant pressure and a GC oven temperature gradient of 100 °C for 30 min, 1 °C/min to 135 °C, 135 °C for 10 min, 10 °C/min to 200 °C, and 200 °C for 5 min. Cyclopropane product yields were determined by generating a calibration curve (Figure S2.10B) from product standard prepared as described previously (47).
### 2.5 Supplementary Information

#### 2.5.1 Supplementary Tables

**Table S2.1.** Primers used in this study.

<table>
<thead>
<tr>
<th>Forward Primers</th>
<th>Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF1</td>
<td>CAGGAAACAGGATCAGCTTACTCCCC</td>
</tr>
<tr>
<td>BM3h_T269_OE_XXX_F*</td>
<td>CTGAATTCGGGTCCCGCAATTAAG</td>
</tr>
<tr>
<td>BM3h_L322_OE_XXX_F*</td>
<td>GCATGGTCTTAACGAAGCGXXXCGTTATGGCCAACTGC</td>
</tr>
<tr>
<td>WIVS_I153_OE_XXX_F*</td>
<td>CACGTTTAACGCTTGATACGXXXGCTTCTTGCGGCCTTAAAC</td>
</tr>
<tr>
<td>WIVS_G265_OE_XXX_F*</td>
<td>CAAAATTTAACATTCTATTCGCGXXXACGGAAACAGATGAAGAAG</td>
</tr>
<tr>
<td>WIVS_F405_OE_XXX_F*</td>
<td>GTGCGTGTATCGGTCAGCAGXXXAGTCTTCTCATGAAGCAACGC</td>
</tr>
<tr>
<td>BM3h_WIVSFM_S406A_Bsal_F</td>
<td>GATGAAGGTCTCTGATGGCTCTCATTGAAGCAACGCCTTGG</td>
</tr>
<tr>
<td>BM3h_WIVSFM_L52I_Bsal_F</td>
<td>GATGAAGGTCTCTCTACATTCAAGTGACGCTTAAATAG</td>
</tr>
<tr>
<td>BM3h_WIVSFM_I366V_Bsal_F</td>
<td>GATGAAGGTCTCTAAGCTGAGGTTGCACGACGATTGTGGAAGA</td>
</tr>
<tr>
<td>BM3h_G265F_OE_F</td>
<td>CAAAATTTAACATTCTATTCGCGTTACAGCAACGCTTGGT</td>
</tr>
<tr>
<td>WIVS_F405M_OE_F</td>
<td>GTGCGTGTATCGGTCAGCAGATGGCTCTTCTATGAAGCAACGC</td>
</tr>
<tr>
<td>BM3h_F405M_OE_F</td>
<td>GTGCGTGTATCGGTCAGCAGATGGCTCTTCTATGAAGCAACGC</td>
</tr>
<tr>
<td>BM3h_A406NNK_Bsal_F</td>
<td>GAATTCAAGGCTTCTTCTTTACNGCTTCTTATGAAGCAACGCTTGG</td>
</tr>
<tr>
<td>WIVSFM_T268A_Bsal_F</td>
<td>GATGTTGGTCTCTCGAACGTAATGTTGGTATCTATTGGG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reverse Primers</th>
<th>Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCWori_rev_new</td>
<td>GCGCACAATTTCCGGAACGAAATGCCAACC</td>
</tr>
<tr>
<td>BM3h_T269_OE_R</td>
<td>TGGTTTCGTGTCGCGCAATTAAG</td>
</tr>
<tr>
<td>BM3h_L322_OE_R</td>
<td>CGCTTCTAAAGACCTAGCC</td>
</tr>
<tr>
<td>WIVS_I153_OE_R</td>
<td>CGTATCAAAGGCTTAAACACGT</td>
</tr>
<tr>
<td>WIVS_G265_OE_R</td>
<td>CGCAATTAAAGGCTTAAACTTG</td>
</tr>
<tr>
<td>WIVS_F405_OE_R</td>
<td>CTGTCTGACACCATACGAC</td>
</tr>
<tr>
<td>BM3h_WIVSFM_S406A_Bsal_R</td>
<td>GATGAAGGTCTCTCTCATCTGCTGACGACACGGACGCTGACCC</td>
</tr>
<tr>
<td>BM3h_WIVSFM_L52I_Bsal_R</td>
<td>GATGAAGGTCTCTGTAGCGCAGTGCCTTATGAAGCAACGC</td>
</tr>
<tr>
<td>BM3h_WIVSFM_I366V_Bsal_R</td>
<td>GATGAAGGTCTCTTGGTTTTATCACGGTGAAGCTGAG</td>
</tr>
<tr>
<td>BM3h_A406NNK_Bsal_R</td>
<td>GAATTCAAGGCTTCTTCTTTACNGCTTCTTATGAAGCAACGCTTGG</td>
</tr>
<tr>
<td>WIVSFM_T268A_Bsal_R</td>
<td>GATGTTGGTCTCTCGAACGTAATGTTGGTATCTATTGGG</td>
</tr>
</tbody>
</table>

* XXX = YWC, GBC, ATK, TGG codons mixed in a ratio of 4:3:2:1, respectively, encoding the amino acids G, W, F, L, I, M, H, A, V, and Y.
**Table S2.2.** Mutations present in hits from first library.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Thr269</th>
<th>Leu272</th>
<th>Leu322</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ala</td>
<td>Trp</td>
<td>Ile</td>
</tr>
<tr>
<td>2</td>
<td>Val</td>
<td>Trp</td>
<td>Ile</td>
</tr>
<tr>
<td>3</td>
<td>Ile</td>
<td>Trp</td>
<td>Ile</td>
</tr>
</tbody>
</table>

**Table S2.3.** Cofactor selectivity of variants used for catalysis in minimal media and rich media.

<table>
<thead>
<tr>
<th>Variant</th>
<th>mol % from HPLC Fe-DPIX (M9 media)</th>
<th>mol % from HPLC heme (TB media)</th>
<th>mol % from HPLC Fe-DPIX (M9 media)</th>
<th>mol % from HPLC heme (TB media)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIVS-FM</td>
<td>&gt;99.9</td>
<td>n.d.</td>
<td>94.9</td>
<td>5.1</td>
</tr>
<tr>
<td>T268A/Fe-DPIX</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table S2.4.** Gradient used for HPLC analysis of purified proteins.*

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>% A (H$_2$O + 0.1% formic acid)</th>
<th>% B (MeCN + 0.1% formic acid)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>20.01</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>25</td>
<td>95</td>
<td>5</td>
</tr>
</tbody>
</table>

* Separations were achieved on a Restek Viva C4 column (5 µm, 2.1 x 150 mm).
<table>
<thead>
<tr>
<th>pdb accession #</th>
<th>WIVS Fe-DPIX</th>
<th>WIVS-FM Fe-DPIX</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P 21 21 21</td>
<td>P 21 21 21</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.000</td>
<td>1.040</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>106.708, 167.405, 228.471</td>
<td>106.057, 166.299, 229.331</td>
</tr>
<tr>
<td>a, b, g (°)</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>32.97 – 2.34 (2.34 – 2.38) *</td>
<td>28.85 – 2.16 (2.16 – 2.20) *</td>
</tr>
<tr>
<td>Rmerge</td>
<td>16.2 (46.2)</td>
<td>10.7 (47.0)</td>
</tr>
<tr>
<td>σI/I</td>
<td>9.0 (4.6)</td>
<td>6.5 (2.6)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (100.0)</td>
<td>98.9 (99.3)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>11.5 (11.4)</td>
<td>4.9 (4.4)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>32.97 – 2.34</td>
<td>28.85 – 2.16</td>
</tr>
<tr>
<td>No. reflections</td>
<td>173212</td>
<td>213435</td>
</tr>
<tr>
<td>Rwork / Rfree</td>
<td>0.21/ 0.28</td>
<td>0.26/0.33</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>28537</td>
<td>28408</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>312</td>
<td>312</td>
</tr>
<tr>
<td>Water</td>
<td>1163</td>
<td>1196</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>23.0</td>
<td>15.7</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>13.9</td>
<td>7.6</td>
</tr>
<tr>
<td>Water</td>
<td>17.9</td>
<td>10.8</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.009</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.149</td>
<td>1.163</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>0.6 %</td>
<td>0.5 %</td>
</tr>
</tbody>
</table>

*All data sets were collected from single crystals. Highest-resolution shell is shown in parentheses*
2.5.2 Supplementary Figures

Figure S2.1. Ferrous-CO bound absorbance spectra of BM3h with Fe-DPIX. Protein expressed in minimal media supplemented with Fe-DPIX in cells expressing ChuA (purple) and cells without ChuA (black).

Figure S2.2. HPLC assay for metalloporphyrin quantitation. A) Sample HPLC traces of variants along the evolutionary trajectory, shown in comparison with Fe-DPIX and heme standards. B) HPLC calibration curves for Fe-DPIX (left) and heme (right) used to calculate the concentration of these cofactors in proteins.
Figure S2.3. Heme vs. Fe-DPIX content in BM3h variants. Mol % Fe-DPIX (purple) and heme (blue) for each variant expressed in minimal media (A) and TB (B) supplemented with Fe-DPIX, as determined by HPLC analysis of purified proteins. Error bars represent standard deviations determined from n=3 independent experiments, with the exception of WIVS-FM, for which n=7 in minimal media and n=8 in TB.
Figure S2.4. Ferrous-CO bound spectra of BM3h variants expressed in rich media. A) WIVSFM expressed in the absence of Fe-DPIX; yields “P420” misligated protein. B) Spectra of variants WIV-FM (blue), FM (orange), S-FM (green) expressed with Fe-DPIX supplementation. C) Thermostable WIVS-FM-IV expressed in the presence of Fe-DPIX.
Figure S2.5. Overall structure of BM3h variants. Overlay of wild type substrate-free BM3h (purple, 2IJ2 chain A), substrate-bound BM3h (blue, 1JPZ chain A) and WIVS-FM (A) or WIVS (B) in green, with the F, G and I helices indicated by arrows.
Figure S2.6. Conformation of the Fe-DPIX porphyrin ring in WIVS and WIVSFm. $2F_o - F_c$ map contoured at $\sigma = 2.0$ Å, showing the electron density around the Fe-DPIX cofactor in WIVS-FM (A) and WIVS (B). Spherical representation of the cofactor and select residues in the active site of WIVS-FM (C). There are steric clashes between the porphyrin methyl group and the side chains of F265 and S406 when the Fe-DPIX cofactor is modeled in the “180° flipped” conformation, shown in blue (D), and additional clashes with W272 when native heme is modeled into the active site, shown in purple (E).
Figure S2.7. I-helix backbone rearrangements in evolved BM3h variants. (A) Hydrogen bonding patterns in the I-helix of wild type (green, left, 2IJ2) WIVS (yellow, middle) and WIVS-FM (cyan, right). Relevant H-bonds and waters are shown as black dashed lines and red spheres, respectively. Only the side chains of residues 265 and 269 are shown for clarity. (B) Cartoon overlay of the I-helix of WIVS (yellow) with substrate-bound (purple, 1JPZ) and substrate-free (green, 2IJ2) wild type structures. (C) The I-helix of WIVS-FM (cyan) overlaid with WIVS (yellow), shown in stick representation. Only the side chain of residue 265 is shown for clarity.
Figure S2.8. Volume of the cofactor binding pocket in BM3h variants. A) The cofactor binding pocket (gray surface) of WIVS (yellow, see wild type and WIVS-FM cavity in Fig. 4E of the main text) B) Voxel point clouds (red) representing the volume of the cofactor binding pocket in wild type BM3h (green, left, 2IJ2) WIVS (yellow, middle) and WIVS-FM (cyan, right).
**Figure S2.9.** Determination of the ratio of cofactor-bound protein to total protein for BM3h variants. A) SDS-PAGE analysis of purified BM3h variants. The 50 kDa and 70 kDa protein markers are labeled. L = ladder, 1 = WIVS-FM/Fe-DPIX, 2 = Wild type BM3h/heme. B) Ratio of cofactor-bound to total protein. *calculated by dividing the concentration of cofactor-bound protein (determined via HPLC analysis) by total protein concentration (determined by $A_{280nm}$).
Figure S2.10. Chiral GC analysis of cyclopropanation reactions. A) GC traces from cyclopropanation reactions of styrene with EDA. Elution times: 3a (34.0 min), 3b (36.8 min), 3c (40.6 min), and 3d (41.5 min). Insets show ferrous CO-bound absorbance spectra of the respective proteins. B) Calibration curve of ethyl 2-phenylcyclopropane-1-carboxylate used to determine product concentrations.
Figure S2.11. 12-pNCA monooxygenation by P450_{BM3} variants. A) Reaction scheme illustrating oxidation of 12-pNCA by P450_{BM3} (driven by NADPH or H$_2$O$_2$) to form the colored product, 4-nitrophenolate. B) Time course of peroxide-driven oxidation of 12-pNCA catalyzed by BM3h variants. Reaction conditions were 0.25 µM BM3h F87A/heme (0.9 µM WIVS-FM F87A/Fe-DPIX), 100 µM 12-pNCA, and 5 mM H$_2$O$_2$. C) Time course of NADPH-driven oxidation of 12-pNCA catalyzed by full length P450_{BM3} variants that include the P450_{BM3} reductase domain in addition to the heme-binding domain. Reaction conditions were 0.1 µM BM3 F87A-R1/heme (3 µM WIVS-FM F87A-R1/Fe-DPIX), 200 µM 12-pNCA, and 870 µM NADPH. R1 = reductase domain. The background absorbance, taken as the zero time point, was subtracted from both plots. The boxed graphs to the right show the ferrous CO-bound absorbance spectra of the enzymes used for catalysis.
REFERENCES


32. Lopez MS, Kliegman Jl, Shokat KM (2014) *The logic and design of analog-sensitive*


3.1 Introduction

Approximately half of all enzymes require a metallocofactor for function (1). These metalloenzymes perform a diverse array of tasks in the cell, enabled by the inherent reactivity of the metal center, which is tuned toward a specific purpose by the surrounding protein scaffold. The ability to modulate the activity and selectivity of metalloenzymes through laboratory evolution has permitted the use of these systems in diverse synthetic transformations (2–5). Furthermore, novel reactions, including abiological carbenoid and nitrenoid transfer reactions, can be accessed through promiscuous activity of metalloenzymes with non-natural substrates (6–9). Despite the breadth of catalysis displayed by natural metalloenzymes, there remain many synthetic transformations that are not accessible with the reactive groups provided by nature.

To further expand biocatalytic reaction space, researchers have sought to create artificial metalloenzymes which utilize non-natural metallocofactors with unique reactivity (1, 10, 11). In particular, many efforts have focused on utilizing existing metallocofactor-binding motifs to introduce new metals into proteins (12–14). These metal-substituted enzymes often possess new activity in abiological reactions that can be optimized through mutation of the protein host (15–18). However, most current techniques for generating artificial metalloenzymes are labor-intensive and require extensive in vitro manipulations, making directed evolution challenging and preventing application of these systems in cells (1, 15, 19).

We recently introduced a new strategy for the introduction of non-natural metallocofactors, in the form of synthetic heme derivatives, into proteins via the evolution of orthogonal enzyme/cofactor pairs (20). Orthogonal enzyme/cofactor pairs comprise a non-
natural cofactor and a complementary binding protein that interact self-sufficiently without crosstalk with native cofactors or enzymes, thus permitting efficient construction of non-natural metalloenzymes in cells. To establish the utility of this approach, the heme domain of a cytochrome P450 from *Bacillus megaterium* (BM3h) was evolved to selectively incorporate a nonproteinogenic metalloporphyrin, Fe-deuteroporphyrin IX (Fe-DPIX), *in vivo*, even in the presence of native heme (20). This work demonstrated that the generation of orthogonal enzyme/cofactor pairs is a promising method to efficiently introduce new chemical functionality into proteins. However, because both Fe-DPIX and heme possess an iron center, the orthogonal Fe-DPIX enzyme displayed a similar reaction scope when compared to the native heme enzyme (20).

