THE UNIVERSITY OF CHICAGO

DESIGN, CONSTRUCTION, AND DIRECTED EVOLUTION OF ARTIFICIAL METALLOENZYMES

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY

CHEN ZHANG

CHICAGO, ILLINIOS

DECEMBER 2015

To my parents Yuanfeng Zhang and Suqin Wang,

and my girlfriend Miaoyan Wang,

for their enduring love and support

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

LIST OF SCHEMES

ACKNOWLEDGEMENTS

Foremost, I thank my research adviser, Professor Jared C. Lewis for his support and advice throughout my graduate studies. I learned how to design scientific experiments, how to conduct presentation, and how to communicate with people. I also thank the members of my thesis committee, Professor Viresh Rawal and Professor Bryan Dickinson for their discussions.

I have worked with many talented people during my time in the Lewis group. Notably, my colleague Hao Yang has provided guidance and kind suggestion. Dr. Poonam Srivastava and Dr. Zhihui Zhang have shared their technique and experience on molecular biology and organic synthesis. I am grateful to have mentored two very talented undergraduates, Andrew Ng and Marissa Parker who are now applying for graduate schools. The past and current of the Lewis group with whom I worked, Landon Durak, James Payne, Yifan Gu, Mary Andorfer, Joe Gair, Ken Ellis-Guardiola, Alan Swartz, David Upp, Dr. Hyun June Park, Dr. Lee Solomon, and Dr. Duo-Sheng Wang have all helped provide thoughtful advice. I would like to thank Prof. Steve Kent and Dr. Michal Avital-Shmilovic in Kent group with assistance on all peptide synthesis work.

I am grateful for my undergraduate adviser, Professor Yuefei Hu at for his excellent training at Tsinghua University. The groundwork and platform built up from Tsinghua University is invaluable for my pursuit of a doctoral degree in chemistry.

The research described in this dissertation has relied heavily on the support of departmental facilities and instruments and I thank Dr. Jin Qin and Dr. Antoni Jurkiewicz for their respective assistance with mass spectrometry and NMR spectroscopy. I also thank Dr.

Vera Dragisich and Melinda Moore for their help in nonresearch related departmental matters.

I thank my family, especially my parents Yuanfeng Zhang and Suqin Wang for their love and support throughout my life. They have offered both financial and mental support which I could never repay.

Lastly, I thank my partner Miaoyan Wang. She has been continuously supportive throughout my graduate career, especially when I was stressful. I look forward to sharing my life with her.

ABSTRACT

This dissertation describes the design, construction, and directed evolution of artificial metalloenzymes. Transition metal catalysts and enzymes possess unique and often complementary properties that have made them important tools for chemical synthesis. In hopes of combining the selectivity and adaptability of enzymes with the reactivity of transition metal catalysts, researchers have explored artificial metalloenzymes (ArMs) in which secondary coordination sphere effects impart selectivity to metal catalysts, accelerate chemical reactions, and are systematically optimized via directed evolution. Our ultimate goal is to achieve non-directed and selective C-H functionalization.

Chapter one describes unnatural amino acid (UAAs) synthesis. UAA incorporation into protein scaffolds could provide a facile method to construct ArMs and metallopeptides. However, neither the synthesis nor the incorporation of catalytically active UAAs have been reported. We developed a general approach for the synthesis of UAAs using 9-borabicyclonon-anyl group which enables simultaneous protection of both the amine and carboxylic acid functionalities. Both bidentate and tridentate Pd complexes were synthesized and confirmed by HRMS-ESI and NMR characterization. These UAAs catalyzed imidate rearrangements, but peptides containing these UAAs did not impart enantioselectivity to this reaction, and we were unable to achieve incorporation of the UAAs into proteins using codon suppression methods. We next pursued incorporation of UAAs with catalytically active organic side chains. Inspired by secondary amine catalyzed reactions, we designed and prepared two UAAs, prolyl-lysine and imidazolidinone-lysine, with strong resemblance to pyrrolysine. Both UAAs catalyzed aldol and Michael reactions under physiological

conditions. This work was conducted in collaboration with Dr. Zhihui Zhang and Hao Yang, and was published in *Organometallics*, **2012**, *31*, 7328-7331.

Chapter two describes the construction of Mn terpyridine ArMs for oxygenation. Mn terpyridine is a robust catalyst for benzylic C-H oxygenation and olefin epoxidation. We designed and synthesized a maleimide-substituted Mn-terpyridine cofactor and demonstrated that this cofactor could be incorporated into two different scaffold proteins to generate the desired ArMs. The structure and reactivity of one of these ArMs was explored, and the broad oxygenation capability of the Mn-terpyridine catalyst was maintained, providing a robust platform for optimization of ArMs for selective hydrocarbon functionalization. This work was conducted in collaboration with Dr. Poonam Srivastava and Ken Ellis-Guardiola, and was published in *Tetrahedron*, **2014**, *70*, 4245-4249.

Chapter three describes the construction and directed evolution of dirhodium ArMs. We first demonstrated that strain-promoted azide–alkyne cycloaddition could be used for rapid and general ArM formation. This approach was used to incorporate an alkyne-substituted dirhodium complex into a prolyl oligopeptidase (POP), and the resulting ArM catalyzed carbenoid N-H insertion with low enantioselectivity. Scaffold mutagenesis was then employed to improve the enantioselectivity of N-H insertion reactions involving a range of aryl amines and donor-acceptor carbene precursors. Despite great success in evolving natural enzymes, no example of directed evolution applied to engineering ArMs has been reported. We established a streamlined, high-throughput protocol for design, expression, purification and screening of ArM libraries. After the process was sufficiently optimized, the first round of directed evolution from an engineered parent protein, F99H328, produced the

mutant G1P1_D5 which achieved 54% ee. This work was conducted in collaboration with Dr. Poonam Srivastava, Hao Yang, Hyun June Park and Ken Ellis-Guardiola.

PREFACE

Compounds are numbered continuously in this dissertation. A given compound may have a different number in different chapters. All experimental details, references, and notes for individual chapters are included at the end of each chapter. All unpublished NMR spectra are included in Appendix Two.

CHAPTER ONE

SYNTHESIS OF AMINO ACIDS WITH CATALYTICALLY ACTIVE SIDE CHAINS

1.1 INTRODUCTION TO UNNATURAL AMINO ACIDS

Catalytic processes form an integral part of modern methods for producing fine chemicals, pharmaceuticals, and a wide range of other materials[1]. The potential to exploit catalyst reactivity for applications has driven the development of catalysts ranging from metal complexes to engineered enzymes that function in complex media[2].

From a chemist's point of view, catalyst selectivity, whether to produce a single enantiomer of a pharmaceutical catalyzed by transition metal complexes or to conduct biocatalysis with a target substrate by enzymes, is essential. Transition metal complexes catalyze a broad range of challenging chemical transformations, including cross-coupling reactions[3], C−H bond functionalization[4], and olefin polymerization[5] and metathesis[6]. In general, such catalysts consist of a metal ion bound to some number of ligands that comprise the primary coordination sphere of the metal[7]. These ligands can be rationally designed to alter the reactivity of the metal center and to have broad substrate scope.

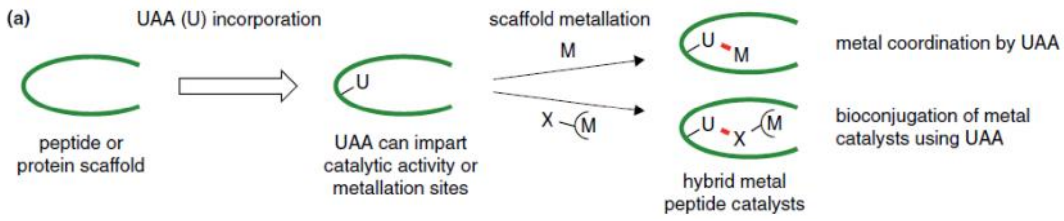
Enzyme catalysis is typically lauded for the extreme levels of rate acceleration and excellent selectivity exhibited by many natural enzymes[8] on their native substrates[9]. Unlike small molecule catalysts, enzymes can recognize specific substrates in active sites through a combination of hydrogen bonding, hydrophobic interactions, electrostatic, steric, and other effects[10]. This level of substrate recognition reduces barriers associated with enzyme catalyzed

reactions. These interactions arose from evolutionary processes[11], which can be mimicked in the laboratory to engineer enzymes with improved efficiency using directed evolution[12,13]. The adaptability of enzymes has been exploited to evolve enzymes with impressive levels of activity and selectivity toward unnatural substrates and reactions[14].

In hopes of combining the selectivity and adaptability of enzymes with the reactivity of transition metal catalysts, researchers have explored different methods to incorporate non-natural metal cofactors into peptides and proteins[10]. Many approaches have thus been developed using peptides and proteins comprised of the 20 natural amino acids. These systems can be roughly categorized as metal coordination involving natural or unnatural amino acids, covalent attachment of substituted catalysts, and non-covalent anchoring of substituted catalysts.

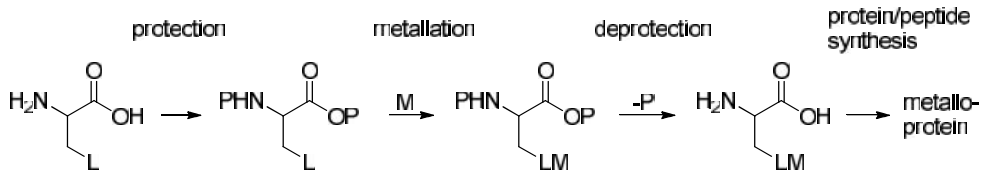
Unnatural amino acids (UAAs) are now routinely incorporated into peptides and proteins to enable functions beyond those accessible using natural amino acids. To date, UAAs have been used to enable new modes of metal coordination or covalent attachment (Figure 1.1). Direct incorporation of organometallic UAAs into proteins[15] has also been demonstrated, albeit without catalytic activity.

Figure 1.1. General scheme for construction of hybrid metal–protein catalysts via coordination or bioconjugation of metal catalysts using a UAA (taken from ref[1])



Although a variety of methods for incorporating unnatural amino acids (UAAs) into proteins have been developed, neither the synthesis nor the incorporation of catalytically active UAAs have been reported. Herein, we describe a general approach for the synthesis of unnatural amino acids that enable direct incorporation of catalytically active side chains (Scheme 1.1)[16]. This work is published[16] and it was done in collaboration with Hao Yang and Zhihui Zhang.

Scheme 1.1. Approach to synthesize organometallic amino acids for preparation of metalloproteins and peptides (adapted from ref[16])

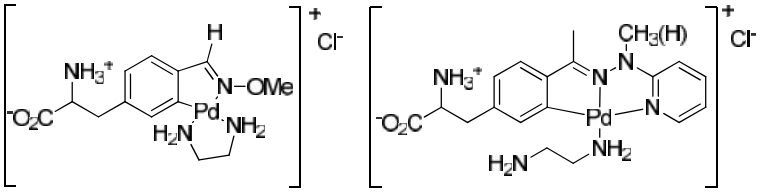


1.2 SYNTHESIS OF METALLOCYCLIC UAAS VIA 9-BBN PROTECTION

With the goal of facilitating the incorporation of transition metal catalysts into peptide scaffolds, we designed several bidentate and tridentate Pd-containing UAAs (Figure 1.2). Importantly, the organometallic amino acids were designed to mimic, as closely as possible, natural amino acids in order to maximize their compatibility with potential in vivo incorporation efforts. Preserving the native amino acid structure was also expected to minimize structural perturbations to peptides, including these amino acids, and to ensure close proximity of the metal center and the chiral peptide scaffold. Prior to this work, no general method existed for the preparation of unprotected amino acids with side chains containing catalytically active transition-metal centers[16], and most protected compounds include extended tethers[17] between the amino acid moiety and the organometallic fragment. Because both of these limitations arise

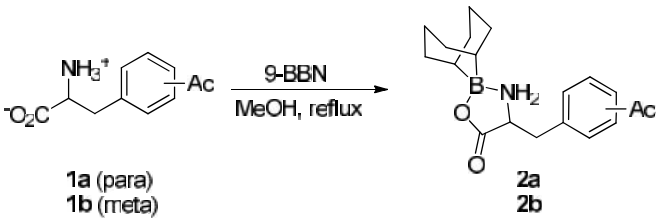
primarily from the well-established metal binding ability of the amino acid moiety, protecting group strategies that would enable installation of a range of metal centers while still allowing amino acid deprotection without compromising the activity of the metal center were required.

Figure 1.2. Proposed bidentate and tridentate Pd-UAA



Particularly attractive in this regard was the 9-borabicyclononanyl (9-BBN) group, which enables simultaneous protection of both the amine and carboxylic acid functionalities and can be readily removed using ethylenediamine[18]. The 9-BBN adducts of *p*- and *m*-acetylphenylalanine* (**1a,b**, respectively) were therefore prepared and purified by washing with hexanes (Scheme 1.2).

Scheme 1.2. Amino acids protection using 9-BBN



The resulting compounds, **2a** and **2b**, were then reacted with a range of amines and hydrazines to generate imines and hydrazines **3** (Scheme 1.3). Reaction of these compounds with either Na_2PdCl_4 or K_2PtCl_4 according to standard literature procedures led to the

* Conducted by Dr. Jared Lewis and Dr. Zhihui Zhang

formation of the 9-BBN protected organometallic amino acids **4**[19]. These were typically isolated as crude solids, and the free amino acids **5** could then be readily prepared by simply stirring these complexes in THF solutions of ethylenediamine. Representative Pd and Pt UAAs are shown in (Table 1.1).

The amino acids precipitated from solution and could be further purified by recrystallization from methanol/ether slow evaporation. The ethylenediamine was typically observed to bind in either a mono- or bidentate manner, depending on the number of available coordination sites. The structures of all UAAs were confirmed by HRMS-ESI, ^{1}H NMR and ^{13}C NMR characterization.

Scheme 1.3. General scheme for Pd-UAA synthesis

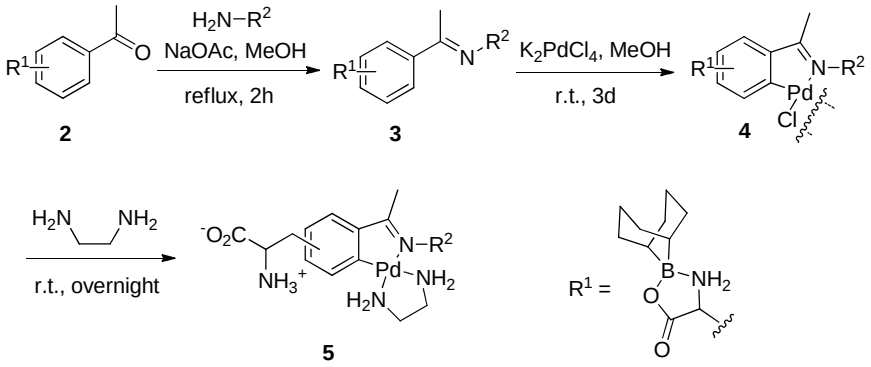
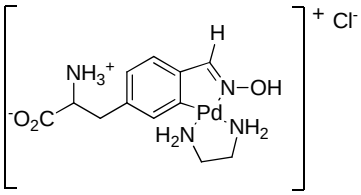
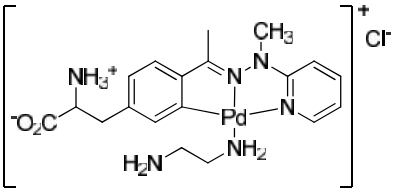
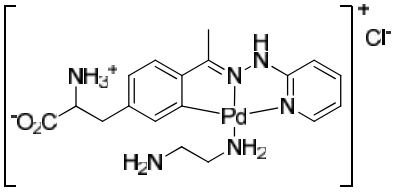
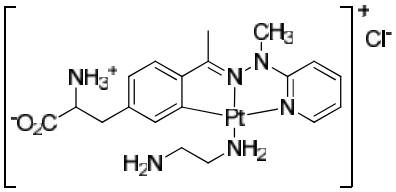


Table 1.1. Representative organometallic amino acids

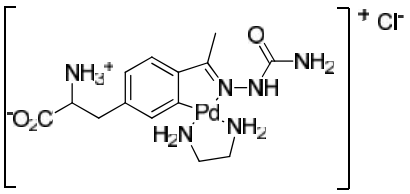
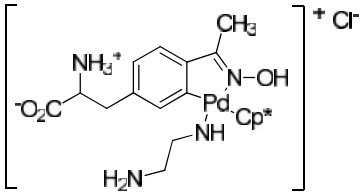
Entry	Structure (**5**)	Yield % **3**	Yield % **5**[1]
1		**3a** (90)	**5a** (58)
2		**3b** (90)	**5b** (50)
3		**3c** (89)	**5c** (30)
4*		**3d** (90)	**5d** (23)

[1]Cumulative yield over metalation and deprotection steps.

* Conducted by Dr. Zhihui Zhang

To demonstrate the generality of the BBN protecting group for the synthesis of unprotected organometallic amino acids, a few additional compounds were prepared by Dr. Zhihui Zhang and Hao Yang (Table 1.2, Scheme 1.4). Specifically, a 9-BBN-protected hydroxyquinoline-based amino acid was reacted with $[Cp^*IrCl_2]_2$ to provide the corresponding Cp*Ir−hydroxyquinoline chloride complex, which was deprotected to yield amino acid **9**. In addition, the known 9-BBN-protected lysine **10** was acylated with 2-pyridinecarboxylic acid chloride and reacted with $[CpRuCO\text{-}(ACN)_2]PF_6$ to provide complex **12**, which was deprotected to yield amino acid **13**[20]. Incorporation of these UAAs into proteins using stop codon suppression methodology was pursued in collaboration with Mike Jewett in Northwestern University, but unfortunately, none of metal-containing UAAs could be incorporated.

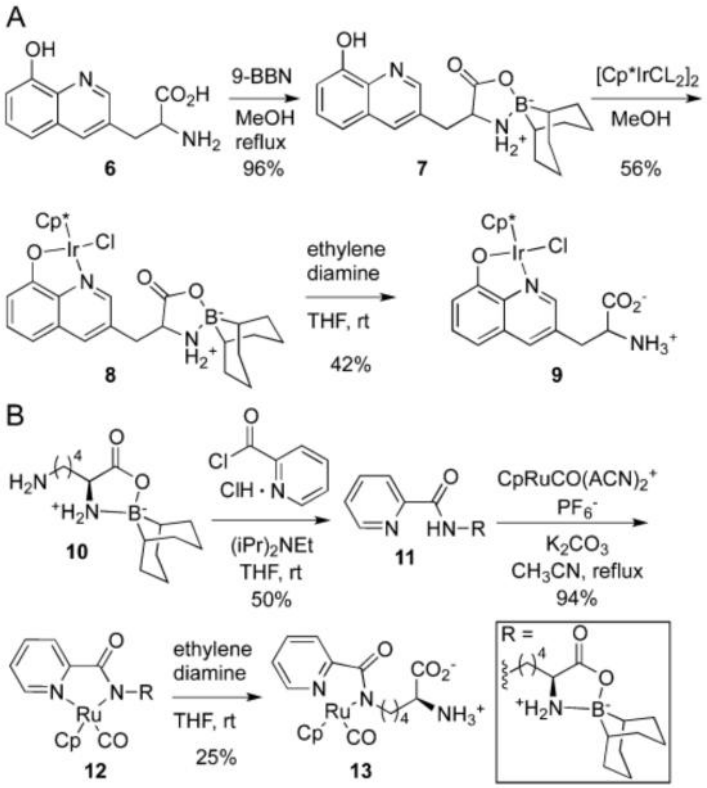
Table 1.2. Other UAAs prepared by Hao Yang and Dr. Zhihui Zhang[†]

Entry	Structure (**5**)	Yield %	
		3	**5**[1]
1		**3e**(89)	**5e**(20)
2		**3f**(93)	**5f**(30)

[1]Cumulative yield over metalation and deprotection steps.

[†] Conducted by Dr. Zhihui Zhang and Hao Yang

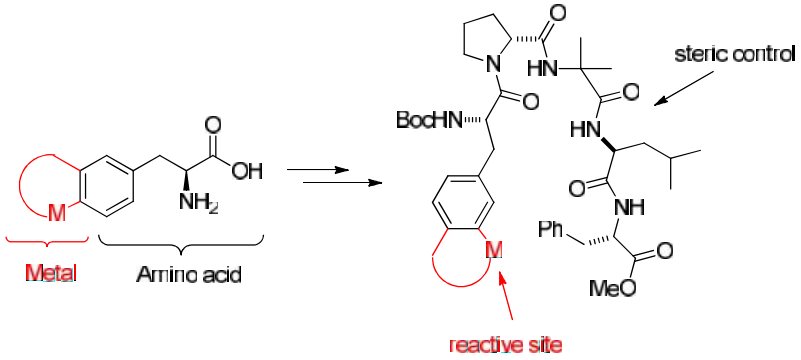
Scheme 1.4. Synthesis of hydroxyquinoline- and lysine-based UAAs 9 and 13[*]



1.3 SYNTHESIS OF METALLOPEPTIDE ON SOLID PHASE

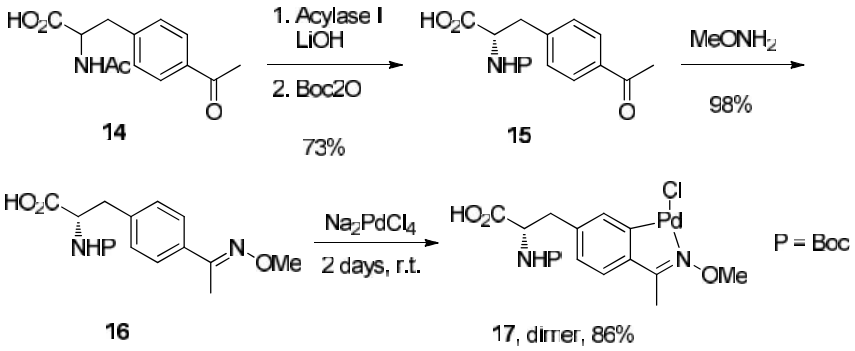
To demonstrate the utility of the UAAs described in section 1.2 for the site-specific introduction of metal centers into peptides, solid phase peptide synthesis (SPPS)[21] was used. The 2-aminoisobutyric acid (Aib)-proline motif in Figure 1.3 is known to enforce a β-turn structure and has been used by many groups, notably Miller and co-workers[22], to generate highly selective catalysts for a range of chemical transformations using organic amino acids. Peptide-based coordination complexes based on β-turn peptides have also been generated[23]. We hypothesized that incorporating organometallic catalysts into this platform would greatly expand the range of catalytic transformations possible using unique selectivity afforded by peptide catalysts. Ideally, the activity and selectivity of the resulting metallopeptides could be tuned by altering the additional amino acids in the peptide.

Figure 1.3. Proposed metallopeptide catalysts



For SPPS, enantiopure amino acids were also required (*E. coli* only incorporates S amino acids into proteins, so racemic samples can be used for *in vivo* studies[15]); therefore, an enzymatic resolution was used to prepare enantiopure acetylphenylalanine **15**[24]. This material was then Boc-protected, converted to the corresponding methoxyimine, and reacted with Na_2PdCl_4 to yield the Boc-protected metallacyclic amino acids **17** (Scheme 1.5). Resolved Pd-UAA was synthesized by Dr. Zhihui Zhang.

Scheme 1.5. Synthesis of resolved, protected Pd-UAA‡



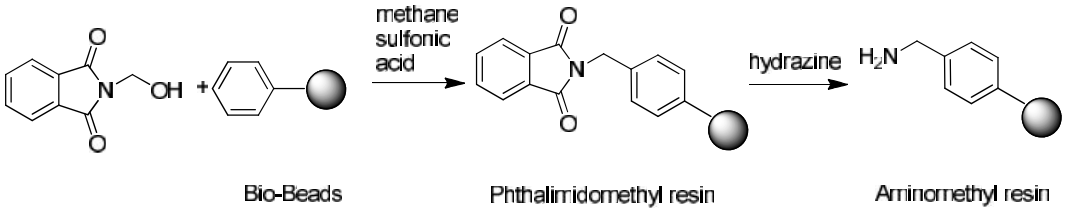
Solid Phase Peptide Synthesis (SPPS) can be defined as a process in which a peptide anchored by its C-terminus to a support polymer is assembled by the successive addition of the protected amino acids constituting its sequence. Each amino acid is added via a sequence

‡ Conducted by Dr. Zhihui Zhang

consisting of: a) cleavage of the N-protecting group; b) washing steps; c) coupling of a protected amino acid; and d) washing steps. Dr. Michal Avital-Shmilovic in Kent group assisted all SPPS work.

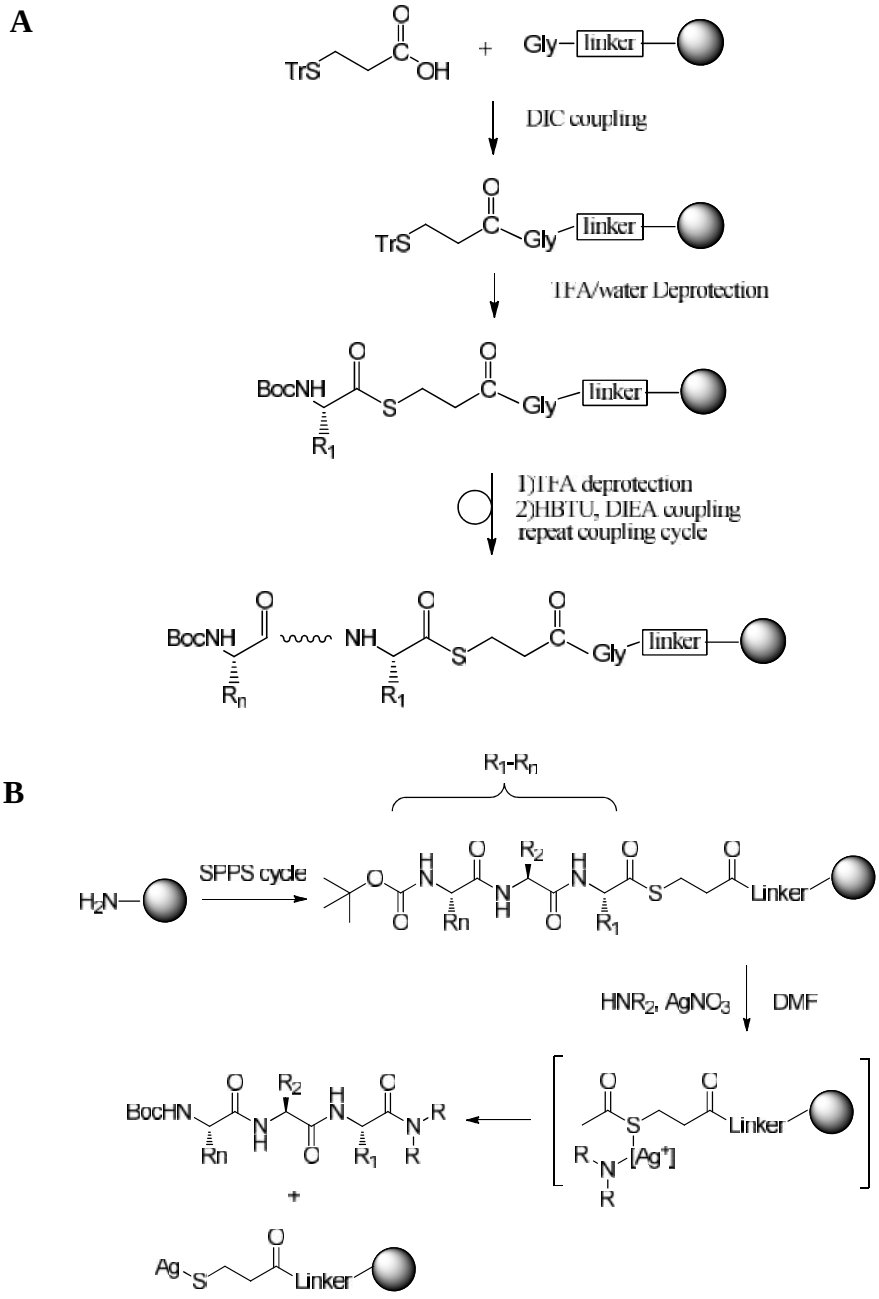
First, aminomethyl resin was prepared using bio-beads (Scheme 1.6). Phenylacetamidomethyl (PAM) linker[§] was then installed to anchor alanine followed by addition of mercaptopropionic acid, the key moiety for future cleavage. The desired peptide was then prepared (Scheme 1.7A). Glycine was chosen for the C-terminus of peptide due that silver-mediated cleavage may introduce racemization of C-terminal residue (Scheme 1.7B). β-turn peptide scaffolds was then reacted with metallocyclic UAAs under standard peptide coupling conditions[25]. NMR and mass spectroscopic analysis indicated that the metal centers remained unperturbed by the peptide coupling conditions. Several analogues, involving substitution of L-proline with D-proline, variation of the terminal phenylalanine residue, and addition of an additional amino acid to form pentapeptides, were also prepared (Table 1.3).

