Designing artificial electron transfer pathways in dioxygen-activating metalloenzymes

by

Rajneesh Kaur Bains

B.Sc., Simon Fraser University, 2014

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Chemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Summer 2016

Approval

Name:	**Rajneesh Kaur Bains**
Degree:	**Master of Science**
Title:	***Designing artificial electron transfer pathways in dioxygen-activating metalloenzymes***
Examining Committee:	**Chair:** Dr. Michael H. Eikerling Professor

Dr. Jeffrey J. Warren
Senior Supervisor
Assistant Professor _____

Dr. Charles J. Walsby
Supervisor
Associate Professor _____

Dr. Erika Plettner
Supervisor
Professor _____

Dr. Tim Storr
Internal Examiner
Associate Professor _____

Date Defended/Approved: July 26, 2016

Abstract

This thesis describes efforts to introduce new redox reactivity into two classes of dioxygen-activating enzymes. First, I investigated modified cytochrome c peroxidase (CcP). Here, a series of Trp residues were introduced between the heme active site and the surface of the enzyme to serve as a hole transfer wire. The addition of two mutations (A193W and Y229W) introduced new oxidation chemistry to CcP, as evaluated using aromatic substrate oxidation assays. This enzyme is a functional model for lignin peroxidase enzymes and provides a strong foundation for the development of new protein-based oxidation catalysts. Second, we investigated cyanobacterial aldehyde deformylating oxygenase (cAdo) enzymes. Here, we characterized and investigated three Ru-cAdo models. To provide the four electrons required for catalysis, we introduced a Ru-tris(diimine) photosensitizer to solvent exposed cysteine residues. Through NMR and GC-MS, we gained an insight into the catalytic activity of Ru-cAdo. This work highlights the nature of protein based electron transfer and points toward other underlying factors that dictate catalytic efficiency.

Keywords: Redox enzymes, oxygen activation, electron transfer, enzyme design, biofuels, oxygen activation, ruthenium-modified enzymes, cyanobacterial aldehyde deformylating oxygenases, cytochrome *c* peroxidase

Acknowledgements

It's hard to find words to describe the impact SFU has had on my life. The last 7 years of my graduate and undergraduate experience has profoundly shaped me into the person I am today. The people in the SFU Department of Chemistry are very special to me. I am thankful for the numerous conversations I had the pleasure of sharing with members of this Department. Whether it was in the hall, at socials, or at Friday morning coffees—thank you for your words. There are acquaintances whose friendship has blossomed into lifelong friendship, and I am thankful for that. Thus, I would like to extend thanks to several individuals that have greatly shaped my graduate studies at SFU.

Firstly—I would like to thank my senior supervisor, Dr. Jeff Warren. Just over two years ago you agreed to take me on as an undergraduate research student. I have learned so much working in your lab. You challenged me to think for myself and come up with my own ideas and pursue them—even if you knew they wouldn't work. I have learned so much working in your lab, and though at times I was overwhelmed and stressed—I would not change a thing. Thank you for being your ridiculous, crazy self and for tolerating my printing on paper towels.

Secondly—I would like to thank the past and present Warren lab members, but specifically Soumyla (Susu) and Jessica. Thank you for dealing with me on a daily basis. Both of you have always supported me and encouraged me when I felt overwhelmed and stressed (which happened more often than I would like to admit…). Susu—I know I scared you at first, but I'm glad that I convinced you to sit next to me (even if it took me over a year)! Jessica—thank you for being that person I could talk to about anything, you are a great friend. Both of you are amazing people and I am grateful that I had an opportunity to work with you!

Thirdly, I would like to thank all the members of the Storr lab. I am grateful that you invited us to join group meetings these past two years. It was a stimulating and beneficial experience. To all the lab groups residing in C7076—thank you for putting up with me (especially Fatima and Thomas). To the Leznoff group, thank you for finally taking my name of the dreaded list. To the famous "Suga palace"—Y'all are great; special shout out

to Cat and Marco – the Harry Potter book club meetings are everything. I am thankful for all my friends in this department—you all mean so much to me.

Lastly—I would like to thank my parents. You mean the world to me and without your love and support I wouldn't be able to get through the past few years of study. I appreciate all the sacrifices you have made in your life which have allowed me to pursue my passions fully. I am forever grateful. To my siblings, Aman, Simar, and Rav, thanks for keeping it real. Baby Mia (Mia sounds like Rhea, right?)—I love you. #peaceout.

Table of Contents

Approval ii
Abstract iii
Acknowledgements iv
Table of Contents vi
List of Tables viii
List of Figures ix
List of Abbreviations xiii

Chapter 1. Dioxygen activation in nature 1
1.1. Oxygen in the environment 1
1.1.1. Metalloenzymes using dioxygen species 2
1.2. Dioxygen activation at heme-iron centres 3
1.2.1. Heme peroxidases 5
1.3. Non-heme iron oxygenases 6
1.3.1. Cyanobacterial aldehyde deformylating oxygenases 8
1.4. Potential applications of biological models 10
1.4.1. Biofuels 11
1.4.2. Lignin: a potential source of biofuels 12
1.4.3. Converting fatty acid pools to alkanes 14
1.5. Electron transfer in biological systems 15
1.6. Research Goals 17
1.7. References 18

Chapter 2. Mobilizing electrons in Cytochrome c peroxidase: designing and testing an electron transfer wire 21
2.1. Cytochrome c peroxidase (CcP) 21
2.1.1. Strategy 1: Introduction of lignin peroxidase activity in CcP 22
2.1.2. Strategy 2: Designing a tryptophan wire in CcP 25
2.2. Materials and Methods 28
2.3. Results 31
2.3.1. Introduction of LiP activity in CcP scaffold 31
2.3.2. Designing a tryptophan wire in C*c*P 32
2.4. Discussion 44
2.5. Conclusions 46
2.6. References 48

Chapter 3. Photocatalytic production of alkanes using Ru-modified cyanobacterial aldehyde deformylating oxygenases 51
3.1. Materials and Methods 55
3.1.1. Protein expression and purification 55
3.1.2. Characterization of Ru-cAdo Models 57
3.2. Results 60
3.2.1. Photochemical reactions 67
3.3. Discussion 72

3.4. Conclusion 73
3.5. References 74

Chapter 4. Summary, Conclusions, and Future Directions 76
4.1. References 79

Appendix A. Supporting Information for Chapter 2 80

Appendix B. Supporting Information for Chapter 3 82

List of Tables

Table 1-1. Summary of metalloenzymatic reactions with dioxygen species (O_2, O_2^- and H_2O_2) 3

Table 1-2. Examples of non heme diiron metalloproteins [22] 7

Table 2-1. ESI-MS and MALDI results of WT CcP and mutants 36

Table 2-2. Reduction potentials of tryptophan and tyrosine amino acids[]. 37

Table 2-3. Data obtained from the VA assay at a pH 8 with all CcP variants 45

Table 3-1. Theoretical and experimentally obtained mass of WT-cAdo and mutants with unlabeled and labeled protein. 62

Table 3-2. Selected fragments of digested proteins from the Trypsin solution digestion 67

List of Figures

Figure 1-1. Molecular orbital diagrams of dioxygen species. From left to right $^3\mathbf{\Sigma g}-$(triplet),$^1\Delta\mathbf{g}$ (singlet) and $^1\mathbf{\Sigma g}+$ (singlet)..... 2

Figure 1-2. Push-pull mechanism employed by heme peroxidases. PyMOL rendering shows the active site of *Saccharomyces cerevisiae* cytochrome c peroxidases (PDB: 2CYP)..... 5

Figure 1-3. Generalized heme peroxidase catalytic cycle. The general reaction catalyzed by peroxidases is $H_2O_2 + RH \rightarrow R(OH) + H_2O$ (Note: Box representative of protoporphyrin IX- Heme B)..... 6

Figure 1-4. Generalized catalytic cycle for methane monooxygenase illustrating the conversion of methane to methanol [5]. 8

Figure 1-5. The generalized catalytic mechanism of *S. elongatus* cyanobacterial aldehyde deformylating oxygenase. [20]..... 10

Figure 1-6. Common monologinols (a) p-hydroxycinnamyl alcohols (b) sinapyl p-hydroxybenzoate (c) sinapyl p-coumarate d) hydroxycinnamic acid amides [48] 13

Figure 1-7. Pymol rendering illustrating the negative microenvironment of catalytic active Trp171 residue of *Phanerochaete chrysosporium* lignin peroxidase (PDB: 1QPA) 14

Figure 1-8. Microbial biosynthetic pathway for the production of long chain alkanes in cyanobacteria 15

Figure 2-1. Schematic of the catalytic cycle of CcP. The amino acid (AA) radical (Trp191) is highlighted; Trp191 is involved with ET with the natural cofactor, Cyt *c*..... 21

Figure 2-2. Heme pocket of CcP illustrating the Trp residues in the heme pocket and Ala170. Ala170 is the proposed site of the introduced surface exposed AA radical..... 24

Figure 2-3. Tertiary structure of CcP illustrated. Insert shows the proposed Trp chain to the surface of CcP abbreviated as, CcP-WW..... 26

Figure 2-4. PyMol rendering of the CcP active site illustrating two single mutants of interest (A) displays the Tyr229Trp mutant abbreviated as CcP-W1 (B) displays the Ala193Trp mutant abbreviated as CcP-W2. 27

Figure 2-5. Coomassie stained SDS gel of MG expression of base variant. Lane: (1) ladder, (2) wild type CcP with lysis buffer, (3) wild type CcP with lysis buffer and 4.1 M urea, (4) CcP-Basic with lysis buffer + 4.1 M urea, (5) purified wild type CcP control 31

Figure 2-6. UV-visible spectra of WT CcP and mutants resting state in 20 mM sodium phosphate pH 6 buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) 33

Figure 2-7. UV-visible spectra of WT CcP and mutants Cmpd I in 20 mM sodium phosphate (pH 6) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) 34

Figure 2-8. UV-visible spectra of WT CcP and mutants coordinated to CN^- in 20 mM phosphate (pH 6) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) 34

Figure 2-9. Far UV CD spectra in 20 mM phosphate (pH 6) CD spectra purified WT CcP and Mutants. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace) and CcP-WW (red trace). 35

Figure 2-10. Monitoring the loss of the Soret peak at 424 nm upon addition of 1 equivalent of H_2O_2 in 20 mM phosphate (pH 6) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) 38

Figure 2-11. Monitoring the loss of the Soret peak at 424 nm upon addition of 1 equivalent of H_2O_2 in 20 mM phosphate (pH 8) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace). 38

Figure 2-12. Two electron oxidation of VA to corresponding aldehyde. In nature, it is postulated that the radical cation produced in Step II can then go on to oxidize large organic compounds, such as lignin. Adapted from [43]. 39

Figure 2-13. PyMol rendering of the CcP-WW active site illustrating two proposed pathways for electron hopping. (1) The red arrows illustrate two electron hopping steps to W191 (2) The green arrows illustrate the alternative one electron hopping step. 40

Figure 2-14. Oxidation assay of veratryl alcohol to veratryl aldehyde by monitoring the absorption at 310 nm. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) Reaction mixture contains 400 μM H_2O_2, 10 μM protein and 2 mM veratryl alcohol in 20 mM sodium phosphate (pH 4-top pH 6 bottom) buffer, control carried out in the absence of enzyme. 41

Figure 2-15. Oxidation assay of veratryl alcohol to veratryl aldehyde by monitoring the absorption at 310 nm. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) Reaction mixture contains 400 μM H_2O_2, 10 μM protein and 2 mM veratryl alcohol in 20 mM sodium phosphate (pH 8) buffer, control carried out in the absence of enzyme. 42

Figure 2-16. Dye decolorization of 4-p-nitrophenylazoresorcinol by monitoring the absorption at 466 nm. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace). Reaction mixtures contained 400 μM H_2O_2, 10 μM protein, and 41.8 μM 4-p-nitrophenylazoresorcinol at pH 8. 43

Figure 2-17. PyMOL rendering of the protein microenvironment of the CcP-WW mutant. Grey AAs represent the mutated residues. Surface Trp not surrounded by a negative surface environment such as LiP. 46

Figure 3-1. Cyanobacterial aldehyde deformylating oxygenase. Insert: diiron active site with ligated amino acids and aldehyde substrate outlining the substrate-binding pocket (pink). Image created in PyMOL. PDB ID: 4RC5) 52

Figure 3-2. Labeling and active site orientations in Ru-cAdo photochemical systems: (A) WT-cAdo has cysteine residue (C70) (red) near the active site; (B) introduction of two mutations R62C and C70S, results in one cysteine residue (red) adjacent to H63 in active site and (C) introduction of a single mutant at R62C produces two solvent exposed Cys (red). The magenta compound in all figures is crystallographically defined bound substrate. 53

Figure 3-3. Cys-reactive brominated Ru(tris-diimine) complex used (4-bromomethyl-4´-methyl-2,2´-bipyridyl)(2,2´-bipyridyl)$_2$Ru^{II}(PF_6)$_2$ 54

Figure 3-4. Proposed photochemical system for shuttling electrons to the active site of cAdo. Upon irradiation from a visible light source, electronically excited *Ru(II) is produced; in the presence of an excess amount of sacrificial reductant, DTC, we can reduce *Ru(II) to Ru(I), which in turn shuttle electrons to the diiron active site. 55

Figure 3-5. UV-visible spectra of apo cAdo and labeled Ru-cAdo. WT-cAdo (black trace), Ru photosensitizer- ((4-bromomethyl-4´-methyl-2,2´-bipyridyl) (2,2´-bipyridyl)$_2$Ru^{II}(PF_6)$_2$ (purple trace), Ru-WT-cAdo (red trace), Ru-R62C/C116S/C70 cAdo (blue trace). 61

Figure 3-6. Stacked MS of R62C-cAdo unlabeled (top) and Ru-R62C-cAdo (bottom) collected from MALDI. The mass shift of approximately 1200 Da is indicative of two Ru-labels attached to the surface of the protein. 62

Figure 3-7. Far UV CD spectra in 50 mM phosphate and 100 mM NaCl (pH 7.5). CD spectra of unlabeled cAdo protein. WT-cAdo (black trace), R62C-cAdo (green trace) and R62C C116S C70S-cAdo (blue trace). 63

Figure 3-8. Far UV CD spectra in 50 mM phosphate and 100 mM NaCl (pH 7.5). CD spectra of labelled cAdo protein WT-cAdo (Black trace), Ru- WT-cAdo (red trace) Ru-R62C-cAdo (green trace) and Ru-R62C C116S C70S-cAdo (blue trace). 64

Figure 3-9. Steady-state fluorescence spectra were generated with an excitation wavelength of 455 nm. The Ru-cAdo models were deoxygenated in 50 mM sodium phosphate and 100 mM NaCl buffer (pH 7.5). Ru-WT-cAdo (green trace), Ru-R62C-cAdo (red trace), and Ru-R62C C70S-C116S cAdo (blue trace). 65

Figure 3-10. Steady state emission spectra of Ru-cAdo models in deoxygenated 50 mM sodium phosphate and 100 mM NaCl, pH 7.5 (black traces). Subsequent emission recorded under identical conditions in the presence of 10 mM DTC quencher (red traces). 66

Figure 3-11. ^{1}H NMR light reaction with undecanal using Ru-WT-cAdo over 40 minutes exposure. Left insert: the growing peak at 8.3 ppm shows the formation of formate over time. Right: insert: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time. 69

Figure 3-12. The putative free radical mechanism for cAdo; the black arrows represent the productive pathway; red arrows represent the unproductive pathway. [24]..... 71

List of Abbreviations

AA	Amino Acid
cAdo	Cyanobacterial aldehyde deformylating oxygenase
CcP	Yeast cytochrome *c* peroxidase
CD	Circular dichroism
Cmpd I	Compound I
CYPs	Cytochrome P450
Cyt	Cytochrome
DEAE	Diethylaminoethanol
DTC	Sodium diethyldicarbamate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EPR	Electron paramagnetic resonance
ET	Electron transfer
GC	Gas chromatography
His tag	Poly histidine tag
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria Broth
LiP	Lignin peroxidase
MALDI	Matrix assisted laser desorption/ionization
MMO	Methane monooxygenase
MS	Mass spectroscopy
MW	Molecular weight
NHE	Normal hydrogen electrode
NMR	Nuclear magnetic resonance
NTA	Nitrlotriacetic acid
OD	Optical density
$OPPH_3$	Triphenylphosphine oxide
PD	Parallel displaced
PMS	Phenazine methosulfate
PMSF	Phenylmethanesulfonyl fluoride
PPH_3	Triphenylphosphine

SDS	Sodium dodecyl sulfate
TB	Terrific broth
TMs	Transition metals
TO	Turnover
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultra-violet
VA	Veratryl alcohol, 3,4-dimethoxybenzyl alcohol
VAd	Veratraldehyde 3,4-dimethoxybenzaldehyde
W•	Trptophanyl radical
W1	Y229W mutant of CcP
W2	A193W mutant of CcP
WW	A193W + Y229W double mutant of CcP

Chapter 1.

Dioxygen activation in nature

1.1. Oxygen in the environment

The Earth's history spans 4.5 billon years and about 2.4 billion years ago, the planet's landscape was forever changed by the 'great oxidation event' [1]. During this event, the primitive reducing atmosphere was transformed to an oxidizing atmosphere, rich in dioxygen (O_2). Since then, dioxygen has become incredibly important for sustaining life on Earth. The four-electron interconversion of O_2 and H_2O is frequently seen in nature (e.g., photosynthesis and aerobic respiration). Activation of dioxygen can be thermodynamically favourable, but a kinetic barrier often prohibits its spontaneous reaction with organic substrates. The molecular orbital energy diagram for O_2 (**Figure 1-1**) shows that the ground state holds a triplet electronic configuration. In contrast, most organic substrates contain paired electrons and therefore are described as singlet species. This difference in ground state electronic configuration means that the spontaneous reaction of singlet organic substrates and triplet dioxygen is formally spin forbidden and therefore not spontaneous. The question then remains: how is nature capable of catalyzing desirable oxidative chemistry using O_2?

Nature's incorporation of environmentally abundant transition metals (TMs) into biological systems has resulted in evolution of a great many enzymes to drive myriad natural processes that make use of O_2 [2]. The terrestrial abundance of first row transition metals, such as manganese, iron, and copper, likely prompted their evolutionary incorporation into biological systems during, and perhaps before, the great oxidation event [3].TMs are ideal candidates for the activation of dioxygen. The notable features of TMs that make them desirable for dioxygen activation include: (1) access to oxidation states with unpaired electrons that can react with O_2, and (2) the potential to access excited states that also can react with O_2 [4].

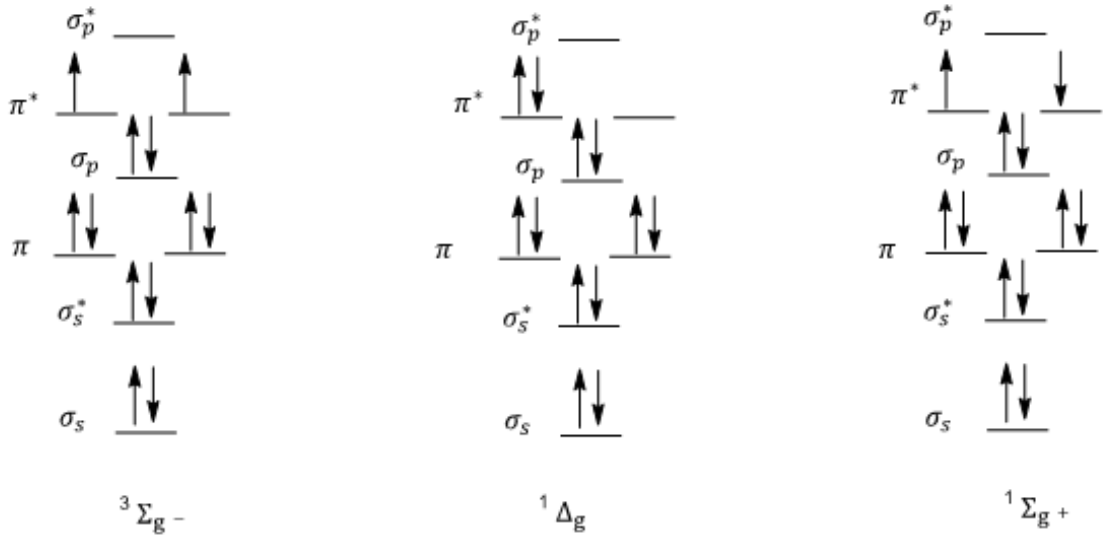

Figure 1-1. Molecular orbital diagrams of dioxygen species. From left to right $^3\Sigma_{g^-}$(triplet),$^1\Delta_g$ (singlet) and $^1\Sigma_{g^+}$ (singlet)

1.1.1. Metalloenzymes using dioxygen species

Evolutionary selection has done an exquisite job of screening for enzymes that incorporate TMs to carry out desirable chemical transformations. By varying the amino acids (AA) that ligate metals and the surrounding (secondary) environment, metalloenzymes are able to fine-tune active sites to carry out a range of chemical transformations. Several general examples are set out in **Table 1-1**. A considerable research effort has gone in to understanding how enzymes activate O_2 to carry out oxidative chemistry [5]. Advances in our understanding of O_2 activation in biological systems can be attributed, in part, to the availability, growth, and refinement of biophysical techniques. A few examples of biophysical techniques include X-ray crystallography [6], time-resolved resonance raman spectroscopy [7], Mössbauer spectroscopy [8], magnetic resonance [9] and stopped flow experiments [10]. Beyond developing a detailed understanding of how O_2 activating enzymes function and malfunction, we can also begin to envision this knowledge as a foundation for building biological catalysts that carry out oxidative chemistry. Desirable applications could include environmental remediation, fine chemical syntheses, and fuel production.

The work described in this thesis is focused on iron-containing enzymes. The geosphere is abundant with iron and it is incorporated into many enzymes that bind and activate O_2. Attractive features that underlie nature's use of iron also include the ability to easily access multiple oxidation states required for catalysis, the ability to be ligated by several AAs and biological ligands (e.g., His, Tyr, Asp/Glu, Cys, Met, and porphyrins), and kinetically facile redox chemistry. Many dioxygen-utilizing enzymes, (**Table 1-1)**, exploit iron as the metal centre. My research focuses on dioxygen activation by iron at heme centres (**Section 1-2 and Chapter 2**) and non-heme oxygenases (**Section 1-3 and Chapter 3**).