Here, we extend the application of orthogonal enzyme/cofactor pairs to reactions not catalyzed by native enzymes through introduction of a non-natural metal center. Metal substitution can greatly expand the reaction scope of heme enzymes (13, 14, 21, 22). For example, recent work has shown that Ir(Me)-containing myoglobin and cytochrome P450 variants catalyze C-H insertion reactions not accessible with the iron-heme proteins (14, 16, 18). Inspired by these results, we sought to create an orthogonal platform for efficient *in vivo* assembly of iridium enzymes. We reasoned that the selective recognition of the deuteroporphyrin IX scaffold achieved with our previously evolved orthogonal enzyme would enable substitution of Fe-DPIX for Ir(Me)-DPIX while maintaining orthogonal recognition of the cofactor (*Figure 3.1*). In validation of this approach, an orthogonal enzyme variant was isolated with >98% enrichment for Ir(Me)-DPIX, from cells grown in rich media producing endogenous heme. Furthermore, the orthogonal Ir(Me)-DPIX enzyme shows activity in challenging abiological cyclopropanation and aziridination reactions that are not catalyzed by native heme enzymes. The ability to generate orthogonal artificial metalloenzymes in cells will facilitate directed evolution of these systems and will open the door for integration of abiological catalysts into biosynthetic pathways.
3.2 Results and Discussion

3.2.1 In vivo incorporation of Ir(Me)-DPIX into BM3h variants

First, Ir(Me)-DPIX was synthesized in two steps and 38% overall yield from deuteroporphyrin IX dimethyl ester, by adapting literature procedures for the preparation of other Ir(Me)-heme analogs (14). With this non-natural cofactor in hand, we then tested for cellular uptake of Ir(Me)-DPIX and its incorporation into proteins in vivo.

Figure 3.1. Generation of an orthogonal metal-substituted enzyme. A) Selective recognition of the non-natural cofactor Fe-DPIX by the orthogonal enzyme WIVS-FM* could be extended to metal-substituted deuteroporphyrin derivatives, namely, Ir(Me)-DPIX, for in vivo generation of protein catalysts for abiological reactions. B) Absorbance spectra of native BM3h/heme (blue) and the orthogonal WIVS-FM/Fe-DPIX pair. C) Absorbance spectra of free Ir(Me)-DPIX (red) and WIVS-FM* T268A (blue), WIVS-FM* T268A C400G (green) and WIVS-FM* T268A C400A (violet) expressed in minimal media supplemented with Ir(Me)-DPIX.
Overexpression of the *Escherichia coli* heme transport protein ChuA permits cellular uptake of heme derivatives with modifications to both the metal center and/or porphyrin scaffold (20, 23, 24). Therefore, we sought to exploit this system for *in vivo* incorporation of an Ir(Me)-DPIX cofactor into proteins (Figure 3.1).

Several variants of BM3h and the orthogonal enzyme, WIVS-FM (see Table S3.1 for a list of amino acid mutations present in each variant), were interrogated for their ability to accommodate the Ir(Me)-DPIX cofactor. Mutation of the conserved axial cysteine to glycine in the cytochrome P450 enzyme CYP119 was shown to improve the abiological carbenoid transfer activity of CYP119 harboring an Ir(Me)-heme derivative, presumably by accommodating the pendent Ir(Me)-DPIX methyl group and allowing orientation of the cofactor in the active site such that it has an open distal coordination site, facilitating substrate activation (see orientation of Ir(Me)-DPIX in enzyme shown in Figure 3.1(16)). Therefore, several different axial amino acids enviroments were interrogated in BM3h and WIVS-FM: the native C400 residue, a C400A variant, and a C400G variant. In addition, the mutation T268A was introduced into all variants, as this mutation has also been shown to be activating in carbenoid transfer reactions (6, 25). Furthermore, as protein stability often diminishes with increasing number of mutations, a variant of WIVS-FM with two thermostabilizing mutations (denoted here as WIVS-FM*) (20, 26) was also examined; the aforementioned axial mutations were also introduced into this parent scaffold. In total, this created 9 variants to screen for Ir(Me)-DPIX incorporation.

These variants, along with ChuA, were expressed in *E. coli* grown in iron-deficient minimal media (to limit heme biosynthesis) that was supplemented with 10 µM Ir(Me)-DPIX. BM3h and thermostable WIVS-FM* variants containing Ir(Me)-DPIX were obtained in good yield (up to 14 mg/L), while WIVS-FM variants gave lower yields (Table 3.1). Expression of the C400A and C400G variants of BM3h, WIVS-FM and WIVS-FM* with Ir(Me)-DPIX yielded protein with a single and distinct Soret band at 402 nm compared to the free cofactor which has a peak absorbance at 386 nm (Figure 3.1C, Figure S3.1 and Table S3.2). An absorbance peak at
approximately 400 nm was previously observed for a BM3h-C400S variant that was reconstituted in vitro with an Ir(Me)-heme analog (14). This absorbance is blue-shifted compared to the native heme and orthogonal Fe-DPIX enzymes, which display peaks at 421 and 413 nm respectively (Figure 3.1B). Variants possessing the native axial Cys residue, on the other hand, displayed two prominent Soret bands at 406 nm and 432 nm when expressed in the presence of Ir(Me)-DPIX (Figure 3.1C, Figure S3.1 and Table S3.2). We hypothesize that the peak at 432 nm results from cysteine ligation of the Ir(Me)-DPIX cofactor, whereas the peak at approximately 400 nm in these variants likely represents protein-bound Ir(Me)-DPIX that is not directly ligated by a protein residue. These results confirm the promiscuous ChuA-mediated transport of Ir(Me)-DPIX and its incorporation into proteins inside of cells.

3.2.2. An orthogonal artificial metalloenzyme

With the development of an effective means of introducing Ir(Me)-DPIX to proteins in cells, we next sought to determine if the orthogonal enzyme maintained selective recognition of the metal-substituted deuteroporphyrin cofactor. Because WIVS-FM* provided higher yields, this was used as the orthogonal scaffold rather than WIVS-FM. The selectivity for Ir(Me)-DPIX vs. heme was compared for the C400, C400A and C400G variants of BM3h T268A and WIVS-FM* T268A expressed in both minimal media (low cellular heme levels) and rich media (TB, normal levels of endogenous heme) in the presence of Ir(Me)-DPIX. The T268A mutation was included due to its importance in enabling carbenoid transfer reactivity, as discussed above. This mutation does not affect selectivity for Fe-DPIX in WIVS-FM (20). Modification of a previously described high-performance liquid chromatography (HPLC) assay (20) allowed for quantification of Ir(Me)-DPIX and heme content in protein preparations (Figure 3.2, Table 3.1, Figure S3.2 and Table S3.3).

All variants expressed in minimal media showed ≥ 95% enrichment for Ir(Me)-DPIX, consistent with results obtained previously for Fe-DPIX (Table 3.1 and Figure S3.2) (20).
Table 3.1. Selectivity of BM3h variants for Ir(Me)-DPIX vs. heme in minimal media and rich media.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Variant</th>
<th>mol %\textsuperscript{b} (M9 media)</th>
<th>mol %\textsuperscript{b} (TB media)</th>
<th>M9 yield (mg/L)\textsuperscript{c}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ir(Me)-DPIX</td>
<td>heme</td>
<td>error</td>
</tr>
<tr>
<td>BM3h T268A</td>
<td>95.0</td>
<td>5.0</td>
<td>3.2</td>
</tr>
<tr>
<td>BM3h T268A C400A</td>
<td>98.4</td>
<td>1.6</td>
<td>2.2</td>
</tr>
<tr>
<td>BM3h T268A C400G</td>
<td>99.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>WIVS-FM* T268A</td>
<td>96.7</td>
<td>3.3</td>
<td>2.1</td>
</tr>
<tr>
<td>WIVS-FM* T268A C400A</td>
<td>99.7</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>WIVS-FM* T268A C400G</td>
<td>100</td>
<td>n.d.</td>
<td>n/a</td>
</tr>
</tbody>
</table>

\textsuperscript{a}All values reported are the average and standard deviations from three independent experiments.
\textsuperscript{b}Determined by HPLC analysis of purified proteins from expressions in the denoted media.
\textsuperscript{c}Yields were determined via HPLC analysis or from absorbance spectra using $\varepsilon_{402\ nm} = 107.7\ \text{mM}^{-1}\text{cm}^{-1}$

Figure 3.2. HPLC assay for metalloporphyrin quantitation. Sample HPLC traces of several BM3h and WIVS-FM* variants, shown in comparison with Ir(Me)-DPIX and heme standards.

Remarkably, WIVS-FM* T268A C400G showed no detectable heme binding when expressed in minimal media. In contrast, when expressed in rich media, BM3h T268A showed a preference for native heme over Ir(Me)-DPIX (Ir(Me)-DPIX/heme = 37.8 mol%:62.2 mol%; Figure 3.2, Table 3.1 and Figure S3.2). WIVS-FM* T268A, although slightly favoring Ir(Me)-DPIX, also contained a considerable amount of contaminating heme (Ir(Me)-DPIX/heme = 51.8 mol%:48.2 mol%.
mol%; **Figure 3.2, Table 3.1 and Figure S3.2**). Compared to our previous results with Fe-DPIX (20), these results show that Ir(Me)-DPIX competes less effectively against heme for binding to these proteins.

Introduction of the C400A or C400G mutations, however, improved Ir(Me)-DPIX selectivity. The C400A and C400G variants of BM3h T268A both displayed increases in Ir(Me)-DPIX selectivity in rich media (80.2 mol% and 81.9 mol% respectively; **Figure 3.2, Table 3.1 and Figure S3.2**). Notably, the WIVS-FM* T268A C400A and C400G variants were highly selective for Ir(Me)-DPIX, with greater than 97% purity in the presence of endogenous heme (**Figure 3.2, Table 3.1 and Figure S3.2**). These data suggest that the identity of the metal center and axial ligand have a significant influence on the efficiency of non-natural cofactor incorporation into hemoproteins. In this case, mutations that provide selective recognition of the deuteroporphyrin scaffold, as well as mutations that allow for accommodation of the axial methyl group of the cofactor, were required to achieve complete orthogonality with WIVS-FM* T268A C400A(G). Due to the high degree of Ir(Me)-DPIX selectivity and high yields obtained with WIVS-FM* T268A C400G, this variant was used for the catalysis experiments described below.

### 3.2.3 Non-natural carbenoid and nitrenoid insertion reactions catalyzed by an orthogonal Ir(Me)-DPIX enzyme

Heme enzymes have been shown to catalyze non-natural reactions involving carbenoid and nitrenoid insertion into olefins and X-H bonds (X = C, N, S, Si) (6, 25, 27–34). Through mutagenesis of the protein scaffold, highly active and selective catalysts have been obtained for these transformations. However, in many of these studies the substrate scope is limited to electron-rich olefins or molecules with activated X-H bonds (9, 35). Recently, Hartwig and coworkers have demonstrated the ability of Ir(Me)-substituted hemoproteins to catalyze carbenoid and nitrenoid insertion into C-H bonds of more difficult substrates not accessible with the native iron proteins (14, 16, 18). To further explore the catalytic potential of artificial Ir(Me)-enzymes, we tested the activity of the orthogonal WIVS-FM* T268A C400G/Ir(Me)-DPIX pair in
abiological carbenoid and nitrenoid-mediated olefin cyclopropanation and aziridination, respectively. Alkyl olefins, such as 1-octene (1), have not been demonstrated as substrates for heme enzymes in these reactions. Therefore, we began with this molecule to determine if orthogonal artificial metalloenzymes possess the ability to expand biocatalytic reaction space.

Figure 3.3. Cyclopropanation of 1-octene with EDA catalyzed by WIVS-FM* T268A C400G/Ir(Me)-DPIX. GC-MS traces of cyclopropanation reactions with 1-octene and EDA. Elution times: 3 (16.9 and 17.4 min), 4 (11.6 and 12.0 min), and internal standard cyclopropyl phenyl ketone (15.1 min).

First, we compared the activity of WIVS-FM* T268A C400G/Ir(Me)-DPIX and BM3h T268A/heme as catalysts for the cyclopropanation of 1-octene with ethyl diazoacetate (EDA, 2) (Figure 3.3). Hartwig and coworkers have shown that Ir(Me)-myoglobin catalyzes this reaction (14), thus this reaction can be used as a benchmark to compare the orthogonal artificial metalloenzyme to the previously developed system. Gas chromatography-mass spectrometry (GC-MS) analysis of the reaction using WIVS-FM* T268A C400G/Ir(Me)-DPIX as the catalyst revealed the formation of a small amount of product with a mass fragmentation pattern consistent with the cyclopropanation product 3, along with EDA dimerization product 4 (Figure 3.3 and Figure S3.3). In contrast the free Ir(Me)-DPIX cofactor and BM3h/heme showed no cyclopropanation activity and only produced the EDA dimerization product (Figure 3.3). While further work is required to confirm formation of the desired product and to quantify turnover.

70
number and selectivity through comparison with authentic product standards, this result is encouraging in that it suggests the orthogonal artificial metalloenzyme possesses expanded activity in non-natural reactions.

We next turned our attention to aziridination of 1-octene with tosyl azide 5 (Figure 3.4). While heme enzymes can utilize tosyl azide as a nitrenoid precursor for aziridination, this reaction was only demonstrated with styrene and other aryl olefins (36). Furthermore, achieving aziridination with native heme-P450s requires reduction of the iron center, necessitating the use of the full length protein consisting of the fused heme and reductase domains and consuming the expensive cofactor NADPH as the reductant (36). In contrast, Ir(Me) enzyme variants can activate sulfonyl azides in intramolecular C-H amination reactions without the need for added reductant, greatly simplifying the system (18). However, no artificial metalloenzymes have been developed for aziridination reactions to date.

![Figure 3.4. Aziridination of 1-octene with tosyl azide catalyzed by WIVS-FM* T268A C400G/Ir(Me)-DPIX. LC-MS trace of the aziridination reaction with 1-octene and tosyl azide. The total ion chromatogram (black) is overlaid with the extracted ion chromatogram for the calculated mass of the expected aziridine product 6 (calc. [M+H]⁺ = 282.1522). The mass spectrum is shown for the peak at 16.8 min which appears in the extracted ion chromatogram.](image-url)
Remarkably, liquid chromatography-mass spectrometry (LC-MS) analysis of the reaction of 1-octene with tosyl azide using WIVS-FM* T268A C400G/Ir(Me)-DPIX as a catalyst showed a peak with a m/z value of 282.1547, consistent with the calculated [M+H]+ value of 282.1522 for the aziridine product 6 (Figure 3.4). No peak in the LC-MS chromatogram could be clearly identified as the reduced sulfonamide byproduct 7, though this was detected in GC-MS analysis of the reaction (data not shown). Further work is required to confirm formation of the aziridine product rather than C-H amination product, and to quantify the results. Also, control reactions with the heme enzyme need to be conducted for comparison of the two systems. In previous work with the BM3/heme enzyme, the mutation I263F was found to be activating for the aziridination of styrene derivatives (36). It will be interesting to see if the introduction of this mutation to WIVS-FM* T268A C400G/Ir(Me)-DPIX improves the aziridination activity of this enzyme. Taken together, these preliminary cyclopropanation and aziridination results demonstrate the potential for orthogonal artificial metalloenzymes to expand the scope of enzymatic catalysis and open the door to exploring the activity of this system with other challenging substrates (Figure 3.5).

![Figure 3.5](image_url)

**Figure 3.5. Additional olefinic substrates to be tested in WIVS-FM* T268A C400G/Ir(Me)-DPIX-catalyzed cyclopropanation and aziridination.** Alkyl and electron-deficient olefins were chosen due to the inability of heme enzymes to catalyze reactions with these substrates and for the opportunity to engineer the enzyme for selective catalysis. For example, we envision generating catalysts for cyclopropanation/aziridination with cyclohexene 8, rather than allylic C-H insertion, and for regioselective cyclopropanation/aziridination of one of the two double bonds in carvone 9.

### 3.3 Conclusions and Future Work

Here, we report the design and catalytic properties of an orthogonal artificial metalloenzyme utilizing Ir(Me)-DPIX. The orthogonal enzyme, WIVS-FM* T268A C400A(G),
was adapted from an enzyme previously evolved for selective recognition of Fe-DPIX, taking advantage of the amino acid substitutions in that protein scaffold which provided specificity towards deuteroporphyrin while introducing an additional mutation to the axial cysteine residue in order to accommodate the axial methyl ligand of the cofactor, thereby achieving high selectivity for Ir(Me)-DPIX. This demonstrates the utility of orthogonal enzyme/cofactor pairs for the efficient introduction of metal-substituted cofactors to proteins inside of cells.

The orthogonal artificial metalloenzyme was active in cyclopropanation and aziridination reactions with substrates that are not accessible with the native heme enzyme. We are continuing work towards complete characterization of the cyclopropanation and aziridination activity of WIVS-FM* T268A C400G/Ir(Me)-DPIX with 1-octene and other challenging substrates shown in Figure 3.5. In addition, efforts are currently underway to obtain a crystal structure of WIVS-FM* C400A(G)/Ir(Me)-DPIX, to better understand the molecular determinants of selectivity for the non-natural cofactor, and to guide further engineering efforts aimed at improving the catalytic properties of the enzyme.

The ability to orthogonally construct artificial metalloenzymes opens the door to utilization of these systems for non-natural catalysis in vivo. This eliminates the need for laborious protein purification and can facilitate directed evolution of these systems to generate highly active abiological catalysts for a broad range of reactions. It can also allow for integration of the orthogonal enzyme into biosynthetic pathways. In pursuing these goals however, care will have to be taken to prevent background activity from any non-natural cofactor present in the cell that is not associated with protein. In some cases, synthetic cofactors are inhibited by cellular components, which could alleviate this issue (15, 17).