Scheme 1.6. Preparation of aminomethyl resin from Bio-beads



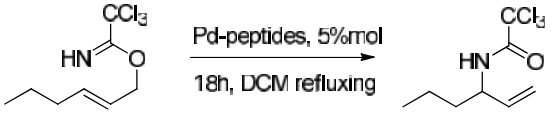
[§] Assisted by Dr. Michal Avital-Shmilovic

Scheme 1.7. (A)Workflow of SPPS elongation and (B) Aminolysis cleavage



Dr. Zhihui Zhang then coupled these peptides to different Pd-UAAs and conducted catalytic reactions employing the resulting metallopeptides (Table 1.3). All Pd-peptides catalyzed the conversion of an allylic trichloroacetimidate into an allylic trichloroacetamide[26]. Unfortunately, the small set of compounds investigated thus far provided no enantio-induction in either of these reactions. Control experiments using $PdCl_2$ as catalysts for this reaction provided no conversion, indicating that release of metal complexes from the peptides was probably not responsible for the observed nonselective reactions. Presumably, the metal fragment in each case adopted a conformation that placed it outside the influence of the peptide.

Table 1.3. Metallopeptides catalyzed Overmann rearrangement reaction



Entry	Catalyst	Conv.[1]
1	PdUAA-Pro-Aib-Leu-Asp-Gly	25
2	PdUAA-Pro-Aib-Leu-Asn-Gly*	20
3	PdUAA-Pro-Aib-Leu-Phe-Gly*	10
4	PdUAA-(D)-Pro-Aib-Leu-Asp-Gly	20
5	PdUAA-(D)--Pro-Aib-Leu-Phe-Gly	20
6	PdUAA-Phe-Pro-Aib-Leu-Phe-Gly	<5
7	PdUAA **3a**	56
8	$PdCl_2$*	0

PdUAA-Pro-Aib-X_1-X_2-Gly

[1]Conversion is calculated by NMR integration

* Conducted by Dr. Zhihui Zhang

1.4 DESIGN AND SYNTHESIS OF UAAS WITH CATALYTICALLY ACTIVE ORGANIC SIDE CHAINS

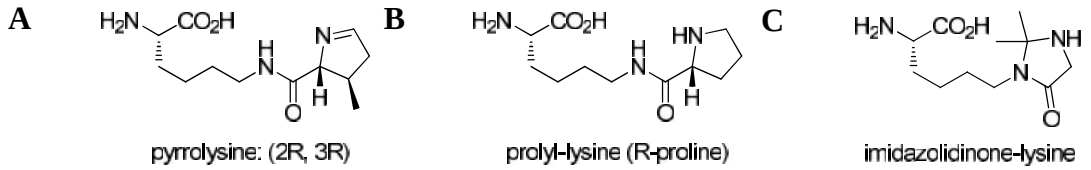
Since metal-containing UAAs could not be incorporated into proteins using codon suppression methods (only a handful of known tRNA synthetases were explored), we next pursued incorporation of amino acids with catalytically active organic side chains. Secondary amines are widely used as organocatalysts for aldol, Diels-Alder and a range of other reactions[5]. Enamines resulting from condensation reaction between amines and ketones have long been recognized as key intermediates in enzyme catalysis[27]. L-Proline-based small peptides have been developed as efficient catalysts for the asymmetric direct aldol reactions of hydroxyacetone with aldehydes[28]. The great synthetic usefulness of the aldol reaction in organic synthesis has driven a rapid development of numerous highly enantioselective chiral catalysts[29]. In nature, many enzymes facilitate amine as key residues. For instance, tautomerase super family enzymes can catalyze aldol reactions by terminal proline[30], while aldolases efficiently catalyze aldol reactions under mild ambient conditions[31]. Furthermore, MacMillan[32] and coworkers have developed several C-C bond-forming reactions using an interwoven activation pathway that combines a photoredox catalyst with an organocatalyst. Incorporation of catalytically active secondary amines into protein scaffolds could be advantageous for merging visible-light photocatalysis and secondary amine organocatalysis in asymmetric chemical synthesis.

Inspired by these reactions, we designed and prepared two UAAs consisting of secondary amine with strong resemblance to pyrrolysine, whose tRNA synthetase accepts a range of structural related amino acids (Figure 1.4). The prolyl side chain of D-Prolyl-lysine

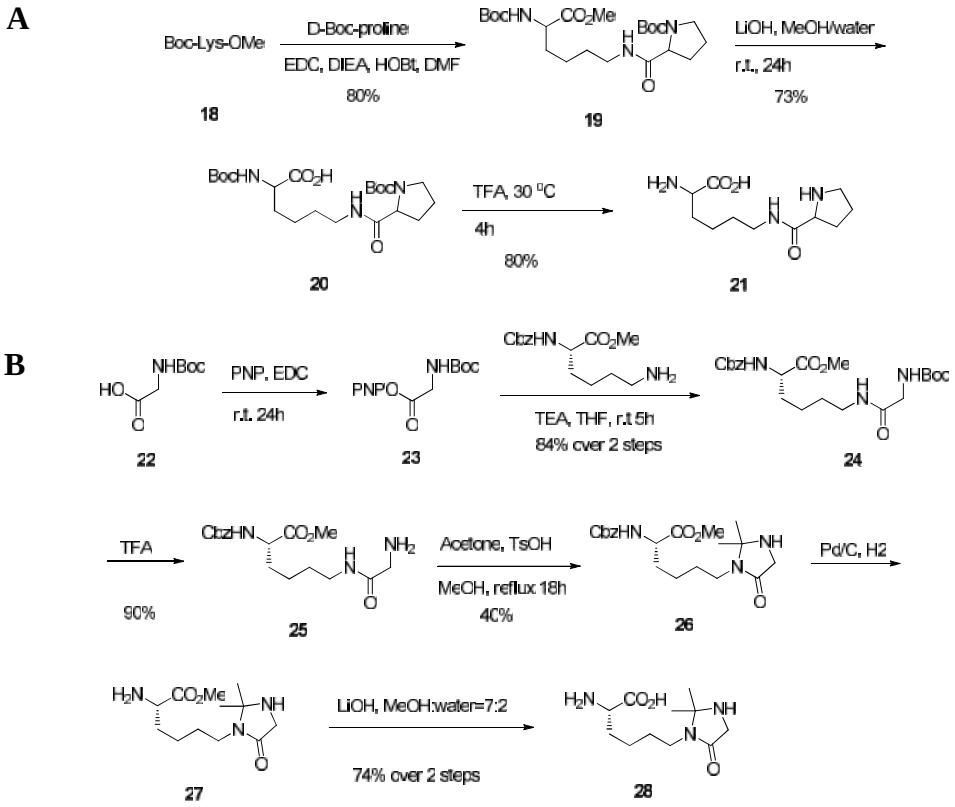
(Figure 1.4B) mimics proline-amide which could catalyze aldol and Michael reactions under physiological conditions[31]. Dieter Soll and coworkers reported in vitro aminoacylation of tRNAPyl by pyrrolysyl-tRNA synthetase with D-prolyl-Lysine, a structure analogue of pyrrolysine[33].

Imidazolidinones have been used as powerful organocatalysts promoting Diels-Alder reactions[34], aldol reactions[35], and 1,3-dipolar cycloaddition[36] through iminium intermediate (Figure 1.4C). We hypothesized that incorporating imidazolidinone-lysine UAA could facilitate secondary structure provided by protein scaffold, and thus introduce enantio and site selectivity, ideally. Two UAAs were synthesized shown in Scheme 1.8.

Figure 1.4. Proposed secondary amine UAAs

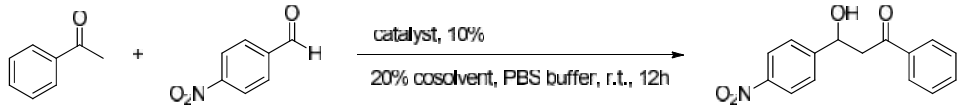


Scheme 1.8. Synthesis of (A) prolyl-lysine and (B) imidazolidinone-lysine



Both compounds can catalyze aldol reactions (Table 1.4). To suppress background reactions, 0.1 equivalent of SDS was added according to previous result[31]. Prolyl-lysine showed positive results in UAA incorporation (with Jewett). Further study of prolyllysine-protein expression and imidazolidinone-lysine incorporation is in progress.

Table 1.4. Two UAAs catalyzed aldol reactions



29 30 31

Entry	Catalyst	Conv.[1]
1	**21**	47
2	**28**	28
3	none	1
4	none[2]	30

[1] Conversion calculated from NMR integration
[2] reaction without SDS

1.5 CONCLUSIONS

In conclusion, we have described synthetic approaches for the preparation of a range of unprotected and *N*-protected organometallic amino acids. These compounds are structural analogues of natural amino acids, making them ideal vehicles for incorporation of organometallic complexes into peptides and potentially proteins. The compatibility of these amino acids with standard peptide coupling conditions allowed the facile preparation of a small set of catalytically active metallopeptides. Two UAAs with catalytically active side chain were synthesized. Both secondary amine based UAAs could catalyze aldol reaction.

1.6 EXPERIMENTAL

Materials:

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. Benzene, dimethylformamide (DMF), acetonitrile (ACN), tetrahydrofuran (THF) and methylene chloride (CH_2Cl_2), were obtained from a PureSolv MD solvent purification system by Innovative Technology (solvent dried over alumina under a N_2 atmosphere). "Extra Dry" grade methanol purchased from Acros was utilized. Diisopropylethylamine (99.5%, redistilled) was purchased from Aldrich and sparged with a N_2 stream for at least 15 min prior to use. Deuterated solvents were obtained from Cambridge Isotope labs. Silicycle silica gel plates (250 μm, 60 F254) were used for analytical TLC and chromatographic purifications were performed on silica gel (40-63 μm). **1a** and **1b**[37†], **6**[38], **10**,[39], **14**[40] and $[CpRu(CO)(CH_3CN)_2]^+PF_6^-$ [41‡] were prepared according to literature procedures.

General Procedures:

Unless otherwise specified, all reactions were prepared in flame or oven-dried glassware under an inert N_2 atmosphere using either syringe and cannula techniques or a nitrogenfilled Vacuum Atmospheres inert atmosphere box. TLC plates were visualized using 254 nm ultraviolet light. ^1H and ^{13}C NMR spectra were recorded at 500 MHz and 125 MHz, respectively, on a Bruker DMX-500 or DRX-500 spectrometer, and chemical shifts are reported relative to residual solvent peaks.6 In some cases, peaks corresponding to the carbon

† Conducted by Dr. Jared Lewis and Dr. Zhihui Zhang
‡ Conducted by Hao Yang

atoms alpha to boron in the ^{13}C NMR spectra for BBN-protected amino acids were not observed, presumably due to reduced signal intensity resulting from boron splitting (typically, either 6 or 8 signals were observed). Infrared spectra were recorded on a KBr plate using a Perkin-Elmer 1600 series Fourier transform spectrometer. High resolution mass spectra were obtained at the University of Chicago Mass Spectrometry Service Center on an Autospec high-resolution double-focusing electrospray ionization/chemical ionization spectrometer with either DEC 11/73 or OPUS software data system. HPLC analysis was conducted using a ZORBAX Eclipse Plus C18 column (4.6 x 150 mm, 3.5 um) with an acetonitrile (ACN)/0.1% aq. TFA gradient (20-100% ACN over 10 min., then 100% ACN for 2 min.).

Compound 2a

1a (1 mmol, 207 mg) was dissolved in MeOH (5 mL), and a solution of 9-BBN (0.5M in THF, 1.5 mmol, 3mL) was added slowly. The resulting solution was heated at reflux for 12 h, cooled to room temperature, and concentrated in vacuo to give an oil which was purified by flash column chromatography (Hex/EtOAc=2/1) to afford **2a** as white solid (0.276 g, 85%). ^{1}H NMR (500 MHz, DMSO): δ 7.91 (d, *J* = 8 Hz, 2H), 7.50 (d, *J* = 8 Hz, 2H), 6.45 (m, 1H), 5.94 (m,1H), 3.90 (m, 1H), 3.25 (dd, *J* = 14.75, 4.5 Hz,1H), 3.02 (dd, *J* = 14. 5, 4.5 Hz,1H), 2.57 (s, 3H), 1.76-1.41 (m, 12H), 1.95-1.76 (s, 2H). ^{13}C NMR (125 MHz, DMSO) δ 197.6, 172.9, 143.3, 135.3, 129.6, 128.2, 55.46, 36.01, 31.3, 31.2, 30.7, 30.5, 26.7, 24.2, 23.9. HRMS-ESI (m/z): calcd for $C_{19}H_{27}BNO_3$ $[M+H]^+$: 328.2084, found: 328.2070.

Compound 2b

1b (1 mmol, 207 mg) was dissolved in MeOH (5 mL) and a solution of 9-BBN (0.5M in THF, 1.5 mmol, 3mL) was added slowly. The resulting solution was heated at reflux for 12 h, cooled to room temperature, and concentrated in vacuo to give an oil which was purified by flash column chromatography (Hex/EtOAc=2/1) to afford **2b** as white solid (0.267 g, 82%). ^{1}H NMR (500 MHz, CDCl3) δ 7.90 (s, 1H), 7.82 (d, *J* = 7.0 Hz, 1H), 7.52 (d, *J* = 7.0 Hz, 2H), 7.41 (t, *J* = 7.0 Hz, 1H), 6.39 (s, 1H), 5.48 (s, 1H), 3.94 (s, 1H), 3.31~3.22 (m, 1H), 3.07 (dd, *J* = 8.5, 15.0 Hz, 1H), 2.52 (s, 3H); 1.74~1.20 (m, 12H), 0.46 (s, 1H), 0.21 (s, 1H); ^{13}C NMR (125 MHz, CDCl3) δ 200.5, 176.4, 138.9, 138.7, 135.6, 130.6, 130.3, 128.4, 57.7, 37.5, 32.7, 32.6, 32.3, 26.9, 25.7, 25.3. HRMS-ESI (m/z): calcd for $C_{19}H_{26}ClBNO_3$ $[M+Cl]^-$: 362.1694, found: 362.1701.

Compound 3a

2b (0.5 mmol, 165 mg) and hydroxylamine hydrochloride (0.75 mmol, 52 mg) were dissolved in 2 mL MeOH. Et_3N (0.75 mmol, 75 mg) was added *via* syringe, the mixture was stirred at room temperature overnight, and the solvent was evaporated. The product was purified by flash chromatography (silica gel, 2:1 hexanes:EtOAc) to afford **3b** as white amorphous solid (0.154 g, 90%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.64 (s, 1H), 7.58 (d, *J* = 6.5 Hz, 1H), 7.37 (d, *J* = 7.5 Hz, 2H), 4.04~4.00 (m, 1H), 3.37 (dd, *J* = 8.5, 15.0 Hz, 1H), 3.12 (dd, *J* = 8.5, 15.0 Hz, 1H), 2.26 (s, 3H); 1.84~1.41 (m, 12H), 0.57 (s, 1H), 0.31 (s, 1H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 176.6, 156.2, 139.3, 138.0, 131.0, 130.1, 128.2, 126.2, 57.7, 37.5, 32.7, 32.6, 32.3, 32.2, 25.7, 25.3, 12.3. HRMS-ESI (m/z): calcd for $C_{19}H_{28}BN_2O_3$ $[M+H]^+$:

343.2193, found: 343.2158.

Compound 3b

Compound **2a** (0.0998 g, 0.305 mmol, 1 equiv), semicarbazide•HCl (0.0425 g, 0.0381 mmol, 1.2 equiv), and NaOAc (0.0306 g, 0.0373 mmol, 1.2 equiv) were added to a round-bottomed flask containing a stir bar. The flask was fitted with a reflux condenser and methanol was added to the flask. The reaction mixture was heated at reflux with stirring for 12 h. The reaction mixture was concentrated onto silica gel and loaded onto a silica gel column equilibrated with 5% MeOH/CH_2Cl_2. The desired product was eluted with 5% MeOH/CH2Cl2 to provide **3b** as an off white solid (0.104 g, 89%). ^{1}H NMR (500 MHz, DMSO) δ 9.29 (s, 1H), 7.78 (d, *J* = 10.0 Hz, 2H), 7.34 (d, *J* = 10.0 Hz, 2H), 6.47 (s, 3H), 5.82 (m, 1H), 3.84 (s, 1H), 3.17 (m, 1H), 2.96 (m, 1H), 2.17 (s, 4H), 1.80~1.15 (m, 12H), 0.47(s, 2H). ^{13}C NMR (125 MHz, MeOD): δ 176.4, 160.5, 147.7, 138.7, 138.6, 130.4, 127.7, 57.5, 37.2, 32.6, 32.5, 32.3, 32.2, 25.6, 24.1, 13.4. HRMSESI (m/z): calcd for $C_{20}H_{31}BN_4O_4$ $[M+H_2O]^+$: 402.2442, found: 402.2439.

Compound 3c

2a (0.5 mmol, 165 mg) and hydroxylamine hydrochloride (0.75 mmol, 52 mg) were dissolved in 2 mL MeOH. Et_3N (0.75 mmol, 75 mg) was added *via* syringe, the resulting mixture was stirred at room temperature overnight, and the solvent was evaporated. The product was purified by flash chromatography (silica gel, 2:1 hexanes:EtOAc) to afford **3b**as white amorphous solid (0.159 g, 93%).^{1}H NMR (500 MHz, MeOD): 7.80 (d, *J* = 8 Hz, 2H),

7.32 (d, *J* = 8 Hz, 2H), 6.45 (t, 1H), 5.94 (t, 1H), 3.91 (m, 1H), 3.21~3.16 (m, 1H), 3.06 (dd, *J* = 14.5 7.0 Hz, 1H), 2.56 (s, 3H), 1.72~1.33 (m, 12H), 0.46 (s, 1H), 0.19 (s, 1H). ^{13}C NMR (125 MHz, MeOD): 176.4, 157.1, 154.3, 139.8, 137.4, 130.7, 128.3, 57.3, 37.2, 32.6, 32.5, 32.3, 32.2, 25.6, 25.2, 14.3. HRMS-ESI (m/z): calcd for $C_{21}H_{31}BN_3O_3$ $[M+CH3CN+H]^+$: 384.2459, found: 384.2443.

Compound 5a

Compound **3a** (0.5006 g, 1.07 mmol, 1 equiv), Na_2PdCl_4•$3H_2O$ (0.3772 g, 1.07 mmol, 1 equiv), NaOAc•$3H_2O$ (0.1460 g, 1.07 mmol, 1 equiv) and a stir bar were added to a 20 dram vial. The solids were dissolved in 10 mL methanol, the vial was capped, and the mixture was stirred at room temperature for three days. The mixture was filtered through a fine frit to yield a yellow solution and a grey solid. Approximately 10 mL water was added to the solution to precipitate the product, **4a**, as a yellow solid, which was collected by filtration (0.4989 g, 71%). Without further purification, compound **4a** (0.1025 g, 0.1061 mmol, 1 equiv) was added to a 20 dram vial containing a stir bar. The solid was dissolved in 10 mL THF, ethylenediamine (100 µL, 1.496 mmol, 14 equiv) was added dropwise, and the vial was capped. A solid immediately precipitated and the mixture was allowed to stir for 24 h. The mixture was filtered to yield the product, **5a**, as a beige solid (0.075 g, 82%, 58% over two steps). ^{1}H NMR (500 MHz, MeOD): 6.89 (s, 1H), 6.79 (d, *J* =7.5Hz, 1H), 6.69 (d, *J* =7.5Hz, 1H), 3.59 (m, 1H), 3.09 (dd, *J1* = 14.5 Hz, *J2* = 4.5 Hz, 1H), 2.79 (dd, *J1* = 14.5 Hz, *J2* = 4.5 Hz, 1H), 2.72-2.63 (m, 4H), 2.61 (s, 3H). ^{13}C NMR (125 MHz, MeOD): 175.0, 165.2, 154.8, 148.9, 133.6, 133.2, 127.6, 125.0, 57.9, 47.3, 47.1, 38.7, 10.8. HRMS-ESI (m/z): calcd for

$C_{13}H_{21}N_4O_3Pd$ [M-Cl]$^-$: 387.0648, found: 387.0650.

Compound 5b

Compound **3b** (0.0504 g, 0.131 mmol, 1 equiv), Na_2PdCl_4•$3H_2O$ (0.0505 g, 0.145 mmol, 1.1 equiv), and a stir bar were added to a 20 dram vial. The solids were dissolved in 1 mL H_2O and 5 mL ethanol, the vial was capped, and the mixture was stirred at room temperature for three days. The resulting solid was filtered through a fine frit to yield the product, **5a**, as a yellow solid, which was rinsed with water and diethyl ether (0.0567 g, 83%). Without further purification, compound **5a** (0.025 g, 0.048 mmol, 1 equiv) was added to a 20 dram vial containing a stir bar. The solid was dissolved in 4.5 mL THF, ethylenediamine (30 μL, 0.45 mmol, 9 equiv) was added, and the vial was capped. A solid immediately precipitated and the mixture was allowed to stir for 24 h. The mixture was filtered to yield the product, **5b**, as a light yellow solid. This material was further purified by precipitation from methanol solution via slow diffusion of ether and filtration (0.0050 g, 23%, 20% over two steps).^{1}H NMR (500 MHz, D2O) δ 7.51 (d, *J*= 8.0 Hz, 1H), 7.11 (d, *J* = 8.0 Hz, 1H), 6.92 (s, 1H), 3.93 (t, *J* = 6 Hz, 1H), 3.08~3.16 (m, 2H), 2.81 (dd, *J* = 5.5, 17.0 Hz, 4H), 2.40 (s, 3H); ^{13}C NMR (125 MHz, D_2O): incomplete data due to poor solubility. δ 194.9, 136.6, 133.8, 129.4, 126.4, 55.8, 45.7, 42.8, 36.8, 13.8. HRMSESI (m/z): calcd for $C_{19}H_{26}ClBNO_3$ [M-Cl]$^+$: 429.0867, found: 429.0851.

Compound 5c

Compound **3c** (0.1339 g, 0.39 mmol, 1.1 equiv), [$IrCl_2$Cp*]2 (0.1409 g, 0.18 mmol,

0.5 equiv), NaOAc (0.0583 g, 0.71 mmol, 2 equiv) and a stir bar were added to a 20 dram vial. The solids were dissolved in 8 mL methanol, the vial was capped, and the mixture was stirred at room temperature for 16 h. Approximately 10 mL water was added to the solution to completely precipitate the product, **4c**, as a yellow solid, which was collected by filtration (0.184 g, 67%). Without further purification, compound **4c** (0.1320 g, 0.1875 mmol, 1 equiv) was added to a 20 dram vial containing a stir bar. The solid was dissolved in 18 mL THF, ethylenediamine (130µL, 1.87 mmol, 10 equiv) was added dropwise, and the vial was capped. A solid gradually precipitated and the mixture was allowed to stir for 24 h. The mixture was filtered to yield the product, **5c**, as a yellow solid. This material was added to a vial and dissolved in a minimal amount of methanol. The vial was gently placed in a larger vial containing diethyl ether, and the product slowly precipitated as a mixture of diastereomers as diethyl ether diffused into the methanol solution. The product was collected by filtration (0.054 g, 45%, 42% over two steps). ^{1}H NMR (500 MHz, D_2O): δ 7.55 (d, *J* =10 Hz, 1H), 7.22 (m, 1H), 6.95 (m, 1H), 3.51 (m, 1H), 3.04 (m, 1H), 2.84 (m, 1H), 2.43 (m, 4H), 2.25 (s, 3H), 1.69 (s, 15H). Because the product diastereomers were inseparable and poorly soluble, acquisition of a useful ^{13}C NMR spectrum was not possible. HRMS-ESI (m/z): calcd for $C_{23}H_{36}IrN_4O_3$ $[M\text{-}Cl+H]^+$: 609.2417, found: 609.2433.

Compound 15

A stirred solution of **14**[40] (1 mmol, 249 mg) in water (2 mL, pH = 8-9 adjusted by 1 M LiOH) was treated with $CoCl_2 \bullet 6H_2O$ (0.01 mmol, 3mg) and *Aspergillus* acylate I (0.7 U/mg, 50 mg), and then heated to 40 ºC. After stirring for 14 h, the reaction mixture was cooled to

room temperature and acidified with 1N HCl to pH 2, and the unreacted 1 was extracted with EtOAc (3 mL x 3). The aqueous layer was heated with charcoal at 90 ºC for 15 min, cooled to room temperature, filtered and concentrated in vacuo to give product **15** (0.165 g, 80%). ^{1}H NMR (500 MHz, MeOD) δ 7.94 (d, *J* = 8.0 Hz, 2H), 7.42 (d, *J* = 8.0 Hz, 2H), 4.03 (t, *J* = 6.5 Hz, 1H), 3.33 (m, 1H), 3.21 (m, 1H), 2.62 (s, 3H); ^{13}C NMR (125 MHz, MeOD) δ 203.6, 173.6, 141.7, 135.7, 129.8, 129.3, 41.5, 36.4, 26.3; HRMS-ESI (m/z): $[M+H]^+$ calcd for $C_{11}H_{14}NO_3$, 208.0973; found, 208.0937.

Compound 16

A stirred solution of **15** (1 mmol, 207 mg) in 1,4-dioxane (2 mL) was added water (2 mL) and $NaHCO_3$ (2 mmol, 168 mg) . The resulting solution was treated with Boc_2O (1.1 mmol, 140 mg) and stirred for 12h at room temperature. The mixture was washed with Et_2O to remove excess Boc_2O and the aqueous layer was acidified with 1 N HCl and extracted with EtOAc (5 mL x 3). The combined organic layer was washed with water and brine, dried over Na_2SO_4, filtered, and concentrated in vacuo to provide white amorphous solid, which was used for next reaction without further purification. The *N*-Boc amino acid (0.75 mmol, 230 mg) and *O*-methy ydroxylamine hydrochloride (1.1 mmol, 92 mg) were dissolved in 2 mL MeOH. Et_3N (1.5 mmol, 150 mg) was added via syringe and after stirring at room temperature overnight the solvent was evaporated. The product was extracted into EtOAc (5 mL) and the organic phase was washed with 1 M HCl (3 mL), water and brine. After drying over $MgSO_4$ the solvent was evaporated. Purification by flash chromatography (silica gel, 1:1 hexanes:EtOAc) afforded the product as white amorphous solid **16** (0.300 mg, 89%). ^{1}H

NMR (500 MHz, $CDCl_3$) δ 7.61 (d, *J* = 8.0 Hz, 2H), 7.22 (d, *J* = 8.0 Hz, 2H), 5.02 (d, *J* = 8.0 Hz, 1H), 4.65 (d, *J* = 6.5 Hz, 1H), 4.01 (s, 3H), 3.25-3.11 (m, 2H), 2.23 (s, 3H); 1.45 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 175.9, 155.5, 154.8, 137.1, 135.4, 129.6, 126.4, 80.4, 61.9, 54.2, 37.6, 28.3, 12.8. HRMS-ESI (m/z): $[M+Na]^+$ calcd for $C_{17}H_{24}N_2O_5Na$, 359.1583; found, 359.1541.

Compound 17

Compound **16** (1 mmol, 336 mg) and $NaPdCl_4$. $3H_2O$ (1 mmol, 348 mg) were dissolved in 10 mL MeOH. NaOAc (2 mmol, 164 mg) was added and the solution was stirred at room temperature for 3 days. The Pd black was filtered off and then 10 mL H_2O was slowly added to the filtrate. The formed yellow precipitate was collected by filtration as product **17** (0.409 g, 86%). ^1H NMR (500 MHz, MeOD) δ 7.34 (s, 1H), 7.15 (d, *J* = 7.5 Hz, 1H), 6.91 (d, *J* = 7.5 Hz, 1H), 4.24 (br, 1H), 3.79 (s, 3H), 3.035-2.75 (m, 2H), 2.27 (s, 3H); 1.30 (s, 9H); ^{13}C NMR (125 MHz, MeOD) δ 179.5, 175.2, 157.8, 153.4, 140.7, 140.4, 135.6, 128.5, 127.2, 80.7, 63.0, 56.2, 39.1, 28.9, 12.2. HRMS-ESI (m/z): $[MCl+ MeOH]^+$ calcd for $C_{18}H_{27}N_2O_6Pd$, 473.0904; found, 473.0873.