Table 1-1. Summary of metalloenzymatic reactions with dioxygen species (O_2, O_2^- and H_2O_2)

Class	Function
Transport	$M + O_2 \leftrightharpoons M(O_2)$ or $M(O)_2$
Oxidase	$O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$
Superoxide dismutase (SOD)	$2O_2^{\cdot -} + 2H^+ \rightarrow H_2O_2 + O_2$
Peroxidase	$H_2O_2 + 2e^- + 2H^+ \rightarrow R(OH) + R'(OH)$
Catalase	$2H_2O_2 \rightarrow O_2 + 2H_2O$
Monooxygenase	$O_2 + RH + 2e^- + 2H^+ \rightarrow R(OH) + H_2O$
Dioxygenase	
(a)Intramolecular	$O_2 + RH \rightarrow R(OH)_2/R(OOH)$
(b)Intermolecular	$O_2 + RH + R'H \rightarrow R(OH) + R'(OH)$

Adapted from reaction mechanism of mononuclear non-heme iron oxygenases [11]

1.2. Dioxygen activation at heme-iron centres

The heme porphyrin ring, protoheme IX (also known as iron protoporphyrin IX and heme B), is a ubiquitous cofactor [12]. Heme enzymes catalyze a range of chemical transformations and play essential roles in O_2 transport (i.e., in hemoglobin and myoglobin). Variations in the amino acids that make up the heme environment have significant effects on reactivity supported by a given enzyme. A considerable amount of research with respect to O_2 activation at heme centres is associated with cytochromes P450 (CYPs) [13]. CYPs have long fascinated chemists due to their ability to catalyze a wide range of transformations that incorporate an O atom from O_2 into a substrate, often into an unactivated C-H bond [14]. The active sites of CYP enzymes contain a heme

cofactor that is ligated by a Cys residue in the proximal pocket. The activation of O_2 occurs after the reduction of the ferric resting state to the ferrous state, which readily binds O_2. Delivery of an additional electron and two protons yields water and the active oxidizing intermediate, called Compound I (Cmpd I) [8,9,15,16]. Cmpd I is typically formulated as $(P^{\bullet+})Fe^{IV}=O$, where $P^{\bullet+}$ is a porphyrin radical cation. Two-electron oxidation of a substrate (e.g., hydroxylation of C-H) regenerates the ferric resting state of the enzyme. The activation of the resting state with hydrogen peroxide results in direct formation of Cmpd I from ferric enzyme (the so-called 'peroxide shunt'), in analogy to reactions of peroxidase enzymes, described in the following paragraph.

Another class of enzymes that are often compared to CYPs are heme peroxidases. Peroxidases utilize peroxides (e.g., hydrogen peroxide) as co-substrates to carry out oxidative chemistry on their substrates in many different organisms. Much like CYPs, peroxidases encompass a large enzyme family that carries out different net reactions using similar catalytic intermediates. Perhaps the most notable difference between CYPs and peroxidases is the identity of the proximal ligand; while heme in CYPs have Cys thiolate in the proximal position, peroxidases have His (often hydrogen bonded to Asp**,** **Figure 1-2**).

Discrete differences in heme peroxidases in relation to CYPs are revealed in analyses of the active oxidants. In heme peroxidases, Cmpd I is a strong oxidant and substrates are usually activated via initial electron transfer (ET). In contrast, Cmpd I in CYP enzymes is both oxidizing and basic; this distinguishing feature is a consequence of the 'thiolate push' afforded by an anionic axial ligand [10]. The basicity of Cmpd I in CYP enzymes means that substrates are typically activated by hydrogen atom ($H^+ + e^-$) abstraction [2,10].

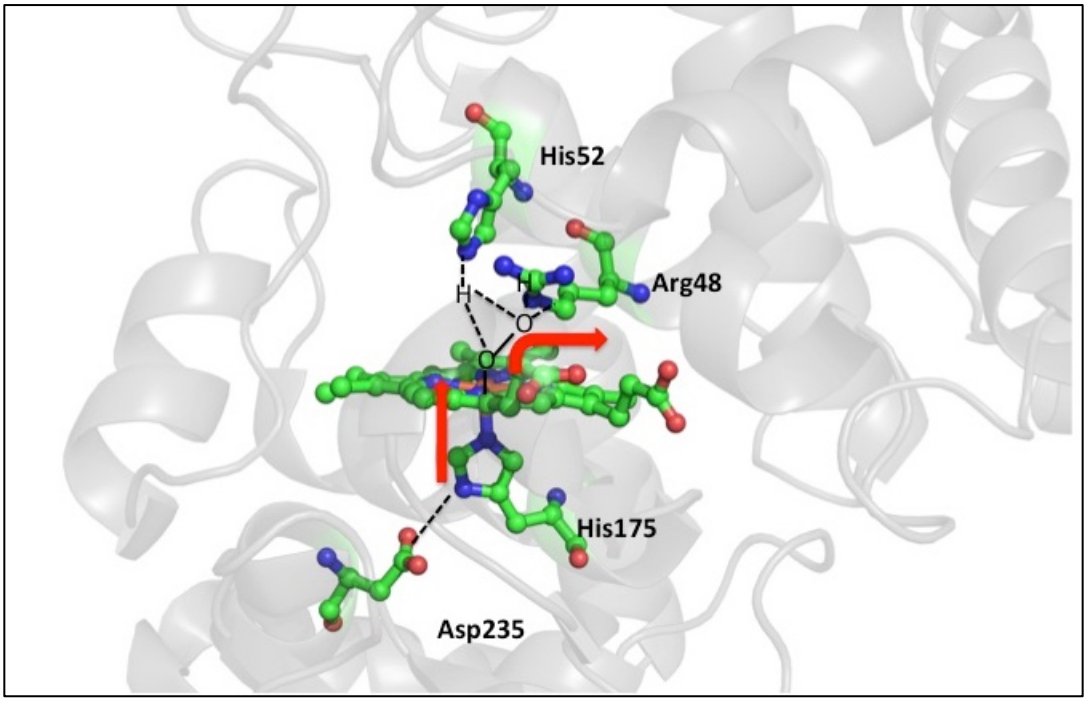

Figure 1-2. Push-pull mechanism employed by heme peroxidases. PyMOL rendering shows the active site of *Saccharomyces cerevisiae* cytochrome c peroxidases (PDB: 2CYP)

1.2.1. Heme peroxidases

Heme peroxidases are found in diverse biological systems. These enzymes typically catalyze the reactions outlined in **(Table 1-1**), using hydrogen peroxide (H_2O_2) as the oxidant. A generic peroxidase catalytic cycle is illustrated in **Figure 1-3**. Catalysis is initiated upon binding of H_2O_2 to the ferric high spin metal centre (labeled 'resting state'). The AAs in the heme's secondary coordination sphere play an important role in the formation of Cmpd I. A 'push-pull' mechanism is commonly described for heme peroxidases **(Figure 1-2)** [17]. In this mechanism, the proximally ligated His175 exhibits enhanced electron-donating properties due to the strong hydrogen bond with a nearby Asp residue. His175 provides a 'push' of electron density toward the ligated peroxide, while the distal pocket provides the 'pull'. The distal pocket contains a conserved His (His52) that acts as a general acid-base catalyst as proposed by Poulos and Kraut [18]. In their mechanism, His52 initially deprotonates the peroxide to form Cmpd 0 and His52-H^+ (**Figure 1-3**). His52-H^+ then serves as an acid, to release water and produce Cmpd I. The oxidizing equivalents are delocalized onto the heme cofactor, and substrates in close

proximity to the heme edge can be oxidized. Following a one-electron reduction, the enzyme forms Cmpd II, and a subsequent one-electron reduction returns the enzyme to the resting state. Following this general catalytic cycle, peroxidases are capable of carrying out a wide array of oxidative chemistry, from oxidation of cytochrome *c* (by cytochrome c peroxidase) to oxidation of Cl^- (by myeloperoxidase) [19].

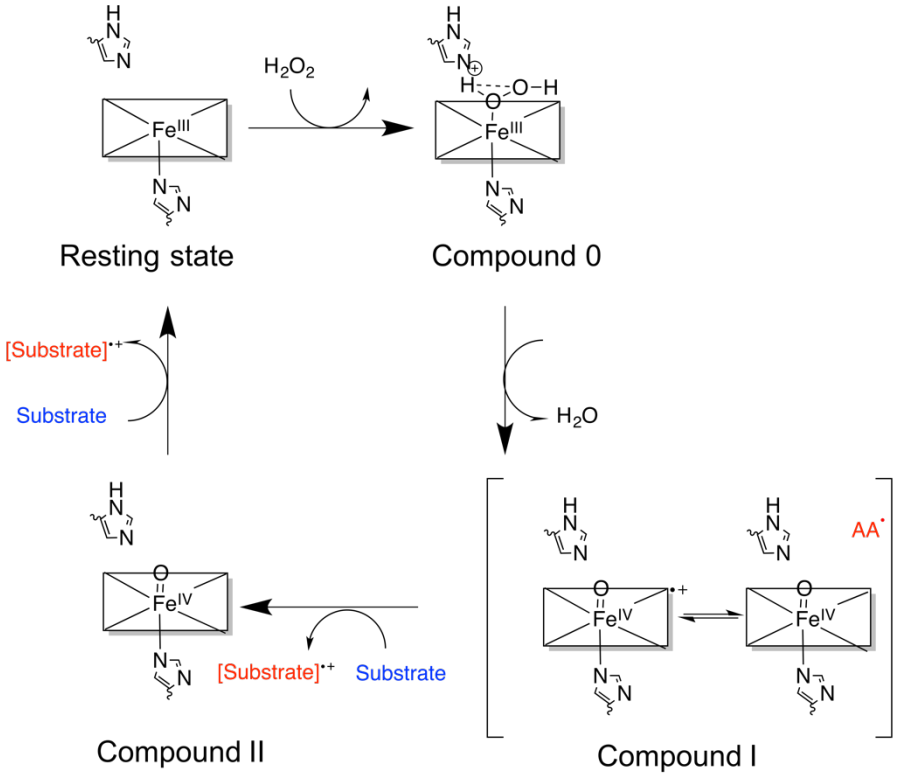

Figure 1-3. Generalized heme peroxidase catalytic cycle. The general reaction catalyzed by peroxidases is H_2O_2 + RH → R(OH) + H_2O (Note: Box representative of protoporphyrin IX- Heme B).

1.3. Non-heme iron oxygenases

Non-heme oxygenases comprise of a large family of enzymes that are capable of binding and activating dioxygen without using a heme cofactor. These enzymes can be divided into two broad categories: active sites containing one iron centre or active sites containing two iron centres. Although mono-iron oxygenases are plentiful in nature [20,21], diiron oxygenases are arguably more intriguing due to the many reactions carried out at the diiron core. **Table 1-2** outlines the diversity of diiron heme oxygenases in nature [22]. The family of non-heme diiron oxygenases exhibits similar active-site structures

where the iron centres are typically coordinated to two glutamates (sometimes bridging), and one histidine. These active sites are usually contained in an antiparallel 4-helix bundle [20], but the entire protein can contain several other subunits. Methane monooxygenase (MMO) is a non-heme diiron oxygenase that has received much attention because it catalyzes the kinetically difficult conversion of methane to methanol in methanotropic bacteria [23]. During the catalytic cycle, the diiron core first binds O_2 as a bridged peroxo-bound diferric state (intermediate P, **Figure 1-4**). The following oxidized intermediates are difficult to detect and work is ongoing in the field. In recent years, research has focused on understanding the catalytic cycle and these key intermediates, where the most compelling proposed intermediate is intermediate Q (**Figure 1-4**) [24]. Though much progress has been made in understanding the catalytic cycle [7] there remain many questions. Notably, many diiron oxygenases enter their catalytic cycles through an initial bridged $Fe(O)_2Fe$ intermediate. As for CYPs and peroxidases, it is remarkable that these structurally related enzymes can carry out diverse chemistry from similar intermediates.

Table 1-2. Examples of non heme diiron metalloproteins [22]

Enzyme	Function
Hemerythrin	Dioxygen carrier
Ribonucleotide reductase	Tyrosyl radical generator
Methane monooxygenase	Methane to methanol oxidation
Ferritin	Iron storage
Toluene monooxygenases	Toluene to cresol oxidation
Rubrerythrin	Putative peroxidase
DMQ monooxygenase	Quinone generation

Adapted from: Synthetic models for non-heme carboxylate-bridged diiron metalloproteins: strategies and tactics [22]

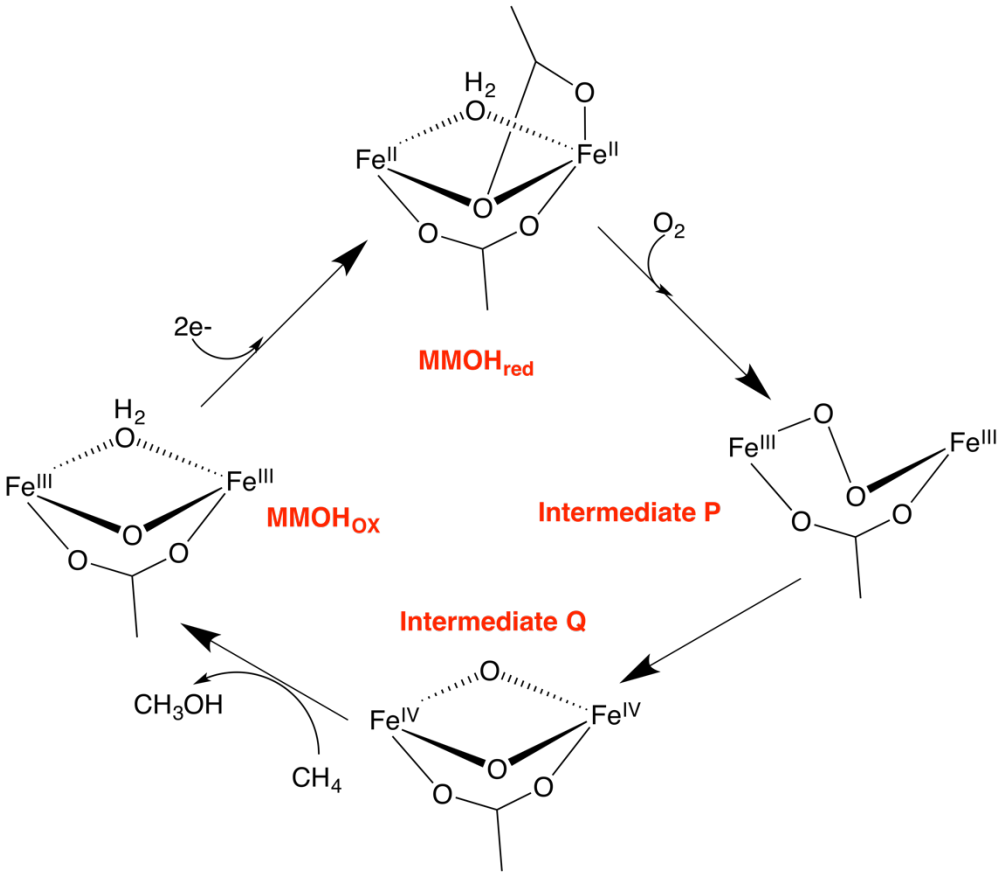

Figure 1-4. Generalized catalytic cycle for methane monooxygenase illustrating the conversion of methane to methanol [25].

1.3.1. Cyanobacterial aldehyde deformylating oxygenases

In 2010, a small biotechnology company, LS9 Inc. (San Francisco, USA), identified a microbial pathway for the production of alkanes in cyanobacteria. [26] One of the key enzymes isolated and identified was a non-heme diiron oxygenase. This enzyme was classified as cyanobacterial aldehyde deformylating oxygenase (cAdo). The primary coordination sphere and the tertiary structure closely resemble other diiron enzymes, but cAdo carries out distinct chemistry. The enzyme pocket accommodates long hydrophobic aliphatic (C_n) aldehydes (n = 15 to 17) and orients the aldehyde-oxygen near the iron site, which catalyzes the overall transformation of C_n aldehydes to the corresponding C_{n-1} alkane and formate. Notably, this redox neutral reaction is coupled to the $4H^+/4e^-$ reduction of O_2 to H_2O [32]. Though the biological function of these enzymes in cyanobacteria are not known, the net function of related enzymes in plants and insects

are known to be involved in the production of waterproof barriers and biosynthesis of pheromones [32]. The discovery of this microbial pathway sparked much interest in the scientific community due to the prospect of generating alkane fuels (gasoline or diesel) from bacteria. Since the initial discovery, several biotech companies have filed patents for microbial fuel-generating technologies [27,28].

A significant amount of effort has gone into understanding the details of the cAdo catalytic cycle. Early research reported that cAdo was an anaerobic enzyme, implying the chemical transformation of aldehyde to alkane was catalyzed independent of O_2 [29]. This observation is consistent with the redox-neutral hydration of an aldehyde to formate and an alkane. Soon thereafter, Bollinger and co-workers demonstrated that cAdo does require O_2 for function and that aldehyde deformylation is actually coupled to reduction of O_2 to H_2O [30]. Further refinements to the O_2-dependent model suggest that microaerobic conditions are best for *in vitro* studies [31]. **Figure 1-5** outlines the most recently proposed mechanism catalyzed by cAdo [32]. Binding of O_2 to the diferrous state of the protein (species I $\rightarrow$ II) initiates catalysis and produces a peroxide-bound diferric state, similar to intermediate P observed in MMO [33]. The nucleophilic addition of this species to the bound aldehyde produces peroxy-hemiacetal intermediate (intermediate III). The injection of an electron promotes the scission of the O–O bond and forms a transient formyl radical. Subsequent fragmentation of this radical produces an alkyl radical and formate (intermediate V). Finally, a proposed proton-coupled electron transfer step produces the alkane product and the diferric enzyme [34]. Native substrates for cAdo are long chain aldehydes, but these substrates have limited water solubility, making catalytic studies difficult. Consequently, Marsh and co-workers demonstrated that short chain aldehydes, with higher water solubility, are also compatible substrates [29]. However, closer investigation of reactions with short chain aldehydes reveal the formation of additional products such as n–1 aldehydes and alcohols [35]. Further discussion on cyanobacterial aldehyde deformylating oxygenases is provided in **Chapter 3**.

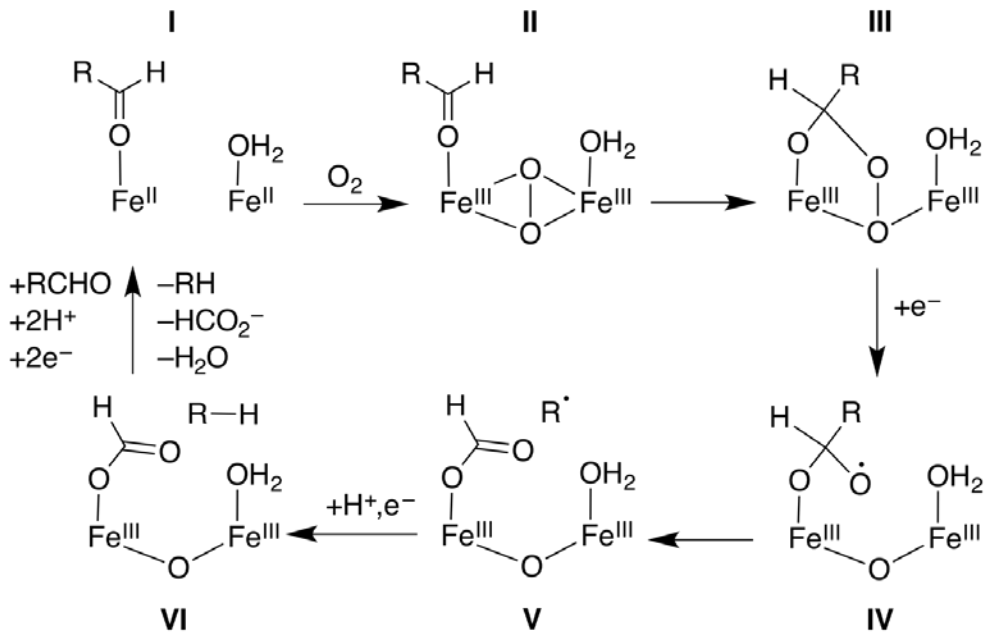

Figure 1-5. The generalized catalytic mechanism of *S. elongatus* cyanobacterial aldehyde deformylating oxygenase. [20]

1.4. Potential applications of biological models

The biotechnology industry has seen extraordinary growth over the past 30 years, which is strongly coupled to the growth and development of modern biochemical techniques (e.g., DNA sequencing and synthesis). As a result, a great many enzymes have been discovered, isolated, and studied in detail. Many enzymes have evolved to carry out difficult chemical transformations needed in natural biosynthetic and metabolic pathways. As a consequence, there is increased interest in exploiting and manipulating these biological systems for practical industrial applications.

Biologically-based systems are advantageous to synthetic small-molecule catalysts due to their inherently good regio-,and stereoselectivity, activity under ambient reaction conditions, and their use of abundant first row TMs [36]. However, there also are some critical drawbacks to using biomolecules as catalysts. First, many enzymes are sensitive to the external environment, such as changes in temperature and/or pH. Targeting robust enzymes or incorporating beneficial mutations (e.g., via directed evolution) to the protein can circumvent these issues [37]. Second, enzymes are evolved to function in water, so integrating these enzymes into industrial systems, where organic

solvents are more prevalent, can be problematic. Mixed solvents, and two-phase solvent systems have proved to be effective methods to avoiding protein denaturation [38], but more work is needed.

In recent years, several examples have been reported where companies have successfully incorporated enzymes into their existing processes. For example, Merck and Codexis, described an enzyme-catalyzed synthesis that improved efficiency of the diabetes drug, Januvia™ (sitagliptin) [39]. Enzyme stability was enhanced using several rounds of directed evolution, allowing it be catalytically active in a mixed solvent system [39]. It is evident that industry has real interest in developing biological systems for industrial processes, especially fine chemical syntheses and biofuel production [27,28]. The primary focus of the research in this thesis was on exploiting two different dioxygen activating metalloenzymes to generate biofuels.

1.4.1. Biofuels

Over the years, there has been increased pressure to search for alternatives to petroleum reservoirs. As energy consumption continues to grow, petroleum reservoirs are continuing to be depleted at a similar rate [40]. The dependence on petroleum fuel sources has resulted in a significant increase in major air pollutants (e.g., CO_2, SO_2, and NO_X). Many argue that an attractive alternative to fossil fuels could include biofuels, which are liquid and/or gaseous fuels derived from biomass. These fuels are attractive as they can readily be integrated into the existing infrastructure, specifically the transportation sector. There are several inherent properties that make biofuels attractive as fuel sources. First, biofuels are derived from renewable resources. Second, there is no net release of carbon dioxide (as biofuels are derived from renewable resources), and sulphur content remains low (somewhat mitigating air pollution) [41]. Finally, biofuels show significant economic potential since petroleum fuel costs are anticipated to rise in coming years. A thorough comparison revealed that energy sources derived from biomass are notably cheaper than some commonly used petroleum sources [42]. It is important to note that though biofuels present an attractive alternative, other fuel sources (e.g., hydrogen gas) should be explored which have little, or no release of air pollutants.

A great deal of research in the past decade has focused on discovering, and/or improving methods for biofuel production [43]. Biopolymers (cellulose and lignin) remain a focus in biofuel production due to the high abundance and low cost. As discussed above, enzymatic pathways for biofuel production are attractive due to the ambient reaction conditions. Recently, there has been increased interest in exploiting cellulases to allow for the enzymatic hydrolyzing of the rigid cellulosic biopolymer [44]. However, improvements to catalytic activity and thermostability need to explored before we begin to see widespread industrial use. In addition, lignocellulose, the more recalcitrant biopolymer in biomass, can be synergistically hydrolyzed by three classes of enzymes which include (1) endo-β-1,4-glucanases, (2) β −glucosidases, and (3) hemicellulases and ligninases. However, the high cost of these enzymatic pathways along with large quantities required allow large room for improvement [45].

1.4.2. Lignin: a potential source of biofuels

Though much focus has been given to biofuel production from cellulosic biomass, lignin biopolymers have emerged as attractive complementary (or alternative) starting materials for production of biofuels [46]. Lignin is an aromatic biopolymer with a highly variable structure and is the source of the rigidity of plant matter [47]. Lignin is primarily composed of different phenolic subunits named monolignols. Three common monolignol subunits found in lignin are p-hydroxphenyl-, guaiacyl-, and syringyl-monolignols (**Figure 1-6**) [48]. The key challenge in using lignin in the production of biofuels is finding enzymes that break down the highly heterogeneous lignin structures into a defined product set.