This work demonstrates that metal-substitution in orthogonal enzymes is a promising strategy to expand the chemical functionality of the cell. In the following chapter, I describe my work in generating orthogonal proteins for the recognition of unique cofactor structures that impart improved cyclopropanation activity through synthetic derivatization of the porphyrin
scaffold. In the future, we envision combining the strategies of alteration of the metal center and modification of porphyrin structure to create a suite of artificial metalloenzymes tuned towards diverse applications.

3.4 Materials and Methods

3.4.1 General methods and materials

Deuteroporphyrin IX dimethyl ester was obtained from Frontier Scientific. Oligonucleotides were obtained from Integrated DNA Technologies. Enzymes and reagents used for cloning were obtained from New England BioLabs. All other chemical reagents and solvents were obtained from chemical suppliers (Acros, Fisher Scientific, or Sigma-Aldrich) and used without further purification.

Reactions requiring inert atmosphere were carried out in oven-dried glassware using standard Schlenk techniques under nitrogen. If required, solvents were degassed by bubbling with nitrogen. Reaction progress was monitored by thin layer chromatography (EMD Millipore TLC silica gel 60 F254). Silica gel chromatography was performed on an automated Biotage Isolera One using 50 g SNAP Ultra columns. Proton magnetic resonance spectra were acquired on 400 MHz Bruker instruments in the Department of Chemistry NMR facility at the University of North Carolina – Chapel Hill. NMR spectra were processed with ACD/NMR Processor (ACD/Labs). Chemical shifts are referenced to residual solvent peaks (37). GC-MS analyses were obtained with Agilent 7820A gas chromatograph equipped with a J&W Scientific HP-5MS UI column (30 m x 0.25 mm ID x 0.25 µm film) and an Agilent 5977E mass spectrometer detector. LC-MS analyses were obtained with an Agilent 6520 Accurate Mass QTOF LC-MS ESI positive in high-resolution mode, equipped with a Thermo Scientific Acclaim RSLC 120 C18 column (2.2 µm particle size, 120 Å pore diameter, 2.1 x 150 mm).

3.4.2 Organic synthesis and characterization

The synthesis of Ir(Me)-DPIX was adapted from procedures for the synthesis of Ir(Me)-mesoporphyrin IX (14). The steps starting from commercial deuteroporphyrin IX dimethyl ester
(Frontier Scientific) are described below.

A. Ir(CO)Cl-deuteroporphyrin IX dimethyl ester

![Chemical Structure](image)

A solution of deuteroporphyrin IX dimethyl ester (95 mg, 0.18 mmol) and [Ir(COD)Cl]$_2$ (155 mg, 0.23 mmol) in xylenes (100 mL) was refluxed under an air atmosphere. Reaction progress was monitored by TLC with 1:1 hexane to ethyl acetate as the mobile phase. The reaction typically went overnight to reach completion. The solvent was removed via rotary evaporation and purified by column chromatography on silica gel. Impurities were first eluted with a gradient from 0 – 25 % ethyl acetate in hexane, then product was eluted as a pink band with a gradient from 25 – 50 % ethyl acetate in hexane. After evaporating the solvent, the purified product was redissolved in approximately 10 mL DCM and diluted with 100 mL hexane. The mixture was concentrated under vacuum resulting in precipitation of the pure product which was isolated by filtration as a red solid yielding 75 mg (53% yield). $^1$H NMR (CDCl$_3$, 400 MHz) $\delta$ 10.38 (d, J = 5.5 Hz, 2H), 10.32 (s, 1H), 10.27 (s, 1H), 9.21 (s, 1H), 9.18 (s, 1H), 4.51-4.46 (m, 4H), 3.83 (d, J = 9.7 Hz, 6H), 3.76-3.66 (m, 12H), 3.39 (t, J = 7.8 Hz, 4H) ppm. Calculated [M+Na]$^+$ = 815.1592 found [M+Na]$^+$ = 815.1553. UV-Vis (CH$_2$Cl$_2$, concentration = 5 x 10$^{-6}$ M) $\lambda_{\text{max}}$ (ε mM$^{-1}$cm$^{-1}$) 400 (95.6), 514 (14.8), 546 nm (23.5)
B. Ir(Me)-deuteroporphyrin IX

To a solution of Ir(CO)Cl-deuteroporphyrin IX dimethyl ester (70 mg, 0.088 mmol) in degassed, anhydrous ethanol (10 mL), a degassed solution of NaBH₄ (17 mg, 0.44 mmol) and NaOH (1 M) in water (3 mL) was added under nitrogen atmosphere. A color change from bright red to dark red-brown was observed upon addition of NaBH₄ to the porphyrin solution. The mixture was stirred for 1 hour at 50 °C in the dark. Following this, the reaction was cooled to room temperature and methyl iodide (6 mL) was added. The reaction was then stirred overnight at room temperature. The reaction mixture was concentrated to ~ 2 mL by rotary evaporation, and potassium hydroxide (300 mg) was added to achieve hydrolysis of the methyl esters. The reaction was stirred at room temperature and progress was monitored by TLC using 3% MeOH in DCM as the mobile phase. Upon completion (~4 hrs), the reaction was diluted in sodium phosphate buffer (10 mL, 0.1 M, pH 5) and acidified with HCl (1 M) to pH ~5. The resulting red precipitant was collected via centrifugation (4000 x g, 15 min, 4 °C) and the supernatant discarded. The solid was resuspended in water (15 mL) and centrifuged again. The resulting solid was dried under high vacuum overnight to yield the product as a dark red powder 49 mg (78% yield). ¹H NMR (CD₃CN, 400 MHz) δ 9.88-9.78 (m, 4H), 8.84 (d, J = 9.4 Hz, 2H), 4.31-4.27 (m, 4H), 3.71 (s, 3H), 3.68 (s, 3H), 3.62 (s, 3H), 3.57 (s, 3H), 3.28-3.24 (m, 4H), -7.94 (s, 3H) ppm. Calculated [M+H]⁺ = 717.2048 found [M+H]⁺ = 717.2007. UV-Vis (CH₃OH, 3.8 x 10⁻⁶) λmax (ε mM⁻¹cm⁻¹) 387 (121.1), 498 (8.5), 526 nm (13.3).
3.4.3 Plasmids and cloning

BM3 variants and the heme transport protein, ChuA, were cloned in the plasmids pCWori and pChuA, respectively, as described elsewhere (23, 38). Site-specific mutants were assembled via scarless restriction ligation of BsaI digested PCR fragments containing the desired mutation. Primer sequences are available upon request. See Table S3.1 for a list of amino acid mutations in the BM3 variants used in this study.

3.4.4 Protein expression

Native heme-BM3 variants were expressed as previously described (25). Expression of BM3 variants containing Ir(Me)-DPIX was accomplished by slight modification of a previously described methodology (20). BM3h variants in pCWori were co-transformed with pChuA into electrocompetent E.coli BL21-AI cells (Novagen). For expressions in minimal media, individual colonies were inoculated into liquid starter cultures of M9 minimal media supplemented with 0.5% glucose, 0.2% casamino acids, 340 µg/mL thiamine hydrochloride, 2 mM MgSO$_4$, 0.1 mM CaCl$_2$, 100 µg/mL kanamycin and 200 µg/mL ampicillin. For rich media expressions, colonies were inoculated into liquid starter cultures of LB media supplemented with 100 µg/mL kanamycin and 200 µg/mL ampicillin. The cultures were grown overnight at 37 °C and 225 rpm. One liter of M9 minimal media or TB supplemented with 100 µg/mL kanamycin and 200 µg/mL ampicillin was inoculated with 5 mL of overnight starter culture (the M9 or LB culture, respectively) and incubated for 3-4 hours at 37 °C and 225 rpm to an OD$_{600nm}$ of 0.6-0.7. The cultures were then induced by adding arabinose (to induce ChuA expression in BL21-AI) and Isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.2% and 1 mM, respectively. Due to observed decomposition of Ir(Me)-DPIX solutions when exposed to light, concentrated stock solutions of Ir(Me)-DPIX were freshly prepared at the time of induction by dissolving the solid in DMSO and diluting 10-fold in 0.1 M potassium phosphate, pH 12. The Ir(Me)-DPIX solution was added to the culture upon induction to a final concentration of 10 µM. Upon induction, the incubation temperature was reduced to 25 °C and expressions were
allowed to continue for 16-20 hours. After harvesting the cells by centrifugation (4 °C, 10 min, 4,000 × g), the cell pellets were frozen at -20 °C for at least 2 hours.

3.4.5 Protein purification

Proteins were purified via 6 × His-tag purification on a Ni-NTA HisTrap HP column (5 mL HisTrap HP, GE Healthcare, Piscataway, NJ) using an ATKA Purifier (GE Healthcare), as described elsewhere (20). Following purification, the protein was exchanged into 0.1 M potassium phosphate buffer, pH 8.0, with a 30 kDa molecular weight cut-off Amicon centrifugal filtration device (EMD Millipore), flash frozen on dry ice, and stored at -20 °C.

3.4.6 HPLC analysis of purified proteins for heme content

The procedure for HPLC analysis of proteins was adapted from a method for analysis of Fe-DPIX enzymes (20). Proteins for HPLC analysis were exchanged into distilled water via dialysis or a 30 kDa molecular weight cut-off Amicon centrifugal filter (EMD Millipore). Proteins were diluted to approximately 5 µM in distilled water and filtered through glass wool plugs prior to HPLC analysis on a 1260 Infinity binary liquid chromatography system (Agilent). Separations were achieved on a Restek Viva C4 column (5 µm, 2.1 x 150 mm) using the gradient described in Table S3.3. Calibration curves (Figure S3.2D) were generated for Ir(Me)-DPIX and heme by monitoring absorbance at 395 nm to determine metalloporphyrin concentrations in the protein sample.

3.4.7 Absorbance spectra of enzymes and the carbon monoxide binding assay

All absorbance measurements for protein samples were carried out in 0.1 M potassium phosphate buffer, pH 8. UV-vis data was obtained on a Tecan M1000 PRO UV/Vis plate reader or Cary 100 UV/Vis spectrometer (Agilent). The carbon monoxide difference spectra for native heme P450 enzymes were obtained as previously described in multiwell plate format (25). The absorbance spectrum of WIVS-FM* T268A C400G/Ir(Me)-DPIX was measured and compared with exact Ir(Me)-DPIX content determined via HPLC as described above. From these data, the extinction coefficient for the Soret peak was determined to be $\varepsilon_{402\text{nm}} = 107.7 \text{mM}^{-1}\text{cm}^{-1}$. 

78
3.4.8 Enzymatic Cyclopropanation and Aziridination reactions

Small scale (400 µL) anaerobic reactions were setup in a similar manner as was described previously (20), using 2 mL crimp vials (Agilent). Distilled water and a 10X concentrated solution of 1 M potassium phosphate buffer (pH 8.0) were degassed in separate sealed vials for at least 10 minutes prior to use. For heme enzyme reactions, this 10X buffer solution also contained 0.1 M sodium dithionite. The headspace of a sealed crimp vial containing enzyme in 0.1 M potassium phosphate buffer (pH 8.0) (or Ir(Me)-DPIX in 0.1 M NaOH) and a stir bar was flushed with argon, taking care to avoid bubbling of the protein solution. Using glass syringes, 40 µL of the 10X buffer solution was added under argon, along with water to a final volume of 380 µL. 10 µL of a 1-octene solution (1.2 M in EtOH) was then added via syringe and allowed to mix for 30 seconds prior to adding 10 µL of ethyl diazoacetate (400 mM in EtOH) for cyclopropanation reactions, or 10 µL tosyl azide (400 mM in EtOH) for aziridination reactions. The final concentrations of the reagents were: 20 µM catalyst, 30 mM styrene, and 10 mM ethyl diazoacetate (or tosyl azide) in 0.1 M potassium phosphate (pH 8.0). Heme enzyme reactions also included 10 mM sodium dithionite. The vial was removed from the argon line, the septum covered with parafilm, and the reaction was allowed to stir overnight at room temperature. Reactions were quenched by the addition of 1 mL of ethyl acetate to precipitate the protein. Then, 20 µL of an internal standard solution (100 mM cyclopropyl phenyl ketone in EtOH) was added. The mixture was transferred to a 1.7 mL microcentrifuge tube, vortexed, and centrifuged at 18,400 x g for 2 minutes. The ethyl acetate layer was then removed and dried over sodium sulfate in a separate microcentrifuge tube. The organic layer was then filtered through a glass wool plug into glass vials for GC-MS analysis, or evaporated off and the residue resuspended in MeOH for LC-MS analysis. Separations on GC-MS were achieved at constant pressure and a GC oven temperature gradient of 40°C for 1 min, 5 °C/min to 300 °C, and 300 °C for 3 min. Products were identified using Enhanced ChemStation Data Analysis software (Agilent) and the NIST Mass Spectral Search Program for the NIST/EPA/NIH Mass
Spectral Library. Separations on LC-MS were achieved with a mobile phase of water and acetonitrile + 0.1% formic acid, with a gradient of acetonitrile of 5% for 4 min, 5 – 95% over 11 min, 95% for 5 min and 5% for 5 min and the data was analyzed with Agilent MassHunter software.
## 3.5 Supplementary Information

### 3.5.1 Supplementary Tables

**Table S3.1.** Amino acid mutations present in BM3h variants used in this study.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Amino Acid Sequence/Mutations relative to Parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM3h (Parent)</td>
<td>MTIKEMPQPQTKFGEKLNPLLTDKPVQALMKIADGIEIFKFEAP GRVTRYLSSRLKIEACDESDFKNDLSQALKFVRFDAQGDQLTS WTBEKNWKAHNILLPSFSQQAMKGHYHAMVAVLQVQKWERN LNADEHIEVPDMTRLTDLTGFCGFNYRFNFYRDQHPFITSVM RALDEANMLKQRANPDDPAYDENKRQFQEDIKVMNLVDKIIADR KASGEQSDDLTHMLNGKDPETGEPLDDENYRIQIIITFLAGHETT SGLSFALYFLVKNPLQKAAEEARVLVDPVQQVKQVKLYV GMVLNEALRLWPTAPAFSLYAKEDTVLGEYPLEKGEDEMLVLIPQ LHRDKTIWGDDEVFRPERFENPSPAIPQHAFKFPGNGQRACIGQ QFALHEATVLGMMLKHDDEHTNYELDKETLTKPEGFVVKAK SKKAPIPLGLQPSPLLEHHHHHH</td>
</tr>
<tr>
<td>BM3-R1 (Parent)</td>
<td>BM3h (-blue portion) + EQSAKKVRKKAENHTPLLVLVGSNMGTAEGTARDLADIAMS KGFAPQVATLDHAGNLPREGAALVTASYNGHPPDNAPQFVDWL DQASADEVKVRYSVFCCGDKNWATYQVKPFIFLDLACKGAE NIAADRGEADASDDFEQGYEEWREHMWSDVAYFNLIDENSEN KSTLSLQFVSDAADMLAKMHGASTNVASKELQOQPSAARSH LEIENPKEASYQEGDHLGVIQPRNYEGIVNRVTRFGOLASGQIRLE AEEKLAHPLAKTSVEEQVETYQVELQDPVTRTQLRAMAKTVC PPHKVELEALLEKQAYEKEVLAKRTMLELEKYPACEMKFSEQUAL LPSIPRYSYISSSPRVDQEOQASITVSVSVEAWSGYEGYKIASN YLAELEQEGDITCFISTPQSEFTPLKDELPMVPGPTGVAPFRG FVQARKQLEKGGQSLGEAWLYFGCRPHSDYLYQEEAIAEQE SIITLHAFSRMPQPKTYQVHMEDGQDKLIEELQDQGAHYICG DGSQMAPAVEATLMKSYADVHVQSEADARLWLQLEEKGRYADCWVG</td>
</tr>
</tbody>
</table>
**Table S3.2.** UV-Vis absorbance of purified BM3h cofactor-substituted variants.