Compound 20

Peptide synthesis was conducted using a modified variant of a procedure described by Miller and Movassaghi[25]. A flask equipped with magnetic stir bar was charged with MeO-Phe-NH_2 .HCl (108 mg, 0.5 mmol), Boc-Aib-OH (101 mg, 0.7 mmol), HOBt·H_2O (95 mg, 0.7 mmol) and EDC·HCl (134 mg, 0.7 mmol). Dry DCM (2.5 mL, 0.2 M), and triethylamine

(0.14 mL, 1.0 mmol) was added in one portion. The resulting solution was stirred at room temperature for 12 h. The reaction mixture was then diluted with 5 mL of DCM, transferred to a separatory funnel and sequentially washed with 5 mL of 0.5 M aqueous solution of citric acid and 5 mL of saturated aqueous solution of $NaHCO_3$. The organic phase was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The white solid was treated with 4 M solution of HCl in dioxane (1.0 mL). After stirring at room temperature for 0.5 h reaction mixture was concentrated under reduced pressure in a rotary evaporator followed by exposing the product to high vacuum for 1 h. To the white powder obtained were added HOBt·H_2O (95 mg, 0.7 mmol), EDC·HCl (134 mg, 0.7 mmol) and Boc-Pro-OH (107 mg, 0.5 mmol), the mixture was suspended in dry DCM (2.5 mL, 0.2 M), and triethylamine (0.14 mL, 1.0 mmol) was added in one portion. Subsequent coupling, work-up, and Boc removal were carried out in a manner analogous to the one described above. To the dry deprotection product (**19**) thus obtained were added HOBt·H_2O (95 mg, 0.7 mmol), EDC·HCl (134 mg, 0.7 mmol) and **17** (239 mg, 0.5 mmol) the mixture was suspended in dry DCM (2.5 mL, 0.2 M), and triethylamine (0.14 mL, 1.0 mmol) was added in one portion. The resulting solution was stirred at room temperature for 12 h. The reaction mixture was then diluted with 5 mL of DCM, transferred to a separatory funnel and sequentially washed with 5 mL of saturated aqueous solution of NH_4Cl and 5 mL of saturated aqueous solution of $NaHCO_3$. The organic phase was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Crude product obtained after the final peptide coupling was purified by flash column chromatography on silica gel (DCM/MeOH = 20/1~10/1) to yield 0.127 g (31% yield) of metallopeptide as a yellow solid. ^{1}H NMR (500 MHz, MeOD) δ 7.24~7.05 (m, 8H),

7.01~6.90 (m, 2H), 6.76 (s, 1H), 4.61~4.34 (m, 2H), 4.21 (t, *J* = 7.5 Hz, 1H), 3.78 (s, 3H), 3.48 (s, 3H), 3.01~2.77 (m, 6H), 2.16 (s, 3H), 2.09~1.82 (m, 4H), 1.39~1.16 (m, 15H); ^{13}C NMR (125 MHz, MeOD) δ 179.2, 176.9, 173.7, 173.6, 171.8, 157.4, 141.5, 139.9, 138.4, 134.0, 130.6, 130.5, 129.7, 129.4, 128.4, 127.8, 80.9, 62.9, 62.5, 58.1, 58.0, 56.1, 54.7, 52.5, 38.6, 38.5, 30.5, 28.8, 28.7, 27.1, 26.5, 25.7, 24.5, 12.0. HRMS-ESI (m/z): calculated for $C_{38}H_{54}N_5O_8Pd$ $[M-Cl]^-$: 784.2538, found: 784.2541.

N-Boc-Gly-PNP (**23**): Followed reported protocol[42]. To a mixture of **1** (1.206 g, 6.88 mmol), 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 1.94g, 10.12mmol) and p-nitrophenol (940mg, 6.76mmol) was added DCM (15 mL). The mixture was stirred at r.t. for 24h. After completion, the resulting solution was washed with H_2O (1×50 mL), and brine (1×50 mL), and dried over $MgSO_4$. Removing the solvent using rotary evaporation provided the desired product as a yellow oil. The compound was applied to next step without further purification. There was residual p-nitrophenol detected by HPLC.

(N-Boc-Glycyl)-N-(Cbz)lysine Methyl Ester (**24**): Followed reported protocol[43] with slight modification. **23** (590 mg) and N-Z-L-Lysine methyl ester hydrochloride (555 mg) were dissolved in 3 mL of anhydrous THF. After addition of triethylamine (150uL), the solution was stirred at room temperature for 4 h and then concentrated to complete dryness. The residue was taken up in AcOEt and washed with 0.2 M HCl. The organic layer was dried with Na_2SO_4, filtered, and concentrated under vacuum. The residue was purified by column chromatography on silica gel eluting with dichloromethane-ethyl acetate (1:1) (R_f =0.3) to

give **24** as a pale yellow oil (760mg, 84%). Redissolved compound with 50mL DCM and washed with 50mL saturated $NaHCO_3$ to remove trace amount of co-eluted p-nitrophenol. ^1H NMR (500 MHz, $CDCl_3$) δ 7.36 (b, 5H) 6.11 (s, 1H), 5.38 (d, 1H), 5.11 (bs, 2H), 4.35 (q, 1H), 3.74 (bs, 5H), 3.25 (t, 1H), 1.61-1.89 (m, 2H), 1.50-1.57 (m, 2H), 1.79 (s, 9H), 1.31-1.41 (m, 2H).

Glycyl-N-(Cbz)lysine Methyl Ester (**25**): Followed reported protocol[43] with slight modification. To **24** (760mg), 6 mL of trifluoroacetic acid was added under ice-bath. The solution was stirred at 30 ºC for 4 h before residual TFA was evaporated under vacuum. The residue was redissolved in CH_2Cl_2 and washed with 1 M NaOH. The organic layer was dried, and concentrated to dryness. **25** was obtained as a white glassy solid (532 mg, 90%). ^1H NMR (500 MHz, $CDCl_3$) δ 7.20-7.40 (b, 5H), 5.85 (s, 1H), 5.38 (d, 1H), 5.04 (m, 2H), 3.55-3.80 (m, 5H), 3.12 (s, 2H), 1.55-1.79 (m, 2H), 1.23-1.50 (m, 4H).

Dimethylimidazolidinone-N-(Cbz)lysine methyl ester (**26**): Followed reported protocol[44] with slight modification. To **25** (532 mg) was added molecular sieves (500mg), TsOH (1.2eq), MeOH (24mL) and dry acetone (4.8mL). The mixture was refluxed for 24h, and then concentrated in vacuo. Resuspended residues with DCM and filtered. DCM was then washed with pH=3-4 HCl solution, dried and concentrated down. **26** was obtained as a yellow solid (237mg, 40%). ^1H NMR (500 MHz, $CDCl_3$) δ 7.20-7.40 (b, 5H) 5.85 (s, 1H), 5.38 (d, 1H), 5.04 (m, 2H), 3.55-3.80 (m, 5H), 3.12 (s, 2H), 1.55-1.79 (m, 2H), 1.23-1.50 (m, 4H).

Imidazolidinone lysine methyl ester (**27**): **26** (237 mg) and Pd/C (10% w/w) (70 mg) were suspended in 15mL of anhydrous MeOH with H2 balloon. The mixture was stirred at room temperature for 4 h and then filtered through celite. Filtrate was concentrated down and got **6** (123mg, 80%) without purification. ^{1}H NMR (500 MHz, MeOD) δ4.06 (t,1H), 3.90 (s, 2H), 3.85 (s, 3H), 1.85-2.05 (m, 2H), 1.40-1.56 (m, 2H). 2H was buried under MeOH solvent peak.

Imidazolidinone lysine (**28**): **27** (123 mg) and LiOH monohydrate (240 mg) were dissolved in 24 mL of 7:2 MeOH/water. The suspension was stirred at room temperature for 24 h and then concentrated to complete dryness under highvac. The residue was taken up in water and acidified to pH 1 to 2 with 0.5 M HCl. The sample was purified via DOWEX 50WX4 resin to give 86mg of **7** as an off-white solid (73%). Please followed[45] for detailed procotol. ^{1}H NMR (500 MHz, D_2O) δ 4.08(s, 2H), 4.00(t, 1H), 3.65(m, 2H), 1.90-2.09(m, 2H), 1.77(s, 6H), 1.40-1.75(m, 4H). ^{13}C NMR (126 MHz, D_2O) δ 173.1, 167.3, 80.7, 53.6, 44.1, 39.9, 29.6, 27.9, 27.7 (may not be separated, gem-dimethyl), 23.8, 21.9. Solubility: partially in MeOH, good in water and DMSO. Stability: tolerate 1M HCl and 1M NH_4OH

1.7 ACKNOWLEDGEMENT

This work was supported by the University of Chicago, Department of Chemistry. I wish to thank Hao Yang for assistance with organic synthesis. I am grateful to Dr. Zhihui Zhang for all initial work and nice suggestions on peptide synthesis. I would like to thank Prof. Steven Kent and Dr. Michal Avital-Shmilovic for assistance with SPPS.

1.8 REFERENCES

(1) Lewis, J. C. *Curr Opin Chem Biol* **2015**, *25*, 27.
(2) Wallace, S.; Balskus, E. P. *Curr Opin Biotechnol* **2014**, *30*, 1.
(3) de Meijere, A., Brase, S., Oestreich, M., Eds.; Wiley-VCH: 2013.
(4) Yu, J.-Q. In *Topics in Current Chemistry*; Shi, Z., Ed. 2010; Vol. 292.
(5) List, B. *Chem Commun (Camb)* **2006**, 819.
(6) Cossy, J., Arseniyadis; S., M., C., Ed.; John Wiley & Sons: 2010.
(7) Hartwig, J. *Organotransition Metal Chemistry: from Bonding to Catalysis*; University Science Books:: CA, 2010.
(8) Bar-Even, A.; Noor, E.; Savir, Y.; Liebermeister, W.; Davidi, D.; Tawfik, D. S.; Milo, R. *Biochemistry* **2011**, *50*, 4402.
(9) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90.
(10) Lewis, J. C. *ACS Catalysis* **2013**, *3*, 2954.
(11) Wagner, A. *Nat. Rev. Genet.* **2008**, *9*, 965.
(12) Bershtein, S.; Tawfik, D. S. *Curr Opin Chem Biol* **2008**, *12*, 151.
(13) Tracewell, C. A.; Arnold, F. H. *Curr Opin Chem Biol* **2009**, *13*, 3.
(14) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **2012**, *485*, 185.
(15) Tippmann, E. M.; Schultz, P. G. *Tetrahedron* **2007**, *63*, 6182.
(16) Zhang, Z.; Yang, H.; Zhang, C.; Lewis, J. C. *Organometallics* **2012**, *31*, 7328.
(17) Vairaprakash, P.; Ueki, H.; Tashiro, K.; Yaghi, O. M. *J Am Chem Soc* **2011**, *133*, 759.
(18) Dent, W. H. R. E., W.; Fields, S. C.; Parker, M. H.; Tromiczak, E. G. *Org Lett* **2002**, *4*, 1249.
(19) Diego A. Alonso, C. N., M Carmen Pacheco *Adv Synth Catal* **2002**, *2*.
(20) Howard Bregman, D. S. W., G. Ekin Atilla, Patrick J. Carroll, and Eric Meggers *J Am Chem Soc* **2004**, *126*, 13594.
(21) Merrifield, B. *Life During a Golden Age of Peptide chemistry : The Concept and Development of Solid-Phase Peptide Synthesis*; American Chemical Society: Washington, DC, 1993.
(22) Miller, S. J. *Acc Chem Res* **2004**, *37*, 601.
(23) Matthew B Francis, T. F. J. a. E. N. J. *Current Opinion in Chemical Biology* **1998**, 422.
(24) H. Keith Chenault, J. D. a. G. M., Whitesides *J Am Chem Soc* **1989**, *111*, 6354.
(25) Kolundzic, F.; Noshi, M. N.; Tjandra, M.; Movassaghi, M.; Miller, S. J. *J Am Chem Soc* **2011**, *133*, 9104.
(26) Larry E. Overman, C. E. O., and Mary M. Pavan *Org Lett* **2003**, *5*, 1809.
(27) Wolfgang Notz, F. T., Carlos F. Barbas, III *Acc Chem Res* **2004**, *37*, 580.
(28) Zhuo Tang, Z.-H. Y., Lin-Feng Cun, Liu-Zhu Gong, Ai-Qiao Mi, and Yao-Zhong Jiang *Org Lett* **2004**, *6*, 2285.
(29) Trost, B. M., Fleming, I., Heathcock, C. H., Eds; In *Comprehensive Organic Synthesis*; Oxford: 1991; Vol. 2.
(30) Zandvoort, E.; Baas, B. J.; Quax, W. J.; Poelarends, G. J. *Chembiochem* **2011**, *12*, 602.
(31) Nobuyuki Mase, Y. N., Naoko Ohara, Hidemi Yoda, Kunihiko Takabe, Fujie Tanaka,‡ and Carlos F. Barbas III *J Am Chem Soc* **2006**, *128*, 734.

(32) Nicewicz, D. A.; MacMillan, D. W. C. *Nature* **2008**, *322*, 77.
(33) Polycarpo, C. R.; Herring, S.; Berube, A.; Wood, J. L.; Soll, D.; Ambrogelly, A. *FEBS Lett* **2006**, *580*, 6695.
(34) MacMillan, D. W. C. *J Am Chem Soc* **2002**, *11*, 2458.
(35) Ian Storer, R.; MacMillan, D. W. C. *Tetrahedron* **2004**, *60*, 7705.
(36) Wendy S. Jen, J. J. M. W., and David W. C. MacMillan *J Am Chem Soc* **2000**, *122*, 9874.
(37) Zhang, Z. S., B. A. C.; Wang, L.; Brock, A.; Cho, C.; Schultz, Peter G. *Biochemistry* **2003**, *42*, 6735.
(38) Hyun Soo Lee, G. S., Peter G. Schultz, Feng Wang *J Am Chem Soc* **2009**, *131*, 2481.
(39) Dent, W. H. E., W. R.; Fields, S. C. *Org. Lett.* **2002**, *4*, 1249.
(40) K. Laumen, K. G., O. *Engineering in Life Sciences* **2006**, *6*.
(41) Gill, T. P. M., K. R. *Organometallics* **1982**, *1*, 485.
(42) Benkovics, T., Guzei, I. A., Yoon, T. P. *Angew. Chem. Int. Ed.* **2010**, *49*, 9153 –9157
(43) Morandeau, L., Benoist, E., Loussouarn, A., Ouadi, A., Lesaec, P., Mougin, M., Faivre-Chauvet, A., Le Boterff, J., F. Chatal, J., Barbet, J., Gestin, J. F. *Bioconjugate Chem.* **2005**, *16*, 184-193
(44) Ahrendt, K. A., Borths, C. J., and MacMillan, D. W. C. *J. Am. Chem. Soc.*, **2000**, *122*, 4243-4244
(45) Link, A. J., Vink, M. K. S., Tirrell, D. A. *Nature Protocols* **2**, 2007, 1879-1883

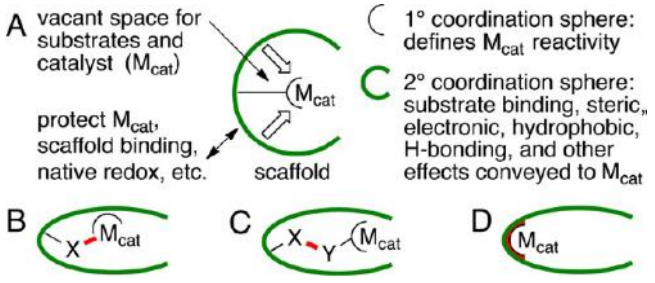
CHAPTER TWO

MANGANESE TERPYRIDINE ARTIFICIAL METALLOENZYMES FOR BENZYLIC OXYGENATION AND OLEFIN EPOXIDATION

2.1 INTRODUCTION TO ARTIFICIAL METALLOENZYMES

This work was published and coauthored with Poonam Srivastava and Ken Ellis-Guardiola[1]. Transition metal catalysts and enzymes possess unique and often complementary properties that have made them important tools for chemical synthesis. Artificial metalloenzymes (ArMs) formed by incorporating synthetic metal catalysts into protein scaffolds have the potential to impart to chemical reactions selectivity that would be difficult to achieve using metal catalysts alone. Researchers have developed systems in which secondary coordination sphere[2] effects impart selectivity to metal catalysts[3,4], accelerate chemical reactions[5], and are systematically optimized via directed evolution[3,4]. In general, methods can be categorized into three types[6]: binding metal ions and complexes, covalent linkage of metal catalysts and ligands, and noncovalent linkage of metal catalysts and ligands (Figure 2.1).

Figure 2.1. (A) General hybrid catalyst structure and structures of hybrid catalysts formed via (B) coordination of metal ions or catalysts, (C) Covalent scaffold modification, and (D) Noncovalent scaffold modification (taken from ref[6])



1) Binding Metal Ions and Complexes: Preparing an ArM typically involves combines a metal catalyst precursor (M) with protein as ligands (L) via one or more metal−ligand bonds (L−M). Proteins, which contain a range of metal-binding N, O, and S functional groups within well-defined, three dimensional, chiral structures, can be viewed as yet another step along this ligand complexity continuum and have therefore attracted attention as scaffolds for ArM formation via metal coordination (Figure 2.1B)[7,8].
2) Covalent Linkage of Metal Catalysts and Ligands: Coordinating metals with proteins can enable ArM formation in much the same way that metals and ligands are mixed together to form small molecule catalysts. This approach is generally limited to coordination by the 20 natural amino acids, and many reactions are catalyzed by metals with ligands not found in nature (e.g., carbenes, phosphines, etc.). Although UAA mutagenesis can be used to expand this scope, this process itself requires extensive engineering for each desired amino acid and is limited to complexes that can be formed in the presence of a protein. To more readily expand the range of catalysts that can be incorporated into proteins, researchers have developed methods to covalently link synthetic, catalytically active transition metal cofactors to proteins (Figure 2.1C)[9]. At a minimum, covalent ArM formation (bioconjugation) requires a scaffold protein containing a uniquely reactive residue (typically a nucleophile) and a cofactor substituted with the corresponding reaction partner (typically an electrophile), both of which present unique synthetic challenges.
3) Noncovalent Linkage of Metal Catalysts and Ligands: This method incorporates metal cofactors into scaffold proteins without the need for direct coordination of the metal by the scaffold (Figure 2.1D). This approach eliminates the need for covalent scaffold

modification while still allowing the use of diverse cofactors for ArM formation, but requires specific scaffold-cofactor interactions that restrict the range of scaffold proteins that can be used.

New catalysts for non-directed hydrocarbon functionalization have great potential for applications in organic synthesis[10,11]. By avoiding the need for directing groups, these species expand the range of substrates on which they can act and eliminate synthetic steps and byproducts associated with directing group installation and removal[12]. A number of catalysts, including dirhodium tetracarboxylate complexes[13] and several different iron[14] and manganese[15,16], complexes, are particularly notable in this regard and are widely used for insertion of carbene, nitrene[17], and oxo[12] fragments into C-H bonds and olefins. While reactivity trends for such catalysts have been outlined, controlling their selectivity[12], particularly on complex substrates,[10,11] remains difficult.

Crabtree and Brudvig have demonstrated that the selectivity of Kemp's-triacid-based Mn-terpyridine complexes for oxygenation of certain carboxylic acid-substituted substrates can be controlled by catalyst-substrate hydrogen bonding[18-21]. The reported examples illustrate the potential for supramolecular interactions to control the selectivity of Mn-terpyridine catalysts but require the use of a carboxylic acid directing group. We hypothesized that incorporating Mn-terpyridine cofactors into protein scaffolds would lead to artificial metalloenzymes (ArMs)[6] in which the selectivity of the Mn cofactor could be controlled by substrate binding to the protein. Mn-terpyridine complexes catalyze a number of oxygenation reactions[22], ethereal α-C-H oxidation[22], and hydrocarbon desaturation[23], suggesting that the proposed ArMs could possess broad reaction scope.

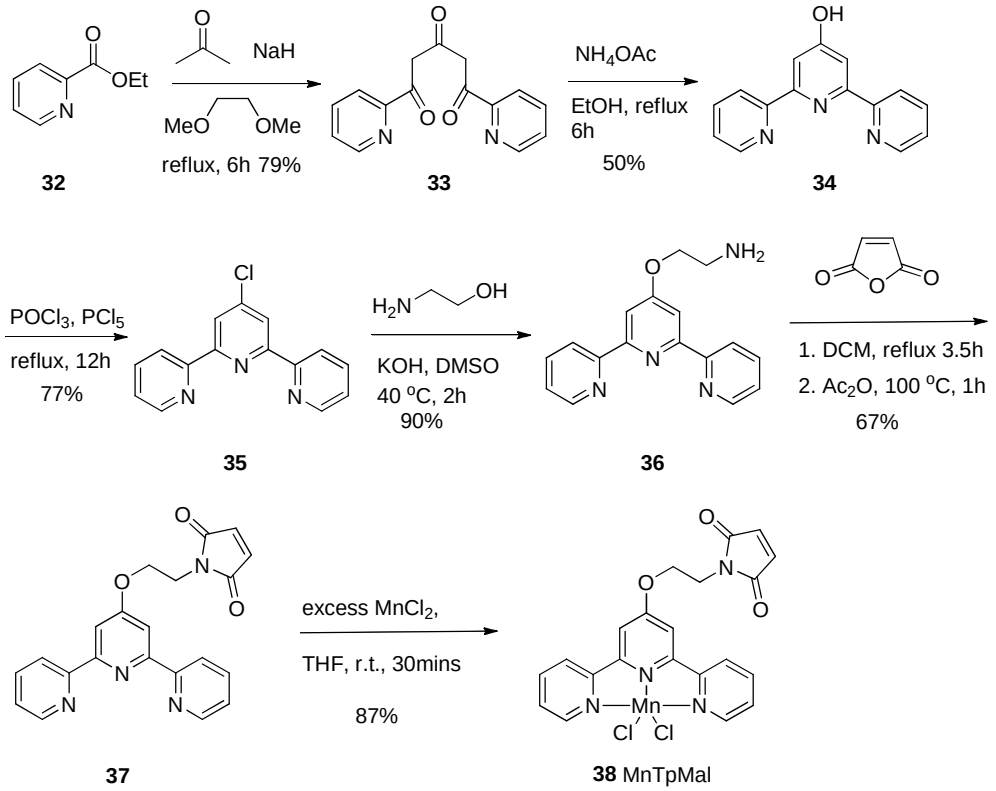
The utility of ArMs for selective catalysis has been demonstrated for a number of different reactions[24]. Hayashi recently demonstrated that myoglobin reconstituted with a Mn-porphycene cofactor catalyzes benzylic oxygenation[25], but we are unaware of any reports of similar reactions using ArMs generated via covalent scaffold modification[6]. While biocatalysts (e.g., cytochromes P450) for this reaction exist, altering their substrate scope and site selectivity can be challenging[26,27]. The potential ease with which a Mn-terpyridine cofactor could be incorporated into different proteins[9] to modulate selectivity led us to develop a robust platform to explore the feasibility of this approach.

2.2. RESULT AND DISCUSSION

A covalent linkage strategy was used to bioconjugate Mn terpyridine to scaffold proteins. Smyth[28] reported in 1964 that *N*-Ethylmaleimide could specifically react with cysteine sulfhydro sidechain to yield a thioether through Michael addition with α,β-unsaturated carbonyl compounds. Maleimide reactions are specific for sulfhydryl groups at pH range 6.5-7.5. At pH=7, the reaction rates are 1000 times higher than other nucleophiles in protein, like side-chain amines on lysine[29].

Maleimide-substituted cofactor **38** was synthesized[30,31] from the 2-acetylpyridine **32**. Terpyridine phenol compound **34** was prepared in two steps[32]. Considering the stability of maleimide cofactor in aqueous solution, an ethanolamine linker was installed as *N*-alkyl group on melaimide can prevent hydrolytic decomposition in water[29]. Maleimide moiety was introduced upon terpyridine amine and thus yield desired cofactor **38** [30,33] (Scheme 2.1).

Scheme 2.1. Synthesis of maleimide-substituted Mn terpyridine cofactor 38

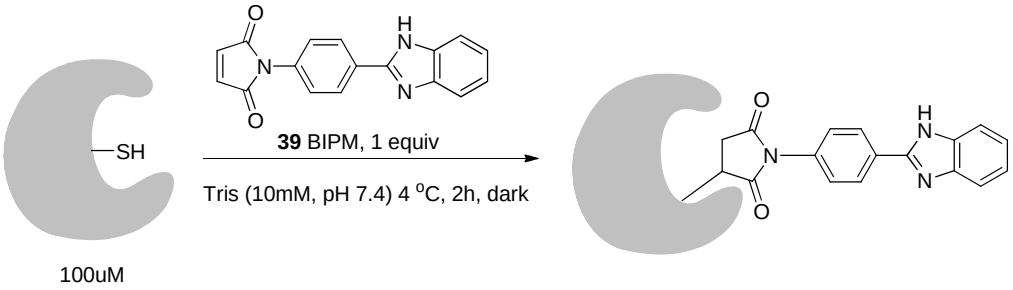


Cysteine mutants of tHisF[34] and apo-nitrobindin (Nb)[35] were then prepared as scaffolds for initial bioconjugation studies. tHisF from *Thermotoga maritima* constitutes the synthase subunit of the glutaminase-synthase bi-enzyme complex that catalyzes the formation of imidazole glycerol phosphate in histidine biosynthesis[34,36]. Monomeric tHisF contains 253 amino acid residues and has a molecular mass of 27.7 kDa. This protein could be an ideal candidate for hosting synthetic catalysts for several reasons. First, protein expression was optimized and could produce 100 mg of protein per liter of culture medium[30]. Second, the pronounced thermostability of tHisF allows a heat treatment and concentration as a straightforward way to parallelize purification. Third, the crystal structure of tHisF[37] suggests that the enzyme has a $(\beta\alpha)_8$-barrel structure which might be a convenient site for covalent anchoring of synthetic transition-metal catalysts. Nitrobindin also has the β-barrel concave

structure, which inherently transports an NO molecule in *Arabidopsis thaliana*[38]. Importantly, both of these proteins have been used as ArM scaffolds by us[39] and by others[30,38] via covalent linkage strategy and therefore serve as ideal substrates for initial studies on cofactor bioconjugation.

Initial bioconjugation was being studied by reacting tHisF-C50[§] with a known thiol inhibitor *N*-[4-(2-Benzimidazolyl)phenyl]maleimide **39** (BIPM). When 1 equiv. **39** was applied, over 80% conversion was detected on ESI-MS in which mass observed is close to calculated mass (Scheme 2.2).

Scheme 2.2. Bioconjugation between tHisF-C50 and BIPM

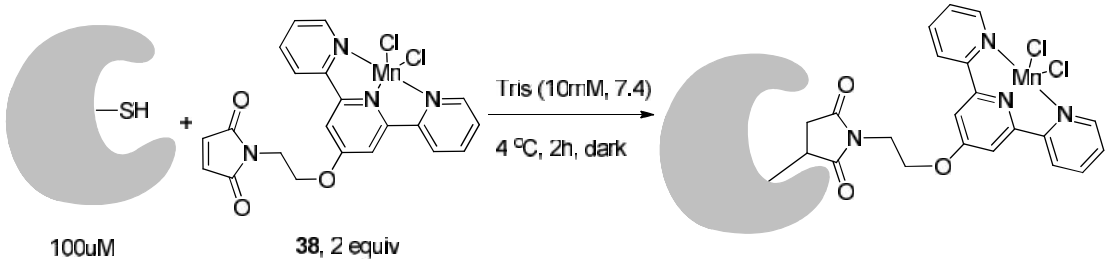


A single cysteine residue was introduced into both tHisF and Nb scaffold. Initial test using tHisF-C50 mutant and 5 equiv. cofactor under 12h incubation showed 100% consumption of protein (Table 2.1, entry 1). However, it appeared to give multiple addition products, a common problem with maleimide bioconjugation reaction. All samples were treated under 4 °C in dark to avoid aggregation. Decreased cofactor loading (entry 4) or reduced reaction time (entry 5) could prevent formation of multiple adducts. Reduced reaction time to 2h in presence of 2 equiv. cofactor (Entry 6) led to about 90% conversion

[§] Proteins were prepared by Dr. Poonam Srivastava

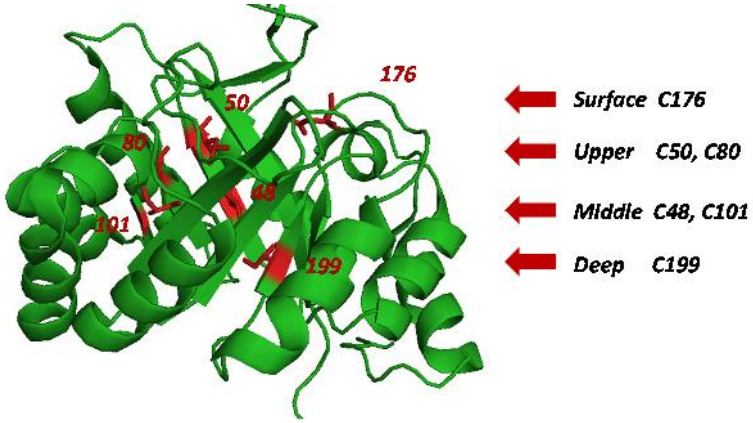
without any multiple adduction. Bioconjugation rates decreased when cysteine mutations deeper in the scaffold proteins were used (entries 7-11). The mutations are mapped in Figure 2.2.