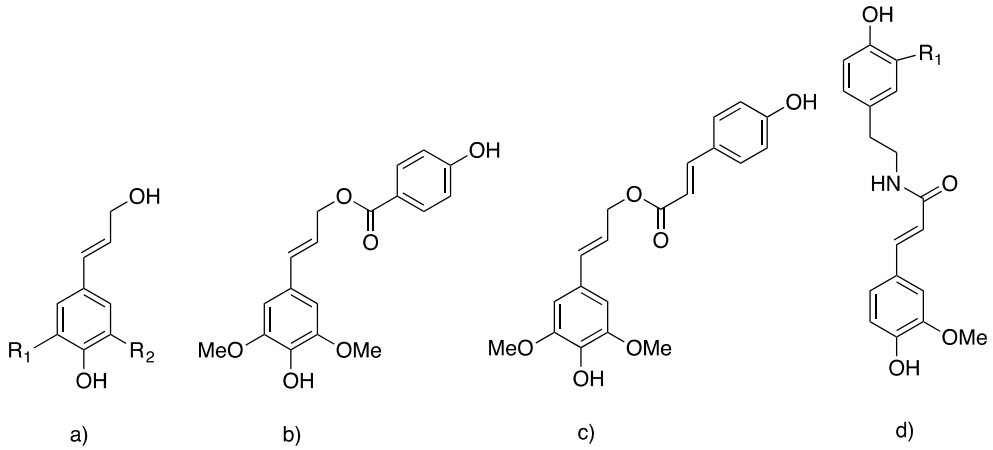

Figure 1-6. Common monologinols (a) p-hydroxycinnamyl alcohols (b) sinapyl p-hydroxybenzoate (c) sinapyl p-coumarate d) hydroxycinnamic acid amides [48]

In 2004, the genome of the lignin-degrading white rot fungus *Phanerochaete chrysosporium* was sequenced [49]. The sequencing of *P. chrysosporium* identified secreted oxidases, peroxidases, and hydrolytic enzymes involved in lignin degradation. This stimulated research in using these enzymatic pathways in lignin depolymerization. Lignin peroxidase (LiP) was among one of the fungal enzymes that gathered much interest in the research community [50]. The function and mechanism of this enzyme puzzled researchers: how was LiP capable of oxidizing high reduction potential substrates such as lignin using an embedded heme cofactor? Subsequent experimental work revealed the importance of Trp171 on the surface of the protein [51]. The catalytic Trp171 is involved with depolymerization of lignin by forming an tryptophanyl radical (W•+) (**Figure 1-7**). The Trp171 microenvironment, which is rich in carboxylate-containing residues, is believed to stabilize the Trp radical and aid in the oxidation veratryl alcohol, which act as diffusible mediators. The oxidized veratryl alcohol is thought to carry oxidizing equivalents towards the bulky lignin biopolymer for depolymerization. These enzymes are industrially attractive due to the ease of breaking down recalcitrant biomaterials, but widespread use of LiP is not yet practical due to very low expression yields in heterologous systems (~ 1 mg per liter of culture) [52]. Recent work has focused on mimicking the surface exposed Trp radical in related enzymes. The successful incorporation of the Trp radical analogous to LiP was carried out in *Corprinus cinereus* peroxidase, with several mutations [53]. We

adopted a similar strategy to introduce LiP reactivity into a more robust scaffold, which will be discussed further in **Chapter 2**.

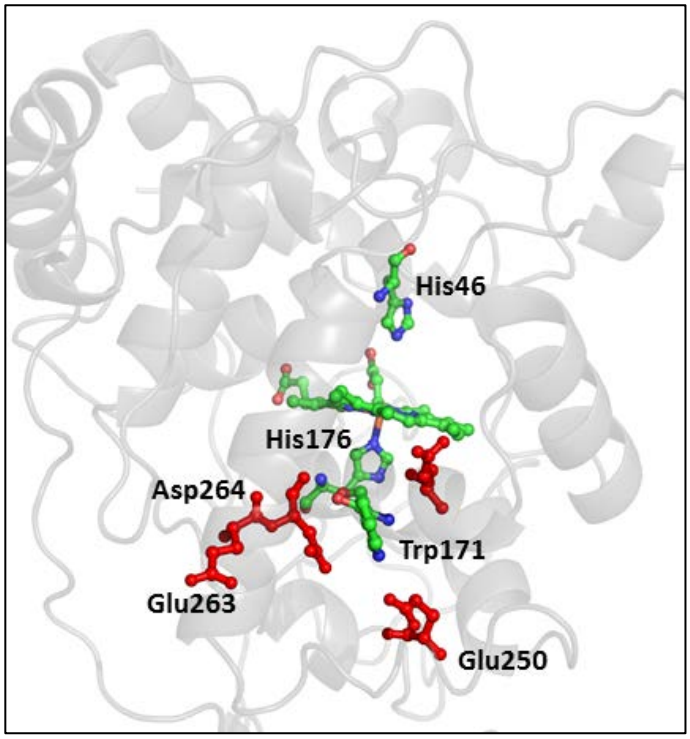

Figure 1-7. Pymol rendering illustrating the negative microenvironment of catalytic active Trp171 residue of *Phanerochaete chrysosporium* lignin peroxidase (PDB: 1QPA)

1.4.3. Converting fatty acid pools to alkanes

A great deal of research exploring routes for production of biofuels has focused on pathways that produce ethanol and/or butanol [54,55,56]. These are still attractive goals, but a recent shift in research has been to explore other products (e.g., alkanes) that are more facile 'drop in' replacements for fossil fuels. To that end, the recent discovery of a cyanobacterial alkane biosynthetic pathway stimulated interest in producing alkanes using microorganisms [57]. This novel pathway has two key enzymes that are involved in the conversion of fatty acids to alkanes, as outlined in **Figure 1.8**. The two-gene pathway includes two distinct enzymes, acyl-ACP reductase, and aldehyde deformylating oxygenase [58]. The conversion of fatty acid pools in *Escherichia coli* (*E. coli*) was demonstrated in organisms bearing plasmids for both the acyl-ACP reductase and the fatty aldehyde decarbonylase. The yields from wild-type organisms are modest, analogous

to that seen in mechanistic and catalytic studies. In later work, a metabolic pathway was engineered for alkane biosynthesis of propane [57]. This new approach integrated catalase into the microbial pathways. The catalase scavenges hydrogen peroxide produced during the alkane formation. An increase in alkane production was observed, but yields were still modest compared to the scale needed for full industrial use. Much work still remains on optimizing reaction conditions for the terminal conversion of aldehyde to alkanes. A major issue that still needs to be addressed is the potential inefficiency of ET reactions that are required in the catalytic cycle. My work focuses on exploring the efficiency of ET in the catalytic cycle of cAdo, which could then increase the efficiency of alkane production. This work is discussed further in **Chapter.3**.

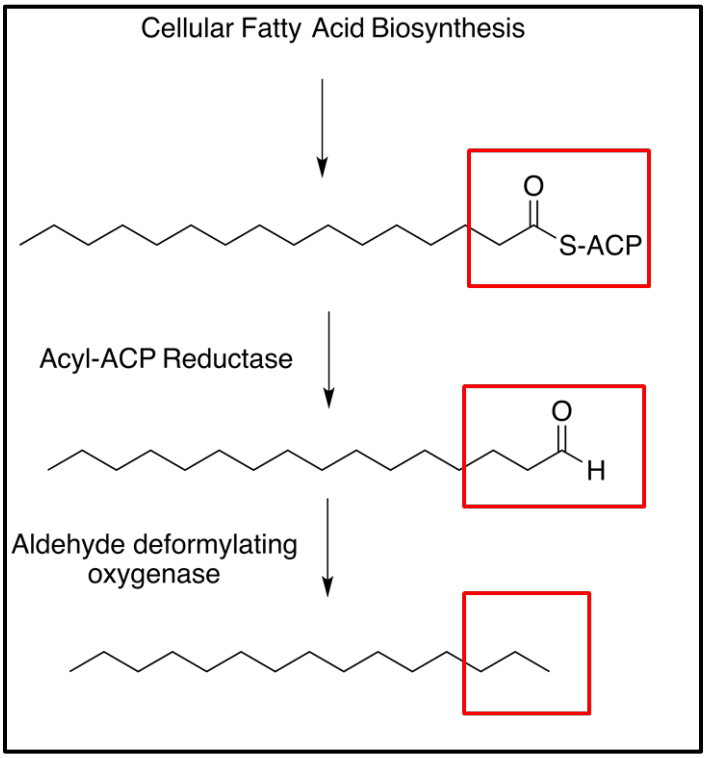

Figure 1-8. Microbial biosynthetic pathway for the production of long chain alkanes in cyanobacteria

1.5. Electron transfer in biological systems

Nature has fine-tuned biological systems to carry out many important ET reactions. Though ET in biological systems is often complex, each discrete step in a large system can be related back to fundamentals of ET. In 1956, Marcus published a paper describing

the relation of the rates of ET reaction between a single donor (A) and acceptor (B) to the thermodynamics of that system [59]. Typically, the expression shown in **(Eqn. 1)** can be used to describe the rate of single step ET between a donor and acceptor, where $\Delta G°$ is the driving force, H_{AB} is electronic coupling between the donor and acceptor, λ is reorganization energy, k_b is Boltzmann's constant, and h is Planck's constant. The driving force ($-\Delta G°$) of the reaction is described through Gibbs free energy shown in **(Eqn. 2).** The λ for biological systems contains two components (1) the energy associated with changes in the inner sphere which include the primary coordination sphere at the redox centre, and (2) the energy associated with changes to outer coordination sphere which include the environment surrounding the redox centre. Finally, H_{AB} is the electronic coupling matrix element that describes interactions between two redox partners. It is a function of distance and exponentially decreases as the distance between the donor and acceptor increases. Electronic coupling between a donor and acceptor is described with **(Eqn. 3),** where β is the decay constant for tunneling, r is the distance between the D and A and r_0 is the limit of close contact [60]. As shown **(Eqn. 1),** the three factors that greatly control ET are $\Delta G°$, λ, and H_{AB} and therefore changes to any one of these variables can greatly influence rates of ET.

$$k_{ET} = \sqrt{\left(\frac{(4\pi^3)}{h^2 \lambda k_b T}\right)} H_{AB}^2 \exp\left(\frac{-(\Delta G° + \lambda)^2}{4 h k_b T}\right) \textbf{ (Eqn. 1)}$$

$$\Delta G^o = -RTln\left(K_{eq}\right) = -nFE^{\circ} \textbf{ (Eqn. 2)}$$

$$H_{AB} = H_{AB}(r_0) * \exp\left(-0.5\beta(r - r_0)\right) \textbf{ (Eqn. 3)}$$

The expression described in **Eqn. 1** is for single step ET between a donor and acceptor. However, in most biological systems, we see electron 'hopping' where electrons (or holes) are transferred between distant cofactors via several intermediate steps. Breaking down ET into such small steps can often make ET transfer occur at a faster rate due to the exponential distance dependence of ET reactions (**Eqn. 3**). However, multi-step ET further complicates the ET rate expression shown in **(Eqn. 1).** An extensive discussion of the theoretical and functional consequences of electron hopping within proteins is reviewed in Ref. [61]. In this thesis, I will investigate ET in two different dioxygen activating enzymes. In Chapter 2, we will investigate ET hopping through several Trp

residues. In Chapter 3, I will investigate single step ET to the active site through the addition of an Ru-photosensitizer. The unifying theme of these projects is to introduce new ET pathways that are designed to increase rates of ET, and therefore support new catalytic activity, by modifying driving forces ($-\Delta G°$) and electronic coupling (H_{AB})

1.6. Research Goals

Dioxygen activating enzymes have fascinated biochemists for decades. As our understanding of these enzymes has developed, we can shift our focus toward exploiting them for useful applications. In this thesis, I will present two approaches to exploiting dioxygen-activating enzymes to produce biofuels. First, I will focus on how to harness highly oxidizing intermediates and manipulate them for the degradation of large stable biopolymers. Second, I will describe methods to improve electron transfer within non-heme oxygenases to produce alkanes. This thesis concludes with a chapter describing the broader impacts of these approaches and my thoughts on future directions for both of the projects.

1.7. References

1. Lyons, T. W.; Reinhard, C. T.; Planavsky, N. J. Nature. 2014, 506, 307–315
2. Que, L.; Tolman, W. B. *Nature*. **2008**, *455*, 333–340
3. Rosenzweig, A. C.; Sazinsky, M. H. *Curr. Opin. Struct. Biol.* **2006,** *16***,** 729–735.
4. Jensen, K. P.; Ryde, U. *J. Biol. Chem*. **2004**, *279*, 14561–14569
5. Klinman, J. P. *Acc. Chem. Res.* **2007**,*40*, 325–333.
6. Poulos, T. L.; Freer, S. T.; Alden, R. A.; Edwards, S. L.; Skoglund, U.; Takio, K.; Xuong, N.-H.; Yonetani, T.; Kraut, J. *J. Biol. Chem.* **1980,** *259***,** 575–580.
7. Banerjee, R.; Proshlyakov, Y.; Lipscomb, J. D.; Proshlyakov, D. A. *Nature*. **2015**, *518*, 431–434.
8. Pandelia, M. E.; Li, N.; Nørgaard, H.; Warui, D. M.; Rajakovich, L. J.; Chang, W.-C.; Booker, S. J.; Krebs, C.; Bollinger, J. M. *J. Am. Chem. Soc.* **2013**, *135*, 15801–15812.
9. Volkov, A. N.; Worrall, J. A. R.; Holtzmann, E.; Ubbink, M. Proc. Natl. Acad. Sci. U.S.A. **2006,** *103*, 18945–18950.
10. Rittle, J.; Green, M. T. *Science*. **2010**,*330*, 933–937.
11. Abu-Omar, M. M.; Loaiza, A.; Hontzeas, N. *Chem. Rev*. **2005**, *105*, 2227–2252
12. Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev*. **1996,** *96* 2841–2888.
13. Montellano, P. R. *Cytochrome P450 Structure, Mechanism, and Biochemistry*; 3rd ed.; Kluwer Academic/Plenum Publishers: New York, 2005.
14. Bell, S. R.; Groves, J. T. *J. Am. Chem. Soc.* **2009**, *131*, 9640–9641
15. Krest, C. M.; Onderko, E. L.; Yosca, T. H.; Calixto, J. C.; Karp, R. F.; Livada, J.; Rittle, J.; Green, M. T. *J. Biol. Chem.* **2013**, 288, 17074–17081
16. Krest, C. M.; Silakov, A.; Rittle, J.; Yosca, T. H.; Onderko, E. L.; Calixto, J. C.; Green, M. T. S*. Nature Chem*. **2015**, 7, 696–702.
17. Poulos, T. L. *Arch. Biochem. Biophys*. **2010**, *500*, 3–12.
18. Poulos, T. L.; Kraut, J.; *J. Biol. Chem.* **1980,** *255*,8199-8205
19. Davies, M. J. J. Clin. Biochem. Nutr. **2010**, *48*, 8–19.
20. Solomon, E. I.; Brunold, T. C.; Davis, M. I.; Kemsley, J. N.; Lee, S.-K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y.-S.; Zhou, J. *Chem. Rev*. **2000**, *100*, 235–350.
21. Butler, C. S.; Mason, J. R. *Adv. Microb. Physiol*. **1996**,*38*, 47–84.
22. Tshuva, E. Y.; Lippard, S. J. *Chem. Rev.* **2004**, *104*, 987–1012.
23. Lee, S. J.; Mccormick, M. S.; Lippard, S. J.; Cho, U.-S. *Nature*. **2013**, *494*, 380–384.
24. Shu, L.; Nesheim, J.C.; Kauffmann, K.; Munck, E.; Lipscomb, J.D.; Que, L. Jr.; *Science*, **1997**,275, 515-518

25. Pazmiño, D.E. Torres; Winkler, M; Glieder, A; Fraaije, M.W. *J. Biotechnol.* **2010**, *146*, 9-24.

26. Schirmer, A.; Rude, M. A.; Li, X.; Popova, E.; Cardayre, S. B. D. *Science*. **2010**, *329*, 559–562

27. Shanklin, J.; Andre, C.; (Brookhaven Science Associates, Llc) Compositions and methods for the relief of inhibition of aldehyde decarbonylase US Patent 20140370563 A, August 26th 2011

28. Schirmer, A.; Rude, M. A.; Brubaker, S.A.; (Ls9, Inc.) Methods and compositions for producing fatty alcohols and fatty aldehydes US Patent US8658404 B2. May 16 2008

29. Das, D.; Eser, B. E.; Han, J.; Sciore, A.; Marsh, E. N. G. *Angew. Chem.* **2011**, *50*, 7286–7290

30. Li, N.; Chang, W.-C.; Warui, D. M.; Booker, S. J.; Krebs, C.; Bollinger, J. M. *Biochemistry* **2012**, *51*, 7908–7916.

31. Rajakovich, L. J.; Nørgaard, H.; Warui, D. M.; Chang, W.-C.; Li, N.; Booker, S. J.; Krebs, C.; Bollinger, J. M.; Pandelia, M.-E. *J. Am. Chem. Soc*. **2015**, *137*, 11695–11709

32. Marsh, E. N. G.; Waugh, M. W. *ACS Catal***. 2013**, 2515–2521

33. Pandelia, M. E.; Li, N.; Nørgaard, H.; Warui, D. M.; Rajakovich, L. J.; Chang, W.-C.; Booker, S. J.; Krebs, C.; Bollinger, J. M. *J. Am. Chem. Soc.* **2013**, *135,* 15801–15812.

34. Waugh, M. W.; Marsh, E. N. G. *Biochemistry* **2014**, *3*,5537–5543.

35. Aukema, K. G.; Makris, T. M.; Stoian, S. A.; Richman, J. E.; Münck, E.; Lipscomb, J. D.; Wackett, L. P. *ACS Catal*. **2013**, *3*, 2228–2238.

36. Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature*. **2012**, *485*, 185–194.

37. Turner, N. J. *Nat. Chem. Biol*. **2009***, 5*,567–573.

38. Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. *Nature.* **2001**, *409*, 258–268.

39. Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. *Science*. **2010**, *329*, 305–309

40. Slini, T.; Giama, E.; Papadopoulos, A. *Int. J. Sustain. Energ.* **2014**, *34,*259–270.

41. Luque, R.; Herrero-Davila, L.; Campelo, J. M.; Clark, J. H.; Hidalgo, J. M.; Luna, D.; Marinas, J. M.; Romero, A. A. *Energy Environ. Sci.* **2008**, *1*, 542–542.

42. Lynd, L. R.; Laser, M. S.; Bransby, D.; Dale, B. E.; Davison, B.; Hamilton, R.; Himmel, M.; Keller, M.; Mcmillan, J. D.; Sheehan, J.; Wyman, C. E. *Nat. Biotechnol.* **2008**, *26*,169–172.

43. Liew, W. H.; Hassim, M. H.; Ng, D. K. *J. Clean. Prod*. **2014**, *71*, 11–29

44. Wilson, D. B. *Curr. Opin.Biotech*. **2009**, *20,* 295–299.

45. Margeot, A.; Hahn-Hagerdal, B.; Edlund, M.; Slade, R.; Monot, F. *Curr. Opin.Biotech*. **2009**, *20*,372–380

46. Alvira, P.; Tomás-Pejó, E.; Ballesteros, M.; Negro, M*. Bioresour. Technol*. **2010**, *101*, 4851–4861

47. Yanase, H.; Sato, D.; Yamamoto, K.; Matsuda, S.; Yamamoto, S.; Okamoto, K. *Appl. Environ. Microb.* **2007**, *73*, 2592–259

48. Weng, J.-K.; Li, X.; Bonawitz, N. D.; Chapple, C. *Curr. Opin.Biotech.* **2008**, *20*, 166–172

49. Martinez, D.; Larrondo, L. F.; Putnam, N.; Gelpke, M. D. S.; Huang, K.; Chapman, J.; Helfenbein, K. G.; Ramaiya, P.; Detter, J. C.; Larimer, F.; Coutinho, P. .; Henrissat, B.; Berka, R.; Cullen, D.; Rokhsar, D. *Nat. Biotechnol*. **2004**, *26*, 899–899

50. Banat, I., Nigam, P., Singh, D., & Marchant, R. *Bioresour. Technol*.**1996**, *58*, 217-227.

51. Doyle, W. A.; Blodig, W.; Veitch, N. C.; Piontek, K.; Smith, A. T. *Biochemistry* **1998**, *37*,15097–15105.

52. Doyle, W. A.; Smith, A. T. *Biochemical J*. **1996**, *315*,15–19

53. Smith, A. T.; Doyle, W. A.; Dorlet, P.; Ivancich, A. *Proc. Natl. Acad. Sci. U.S.A*. **2009**, *106*, 16084–16089

54. Shen, C.; Liao, J. Metabolic Engineering of Escherichia Coli for 1-Butanol and 1-Propanol Production via the Keto-Acid Pathways. *Metab. Eng*. **2008**, *10*, 312–320.

55. Green, E. M. *Curr. Opin. Biotechnol*. **2011**, *22*, 337–343.

56. Daroch, M.; Geng, S.; Wang, G. Recent Advances in Liquid Biofuel Production from Algal Feedstocks. *Appl. Energy*. **2013**, *102*, 1371–1381

57. Kallio, P.; Pásztor, A.; Thiel, K.; Akhtar, M. K.; Jones, P. R. *Nat. Commun*. **2014**, *5*, 4731–4731.

58. Howard, T. P.; Middelhaufe, S.; Moore, K.; Edner, C.; Kolak, D. M.; Taylor, G. N.; Parker, D. A.; Lee, R.; Smirnoff, N.; Aves, S. J.; Love, J. *Proc. Natl. Acad. Sci. U.S.A*. **2013**,*110*, 7636–7641

59. Marcus, R. A. *J. Chem. Phys.* **1956**, *24*, 966-978.

60. Saen-Oon, S.; Lucas, M. F.; Guallar, V. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15271-15285

61. Warren, J. J.; Ener, M. E.; Vlček, A.; Winkler, J. R.; Gray, H. B. *Coord. Chem. Rev.* **2012**, *256*, 2478–2487.

Chapter 2.

Mobilizing electrons in Cytochrome c peroxidase: designing and testing an electron transfer wire

2.1. Cytochrome c peroxidase (CcP)

Originally isolated from baker's yeast, CcP is a part of the heme peroxidase superfamily [1]. In nature, CcP catalyzes the oxidation of ferrocytochrome *c* to ferricytochrome *c* with the accompanying reduction of hydrogen peroxide to water [2]; **Figure 2-1** illustrates the overall catalytic reaction of CcP. Following the publication of the X-ray structure of CcP, many questions concerning the nature of electron transfer (ET) during the catalytic cycle emerged [3]. Specifically, researchers asked how CcP can oxidize another macromolecule, like cytochrome *c* (cyt c), using an embedded heme active site?