<table>
<thead>
<tr>
<th>Variant*</th>
<th>Condition</th>
<th>Soret band (nm)</th>
<th>Q bands (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM3h/heme</td>
<td>-</td>
<td>421</td>
<td>540, 576</td>
</tr>
<tr>
<td></td>
<td>+ Na₂S₂O₄</td>
<td>413</td>
<td>551 (broad)</td>
</tr>
<tr>
<td></td>
<td>+ Na₂S₂O₄ + CO</td>
<td>450</td>
<td>554 (broad)</td>
</tr>
<tr>
<td>WIVS-FM/Fe-DPIX</td>
<td>-</td>
<td>413</td>
<td>526, 552</td>
</tr>
<tr>
<td></td>
<td>+ Na₂S₂O₄</td>
<td>409</td>
<td>532 (broad)</td>
</tr>
<tr>
<td></td>
<td>+ Na₂S₂O₄ + CO</td>
<td>430</td>
<td>535 (broad)</td>
</tr>
<tr>
<td>BM3h T268A/Ir(Me)-DPIX</td>
<td>-</td>
<td>402, 438</td>
<td>518, 546</td>
</tr>
<tr>
<td>BM3h T268A C400A/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>508, 534</td>
</tr>
<tr>
<td>BM3h T268A C400G/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>508, 534</td>
</tr>
<tr>
<td>WIVS-FM T268A/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>508, 534</td>
</tr>
<tr>
<td>WIVS-FM T268A C400A/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>508, 534</td>
</tr>
<tr>
<td>WIVS-FM T268A C400A/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>508, 534</td>
</tr>
<tr>
<td>WIVS-FM T268A C400G/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>508, 534</td>
</tr>
<tr>
<td>WIVS-FM T268A C400G/Ir(Me)-DPIX</td>
<td>-</td>
<td>402</td>
<td>507, 534</td>
</tr>
</tbody>
</table>

*Variants with only one condition reported showed no change in the absorbance spectra upon addition of dithionite and exposure to CO.

**Table S3.3.** Gradient used for HPLC analysis of purified proteins.*

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>% A (H₂O + 0.1% formic acid)</th>
<th>% B (MeCN + 0.1% formic acid)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>10.01</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>17.01</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

* Separations were achieved on a Restek Viva C4 column (5 µm, 2.1 x 150 mm) with flow rate 0.4 mL/min.
3.5.2 Supplementary Figures

**Figure S3.1. Absorbance spectra of Ir(Me)-DPIX-containing proteins.** UV-Vis absorbance data is shown for the C400 (blue), C400A (violet) and C400G (green) variants of BM3h T268A (A) and WIVS-FM T268A (B). See Table S3.2 for the exact peak wavelengths.
Figure S3.2. HPLC analysis of Ir(Me)-DPIX content in protein preparations. Mol % Ir(Me)-DPIX (red) and heme (blue) for each variant expressed in minimal media (A) and TB (B) supplemented with Ir(Me)-DPIX, as determined by HPLC analysis of purified proteins. Error bars represent standard deviations determined from n=3 independent experiments. C) Sample HPLC traces of BM3h T268A C400A and WIV-FM* T268A C400A. See Figure 3.2 for HPLC traces of other variants. D) HPLC calibration curves for Ir(Me)-DPIX and heme (E) used to calculate the concentration of these cofactors in proteins.
Figure S3.3. GC-MS analysis of cyclopropanation reaction of 1-octene with EDA catalyzed by WIVS-FM* T268A C400G/Ir(Me)-DPIX. GC-MS chromatogram for the reaction is shown, along with the mass spectrum for the peaks at 16.9 and 17.4 min, which display a fragmentation pattern consistent with the cyclopropanation product 3.
Figure S3.4. $^1$H NMR spectrum of Ir(CO)Cl-deuteroporphyrin IX dimethyl ester.
Figure S3.5. $^1$H NMR spectrum of Ir(Me)-deuteroporphyrin IX.
REFERENCES


4.1 Introduction

Metal-catalyzed carbenoid transfer reactions have emerged as a powerful means of constructing carbon-carbon and carbon-heteroatom bonds in organic synthesis (1, 2). Among the many transition metal catalysts developed for these transformations, metalloporphyrins have received a great deal of attention due to their high turnover and selectivity (3–14). Inspired by the synthetic versatility of metalloporphyrins, protein engineers have demonstrated that enzymes utilizing the naturally-occurring metalloporphyrin cofactor, heme, can also catalyze carbenoid transfer reactions (15–17). Exploitation of this promiscuous enzymatic activity has led to the development of biocatalysts for non-natural reactions such as carbenoid-mediated olefin cyclopropanation and X-H insertion (X=N, S, Si) (15–23).

The catalytic properties of metalloporphyrins can be tuned by changing the metal center or through synthetic modification of the porphyrin ligand (5, 6, 9). On the other hand, protein engineers have traditionally relied on mutation of the protein scaffold to tune the activity and selectivity of enzymes (24, 25). Recently, some research groups have sought to combine both strategies by introducing synthetic heme analogs to proteins to create artificial metalloenzymes, the catalytic properties of which can be optimized through laboratory evolution (26, 27). For example, modifications to the heme protoporphyrin IX ligand and metal center have been shown to improve myoglobin sulfoxidation and C-H hydroxylation activity (28–30) and create P450 variants for carbenoid insertion into C-H bonds (31). Despite these successes, most methods for heme cofactor replacement require laborious and time-consuming *in vitro* manipulations, limiting the application of these systems. While several techniques for *in vivo* introduction of
heme analogs to proteins have been developed more recently, these efforts have thus far not produced artificial metalloenzymes with improved activity in non-natural reactions (32–38).

We have recently introduced a new methodology for the generation of cofactor-substituted hemoproteins in cells through the evolution of orthogonal enzyme/cofactor pairs (38). In this work, the nonproteinogenic metalloporphyrin, iron deuteroporphyrin IX (Fe-DPIX), was selectively incorporated in place of heme in an evolved cytochrome P450 scaffold. Although this was shown to be an effective means for constructing a non-natural metalloprotein \textit{in vivo}, the altered electronics of the deuteroporphyrin IX ligand resulted in decreased activity of the Fe-DPIX enzyme in carbenoid transfer reactions relative to the native heme enzyme (38). To overcome this challenge, the orthogonal enzyme was adapted to create an orthogonal artificial metalloenzyme harboring Ir(Me)-DPIX, thereby enabling expanded catalysis of carbenoid and nitrenoid transfer reactions (see Chapter 3).

While metal-substitution of the orthogonal enzyme/cofactor pair proved to be a useful strategy for improving non-natural catalysis, we envisioned that modifications to the porphyrin ligand could also serve to tune the catalytic properties of the cofactor, as has been achieved with synthetic metalloporphyrins (6, 7, 11). At the same time, these alternative porphyrin scaffolds could enable discrimination from native heme and the evolution of an orthogonal enzyme.

In validation of this approach, we demonstrate that aryl-modified heme derivatives (Figure 4.1) are improved carbenoid transfer catalysts relative to heme. These phenyl heme\textsuperscript{3} cofactors can be synthesized readily from heme and various substituted styrenes using a simple cross metathesis protocol. Furthermore, these heme analogs can be transported into the cell and incorporated into hemoproteins \textit{in vivo}. The bulky aryl-substituents provide the opportunity

\textsuperscript{3}To simplify writing and reading of this manuscript, we refer to the aryl-modified heme derivatives used in this study as phenyl hemes. Complete chemical names for these compounds based on IUPAC porphyrin nomenclature recommendations (63) are listed in Table S4.1.
to evolve orthogonal protein scaffolds for selective incorporation of these synthetic cofactors. This work further expands the diversity of cofactor structures that can be incorporated into hemoproteins, and represents a new paradigm in tuning heme enzyme chemistry for non-natural catalysis.

Figure 4.1. Aryl-substituted heme derivatives used in this study.

4.2 Results and Discussion

4.2.1 Aryl-modified hemes as non-natural cofactors for expanding heme enzyme catalysis

The aryl-modified heme analogs shown in Figure 4.1 were designed to meet several criteria, including the ability to incorporate the non-natural cofactors in vivo and tune the carbenoid transfer activity of heme enzymes. There have several successful examples of the incorporation of vinyl-modified hemes into proteins both in vitro and in vivo (26, 27, 33, 38–41). Therefore, we reasoned that introducing new functionality at the vinyl positions of heme would be ideal to enable cellular uptake of the cofactor analogs via the heme transport protein ChuA (which has been demonstrated to be promiscuous for alternative porphyrin scaffolds (38)) and incorporation into hemoproteins. We also drew inspiration from highly active synthetic porphyrin catalysts which are decorated with phenyl substituents at porphyrin meso-carbons (6, 11). By attaching phenyl groups at the vinyl positions of heme, we hoped to modulate catalysis in a similar manner. Furthermore, our design maintains conjugation between the styryl-substituents and the porphyrin macrocycle, allowing for tuning of the electronic properties of the cofactor by addition of electron-donating or –withdrawing groups to the phenyl ring. The increased steric bulk of the modified heme should also allow for engineering of proteins which utilize this
cofactor selectively, similar to other “bump-hole” strategies used to study kinases and methyltransferases (42, 43).

4.2.2 Synthesis of aryl-modified hemes

**Table 4.1.** Reaction optimization for cross-metathesis of metalated protoporphyrin IX dimethyl ester with styrene derivatives.

<table>
<thead>
<tr>
<th>entry</th>
<th>M</th>
<th>styrene derivative</th>
<th>styrene equivalents</th>
<th>catalyst (loading)</th>
<th>time (hr)</th>
<th>total conversion (%)</th>
<th>%monoaryl</th>
<th>% bisaryl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fe(Cl)</td>
<td>styrene</td>
<td>40</td>
<td>1 (25 mol%)</td>
<td>1</td>
<td>11</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Fe(Cl)</td>
<td>styrene</td>
<td>40</td>
<td>1 (25 mol%)</td>
<td>16</td>
<td>32</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Fe(Cl)</td>
<td>styrene</td>
<td>40</td>
<td>2 (25 mol%)</td>
<td>16</td>
<td>40</td>
<td>37</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Fe(Cl)</td>
<td>styrene</td>
<td>4</td>
<td>2 (25 mol%)</td>
<td>16</td>
<td>88</td>
<td>43</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>Fe(Cl)</td>
<td>styrene</td>
<td>4</td>
<td>2 (10 mol%)</td>
<td>16</td>
<td>32</td>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Fe(Cl)</td>
<td>4-OMe</td>
<td>4</td>
<td>2 (25 mol%)</td>
<td>16</td>
<td>21</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Fe(Cl)</td>
<td>4-CF₃</td>
<td>4</td>
<td>2 (25 mol%)</td>
<td>16</td>
<td>29</td>
<td>26</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>Fe(Cl)</td>
<td>2,3,4,5,6-pentafluoro</td>
<td>4</td>
<td>2 (25 mol%)</td>
<td>16</td>
<td>33</td>
<td>20 (13)²</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Zn</td>
<td>styrene</td>
<td>4</td>
<td>2 (25 mol%)</td>
<td>6</td>
<td>96</td>
<td>27</td>
<td>69</td>
</tr>
</tbody>
</table>

²conversion was determined by HPLC analysis of the crude reaction mixture, comparing the relative peak areas of porphyrin starting material and products.

²along with desired the pentafluorostyryl-substituted product, cross-metathesis product for the reaction between heme and the 1-isoproxy-2-vinylbenzene catalyst ligand was also detected by LC-MS.

With these design considerations established, we next searched for suitable procedures for the synthesis of phenyl heme cofactors. Bisphenyl heme (denoted as BPH, see Table 4.1 for chemical structure) derivatives have been prepared previously through Heck-type reactions (44,
45). However, these required long reaction times, produced a complex mixture of products that were difficult to separate, and did not provide access to monophenyl heme (denoted as MPH, see Table 4.1 for chemical structure) derivatives, which we wished to obtain as well to provide greater flexibility in future protein engineering work. Alternatively, several groups have demonstrated that vinyl-porphyrins are suitable substrates for cross-metathesis with a variety of olefins (46–49). These reactions were shown to be highly $E$ selective due to the steric hindrance provided by the porphyrin macrocycle (46–48), thereby reducing the complexity of the product mixture. This strategy could also provide access to both MPH and BPH in a single reaction. Therefore, we sought to develop a cross-metathesis procedure for the synthesis of phenyl hemes (Table 4.1).

To begin, iron(III) protoporphyrin IX dimethyl ester chloride was obtained in quantitative yield from commercially available hemin by treatment with trimethyloxonium tetrafluoroborate (50). This was then used in cross-metathesis reactions with various styrene derivatives. Using previously reported reaction conditions (48) with the Grubb’s 2nd generation metathesis catalyst 1 provided MPH with 10% conversion (Table 4.1, entry 1). Monitoring the reaction time course revealed that total conversion improves upon continuing the reaction overnight (Table 4.1, entry 2 and Figure S4.1B). Use of the Hoveyda-Grubb’s 2nd generation catalyst 2 led to a further increase in conversion, and, upon lowering the equivalents of styrene, provided 88% conversion overall with approximately equal amounts of MPH and BPH (Table 4.1, entries 3 and 4, and Figure S4.1A). Attempts to lower the catalyst loading from 25 mol% to 10 mol% resulted in lower conversions (Table 4.1, entry 5). Following the reaction, MPH and BPH products could be separated via reverse phase chromatography. The two regioisomers of MPH that result from reaction on either of the two heme vinyl groups were not separated, though we envision that engineering of enzymes for selective recognition of MPH cofactors will provide enrichment for one isomer over the other during protein incorporation.
With these optimized reaction conditions, we then attempted cross-metathesis between hemin dimethyl ester and the aryl olefins 4-methoxystyrene, 4-trifluoromethylstyrene and 2,3,4,5,6-pentafluorostyrene. These substituents on the phenyl ring provide a range of electron-donating or –withdrawing effects to tune the electronics of the cofactor. Cross-metathesis with these styrene derivatives was less effective than with unsubstituted styrene, with conversions between 21-33% (Table 4.1, entries 6-8). Again, the products were purified by reverse phase chromatography, although the bis-substituted derivatives could not be isolated in high enough yield for characterization. The methoxy-, trifluoromethyl-, and pentafluorophenyl-substituted heme derivatives (denoted as OMe-MPH, CF₃-MPH and F₅-MPH, respectively) display shifted Soret absorbance bands compared to heme and unsubstituted MPH and BPH (Figure 4.2), suggesting that these substituents are influencing the electronic properties of the porphyrin.

![Absorbance spectra of phenyl heme derivatives.](image)

**Figure 4.2. Absorbance spectra of phenyl heme derivatives.** UV-Vis absorbance spectra of the pyridine hemochrome species of heme derivatives used in this study, with the peak wavelength of the Soret band given in parentheses in the legend.

Finally, we investigated the reactivity of zinc(II) protoporphyrin IX dimethyl ester in cross-metathesis with styrene, as excellent conversions had been observed in cross-metathesis reactions with this metalloporphyrin previously (46–48). As expected, the reaction with the zinc porphyrin proceeded more readily than with the iron counterpart, reaching 96% conversion in six
hours (Table 4.1, entry 9 and Figure S4.1C). Despite the increase in conversion observed with zinc(II) protoporphyrin IX, these zinc porphyrins cannot catalyze carbenoid transfer reactions and thus further steps would be required to replace the zinc center with iron or other metals. The final step in the synthesis of phenyl hemes was the hydrolysis of the methyl esters, yielding phenyl heme dicarboxylic acid species for use in catalysis and biological experiments.

4.2.3 Phenyl hemes are active catalysts for carbenoid-mediated olefin cyclopropanation

The generation of a panel of aryl-substituted heme analogs permits the investigation of the role of heme electronics in catalysis of non-natural reactions, a topic which thus far has not been systematically studied. Therefore, we aimed to assess the catalytic properties of phenyl heme cofactors in non-natural carbenoid-mediated olefin cyclopropanation reactions, a reaction which heme has been shown to promiscuously catalyze (15). The observance of improved carbenoid transfer activity with the free phenyl heme catalysts would suggest that enzymes harboring these modified cofactors would also possess expanded activity, and justify work towards protein incorporation of phenyl hemes and evolution of an orthogonal enzyme.

While we observed good conversions in cross-metathesis reactions, the isolated yields of pure phenyl hemes was often low (≤ 10 mg) due to challenges with the purification and the high cost of the metathesis catalyst, which limited the scale of the reaction. This, along with the notorious ability of porphyrins to retain solvent and other minor impurities (51), made obtaining accurate catalyst loading through mass measurements difficult. However, we reasoned that the concentration of phenyl heme solutions could be determined by NMR, by comparing peak areas for the porphyrin to that of an internal standard of known concentration. The measured concentrations by NMR could then be used to calculate the extinction coefficients of phenyl hemes for future concentration determination via absorbance measurements. This work is ongoing, but preliminary experiments have yielded reasonable extinction coefficient values for MPH and BPH (see Materials and Methods). Although additional work is needed to confirm these values and determine the extinction coefficients for the other phenyl heme derivatives, we
moved forward with determining the catalytic properties of MPH and BPH in carbenoid-mediated olefin cyclopropanation.

Figure 4.3. Carbenoid-mediated olefin cyclopropanation catalyzed by phenyl heme derivatives. Gas chromatogram for the cyclopropanation of styrene with EDA catalyzed by heme (red), MPH (blue) and BPH (green).