Table 2.1. Bioconjugation study of tHisF variants



entry	protein*	dark	T /°C	time	cof. equiv	conv.	comment
1	tHisF-C50	Y	4	12h	5	100	90% multiple adducts
2	tHisF-C50	N	4	12h	5	80	protein crashed out
3	tHisF-C50	Y	25	12h	5	N/A	protein crashed out
4	tHisF-C50	Y	4	12h	2	100	20% multiple adduct
5	tHisF-C50	Y	4	12h	0.8	60	5% multiple adduct
6	tHisF-C50	Y	4	2h	2	90	
7	tHisF-C80	Y	4	2h	2	50	
8	tHisF-C101	Y	4	2h	2	10	
9	tHisF-C199	Y	4	2h	2	0	
10	tHisF-C176	Y	4	2h	2	100	5% multiple adduct
11	tHisF-C48	Y	4	2h	2	40	
12	tHisF-C48A50	Y	4	2h	2	80	
13	Nb-C96	Y	4	2h	2	90	
14	Nb-C125	Y	4	2h	2	95	

Figure 2.2. Mutation sites of tHisF variants



* Mutagenesis was conducted by Dr. Poonam Srivastava

Reacting tHisF-C48 with **38** provided modest conversion (ca. 40%) to tHisF-C4-**38** within 2 h (entry 11), but we found that mutating a leucine residue (L50)* located directly above C48 in the β-barrel of tHisF[34] led to a marked increase in bioconjugation rate and selectivity, and over 80% conversion to tHisF-C48A50-**38** was achieved within 2 h (entry 12). These improvements presumably resulted from decreased steric hindrance proximal to the cysteine nucleophile.

Nitrobindin is a heme protein with a β-barrel structure and a large hydrophobic pocket for heme binding. Cysteine mutations at suitable locations within the β-barrel can be used to position cofactors within the heme binding pocket of the apo enzyme (Nb)[38]. Bioconjugation of two cysteine mutants of Nb, C96† within the β-barrel and C125 on the protein surface, were examined, and both of these provided greater than 90% conversion to the corresponding Nb-C-**38** bioconjugates (Table 2.1, entries 13 and 14). These results, in addition to those noted above using tHisF, highlight two key advantages of covalent bioconjugation for ArM formation[9]: the nature of both the scaffold itself and the local environment of the bioconjugation site can be readily manipulated to facilitate cofactor installation[6].

Bioconjugated tHisF- and Nb-**38** samples were purified by 4 rounds of centrifugal filtration into buffer containing 10% ACN. The isolated ArM/scaffold mixture was then characterized by ESI mass spectrometry to confirm the composition of the ArM (Figure. 2.3A). It should be noted that $MnCl_2$ deletion was detected, potentially due to metal-ligand bond dissociation in mass spectrometry. UV-vis spectroscopy revealed that characteristic peaks at 310nm and 320nm were detected in both cofactor 8 and bioconjugated ArMs(Figure.

† Proteins were prepared by Dr. Poonam Srivastava

2.3B). This observation confirmed the coordination environment of Mn terpyridine. CD spectroscopy[*] indicated that little structural perturbation resulted from cofactor incorporation (Figure. 2.3C). We built up a pymol model from a crystal structure of Nb (PDB ID 3EMM[35]) with overlaid DFT-optimized structure of **38** (red) onto the native heme (gray) (Figure. 2.3D).

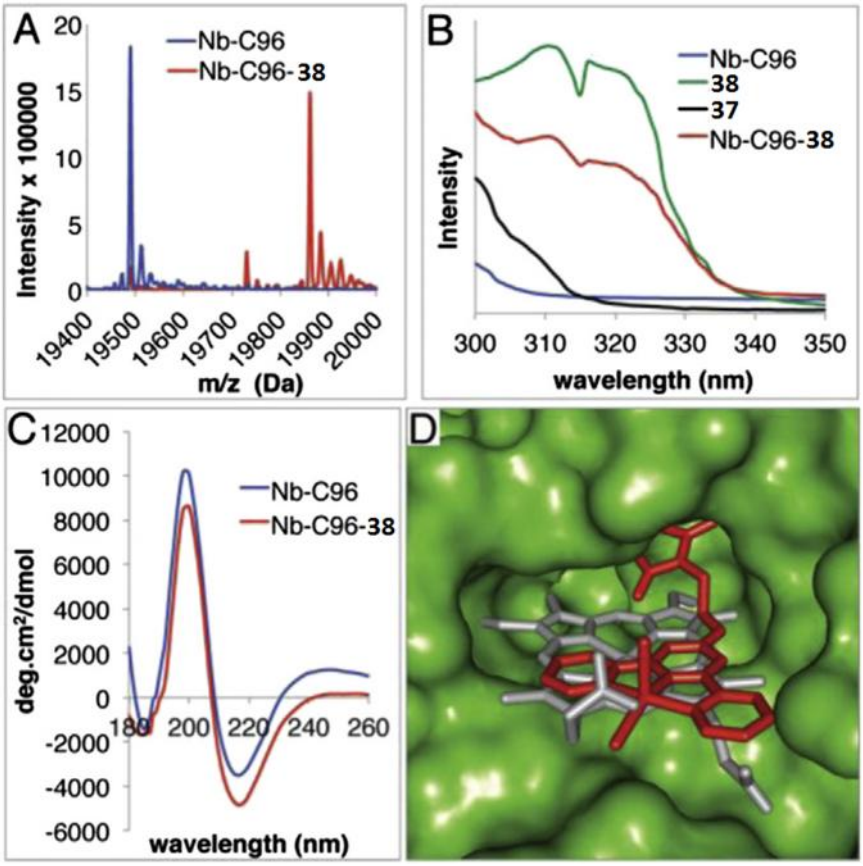
Table 2.2. Mass spectrometry data for ArMs

Entry	Scaffold[†]	calculated MW[a]	observed MW
1	tHisF-C50	28754	28754
2	tHisF-C50-**38**	29126[b]	29128
3	tHisF-C48A50	28712	28710
4	tHisF-C48A50-**38**	29084[b]	29082
5	Nb-C96	19490	19489
6	Nb-C96-**38**	19862[b]	19862
7	Nb-C125	19547	19546
8	Nb-C125-**7**	19919[b]	19919

[a]Protein MW calculated at http://www.scripps.edu/wcdputnam/ protcalc.html.
[b]The $MnCl_2$ fragment not detected.

[*] Conducted by Ken Ellis-Guardiola
[†] Proteins were prepared by Dr. Poonam Srivastava

Figure 2.3. (A) HR ESI-MS spectra of Nb-C96 and Nb-C96-38. (B) UV spectra of Nb-C96, Nb-C96-38, 37, and 38 showing the presence of $MnCl_2$ in Nb-C96-38. (C)* CD spectra of Nb-C96 and Nb-C96-38 showing proper Nb folding following bioconjugation. (D) Pymol model of Nb-C96 and Nb-C96-38. (taken from ref[1])



The reactivity of this ArM under a variety of conditions was explored to examine the impact of confining the Mn center within the heme binding pocket. Benzylic oxygenation of ibuprofen methyl ester (**40a**) was initially investigated based on previous reports from Crabtree and Brudvig[21]. As previously noted, these researchers reported that ibuprofen (**41a**) oxygenation can be catalyzed by Kemp's-triacid-based Mn-terpyridine complexes to provide a mixture of **41b** and **41c**, the latter resulting from decarboxylation following oxygenation (Table 2.4). The selectivity of this reaction could be controlled using hydrogen bonding between carboxylic acid groups on ibuprofen and the catalyst, with the presence of the

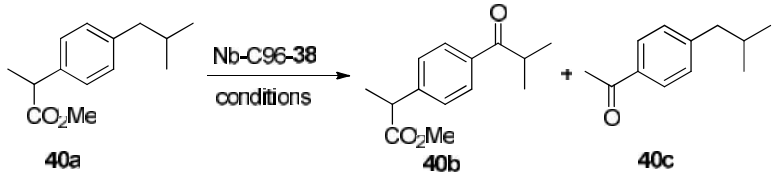
* Conducted by Ken Ellis-Guardiola

carboxylic acid on the catalyst improving the **41b**:**41c** ratio from 77:23 to 98.5:1.5.

We found that similar benzylic oxygenation and oxygenation/decarboxylation of **40a** was catalyzed by Nb-C96-**38** (Table 2.3). Oxone, peracids, and a number of additional oxidants are known to promote Mn-terpyridine catalyzed oxygenation, and both oxone and peroxyacetic acid proved to be compatible with Nb-C96-**38**. Crucial to this reactivity, however, was the use of phosphate buffer (Table 2.3, entries 1-3). A modest excess of peroxyacetic acid was sufficient to achieve maximum conversion, and this helped to minimize pH changes over the course of the reaction (entries 3-5). While the Nb scaffold can tolerate 20% v/v acetonitrile, greater than this amount led to protein precipitation and dramatically reduced conversion. Maximum conversion was observed using 10% v/v acetonitrile, and omitting acetonitrile entirely led to reduced conversion, presumably due to poor substrate solubility (entries 6-8). Conducting the reactions at room temperature or decreasing ArM loading to 2.5 mol % both led to significant decreases in conversion (entries 9-10). All reactions in Table 2.3 provided a similar **40b**:**40c** ratio of 88:12. Nb-C125-**38** provided this same ratio with slightly higher conversion, suggesting that improved substrate access to **38** attached to the protein surface (in Nb-C125-**38**) offset any potential scaffold acceleration[5] afforded by the heme binding pocket (in Nb-C96-**38**).

The optimized conditions identified in Table 2.3, entry 7 was then used to explore the substrate scope of Nb-C96-**38** catalyzed oxygenation[22]. To prove the catalytic activity, $MnCl_2$, Nb scaffold and controls lacking catalyst were conducted on **40a**. Results showed that no activity was detected without Mn or terpyridine ligand (Table 2.4).

Table 2.3. Optimization of ArM catalyzed oxidation[a]



40a 40b 40c

Entry	Solvent	Oxidant	Conv[b]. (%)
1	Tris (25mM, 7.0), 20%ACN	Oxone, 10eq	0
2	NaPi (100mM, 7.0), 20%ACN	Oxone, 10eq	77
3	NaPi (100mM, 7.0), 20%ACN	AcOOH, 10eq	85
4	NaPi (100mM, 7.0), 20%ACN	AcOOH, 2.5eq	83
5	NaPi (100mM, 7.0), 20%ACN	AcOOH, 2eq	76
6	NaPi (100mM, 7.0), 30%ACN	AcOOH, 2.5eq	12[e]
7	NaPi (100mM, 7.0), 10% ACN	AcOOH, 2.5eq	91
8	NaPi (100mM, 7.0)	AcOOH, 2.5eq	24
9[c]	NaPi (100mM, 7.0), 10% ACN	AcOOH, 2.5eq	45[e]
10[d]	NaPi (100mM, 7.0), 10% ACN	AcOOH, 2.5eq	58

[a]Reaction conditions: 1 mM substrate, 5% ArM, 4 °C, overnight in dark.
[b]Determined by GC analysis of DCM extract.
[c]Reaction conducted at room temperature
[d]2.5% ArM loading used.
[e]Protein precipitation was observed.

Table 2.4. Nb hybrid catalyzed 40a epoxidation

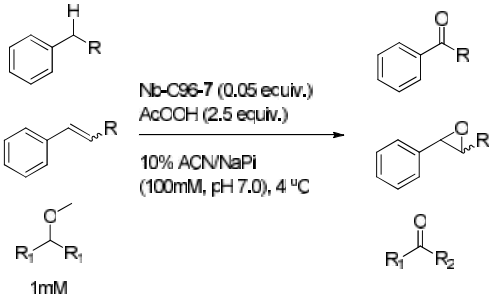
Entry	Catalyst	Conversion (%)[a]
1	Nb-C96-**38**	91
2	Nb-C125-**38**	92
3	**38**	95
4	Nb-C96	<1
5	$MnCl_2$	<1
6	-	<1

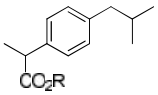
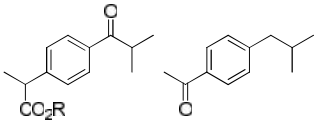
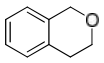
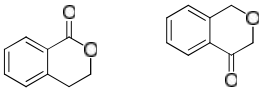
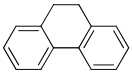
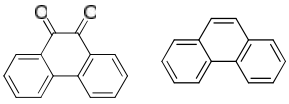
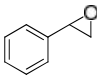
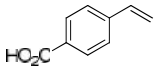
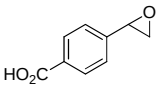
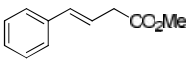
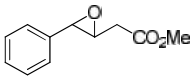
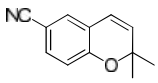
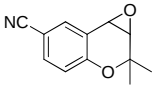
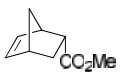
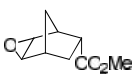
[a]Conversion was detected by RP-HPLC.

Both ibuprofen methyl ester and ibuprofen provided high conversions (Table 2.5, entry 1), and benzylic oxygenation **40**/**41b** and decarboxylation **40**/**41c** products were formed in similar ratios. Isochroman **42a** underwent benzylic oxygenation with a preference for the more activated 1-position (entry 2), while 9,10-dihydrophenan-threne **43a** was oxidized to a mixture of diketone **43b** and phenanthrene **43c** (entry 3). Epoxidation of different olefins was also examined. Styrene, para-substituted styrene, and the β-substituted styrylacetic acid methyl ester were both epoxidized in moderate to good yield (entries 4-6). Cyclic cis-internal olefins were particularly reactive substrates, and nearly quantitative epoxidation of both chromene **47a** and norbornene-2-carboxylate methyl ester **48a** was observed (entries 7 and 8). Finally, ethereal α-C-H oxidation of cyclohexyl methylether **49a**, which proceeds via C-H hydroxylation followed by hemiacetal decomposition, was also observed in moderate yield (entry 9).

The substrate scope outlined in Table 2.5 shows that the broad reactivity of Mn-terpyridine complexes toward hydrocarbon oxidation[22] is not compromised in the context of a protein scaffold. In general, a slight decrease in reaction rate was observed using Nb- C96-**38** relative to **38** (Table 2.4), but otherwise similar reactivity was observed. In entries 1-3, product mixtures were obtained, and while no difference between the selectivities of **38** and Nb-C96-**38** was observed, these substrates will serve as valuable probes to explore the impact of scaffold mutations on oxygenation selectivity in future studies[40,41].

Table 2.5. Substrate scope for ArM catalyzed oxidation



Entry	Substrate	Products (% conv.)
1[a]	 **40a** (R=Me) **41a** (R=H)	 **40b** (89) **41b** (87) **40c** (7) **41c** (4)
2[b]	 **42a**	 **42b** (82) **42c** (15)
3[b]	 **43a**	 **43b** (17) **43c** (11)
4[a]	 **44a**	 **44b** (65)
5[c]	 **45a**	 **45b** (20)
6[c]	 **46a**	 **46b** (80)
7[a]	 **47a**	 **47b** (93)
8[a]	 **48a**	 **48b** (99)
9[a]	 **49a**	 **49b** (59)

[a] Conversion determined using GC integration for starting materials and products

[b] Conversion determined using ^{1}H NMR signals for starting materials and products

[c] Conversion determined using HPLC integration for starting materials and products

2.3. CONCLUSIONS

Maleimide-substituted manganese terpyridine cofactor **38** was prepared and reacted with cysteine mutants of tHisF and Nb scaffold proteins to generate the desired ArMs in good to excellent yield based on HR ESI-MS. Spectroscopic characterization (CD and UV) of Nb-C96-**38** provided additional details on the structure of this bioconjugate, confirming both the nature of the scaffold fold and the presence of the $MnCl_2$ fragment. Nb-C96-**38** catalyzed hydrocarbon oxygenation in the presence of either oxone or AcOOH, and optimized reaction conditions enabled benzylic oxygenation, olefin epoxidation, and ethereal α-C-H oxidation[22]. The substrates examined, in conjunction with the robust ArM reactivity observed, will enable exploration of the impact of scaffold structure on the regio-, enantio-, and chemo-selectivity of **38**. Toward this end, we are currently examining a range of cofactor linkage sites within the Nb scaffold as well as mutagenesis of residues proximal to the Mn center in Nb-C96-**38**. The generality of the covalent linkage method employed[9] will also enable facile exploration of additional scaffold proteins to identify systems with unique hydrocarbon oxygenation selectivity.

2.4. EXPERIMENTAL

Materials

E. coli DH5α and BL21 (DE3) cells were purchased from Invitrogen (Carlsbad, CA). Nde I, Xho I restriction enzyme, T4 DNA ligase, Taq DNA polymerase and Phusion HF polymerase (Cat# 530S) were purchased from New England Biolabs (Ipswitch, MA). Luria broth (LB) and rich medium (2YT) and Agar media were purchased from Research Products International (Mt. Prospect, IL). Qiagen DNA extraction kit (Cat# 28706) and plasmid isolation kit (Cat# 27106) were purchased from QIAGEN Inc. (Valencia, CA) and used according to the manufacturer's instructions. DNA purification kit (Zymo, Cat# D4004) was purchased from Zymo research (Irvine, CA) and used as recommended. All genes were confirmed by sequencing at the University of Chicago Comprehensive Cancer Center DNA Sequencing & Genotyping Facility (900 E. 57th Street, Room 1230H, Chicago, IL 60637). Electroporation was carried out on a Bio-Rad MicroPulser using method Ec2. Ni-nitrilotriacetic acid (Ni-NTA) resin and Pierce® BCA Protein Assay Kits (Cat# 23225) were purchased from Fisher Scientific International, Inc. (Hampton, NH), and the manufacturer's instructions were followed when using both products (for Ni-NTA resin, 8 mL resin were used, with buffers delivered by a peristaltic pump at a rate of 1 mL/min, in a 4 °C cold cabinet). Amicon® 10 kD spin filters for centrifugal concentration were purchased from EMD Millipore (Billerica, MA) and used at 4,000 g at 4 °C. PD-10 desalting columns (Cat# 17-0851-01) and Hitrap desalting columns (Cat# 11-0003-29) were purchased from GE Healthcare (Pittsburg, PA).

General procedures

Unless otherwise specified, all reactions were prepared in flame or oven-dried glassware under N_2 atmosphere using syringe transfer. TLC plates were visualized under 254 nm light. Flash column chromatography was carried out using Silicycle 230-400 mesh silica gel. ^{1}H and ^{13}C NMR spectra were recorded at 500 MHz and 125 MHz, respectively, on a Bruker DMX-500 or DRX-500 spectrometer, and chemical shifts are reported relative to residual solvent peaks. Chemical shifts are reported in ppm and coupling constants are reported in Hz. Conversions were calculated by integrating appropriate signals from GC/HPLC/NMR spectra, and they were reported as the average of two trials set up in parallel. High resolution ESI mass spectra were obtained from the University of Chicago mass spectrometry facility using an Agilent Technologies 6224 TOF LC/MS. Amicon® 15 mL 10 kD cutoff centrifugal filters were used to concentrate protein solutions. CD spectra were obtained using an AVIV-202 CD spectrophotometer (AVIV Biomedical, Inc.) between 260 and 180 nm at room temperature.

Cofactor (**38**) synthesis

N-[2-([2,2':6',2"-Terpyridin]-4'-yloxy) ethyl maleimide (**37**). To a solution of **36**[32,33] (292 mg, 1 mmol) in CH_2Cl_2 (10 mL) was added maleic anhydride (490 mg, 5 mmol). The mixture was heated at reflux for 3.5 h and subsequently cooled to room temperature. A white precipitate formed, was collected by filtration, and was washed with cold CH_2Cl_2. The solid was then added to a suspension of NaOAc (1 g) in Ac2O (10 mL), and the resulting mixture was heated at reflux for 1 h. The mixture was cooled to room temperature, poured into ice-

water (40 mL), and stirred for 2 h. The mixture was extracted with CH_2Cl_2 (3×100 mL), and the combined organic extracts were washed sequentially with saturated NaHCO3 (3×50 mL), H2O (1×50 mL), and brine (1×50 mL), and dried over $MgSO_4$. Removing the solvent using rotary evaporation provided the desired product as a white solid; yield: 219 mg (0.589 mmol, 59%). 1H NMR (500 MHz, $CDCl_3$): δ 8.72 (d, *J*=4.6 Hz, 2H), 8.64 (d, *J*=7.9 Hz, 2H), 8.02 (s, 2H), 7.88 (td, *J*=7.7, 1.7 Hz, 2H), 7.37 (dd, *J*=6.9, 5.3 Hz, 2H), 6.79 (s, 2H), 4.45 (t, *J*=5.6 Hz, 2H), 4.08 (t, *J*=5.6 Hz, 2H). 13C NMR (500 MHz, $CDCl_3$): d 170.5, 166.6, 157.3, 156.0, 149.2, 137.0, 134.4, 124.0, 121.5, 107.4, 64.9, 37.1. HRMS-ESI (m/z): for $C_{21}H_{16}N_4O_3$, calculated: 373.1295, found: 373.1290.

*Manganese(II) terpyridine complex (**38**)* Dissolved 212 mg (1.07 mmol) manganese chloride tetrahydrate in 10 mL dry THF by sonication. This solution was added to a stirring solution of 50 mg (0.134 mmol) ligand **2** dissolved in 1.07 mL dry THF. A fine yellow precipitate formed immediately upon addition. Collected yellow solid on a fine frit and washed with copious THF to obtain 38 mg of product (0.076 mmol, 57% yield). IR n_{max} cm^{-1} (KBr pellet): 3486 (br), 3074 (w), 2062 (w), 1712 (s), 1598 (s), 1474 (m), 1435 (m), 1404 (m), 1384 (s), 1355 (m), 1255 (w), 1220 (s), 1161 (m), 1098 (m), 1063 (m), 1041 (w), 1014 (s), 982 (w), 947 (w), 837 (s), 799 (s), 729 (w), 694 (s), 668 (w), 660 (w), 639 (w), 621 (m). UV (1:1 ACN/water) λ_{max}, nm: 310, 316. ESI-MS (m/z): 462.0298.

Cloning, Expression and Protein Purification

A codon optimized gene for Apo Nitrobindin[35] was obtained from GenScript USA Inc.

(Piscataway, NJ) and cloned into pET22B using NdeI and XhoI restriction sites to install a C-terminal hexa-histidine tag for Ni-NTA affinity chromatography. A gene encoding the cyclase subunit of the imidazole glycerol phosphate synthase enzyme complex from *Thermotoga maritima* (tHisF) was amplified from pET11c-tHisF20[30] by PCR using gene specific primers containing NdeI (forward) and XhoI (reverse) restriction sites. Cysteine mutations were introduced into the Nitrobindin gene at positions Q96 and A125 and for tHisF at position V48 by site directed overlap extension[42] PCR. To introduce mutations, two separate polymerase chain reactions were performed, each using a perfectly complementary flanking primer at the 5' and 3' end of the sequence and a mutagenic primer. The PCR conditions were as follows: Phusion HF buffer 1x, 0.2 mM dNTPs each, 0.5 μM forward primer, 0.5 μM reverse primer, 0.02 U/μL Phusion polymerase and 1 ng/mL template plasmid. The resulting two overlapping fragments that contained the base pair substitution were then assembled in a second PCR using the flanking primers resulting in the full-length mutated gene. The double mutant of tHisF was generated using C48 as template and primers to install L50A using overlap extension PCR. Nucleotide sequences for the all the primers are summarized in Table 2.6.

Table 2.6. Nucleotide sequences for the primers

#	Primer name	Primer sequence
1	T7 for	5'-GCG AAA TTA ATA CGA CTC ACT ATA-3'
2	T7 rev	5'-TTA TGC TAG TTA TTG CTC AGC GG-3'
5	Q96C for	5'- GTT ATC GCA TGC TCG ACC GGT-3'
6	Q96C rev	5'-ACC GGT CGA GCA TGC GAT AAC-3'
7	A125C for	5'-G GTG GGC AAC TGC TCC AAA GTT-3'
8	A125C rev	5'-AAC TTT GGA GCA GTT GCC CAC C-3
13	V48C for	5'-GAC GAA CTC TGC TTT CTG GAT ATC-3'
14	V48C rev	5'-GAT ATC CAG AAA GCA GAG TTC GTC-3'
15	C48A50 for	5'-GAC GAA CTC TGC TTT GCG GAT ATC ACC-3'
16	C48A50 rev	5'-GGT GAT ATC CGC AAA GCA GAG TTC GTC-3'

PCR amplified fragments and plasmid vector pET22b were digested with Nde I and Xho I enzymes in recommended buffers at 37 °C for 2 hours. Digested DNA was purified by agarose gel extraction before ligation. Ligation reactions were set-up with a molar ratio of 1:3 (plasmid: insert) in 10 µL reaction mix. Typically a ligase reaction mix had 1 ng/L digested plasmid vector, 9 ng/mL of the insert, 1 µL 10X ligase buffer and 1 U/mL ligase. Reactions were incubated at 16 °C overnight, purified using DNA purification kits, and transformed into *E. coli* DH5 cells. Cells were recovered in SOC medium for 1 hour at 37 °C and then spread on LB ampicillin plates (6.25 g LB powder mix, 4 g agar, 250 mL DDI water, 0.1 mg/mL ampicillin). Plates were incubated at 37 °C overnight; individual colonies that appeared next day were tested for gene fragments by colony PCR. Clones that showed amplification of desired fragments were used to inoculate LB medium containing 0.1 mg/mL ampicillin and grown overnight at 37 °C, 250 rpm. Recombinant plasmid from these overnight grown cultures were isolated via miniprep (Qiagen, Valencia, CA) and sequenced. Plasmid sequencing was done by the U Chicago sequencing facility staff and T7 for and T7 rev primers were used for sequencing reactions.

Expression and purification protocol

pET22b-NB or pET22b-tHisF (mutants thereof) was transformed into electrocompetent *E. coli* BL21 (DE3) medium (37 °C, 50 min), then plated onto LB amp+ agar plates (6.25 g LB powder mix, 4 g agar, 250 mL DDI water, 0.1 mg/mL ampicillin), and incubated at 37 °C for 16 h. Several colonies appeared on overnight-incubated plates; a single colony from this plate was inoculated in 5 mL 2YT medium having antibiotics with the same

concentrations as above. The culture was incubated overnight at 37 °C with constant shaking at 250 rpm. On the following day, 5 mL of the overnight cultures was used to inoculate 500 mL of fresh 2YT media having the same antibiotic in 1 L Erlenmeyer flask. The culture was incubated at 37 °C, 250 rpm and protein expression was induced by adding 1mM IPTG when OD_{600} reached 1. The induced culture was allowed to grow for 12 hours, and then the cells were harvested by centrifugation at 4 °C, 3000 x g for 20 minutes. Cell pellets were re-suspended in 30 mL PBS (pH 7.5) and sonicated (40 amplitude, 30 second burst, 10 minute total process). Lysed culture was then clarified at 16000 x g, 4 °C for 30 minutes and supernatant thus obtained was purified by Ni-NTA resin using manufacturer's instructions. Purified protein was buffer exchanged to 10 mM Tris (pH 7.5) and measured by Pierce® BCA Protein Assay Kit as recommended.

ESI-MS of tHisF-C48 and tHisF-C48A50 and Nb-C125 scaffold and hybrid

Protein samples were buffer exchanged into 1:1 water/ACN mixture. Then 1% vol. acetic acid(tHisF) or formic acid(Nb) was added. 10uL sample was injected.

Bioconjugation and purification

50 µL of cofactor **38** solution (1.99 mg/mL in 1:1 water/acetonitrile, 2.0 equiv) were added to a solution of protein scaffold (950 µL, 100 µM protein in Tris buffer, 25 mM pH 7.5 for Nb and 50mM pH 7.2 for tHisF) in microcentrifuge tube and shaken in the dark at 4°C for 1h. The final concentrations were: 100 µM protein, 200 µM cofactor **38**, 2.5 vol% acetonitrile. The hybrid was purified by the 4 times centrifugal filtration using 10 kDa cutoff

spin filters (Millipore) with NaPi buffer (100 mM, pH 7.0 with 10% ACN). The removal of excess cofactor was confirmed by HPLC analysis, and the conversion was estimated by ESI analysis.

Genreal procedure for biocatalysis

Hybrid solution (500 μL, 100 μM), 500 μL NaPi buffer (100 mM, pH 7.0 with 10% ACN), and 12.5 μL substrate (80 mM stock in ACN) were added to 1.5 mL microcentrifuge tube, and incubated at 4°C for 15 mins. 5 μL of peracetic acid aqueous solution (1M) was added to the mixture. The resulting solution kept shaking at 4℃ overnight. The final concentrations was: 1 mM substrate, 2.5 mM AcOOH, and 50 hybrid. The reaction was quenched by adding 350 μL DCM. pH was adjusted to 2 for ibuprofen case. The closed vials was vortexed for 30s and spin down at 15000 rpm for 3 mins. The organic layer was isolated. 10 μL sample was injected on HPLC.