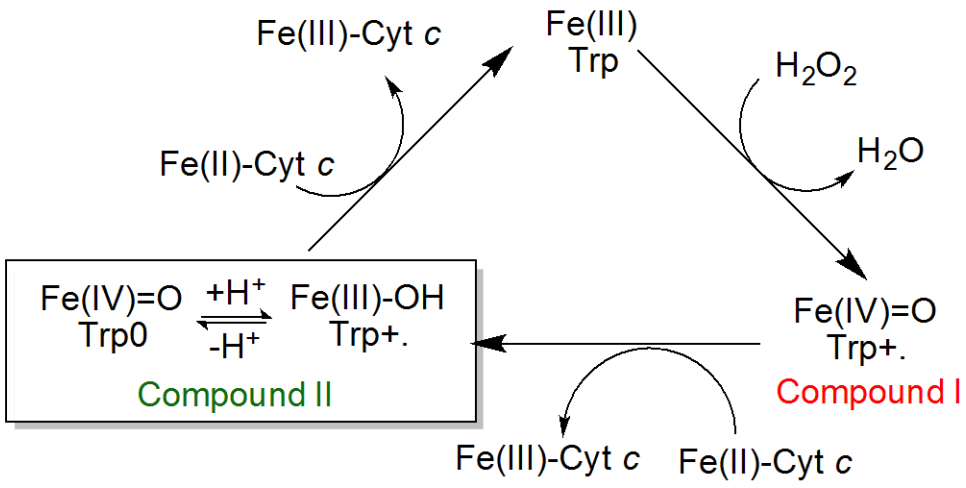

Figure 2-1. Schematic of the catalytic cycle of CcP. The amino acid (AA) radical (Trp191) is highlighted; Trp191 is involved with ET with the natural cofactor, Cyt *c*.

We now know that protein embedded radicals are involved in many ET roles in biological systems [4,8] the most common examples are for the aromatic amino acids Trp and Tyr [5]. CcP was among the first proteins where an embedded amino acid radical was identified as part of the enzyme's function. In CcP, the key residue involved with ET between cyt c was found to be a tryptophanyl radical (W•+191) located in the proximal heme pocket [6]. The radical has a remarkably long half-life ($t_{1/2}$ ~ 6.6 hr) [7]. The formal reduction potential for Trp in solution ($E^{\circ\prime}(Trp^{\bullet+/0})$) ranges from 1.05-1.15 V [8], but factors such as local environment can greatly modulate $E^{\circ\prime}$, making AA radicals accessible under biological conditions. The local environment of the Trp radical is crucial to stabilizing the radical. As such, numerous studies have been conducted investigating the local environment of Trp191 in CcP [9,10,11,12].

Other heme peroxidases also contain transient tryptophanyl radicals. Perhaps the most notable example is *Phanerochaete chrysosporium* lignin peroxidase (LiP). LiP produces a radical centered on Trp171, which is located on the surface of the protein [13].The Trp171 radical is crucial to carrying out oxidative chemistry of organic compounds at the surface of the enzyme, rather than direct oxidation at the buried heme site. Much interest has been given to artificial proteins that harness this distinct feature of LiP as a means to degrade lignin [14].The work presented in this chapter uses CcP as a scaffold to introduce new amino acid radical chemistry that is reminiscent of LiP. CcP is an optimal scaffold as it provides high expression yields and already contains an AA radical in the proximal pocket. In this chapter, I will discuss two distinct strategies for the introduction of surface AA radicals.

2.1.1. Strategy 1: Introduction of lignin peroxidase activity in CcP

As mentioned above, our goal is to imitate LiP lignin oxidation reactivity in CcP. This approach involves integrating key structural and functional features of LiP into the protein backbone of CcP. The most important feature of LiP that we tried to emulate is the surface exposed Trp radical. The microenvironment surrounding the radical is also important to consider as we refine designs [15]. Related attempts at bioengineering surface-exposed AAs have proven successful. For example, a surface exposed AA radical was installed into *Corprinus cinereus* peroxidase by incorporating three mutations

(Asp179Trp + Arg258Glu + Arg272Asp) [16]. The goal in this thesis was to do the same in CcP by introducing: (1) modifications to the heme pocket; (2) a redox-active surface-exposed Trp; (3) a negatively charged surface patch around Trp; and (4) a cation binding loop near the heme.

The heme pocket contains Trp residues in both the distal and proximal pocket, as shown in **Figure 2-2**. Though Trp191 is the catalytically active residue CcP, it has been experimentally confirmed that Trp51 can also form a radical in the absence of its natural redox partner, cyt *c* [17]. Thus, by replacing Trp51 and Trp191 in CcP with Phe, we can prevent the formation of those radicals at the active site [18]. Next, introduction of a surface Trp within electron tunnelling distance from the heme active site is required for catalysis. This residue will be introduced by mutating Ala170 to Trp. Analysis of the crystal structures of LiP and CcP shows that CcP-Ala170 is in a position analogous to redox-active LiP-Trp171. Further examination of the CcP-Ala170 environment shows two nearby Tyr, which also can be redox active. These two residues are mutated to Phe to minimize the possibility for redox side reactions. In wild type LiP, the surface exposed AA radical is surrounded by five negatively charged residues (**Figure 1-7**). This charged patch plays a role in catalytic activity, as demonstrated by site directed mutagenesis studies [19]. In these studies, it was found that three specific residues are crucial for the rate of catalysis: Glu168, Glu250, and Asp264. This negatively charged surface patch can be incorporated into CcP by mutating a positively charged Arg to Leu (CcP-Arg166Leu). A negatively charged residue (CcP-Asp261) exists on the scaffold, and an additional negative residue can be added (CcP-Ile247Glu) to the microenvironment of the surface Trp.

Additional refinements to the CcP heme pocket include swapping the proximally located Met172 to Leu. We propose that the introduction of this mutation will remove the distortion at the heme plane. It is believed that the distorted heme plane is a direct result of the increased electron density and large van der Waals radius of the sulphur atom on the methionine residue. By mutating this residue to Leu, we predict this distortion will be diminished and consequently change the heme reduction potential.

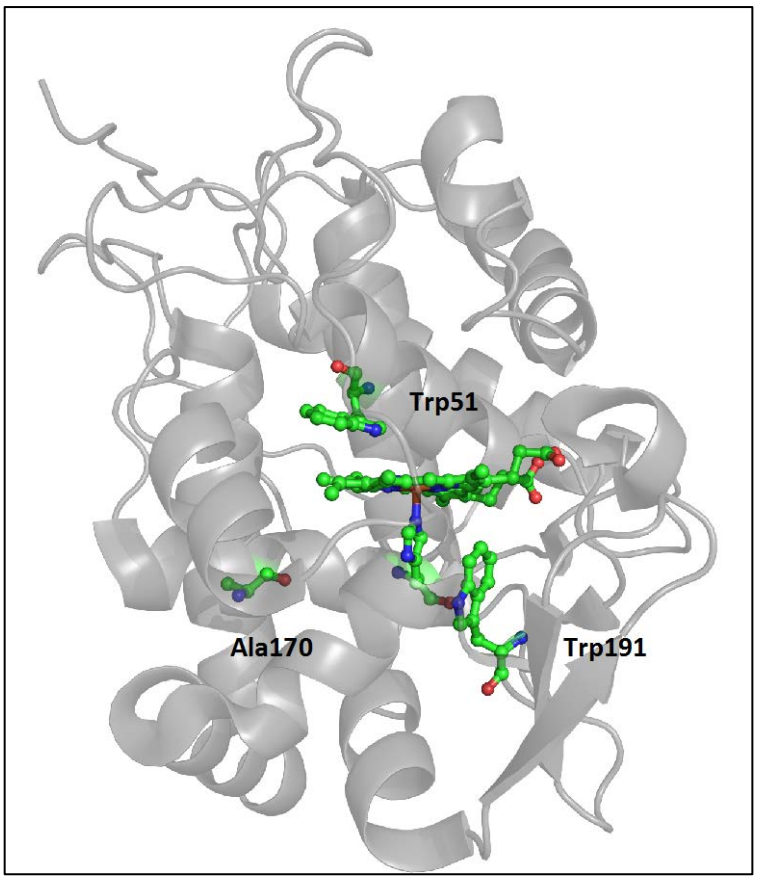

Figure 2-2. Heme pocket of CcP illustrating the Trp residues in the heme pocket and Ala170. Ala170 is the proposed site of the introduced surface exposed AA radical.

Cation binding loops can be found in all peroxidases. However, CcP is unique because it binds monovalent K^+ as opposed to divalent Ca^{2+}, which is found in other peroxidases [1].Electrostatic arguments suggest that the presence of Ca^{2+} disfavors the formation of a cation radical at Trp191 (the site of catalysis in CcP) [20]. Native LiP has two Ca^{2+} binding loops located on both proximal and distal sides of the heme. The presence of such Ca^{2+} binding loops play important roles in maintaining the integrity of heme active sites [21,22]. By introducing several conservative mutations, a Ca^{2+}-binding loop can be integrated into CcP, analogous to the binding loop in native LiP [20]. I hypothesize that the introduction of this modified binding loop will promote formation of the surface Trp radical required for catalysis.

The desired mutations were introduced to our scaffold by designing and obtaining synthetic genes from commercial sources (see experimental section). Two synthetic

genes were obtained: one contained all the desired mutations (called CcP-Full) and the second contained a more conservative group of mutations, including only the imperative Trp to Phe and catalytic Trp mutations (called CcP-Basic). Work on these two scaffolds is discussed in further detail below.

2.1.2. Strategy 2: Designing a tryptophan wire in CcP

Redox active AAs, such as Tyr and Trp, are vital in biological redox reactions and energy transduction. Electron hopping through proteins requires redox-active cofactors to shuttle electrons from one part of the biomolecule to another, and Tyr/Trp are often able to maintain the free energy of holes above 1 V because of their intrinsically high reduction potentials. Recently, Winkler and Gray used a systematic search of the Protein Data Bank (PDB) to propose that chains of Trp and/or Tyr amino acids in enzymes can act as a protective measure for maintaining the longevity of redox enzymes, especially oxidoreductases like CYPs [23]. The Tyr/Trp chains serve as a protective mechanism against oxidative damage from cofactors that are activated in the absence of substrate (e.g., high valent heme in CYPs) [24,25]. The precise location of these Trp/Tyr chains are crucial, because the proposed protective mechanism must not overtake the enzyme's natural biological function. Though these Tyr/Trp chains appear to have been evolved to protect the enzyme against oxidative damage, our goal is to use artificial chains as a strategy to transfer energetic holes from activated cofactors to enzyme surfaces to carry out oxidative chemistry.

Electron hopping for Tyr is more complex than other redox-active AAs because oxidation of phenols requires loss of both an electron and a proton. Recent investigations have revealed that ET is favoured when there is hydrogen bonding to the phenolic proton [4]. ET is more likely for Tyr when the phenolic proton is proximal to AAs such as Glu, Asp, and His. Examples include the redox active Tyr in photosystem II and ribonucleotide reductase. [26]. Conversely, the protonated Trp radical (Trp•+) is more stable (less acidic than the Tyr radical) and can participate in pure ET reactions. Note that it also can be found near proton acceptors (e.g., Glu, Asp), as highlighted above for LiP. Therefore, designing artificial electron wires with Trp residues are an ideal starting point.

As outlined above, Ccp-Trp191 forms a stable AA radical in the proximal heme pocket. Though Trp191 is not directly on the surface of the protein, it is involved with long-range ET with CcP's redox partner, cyt *c*. In my modified CcP proteins, Trp191 will shuttle holes between surface residues and the heme. Introduction of two Trp residues (Ala193Trp and Tyr229Trp, **Figure 2-3**) will complete the designed Trp chain between the heme and the CcP surface. The corresponding single mutants (**Figure 2-4**) will also be evaluated. The ability of this Trp chain to deliver holes to the surface of the C*c*P, where oxidation of organic substrates can occur, is described below. This mode of reactivity is analogous to LiP function.

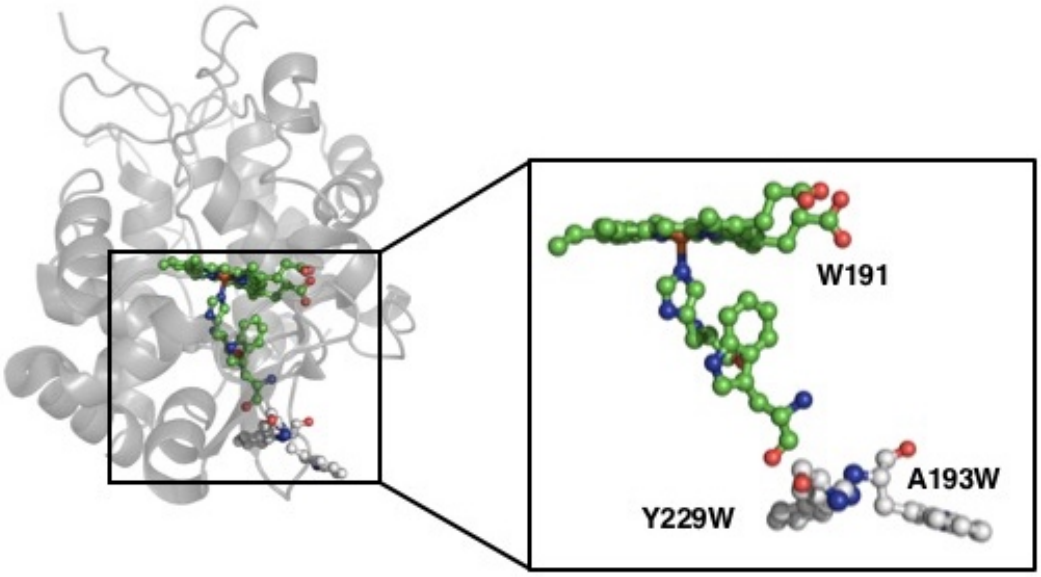

Figure 2-3. Tertiary structure of CcP illustrated. Insert shows the proposed Trp chain to the surface of CcP abbreviated as, CcP-WW.

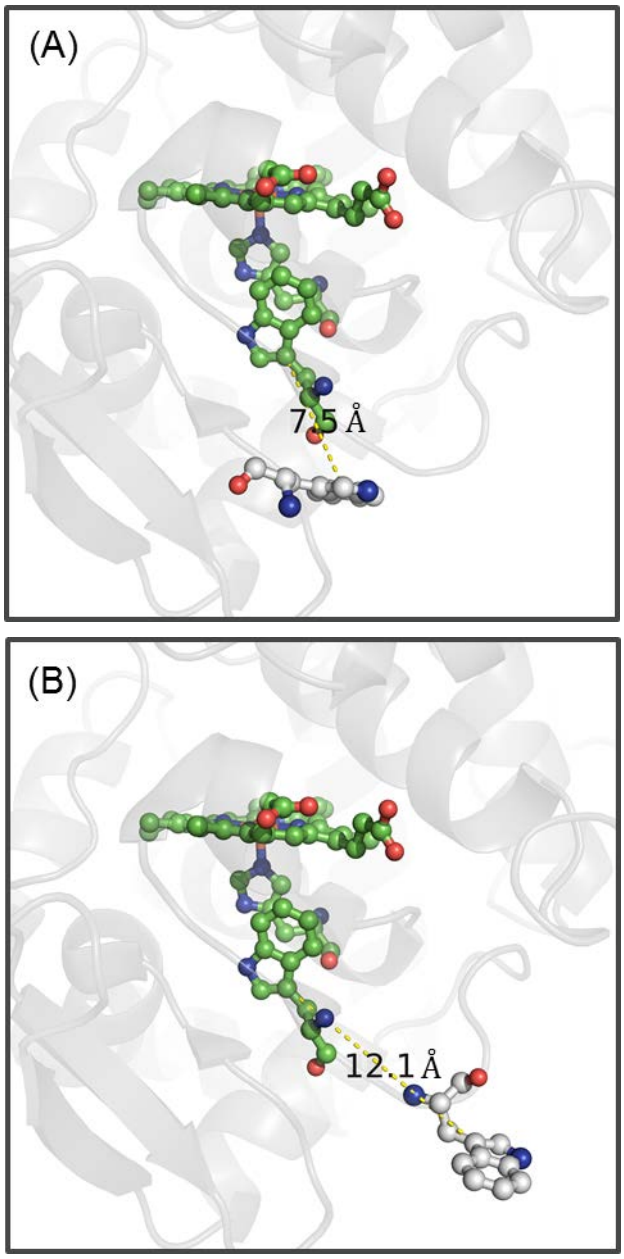

Figure 2-4. **PyMol rendering of the CcP active site illustrating two single mutants of interest (A) displays the Tyr229Trp mutant abbreviated as CcP-W1 (B) displays the Ala193Trp mutant abbreviated as CcP-W2.**

2.2. Materials and Methods

Site directed mutagenesis

Plasmids encoding the *Saccharomyces cerevisiae* CcP gene were provided by Alexander N. Volkov from Universiteit Brussel [33]. The gene was cloned into pET24a(+) NdeI and HindIII restriction sites to introduce a C-terminal His_6 tag. The plasmid also included the addition of three N-terminal residues, MKT, which is known to increase CcP expression yields [27]. All CcP mutants were prepared by PCR using primers purchased from Eurofins Operon (primers used are outlined in **Appendix A**). Standard protocols were followed for PCR reactions. Site directed mutant plasmids were obtained with the Quikchange protocol (Agilent Technology). DNA polymerase (Q5) and DpnI were obtained from New England Biolabs. PCR reaction mixtures were transformed into chemically competent DH5α *E. coli* plated on kanamycin-supplemented agar plates, and incubated overnight (37°C). Single colonies were grown overnight and plasmid DNA was purified using a Qiagen plasmid spin MiniPrep kit. Plasmid DNA was sequenced by Eurofins Operon Simple Sequence read service to verify mutations

Protein expression

Standard transformation protocols were followed using *E. coli* BL21(DE3) on kanamycin-supplemented plates (50 µg/mL). A single colony was incubated for five hours (37 ℃, 180 rpm shaking) in 25 mL luria broth (LB) with 50 µg/mL kanamycin. Large-scale expression was initiated by inoculating 1 L of terrific broth (TB) containing 50 µg/mL kanamycin with 1 mL of the starter culture. Cultures were grown to an OD_{600} of approximately 0.6-0.8 at which time IPTG was added to a final concentration of 1 mM. Cultures were grown overnight (32℃, 180 rpm shaking) after which bacteria were pelleted via centrifugation.

Purification

Cell pellets were resuspended into 50 mM Tris•HCl,100 mM NaCl, 5 mM imidazole, and 1mM PMSF (pH 8). The resulting suspension was lysed by sonication on ice for 15 minutes and the supernatant was collected by centrifugation (30 minutes at 14000 rpm, 4℃). The pH of the supernatant was adjusted to 8 and then loaded directly onto a Ni-NTA

affinity column. The column was washed with at least 10 column volumes of buffer containing 20 mM Tris•HCl, 5 mM imidazole, and 500 mM NaCl (pH 8). Increasing the concentration of imidazole to 250 mM eluted the protein of interest. The buffer was exchanged to 100 mM sodium phosphate (pH 7.5). A hemin solution was then prepared in 20 mM sodium phosphate buffer (pH 6), containing approximately 1.2 equivalents of hemin chloride and then added to the protein solution. The protein solution containing hemin was dialyzed with SnakeSkin™ Dialysis tubing (molecular weight cutoff 3.5 kDa) against 20 mM sodium phosphate buffer (pH 6) at 4℃ overnight. The protein was then concentrated using a 10 kDa centrifugal filter. The buffer was then exchanged into 5 mM sodium phosphate (pH 6) using a PD10 desalting column and loaded directly onto an FPLC DEAE column. The protein was purified by increasing salt concentration (buffer b: 5mM sodium phosphate+ 500 mM NaCl, pH 6). Fractions containing labeled protein were pooled and concentrated and stored at 4℃. Protein that was not intended for immediate use was flash-frozen in liquid nitrogen with 20% glycerol and stored at -80℃.

Optical spectroscopy

The optical (UV-Visible) spectra were obtained by either a Photon Control SPM-002-EH CCD with a SPLC 1DH deuterium/tungsten light source or a Cary 100-Bio UV-visible spectrometer. The optical spectra for WT CcP and mutants were collected in 20 mM sodium phosphate buffer (pH 6). The extinction coefficient of the Soret was found by using the pyridine hemochrome assay [28]. After collecting the resting state spectra, Cmpd I was generated by addition of approximately one equivalent of hydrogen peroxide, and spectra were collected immediately. To generate CN^- bound CcP, approximately one equivalent of CN^- was added to ferric CcP solutions and spectra were collected immediately.

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were collected using a Chirascan qCD spectrometer scanning from 200 nm to 240 nm. A total of five scans were taken and averaged with a step size of 1 nm and 0.5 seconds per point. Raw data was converted using the following formula $[\theta] = \theta M_r/(nCl)$ where θ, M_r, n, C, and l represent ellipticity in millidegrees, molecular weight (Da), number of residues, concentration in mg/ml and

the cuvette path length respectively. Values are reported as (deg^2•cm•mol^{-1}) against wavelength (nm). All scans were collected in 20 mM sodium phosphate buffer (pH 6).

Mass spectrometry

Electrospray ionisation (ESI) and matrix-assisted laser desorption/ionization (MALDI) were routinely used to confirm identity and purity of WT CcP and all mutants (CcP-WW, CcP-W1, CcP-W2). ESI data were collected on an Agilent 6210 TOF ESI-MS instrument and MALDI data were collected using a Bruker microFLEX MALDI-TOF instrument with a 48 well ground steel target. Sinapinic acid was used as a MALDI matrix and mixed in equal amounts with desalted protein solution using the dried droplet method. Theoretical molecular weights were obtained from the protein sequence (sequenced via Eurofin Operon) using the ExPasy Compute Pi/MW tool [29].

Compound I stability

Cmpd I stability measurements were performed using a Cary 100-Bio-UV-visible spectrometer. To assess stability of Cmpd I, the absorbance at 424 nm was monitored over 20 minutes. The reaction mixture contained 6.5 µM of ferric CcP protein and 1 equivalent of H_2O_2. Data collection was initiated immediately upon addition of hydrogen peroxide. All protein samples were run in 20 mM sodium phosphate buffer at pH values of 6 or 8.

Oxidation assays

All oxidation assays were conducted on a Cary 100-Bio-UV-visible spectrometer. For the veratryl alcohol (VA) activity assay, the oxidation of VA to veratryl aldehyde was followed by monitoring the absorbance at 310 nm, following published protocols [30]. Enzyme assays contained 2 mM VA, 400 µM H_2O_2 and 10 µM protein. VA oxidation was investigated at pH 4, 6, and 8 in 20 mM sodium phosphate buffer. Decolourization of 4-p-nitrophenylazoresorcinol was investigated at pH 8 by monitoring the absorbance feature at 466 nm. Reactions contained 41.8 µM 4-p-nitrophenylazoresorcinol, 400 µM H_2O_2, and 10 µM protein in 20 mM sodium phosphate buffer.

2.3. Results

2.3.1. Introduction of LiP activity in CcP scaffold

Protein expression

In these experiments, I focused on mimicking the key features of LiP through several rational designs denoted CcP-Full and CcP-Basic. Synthetic genes for CcP-Full and CcP-Basic were obtained from Eurofins Operon and were subcloned into pET 20b(+) and pET 22b(+) expression vectors. Several methods to promote overexpression of the highly modified CcP enzymes were investigated. However, all attempts to isolate solubilized fractions were unsuccessful. Analysis of protein pellets using sodium dodecyl sulfate (SDS) gel electrophoresis demonstrated the proteins were expressed as inclusion bodies. **(Figure 2-5)**. Several other strategies were explored to solubilize and refold the proteins (e.g., on-column refolding) [31], but again all methods proved unsuccessful. Perhaps a more tractable approach to introducing LiP reactivity would involve a step-by-step introduction of mutations. This would allow us to investigate the effect of each mutation, and determine which mutation, or mutations, result in a loss of structural integrity.