Examination of the activity of heme, MPH and BPH in the model cyclopropanation reaction between styrene and ethyl diazoacetate (EDA) revealed that MPH and BPH catalyze the reaction with an approximately 4- and 5-fold enhancement in turnover number relative to heme, respectively (Figure 4.3). While modest, this enhancement is encouraging in that it shows that modifying the porphyrin architecture is a viable means of tuning non-natural catalysis with heme cofactors. Furthermore, the introduction of phenyl heme cofactors to proteins will provide additional opportunities to enhance activity through mutagenesis. Through this approach, we hope to access new reactivity not accessible with native heme, for example, cyclopropanation with more challenging acceptor/acceptor diazo compounds (10, 52–54).

Future work will include characterizing the cyclopropanation activity of OMe-MPH, CF₃-MPH, and F₅-MPH to determine the effects of the electron-donating and –withdrawing groups on catalysis.

4.2.4 In vivo incorporation of phenyl heme cofactors into hemoproteins

Encouraged by the preliminary results in the catalysis of non-natural cyclopropanation
reactions, we next attempted *in vivo* incorporation of phenyl hemes into hemoproteins. Cytochrome P450\textsubscript{BM3}, CYP119, and myoglobin (Mb) were tested for their ability to accommodate MPH and BPH, as these hemoproteins have been used previously as scaffolds to harbor non-natural cofactors (29–31, 38). In the case of CYP119 and Mb, mutations were introduced that have previously been found to increase carbenoid transfer activity of these proteins (20, 31). To test for cellular uptake of phenyl hemes and incorporation into proteins, the heme transport protein ChuA and the hemoprotein of interest were coexpressed in *E. coli* grown in iron-deficient minimal media, supplemented with either 10 µM MPH or BPH. The proteins isolated from these expressions were analyzed by high performance liquid chromatography (HPLC) to determine phenyl heme and heme content. Although BPH was poorly incorporated (*Figure S4.2C*), all three proteins showed ≥ 84% incorporation of MPH under these conditions (*Figure 4.4A and Figure S4.2*).

![Figure 4.4. Incorporation of MPH into hemoproteins.](image)

A) Sample HPLC traces of hemoproteins expressed in minimal media in the presence of MPH, shown in comparison with MPH and heme standards. B) Absorbance spectra of purified ferrous CO-bound CYP119 T213A in complex with heme (blue) or MPH (green). C) Absorbance spectra of CYP119 T213A in complex with ferric heme (blue) or Zn-MPH (gold).
These results demonstrate that ChuA can import bulky vinyl-modified heme analogs and that wild type heme binding pockets are capable of accommodating the MPH cofactor, providing a starting point to evolve orthogonal protein scaffolds.

4.2.5 Evolution of a MPH-specific protein

Due to the high thermostability of CYP119 (55), this hemoprotein was chosen as the starting scaffold for evolution of an orthogonal MPH enzyme. The mutation T213A was included as this was shown to improve carbenoid transfer activity of the heme enzyme (31). In order to screen for variants with improved MPH selectivity, we aimed to develop a spectroscopic assay based on the shift in UV-Vis absorbance between the heme and MPH cofactor (Figure 4.2), similarly to the screen used in the evolution of an orthogonal Fe-DPIX enzyme (38).

Reduced, CO-bound CYP119 T213A/MPH displays a Soret band absorbance at 452 nm that is slightly red-shifted from CYP119 T213A/heme, which has a peak at 448 nm (Figure 4.4B). While this bathochromic shift is consistent with that observed with the free cofactors (Figure 4.2), the magnitude of the shift in the protein absorbance spectra is too small to distinguish between heme- and MPH-bound species in a mixture. However, we reasoned that changing the iron metal center of MPH could enable deconvolution of MPH- and heme protein spectra. For example, cytochrome P450 reconstituted with Zn-protoporphyrin IX shows an approximately 30 nm red-shift in Soret band absorbance compared to the unreduced heme enzyme (34). We hypothesized that incorporation of zinc monophenyl heme (Zn-MPH) into CYP119 would produce a similar shift in the absorbance spectrum of the enzyme and allow us to better distinguish Zn-MPH-bound CYP119 from the native heme protein.

To test this hypothesis, Zn-MPH was prepared and incorporated into CYP119 T213A in vivo. The Zn-MPH enzyme shows a Soret band absorbance at 456 nm that is shifted 35 nm relative to the ferric heme enzyme (Figure 4.4C). The shoulder at 431 nm in the absorbance spectrum of CYP119 T213A/Zn-MPH likely represents a misligated form of the enzyme in which a histidine residue is coordinating the Zn cofactor rather than the native cysteine, as this peak
wavelength is similar to zinc myoglobin variants which natively possess an axial histidine residue (56). Nonetheless, the large Soret band shift observed in CYP119 T213A/Zn-MPH provides a convenient spectroscopic assay to identify variants with improved selectivity for the MPH cofactor.

The vinyl pockets of CYP119 are composed primarily of hydrophobic amino acids which provide a complementary binding surface for native heme. We reasoned that mutations in these pockets to small hydrophobic or aromatic amino acids would result in better accommodation of the bulky phenyl substituent of MPH. Therefore, we created limited diversity libraries in these pockets, randomizing residues to 10 amino acids – Ala, Gly, Leu, Ile, Met, Val, His, Phe, Tyr, Trp – at each position. Because the phenyl group of MPH could be oriented in either vinyl pocket, we focused libraries to two positions on each side of the active site (Figure 4.5A).

![Figure 4.5. Engineering CYP119 for selective recognition of MPH. A) Crystal structure of wild type CYP119 (green, PDB: 1IO7 (57)) highlighting side chains targeted for mutation (blue) as well as the heme cofactor (yellow). B) HPLC analysis of purified CYP119 I130A T213A and CYP119 T213A isolated from expressions in TB media supplemented with 10 µM MPH.](image)

Out of this 10,000 variant library, only ~500 clones were screened as a result of possessing only limited amounts of the Zn-MPH cofactor. These clones were grown in minimal media supplemented with 10 µM Zn-MPH and 1 µM heme, which was determined to be an ideal ratio of the cofactors to obtain roughly equal amounts of heme- and Zn-MPH-bound protein. Following protein expression, absorbance measurements revealed several variants with a small increase in Zn-MPH binding (Figure S4.3). Sequencing showed that all of these variants
converged on the mutation I130A (Table S4.2). HPLC analysis of the CYP119 I130A T213A variant expressed in rich media (i.e., in the presence of endogenous heme) supplemented with 10 µM MPH showed an increase in MPH incorporation compared to CYP119 T213A (54% vs. 65% MPH, Figure 4.5B), corroborating the absorbance data obtained from expressions with Zn-MPH. These initial results represent a promising step forward in the quest for an orthogonal MPH enzyme. Future work will aim to continue improving MPH incorporation by focusing libraries in the pocket around the I130A mutation (Figure 4.6).

**Figure 4.6. Library design for improving MPH selectivity.** The I130A (purple) modeled into the wild type CYP119 structure (green, PDB: 1IO7), with the surrounding amino acid residues targeted for mutation highlighted in blue. One regioisomer of the MPH heme cofactor (yellow) is shown modeled into the structure as well.

### 4.3 Conclusions and Future Work

With the goal of expanding nature’s catalytic repertoire, we have designed novel aryl-substituted heme derivatives with improved activity in non-natural carbenoid transfer reactions. Inspired by synthetic catalysts, the modifications introduced to the porphyrin scaffold provide a means of tuning the activity of the cofactor through the addition of electron-withdrawing and donating groups. In contrast to other non-natural heme analogs that require laborious, multi-step synthetic routes starting from pyrrole building blocks (26, 28, 29), the phenyl heme derivatives we developed are readily synthesized via cross-metathesis reactions using commercially available hemin. Furthermore, the synthesis is modular, with the ability to use a variety of
vinylaryl compounds as the olefin partner for cross-metathesis. Currently, work is underway to prepare a panel of phenyl heme cofactors and characterize their activity in non-natural carbenoid transfer reactions.

We found that MPH is imported into cells by the heme transport protein ChuA and incorporated into proteins. This shows that ChuA can provide broadened access to heme analogs with bulky substituents at the vinyl positions of heme. The bis-substituted cofactor, however, was not found to be incorporated into proteins in vivo. Presumably, the additional steric bulk of BPH presents challenges to either uptake by ChuA or incorporation into the heme-binding pocket of the enzyme. Once the cause of poor BPH incorporation is identified, engineering of ChuA, the hemoprotein, or both can be performed in order to enable more efficient construction of BPH enzymes in cells.

The ability to generate MPH enzymes in vivo allowed us to begin evolving an orthogonal protein scaffold for selective MPH recognition. Using an absorbance-based screen with a Zn-substituted MPH cofactor, a variant was identified with an improvement in MPH binding. The mutation introduced, I130A, presumably favors MPH incorporation by expanding the pocket to accommodate the phenyl group. Future work will include screening additional libraries in the active site to attempt to enhance MPH selectivity. Further expansion of the pocket, however, may limit the extent to which the enzyme can discriminate between heme and MPH. A potential solution to this problem could be modification of the other vinyl group of MPH to create an asymmetrically-substituted heme analog with two different handles for engineering selectivity. In any case, this work is encouraging in that it suggests that the generation of orthogonal enzyme/cofactor pairs is a general strategy for expanding cofactor diversity that can be expanding to a variety of cofactor structures and enzymes.

The introduction of new functionality to a non-natural cofactor that simultaneously tunes the catalytic activity and provides a discriminating feature for selective protein recognition is a powerful strategy for expanding heme enzyme catalysis. We envision this methodology can be
further developed to provide a panel of orthogonal enzyme/cofactor pairs that are each tuned
towards a specific purpose. The ability to substitute the metal center of the cofactor provides an
additional degree of freedom for diversifying the function of these systems. These orthogonal
enzyme/cofactor pairs may then be combined into biosynthetic pathways in cells, or further
evolved to create highly active and selective non-natural biocatalysts.

4.4 Materials and Methods

4.4.1 General methods and materials

Hemin and zinc(II) protoporphyrin IX were obtained from Frontier Scientific.
Oligonucleotides were obtained from Integrated DNA Technologies. Enzymes and reagents
used for cloning were obtained from New England BioLabs. All other chemical reagents and
solvents were obtained from chemical suppliers (Acros, Fisher Scientific, or Sigma-Aldrich) and
used without further purification.

Reactions requiring inert atmosphere were carried out in oven-dried glassware using
standard Schlenk techniques under nitrogen. Reverse phase chromatography of phenyl hemes
was performed on an automated Biotage Isolera One using 120 g SNAP Ultra C18 columns.

Iron(III) porphyrin samples were prepared for NMR by dissolving the porphyrin (2-4 mg)
in pyridine-d$_5$ in a 2 mL crimp vial under argon. A small amount of SnCl$_2$ (~1-2 mg) was then
dissolved in pyridine-d$_5$ in a separate vial under argon. The SnCl$_2$ solution was added to the
porphyrin vial via syringe. The mixture was then transferred to a dry NMR tube under argon via
syringe and the tube was sealed with parafilm. Proton magnetic resonance spectra were
acquired on 400 MHz Bruker instruments in the Department of Chemistry NMR facility at the
University of North Carolina – Chapel Hill. NMR spectra were processed with ACD/NMR
Processor (ACD/Labs). Chemical shifts are referenced to residual solvent peaks (58).

GC analyses were carried out on a 7820A gas chromatograph with flame ionization
detector (Agilent), using a J&W Scientific Cyclosil-B column (30 m x 0.25 mm x 0.25 µm film).
LC-MS analyses were obtained with an Agilent 6520 Accurate Mass QTOF LC-MS ESI positive
in high-resolution mode, equipped with a Restek Ultra biphenyl column (5 µm particle size, 3 x 50 mm), using a gradient of acetonitrile in water + 0.1% formic acid from 5-85% over 10 min, 85% for 5 min, 95% for 4 min, and 5% for 1 min.

4.4.2 Organic synthesis and characterization

A. Synthesis of iron(III) protoporphyrin IX dimethyl ester chloride (hemin dimethyl ester) and zinc(II) protoporphyrin IX dimethyl ester

Iron(III) protoporphyrin IX dimethyl ester chloride was prepared by treatment of commercially available hemin (Frontier Scientific) with trimethyloxonium tetrafluoroborate, as previously described (50). Zinc(II) protoporphyrin IX dimethyl ester was prepared in a similar way starting from commercially available zinc(II) protoporphyrin IX (Frontier Scientific).

B.i. General procedure for preparation of aryl-modified heme derivatives.

Procedures for the synthesis of aryl-modified heme derivatives were adapted from Dolphin and coworkers (48). Iron(III) protoporphyrin IX dimethyl ester chloride (100 mg, 0.15 mmol) was added to an oven-dried flask with condenser under nitrogen. In a separate oven-dried flask, the Hoveyda-Grubb’s 2nd generation catalyst 2 (23 mg, 0.037 mmol) was dissolved in anhydrous THF (3 mL). The catalyst solution was added to the porphyrin via syringe and the mixture was stirred at room temperature. The styrene derivative (0.6 mmol) was immediately added to the stirred mixture via syringe, and the reaction was gently refluxed under nitrogen overnight. Following the reaction, solvent was removed by rotary evaporation. The compounds were purified by reverse phase chromatography. The mobile phase for separation consisted of water + 0.1% formic acid (Solvent A) and acetonitrile + 0.1% formic acid (Solvent B), with iron phenyl hemes being separated using a gradient of 5 - 65% B over 1 CV, 65 – 90% over 11 CV, and 100% B for 3 CV. Zinc phenyl hemes were separated using a gradient of 5 – 80% B over 1 CV, 80 – 100% B over 12 CV, and holding at 100% B for 5 CV. Hydrolysis of the methyl esters was performed prior to biological or catalytic experiments.
B.ii. Synthesis of Fe(III) 3-styryl-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-styryl-deuteroporphyrin dimethyl ester chloride and Fe(III) 3,8-distyryl-deuteroporphyrin dimethyl ester chloride (monophenyl heme (MPH) and bisphenyl heme (BPH))

Following the general reaction procedure, Fe(III) protoporphyrin IX dimethyl ester chloride (100 mg, 0.15 mmol), catalyst 2 (23 mg, 0.037 mmol) and styrene (62 mg, 69 µL, 0.6 mmol) were reacted overnight. The reaction mixture was purified by reverse phase chromatography to separate unreacted hemin dimethyl ester, monophenyl heme and bisphenyl heme. Separation of 3- vs. 8-styryl-substituted regioisomers of monophenyl heme could not be achieved. Following purification, solvent was evaporated to yield pure products as a black-brown solid, typically ≤ 10 mg.

Fe(III) 3-styryl-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-styryl-deuteroporphyrin dimethyl ester chloride. \( ^1H \text{NMR (pyridine-}d_5 + \text{SnCl}_2, 400 \text{MHz)} \delta 10.42 \text{ (s, 1H), 10.32 (s, 1H), 10.28 (s, 1H), 10.16 (s, 1H), 10.08 (d, J = 3.0 Hz, 2H), 10.00 (d, J = 8.8 Hz, 2H), 9.09-8.96 (m, 2H), 8.59-8.48 (m, 2H), 8.08 (d, J = 7.0 Hz, 4H), 7.89 (d, J = 6.8 Hz, 1H), 7.85 (d, J = 6.5 Hz, 1H), 7.44-7.34 (m, 6H), 6.39-6.30 (m, 2H), 6.04-5.98 (m, 2H), 4.45-4.39 (m, 8H), 3.76 (d, J = 6.5 Hz, 6H), 3.68-3.65 (m, 6H), 3.52-3.47 (m, 24H), 3.43-3.40 (m, 8H) \text{ppm.} \)

Fe(III) 3,8-distyryl-deuteroporphyrin dimethyl ester chloride. $^1$H NMR (pyridine-d$_5$ + SnCl$_2$, 400 MHz) $\delta$ 10.45 (s, 1H), 10.29 (s, 1H), 10.08 (s, 1H), 10.02 (s, 1H), 9.11-8.99 (m, 2H), 8.09 (d, J = 7.5 Hz, 2H), 8.05 (d, J = 7.3 Hz, 2H), 7.91-7.84 (m, 2H), 7.55 (m, 2H), 7.44-7.37 (m, 4H), 4.42 (q, J = 7.4 Hz, 4H), 3.76 (d, J=9.5 Hz, 6H), 3.52-3.47 (m, 12H), 3.45-3.39 (m, 4H) ppm.