Figure 2.4. Top left: tHisF-C48; Top right: tHisF-C48-3; Bottom: deconvolution

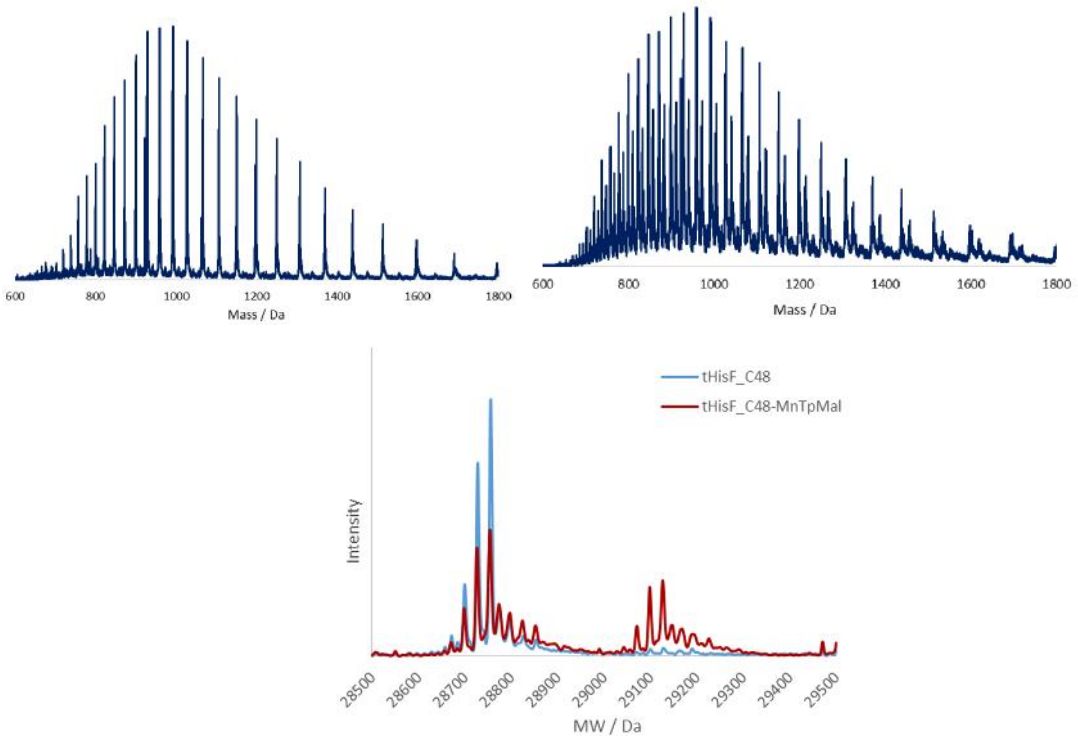


Figure 2.5. Top left: tHisF-C48A50; Top right: tHisF-C48A50-3; Bottom: deconvolution

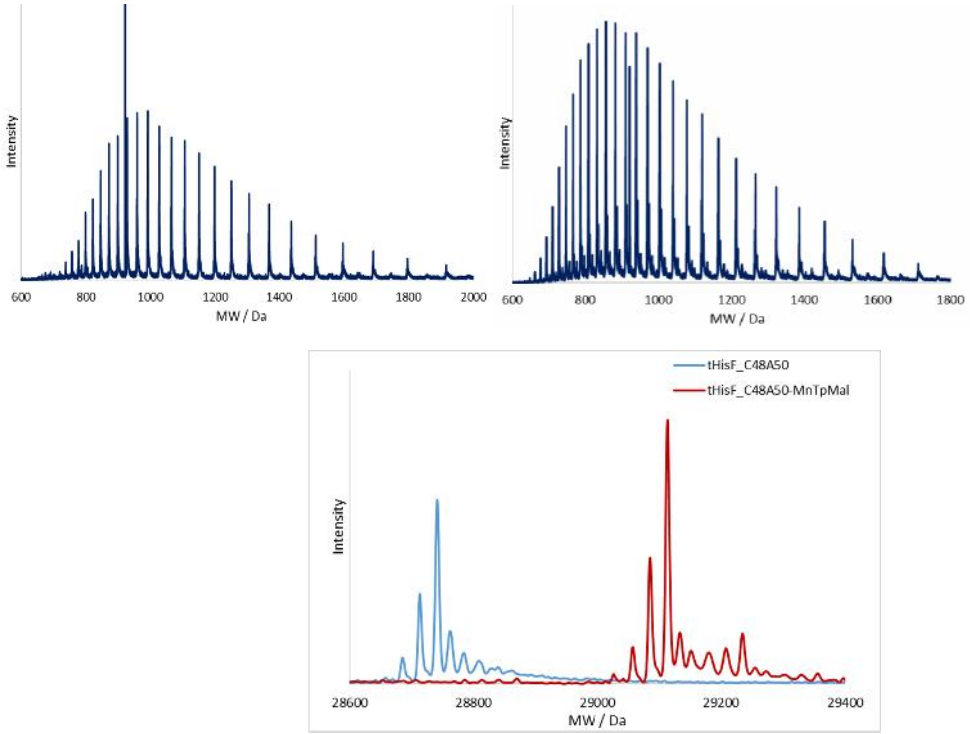
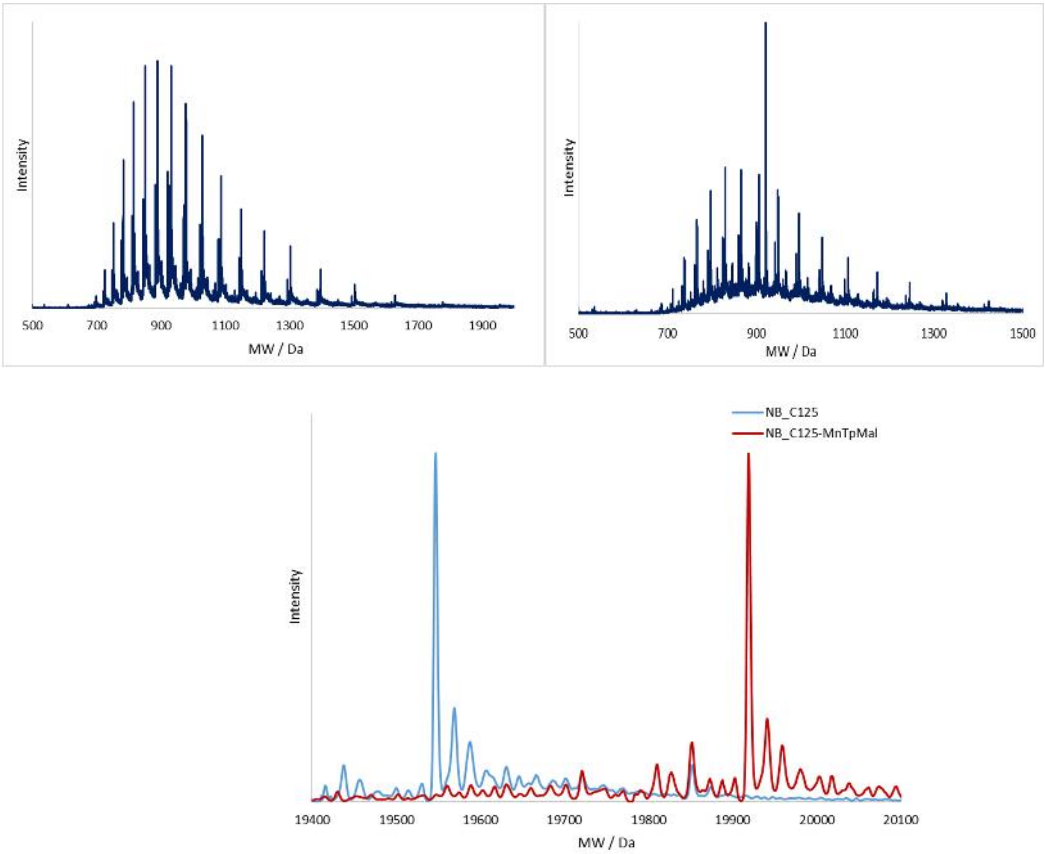


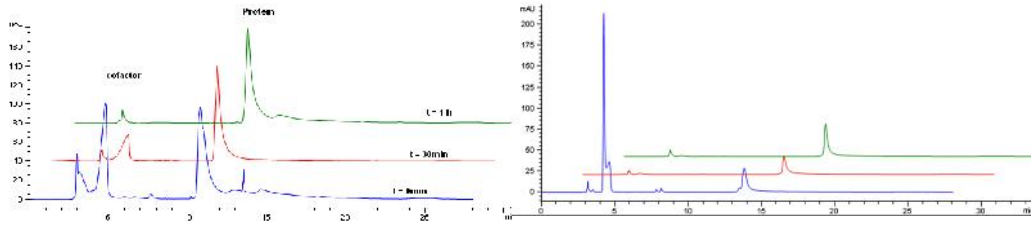
Figure 2.6. Top left: Nb-C125; Top right: Nb-C125-3; Bottom: deconvolution



HPLC trace of bioconjugation

HPLC analysis was used to monitor bioconjugation when limiting amount of cofactor (0.8 equiv) was used. Bioconjugation samples at t = 0 min, 30min and 1h for Nb-C96 were analyzed on HPLC. The trace demonstrated the consumption of cofactor.

Figure 2.7. HPLC trace of Nb-C96 (left) and tHisF-C48A50 (right) bioconjugation when limiting amount of cofactor (0.8 equiv) was used.



Circular Dichroism (CD) and Fluorescence Analysis

CD spectra were obtained using an AVIV-202 CD spectrophotometer (AVIV Biomedical, Inc.) between 260 and 180 nm at room temperature. CD samples contained 25 uM protein (as determined by A280) in 10 mM sodium phosphate buffer at pH 8.0.

Substrates and standard product synthesis

All substrates and standard products were prepared according to reported procedure: **40a**, **45a**, **47a**[43]; **40b**, **40c**, **41b**, **41c**[18]; **42b**[44]; **42c**[45]; **43b**, **43c**[23]; **44b**[46]; **45b**[47]; **46b**[48]; **47b**[49]; **48b**[50].

Biocatalysis data

Table 2.7. Biocatalysis on substrate 40a-45a and 47a

	Conversion (%)		
Substrate	Nb-C96-**38**	Nb-C125-**38**	**38**
40a	96	95	99
41a	91	92	98
42a	97	97	99
43a	28	33	54
44a	65	69	75
45a	80	64*	81
47a	99	99	99

*Protein precipitation was observed.

2.5 ACKNOWLEDGEMENTS

This work was supported by the University of Chicago, Department of Chemistry. I wish to thank Dr. Poonam Srivastava for assistance with the molecular biology works of protein engineering and expression. I am grateful to Ken Ellis-Guardiola for his assistance with mass and CD characterizations.

2.6. References

(1) Zhang, C.; Srivastava, P.; Ellis-Guardiola, K.; Lewis, J. C. *Tetrahedron* **2014**, *70*, 4245.
(2) Davies, C. L.; Dux, E. L.; Duhme-Klair, A. K. *Dalton Trans* **2009**, 10141.
(3) Reetz, M. T. In *Top Organomet Chem* 2009; Vol. 25, p 63.
(4) Letondor, C.; Ward, T. R. *Chembiochem* **2006**, *7*, 1845.
(5) Collot, J.; Humbert, N.; Skander, M.; Klein, G.; Ward, T. R. *Journal of Organometallic Chemistry* **2004**, *689*, 4868.
(6) Lewis, J. C. *ACS Catalysis* **2013**, *3*, 2954.
(7) Lu, Y.; Berry, S. M.; Pfister, T. D. *Chem Rev* **2001**, *101*, 3047.
(8) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. In *Topics in Organometallic Chemistry*; Ward, T. R., Ed. Berlin, Heidelberg, 2009; Vol. 25, p 45.
(9) Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. *Chem Rev* **2001**, *101*, 3081.
(10) Davies, H. M.; Morton, D. *Chem Soc Rev* **2011**, *40*, 1857.
(11) Doyle, M. P.; Duffy, R.; Ratnikov, M.; Zhou, L. *Chem Rev* **2010**, *110*, 704.
(12) Chen, M. S.; White, M. C. *Science* **2006**, *318*, 783.
(13) Kuhl, N.; Hopkinson, M. N.; Wencel-Delord, J.; Glorius, F. *Angew Chem Int Ed Engl* **2012**, *51*, 10236.
(14) Costas, M.; Chen, K.; Jr., L. Q. *Coordination Chemistry Reviews* **2000**, *200-202*, 517.
(15) Talsi, E. P.; Bryliakov, K. P. *Coordination Chemistry Reviews* **2012**, *256*, 1418.
(16) Tanase, S.; Bouwman, E. In *ADVANCES IN INORGANIC CHEMISTRY* 2006; Vol. 58, p 29.
(17) Davies, H. M.; Manning, J. R. *Nature* **2008**, *451*, 417.
(18) Das, S.; Brudvig, G. W.; Crabtree, R. H. *J Am Chem Soc* **2008**, *130*, 1628.
(19) Tagore, R.; Crabtree, R. H.; Brudvig, G. W. *Inorg Chem* **2008**, *47*, 1815.
(20) Balcells, D.; Moles, P.; Blakemore, J. D.; Raynaud, C.; Brudvig, G. W.; Crabtree, R. H.; Eisenstein, O. *Dalton Trans* **2009**, 5989.
(21) Das, S.; Incarvito, C. D.; Crabtree, R. H.; Brudvig, G. W. *Science* **2006**, *312*, 1941.
(22) Inoue, M.; Kamijo, S.; Amaoka, Y. *Synthesis* **2010**, *2010*, 2475.
(23) Hull, J. F.; Balcells, D.; Sauer, E. L. O.; Raynaud, C.; Brudvig, G. W.; Crabtree, R. H.; Eisenstein, O. *J Am Chem Soc* **2010**, *132*, 7605.
(24) Steinreiber, J.; Ward, T. R. In *Top Organomet Chem* 2009; Vol. 25, p 93.
(25) Oohora, K.; Kihira, Y.; Mizohata, E.; Inoue, T.; Hayashi, T. *J Am Chem Soc* **2013**, *135*, 17282.
(26) Lewis, J. C.; Arnold, F. H. *Chimia* **2009**, *63*.
(27) Fasan, R. *ACS Catalysis* **2012**, *2*, 647.
(28) Smyth, D. G. *Biochem. J.* **1964**, 589.
(29) Hermanson, G. T. *Bioconjugate Technique*; Academic: San Diego, CA 92101, USA, 1996.
(30) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Taglieber, A.; Hollmann, F.; Mondiere, R. J.; Dickmann, N.; Hocker, B.; Cerrone, S.; Haeger, M. C.; Sterner, R. *Chembiochem* **2008**, *9*, 552.
(31) Chen, H.; Tagore, R.; Das, S.; Incarvito, C.; Faller, J. W.; Crabtree, R. H.; Brudvig, G. W. *Inorg Chem* **2005**, *44*, 7661.

(32) Constable, E. C.; Ward, M. D. *Journal of the Chemical Society, Dalton Transactions* **1990**, 1405.
(33) Maayan, G.; Yoo, B.; Kirshenbaum, K. *Tetrahedron Lett* **2008**, *49*, 335.
(34) Beismann-Driemeyer, S.; Sterner, R. *J Biol Chem* **2001**, *276*, 20387.
(35) Bianchetti, C. M.; Blouin, G. C.; Bitto, E.; Olson, J. S.; Phillips, G. N., Jr. *Proteins* **2010**, *78*, 917.
(36) Douangamath, A.; Walker, M.; Beismann-Driemeyer, S.; Vega-Fernandez, M. C.; R. Sterner, M. W. *Structure* **2002**, *10*, 185.
(37) Lang, D.; Thoma, R.; Henn-Sax, M.; Sterner, R.; Wilmanns, M. *Science* **2000**, *289*, 1546.
(38) Onoda, A.; Fukumoto, K.; Arlt, M.; Bocola, M.; Schwaneberg, U.; Hayashi, T. *Chem Commun (Camb)* **2012**, *48*, 9756.
(39) Yang, H.; Srivastava, P.; Zhang, C.; Lewis, J. C. *Chembiochem* **2014**, *15*, 223.
(40) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondière, R.; Taglieber, A. *Tetrahedron* **2007**, *63*, 6404.
(41) Creus, M.; Ward, T. R. *Org Biomol Chem* **2007**, *5*, 1835.
(42) Heckman, K. L.; Pease, L. R. *Nat Protoc* **2007**, *2*, 924.
(43) Hashimoto, N.; Aoyama, T.; Shioiri, T. *Chem. Pharm. Bull.* **1981**, *29*, 1475.
(44) Yin, L.; Wu, J.; Xiao, J.; Cao, S. *Tetrahedron Letters* **2012**, *53*, 4418.
(45) Anzalone, L.; Hirsch, J. A. *J Org Chem* **1985**, *50*, 2128.
(46) Lane, B. S.; Vogt, M.; DeRose, V. J.; Burgess, K. *J Am Chem Soc* **2002**, *124*, 11946.
(47) Izquierdo, J.; Rodriguez, S.; Gonzalez, F. V. *Org Lett* **2011**, *13*, 3856.
(48) Lee, N. H.; Muci, A. R.; Jacobsen, E. N. *Tetrahedron Letters* **1991**, *32*, 5055.
(49) Afarinkia, K.; Mahmood, F. *Tetrahedron Letters* **2000**, *41*, 1287.
(50) Ogawa, H.; Hagiwara, T.; Chihara, T.; Teratani, S.; Taya, K. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 627.

CHAPTER THREE

ENGINEERING AND DIRECTED EVOLUTION OF DIRHODIUM ARTIFICIAL METALLOENZYMES FOR ENANTIOSELECTIVE N-H INSERTION

3.1 INTRODUCTION

As previously discussed in this dissertation, a key goal in catalytic C-H functionalization is to develop methods in which the site-selectivity observed is not dependent on a directing group or a substrate's electronic bias[1,2]. Artificial metalloenzymes (ArMs) formed by incorporating synthetic metal catalysts into protein scaffolds have the potential to impart to chemical reactions selectivity that would be difficult to achieve using metal catalysts alone. In Chapter Two, a Mn-terpyridine artificial metalloenzyme was explored. Although robust oxygenation reactivity was observed, the lack of selectivity motivated us to revise this system. This lack of selectivity could arise from diffusible reactive oxygen species (ROS) generated from the Mn-terpyridine cofactor or from poor catalyst-substrate recognition within the ArM. If the latter is occurring, alternate protein scaffolds could improve this problem. Unfortunately, a range of sidechains on proteins are potentially reactive toward the nucleophilic cofactor anchoring approach used in Chapter Two. This lack of modification specificity constrains the choice of scaffold proteins that can be explored. Furthermore, because scaffold engineering is required to achieve high selectivity, expression level, thermostability, and organic solvent tolerance, a novel method providing orthogonal and site-specific bioconjugation with a broad range of scaffold scope offers a significant advantage.

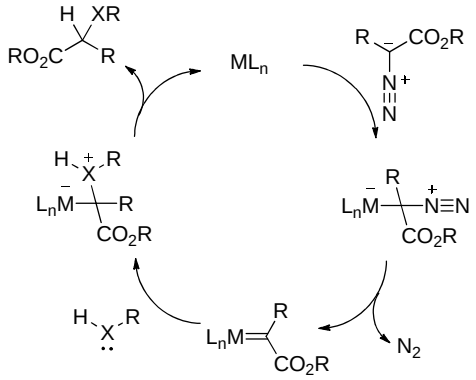
Herein, we demonstrate that strain-promoted azide-alkyne cycloaddition (SPAAC)[3,4]

provides a simple means to introduce metal catalysts into proteins with the broad scaffold scope of covalent modification while eliminating the limitations of naturally occurring anchor residues (poor reactivity, selectivity, etc.). The high efficiency of this reaction allows rapid ArM formation when using Az residues within the scaffold protein in the presence of various reactive components of cellular lysate.

ArM formation from various cofactors, including the esp-based[5] dirhodium cofactor, was demonstrated with multiple protein scaffolds. Dirhodium(II) tetracarboxylates are exceptionally active catalysts for the reaction of donor/acceptor rhodium carbenes, capable of turnover frequencies of 300 s^{-1} and turnover numbers of >1,000,000[6]. Robust catalytic reactions toward a range of carbene insertion reactions[7] that tolerate both air and water were developed[8].

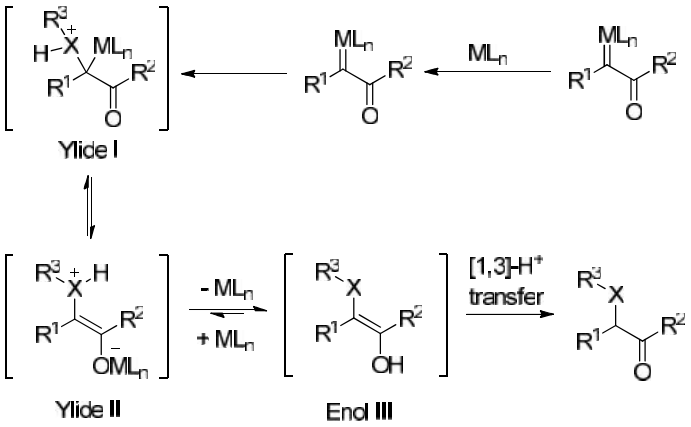
Carbenoid insertion of N-H bonds is an efficient method for preparing amino esters, amino ketones, and nitrogen-containing heterocycles[9]. This reaction was first described by Casanova and Reichstein in 1950,[10] and Merck demonstrated that N-H insertion could be used industrially for synthesis of (+)-Thienamycin in 1980[11]. Enantioselective variants of this reaction provided only low levels of enantioselectivity until Zhou and coworkers reported a method involving a spirophosphoric acid (SPA) co-catalyst in 2011[12,13]. Reactions between diazoacetates (or diazoketones) and Boc-amines catalyzed by dirhodium tetraacetates and SPAs provided up to 90% conversion and 98% enantiomeric excess.

Figure 3.1. Stepwise mechanism (adapted from ref[14])



The mechanisms of rhodium(II) catalyzed carbenoid insertions into polar X-H bond have been investigated using DFT calculations by Yu group[15]. For non-polar bond insertion (C=C, C-H, Si-H), a concerted mechanism is generally accepted. However, a stepwise ylide mechanism is believe to be operative for polar X-H insertions (NH, OH, SH)[14] (Figure 3.1). Further study by Cao[16] indicated that free enol III formed during the reaction (Figure 3.2), and that the rhodium catalyst is thus not present during the chirality-determining proton transfer step. Ylide II dissociates from the metal and undergoes an intramolecular [1,3]-proton shift to give the final N-H insertion product. Calculations suggest that the chiral SPA catalyst promotes the [1,3]-proton shift of enamine to impart enantioselectivity to the N-H insertion reaction. Herein, a dirhodium ArM constructed with the hope that the enantiodetermining proton transfer within a chiral protein scaffold would give rise to enantioselective N-H insertion.

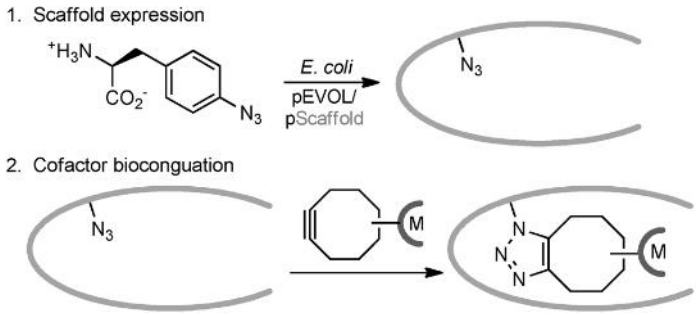
Figure 3.2. Proposed pathway of X-H insertions of α-diazoketones (adapted from ref[13])



3.2 A GENERAL METHOD FOR ARMS FORMATION THROUGH STRAIN-PROMOTED AZIDE ALKYNE CYCLOADDITION

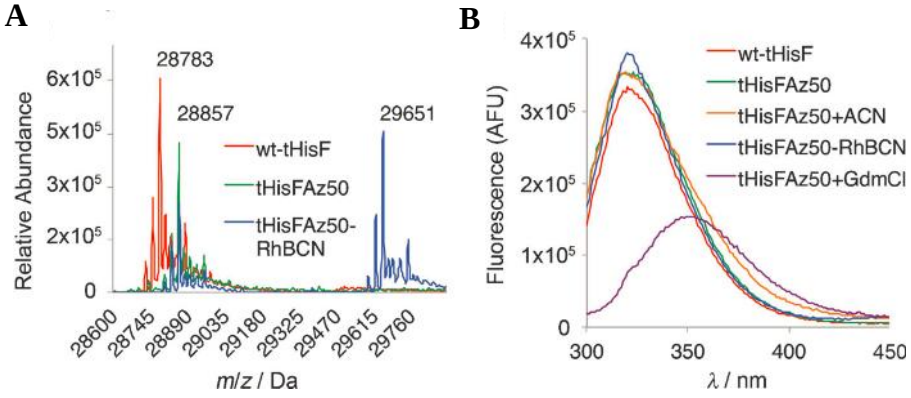
This work is published[17] and was conducted in collaboration with Hao Yang and Dr. Poonam Srivastava. As discussed above, a novel method providing orthogonal and site-specific bioconjugation with a broad range of scaffold scope is required. The high rate, bioorthogonality[18], and site-specificity of SPAAC have led to its wide adoption in chemical biology, but incorporating metal catalysts into proteins has not been described. Herein, we developed SPAAC strategy to generate ArMs from scaffold proteins containing a *p*-azido-L-phenylalanine (Az or Z) residue and catalytically active bicyclononyne (BCN)-substituted metal complexes (Scheme 3.1).

Scheme 3.1. ArM formation through SPAAC (taken from ref[17])



A thermostable α,β-barrel protein tHisF[19] from *Thermotoga maritima* was employed as protein scaffold. We used amber stop codon suppression to Az at representative positions 176, 50, and 199 in tHisF. High levels of scaffold expression and UAA incorporation was observed[*] (Figure 3.3A). The His_6-tagged scaffold proteins were purified by Ni-affinity chromatography and fluorescence[*] (Figure 3.3B) indicated that little structural perturbation resulted from UAA incorporation.

Figure 3.3. A) HR-ESI-MS of wt-tHisF, tHisF-Az50, and tHisFAz50-RhBCN (taken from ref[17]). B) Fluorescence spectra (290 nm) of wt-tHisF, tHisF-Az50 (in buffer, 60% CH_3CN, 6M guanidinium chloride [GdmCl])**, and tHisF-Az50-RhBCN** (CH_3CN, GdmCl) (taken from ref[17])

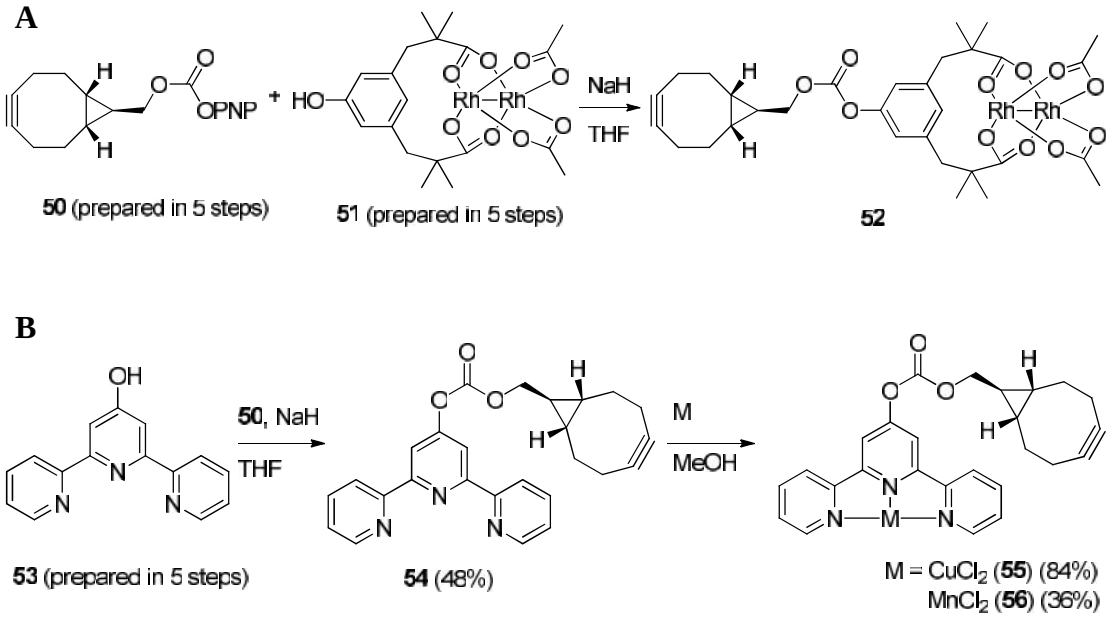


We next developed a modular approach to synthesize metal cofactors containing BCN[20] moiety, due to its small size and high SPAAC rates[3]. Given the high activity of these complexes toward carbene insertion reactions[7] that tolerate both air and water[8], a dirhodium tetracarboxylate cofactor was designed. Inspired by Du Bois and co-workers using tetramethyl *m*-benzenedipropionic acid ligands (esp)[5], we prepared the mixed esp/diacetate complex **52**[21] (Scheme 3.2 A)[†]. The carbonate linkage was not hydrolyzed, even after extended room

[*] Conducted by Dr. Poonam Srivastava
[†] Synthesized by Hao Yang

temperature incubation in various aqueous buffers. Two additional BCN cofactors, **55** and **56**, containing Cu^{22} and Mn^{23} terpyridine complexes, were prepared by metallating BCN-terpyridine **54**, which was formed from phenol **53** and carbonate **50** (Scheme 3.2B).