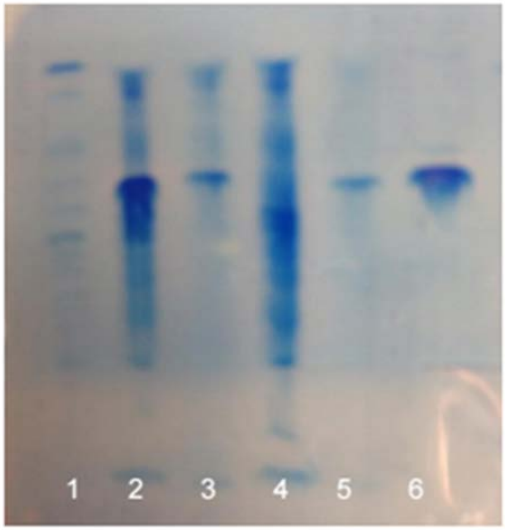

Figure 2-5. Coomassie stained SDS gel of MG expression of base variant. Lane: (1) ladder, (2) wild type CcP with lysis buffer, (3) wild type CcP with lysis buffer and 4.1 M urea, (4) CcP-Basic with lysis buffer + 4.1 M urea, (5) purified wild type CcP control

2.3.2. Designing a tryptophan wire in C*c*P

The series of artificial proteins designed to include Trp wires were all expressed as soluble protein and isolated according to literature protocols (see above). The following sections describe my work to characterize the structure/fold and identity of each protein, to characterize the ferric states of each variant, and to evaluate each proteins' ability to support high oxidation state intermediates.

Optical spectroscopy

The heme environment is highly sensitive to perturbations to the surrounding protein environment. Changes in oxidation state, spin state, or heme environment can affect the cofactor's electronic structure and this is often reflected in the electronic absorption spectra [32]. Three different coordination states of CcP and its mutants were investigated: the resting state, Cmpd I, and the ferric-CN complex. The resting state of WT CcP has absorbance features consistent with a five coordinate heme species, (**Figure 2-6 black trace**), in agreement with literature reports [33]. All of the new CcP mutants (CcP-WW, CcP-W1, CcP-W2) have similar spectra with some slight changes in the R/Z (A_{Soret}/A_{280}) ratio. These variations in the R/Z ratio are due to the mutations to the aromatic residues (i.e., addition of Trp) which directly affect the protein-based absorbance at 280 nm [34].

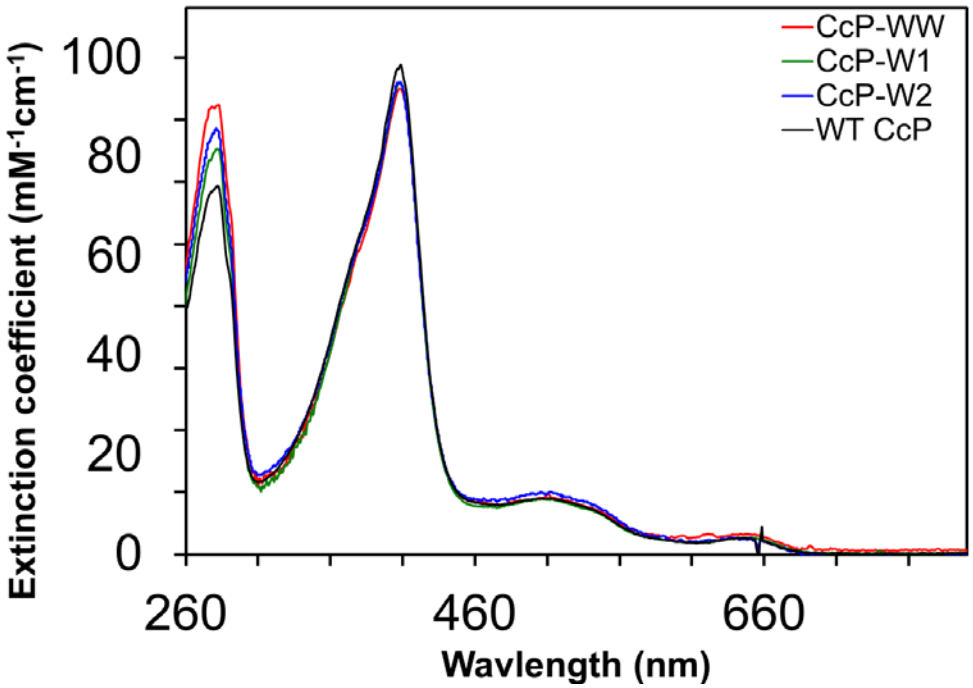

Figure 2-6. UV-visible spectra of WT CcP and mutants resting state in 20 mM sodium phosphate pH 6 buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace)

The notable stability of Cmpd I in CcP at room temperature offers a means to probe the relative reactivity of WT CcP and mutants described here (CcP-WW, CcP-W1, CcP-W2). Cmpd I was generated by adding 1 equiv. of hydrogen peroxide. WT CcP forms the characteristic Cmpd I, as indicated by a distinct red shift of the Soret peak and changes in the Q bands (**Figure 2-7,** black trace). A Soret band centered at 419 nm is indicative of the $(P^{\bullet+})Fe^{IV}$=O species [35]. This result was obtained for all our mutants (CcP-WW, CcP-W1, CcP-W2).

Alternatively, changes to the heme spin state can be elicited by coordinating ligands such as fluoride (to produce high spin heme) or cyanide (to produce low spin heme) [36]. Here, sodium cyanide was added to ferric CcP solutions to generate the spectra set out in **Figure 2-8**. A characteristic red shift in the Soret peak was observed for WT and all CcP mutants (CcP-WW, CcP-W1, CcP-W2). Taken together, these results confirm the integrity of the heme site for the surface Trp mutants and the ability of these proteins to support the production of high-valent intermediates.

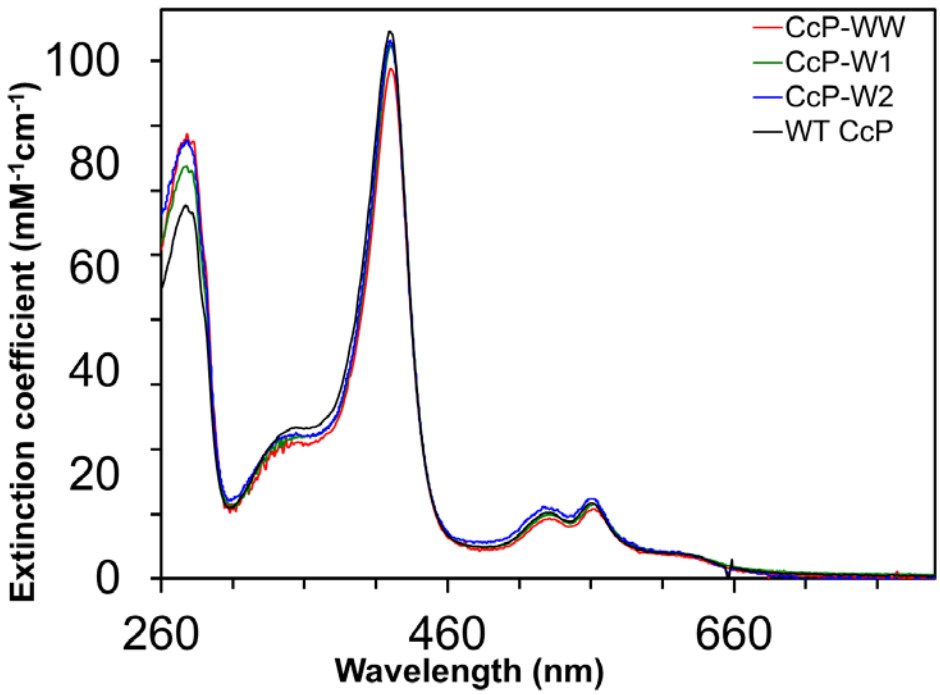

Figure 2-7. **UV-visible spectra of WT CcP and mutants Cmpd I in 20 mM sodium phosphate (pH 6) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace)**

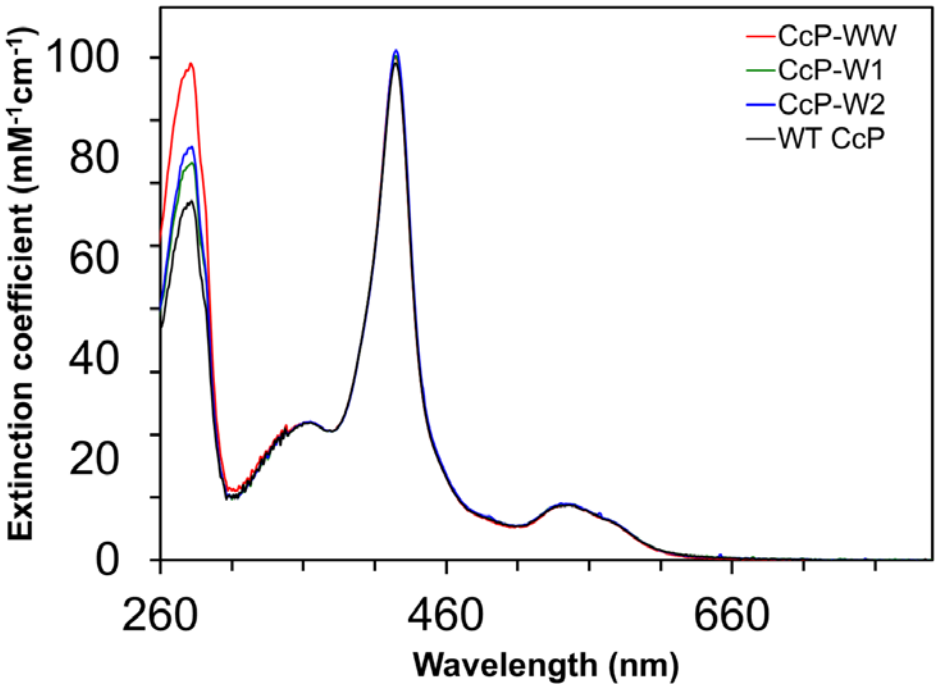

Figure 2-8. **UV-visible spectra of WT CcP and mutants coordinated to CN^- in 20 mM phosphate (pH 6) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace)**

Circular dichroism spectroscopy

CD spectroscopy is a useful technique for confirming that the integrity of secondary structure is conserved when a protein is subject to mutations. The structure of CcP is dominated by alpha helices **(Figure 2-3)** and such proteins exhibits CD bands at 222 and 208 nm [37]. The far-UV CD spectra of holo WT CcP, CcP-WW, CcP-W1, and CcP-W2 were obtained (**Figure 2-9).** In all cases the CD spectra are similar to those reported in the literature [38]. As there are no significant perturbations in the CD spectra (**Figure 2-9**), we conclude that our mutants (CcP-WW, CcP-W1, CcP-W2) hold analogous CcP secondary structures.

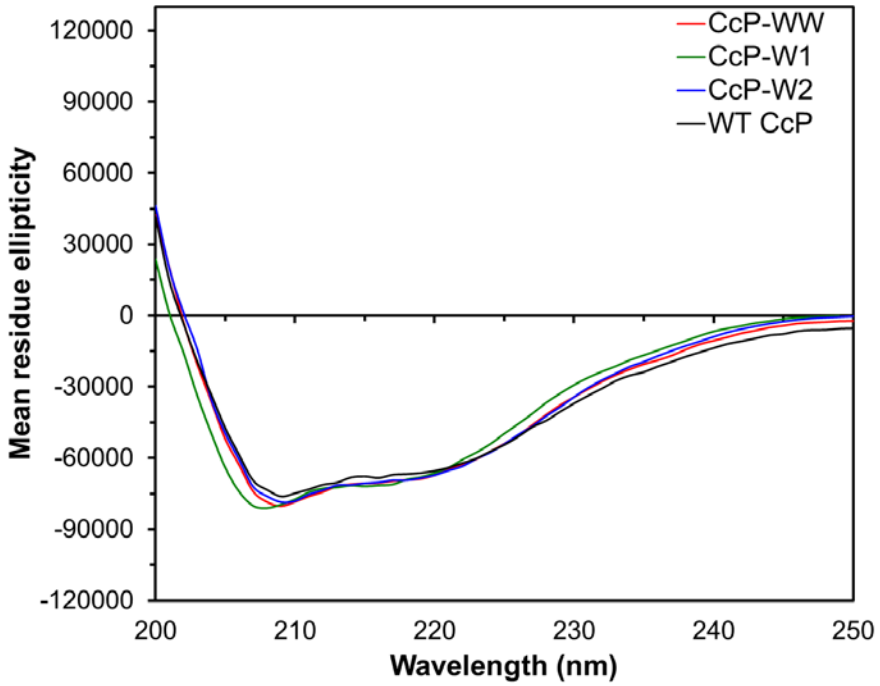

Figure 2-9. Far UV CD spectra in 20 mM phosphate (pH 6) CD spectra purified WT CcP and Mutants. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace) and CcP-WW (red trace).

Mass spectrometry

Mass spectrometry (MS) was used to confirm the identity of our modified CcP proteins. **Table 2-1** contains the theoretical and experimental masses (ESI and MALDI) for WT CcP, CcP-WW, CcP-W1, and CcP-W2. MALDI was often preferred as a rapid method to evaluating expression with the notable drawback of increased uncertainty as the molecular weight of the protein increased. This is exemplified for CcP-W2, where the

mass fluctuates around 60 Da. The fluctuations in the observed mass can be attributed to the peak width of the protein mass manifold. Consequently, ESI-MS was also used to characterize all CcP proteins (**Table 2-1**). The observed and calculated masses are in good agreement.

For both CcP-WW and CcP-W2, the theoretical and observed masses indicate a mass difference of approximately 14-16 da. This suggests that perhaps the surface Trp193 has been modified. This mass difference could be due to addition of an oxygen atom to the protein. Both CcP-WW and CcP-W2 display this mass difference, which is consistent with modification of Trp193. Further investigations should include performing a trypsin digestion (methods described in Chapter 3), which can confirm the modification of the Trp193 residue. The addition of a hydroxyl group to Trp residues is known in LiP, where the surface Trp171 undergoes the autocatalytic formation of a hydroxyl group at the Cβ of the Trp residue [39]. A model of such a modified Trp residue, N-acetyl-α,β −didehydrotrptophanamide, displays distinct UV bands where the absorbance band is red-shifted to 333 nm. A HPLC coupled to UV-vis detection can be used on digested proteins to further confirm this modification [40]. Likewise, protein crystals can be obtained in which the presence of significant electron density differences around Trp193 would confirm Trp hydroxylation [39]. Overall, the mass change in our CcP-WW and CcP-W2 mutants raises some intriguing question regarding the origins, and requires further investigation.

Table 2-1. ESI-MS and MALDI results of WT CcP and mutants

Name	Theoretical (Da)	Experimental (Da) (ESI)	Experimental (Da) (MALDI)
WT CcP (C-terminus His Tag)	34416	34419	34425
CCP-WW (A193W+ Y229W)	34554	34571	34542
CCP-W1 (Y229W)	34439	34442	34431
CCP-W2 (A193W)	34531	34546	34593

Compound I stability

We hypothesize that Cmpd I generated for the mutants will display decreased stability with respect to WT Cmpd I due to the addition of alternative ET pathways (via the new Trp residues). Note that the stability of Cmpd I is pH dependent, decaying more

rapidly at pH > 8 [41]. In addition, the reduction potentials of aromatic residues, such as Tyr and Trp, also are pH dependent (**Table 2-2**). The formal reduction potentials decrease as pH increases, in accord with predictions from the Nernst equation. To evaluate the stability of Cmpd I in CcP-WW, CcP-W1, and CcP-W2 UV-visible spectroscopy was used to monitor changes in absorbance to the Soret band at 424 nm. Experiments were carried out using 6.5 µM of ferric CcP protein and 1 molar equivalent of H_2O_2 in 20 mM sodium phosphate buffer at pH 6 **(Figure 2-10)** and 8 **(Figure 2-11).**

At pH 6, Cmpd I in WT CcP decays by about 10% over 20 minutes. This is similar to the amount of decay observed for CcP-W1. For both CcP-W2 and CcP-WW, we have Trp193, decay by closer to 15%. Interestingly, all four proteins show only minimal decay in the first 5-7 minutes, after which time the rates of decay become more variable. At pH 8, WT CcP has an initial decay rate that is slower than the other three mutants, and the Soret peak loses ca. 15% intensity over the course of the experiment. CcP-W1 loses about 15% intensity and as observed at pH 6, CcP-W2 and CcP-WW behave similarly, losing about 20% Soret intensity. The ending optical spectra for all four proteins display a similar trend. The Soret peak shift to a lower wavelength, reminiscent of the five coordinate resting state enzyme. These qualitative observations demonstrate the addition of surface Trp residues to CcP affects the stability of Cmpd I, consistent with the introduction of a new ET pathway.

Table 2-2. Reduction potentials of tryptophan and tyrosine amino acids[42].

Amino acids	Reduction potential in aqueous solutions E°′, V vs NHE		
	pH 2	pH 7	pH 13
Tryptophan	1.15	1.02	0.65
Tyrosine	1.22	0.93	0.72

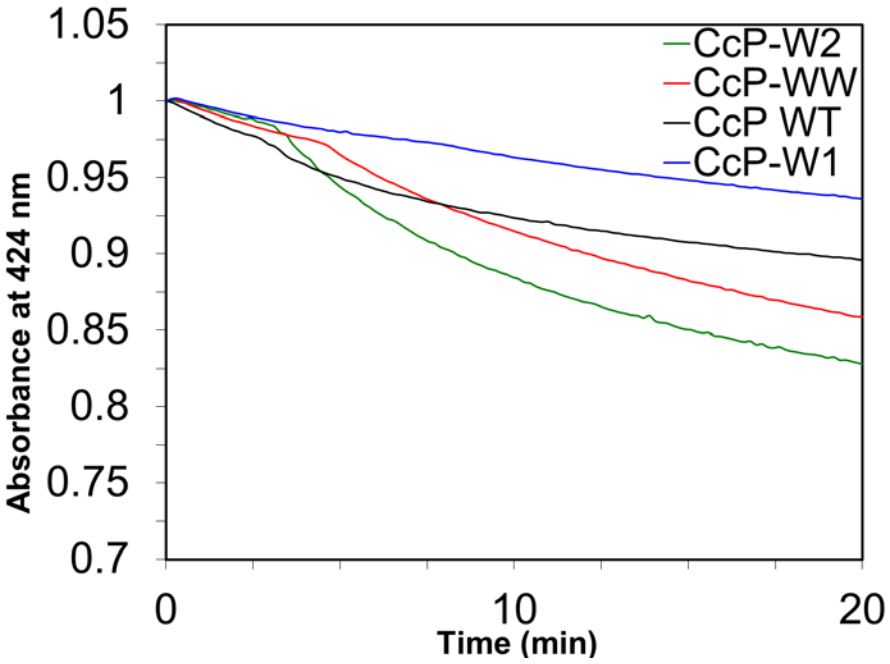

Figure 2-10. **Monitoring the loss of the Soret peak at 424 nm upon addition of 1 equivalent of H_2O_2 in 20 mM phosphate (pH 6) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace)**

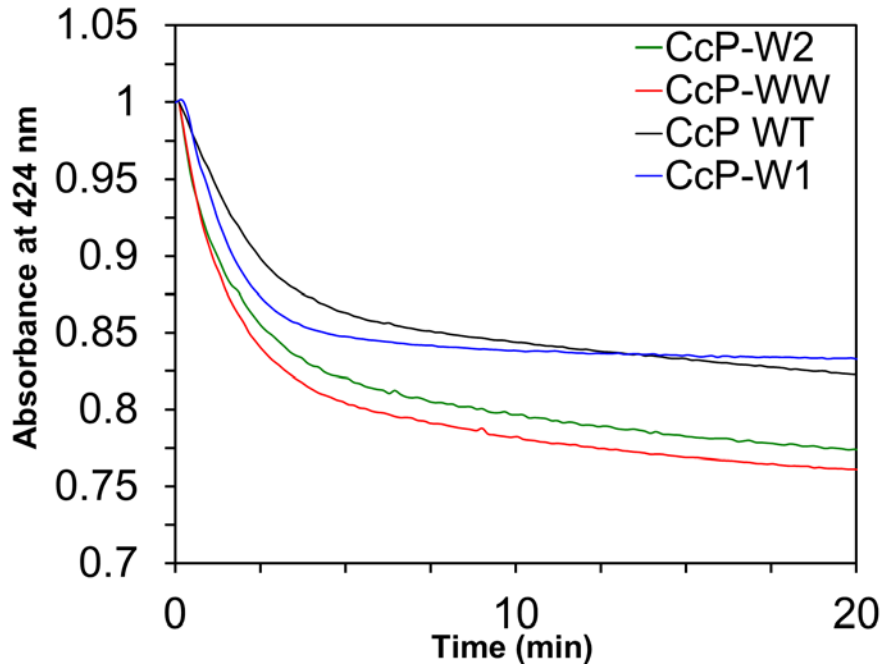

Figure 2-11. **Monitoring the loss of the Soret peak at 424 nm upon addition of 1 equivalent of H_2O_2 in 20 mM phosphate (pH 8) buffer. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace).**

Oxidation assays

Enzyme activity for WT CcP and all three mutants was assayed using a canonical activity assay used for LiP, where I monitored oxidation of VA to veratryl aldehyde **(Figure 2-12**) [15]. VA is proposed to be a secondary metabolite that is involved with the degradation of lignin. **Figure 2-13** shows two potential pathways for hole flow to the surface of CcP-WW where VA oxidation can occur. We propose that electrons flow via the red pathway, as opposed to the pathway which hops directly from A193W mutant (green pathway **Figure 2-13**). Each reaction mixture contained 2 mM VA, 400 µM H_2O_2 and 10 µM ferric protein. At a pH of 4 and 6 **(Figure 2-14 top** and **bottom**, respectively**)** the oxidation reaction, monitored using the appearance of the veratryl aldehyde peak at 310 nm, is negligible for all proteins assayed. In contrast, at pH 8 **(Figure 2-15),** a substantial increase in VA oxidation yield is observed for all enzymes. Consistent with studies of Cmpd I stability in the absence of substrate, the WT enzyme was the least reactive. CcP-WW oxidized almost twice as much VA as the other two mutants and three times as much as WT CcP over the course of 3 minutes (Table 2-3). These results suggest that the additional Y229W mutation provides an alternative pathway for holes to flow from heme to substrate.