**B.iii. Synthesis of Fe(III) 3-(4-methoxystyryl)-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-(4-methoxystyryl)-deuteroporphyrin dimethyl ester chloride (OMe-MPH)**

Following the general reaction procedure, Fe(III) protoporphyrin IX dimethyl ester chloride (100 mg, 0.15 mmol), catalyst 2 (23 mg, 0.037 mmol) and 4-methoxystyrene (81 mg, 80 $\mu$L, 0.6 mmol) were reacted overnight. The reaction mixture was purified by reverse phase chromatography to separate unreacted hemin dimethyl ester and OMe-MPH. Following purification, solvent was evaporated to yield pure products as a black-brown solid, typically ≤ 10 mg. The product was contaminated with some of the stilbene byproduct, which made the $^1$H NMR analysis challenging. Future work will include confirming chemical shifts with pure samples.

Fe(III) 3-(4-methoxystyryl)-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-(4-methoxystyryl)-deuteroporphyrin dimethyl ester chloride. $^1$H NMR (pyridine-d$_5$ + SnCl$_2$, 400 MHz) $\delta$ 10.41 (s, 1H), 10.28 (d, J = 9.8 Hz, 2H), 10.14 (s, 1H), 10.06 (m, 2H), 9.99-9.97 (m, 2H),
8.92 (t, 2H), 8.57-8.43 (m, 2H), 8.03 (d, J = 8.0 Hz 4H), 7.83 (dd, J = 16.4 Hz, 2H), 7.59 (m, 4H), 6.37-6.28 (m, 2H), 6.01-5.95 (m, 2H), 4.41-4.38 (m, 8H), 3.79-3.62 (m, 21H), 3.50-3.45 (m, 21H), 3.41-3.37 (m, 21H), ppm. Calculated [M-Cl]$^+$ = 750.2505 found [M-Cl]$^+$ = 750.2474.

B.iv. Synthesis of Fe(III) 3-(4-trifluoromethylstyryl)-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-(4-trifluoromethylstyryl)-deuteroporphyrin dimethyl ester chloride (CF$_3$-MPH)

Following the general reaction procedure, Fe(III) protoporphyrin IX dimethyl ester chloride (100 mg, 0.15 mmol), catalyst 2 (23 mg, 0.037 mmol) and 4-trifluoromethylstyrene (103 mg, 89 µL, 0.6 mmol) were reacted overnight. The reaction mixture was purified by reverse phase chromatography to separate unreacted hemin dimethyl ester and CF$_3$-MPH. Following purification, solvent was evaporated to yield pure products as a black-brown solid, typically ≤ 10 mg. The product was contaminated with some of the stilbene byproduct, which made the $^1$H NMR analysis challenging. Future work will include confirming chemical shifts with pure samples.

Fe(III) 3-(4-trifluoromethylstyryl)8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-(4-trifluoromethylstyryl)-deuteroporphyrin dimethyl ester chloride. $^1$H NMR (pyridine-d$_5$ + SnCl$_2$, 400 MHz) δ 10.39 (s, 1H), 10.29 (s, 1H), 10.25 (s, 1H), 10.11 (s, 1H), 10.04 (d, J = 1.8 Hz, 2 H) 9.99 (s, 1H), 9.94 (s, 1H), 9.15-9.07 (m, 2H), 8.48-8.40 (m, 2H), 8.10 (d, J = 8.0 Hz, 2H), 7.90
(d, J = 11.0 Hz, 2H), 7.82-7.77 (m, 4H), 7.74-7.70 (m, 2H), 6.35-6.27 (m, 2H), 6.00-5.95 (m, 2H), 4.39-4.35 (m, 8H), 3.74 (d, J = 7.5 Hz, 6H), 3.63 (d, J = 2.5 Hz, 6H) 3.48-3.3.43 (m, 18H), 3.39-3.35 (m, 8H) ppm.

$^{19}$F NMR (pyridine-$d_5$ + SnCl$_2$ + 0.1% TFA, 500 MHz) δ -161.77 (s, 3F), -162.00 (s, 3F) ppm.


B.v. Synthesis of Fe(III) 3-(2,3,4,5,6-pentafluorostyryl)-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-(2,3,4,5,6-pentafluorostyryl)-deuteroporphyrin dimethyl ester chloride ($F_5$-MPH)

Following the general reaction procedure, Fe(III) protoporphyrin IX dimethyl ester chloride (100 mg, 0.15 mmol), catalyst 2 (23 mg, 0.037 mmol) and 2,3,4,5,6-pentafluorostyrene (116 mg, 83 µL, 0.6 mmol) were reacted overnight. The reaction mixture was purified by reverse phase chromatography to separate unreacted hemin dimethyl ester and $F_5$-MPH. Following purification, solvent was evaporated to yield pure products as a black-brown solid, typically ≤ 10 mg.

Fe(III) 3-(2,3,4,5,6-pentafluorostyryl)-8-vinyl-deuteroporphyrin dimethyl ester chloride/Fe(III) 3-vinyl-8-(2,3,4,5,6-pentafluorostyryl)-deuteroporphyrin dimethyl ester chloride. $^1$H NMR (pyridine-$d_5$ + SnCl$_2$, 400 MHz) δ 10.35 (s, 1H), 10.31 (s, 1H), 10.21 (s, 1H), 10.11 (s, 1H), 10.04 (d, J = 2.8 Hz, 2H), 10.01 (s, 1H), 9.94 (s, 1H), 9.19 (dd, J = 16.7 Hz, 10.7 Hz; 2H), 8.55-8.42 (m, 2H),
7.79-7.75 (m, 2H), 6.35-6.29 (m, 2H), 6.00-5.95 (m, 2H), 4.40-4.36 (m, 8H), 3.76 (d, J = 10.8 Hz, 6H), 3.68 (s, 3H), 3.63-6.58 (m, 9H), 3.51-3.43 (m, 18H), 3.39-3.35 (m, 8H), ppm. Calculated [M-Cl]+ = 810.1928 found [M-Cl]+ = 810.1927.

B.vi. Synthesis of Zn(II) 3-styryl-8-vinyl-deuteroporphyrin dimethyl ester chloride/Zn(II) 3-vinyl-8-styryl-deuteroporphyrin dimethyl ester chloride and Zn(II) 3,8-distyryl-deuteroporphyrin dimethyl ester chloride (monophenyl heme (MPH) and bisphenyl heme (BPH))

Following the general reaction procedure, Zn(II) protoporphyrin IX dimethyl ester chloride (100 mg, 0.15 mmol), catalyst 2 (23 mg, 0.037 mmol) and styrene (62 mg, 69 µL, 0.6 mmol) were reacted overnight. The reaction mixture was purified by reverse phase chromatography to separate unreacted Zn protoporphyrin IX dimethyl ester, Zn-MPH, and Zn-BPH. Following purification, solvent was evaporated to yield pure products as a purple-black solid, typically ≤ 10 mg.

Zn(II) 3-styryl-8-vinyl-deuteroporphyrin dimethyl ester chloride/Zn(II) 3-vinyl-8-styryl-deuteroporphyrin dimethyl ester chloride. 1H NMR (CDCl3, 400 MHz) δ 9.41-8.98 (m, 8H), 8.10-7.96 (m, 6H), 7.68 (t, J = 7.4 Hz, 5H), 7.55-7.40 (m, 5H), 6.26-6.06 (m, 4H), 4.51-4.46 (m, 4H), 4.16-4.11 (buried under EtOAc, 8H), 3.66 (d, J = 4.8 Hz, 12H), 3.41-3.40 (m, 6H), 3.34-3.28 (m, 18H), 3.09-3.03 (m, 8H) ppm. Calculated [M]+ = 728.2340 found [M]+ = 728.2338.
Zn(II) 3,8-distyryl-deuteroporphyrin dimethyl ester chloride. $^1$H NMR (CDCl$_3$, 400 MHz) $\delta$ 9.33 (s, 1H), 9.22 (s, 1H), 9.07 (s, 1H), 8.96 (s, 1H), 9.18 (s, 1H), 8.34-8.30 (m, 1H), 8.10-8.06 (m, 1H), 8.02-7.96 (m, 4H), 7.68 (t, $J$ = 7.5 Hz, 3H), 7.57-7.51 (m, 3H), 7.45-7.36 (m, 2H), 4.19-4.13 (m, 4H), 3.68-3.66 (m, 6H), 3.36-3.33 (m, 9H), 3.24 (s, 3H), 3.11-3.06 (m, 4H) ppm. Calculated [M$^+$] = 804.2654 found [M$^+$] = 804.2635

C. Saponification of porphyrin dimethyl esters

Saponification of the dimethyl esters was achieved by dissolving the purified aryl-substituted heme dimethyl ester (10 mg) in a 95:5 MeOH/water mixture (10 mL). Potassium hydroxide (350 mg), was added and the reaction was gently refluxed with stirring under a nitrogen-filled balloon overnight. The reaction was diluted in DCM (20 mL) and washed with 2 x 50 mL 0.7 M HCl. The organic layer was collected and washed with water (50 mL) and brine (50 mL), then dried over sodium sulfate. The solvent was evaporated and the residue was dissolved in 0.1 M NaOH or DMSO for catalysis and biological experiments.

Hydrolysis of zinc(II) porphyrin dimethyl esters was performed in a similar manner, however, the sensitivity of zinc porphyrins to demetalation under acidic conditions necessitated a different workup protocol. In this case, solvent was evaporated following the reaction and the resulting residue was redissolved in water for downstream use.

4.4.3 Pyridine hemochrome assay

Heme derivatives dissolved in 0.1 M NaOH were diluted in a 3:17:10 mixture of 1 M NaOH, H$_2$O, and pyridine, to a concentration of approximately 5 µM. To this solution, 10 µL of 400 mM sodium ditionite solution in water was added. The absorbance was measured on a Cary 100 UV/Vis spectrometer (Agilent).

4.4.4 Determination of phenyl heme extinction coefficients

Phenyl heme derivatives were prepared for NMR as described in the general methods, with the addition of a known concentration of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as an internal standard. Comparison of the integration values for the DSS standard and the
phenyl heme peaks allowed for determination of phenyl heme concentration in the NMR sample. Following NMR analysis, the absorbance of the sample was measured by the pyridine hemochrome assay described above. These data were then used to calculate molar extinction coefficient values: MPH: \( \varepsilon_{427 \text{ nm}} = 198 \text{ mM}^{-1}\text{cm}^{-1} \), BPH: \( \varepsilon_{434 \text{ nm}} = 206 \text{ mM}^{-1}\text{cm}^{-1} \). These values are comparable to the literature value for heme: \( \varepsilon_{419 \text{ nm}} = 191.5 \text{ mM}^{-1}\text{cm}^{-1} \) (39).

4.4.5 Plasmids and library construction

Genes for P450 variants were cloned, as previously described (59) into plasmid pCWori. Myoglobin variants were cloned into pET21c(+) (available from Addgene). Heme transport protein, ChuA, was expressed from plasmid pChuA (available from Addgene) under the control of a T7 promoter, or from the plasmid pEvol (60) under the control of an arabinose-inducible promoter. The pEvol-ChuA construct replaces both copies of the tRNA synthetase and tRNA genes in the original pEvol vector with the ChuA gene.

Hydrophobic pocket libraries were constructed through standard overlap extension PCR assembly of PCR fragments. Forward primers containing library mutations were designed using DC Analyzer (61) to include degenerate codons that encode only hydrophobic residues (G, W, F, L, I, M, H, A, V, and Y). Site-specific mutants and single site-saturation libraries were assembled via scarless restriction ligation of BsaI digested PCR fragments containing the desired mutation or an NNK codon (codes for all 20 amino acids and a TAG stop codon). Primer sequences are available upon request.

4.4.6 Protein expression

Native hemoproteins were expressed as previously described (16). Expression of BM3 variants containing phenyl heme cofactors was accomplished by slight modification of a previously described methodology, using the vectors pCWori and pChuA (38). In the case of Mb and CYP119 variants, ideal expression was found to occur with the vectors pET21c(+) (for Mb), or pCWori (for CYP119), in combination with a pEvol-ChuA vector construct that is lower copy.
than pChuA. The expression protocol for CYP119/MPH is detailed below, but can be adapted for other combinations of hemoproteins and MPH/BPH cofactor derivatives.

CYP119 variants in pCWori were co-transformed with pEvol-ChuA into electrocompetent *E.coli* BL21 (DE3) Gold cells (Novagen). For expressions in minimal media, individual colonies were inoculated into liquid starter cultures of M9 minimal media supplemented with 0.5% glucose, 0.2% casamino acids, 340 µg/mL thiamine hydrochloride, 2 mM MgSO₄, 0.1 mM CaCl₂, 40 µg/mL chloramphenicol and 100 µg/mL ampicillin. For rich media expressions, colonies were inoculated into liquid starter cultures of LB media supplemented with 40 µg/mL chloramphenicol and 100 µg/mL ampicillin. The cultures were grown overnight at 37 °C and 225 rpm. One liter of M9 minimal media or TB supplemented with 40 µg/mL chloramphenicol and 100 µg/mL ampicillin was inoculated with 5 mL of overnight starter culture (the M9 or LB culture, respectively) and incubated for 3-4 hours at 37 °C and 225 rpm to an OD₆₀₀nm of 0.6-0.7. The cultures were then induced by adding arabinose (to induce ChuA expression) and Isopropyl β-D-thiogalactopyranoside (IPTG, to induce expression of CYP119) to a final concentration of 0.2% and 0.5 mM, respectively. A MPH stock solution in 0.1 M NaOH was added to the culture upon induction to a final concentration of 10 µM. Upon induction, the incubation temperature was reduced to 25 °C and expressions were allowed to continue for 16-20 hours. After harvesting the cells by centrifugation (4 °C, 10 min, 4,000 × g), the cell pellets were frozen at -20 °C for at least 2 hours.

4.4.7 Protein purification

Proteins were purified via 6 × His-tag purification on a Ni-NTA HisTrap HP column (5 mL HisTrap HP, GE Healthcare, Piscataway, NJ) using an ATKA Purifier (GE Healthcare), as described elsewhere (38). Following purification, the protein was exchanged into 0.1 M potassium phosphate buffer, pH 8.0, with a 30 kDa molecular weight cut-off Amicon centrifugal filtration device (EMD Millipore), flash frozen on dry ice, and stored at -20 °C.

113
4.4.8 HPLC analysis of purified proteins for heme content

The procedure for HPLC analysis of proteins was adapted from a method for analysis of Fe-DPIX enzymes (38). Proteins for HPLC analysis were exchanged into distilled water via dialysis or a 30 kDa molecular weight cut-off Amicon centrifugal filter (EMD Millipore). Proteins were diluted to approximately 5 µM in distilled water and filtered through glass wool plugs prior to HPLC analysis on a 1260 Infinity binary liquid chromatography system (Agilent). Separations were achieved on a Restek Viva C4 column (5 µm, 2.1 x 150 mm) using a gradient of acetonitrile in water + 0.1% formic acid from 5% for 5 min, 5-100% over 10 min, and 100% for 5 min at a flow rate of 0.4 mL/min.

4.4.9 Absorbance spectra of enzymes and the carbon monoxide binding assay

All absorbance measurements for protein samples were carried out in 0.1 M potassium phosphate buffer, pH 8. UV-vis data was obtained on a Tecan M1000 PRO UV/Vis plate reader or Cary 100 UV/Vis spectrometer (Agilent). The carbon monoxide difference spectra for native heme P450 enzymes were obtained as previously described in multiwell plate format (16).

4.4.10 High-throughput screen for MPH-selective variants

The screening procedure was adapted from a previous screen for Fe-DPIX-selective P450BM3 variants (38). Variant libraries of CYP119 in the vector pCWori were co-transformed with the plasmid pEvol-ChuA into BL21(DE3) Gold electropotent cells and selected on LB Agar supplemented with 40 µg/mL chloramphenicol and 100 µg/mL ampicillin. Single colonies were inoculated into 96 well blocks containing 1 mL of M9 minimal media supplemented with 0.5% glucose, 0.2% casamino acids, 340 µg/mL thiamine hydrochloride, 2 mM MgSO4, 0.1 mM CaCl2, 40 µg/mL chloramphenicol and 100 µg/mL ampicillin. One to three wells of each block were used for controls harboring wild type CYP119 as well as an uninoculated well to control for contamination. Blocks were sealed with M3 medical tape to permit aeration during cell growth. Cultures were grown to saturation overnight at 37 °C, 990 rpm and 80% humidity in a Multitron Pro 3 mm orbital incubation shaker (Infors HT). The following day, 100 µL of the cultures was
mixed with 100 µL of 50% sterile glycerol in 96 well plates for storage at -80 °C. Another 50 µL of culture was used to inoculate 96 well blocks containing 1 mL of M9 media prepared as above. Cells were grown for 3 hours at 37 °C, 990 rpm and 80% humidity to an OD$_{600nm}$ of approximately 0.6-0.7, and then induced with 0.2% arabinose, 0.5 mM IPTG, 1 µM heme, and 10 µM MPH. Proteins were expressed overnight at 25 °C, 990 rpm, and 40% humidity. Following overnight expression, cells were harvested at 4000 × g for 10 minutes. The supernatant was discarded and the pellets were frozen at -20 °C for at least 2 hours. Thawed pellets were resuspended in 400 µL of 0.1 M potassium phosphate buffer, pH 8.0, supplemented with 0.5 mg/mL lysozyme and a small amount (approximately 0.01 mg/mL) of DNaseI and incubated at 37 °C, 990 rpm for 1 hour. The lysate was cleared of cell debris by centrifugation at 4000 × g for 10 minutes. 200 µL of the clarified lysate was transferred to microtiter plates for high-throughput absorbance measurements on a Tecan M1000 PRO UV/Vis plate reader from 380 – 500 nm. The relative contribution of MPH-bound and heme-bound protein was estimated by fitting the data to a weighted sum of gaussians, using a non-linear least squares algorithm (62) in Matlab (Mathworks). We used the weighting of the Gaussians ($\omega$) to calculate an approximate percent MPH in the enzyme as:

$$\% MPH = \frac{\omega_{456nm}}{\omega_{456nm} + \omega_{421nm}}$$

We then compared the % MPH between variants and parent proteins to identify hits with improved MPH selectivity.