Scheme 3.2. (A) Syntheses of cofactor 52 and (B) Syntheses of cofactors 55 and 56

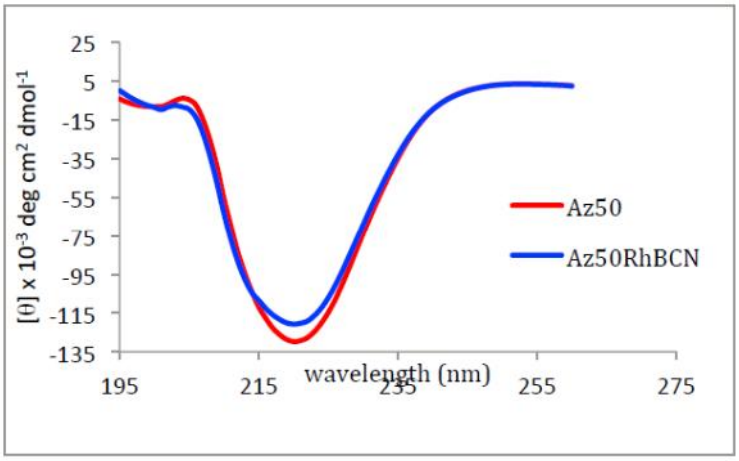


The reactivity of cofactor **52** toward tHisF-Az50 was explored[‡]. A solution of **52** (5 equiv) in 20% (v/v) CH_3CN/Tris buffer (50mM, 7.4) was added to a solution of each tHisF mutant (60 mM), and the reactions were incubated at 4 ºC. After maximum conversion was observed by MALDI mass spectrometry, the ArM/scaffold mixture was purified by preparative HPLC to remove all traces of unreacted cofactor. The isolated ArM/scaffold mixture was then characterized by ESI mass spectrometry (Figure 3.3A)[*]. Fluorescence and circular dichroism (CD) spectroscopy also showed no noticeable difference from a sample of pure scaffold, indicating little perturbation to the scaffold structure as a result of bioconjugation (Figures 3.3B

[‡] Conducted by Hao Yang and Dr. Poonam Srivastava

and 3.4).

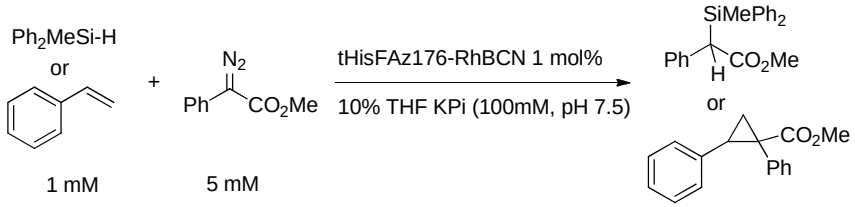
Figure 3.4. CD characterization of tHisF-Az50 and -Az50RhBCN (taken from ref[17])



To demonstrate the generality of our SPAAC ArM approach, we also covalently linked cofactor **52** to Az104 in the central pore of the engineered phytase scaffold[§]. ArM was confirmed by MALDI mass spectrometry following purification. We further established that dirhodium ArMs can catalyze the decomposition of diazo compounds and both Si-H and olefin insertion reactions involving these compounds (Table 3.1).

[§] Conducted by Dr. Poonam Srivastava

Table 3.1. Silane insertion catalyzed by ArM and small molecule dirhodium catalysts[**]**



Catalyst	Si-H insertion (%)[a]	Cyclopropanation (%)[a]
tHisFAz50-RhBCN	28	69 (cis: trans = 1: 1.8)
51-OAc	80	99 (cis: trans = 1: 1.8)

[a]All the yields were calculated relative to silane in the reactions by analysis of HPLC traces for crude reaction mixtures.

3.3 INITIAL STUDY ON DIRHODIUM ARMS CATALYZED ENANTIOSELECTIVE N-H CARBENOID INSERTION

The aforementioned dirhodium ArMs generated using SPAAC bioconjugation were being studied by Hao Yang and Dr. Poonam Srivastava for cyclopropanation reactions. Unfortunately, no selectivity was observed in reactions catalyzed by these systems. We attributed this lack of selectivity to the inability of the protein scaffolds selected for bioconjugation method development to fully encapsulate the cofactors selected for catalysis[24]. We sought to identify a scaffold protein that could impart selectivity to **52**. This would validate our hypothesis regarding the poor selectivity of our initial ArM designs, illustrate the importance of scaffold selection in ArM design and provide a platform for the directed evolution optimization.

An extensive search of different protein X-ray structures in the protein data bank (PDB) led to the identification of several members of the prolyl oligopeptidase (POP) family as potential ArM scaffolds because of their roughly cylindrical shapes (30-60Å) and large internal

[**] Conducted by Hao Yang

volumes (5-8×10^3 $Å^3$) for cofactor enclosure[25 ††]. This family includes POPs, dipeptidyl peptidases IV, oligopeptidases B and acylaminoacyl peptidases. All of these enzymes share a common fold comprising an α/β hydrolase domain, which contains a Ser-Asp-His triad for amide bond hydrolysis, capped by a β-barrel domain[24].

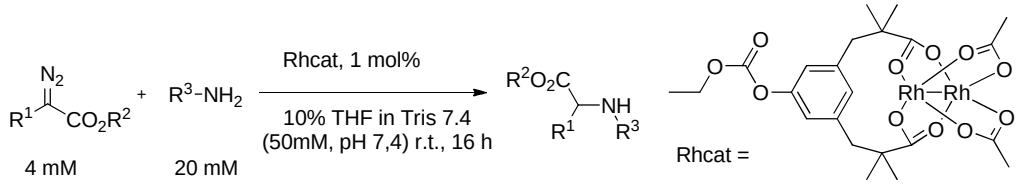
We initially selected a POP from *Pyrococcus furiosus* (*Pfu*) as a scaffold for ArM formation because of its high thermal stability[26]. An amber codon was introduced into the POP gene to replace the catalytically active serine (S477) with a Z residue (Z477). A POP gene whose codon usage was optimized for expression in *E. coli* was used as a template for genetic manipulation, and the resulting scaffold, POP-Z, was expressed in high yield[‡‡] (100mg/L). However, no reaction occurred between POP-Z and **52**. Homology model[27] revealed that four residues (E104, F146, K199 and D202) were largely responsible for blocking access to the active site. Mutated these four residues in POP-Z to alanine, and the resulting protein, POP-ZA4, underwent rapid bioconjugation in the presence of cofactor **52** at 4ºC to form POP-ZA4-**52**[†].

Hao Yang found POPZA4-**52** ArM can catalyze asymmetric cyclopropanation, although low selectivity (20%) was observed. While Hao Yang and Dr. Poonam Srivastava focused on cyclopropanation, I started working on enantioselective N-H insertions. Initial studies on N-H insertion were conducted in aqueous solution using dirhodium esp catalyst **57**[§§]. A numbers of diazo and amine compounds were tested (Table 3.2), and the results of these reactions showed that the dirhodium ArM can promote the desired N-H insertion of donor-

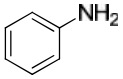
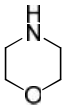
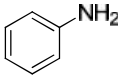
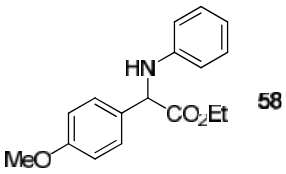
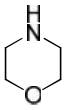
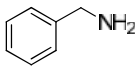
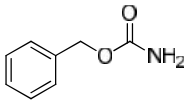
†† Model was established by Dr. Jared Lewis and Ken Ellis-Guardiola
‡‡ Conducted by Dr. Poonam Srivastava
§§ Prepared by Dr. Poonam Srivastava

acceptor carbenes (e.g., methyl 4-methoxylphenyldiazoacetate) with aniline-type substrates. H-acceptor carbenes (e.g., ethyl diazoacetate) or alkyl amines provided no conversion.

Table 3.2. Rhcat catazlyed N-H insertion scope



57

Entry	Diazo		Amine	Conv.[a]	Product[b]
1	R1	R2			
2	H	Et		<1	-
3	H	Et		<1	-
4	4-MeO-Ph	Me		95	 58
5	4-MeO-Ph	Me		2	-
6	4-MeO-Ph	Me		<1	-
7	4-F-Ph	Me		<1	-
8	4-CF3-Ph	Me		<1	-
9	H	Et		<1	-
10	4-MeO-Ph	Me		<1	-
11	4-F-Ph	Me		<1	-
12	4-CF3-Ph	Me		<1	-
13	H	Et		<1	-

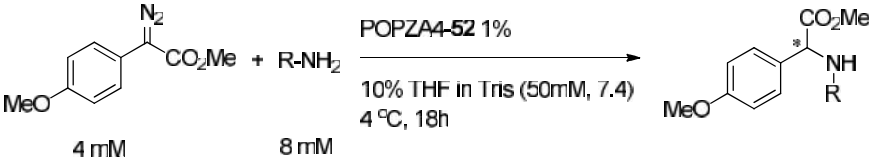
[a]Conversion iscalculated by (consumed amine mol)/(diazo mol)
[b]Product identified by HPLC-MS

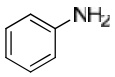
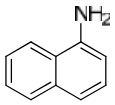
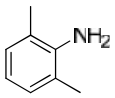
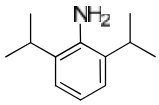
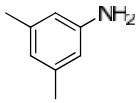
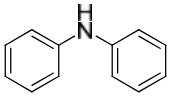
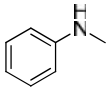
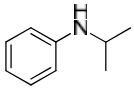
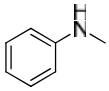
Based on the reactivity between donor-acceptor carbenes and anilines, we moved to biocatalysis using POPZA4-**52**. A few aryl amines were examined (Table 3.3, entries 1-8). The best result was from *N*-methyl aniline which provided 6% ee (Entry 7). Hao Yang observed that high concentrations of NaBr provided a marked increase in enantioselectivity of cyclopropanation reactions[24]. In the N-H insertion case, introducing 1M NaBr resulted in 11% ee, and slightly higher conversion (Table 3.3, entry 9).

Several researchers have hypothesized that cofactor movement within protein scaffolds can reduce ArM selectivity[28], and various strategies intended to decrease cofactor movement, including two-point covalent attachment[29] and metal coordination by histidine[30], have been shown to increase ArM selectivity. The latter strategy was used to improve POP-ZA4-**52** for cyclpropanation, given the established success of this method in peptide-based dirhodium catalysts[31]. Histidine and phenylalanine mutations were individually introduced at several residues within POP-ZA4 that projected towards the POP active site cavity (Figure. 3.5)[***]. The biocatalysis of the resulting ArMs showed improved e.e. in cyclopropanation reaction[†††]. These mutants were also applied in my N-H insertion reactions. H328 (Table 3.4, entry 1), F99 (Entry 2) and H99 (Entry 7) provided increased enantioselectivity.

[***] Conducted by Dr. Poonam Srivastava
[†††] Conducted by Hao Yang

Table 3.3. Amine susbstrates scope of POP-ZA4-52 catalyzed N-H insertion



Entry	Amine	Additive	conversion	e.e.
1		-	73	<1
2		-	41	4
3		-	44	2
4		-	<1	-
5		-	49	3
6		-	<1	-
7		-	70	6
8		-	63	5
9		1M NaBr	77	11

[a]Conversion calculated by (consumed amine mol)/(diazo mol)

[b]Enantiomeric excess identified by chiral HPLC

Figure 3.5. Homology model[27] of *Pfu* POP. The hydrolase domain is shown in green, the propeller domain is shown in grey and cofactor **52** linked at Z477 is shown in red. Sites of different mutations introduced into *Pfu* POP are shown as coloured spheres[24] (taken from ref[24]).

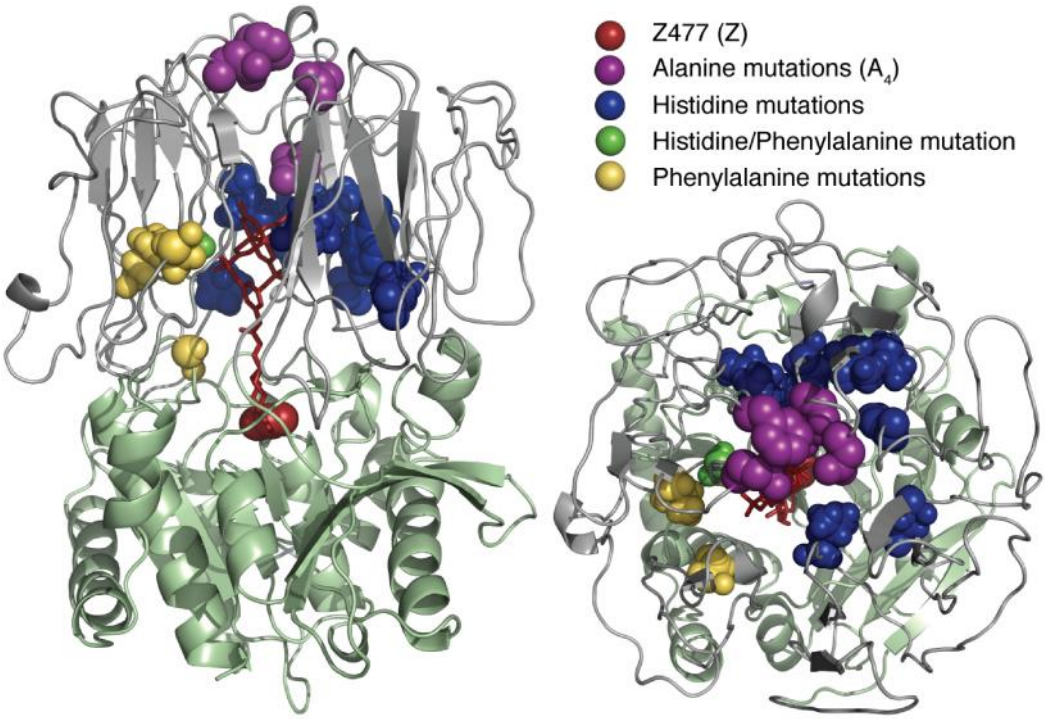
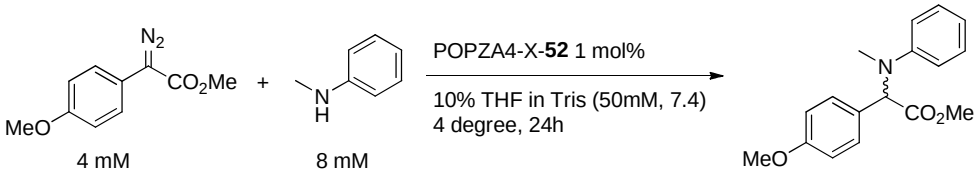


Table 3.4. Optimization of active site mutations



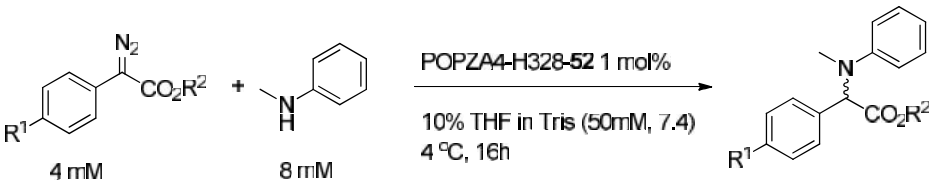
Entry	Variant (X)*	Conversion[a]	e.e.[b]
1	H328	80	36
2	H99	79	31
3	H139	67	17
4	H197	55	16
5	F594	69	9
6	F251	68	19
7	F99	77	28

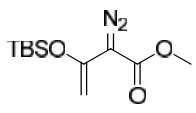
[a] Conversion calculated by (consumed amine mol)/(diazo mol) on HPLC
[b]Enantiomeric excess identified by normal phase chiral HPLC

* Mutants were prepared by Dr. Poonam Srivastava

To evaluate the scope of our dirhodium ArMs catalyzed N-H insertion, we performed initial studies on additional diazos and aryl amines. For *p*-CF_3, Cl, and Br substituted phenyldiazoacetates (Table 3.5, entries 1-3), only trace conversion was detected. *p*-phenyl substituted phenyldiazoacetate provided around 10% conversion and 5% ee, while more electron donating *p*-methoxy phenyldiazoacetate gave almost quantitative conversion and 36% ee (Entry 5). Changing R^2 group to ethyl led to slightly higher ee but only 70% conversion (Entry 6). Increasing the size of R^2 to *tert*-butyl and benzyl resulted in decreased selectivity and conversion (Entries 7 and 8). Switching carbene precursors completely shut down the reaction (Entry 9).

Table 3.5. Initial studies on expanded scope for diazo compounds

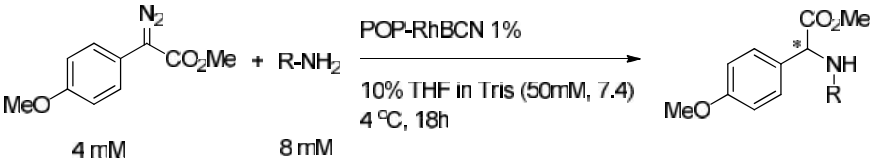


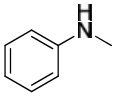
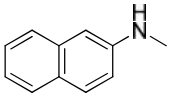
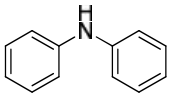
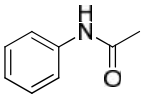
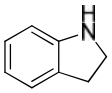
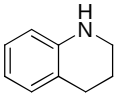
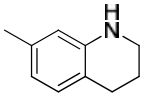
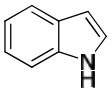
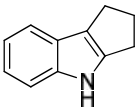
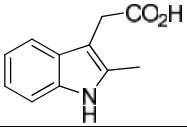
Entry	R1	R2	conversion[a]	e.e. [b]
1	CF_3	Me	<2%	N/A
2	Cl	Me	<2%	N/A
3	Br	Me	<2%	N/A
4	Ph	Me	<10%	5
5	OMe	Me	98	36
6	OMe	Et	70	37
7	OMe	*t*-Bu	3	N/A
8	OMe	Bn	50	13
9			N. R.	N/A

[a]Conversion calculated by (consumed amine mol)/(diazo mol) on HPLC
[b]Enantiomeric excess identified by normal phase chiral HPLC

POPZA4-H328 also catalyzed N-H insertion of a series of amine compounds. The aforementioned *N*-methyl aniline prodived 36% ee (Table 3.6, Entry 1). Initial experiments indicated that *N*-methyl-*N*-naphthylamine underwent slow reactions with 6% ee (Entry 2), whereas *N*,*N*-diphenylaniline afforded trace conversion (Entry 3). The less electron rich substrate *N*-acetyl aniline led to negligible conversion (Entry 4). Two aryl cyclic anilines, dihydroindole and tetrahydroquinoline, generated 11% and 42% ee, respectively. Further study on 5-methyl tetrahydroquinoline showed 40% enantioselectivity (Entried 5-7). Exploration of more electron rich N-H substrates (e.g. indole) led to intractable product mixtures (Entry 8). We hypothesized that indole underwent both N-H insertion and cyclopropanation process, so that introducing substitution groups on C2 and C3 of indole could eliminate cyclopropanation side reactions. Thus, two 2,3-disubstituted indoles completely excluded side reactions. Entry 9 yielded 60% N-H insertion product with 6% ee, while entry 10 produced only trace conversion.

Table 3.6. Initial studies on expanded scope for amines



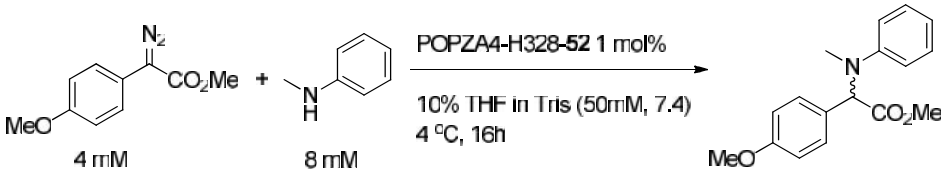
Entry	Amines	Conversion[a]	e.e. [b]
1		98	36
2		25	6
3		<1	N/A
4		<1	N/A
5		89	11
6		98	42
7		93	40
8		100	N/A[c]
9		60	6
10		<1	N/A

[a]Conversion calculated by (consumed amine mol)/(diazo mol) on HPLC
[b]Enantiomeric excess identified by normal phase chiral HPLC
[c]Multiple products detected.

Reaction parameters (buffer, co-solvent, pH, and additives) were also examined for further optimization (Table 3.7). pH ranging from 8.0 to 6.0 was studied using three buffers.

All of them showed same trend: optimal selectivity was obtained at pH 7.5 (Entries 1-11). A range of buffer types including Tris, NaPi, PIPES, HEPES, and KPi were tested. All showed similar level of selectivity and conversion (Entries 2, 4, 8, 12,13 and 14). Further increasing NaBr to 1.5M slightly improved ee to 45%.

Table 3.7. Optimization of biocatalysis conditions



Entry	Buffer	pH	NaBr	Conv.[a]	e.e. [b]
1	Tris	8	1M	95	37
2	Tris	7.5	1M	98	42
3	Tris	7.0	1M	98	40
4	NaPi	7.5	1M	95	42
5	NaPi	7	1M	93	41
6	NaPi	6.5	1M	92	40
7	NaPi	6	1M	92	35
8	Bistris	7.5	1M	90	41
9	Bistris	7	1M	90	40
10	Bistris	6.5	1M	89	40
11	Bistris	6	1M	87	34
12	Pipes	7.5	1M	92	42
13	Hepes	7.5	1M	95	41
14	KPi	7.5	1M KBr	95	40
15	Tris	7.5	1.5M	96	45

[a]Conversion calculated by (consumed amine mol)/(diazo mol) on HPLC
[b]Enantiomeric excess identified by normal phase chiral HPLC

With this optimized biocatalysis condition in hand (Table 3.7, entry 15), we sought to employ protein engineering that can systematically optimize selectivity. Previous results demonstrated that scaffold mutagenesis allows for improvement of enantioselectivity of this reaction (Table 3.4). A number of mutants combining H328 and phenylalanine mutations were thus expressed by Dr. Poonam Srivastava and Hao Yang. The biocatalysis using combined mutants is shown below (Table 3.8). While only H328 improved enantioselectivity, the H328/F99 double mutant[†] and H328/F99/F594 triple mutant provided further improvements, ultimately leading to N-H insertion with 47% ee (Table 3.8, entries 3 and 7). The observed moderate selectivity provide a platform for the directed evolution optimization.

Table 3.8. Optimization of combined active site mutations

Entry	Mutant*	Conversion	e.e.
1	H328	98	42
2	F97H328	87	34
3	F99H328	99	47
4	F251H328	88	18
5	F594H328	90	40
6	F99F594	90	17
7	F99F594H328	97	47
8	F64H328	89	23

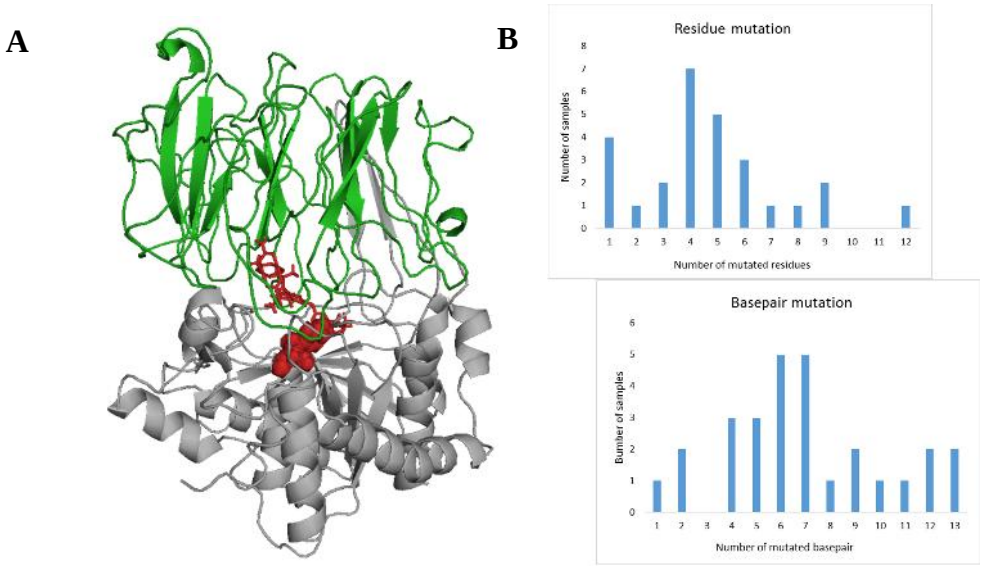
[a]Conversion calculated by (consumed amine mol)/(diazo mol) on HPLC
[b]Enantiomeric excess identified by normal phase chiral HPLC

[†] Prepared by Dr. Poonam Srivastava and Hao Yang for cyclopropanation reactions

3.4. DIRECTED EVOLUTION PROTOCOLS DEVELOPMENT AND OPTIMIZATION FOR DIRHODIUM ARMS

While the enantioselectivity exhibited by POPZA4-F99H328-**52** is modest (47% ee), it provides a good starting point for subsequent directed evolution.[32] No examples in which directed evolution, involving random mutations throughout a scaffold, have been reported to improve the function of an ArM. In terms of library design, we targeted the POP β domain (Figure. 3.6A green), which encapsulates the dirhodium cofactor. While mutations could be made throughout the entire POP structure, we reasoned that those in the hydrolase domain (Figure. 3.7A gray) would be less likely to impact selectivity. The mutations were made through the use of error-prone PCR[33]. One key parameter in error prone PCR is the $MnCl_2$ concentration, which controls the mutation rate. A range of $MnCl_2$ concentrations was examined from 100uM to 1000uM. We observed that higher concentration lead to higher mutation rate but reduced DNA concentration. Using 300uM $MnCl_2$ produced six base pair mutations resulting four residue changes on average (Figure 3.6 B).

Figure 3.6. (A) POP β domain highlighted in green and (B) residue/basepair mutations distribution using 300uM $MnCl_2$



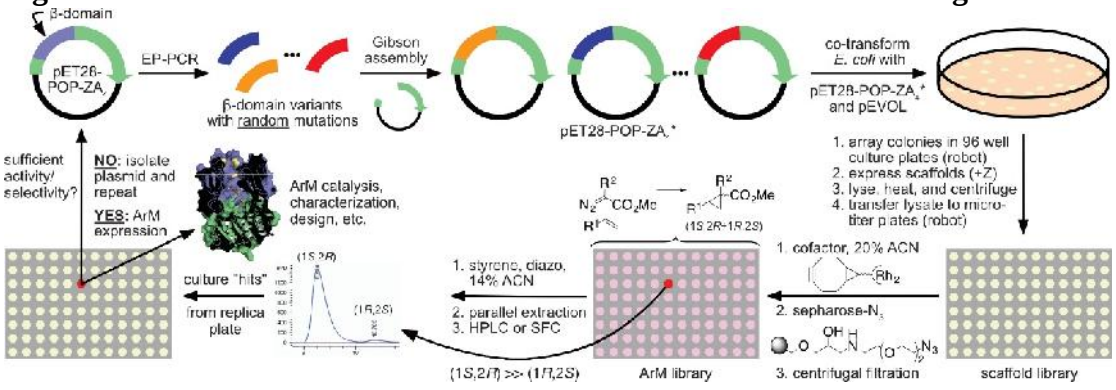
Gibson assembly was used to reassemble the plasmid. Many parameters, including molar ratio of error-prone/backbone fragment (PCR amplified), incubation time, and backbone fragment concentration, were assessed by number of colonies upon transformation of *E. coli.* by electroporation. Efficient Gibson assembly required at least 1h incubation time (Table 3.9, entry 3). Optimal condition is using 100ng and 5:1 or 10:1 ratio of error-prone/backbone fragment, and 1h incubation under 50 ºC (Entries 5 and 6). This protocol is able to consistently generate over 200 colonies/plate on average.

Table 3.9. Optimization of Gibson assembly

Entry	1	2	3	4	5	6
Molar ratio	3:1	3:1	3:1	3:1	5:1	10:1
Time	30min	30min	1h	1h	1h	1h
Backbone conc.	50ng	100ng	50ng	100ng	100ng	100ng
Colony numbers	-	-	+	++	+++	+++

Dr. Poonam Srivastava initialized directed evolution protocol development. She introduced two additional steps, bioconjugation and resin scavenging, into a standard directed evolution workflow (Figure 3.7).

Figure 3.7. Directed evolution workflow of an ArM constructed through SPAAC



POP expression was conducted in 24-well plates in order to obtain higher protein concentrations. Optimized conditions allowed production of large quantities of protein in 24-well plates. OD measurement revealed that final cells amount expressed in 24-well plate was 10-fold excess compared with 96-well plate (Table 3.10, entries 1 and 2). This is potentially due to better aeration in 24-well plate. Increasing volume to 6mL could further improve total cell amount (Entry 3). Neither adding pipette tip nor doubled broth nutrition were able to further improve expression levels (Entries 4 and 5).