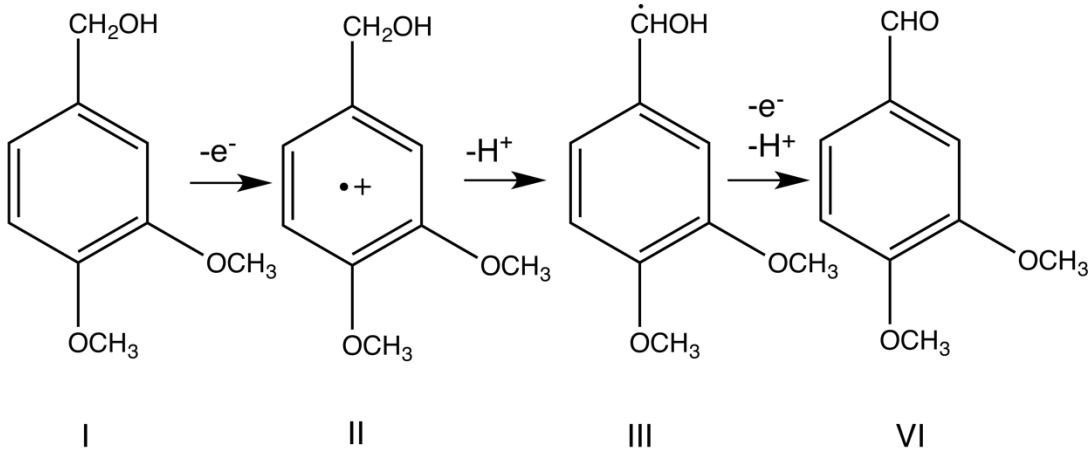

Figure 2-12. Two electron oxidation of VA to corresponding aldehyde. In nature, it is postulated that the radical cation produced in Step II can then go on to oxidize large organic compounds, such as lignin. Adapted from [43].

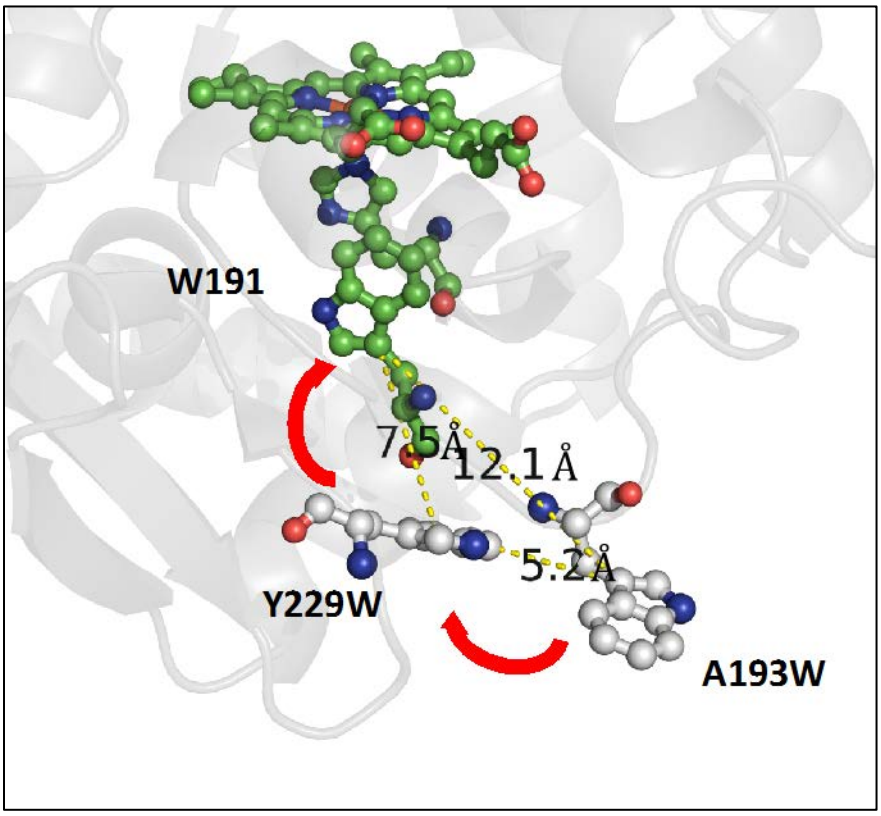

Figure 2-13. PyMol rendering of the CcP-WW active site illustrating two proposed pathways for electron hopping. (1) The red arrows illustrate two electron hopping steps to W191 (2) The green arrows illustrate the alternative one electron hopping step.

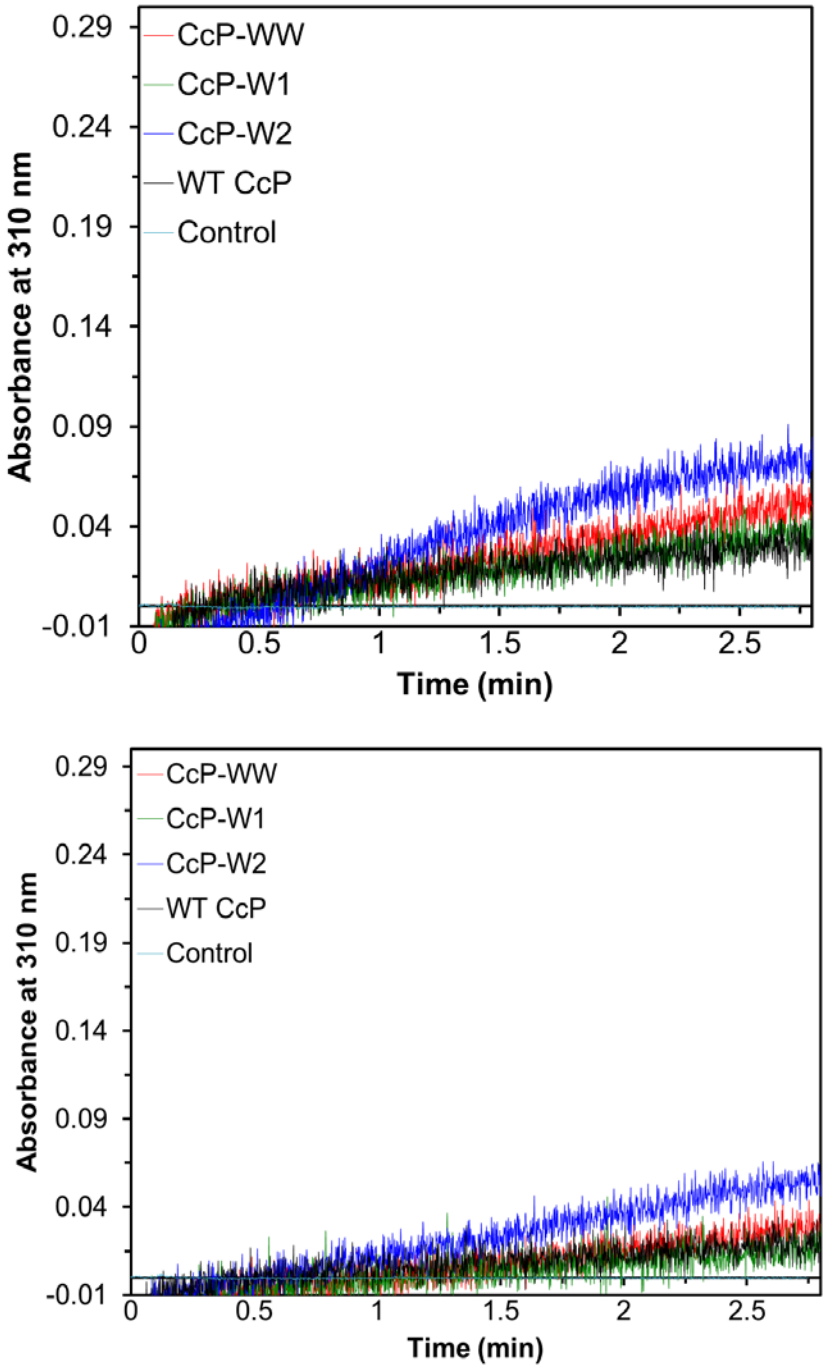

Figure 2-14. **Oxidation assay of veratryl alcohol to veratryl aldehyde by monitoring the absorption at 310 nm. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) Reaction mixture contains 400 μM H_2O_2, 10 μM protein and 2 mM veratryl alcohol in 20 mM sodium phosphate (pH 4-top pH 6 bottom) buffer, control carried out in the absence of enzyme.**

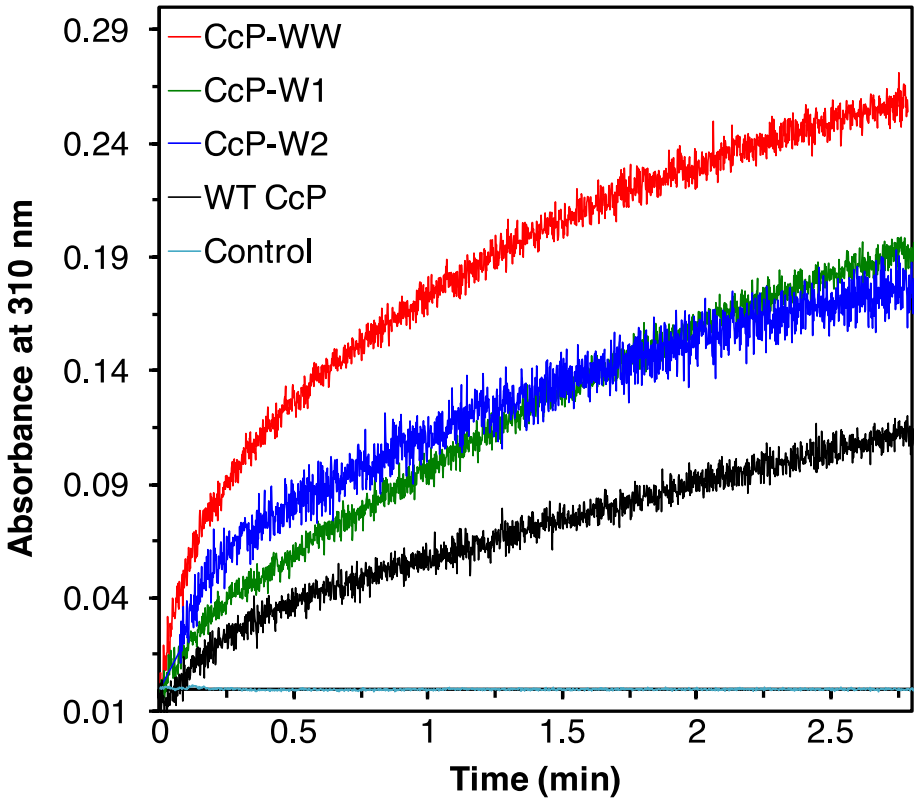

Figure 2-15. Oxidation assay of veratryl alcohol to veratryl aldehyde by monitoring the absorption at 310 nm. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace) Reaction mixture contains 400 µM H_2O_2, 10 µM protein and 2 mM veratryl alcohol in 20 mM sodium phosphate (pH 8) buffer, control carried out in the absence of enzyme.

In addition to VA oxidation, decolourization of the azo dye 4-p-nitrophenylazoresorcinol was explored. Peroxidase-catalyzed dye decolourization is known for other enzymes, including LiP [44, 45], but is less common for CcP. **Figure 2-16** shows the decolorization 4-p-nitrophenylazoresorcinol by the 4 CcP variants, and a trend similar to that described for VA oxidation is observed. CcP-WW (red trace) has the highest initial slope, followed by CcP-W2, CcP-W1 and then WT CcP. Interestingly, the initial reactivity of CcP-W2 is higher than CcP-W1, followed by a shift in relative oxidation rates after approximately one minute. This behaviour was observed for both the dye and VA oxidation assays. One possible explanation is a loss of reactivity of the surface Trp. It should be noted that our CcP mutants lack the negative microenvironment, that is thought to stabilize the radical cation species formed (**Figure 2-12, Step II**).

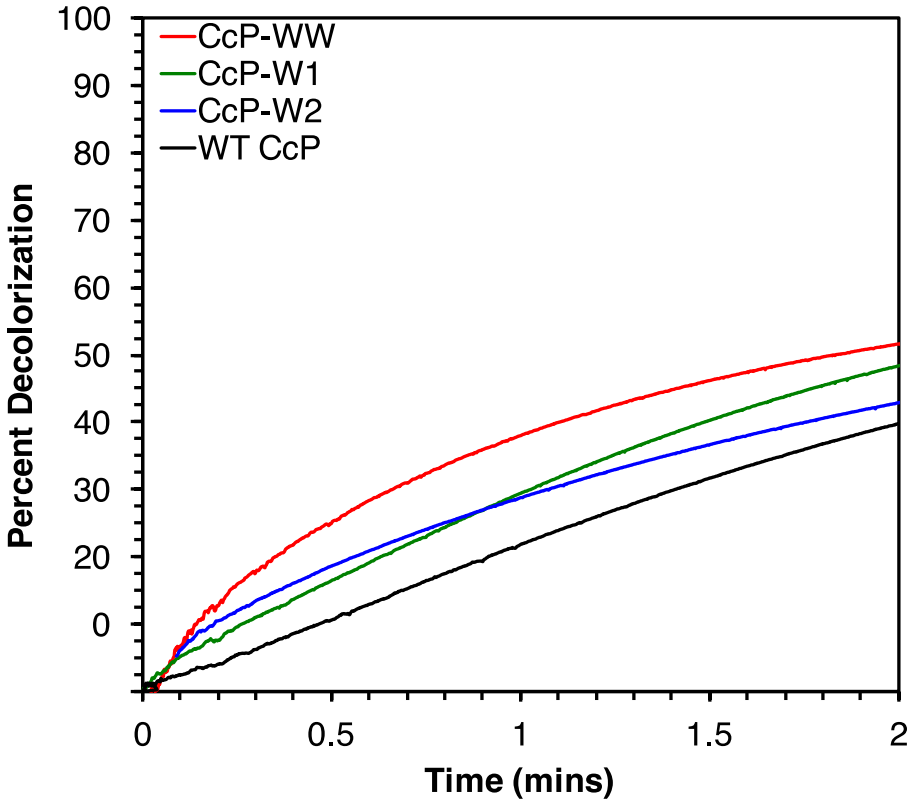

Figure 2-16. Dye decolorization of 4-p-nitrophenylazoresorcinol by monitoring the absorption at 466 nm. WT CcP (black trace), CcP-W2 (blue trace), CcP-W1 (green trace), and CcP-WW (red trace). Reaction mixtures contained 400 µM H_2O_2, 10 µM protein, and 41.8 µM 4-p-nitrophenylazoresorcinol at pH 8.

To test for covalent modification of Trp193 and/or Trp229, a reaction mixture of CcP-WW was concentrated using 3 kDa molecular weight cutoff filter and MALDI spectra were collected. The MALDI **(Appendix A:** Figure A.2**)** shows a shift in the MW (approximately 1065 Da) of CcP-WW after the oxidation assays, suggesting covalent modification of the protein, possible via attachment of a reactive oxidized intermediate. The large peak width in the MALDI spectrum of reaction mixtures containing CcP-WW might contains several proteins modified at different levels. The mass change of 1065 Da, corresponds to modification by an average of six oxidized VA intermediates. However, the precise identity of the modified proteins still needs to be confirmed this experimentally. The deactivation of the surface Trp193, then prompts oxidation to occur from other Trp residues (Trp191). For example, for CcP-W2, initially electrons can transfer from Trp191 to Trp193 in a single step ET event. Following the deactivation of the surface Trp193,

subsequent ET originates from the embedded Trp191 residue, which appreciably slows down oxidation.

Evaluation of Michaelis-Menten kinetics parameters for oxidation of VA by the three CcP mutants were attempted at pH 8. However, when varying concentrations of VA we observed little change in the initial velocity of the CcP mutant. The initial rates changed by less than 5% between 50 µM and 10 mM added VA. To check for association of VA with the CcP active site, UV-Visible spectra were collected at 2 mM VA concentrations. In all cases the spectra of the CcP mutants did not exhibit measurable changes in the heme Soret peak after this incubation, suggesting that VA is not binding near, or inside, the pocket. Future work to probe oxidation reaction dynamics should include other known substrates used for LiP. For example, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) oxidation, which is often used as an assay for peroxidase reactivity [46].

2.4. Discussion

Designing a tryptophan electron hopping chain in CcP

The approach used to move holes away from the active site of CcP included the introduction of strategically placed Trp residues near the protein surface, and within 15 Å of redox active Trp191. As discussed in this chapter, I designed three mutants (CcP-WW, CcP-W1, CcP-W2) to test reactivity. I found that our CcP mutants did demonstrate different reactivity distinct from WT CcP in oxidation assays for two different substrates. As discussed above, the reduction potentials of Trp residues are pH dependent (**Table 2-2**). For electron hopping in any system to occur, the difference in the reduction potentials of all cofactors should be no more than 200 meV [47]. Therefore, as the reduction potential decreases as pH increase, hopping becomes more likely.

The reported activity of LiP varies greatly and depends on the method of protein isolation. Activity ranges anywhere from 2.2 to 31 U mg^{-1} to produce 1 µmol veratryl aldehyde min^{-1} [48,49,50,51]. Comparing our CcP mutants (CcP-W1, CcP-W2, CcP-WW) to the low end of LiP activity, CcPWW is about 50 times slower than LiP. While the newly installed surface Trp residues enhances oxidation reactions with two organic substrates, they do not afford a highly efficient enzyme.

Table 2-3. Data obtained from the VA assay at a pH 8 with all CcP variants

Name	Initial slope rate (M/s)	Veratryl aldehyde formed in 1 min per mg of enzyme
CcPWT	1.5E-7	0.0174 μmol
CcP-WW	5.5E-7	0.0560 μmol
CcP-W1	2.4E-7	0.0291 μmol
CcP-W2	3.5e-7	0.0360 μmol

CcP-WW has the highest reactivity towards oxidizing aromatic organic molecules at the surface of the protein. **Figure 2-13** shows the PyMOL rendering of CcP-WW showing two potential ET pathways. Following the formation of a radical at Trp191, the electron can transfer directly from the A193W to the radical (green arrow), or the electron can be transferred between two individually shorter distances (red arrows). ET over long distances can occur much faster when multiple shorter steps are involved, so called 'hopping.' Workers have used semiclassical (Marcus) theory to construct hopping maps that graphically demonstrate the 'hopping advantage' for multistep processes [47].

The VA and azo dye oxidation assays suggest that multi-step ET (substrate → surface Trp → Trp191 [→ heme]) is preferred, as opposed to the single step ET (substrate → Trp191). Based on initial rates, CcP-WW mutant oxidizes substrates faster than CcPW2. CcPW2 lacks one redox site, Trp229, and therefore reactions are limited to single step ET between Trp191 and Trp193. The consequence of this limitation is that ET occurs on a slower timescale. In contrast to the modest reactivity for my CcP mutants, LiP is very efficient at oxidizing large organic compounds via a surface radical at Trp171. As noted above, the negatively charged microenvironment of Trp171 is thought to stabilize the Trp radical and oxidized substrate products, such as VA. **Figure 2-13** shows the proposed two-electron oxidation process of VA. The lack of the negative surface patch can result in loss of stability of the intermediate labeled II (**Figure 2-11**). Recent work on *Coprinus cinereus* peroxidase illustrated the importance of the negative surface patch [16]. To create LiP catalytic reactivity, the negative environment was imperative, and was achieved by producing a triple mutant, D179W + R258E + R272D, which increased substrate oxidation reactivity. The protein microenvironment of CcP around the site of our mutations (**Figure 2-17**) shows a distinct lack of negative residues (compared to **Figure 1-7**). It is

likely that this is a key contributor to the modest oxidation activity of our CcP mutants. Future work could include generating mutants that have a negative microenvironment, like that found in LiP.

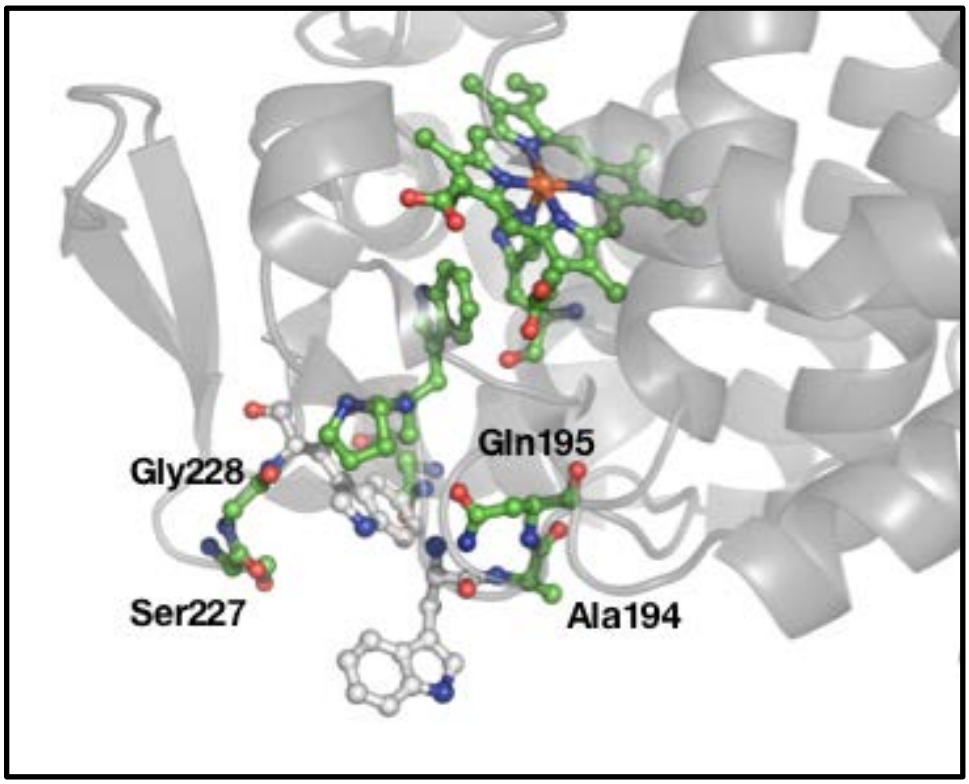

Figure 2-17. PyMOL rendering of the protein microenvironment of the CcP-WW mutant. Grey AAs represent the mutated residues. Surface Trp not surrounded by a negative surface environment such as LiP.

2.5. Conclusions

The work presented in this chapter describes two approaches for introducing Trp radicals at the surface of peroxidase enzymes. First, I described attempts to introduce Trp radicals by directly mimicking LiP structures, as outlined in **Sec 2.1.1.** This approach proved to be unsuccessful due to the formation of inclusion bodies during protein expression. These inclusion bodies could not be refolded to yield active proteins and this strategy was abandoned. Second, I designed a Trp redox chain in CcP with two simple mutations; this work conceptually mimicked LiP activity by providing an ET route from the heme cofactor to the enzyme surface. I evaluated our Trp chain by characterizing, and

investigating the reactivity of three CcP mutants: CcP-WW, CcP-W1, and CcP-W2. The reactivity of the designed Trp chains is pH dependent, with notable differences in catalytic oxidation activity toward organic substrates. In all cases, catalytic oxidations were only observed at pH 8. Based on initial rates, the order of increasing catalytic reactivity for CcP mutants was, CcP-WW > CcP-W2 > CcP-W1 > WT CcP. I postulate that oxidations initiated by H_2O_2 in the CcP-WW are following a multistep ET pathway, as opposed to the other mutants which react more sluggishly, suggesting a different ET pathway. Compared to native LiP, the reactivity of CcP-WW is still 50-fold lower, so many improvements can still be made.