4.4.11 Cyclopropanation Reactions

Small scale (400 µL) anaerobic reactions were setup in 2 mL crimp vials (Agilent). Distilled water and a 10X concentrated solution of 0.1 M sodium dithionite in 1 M potassium phosphate buffer (pH 8.0) were degassed in separate sealed vials for at least 10 minutes prior to use. The headspace of a sealed crimp vial containing the heme derivative in 0.1 M NaOH and a stir bar was flushed with argon. Using glass syringes, 40 µL of the 10X buffer/dithionite
solution was added under argon, along with water to a final volume of 380 µL. 10 µL of a styrene solution (1.2 M in EtOH) was then added via syringe and allowed to mix for 30 seconds prior to adding 10 µL of ethyl diazoacetate (400 mM in EtOH). The final concentrations of the reagents were: 20 µM catalyst, 10 mM sodium dithionite, 30 mM styrene, and 10 mM ethyl diazoacetate in 0.1 M potassium phosphate (pH 8.0). The vial was removed from the argon line, the septum covered with parafilm, and the reaction was allowed to stir overnight at room temperature. Reactions were quenched by the addition of 30 µL of 3 M HCl. The vials were opened and 20 µL of an internal standard solution (100 mM cyclopropyl phenyl ketone in EtOH) was added. The reaction was extracted with 1 mL of ethyl acetate, transferred to a 1.7 mL microcentrifuge tube, vortexed, and centrifuged at 18,400 x g for 2 minutes. The organic layer was then removed and dried over sodium sulfate in a separate microcentrifuge tube. The organic layer was then filtered through a glass wool plug into glass vials for chiral gas chromatography analysis. GC separations were achieved at constant pressure and a GC oven temperature gradient of 100 °C for 30 min, 1 °C/min to 135 °C, 135 °C for 10 min, 10 °C/min to 200 °C, and 200 °C for 5 min.
## 4.5 Supplementary Information

### 4.5.1 Supplementary Tables

**Table S4.1.** Naming of synthetic porphyrins used in this study.

<table>
<thead>
<tr>
<th>Name used in this study</th>
<th>Abbreviation</th>
<th>IUPAC semisystematic name</th>
</tr>
</thead>
<tbody>
<tr>
<td>monophenyl heme</td>
<td>MPH</td>
<td>Fe(III) 3-styryl-8-vinyl-deuteroporphyrin/Fe(III) 3-vinyl-8-styryl-deuteroporphyrin</td>
</tr>
<tr>
<td>bisphenyl heme</td>
<td>BPH</td>
<td>Fe(III) 3,8-distyryl-deuteroporphyrin</td>
</tr>
<tr>
<td>mono(4-methoxy)phenyl heme</td>
<td>OMe-MPH</td>
<td>Fe(III) 3-(4-methoxystyryl)-8-vinyl-deuteroporphyrin/Fe(III) 3-vinyl-8-(4-methoxystyryl)-deuteroporphyrin</td>
</tr>
<tr>
<td>mono (4-trifluoromethyl)phenyl heme</td>
<td>CF₃-MPH</td>
<td>Fe(III) 3-(4-trifluoromethylstyryl)-8-vinyl-deuteroporphyrin/Fe(III) 3-vinyl-8-(4-trifluoromethylstyryl)-deuteroporphyrin</td>
</tr>
<tr>
<td>mono (pentafluoro)phenyl heme</td>
<td>F₅-MPH</td>
<td>Fe(III) 3-(2,3,4,5,6-pentafluorostyryl)-8-vinyl-deuteroporphyrin/Fe(III) 3-vinyl-8-(2,3,4,5,6-pentafluorostyryl)-deuteroporphyrin</td>
</tr>
</tbody>
</table>

**Table S4.2.** Mutations present in hits from from first library.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Ile130</th>
<th>Leu217</th>
<th>Leu248</th>
<th>Leu322</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G7</td>
<td>Ala</td>
<td>Trp</td>
<td>Leu</td>
<td>n.d.</td>
</tr>
<tr>
<td>2G8</td>
<td>Ala</td>
<td>Trp</td>
<td>Leu</td>
<td>Leu</td>
</tr>
<tr>
<td>3B7</td>
<td>Ala</td>
<td>Val</td>
<td>Leu</td>
<td>Ile</td>
</tr>
</tbody>
</table>
4.5.2 Supplementary Figures

Figure S4.1. Analysis of cross-metathesis reactions between protoporphyrin IX derivatives and styrene. A) Representative HPLC trace for the crude product mixture of the cross-metathesis reaction between hemin dimethyl ester and styrene. B) Reaction time course for the cross-metathesis of hemin dimethyl ester with styrene. C) Reaction time course for the cross-metathesis of zinc protoporphyrin IX with styrene, following total conversion (black) as well as the amount of starting material (blue) and monoaryl (orange) and bisaryl (yellow) products.
Figure S4.2. Incorporation of MPH and BPH into proteins *in vivo*. A) Absorbance spectra of purified ferrous CO-bound BM3h in complex with heme (blue) or MPH (green). B) Absorbance spectra of purified ferric Myoglobin H64V V68A in complex with MPH. C) Sample HPLC trace of BM3h expressed in minimal media in the presence of BPH; there is very little BPH incorporation relative to heme and even some MPH incorporation, mostly likely due to small amount of MPH impurity in the BPH preparation.
Figure S4.3. Absorbance spectra of hits from library screening with Zn-MPH. The composite absorbance spectra (black) of variants expressed in minimal media supplemented with 10 µM Zn-MPH and 1 µM heme. The contributions of the heme-bound enzyme (red-dotted line) and the Zn-MPH-bound enzyme (blue-dotted line) were approximated by fitting the data to a sum of Gaussians using Excel’s Solver function.
Figure S4.4. $^1$H NMR spectrum of MPH.
Figure S4.5. $^1$H NMR spectrum of BPH.
Figure S4.6. $^1$H NMR spectrum of OMe-MPH.
Figure S4.7. $^1$H NMR spectrum of CF$_3$-MPH.
Figure S4.8. $^{19}$F NMR spectrum of CF$_3$-MPH.
Figure S4.9. $^1$H NMR spectrum of F$_5$-MPH.
Figure S4.10. $^1$H NMR spectrum of Zn-MPH.
Figure S4.11. $^1$H NMR spectrum of Zn-BPH.
REFERENCES


CHAPTER 5: GADOLINIUM-SUBSTITUTED PROTEINS AS HIGHLY SENSITIVE MRI BIOSENSORS

5.1 Introduction

Magnetic resonance imaging (MRI) has emerged as a powerful tool for studying dynamic physiological processes in animals, such as neuronal signaling in the brain (1). The signal observed in MRI results from spin-excitation of hydrogen nuclei (which in animal tissue are primarily in the form of water protons) in a constant magnetic field, and measurement of the subsequent longitudinal ($T_1$) and transverse ($T_2$) relaxation times (2). Traditionally, the contrast in these images arises from differences in water concentration and the chemical environment in different tissues (3). Paramagnetic metal complexes can be used to enhance contrast in MRI by facilitating the relaxation of water protons in close proximity to the metal (3, 4). The effectiveness of a contrast agent is expressed as relaxivity ($r_1$ or $r_2$), or the change in relaxation rate caused by the contrast agent divided by its concentration (4). The spin of the metal is an important factor governing the relaxivity of a contrast agent (5). For this reason, high spin gadolinium complexes are the most frequently used contrast agents, with more than 10 million MRI studies performed with gadolinium annually (4).

Protein-based MRI contrast agents have emerged as useful alternatives to traditional small molecule contrast agents (6). These protein-based contrast agents offer several advantages over small molecule metal complexes, such as increased stability and solubility, increased tissue retention time, and increased rotational correlation times, which enhances contrast (7). Furthermore, proteins can be used as ligand-responsive MRI contrast agents, the molecular specificity of which can be tuned through evolution (1, 6, 8–10). For example, the enzyme cytochrome P450$_{BM3}$ possesses a paramagnetic iron heme center, which in the ligand-free state is bound to a water molecule. Upon ligand binding, however, the water molecule is
displaced from the iron, resulting in a ligand-dependent change in relaxivity. Taking advantage of this phenomenon, cytochrome P450_{BM3} has been engineered as an MRI biosensor for neuronal signaling molecules such as dopamine and serotonin (1, 9, 10), enabling the spatiotemporal mapping of brain activity in mice (11).

One challenge with the use of metalloproteins as contrast agents, however, is the relatively low relaxivity exhibited by naturally-occurring metal complexes (5, 7). This results in low analyte sensitivity and ligand buffering, limiting the application of these MRI biosensors (5, 11). One strategy to enhance the relaxivity of protein-based MRI contrast agents is to incorporate a non-natural metal center with higher spin into the protein. For example, a Mn-heme derivative was introduced to P450_{BM3}, enhancing the relaxivity of the contrast agent approximately 3-fold due to the higher spin of Mn (S = 2) compared to Fe-heme (S=1/2) (5). Ideally, the introduction of Gd (S = 7/2) to proteins would provide even greater gains in relaxivity. However, Gd-porphyrin complexes are unstable due to the large atomic radius of Gd, which causes the metal to be poorly accommodated in the porphyrin macrocycle (12–14). While it should be noted that Marletta and coworkers recently reported the incorporation of Gd-protoporphyrin IX into proteins \textit{in vivo}, no evidence was provided to confirm Gd remained bound to the porphyrin upon protein incorporation (15). Alternatively, expanded porphyrin-like macrocycles, such as “texaphyrins” (Texas-sized porphyrins) developed by Sessler and coworkers, have been shown to stably accommodate Gd and other lanthanide metals, and have been used as MRI contrast agents (14, 16–19). Therefore, we reasoned that incorporation of Gd-texaphyrin complexes into proteins could create stable, highly-sensitive protein-based MRI biosensors.

Towards this goal, we report the design and synthesis of a novel Gd-texaphyrin complex, 3 (denoted as Gd-Tex, \textbf{Figure 5.1A}). Preliminary modeling studies suggest that the structure can be accommodated within the P450_{BM3} heme-binding pocket (\textbf{Figure 5.1B}). Gd-Tex dimethyl ester 9 is readily synthesized in 5 steps from commercially available and
previously described pyrroles, following established protocols (Figure 5.2) (14, 17, 20). Current efforts are focused on optimizing dimethyl ester deprotection protocols for downstream biological experiments. The introduction of a non-natural Gd-cofactor to the protein will enhance the relaxivity of the MRI biosensor and expand the utility of these systems in the study of dynamic biological processes.

Figure 5.1. Design of a Gd-texaphyrin cofactor. A) The chemical structure of heme and several Gd-texaphyrin derivatives. Complex 3 was chosen for this study. B) Modeling of Gd-Tex 3 (purple) into wild type P450BM3 (green, PDB ID: 1JPZ (21)). Potential hydrogen-bonding interactions are shown as black-dotted lines.

5.2 Results and Discussion

5.2.1 Design of a Gd-texaphyrin cofactor

Several considerations were weighed in our design of a Gd-texaphyrin cofactor. First, we sought to prepare texaphyrin derivatives with similar structural features to heme, to enable cellular uptake by the heme transport protein ChuA and incorporation into the heme binding pocket of the enzyme P450BM3 or other hemoprotein scaffolds. Furthermore, we sought a
texaphyrin derivative that could be synthesized in a limited number of chemical steps. While our initial design of the heme-like Gd-texaphyrin complex 1 maintained many of the heme functional groups such as the vinyl and propionate moieties (Figure 5.1A), the asymmetric nature of this structure necessitates a complex synthetic route. Alternatively, Sessler and colleagues have previously reported a straightforward synthesis for the preparation of Gd-texaphyrin complex 2 (Figure 5.1A)(14). We reasoned that replacement of the hydroxyl groups of this structure with carboxylic acid moieties to generate Gd-Tex 3 (Figure 5.1A) would afford a structure with sufficient similarities to heme to enable incorporation into proteins in vivo while also minimizing the synthetic effort required.

To investigate the likelihood of Gd-Tex being accommodated in the P450BM3 active site, we performed a crude molecular docking of this cofactor into the structure of the wild type heme domain (BM3h) (21), using the software programs COOT and PyMol (Figure 5.1B). In this model, Gd-Tex appears to fit well within the heme-binding pocket of BM3h with little steric clash. One of the carboxylate groups of Gd-Tex binds in the same position as a propionate group of heme, maintaining native hydrogen bonding interactions with the arginine and tryptophan side chains of residues 398 and 96, respectively (Figure 5.1B). In this orientation, the other carboxylate group of Gd-Tex fortuitously takes the place of a conserved water molecule observed in the wild type structure, in good position to make hydrogen bonding interactions with the backbone amides of residues 327 and 328 (Figure 5.1B). Encouraged by this model, we moved forward with the synthesis of Gd-Tex.

5.2.2 Synthesis of Gd-Tex and test incorporation into BM3h

The synthesis of Gd-Tex (Figure 5.2) was adapted from standard texaphyrin syntheses reported by Sessler (14, 16). In collaboration with Dr. Kathy Matera (Elon University), we prepared over 100 mg of the unmetalated texaphyrin 8. Metalation of macrocycle 8 proceeds readily by reaction with Gd(NO₃)₃, and the product displays an absorbance spectrum consistent with other Gd-texaphyrins (Figure 5.3).
Figure 5.2. Synthesis of Gd-Tex dimethyl ester.
Successful production of Gd-Tex dimethyl ester 9 was further confirmed by LC-MS analysis of the purified complex (Figure S5.1).

**Figure 5.3. Absorbance spectra of Gd-Tex dimethyl ester and the base hydrolysis product.** Gd-Tex dimethyl ester (green) displays a Soret-like absorbance band at 436 nm and an additional peak at 770 nm. These features are lost in the absorbance spectrum of the product following reaction under basic conditions (blue), suggesting decomposition of the texaphyrin.

With Gd-Tex dimethyl ester complex 9 in hand, we moved forward with hydrolysis of the methyl ester groups to prepare the target compound, Gd-Tex 3, for protein incorporation experiments. However, attempts to perform the hydrolysis reaction under basic conditions resulted in decomposition of the Schiff-base complex as evidenced by UV-Vis absorbance and LC-MS analyses (Figure 5.3 and Figure S5.2). As an alternative, we sought to exploit lipase enzymes to achieve methyl ester hydrolysis. In nature, lipases catalyze the hydrolysis of fatty acid esters and are frequently used in chemical synthesis for ester hydrolysis/formation with diverse substrates (22–24). Lipases can promote ester hydrolysis under mild conditions (22), thus, we reasoned these enzymes could be used to prepare Gd-Tex without decomposition of the texaphyrin macrocycle.

Comparison of two commonly used lipase enzymes, lipase from *Candida rugosa* and porcine pancreatic lipase, in a range of buffers at different pH and temperatures revealed that porcine pancreatic lipase performed the reaction most efficiently in 0.1 potassium phosphate
buffer, pH 8, at 37 °C. The lipase began to precipitate out of solution after overnight reaction under these conditions. HPLC analysis of the reaction mixture at this point showed the appearance of a new peak with an absorbance profile consistent with Gd-texaphyrin (peak 1, Figure 5.4). LC-MS analysis also revealed a mass peak consistent with the dicarboxylic acid species, Gd-Tex 3 (Figure S5.3).

![HPLC analysis of lipase-catalyzed hydrolysis of Gd-Tex dimethyl ester.](image)

**Figure 5.4.** HPLC analysis of lipase-catalyzed hydrolysis of Gd-Tex dimethyl ester. The top panel shows HPLC traces of purified Gd-Tex dimethyl ester and the crude reaction product from the lipase-catalyzed hydrolysis reaction. Peak 1 shows decreased retention time and UV-Vis absorbance spectrum consistent with Gd-Tex, which suggests it is the dicarboxylic acid species Gd-Tex 3.