Table 3.10. POP expression optimization

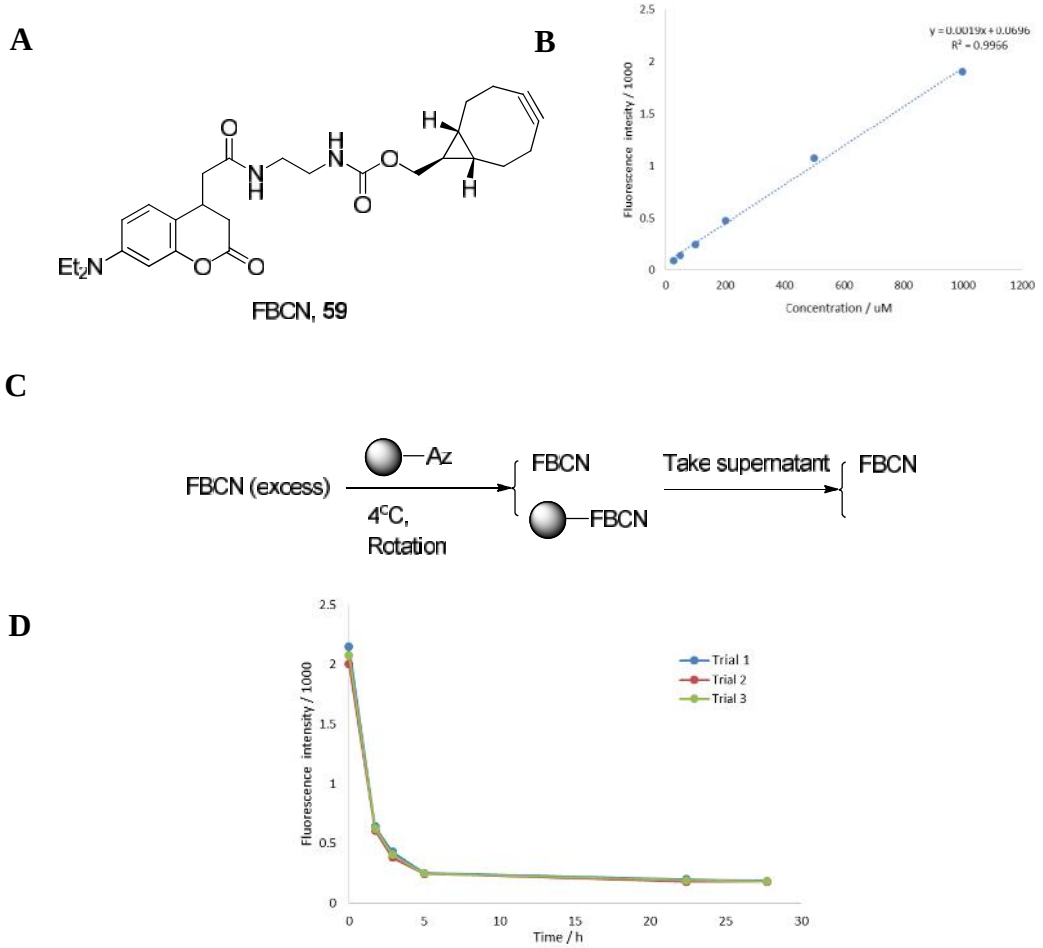
Entry	Plate type	Media broth	volume/mL	Shaking RPM	Final OD[a]
1	96-well	2YT	1	250	2.6
2	24-well	2YT	4	250	6.3
3	24-well	2YT	6	210	5.3
4[b]	24-well	2YT	6	210	5.0
5	24-well	double 2YT[c]	6	210	5.2

[a]OD detected by plastic cuvette under 600nm
[b]Autoclaved pipette tip added in each well
[c]doubled 2YT nutrition

Cell lysis, heat treatment, and centrifugation produced cell lysates containing POP variants essentially free of other proteins, allowing high yields in the following bioconjugation reaction. Bioconjugation was conducted under 4 ºC with 20% ACN as cosolvent. Previous ArM purification required preparative HPLC which is not feasible for high-through put screening in 96-well plates. Thus, Hao Yang prepared azide-substituted sepharose resin. We hope that addition of resin upon bioconjugation could effectively scavenge residual RhBCN cofaotor **52**. With the aim of identifying quality of self-prepared

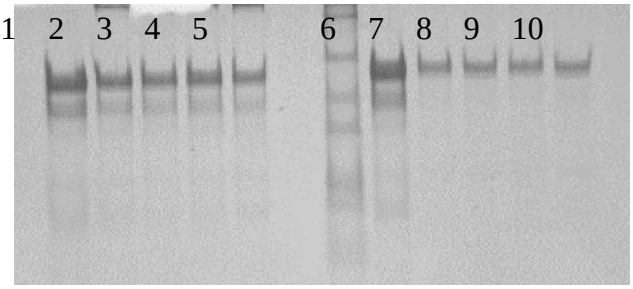
resin[‡], I prepared a coumarin-containing fluorescent BCN (FBCN) cofactor **59**[§] (following synthesis protocol by Jeff Montgomery) (Figure 3.8 A) to monitor effective Azide loading. This cofactor has maximum fluorescence excitation at 325nm as well as emission at 500nm. A fluorescence-based assay was then developed by measuring residual **59** after overnight resin incubation (Figure 3.8 B). Assay is sensitive to quantify **59** concentration between 1uM and 100uM. A qualified resin should have azide loading no less than 35mM.

Figure 3.8 (A) FBCN cofactor 59, (B) calibration curve of 59, (C) schematic illustration of azide loading assay and (D) fluorescence intensity trace of azide loading assay



[‡] Prepared by Hao Yang
[§] Initial synthesis was done by Jeff Montgomery

Figure 3.9. Protein loss monitored by gel.

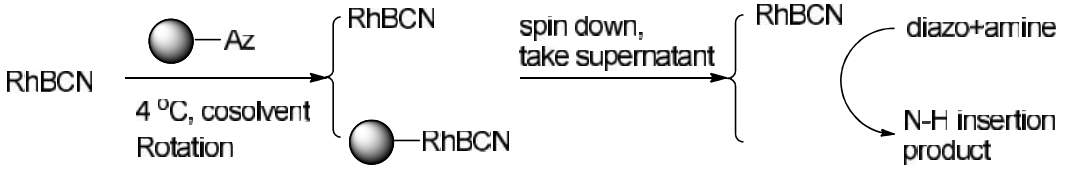


Lane 1-5: incubate sepharose resin with POP solution, and check time point of 0h, 1h, 2h, 3h and 4h, respectively; Lane 6-10: incubate azide sepharose resin with POP solution, and check time point of 0h, 1h, 2h, 3h and 4h, respectively.

One major issue introduced from resin scavenging step is protein loss. We found that the protein concentration decreased after incubation with both regular sepharose resin and azide-sepharose resin (Figure 3.9). This indicated that protein loss is resulting from POP-sepharose interaction, not azide modification on resin.

On one hand, a certain amount of resin is required to eliminate residual cofactor catalyzed background reaction; on the other hand, overloaded resin could quench robust ArMs and then dramatically reduce enzyme loading. To address this dilemma, a range of parameters were examined. ACN as cosolvent (Table 3.11, entry 2), increased resin volume (Entry 3 and 4), elongated incubation time (Entry 7), and additional cosolvent (Entries 5 and 8) improved resin removal. However, 40% ACN led to severe disruption of protein structure so that protein aggregates were observed (Entry 4). Ultimately, single digit of background reaction was detected when 200uL resin was used with ACN (20%) and 24h incubation (Entry 8). Protein loss was monitored by gel (Figure 3.10). Upon completion of resin scavenging, a clear POP band indicated enough purified ArM proteins for upcoming biocatalysis. Dr. Hyun June Park is currently working on further optimization of resin amount.

Table 3.11. Optimization of resin scavenging protocol

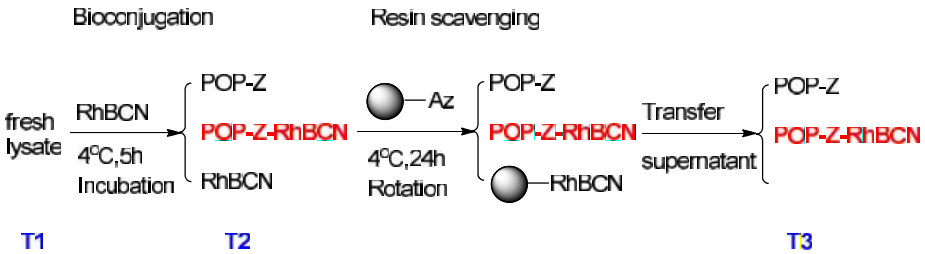


120uL, 100uM

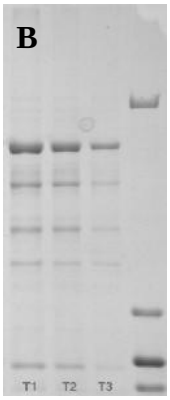
Entry	Cosolvent (%)	Resin volume	Incubation time	Additional cosolvent	Protein present?[a]	Background conversion
1	THF(10)	100uL	12h	-	+	100
2	ACN(20)	100uL	12h	-	+	97
3	ACN(20)	200uL	12h	-	+	57
4	ACN(20)	300uL	12h	-	-	20
5	ACN(30)	200uL	12h	-	+	30
6	ACN(40)	200uL	12h	-	-	17
7	ACN(20)	200uL	24h	-	+	15
8	ACN(20)	200uL	24h	+	+	6

Figure 3.10. (A) Schematic illustration of resin scavenging protocol and (B) monitored protein concentration change by gel

A



B



Methods utilitzing automation were developed in colony picking, liquid handling, lysate transfer and resin removal steps. After the process was sufficiently optimized, the first round of directed evolution from an engineered parent protein, F99H328, produced the mutant G1P1_D5 which achieved 54% ee. Further work on both the enhancing of the selectivity of this system through directed evolution and the optimization of the protocol are currently in progress, conducted by Dr. Hyun June Park and David Upp.

3.5 CONCLUSIONS

In summary, we have designed and constructed a dirhodium ArM via novel SPAAC bioconjugation method. This ArM preforms robust catalytic activity on carbenoid cyclopropanation, Si-H insertion and N-H insertion. Enantioselective N-H insertion has been develop and optimized through optimization of reaction conditions and engineering of POP. Then we have established ArM directed evolution protocol to improve selectivity.

3.6 EXPERIMENTAL

Materials

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. Deuterated solvents were obtained from Cambridge Isotope labs. Silicycle silica gel plates (250 mm, 60 F254) were used for analytical TLC, and preparative chromatography was performed using SiliCycle SiliaFlash silica gel (230-400 mesh). Azide Agarose was purchased from Click Chemistry Tools LLC. Labquake™ Tube Shaker/Rotators was purchased from Thermo Scientific (Catalog# 4002110Q). Oligonucleotides were purchased from Integrated DNA Technologies (San Diego, CA). Plasmid pEVOL-pAzF was provided by the Schultz group of the Scripps Research Institute, CA[S1]. *E. coli* DH5α and BL21 (DE3) cells were purchased from Invitrogen (Carlsbad, CA). Nco I, Xho I restriction enzyme, T4 DNA ligase, Taq DNA polymerase and Phusion HF polymerase (Cat# 530S) were purchased from New England Biolabs (Ipswitch, MA). Luria broth (LB), rich medium (2YT) and Agar media were purchased from Research Products International (Mt. Prospect, IL). Qiagen DNA extraction kit (Cat# 28706) and plasmid isolation kit (Cat# 27106) were purchased from QIAGEN Inc. (Valencia, CA) and used according to the manufacturer's instructions. DNA purification kit (Zymo, Cat# D4004) was purchased from Zymo research (Irvine, CA) and used as recommended. Library colonies were picked using an automated colony picker (Norgren Systems). All genes were confirmed by sequencing at the University of Chicago Comprehensive Cancer Center DNA Sequencing & Genotyping Facility (900 E. 57th Street, Room 1230H, Chicago, IL 60637). Electroporation was carried out on a Bio-Rad MicroPulser using method Ec3.

Ninitrilotriacetic acid (Ni-NTA) resin and Pierce® BCA Protein Assay Kits were purchased from Fisher Scientific International, Inc. (Hampton, NH), and the manufacturer's instructions were following when using both products (for Ni-NTA resin, 5 mL resin were used, with buffers delivered by a peristaltic pump at a rate of 1 mL/min, in a 4 °C cold cabinet). Amicon® 30 kD spin filters for centrifugal concentration were purchased from EMD Millipore (Billerica, MA) and used at 4,000 g at 4 °C. Biotage reverse phase columns (SNAPKP-C18-HS) were purchased from Biotage.

General procedures

Unless otherwise specified, all reactions were prepared in flame or oven-dried glassware under an inert N_2 atmosphere using either syringe or cannula techniques. TLC plates were visualized using 254 nm ultraviolet light. Flash column chromatography was carried out using Silicycle 230-400 mesh silica gel. ^{1}H and ^{13}C NMR spectra were recorded at 500 MHz and 126 MHz, respectively, on a Bruker DMX-500 or DRX-500 spectrometer, and chemical shifts are reported relative to residual solvent peaks. Chemical shifts are reported in ppm and coupling constants are reported in Hz. Yields were determined by HPLC with 1,3,5-trimethoxybenzene as the internal standard and reported as the average of two trials from the same batch of ArM set up in parallel. High resolution ESI mass spectra were obtained using an Agilent Technologies 6224 TOF LC/MS. Low resolution ESI mass spectra were obtained using Agilent 6130 LC-MS. Amicon® 50 mL 30 kD cutoff centrifugal filter was used to concentrate or wash protein solutions. Protein concentrations were measured using the

Pierce® BCA Protein Assay Kit and protein stocks were then stored at -80 °C until use. Circular dichroism (CD) spectra were obtained on a JASCO J-1500 CD Spectrometer.

Standard molecular cloning procedures were followed. To introduce single mutations, the same procedure using site directed overlap extension PCR reported previously[17,24].

tHisF cloning

A gene encoding the cyclase subunit (tHisF) of the imidazole glycerol phosphate synthase enzyme complex from Thermotoga maritima was amplified from pET11c-tHisF20 by PCR using gene specific primers containing NdeI (forward) and XhoI (reverse) restriction sites. The gene was cloned into the NdeI and XhoI sites of pET22b so that scaffolds would be expressed with a C-terminal hexa-histidine tag for Ni-NTA affinity chromatography. Amber mutations were introduced into the tHisF gene at positions L50, G176 and I199 by site directed overlap extension (SOE) PCR7. For each mutation, two separate polymerase chain reactions were performed, each using a perfectly complementary flanking primer at the 5' and 3' end of the sequence and a mutagenic primer. The PCR conditions were as follows: Phusion HF buffer 1x, 0.2 mM dNTPs each, 0.5 µM forward primer, 0.5 µM reverse primer, 0.02 U/µL Phusion polymerase and 1 ng/mL template plasmid. The resulting two overlapping fragments that contained the base pair substitution were then assembled in a second PCR using the flanking primers resulting in the full-length mutated gene. For phytase gene cloning, genomic DNA was isolated from *Bacillus amyloliquefaciens* (ATCC#23350) and genomic PCR was done using gene specific primers having above restriction sites on the flanking region. Genomic PCR was performed in one step using the same conditions as

above, except template (genomic DNA) concentration was increased to 500 ng/mL. An amber stop codon (Y104Az) was introduced for ArM formation, and three alanine mutations (N99A, N100A, D101A) were introduced to improve access to the site of Az mutation within the scaffold. Nucleotide sequences for the all the primers are summarized below.

T7 for 5'-GCG AAA TTA ATA CGA CTC ACT ATA-3'

T7 rev 5'-TTA TGC TAG TTA TTG CTC AGC GG-3'

tHisF L50Az for 5'-GAA CTC GTT TTT TAG GAT ATC ACC GCG-3'

tHisF L50Az rev 5'-CGC GGT GAT ATC CTA AAA AAC GAG TTC-3'

tHisF G176Az for 5'-AGT ATC GAC AGA TAG GGC ACA AAA TCG-3'

tHisF G176Az rev 5'-CGA TTT TGT GCC CTA TCT GTC GAT ACT-3'

tHisF I199Az for 5'-ACA CTT CCC ATC TAG GCT TCC GGT GGT-3'

tHisF I199Az rev 5'-ACC ACC GGA AGC CTA GAT GGG AAG TGT-3'

tHisF C9ala for 5'-AGA ATA ATC GCG GCG CTC GAT GTG AAA-3'

tHisF C9ala rev 5'-TTT CAC ATC GAG CGC CGC GAT TAT TCT-3'

tHisF K243C for 5'-GAG TAC CTC AAA TGC CAC GGA GTG AAC-3'

tHisF K243C rev 5'-GTT CAC TCC GTG GCA TTT GAG GTA CTC-3'

tHisF D174C for 5'-C TCACC AGT ATC TGC AGA GAC GGC-3'

tHisF D174C rev 5'-GCC GTC TCT GCA GAT ACT GGTGA G-3'

Expression and purification protocol:

pET22b-tHisF (or mutants thereof) and pEVOL-pAzF were co-transformed into electrocompetent *E. coli* BL21 (DE3)8, these cells were allowed to recover in SOC medium (37 °C, 50 min), then plated onto LB amp+Cm agar plates (6.25 g LB powder mix, 4 g agar, 250 mL DDI water, 0.1 mg/mL ampicillin, 0.05 mg/mL chloramphenicol), and incubated at 37 °C for 16 h. Several colonies appeared on overnight-incubated plates; a single colony from this plate was inoculated in 5 mL 2YT medium having antibiotics with the same concentrations as above. The culture was incubated overnight at 30 °C with constant shaking at 250 rpm. On the following day, 3 mL of the overnight cultures was used to inoculate 300 mL of fresh 2YT media having the same antibiotics, in 1 L Erlenmeyer flask. The culture was incubated at 30 °C, 250 rpm, and protein expression was induced by adding 1mM IPTG, 2mM 4-Azido-phenyl alanine and 1% (w/v) Larabinose when OD600 reached 1. The induced culture was allowed to grow for 12 hours, and then the cells were harvested by centrifugation at 4 °C, 3000 x g for 20 minutes. Cell pellets were re-suspended in 30 mL PBS (pH 7.5) and sonicated (40 amplitude, 30 second burst, 10 minute total process). Lysed culture was then clarified at 16000 x g, 4 °C for 30 minutes and supernatant thus obtained was purified by Ni-NTA resin using manufacturer's instructions. Purified protein was buffer exchanged to 10 mM Tris (pH 7.5) and measured by Pierce® BCA Protein Assay Kit as recommended.

Error-prone library construction

A codon optimized gene for Prolyl oligopeptidase (POP) was obtained from GenScript USA Inc (Piscataway, NJ) and cloned into pET28a plasmid vector using NcoI and XhoI restriction sites. The gene was cloned upstream of a C-terminal hexa-histidine tag for

Ni-NTA affinity chromatography. Mutant libraries were constructed by conducting error-prone PCR for the β-domain and regular PCR for the vector from a template plasmid (the library parent). The error-prone PCR conditions for mutating the target domain were as follows: Taq polymerase buffer 10x, 0.2 mM dNTPs each, 0.5 μM forward primer (ep-forward), 0.5 μM reverse primer (ep-reverse), 0.3 mM $MnCl_2$, 0.05 U/μL Taq polymerase and 1.0 ng/mL template plasmid.

Thermal cycler was programmed as:

1) 98 °C-120 seconds

2) 95 °C-45 seconds

3) 52 °C-30 seconds

4) 68 °C- 90 seconds

5) Repeat cycles from 2) to 4) 25 times

6) 68 °C- 10 mins

7) 4 °C- hold

The regular PCR conditions for amplifying the vector were as follows: Phusion polymerase buffer 1x, 0.2 mM dNTPs each, 0.5 μM forward primer (vec-forward), 0.5 μM reverse primer (vec-reverse), 0.02 U/μL Phusion HF polymerase and 1.0 ng/mL template plasmid.

Thermal cycler was programmed as:

1) 99 °C-60 seconds

2) 95 °C-30 seconds

3) 54 °C-45 seconds

4) 72 °C- 130 seconds

5) Repeat cycles from 2) to 4) 30 times

6) 72 °C- 10 mins

7) 4 °C- hold

The resulting two overlapping fragments (the mutated domain and the vector) were then assembled using Gibson assembly kit resulting in the full-length mutated gene. The reaction conditions were as follows: in a 200 µL PCR tube, prepare a 5:1 (molar ratio) mixture of mutated domain and vector (total mass of the vector is ~50 ng) in 10 µL molecular grade water. Add 10 µL 2X Gibson Assembly Master Mix in a final volume of 20 µL and incubate the reaction at 50°C for 60 minutes. The reaction mixture was cleaned using DNA purification kits and transformed into *E. coli* BL21 cells with pEVOL-pAzF plasmid. Cells were spread on LB kanamycin plates (6.25 g LB powder mix, 4 g agar, 250 mL DDI water, 0.05 mg/mL kanamycin) before recovering in SOC medium for 1 hour at 37 °C. Plates were incubated at 37 °C overnight; typically, greater than 100 colonies were observed when 1/5 of the outgrowth was used. To examine the mutant rates for a library, 30 colonies were inoculated on LB media (with 0.05 mg/mL kanamycin, 0.02 mg/mL chloramphenicol) and grown overnight at 37 °C, 250 rpm. Mutant plasmid from these overnight grown cultures were isolated using kit from Qiagen (Valencia, CA) and given for sequencing. Plasmid sequencing was done at the U Chicago sequencing facility, and partB-for and partC-rev primers were used for sequencing reactions. Nucleotide sequences for all the primers used are shown below:

ep-forward	5’- CTG AGT GAT AAA CTG TTT CCG GAA GTG TG -3’
ep-reverse	5’- AGA CGA TAC GGA ATC GTA AAC GAG GTG T -3’
vec-forward	5’- GCG CTA CAC CTC GTT TAC GAT TCC GTA TC -3’
vec-reverse	5’- GGG AGA ACT GTT CCC ACA CTT CCG GAAA -3’
partB	5’- TCT GGA TGG AAA ACC TGG AA -3’
partC	5’- TGC AAT GAA GTC ATC GAA CA -3’

Resin azide loading measurement

FBCN **59** was dissolved in DMF to final concentration 100mM. 50uL FBCN solution, 400uL Tris buffer (50mM, pH 7.4) and 50uL resin were mixed and rotated on the Labquake™ Tube Shaker/Rotator in a 4 °C cold cabinet for 24 hours. Upon completion, the sample was spun down under 10,000rpm for 5mins, after which 300uL supernatant was transferred to a microtiter plate. Measurement was conducted by using Tecan Plate Reader Pro200. Excitation: 325nm, emission 500nm.

Library expression, lysis and purification

Library colonies were picked using an automated colony picker (Norgren Systems) and arrayed in 1-mL 96-well plates containing 300 µL LB media with 50 µg/mL kanamycin and 20 µg/mL chloramphenicol. In each plate, 2~4 wells were saved for picking parent colonies as positive controls, and 2 wells were left blank as negative controls. Cells were grown overnight for 14~16 hours at 37 ^{o}C, 250 rpm. 50 µL of the overnight culture was used to inoculate 6 mL 2YT media (with 50 µg/mL kanamycin and 20 µg/mL chloramphenicol) in

10-mL 24-well plates. Following growth at 37 ºC, 200 rpm for about 5 hours, to an OD_{600} = 1.0, enzyme expression was induced by adding 1mM IPTG, 2mM 4-Azido-phenyl alanine and 1% (w/v) L-arabinose. After growth at 37 ºC, 200 rpm for about 16 hours, the cultures were harvested by centrifugation at 2000 rpm, 20 mins and the supernatants were discarded. The cell pellets were washed by adding 4 mL Tris-Cl buffer (50 mM, pH 7.4) to each well and incubating the plate at 37 ºC, 200 rpm (for 10 mins) until the pellets were loosen up, after which the suspension was submitted to centrifugation at 2000 rpm, 10 min. The supernatants were discarded and the above washing process was repeated again.

The cell pellets were then suspended in 600 µL Tris-Cl buffer (50 mM, pH 7.4) containing 0.75 mg/mL lysozyme. After incubation plate at 37 ºC, 250 rpm for 60 mins, plate was flash frozen in liquid nitrogen for 10 mins and thawed in 37 ºC water bath. 60 µL Tris-Cl buffer (50 mM, pH 7.4) containing 1.0 mg/mL lysozyme was added and the plate was incubated at 37 ºC, 250 rpm for 30 mins. The lysed cells were submitted to a heat treatment for 15 mins in 75 ºC water bath, after which the plate was centrifuged at 3600 rpm, 20 mins to remove cell debris. 480 µL of the lysate supernatant was transferred to a 2-mL 96-well plates for bioconjugation. Additionally, 10 µL of the lysate supernatants from randomly-picked wells were saved in a 1.5-mL microcentrifuge tube for future SDS-PAGE analysis.

Bioconjugation

For bioconjugation, 120 µL of cofactor **52** in acetonitrile (93.75 µM, 0.0819mg/mL) was added to lysate of each mutant and the reaction mixture was incubated at 4 ºC, 600 rpm

overnight. 10 μL of the bioconjugation mixtures from wells picked above were saved in a 1.5-mL microcentrifuge tube for SDS-PAGE analysis.

Resin scavenging

50 μL azide agarose resin was then added. The plate was sealed tightly with a cap mat (Arctic white LLC, cat#AWSM-1003S), wrapped with aluminum foil, and rotated on the Labquake™ Tube Shaker/Rotator in a 4 °C cold cabinet for 24 h. The resulting suspension was centrifuged at 3600 rpm, 10 mins. 300 μL of the supernatant was carefully transferred to a new 2-mL 96-well plate with a Microlab NIMBUS liquid handling robot. Additionally, 10 μL of the supernatants from wells picked above were saved in a 1.5-mL microcentrifuge tube for SDS-PAGE analysis.

Biocatalysis

150 μL Tris-Cl buffer (50 mM, pH 7.4, with 4.5 M NaBr) was added to each well and the plate was shaked at 4 °C, 20 mins. A combined stock solution containing 1,2,3,4-tetrahydroquinoline (2.5 mM final concentration) and methyl *p*-methoxyphenyldiazoacetate (0.5 mM final concentration) was added to initiate biocatalysis. The plate was sealed with a plastic lid and incubated overnight at 4 °C, 600 rpm. The reaction was quenched by adding 300 μL hexane and incubated at 600 rpm, 10 mins. The biphasic mixture was clarified by centrifugation at 2000 rpm, 10 mins, after which the plate was flash frozen with liquid nitrogen for 10 mins and kept at -20 °C for 10 mins. 300 μL from the top hexane layer was carefully transferred to a 200-μL 96-well polypropylene plate without interrupting the frozen

aqueous layer. Hexane was evaporated in a vacuum desiccator at room temperature and 120 µL 50% isopropanol/hexane was added to redissolve the mixture. The reactions were filtered through a 96-well filter plate (PTFE, 0.45 µm) and analyzed for cyclopropanation on HPLC. The chiral-HPLC to determine enantioselectivities was performed on Agilent 1100 Series HPLC system using a Phenomenex Lux® 3u Cellulose-1 column (1000 Å, 3 µM, 4.6 mm i.d. x 250 mm), with a flow rate of 1.0 mL/min and detection wavelength set at 254 nm. The following gradient was used: 5% A and 95% B isocratic, keep 30mins. A = isopropanol with 0.1% AcOH; B = hexane with 0.1% AcOH.

Enzyme purification

The putative hits identified from the library screening was grown, expressed, lysed, and purified according to a previous report[24]. A colony for the selected mutant was inoculated in 5 mL 2YT medium having antibiotics with the same concentrations as above. The culture was incubated overnight at 37 °C with constant shaking at 250 rpm. On the following day, 5 mL of the overnight cultures was used to inoculate 500 mL of fresh 2YT media having the same antibiotics, in 5 L Erlenmeyer flask. The culture was incubated at 37 °C, 250 rpm, and protein expression was induced by adding 1mM IPTG, 2mM 4-Azido-phenyl alanine and 1% (w/v) L-arabinose when OD_{600} reached 1. The induced culture was allowed to grow for 12 hours, and then the cells were harvested by centrifugation at 4 °C, 3000 x g for 20 minutes. Cell pellets were re-suspended in 30 mL PBS (pH 7.5) and sonicated (40 amplitude, 30 second burst, 10 minute total process). Lysed culture was clarified by centrifugation at 16000 x g, 4 °C for 30 minutes and supernatant thus obtained was purified by Ni-NTA resin using

manufacturer's instructions. Purified protein was buffer exchanged to 10 mM Tris (pH 7.5) and measured by Pierce® BCA Protein Assay Kit as recommended.

General procedure for analytical bioconversions

Solutions of aryldiazoacetates (25 µL, 96 mM, in THF), amine substrates (25 µL, 480 mM, in THF), and POP-ZA4-X-**52** solution (500 µL, the effective ArM concentration adjusted to 48 µM with respect to the dirhodium cofactor according to the aforementioned method) were added to a 1.5 mL microcentrifuge tube. The final concentrations of the reagents were: 20 mM amine substrate, 4 mM diazo substrate, 44 µM POP- ZA4-X-**52**. The resulting mixture was left shaking at 750 rpm at 4 °C overnight. The reaction was quenched by adding 20 µL 1,3,5-trimethoxybenzene solution (30 mM, in THF) and 600 µL ethyl acetate. The mixture was vortexed and centrifuged (15,000 x g, 3 min). The top organic layer was collected and the bottom aqueous layer was extracted with 600 µL ethyl acetate twice. The organic extracts were combined, evaporated and re-dissolved in 200 µL THF. 4 µL THF solution of the crude product was analyzed on RP-HPLC to determine conversions; 50 µL THF solution of the crude product was purified on preparative-HPLC to isolate the cyclopropane product, which was analyzed on NP-HPLC to determine enantioselectivities. The conversions and enantioselectivities were reported as the average of two trials from the same batch of ArM set up in parallel. The RP-HPLC to determine conversions was performed on an Agilent 1100 Series HPLC system using an Agilent Eclipse Plus C18 column (95 Å,

3.5 µM, 4.6 mm i.d. x 150 mm), with with a flow rate of 1.0 mL/min and detection wavelength set at 254 nm. The preparative HPLC used the same method as above. The normal phase HPLC to determine enantioselectivity was performed on Agilent 1100 Series HPLC system using a Phenomenex Lux® 3u Cellulose-1 column (1000 Å, 3 µM, 4.6 mm i.d. x 250 mm), with a flow rate of 1.0 mL/min and detection wavelength set at 254 nm.