2.6. References

1. Dunford, H. B. Peroxidases and catalases: Biochemistry, biophysics, biotechnology, and physiology. 2010 Hoboken, N.J: John Wiley & Sons
2. Potasek, M *Science*. **1978**, *201*, 151–153.
3. Finzel, B.C, Poulos, T L, Kraut, *J. Biol. Chem*. **1984**, *259*, 13027–13036.
4. Warren, J. J.; Winkler, J. R.; Gray, H. B. FEBS Lett. **2011**, 586, 596–602.
5. Faraggi, M.; Defelippis, M. R.; Klapper, M. H. *J. Am. Chem. Soc*. **1989**, 111, 5141–5145
6. Miller, M. A.; Vitello, L.; Erman, J. E. *Biochemistry* **1995**, *34*, 12048–12058.
7. Foshay, M.; Vitello, L.B.; Erman J.E.; *Biochim. Biophys. Acta* **2011**, *5*, 525-535
8. Stubbe, J.; A. van der Donk W.; *Chem. Rev.* **1998**, *98*, 705-762
9. Miller, M.A.; Han, G.W.; Kraut, J. *Proc. Natl. Acad. Sci. U.S.A*. 1994, *91*, 11118-111122
10. Fishel, L. A.; Farnum, M. F.; Mauro, J. M.; Miller, M. A.; Kraut, J.; Liu, Y.; Tan, X. L.; Scholes, C. P. Biochemistry **1991**, *30*, 1986–1996
11. Miller, M.; Han, G.; Kraut, J. A. *Proc. Natl. Acad. Sci. U.S.A*. **1994**, 91, 11118-11122.
12. Goodin, D.; Mcree, D. Biochemistry, **1993**, *32*, 3313–3324
13. Blodig, W., Doyle, W. A., Smith, A. T., Winterhalter, K., Choinowski, T., Piontek, K. *Biochemistry* **1988**, 37, 8832-8838
14. Martinez, D.; Larrondo, L. F.; Putnam, N.; Gelpke, M. D. S.; Huang, K.; Chapman, J.; Helfenbein, K. G.; Ramaiya, P.; Detter, J. C.; Larimer, F.; Coutinho, P. M.; Henrissat, B.; Berka, R.; Cullen, D.; Rokhsar, D. *Nat. Biotechnol*. **2004**, *22*, 899–899
15. Doyle, W. A.; Blodig, W.; Veitch, N. C.; Piontek, K.; Smith, A. T. *Biochemistry* **1998**, *37*, 15097–15105.
16. Smith, A. T., Doyle, W. A., Dorlet, P., Ivancich, A., & Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A*. **2009,** *106*, 16084-16089.
17. Goodin, D. B.; Davidson, M. G.; Roe, J. A.; Mauk, A. G.; Smith, M. *Biochemistry* **1991**, *30*, 4953–4962.
18. Mauro, J. M., Fishel, L. A., Hazzard, J. T., Meyer, T. E., Tollin, G., Cusanovich, . A., & Kraut, J. *Biochemistry* **1988**, *27*, 6243-6256.
19. Smith AT, Doyle WA *Engineered peroxidases with veratryl alcohol oxidase activity* Patent **2006** (International) WO/2006-11461.
20. Bhaskar, B., Bonagura, C. A., Li, H., & Poulos, T. L. *Biochemistry* **2002**, *41*, 2684
21. Sutherland, G. R., & Aust, S. D. *Biochemistry* **1997**, *36*, 8567-8573.
22. Sutherland, G. R., Zapanta, L. S., Tien, M., & Aust, S. D. *Biochemistry* **1997**, *36*, 3654-3662.
23. Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. U.S.A.*. **2015**, *112*, 10920–10925.

24. Gray, H.B.; Winkler, J.R. *Proc. Natl. cad. Sci. U.S.A* **2015**, *112*, 10920-10925

25. Polizzi, N. F.; Migliore, A.; Therien, M. J.; Beratan, D. N. *Proc. Natl. Acad. Sci. U.S.A.*. **2015**,*112*, 10821–10822.

26. Reece, S.Y.; Hodgkiss, J.M.; Stubbe, J.; Nocera, D. G. *Phil. Trans. R. Soc. B.* **2006**, *361*, 1351-1364

27. Goodin, D. B.; Davidson, M. G.; Roe, J. A.; Mauk, A. G.; Smith, M. *Biochemistry* **1991**, *30*, 4953–4962.

28. Berry, E.A.; Trumpower B.L. *Anal. Biochem*. **1987**, *161*, 1-15

29. ExPASy-Compute Pi/MW tool http://web.expasy.org/compute_pi/ (accessed May 20th 2016)

30. Doyle, W.; Blodig, W.; Veitch, N.; Piontek, K.;Smith, A.; *Biochemistry*, **1998**,*37*, 15097-15105

31. Oganesyan, N., Kim, S., & Kim, R. *J. Struct. Funct. Genomics*, **2005**, 6, 177-182

32. Dunford, H. B. Peroxidases and catalases: Biochemistry, biophysics, biotechnology, and physiology (2nd ed.). **2010** Hoboken, N.J: John Wiley & Sons.

33. Volkov, A.N; Wohlkonig, A.; Soror, S.H.;Van Nuland, N.A; *Biochemistry*,**2013**, 52, 2165-2175

34. Gill, S. C.; Hippel, P. H. V. *Anal. Biochem*.**1989**, *182*, 319–326

35. Yonetani, T.; Ray, G.S.; *J. Biol. Chem.* **1965**,*240*, 4503-4508

36. Iizuka, T.; Kotani, M, Ray, G.S.; *J. Biol. Chem.* **1971**,*246*, 4731-4736

37. Greenfield, N. J. *Nat. Protoc*. **2007**, *1*, 2876–2890.

38. Sievers, G. *Biochim. Biophys. Acta* **1978**, *536*, 212–225.

39. Blodig, W.; Doyle, W. A.; Smith, A. T.; Winterhalter, K.; Choinowski, T.; Piontek, K. *Biochemistry* **1998**, *37*, 8832–8838.

40. Choinowski, T.; Blodig, W.; Winterhalter, K. H.; Piontek, K. *J. Mol. Biol.* **1999**, *286*, 809–827

41. Loo, S.; Erman, J. E. *Biochemistry* **1975**, *14*, 3467–3470.

42. Harriman, A.; *J. Phys. Chem*. **1987**, *91*, 6102-6104

43. Díaz Torres, E., Ayala, M., SpringerLink *Biocatalysis based on heme peroxidases: Peroxidases as potential industrial biocatalysts*. **2010** Heidelberg: Springer-Verlag

44. Young, L. *Water Res.* **1997,** *31*, 1187–1193.

45. Ollikka, P.; Alhonäki, K.; Leppänen, V.M.; Glumoff, T.; Raijola, T.; Suominen, I.; *Appl. Environ Microbiol***. 1993**,59, 4010–4016.

46. Miki, Y.; Calviño, F. R.; Pogni, R.; Giansanti, S.; Ruiz-Dueñas, F. J.; Martínez, M. J.; Basosi, R.; Romero, A.; Martínez, A. T. J. Biol. Chem. **2011**, *286*, 15525–15534

47. Warren, J. J.; Ener, M. E.; Vlček, A.; Winkler, J. R.; Gray, H. B. *Coord. Chem. Rev.* **2012**, *256*, 2478–2487.

48. Johansson, T.; Welinder, K.; Nyman, P. Arch. Biochem. Biophys. **1993**, *300*, 57–62.

49. Paszczynski, A., Huynh, V.B, Crawford, R. *Arch. Biochem. Biophys*. **1986**, *244*, 750-765

50. Rennaganathan, V., Miki, K., Gold, M.H. *Arch Biochem. Biophys*. **1985**, 241,304-314

51. Farrell, R.L, Murtagh, K.E., Tien, M., Mozuch, M.D., Kirk, T.K *Enzyme Microb. Technol.* **1989**, *11*, 322-328

Chapter 3.

Photocatalytic production of alkanes using Ru-modified cyanobacterial aldehyde deformylating oxygenases

First reported in 2010 [1], cyanobacterial aldehyde deformylating oxygenase (cAdo) enzymes (**Figure 3-1**) are a new addition to the non-heme diiron 4-helix bundle protein superfamily. These cAdo enzymes attracted a great deal of attention due their potential use in the production of biofuels. The biological function of cAdo are still unknown, although the function of related enzymes in plants (to produce waterproof barriers) [2] and in insects (in the biosynthesis of pheromones) [3] are known. Because of the comparative ease of expression of bacterial cAdo proteins, with respect to related plant and insect enzymes, research into the mechanism of aldehyde decarbonylation by cAdo has proceeded rapidly. Here, I describe work done to better understand electron flow in cAdo.

The production of alkanes in cyanobacteria requires two enzymes. First fatty-acyl-CoA esters are converted to fatty aldehydes by an acyl-ACP reductase [4], followed by the conversion of the C_n aldehyde to the corresponding C_{n-1} alkane by cAdo **(Eqn.1**). Though formally redox neutral, the overall decarbonylation reaction is coupled to the $4H^+/4e^-$ reduction of O_2 to H_2O [5]. Remarkably, cAdo can be integrated into synthetic metabolic pathways to produce alkanes of various chain lengths [6,7]. However, these enzymatic systems are limited due to their slow turnover (~ 1 min^{-1}) [8], as assayed using either chemical [9] or protein [10] reducing systems. Slow turnover has typically been attributed to poor electronic coupling between the diiron active site and the reductant. Though the natural reductase for cAdo remains unknown, a recent report of a cyanobacterial [2Fe-2S] ferredoxin has demonstrated rapid reduction of the di-ferric active site [11]. These research efforts highlight ongoing interest in mechanistic investigation and optimization of electron transfer (ET) to the diiron active site of cAdo.

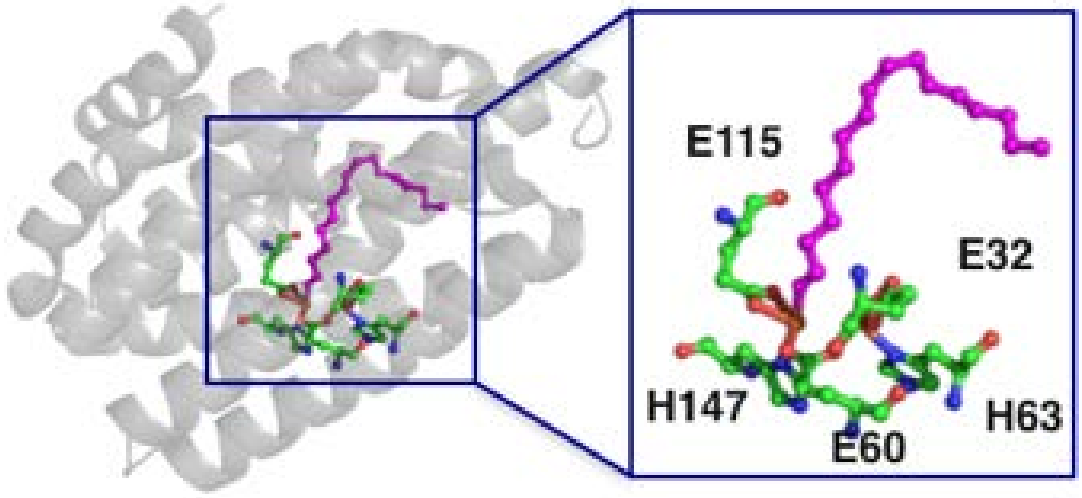

Figure 3-1. Cyanobacterial aldehyde deformylating oxygenase. Insert: diiron active site with ligated amino acids and aldehyde substrate outlining the substrate-binding pocket (pink). Image created in PyMOL. PDB ID: 4RC5)

$$RCH_2CHO + O_2 + 4H^+ + 4e^- \rightarrow RCH_3 + HCOOH + H_2O \qquad \text{(Eqn. 1)}$$

To better couple ET between a reductant and the embedded cAdo diiron site, I integrated a Ru-photosensitizer to the surface of cAdo. Ruthenium photosensitizers have routinely been used to study ET in many enzymes [12,13,14]. Attachment of a photosensitizer requires the presence of a solvent exposed cysteine or histidine residue. Here, Cys is preferred due to the low number of exposed Cys residues in cAdo. To label the protein, a Cys-reactive Ru(tris-diimine) complex was used (4-bromomethyl-4'-methyl-2,2´-bipyridyl)(2,2'-bipyridyl)$_2$Ru^{II}(PF$_6$)$_2$, as shown in **Figure 3-3** [15]. Analysis of the X-ray structure reveals one surface cysteine (Cys70) adjacent to E115 in the active site, as illustrated in **Figure 3-2 (A).** Further, two mutations, R62C and C70S, produce a mutant with one solvent exposed Cys residue (**Figure 3-2 (B)**). Finally, a mutation to R62C introduces an additional Cys along the same peptide segment as Fe-coordinating H63, giving a total of two solvent exposed Cys residues (**Figure 3-2 (C)**). Overall, I designed three different Ru-cAdo models and investigated their catalytic reactivity for aliphatic aldehyde deformylation.

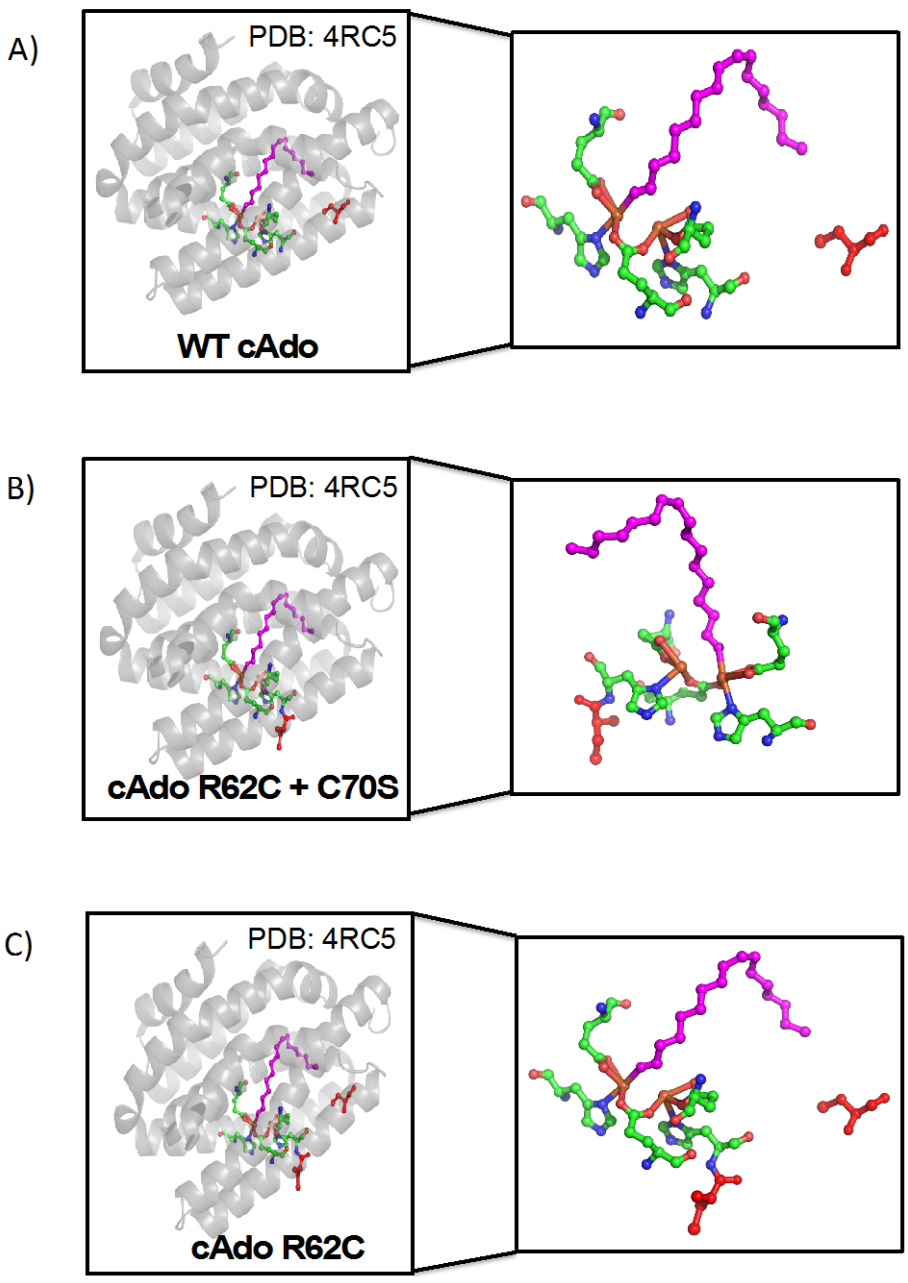

Figure 3-2. **Labeling and active site orientations in Ru-cAdo photochemical systems: (A) WT-cAdo has cysteine residue (C70) (red) near the active site; (B) introduction of two mutations R62C and C70S, results in one cysteine residue (red) adjacent to H63 in active site and (C) introduction of a single mutant at R62C produces two solvent exposed Cys (red). The magenta compound in all figures is crystallographically defined bound substrate.**

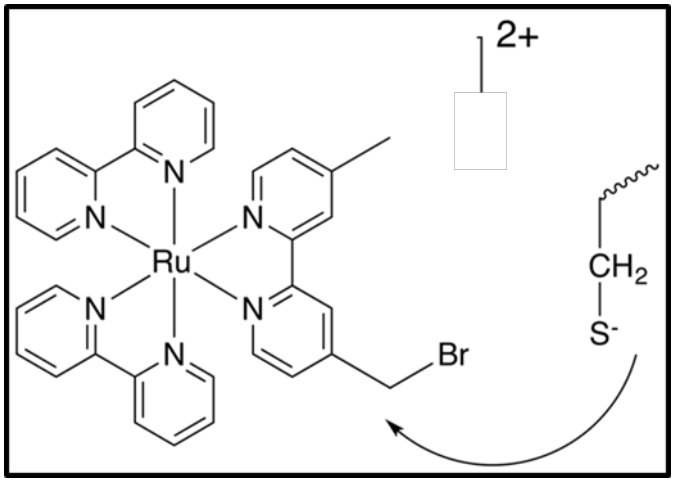

Figure 3-3. Cys-reactive brominated Ru(tris-diimine) complex used (4-bromomethyl-4´-methyl-2,2´-bipyridyl)(2,2´-bipyridyl)$_2$Ru^{II}(PF$_6$)$_2$

I used the photochemical flash-quench system outlined in **Figure 3-4** to reduce cAdo. A visible light source (e.g., a halogen lamp) is required to excite the Ru(II) to a long-lived electronically excited state, *Ru(II). In the presence of a quencher, *Ru(II) can be reduced to (formally) Ru(I) [16]. In this process, sodium diethyldithiocarbamate (DTC) was used to quench Ru(II)*. The high water solubility and low cost of DTC make it an ideal sacrificial reductant for our photochemical systems [17]. The flash-quench generated Ru(I) is a strong reductant, with $E°$ = –1.26 V vs. NHE [18]. This potential is much more reducing than that of any chemical or biochemical-reducing systems commonly used with cAdo, providing a valuable test of whether or not low driving force ET is limiting the rate of cAdo catalysis. This also makes it an effective method to shuttle electrons to the active site. A total of four cycles from Ru(I) to Ru(II) are required to complete one cAdo cycle. In this chapter, I describe the characterization and implementation of photochemical Ru-cAdo systems.

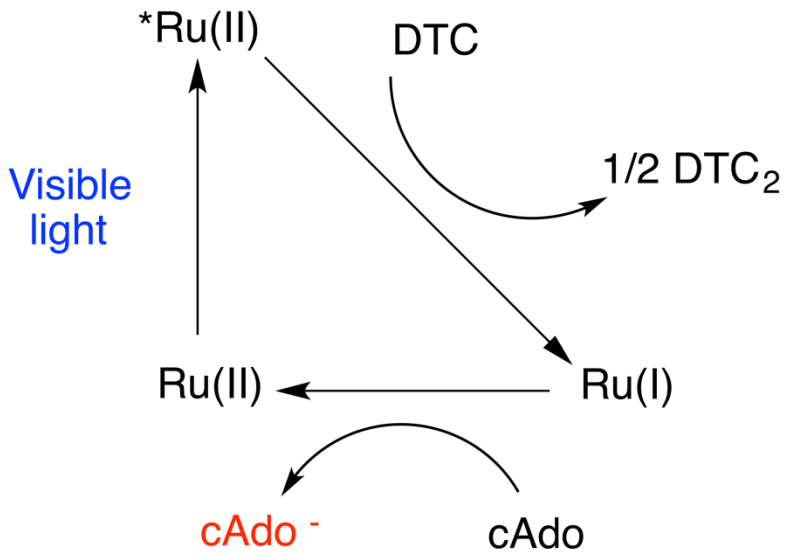

Figure 3-4. Proposed photochemical system for shuttling electrons to the active site of cAdo. Upon irradiation from a visible light source, electronically excited *Ru(II) is produced; in the presence of an excess amount of sacrificial reductant, DTC, we can reduce *Ru(II) to Ru(I), which in turn shuttle electrons to the diiron active site

3.1. Materials and Methods

3.1.1. Protein expression and purification

Site directed mutagenesis

Plasmids encoding the Synechococcus *elongatus* cyanobacterial aldehyde deformylating oxygenase gene were a gift from Prof. David Baker (University of Washington). The gene was cloned into pET29(b+) between NdeI and Xho1 restriction sites, which included a polyhistidine (His_6) tag at the C-terminus. All cAdo (R62C, C116S + C70S) mutants were prepared by PCR using primers purchased from Eurofins Operon **(Appendix B).** Site directed mutant plasmids were obtained with the Quickchange protocol (Agilent Technology). DNA polymerase (Q5) and DpnI were obtained from New England Biolabs. PCR mixtures were transformed with chemically competent DH5α *E. coli* and plated on kanamycin-supplemented Luria broth agar plates. Single colonies were grown overnight. Plasmid DNA was isolated from overnight cultures according to a Qiagen Spin MiniPrep kit and sequenced by Eurofins Operon simple sequence read service to verify mutations.

Protein expression

Standard transformation protocols were followed using *E. coli* BL21(DE3) on kanamycin-supplemented agar plates (50µg/mL). A single colony was incubated overnight (37℃, 180 rpm shaking) in 25 mL Luria broth (LB) supplemented with 50 µg/mL kanamycin. Large-scale expression was initiated by inoculating 1 L of terrific broth (TB) containing 50µg/mL kanamycin with 1 mL of the starter culture. Cultures were grown to an OD_{600} of approximately 0.6-08. IPTG was then added to a final concentration of 500 μM and grown for an additional 5 hours. Cells were harvested by centrifugation for 30 minutes at 5000 rpm.

Purification and labeling procedure

Cell pellets were resuspended in 50 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole, and 1 mM PMSF buffer (pH 8). The resulting suspension was lysed by sonication on ice for 15 minutes and the supernatant was collected by centrifugation for 30 minutes at 14000 rpm (4℃). The pH of the supernatant was adjusted to pH 8 and then loaded directly onto a Ni-NTA affinity column. The column was washed with at least 10 column volumes of buffer containing 20 mM sodium phosphate, 5 mM imidazole, and 500 mM NaCl buffer (pH 8). Increasing the concentration of imidazole to 250 mM eluted the protein of interest. The buffer was exchanged to 20 mM Tris•HCl buffer (pH 8) containing 20 mM DTT and incubated at room temperature for at least 45 minutes. DTT was removed by desalting on a PD-10 desalting column.

To prepare labeled cAdo, varying concentrations of the Ru-photosensitizer ((4-bromomethyl-4′-methyl-2,2′-bipyridyl)(2,2′-bipyridyl)$_2$Ru^{II}(PF_6)$_2$) were added to protein mixtures ranging from 1.2-2.4 equivalents. The Ru-photosensitizer is added to 60 μM protein solution in 20 mM Tris (pH 8) buffer and stirred for at least 4 hours in the dark at 4℃; the longest reactions times were overnight (12 hours). Protein solutions were then exchanged into 5 mM HEPES buffer (pH 7.6) and loaded onto a FPLC DEAE Hi-Prep column and purified by increasing salt concentration (Buffer B: 5mM HEPES + 500 mM NaCl pH 7.6). Fractions containing labeled protein were pooled and exchanged into a 50 mM sodium phosphate buffer solution with 100 mM NaCl (pH 7.5). Protein samples were stored in the dark at 4℃ until needed.