Although conversion appeared to be low in this reaction, we attempted to use the product recovered from this reaction for *in vivo* incorporation experiments. The proteins BM3h and ChuA were coexpressed in *E. coli* cells grown in minimal media supplemented with approximately 10 µM of the Gd-Tex mixture. Following protein expression, the cells were lysed and the absorbance was measured prior to, and following, addition of reductant and exposure to
CO. This revealed reduced, CO-bound protein with an absorbance peak at 450 nm, consistent with the heme-bound enzyme (Figure S5.4). Unfortunately, no peaks that corresponding to a Gd-Tex-bound enzyme were detected in the absorbance spectrum. Current work is aimed at improving conversion for the lipase hydrolysis reaction and investigating the cause of poor Gd-Tex incorporation into proteins.

5.3 Conclusions and Future Work

We report the design and synthesis of a novel Gd-texaphyrin cofactor to generate improved protein-based MRI biosensors. While the Gd-Tex dimethyl ester derivative 9 is readily accessed in 5 steps from common pyrrole building blocks, the hydrolysis step to obtain the desired dicarboxylic acid product 3 requires further optimization. Though we have initially adopted a lipase-catalyzed ester hydrolysis strategy, we are also pursuing an alternative approach involving hydrolysis of the methyl ester groups prior to the ring-formation reaction.

We assume that Gd-Tex dimethyl ester will not be efficiently incorporated into proteins, due to the additional steric bulk of the methyl groups and the importance of hydrogen bonding interactions between the protein and heme carboxylates for heme recognition. Therefore, our initial failure to observe incorporation of Gd-Tex into proteins is perhaps unsurprising given that we used a Gd-Tex preparation containing large amounts of the Gd-Tex dimethyl ester. If incorporation of pure Gd-Tex 3 into proteins also fails, engineering of ChuA or the hemoprotein may be required. Alternatively, the synthesis of a more complex, heme-like Gd-texaphyrin derivative, such as 1, could be undertaken to determine if these modifications to the texaphyrin structure improve incorporation. Despite the initial challenges encountered, we remain confident that the preparation of Gd-texaphyrin-containing proteins is a promising means of generating highly-sensitive protein-based MRI contrast agents.
5.4 Materials and Methods

5.4.1 General methods and materials

All chemical reagents and solvents were obtained from chemical suppliers (Acros, Fisher Scientific, Sigma-Aldrich, or Frontier Scientific) and used without further purification. Lipase enzymes were obtained from Sigma-Aldrich.

Reactions requiring inert atmosphere were carried out in oven-dried glassware using standard Schlenk techniques under nitrogen. Reaction progress was monitored by thin layer chromatography (EMD Millipore TLC silica gel 60 F254). Silica gel chromatography was performed on an automated Biotage Isolera One using 25 g SNAP Ultra columns. Proton magnetic resonance spectra were acquired on 400 MHz Bruker instruments in the Department of Chemistry NMR facility at the University of North Carolina – Chapel Hill or on NMR instruments at Elon University. NMR spectra were processed with ACD/NMR Processor (ACD/Labs). Chemical shifts are referenced to residual solvent peaks (25). LC-MS analyses were obtained with an Agilent 6520 Accurate Mass QTOF LC-MS ESI positive in high-resolution mode, equipped with a Restek Ultra biphenyl column (5 µm particle size, 3 x 50 mm), using a gradient of acetonitrile in water + 0.1% formic acid from 5-85% over 10 min, 85% for 5 min, 95% for 4 min, and 5% for 1 min. Small molecule absorbance spectra were taken on a on a Tecan M1000 PRO UV/Vis plate reader or Cary 100 UV/Vis spectrometer (Agilent) in MeOH. Protein absorbance spectra were also measured with a Tecan M1000 PRO UV/Vis plate reader, in 0.1 M potassium phosphate buffer, pH 8.

5.4.2 Organic synthesis and characterization

The synthesis of Gd-Tex dimethyl ester 9 was adapted from published procedures for the synthesis of 2 (14). The starting pyrrole, 2-(acetoxyethyl)-5-((benzyloxy)carbonyl)-4-methyl-3-((methoxycarbonyl)ethyl)pyrrole (4) was prepared by reaction of commercially available methyl 5-((benzyloxy)carbonyl)-2,4-dimethyl-3-pyrrolepropionate (Sigma-Aldrich) with
lead acetate as described previously (20). The 3,4-diethylpyrrole (5) was purchased from Frontier Scientific.

A. Synthesis of 2,5-bis[(5-((benzoxyl)carbonyl)-3-((methoxycarbonyl)ethyl)-4-methylpyrrol-2-yl)methyl]-3,4-diethylpyrrole (6)

Pyrrole 4 (1.21 g, 3.2 mmol) and pyrrole 5 (0.2 g, 1.6 mmol) were combined in EtOH (40 mL) in a round bottom flask. The solution was heated to reflux and stirred under nitrogen until pyrrole 4 was completely dissolved. Following this, p-toluenesulfonic acid monohydrate (10 mg, 0.06 mmol) was added the reaction was refluxed under a N2-filled balloon and stirred for 15 hr. The solvent was evaporated to half-volume under vacuum and the resulting slurry was placed at -20 °C for 6 hours. The product was collected by filtration and washed with cold ethanol to yield the dibenzyloxytripyrane 6 as an off-white solid, 516 mg (43% yield). The characterization data match that of the previously reported molecule (14).

B. Synthesis of 2,5-bis[5-formyl-3-((methoxycarbonyl)ethyl)-4-methylpyrrol-2-yl)methyl]-3,4-diethylpyrrole (7)

The dibenzyloxytripyrane 6 (500 mg, 0.67 mmol) was dissolved in anhydrous THF (25 mL) in a round bottom flask under nitrogen. While stirring, triethylamine (∼3-4 drops) was added, along with 10% Pd on carbon (100 mg). The reaction was then stirred at room temperature under a H2-filled balloon overnight. The reaction was filtered through Celite to
remove the catalyst, and the solvent evaporated to yield a light pink solid. The resulting dicarboxyltripyrrene was then carried on immediately to the next step. In a round bottom flask containing this product under nitrogen, trifluoroacetic acid (2 mL) was added dropwise via syringe, with visible evolution of a gas (CO$_2$). The resulting solution was stirred at room temperature for 15 min, then cooled to 0 °C in an ice water bath. Triethylorthoformate (1 mL) was added, and the reaction stirred for 10 min at 0 °C. Following this step, the mixture was placed back at room temperature, and 50 mL of a saturated sodium bicarbonate solution was added. The product was then extracted into 20 mL of DCM. The DCM layer was collected and dried over sodium sulfate. The solvent was evaporated and then the residue was redissolved in 95% MeOH (25 mL). This solution was heated to 60 °C and then slowly cooled to room temperature, followed by evaporation of the solvent to half-volume under reduced pressure. The slurry was placed in the freezer for several hours. The product was collected by filtration and rinsed with water to yield the diformyltripyrrene 7 as a brown solid, 269 mg (75% yield). $^1$H NMR (CDCl$_3$, 400 MHz) δ 9.61 (s, 2H), 8.89 (s, 1H), 3.84 (s, 4H), 3.61 (s, 6H), 2.73-2.70 (m, 4H), 2.44-2.35 (m, 8H), 2.23 (s, 6H), 1.10-1.07 (m, 6H) ppm.

C. Synthesis of 4,5-diethyl-10,23-dimethyl-9,24-bis(3-(methoxycarbonyl)ethyl)-13,20,25,26,27-pentaazapentacyclo-[20.2.1.1$^3$.6.1$^8$.11.0$^{14}.19$]heptacosa-3,5,8,10,12,14,16,18,20,22,24-undecaene (8)

![Diagram of the molecule](image)

The diformyltripyrrene 7 (250 mg, 0.47 mmol) was added to a solution of toluene (95 mL) and MeOH (20 mL) in a round bottom flask under nitrogen with stirring. To this mixture, o-
phenylenediamine (51 mg, 0.47 mmol) was added. Concentrated HCl (0.25 mL) was added and the reaction was heated to reflux under nitrogen. After 5 hours the reaction was cooled to room temperature and the solvent was evaporated under reduced pressure. The red product was recrystallized from DCM/hexane and yielded 287 mg (quantitative). $^1$H NMR (CDCl$_3$, 400 MHz) δ 7.89 (s, 2H), 7.38-7.35 (m, 4H), 4.04 (s, 4H), 3.68 (s, 6H), 2.83-2.80 (m, 4H), 2.51-2.49 (m, 4H), 2.42-2.39 (m, 4H), 2.25 (s, 6H), 1.10-1.07 (m, 6H) ppm.

D. Synthesis of gadolinium(III) 4,5-diethyl-10,23-dimethyl-9,24-bis(3-(methoxycarbonyl)ethyl)-13,20,25,26,27-pentaazapentacyclo-[20.2.1.1$^{3,6}$]heptacosa-1,3,5,7,9,11(27),12,14,16,18,20,22(25),23-tridecaene (9)

The macrocycle 8 (150 mg, 0.25 mmol) was dissolved in MeOH (130 mL) in a round bottom flask open to air. Next, Gd(NO$_3$)$_3$·6H$_2$O (330 mg, 0.75 mmol) and trimethylamine (0.5 mL) were added, and the reaction was heated to reflux under air with stirring for 5 hours. Following the reaction, the solvent was evaporated off and the crude reaction mixture dissolved in 5% MeOH in DCM and loaded onto a silica gel column in DCM. Orange-colored impurities were eluted with DCM. Increasing the strength of the mobile phase to 5% MeOH in DCM resulted in elution of a green impurity, followed by elution of the product as another green band displaying absorbance peaks at 436 nm and 770 nm, consistent with metalated Gd-texaphyrin. The product fractions were collected and solvent was removed via rotary evaporation to yield the Gd-texaphyrin dimethyl ester 9 as a dark green solid, 157 mg (~75%). The metal complex likely has two nitrate counterions, however, elemental analysis must be performed to confirm this. The product was characterized by LC-MS analysis (Figure S5.1), and shows a mass
consistent with the calculated mass of the Gd-texaphyrin complex with a formate ligand coming from the mobile phase. Calculated \([M+\text{CHO}_2]^+ = 807.2141\) found \([M+\text{CHO}_2]^+ = 807.2176\)

**E. Base-catalyzed methyl ester hydrolysis**

The Gd-texaphyrin dimethyl ester complex 9 (150 mg) was dissolved in 95% MeOH (10 mL) and potassium hydroxide (300 mg) was was added. The reaction was heated to 50 °C with stirring under nitrogen. Following overnight reaction, UV-Vis and LC-MS analysis showed decomposition of the Schiff-base complex (see Figure 5.3 and Figure S5.2). Similarly, milder hydrolysis conditions at room temperature resulted in loss of the green color of the solution after several hours at basic pH, indicating decomposition of the Gd-complex.

**F. Lipase-catalyzed methyl ester hydrolysis**

The Gd-texaphyrin dimethyl ester complex 9 (20 mg) was dissolved in 500 µL MeOH then diluted up to 10 mL in 0.1 M potassium phosphate buffer, pH 8.0. Porcine pancreatic lipase (90 mg) was added, forming a brown suspension. The reaction was heated to 37 °C with stirring overnight. The reaction mixture was filtered through a 0.2 µm syringe filter to remove precipitated protein. The filtered reaction mixture was then used directly in biological experiments. HPLC analysis of the reaction mixture (Figure 5.4 and Figure S5.3A) on a 1260 Infinity binary liquid chromatography system (Agilent) was performed to determine approximate conversions. Separations were achieved on a Restek Viva C4 column (5 µm, 2.1 x 150 mm) using a gradient of acetonitrile in water + 0.1% formic acid from 5% for 5 min, 5-95% over 10 min, 95% for 5 min, and 5% for 5 min. LC-MS analysis of the product (Figure S5.3B) using the same column and gradient provided a mass value in close agreement with that of the Gd-texaphyrin complex 3 with a formate ligand coming from the mobile phase. Calculated \([M+\text{CHO}_2]^+ = 779.1828\) found \([M+\text{CHO}_2]^+ = 779.3332\).

**5.4.3 Small-scale test incorporation of Gd-Tex into BM3h variants**

BM3h variants were cloned into the plasmid pCWori (9), and the heme transport protein ChuA was expressed from the plasmid pChuA (available from Addgene). These
plasmids were co-transformed into electrocompetent *E. coli* BL21-AI cells (Novagen). Individual colonies from these transformations were inoculated into liquid starter cultures of M9 minimal media supplemented with 0.5% glucose, 0.2% casamino acids, 340 µg/mL thiamine hydrochloride, 2 mM MgSO$_4$, 0.1 mM CaCl$_2$, 100 µg/mL kanamycin and 200 µg/mL ampicillin. The cultures were grown overnight at 37 °C and 225 rpm. Five milliliters of M9 minimal media supplemented with 100 µg/mL kanamycin and 200 µg/mL ampicillin was inoculated with 100 µL of overnight starter culture and incubated for 2-3 hours at 37 °C and 225 rpm to an OD$_{600\text{nm}}$ of 0.6-0.7. The cultures were then induced by adding arabinose (to induce ChuA expression in BL21-AI) and Isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.2% and 1 mM, respectively. A concentrated stock solution of Gd-Tex in 0.1 M potassium phosphate buffer, pH 8, was added to the culture upon induction to a final concentration of approximately 10 µM. The concentration of the Gd-Tex stock solution was approximated by measuring the absorbance and calculating concentration using the reported extinction coefficient for a similar Gd-texaphyrin complex in MeOH of ε$_{416\text{nm}}$ = 52,481 M$^{-1}$cm$^{-1}$ (14). Upon induction, the incubation temperature was reduced to 25 °C and expressions were allowed to continue for 16-20 hours. Following expression, the cells were harvested by centrifugation (4 °C, 10 min, 4,000 × g) and the cell pellets were frozen at -20 °C for at least 2 hours. The pellets were thawed and resuspended in 400 µL of lysis buffer consisting of 0.1 M potassium phosphate buffer, pH 8, with 0.5 mg/mL lysozyme (Sigma) and 0.1 mg/mL of DnaseI (Sigma). The cells were then incubated at 37 °C and 225 rpm for 1 hour. Cellular debris was removed by centrifugation (4 °C, 10 min, 4,000 × g) and a CO-binding assay in multiwell plate format was performed on of 200 µL of the cleared lysate using a Tecan M1000 PRO UV/Vis plate reader, as previously described (26).
5.5 Supplementary Information

5.5.1 Supplementary Figures

Figure S5.1. LC-MS analysis of Gd-Tex dimethyl ester 9. A) Total ion chromatogram (black) for the purified Gd-Tex dimethyl ester 9, along with the UV chromatogram monitoring 430 nm (green). B) Mass spectrum of the peak at 6.253 min in the TIC, with a mass corresponding to the Gd-Tex dimethyl ester complex + formate. C) UV trace of the peak at 6.253 min in the UV chromatogram, with an absorbance band at ~435 nm closely matching the absorbance spectrum of Gd-Tex dimethyl ester 9 shown in Figure 5.3.
Figure S5.2. LC-MS analysis of base-catalyzed Gd-Tex dimethyl ester hydrolysis reaction. The top panel shows the total ion chromatogram (black) and extracted ion chromatogram (m/z = 505.2, blue) for the crude reaction mixture analyzed by LC-MS. The mass spectrum for the peak at 10.02 min in the EIC is shown below, with the major peak closely matching the mass of the dialdehyde decomposition product.
Figure S5.3. LC-MS analysis of lipase-catalyzed Gd-Tex dimethyl ester hydrolysis reaction. A) Total ion chromatogram of the crude reaction mixture analyzed by LC-MS. B) Mass spectrum of peak 1 in the TIC, with a mass peak closely matching the calculated mass of the Gd-Tex 3 [M+H]^+ species. C) Mass spectrum of peak 2 in the TIC, with a mass peak closely matching the calculated mass of the Gd-Tex dimethyl ester 9 [M+Na]^+ species.
**Figure S5.4. Test incorporation of Gd-Tex into BM3h variants.** The UV-Vis absorbance spectra BM3h (blue) and BM3h L272W (orange) variants expressed in the presence of Gd-Tex. The absorbance peak at 450 nm observed for BM3h and BM3h L272W is indicative of CO-binding to the reduced heme cofactor of the enzyme, whereas we would expect a peak at approximately 436 nm if Gd-Tex was bound to the enzyme, based on the absorbance of Gd-Tex dimethyl ester (green), shown for reference.
Figure S5.5. $^1$H NMR of dibenzylxytripyrrane 6.
Figure S5.6. $^1$H NMR of diformyltripyrane 7.
Figure S5.7. $^1$H NMR of unmetalated texaphyrin macrocycle 8.
REFERENCES