3.7 ACKNOWLEDGEMENT

This work was supported by the University of Chicago, Department of Chemistry. I wish to thank Dr. Poonam Srivastava for assistance with all molecular biology works and lively discussions. I am grateful to Hao Yang for all initial work and nice suggestions on dirhodium ArM work. I would like to thank Ken Ellis-Guardiola who identified the POP scaffold and supported all spectroscopic characterizations.

3.8 REFERENCES

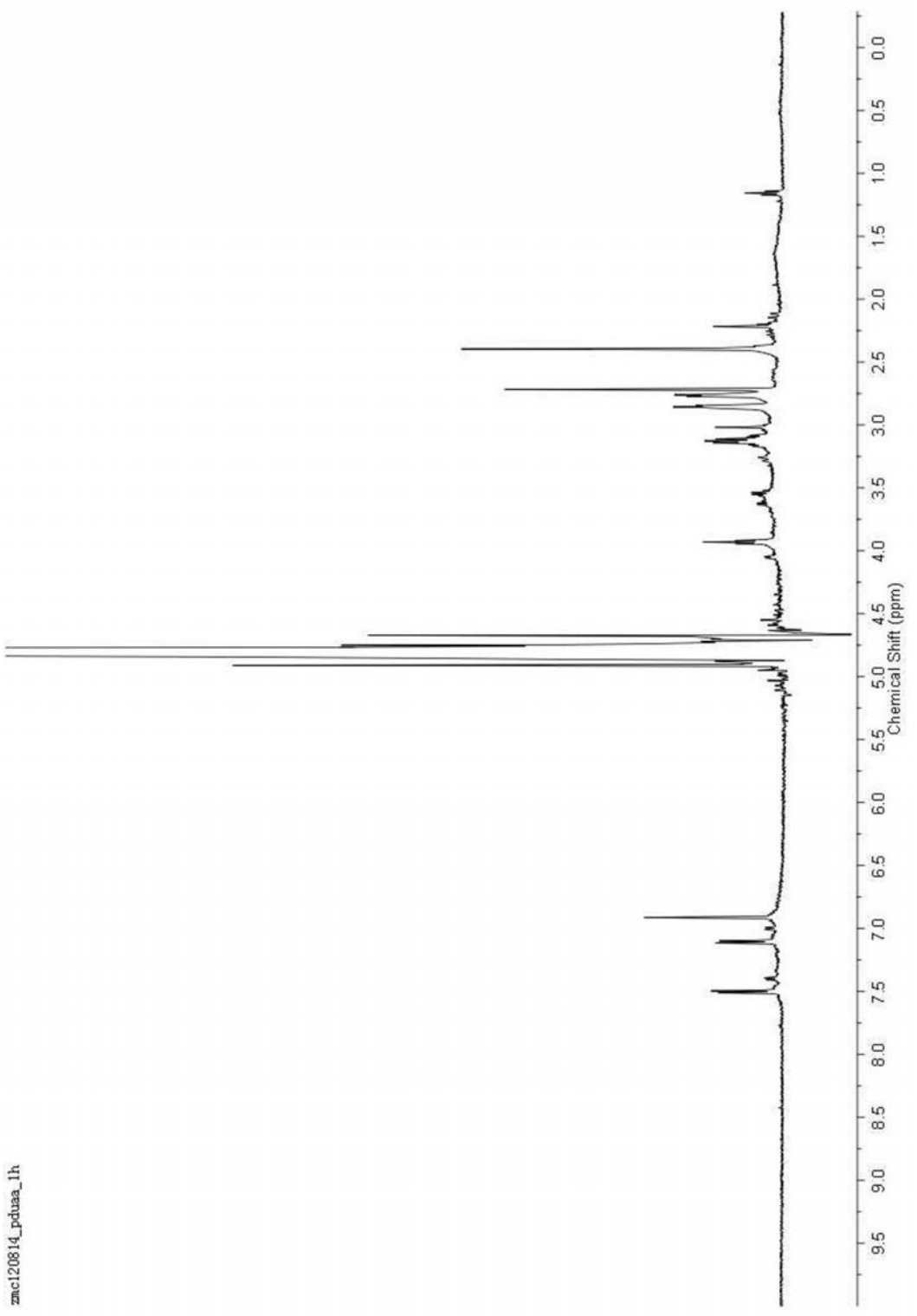
(1) Chen, X.; Engle, K. M.; Wang, D. H.; Yu, J. Q. *Angew Chem Int Ed Engl* **2009**, *48*, 5094.
(2) Lyons, T. W.; Sanford, M. S. *Chem Rev* **2010**, *110*, 1147.
(3) Jewett, J. C.; Bertozzi, C. R. *Chemical Society Reviews* **2010**, *39*, 1272.
(4) DEBETS, M. F.; BERKE, S. S. V.; DOMMERHOLT, J.; DIRKS, A. T. J.; RUTJES, F. P. J. T.; DELFT, F. L. V. *Accounts of Chemical Research* **2011**, *44*, 805.
(5) Espino, C. G.; Fiori, K. W.; Kim, M.; Bois, J. D. *J Am Soc Chem* **2004**, *126*, 15378.
(6) Pelphrey, P.; Hansen, J.; Davies, H. M. L. *Chemical Science* **2010**, *1*, 254.
(7) Davies, H. M. L.; Beckwith, R. E. J. *Chem Rev* **2003**, *103*, 2861.
(8) Tishinov, K.; Schmidt, K.; Haussinger, D.; Gillingham, D. G. *Angew Chem Int Ed Engl* **2012**, *51*, 12000.
(9) Zhu, S.-F.; Zhou, Q.-L. *Accounts of Chemical Research* **2012**, *45*, 1365.
(10) Casanova, R.; Reichstein, T. *Helv. Chim. Acta* **1950**, *33*, 417.
(11) Salzmann, T. N.; Ratcliffe, R. W.; Christensen, B. G.; Bouffard, F. A. *J Am Soc Chem* **1980**, *102*, 6163.
(12) Xu, B.; Zhu, S. F.; Xie, X. L.; Shen, J. J.; Zhou, Q. L. *Angew Chem Int Ed Engl* **2011**, *50*, 11483.
(13) Xu, B.; Zhu, S. F.; Zuo, X. D.; Zhang, Z. C.; Zhou, Q. L. *Angew Chem Int Ed Engl* **2014**, *53*, 3913.
(14) Gillingham, D.; Fei, N. *Chem Soc Rev* **2013**, *42*, 4918.
(15) Liang, Y.; Zhou, H.; Yu, Z.-X. *J Am Soc Chem* **2009**, *131*, 17783.
(16) Wang, X.-C.; Song, X.-S.; Guo, L.-P.; Qu, D.; Xie, Z.-Z.; Verpoort, F.; Cao, J. *Organometallics* **2014**, *33*, 4042.
(17) Yang, H.; Srivastava, P.; Zhang, C.; Lewis, J. C. *Chembiochem* **2014**, *15*, 223.
(18) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664.
(19) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Taglieber, A.; Hollmann, F.; Mondiere, R. J.; Dickmann, N.; Hocker, B.; Cerrone, S.; Haeger, M. C.; Sterner, R. *Chembiochem* **2008**, *9*, 552.
(20) Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J.; Rutjes, F. P.; van Hest, J. C.; Lefeber, D. J.; Friedl, P.; van Delft, F. L. *Angew Chem Int Ed Engl* **2010**, *49*, 9422.
(21) Sambasivan, R.; Ball, Z. T. *J Am Soc Chem* **2010**, *132*, 9289.
(22) Suntharalingam, K.; White, A. J.; Vilar, R. *Inorg Chem* **2010**, *49*, 8371.
(23) Chen, H.; Tagore, R.; Das, S.; Incarvito, C.; Faller, J. W.; Crabtree, R. H.; Brudvig, G. W. *Inorg Chem* **2005**, *44*, 7661.
(24) Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C. *Nat Commun* **2015**, *6*, 7789.
(25) Polgár, L. *Cell. Mol. Life Sci.* **2002**, *59*, 349.
(26) Harwood, V. J.; Denson, J. D.; Robinson-Bidle, K. A.; Schreier, H. J. *Journal of Bacteriology* **1997**, *179*, 3613.
(27) Harris, M. N.; Madura, J. D.; Ming, L. J.; Harwood, V. J. *J Biol Chem* **2001**, *276*, 19310.
(28) Lewis, J. C. *ACS Catalysis* **2013**, *3*, 2954.
(29) Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z.; Guo, Z.; Lu, Y. *J Am Soc Chem* **2004**, *126*, 10812.
(30) Zimbron, J. M.; Heinisch, T.; Schmid, M.; Hamels, D.; Nogueira, E. S.; Schirmer, T.;

Ward, T. R. *J Am Chem Soc* **2013**, *135*, 5384.
(31) Sambasivan, R.; Zheng, W.; Burya, S. J.; Popp, B. V.; Turro, C.; Clementi, C.; Ball, Z. T. *Chemical Science* **2014**, *5*, 1401.
(32) Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. *Proc Natl Acad Sci U S A* **2006**, *103*, 5869.
(33) Payne, J. T.; Poor, C. B.; Lewis, J. C. *Angew Chem Int Ed Engl* **2015**, *54*, 4226.

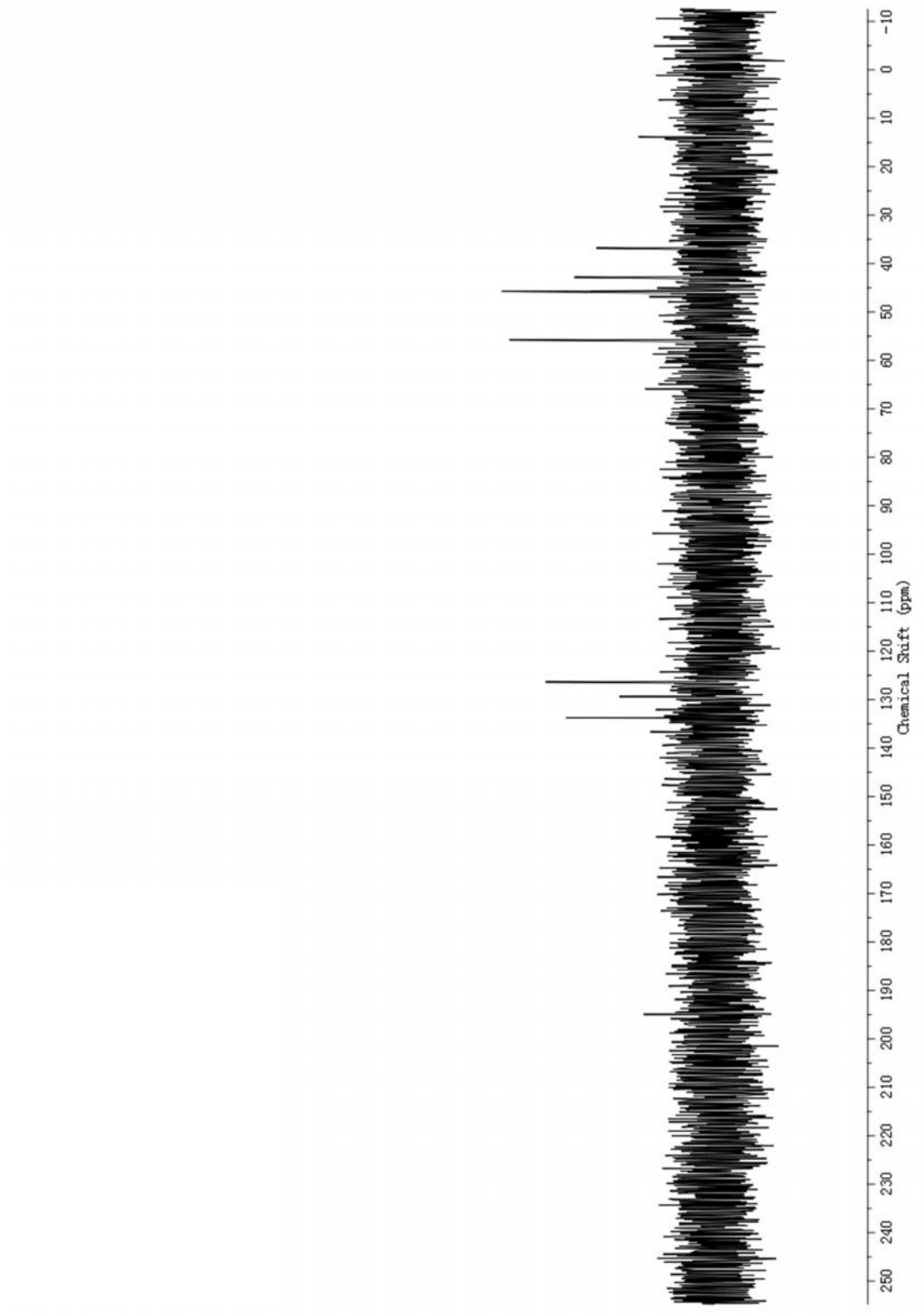
APPENDIX NMR Spectra

NMR data of previously published compounds can be found in corresponding references. This session only includes unpublished compounds that are first prepared/characterized by myself.

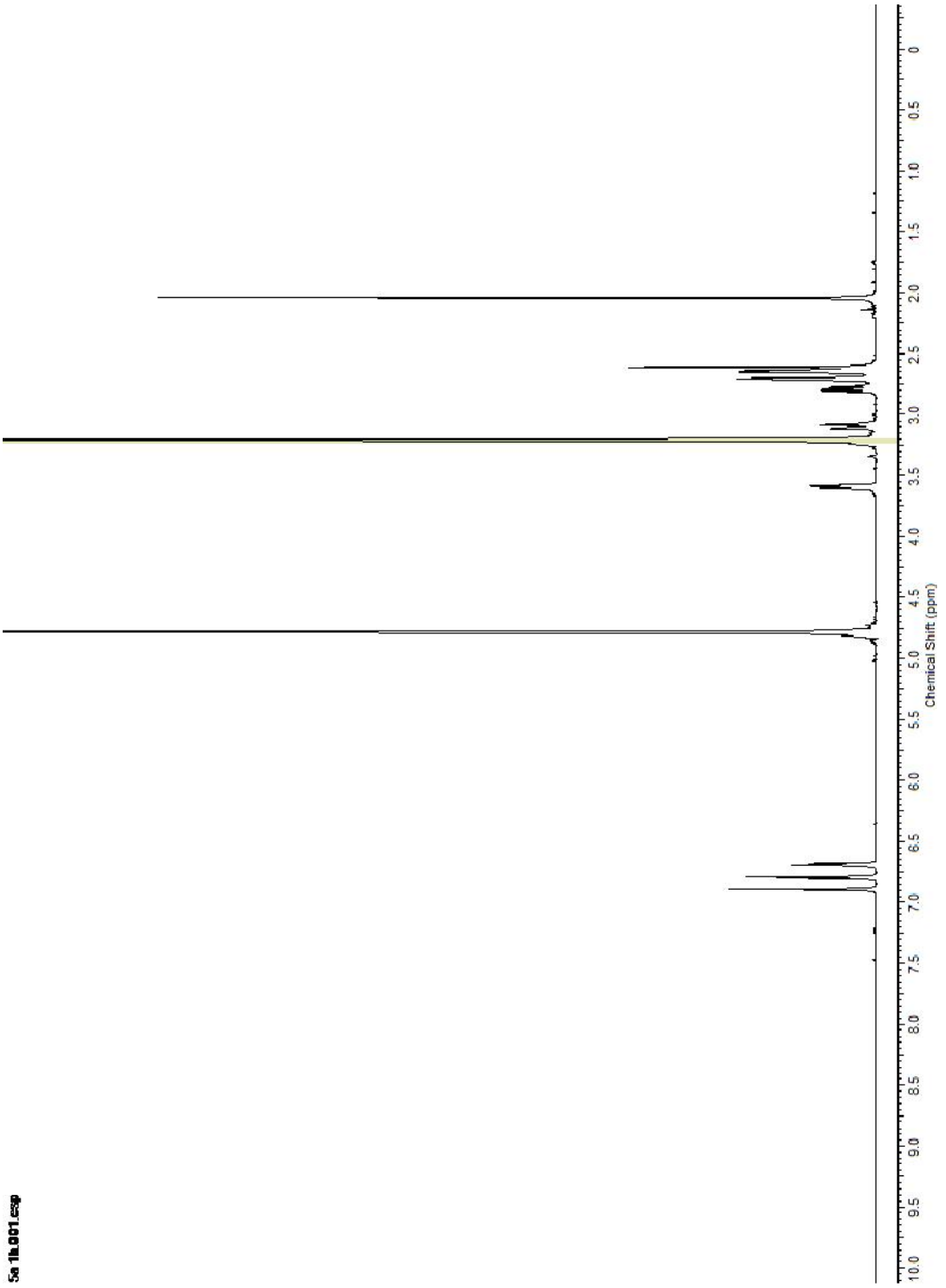
^{1}H NMR Spectrum of 5d



^{13}C NMR Spectrum of 5d

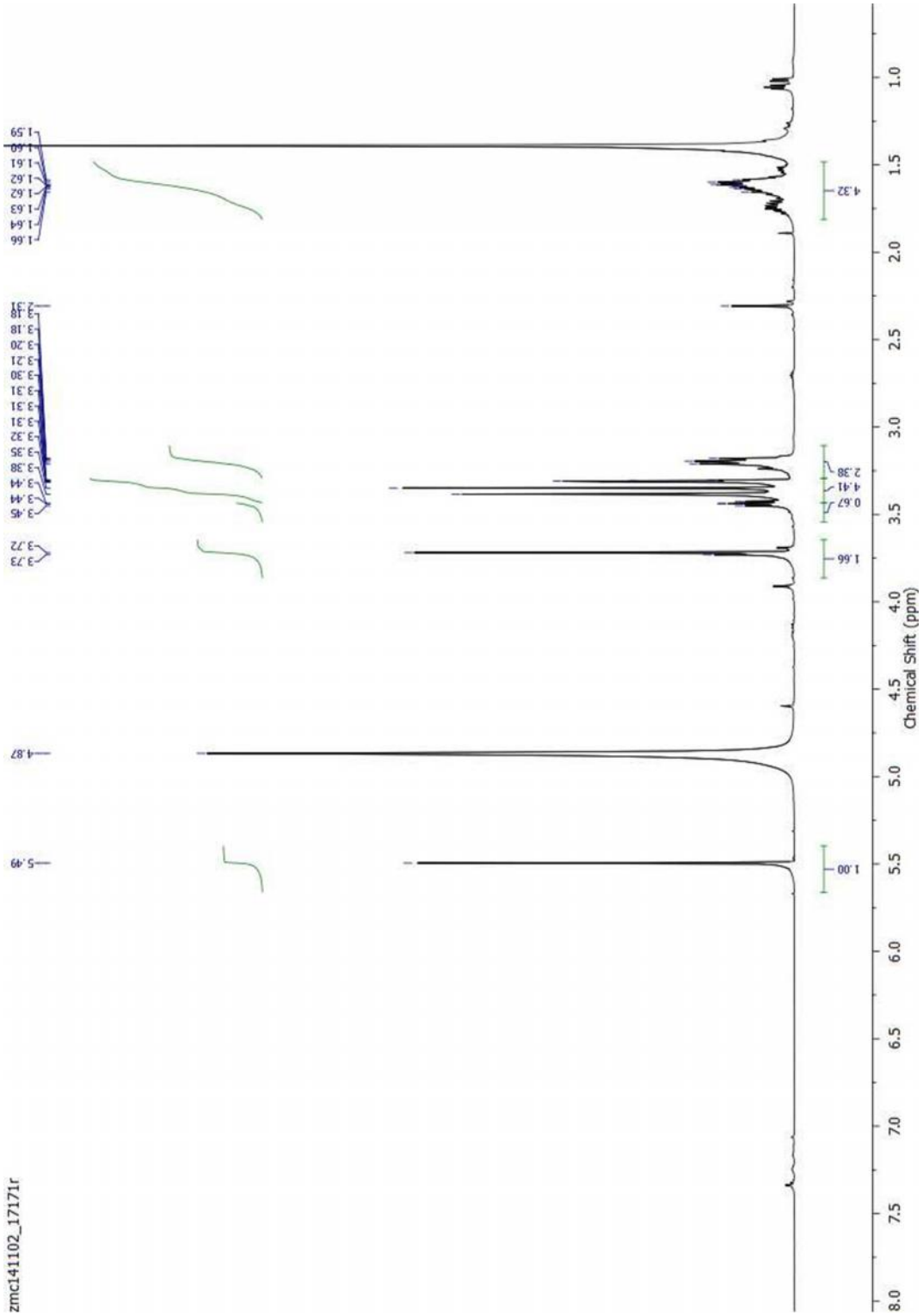


^{1}H NMR Spectrum of 5a

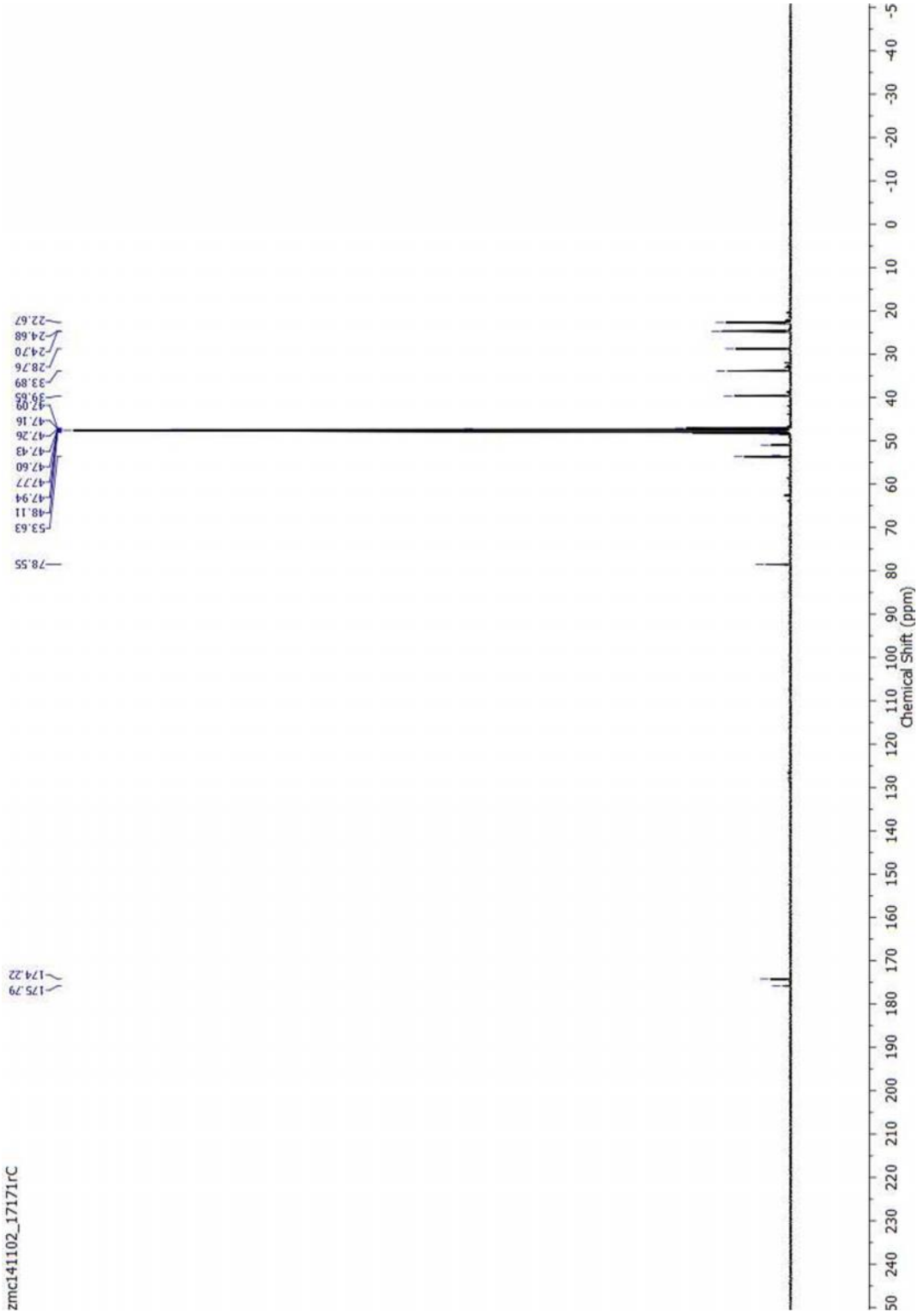


^{13}C NMR Spectrum of 5a

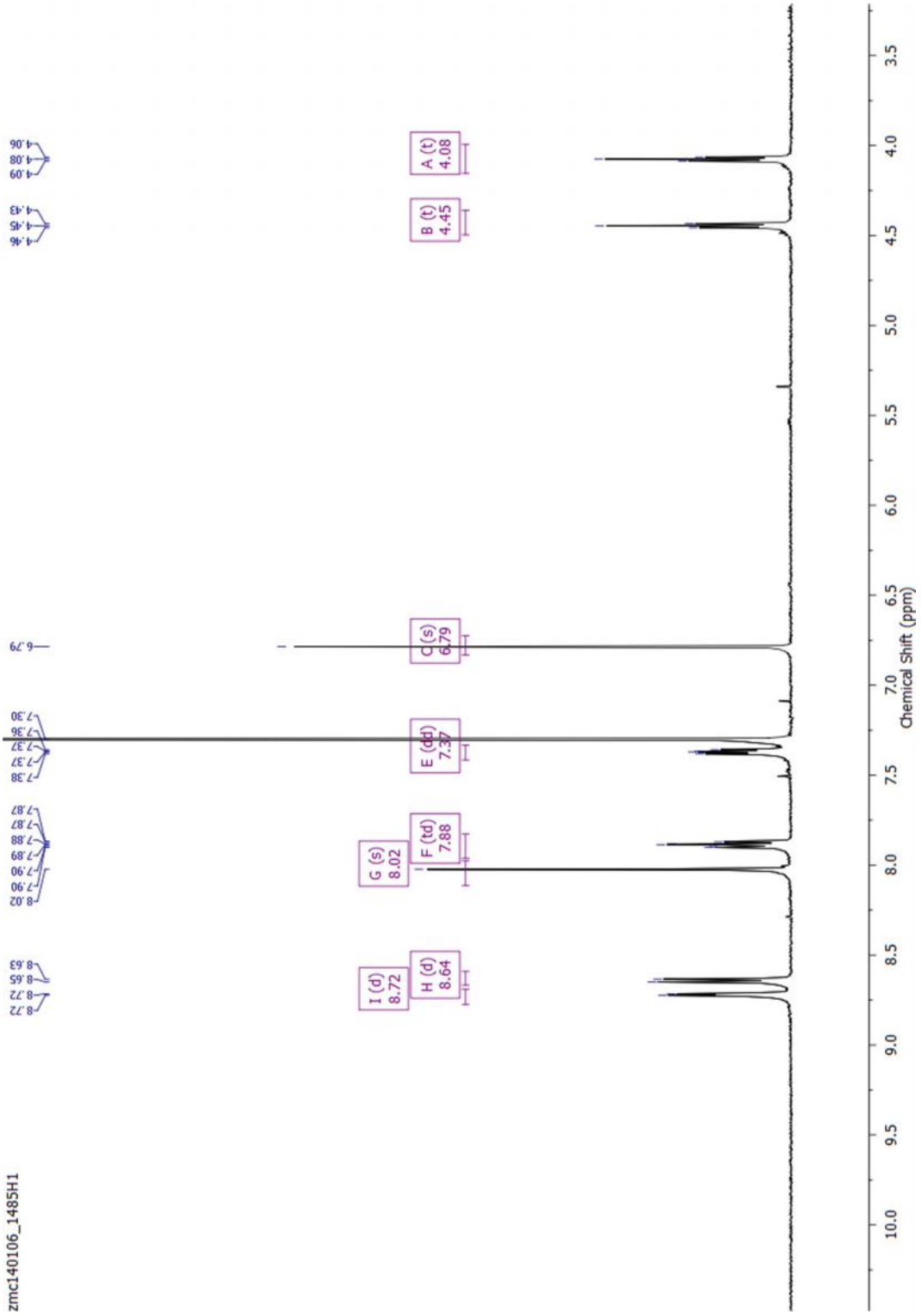
^{1}H NMR Spectrum of 28



^{13}C NMR Spectrum of 28



^{1}H NMR Spectrum of 37



^{13}C NMR Spectrum of 37