3.1.2. Characterization of Ru-cAdo Models

Optical spectroscopy

The UV-Visible spectra of the proteins were obtained with either a Photon Control SPM-002-EH CCD with a SPLC 1DH deuterium/tungsten light source, or a Cary 100-Bio UV-visible spectrometer. The optical spectra for all proteins, including Ru-modified proteins, were collected in 50 mM sodium phosphate, 100 mM NaCl buffer (pH 7.5).

Mass spectrometry: MALDI Characterization

MALDI was routinely used to confirm the identity and purity of Apo protein and Ru-labeled proteins. A Bruker microFLEX MALDI-TOF instrument with a 48 well ground steel target was used for all MS experiments. The dried droplet method was used with sinapinic acid as the matrix for all samples. Spectra are provided in the **Appendix B**. Theoretical molecular weights are obtained from the protein sequence using ExPasy: Compete Pi/MW tool [19].

Circular Dichroism

Circular dichroism spectra were collected using a Chirascan qCD spectrometer scanning from 200 nm to 250 nm. A total of five scans were taken and averaged with a step size of 1 nm and 0.5 seconds per point. Raw data was converted using the following formula $[\theta] = \theta M_r/(nCl)$ where θ, M_r, n, C, and l represent ellipticity in millidegrees, molecular weight (Da), number of residues, concentration in mg/ml and the cuvette path length respectively. Values are reported as (deg^2• cm•mol^{-1}) against wavelength (nm). All scans were collected in 50 mM sodium phosphate buffer, 100 mM NaCl, pH 7.5.

Steady state fluorescence

All fluorescence spectra were collected on Jobin-Yvon FluoroLog Spectrofluorometer. Samples were deoxygenated with several pump-purge cycles with argon. Deoxygenated samples were excited at 455 nm and scanned from 480 to 800 nm. Each sample contained 10 μM of Ru-modified protein in 50 mM sodium phosphate, 100 mM NaCl buffer (pH 7.5). The buffer composition was identical for scans containing 10 mM DTC quencher.

Trypsin digestion

A trypsin digestion carried out for Ru-R62C-cAdo and Ru-WT-cAdo along with the unlabeled proteins as controls. The enzyme was desalted into 100 mM ammonium bicarbonate buffer (pH 8) using a PD10 desalting column. Trypsin from bovine pancreas was dissolved in 1 mM HCl to a concentration of 1 mg/mL. The trypsin was then added directly to the protein solution to give a final trypsin: substrate ratio of 1:20. The digestion was incubated for 36 hours at 37℃. The solution was then directly spotted onto a 48 well MALDI stainless steel target using the dried droplet method, with 2,5-dihydroxybenzoic acid as the matrix. To calculate the cleavage fragments from the trypsin digest an online calculator, Peptidecutter [20].

Photochemical reaction conditions

To determine ideal reaction conditions, the concentration of the quencher, Ru-cAdo concentration, time, and temperature were evaluated. Optimal conditions were found to be 10 μM of cAdo, 10 mM DTC in 50 mM sodium phosphate, 100 mM NaCl buffer (pH 7.5). For all reactions, half of a master reaction mixture was placed in the dark as a control. The other half of the reaction mixture was exposed to visible light from a KODAK extra bright lamp projector source, stirring in a water bath at room temperature. The reactivity of a range of long-chain aldehydes were assessed under saturating conditions of substrate.

Proton NMR

All NMR experiments were run on the Bruker AVANCE III 500 MHz running IconNMR under TopSpin 2.1. The Simpson water suppression pulse program was used to analyze the products of all photochemical reactions. In addition to reaction conditions described above, NMR samples were run with 10% added D_2O. For NMR samples, the reaction solution was centrifuged for 5 minutes at 16000 g to remove insoluble oxidized DTC. The solution was then transferred to NMR tubes (5 mm diameter) and NMR spectra were collected.

Alkane detection using GC-MS

Alkane formation was quantified through GC-MS (Agilent 6890 series with 5973 Network Mass Selective Detector). A capillary DB-5 column (0.25mm x 30m x 0.25 μm) was used to separate hydrocarbons. Polypropylene Eppendorf microcentrifuge tubes were used to extract decane from the reaction mixture. Photochemical reactions (1000 μL) were extracted with ethyl acetate (500 μL). The mixture was vortexed for 2 minutes and centrifuged for 30 minutes at 16000 g. A 150 μL aliquot of the organic fraction was extracted and dodecane was added as an internal standard (5 μM).

For GC-MS experiments, 1 μL was injected using a 2:1 split ratio with a flow rate of 3 mL min^{-1}. Hydrogen gas was used as the carrier gas with a flow rate of 1.5 mL min^{-1}. Upon injection, the oven temperature was set to 70℃ and the inlet temperature was set at 280℃. The oven temperature was held at 70℃ for 3 minutes, and then ramped to 145℃ at 20℃ min^{-1}, ending with a ramp to 280 ℃ at 30 ℃ min^{-1}. To detect decane peaks, we used SIM mode; selected ions include 43,57,85,142 m/z. A single point calibration (100 μM decane and 5 μM dodecane) was used to determine decane concentrations.

Control reactions

The activity of heterogeneously expressed protein was confirmed by performing a standard activity assay with minor modifications [21]. A 500 μL solution containing 250 μM of oxidized cAdo protein in 50 mM sodium phosphate, 100 mM NaCl buffer (pH 7.5) with 10% D_2O was added to equal volume of 250 μM phenazine methosulfate (PMS) and sodium dithionite prepared in 50 mM sodium phosphate, 100 mM NaCl, buffer (pH 7.5) with 10% D_2O. Reaction mixtures were kept under argon atmosphere and monitored using ^{1}H NMR, as described above.

A number of controls were run in order to confirm the efficacy of the photochemical reactions. Controls with non-covalently attached tris(2, 2′-bipyridine)ruthenium(II) chloride ($Ru(bpy)_3^{2+}$) were investigated. Reaction mixtures contained 10 μM $Ru(bpy)_3^{2+}$, 10 μM WT-cAdo, with 10 mM DTC in 50 mM sodium phosphate, 100 mM NaCl, buffer (pH 7.5). Additionally, controls without DTC were investigated. These reaction mixtures contained 10 μM Ru-labeled cAdo in 50 mM sodium phosphate, 100 mM NaCl, buffer (pH 7.5). For all controls, half of the reaction mixture was placed in the dark as a control. The other half

of the reaction mixture was irradiated with visible light from a KODAK extra bright lamp projector while stirring in a water bath held at 22℃. Reactions were monitored using ^{1}H NMR as described above. All controls used octanal as the aldehyde substrate under saturating conditions of substrate.

3.2. Results

UV-visible spectroscopy

The UV-visible spectrum of the iron-loaded cAdo protein exhibits a weak shoulder at 350 nm indicative of ferric iron coordination in the active site (**Figure 3-5**, black trace). The Ru photosensitizer in sodium phosphate buffer (**Figure 3-5**, purple trace) has distinct absorption bands in the 400-500 nm region that are attributed to the metal-to-ligand charge transfer MLCT transition of Ru-bipryidyl complexes. Exposure to the Cys-reactive brominated Ru-photosensitizer results in Ru-modified WT-cAdo (Ru-Cys70) and related mutants (Ru-R62C/C116S/C70 cAdo and Ru-R62C-cAdo). FPLC purification removes unlabelled protein and provides purified Ru-labeled proteins. Distinct Ru MLCT transitions are observed for all three Ru-cAdo models. At a given concentration, the absorption bands in the visible region are more intense for Ru-R62C-cAdo, consistent with the attachment of two Ru photosensitizers to the protein. The absorption bands for Ru-WT-cAdo and Ru-R62C/C116S/C70S-cAdo in the visible region are similar due to the coordination of a single Ru-label on the surface (**Figure 3-5**, blue and red trace).

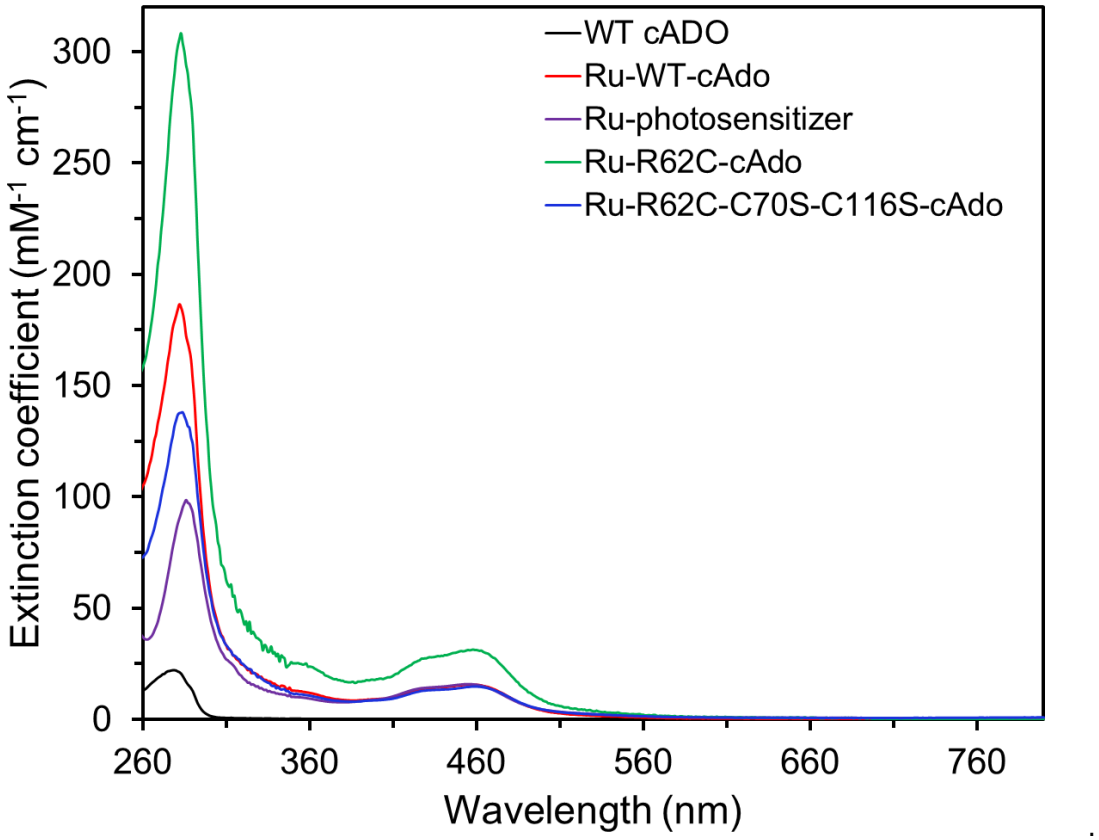

Figure 3-5. UV-visible spectra of Apo cAdo and labeled Ru-cAdo. WT-cAdo (black trace), Ru photosensitizer- ((4-bromomethyl-4´-methyl-2,2´-bipyridyl) (2,2´-bipyridyl)$_2$Ru^{II}(PF$_6$)$_2$ (purple trace), Ru-WT-cAdo (red trace), Ru-R62C/C116S/C70 cAdo (blue trace).

Mass spectroscopy: MALDI characterization

MALDI was used to confirm the identity WT-cAdo and mutants. **Table 3-1** shows the theoretical and experimentally obtained masses of purified mutants. MALDI was also used to confirm the attachment of the Ru-photosensitizer to the surface of the protein. The Ru-photosensitizer has a mass of 677 Da; upon coordination to the solvent exposed Cys residue the mass is reduced to 597 Da due to loss of the bromide leaving group. Therefore, a mass difference of approximately 597 Da is expected for each Ru-label attached to the protein. **Figure 3-6** shows the difference between unlabelled and labeled protein Ru-R62C-cAdo. The clear shift in the peaks indicates that two Ru-labels are attached to the protein.

Table 3-1. Theoretical and experimentally obtained mass of WT-cAdo and mutants with unlabeled and labeled protein.

Name	Theoretical (Da)	Experimental (Da)	Labeled: theoretical (Da)	Labeled: experimental theoretical (Da)
WT	28558	28558	29155	29160
R62C	28505	28504	29699	29709
R62C/C116S/C70S	28473	28482	29070	29036

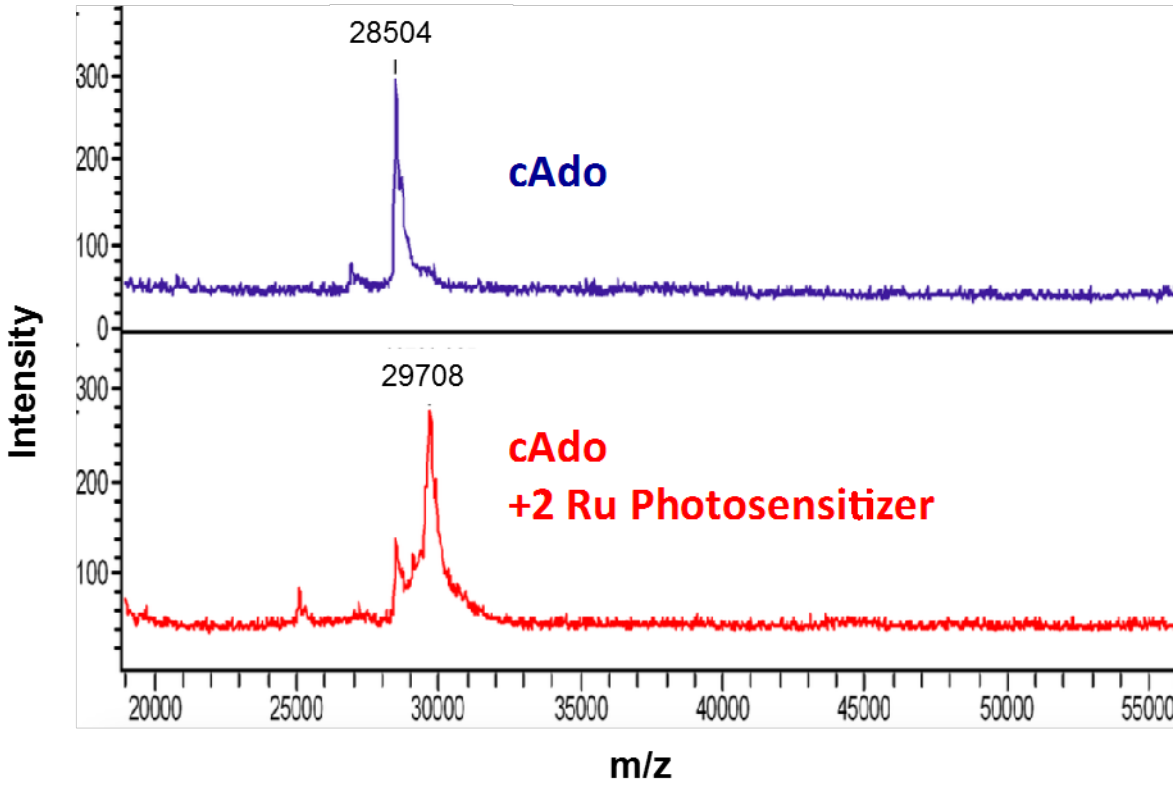

Figure 3-6. Stacked MS of R62C-cAdo unlabeled (top) and Ru-R62C-cAdo (bottom) collected from MALDI. The mass shift of approximately 1200 Da is indicative of two Ru-labels attached to the surface of the protein.

Circular dichroism

The far-UV CD spectra is useful for evaluating whether the secondary structure is conserved upon modification (mutations and/or labelling protein). The X-ray structure of cAdo (PDB: 4RC5) shows that the protein adopts an antiparallel four helix bundle

structure. **Figure 3-7** shows that WT-cAdo (black trace) has a CD spectrum indicative of an alpha helix secondary structure [22], in agreement with the reported structure. Far-UV CD spectra of R62C-cAdo (green trace), and R62C/C116S/C70S-cAdo (blue trace) show modest changes, and suggest that secondary structure is conserved. Additionally, the CD spectra of the Ru-labeled proteins were collected. **Figure 3-8**, shows that upon addition of the Ru-photosensitizer the protein exhibits similar CD spectra. This demonstrates that addition of the Ru-photosensitizer does not significantly perturb the protein's secondary structure.

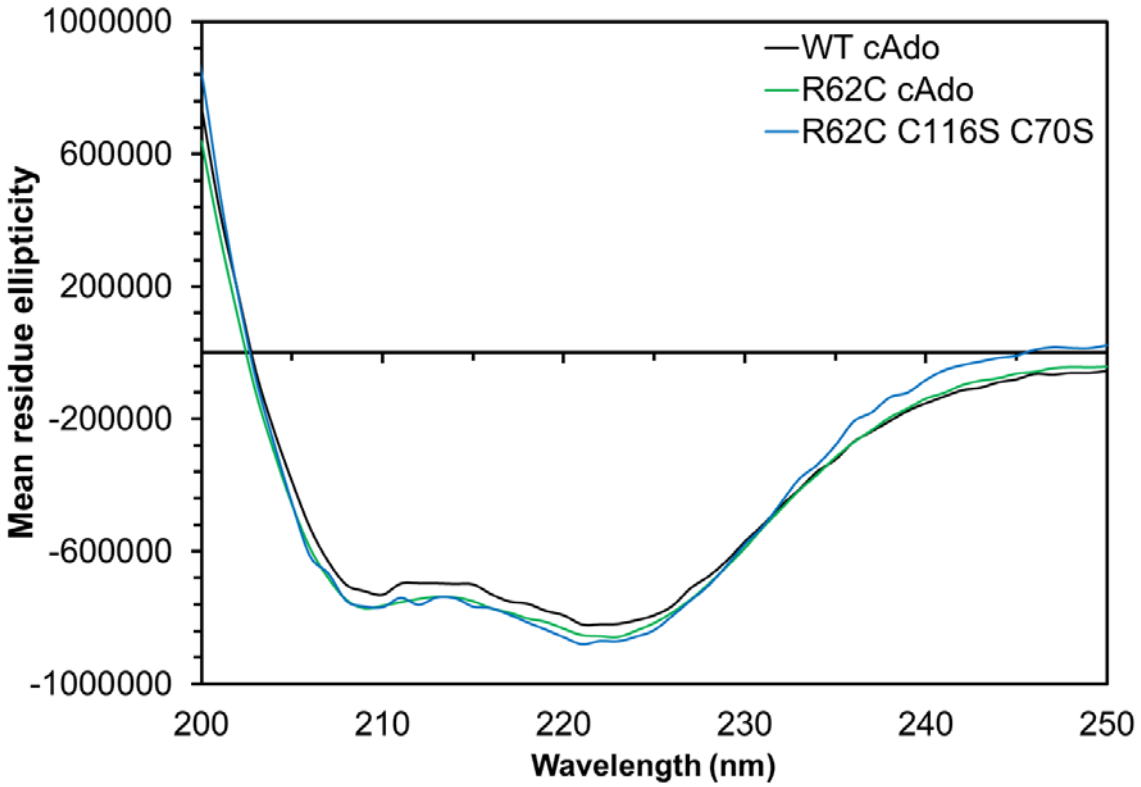

Figure 3-7. Far UV CD spectra in 50 mM phosphate and 100 mM NaCl (pH 7.5). CD spectra of unlabeled cAdo protein. WT-cAdo (black trace), R62C-cAdo (green trace) and R62C C116S C70S-cAdo (blue trace).

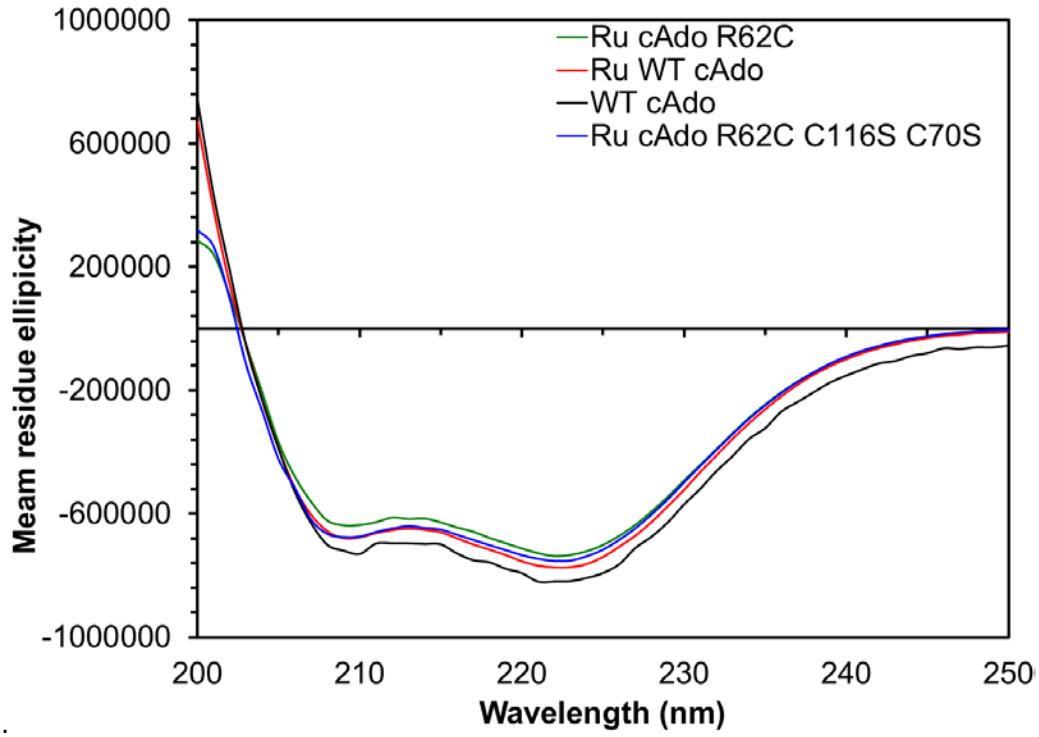

Figure 3-8. **Far UV CD spectra in 50 mM phosphate and 100 mM NaCl (pH 7.5). CD spectra of labelled cAdo protein WT-cAdo (Black trace), Ru- WT-cAdo (red trace) Ru-R62C-cAdo (green trace) and Ru-R62C C116S C70S-cAdo (blue trace).**

Steady-state fluorescence

Ruthenium-labeled proteins are known to exhibit a characteristic emission peak centered at approximately 600 nm [23]. Upon excitation of all Ru-cAdo proteins at 455 nm, a distinct emission peak is observed. **Figure 3-9** shows the normalized emission spectra observed for each Ru-modified mutant. As expected, the emission peak for Ru-R62C-cAdo (red trace) is about twice as strong as for the other Ru-labeled proteins due to the presence of two Ru-photosensitizers. Likewise, the emission spectra for Ru-WT-cAdo (green trace) and Ru-R62C/C116S/C70S-cAdo (blue trace) are less intense due to the coordination of only one Ru-label.

In addition, the emission spectra of the Ru-modified protein in the presence of DTC quencher (**Figure 3-10**) were collected. The emission of the Ru-labeled protein is reduced by more than 50% in the presence of quencher, indicating that DTC is capable of

quenching the excited *Ru^{II} species and confirming that it is a suitable quencher for our photochemical reactions.

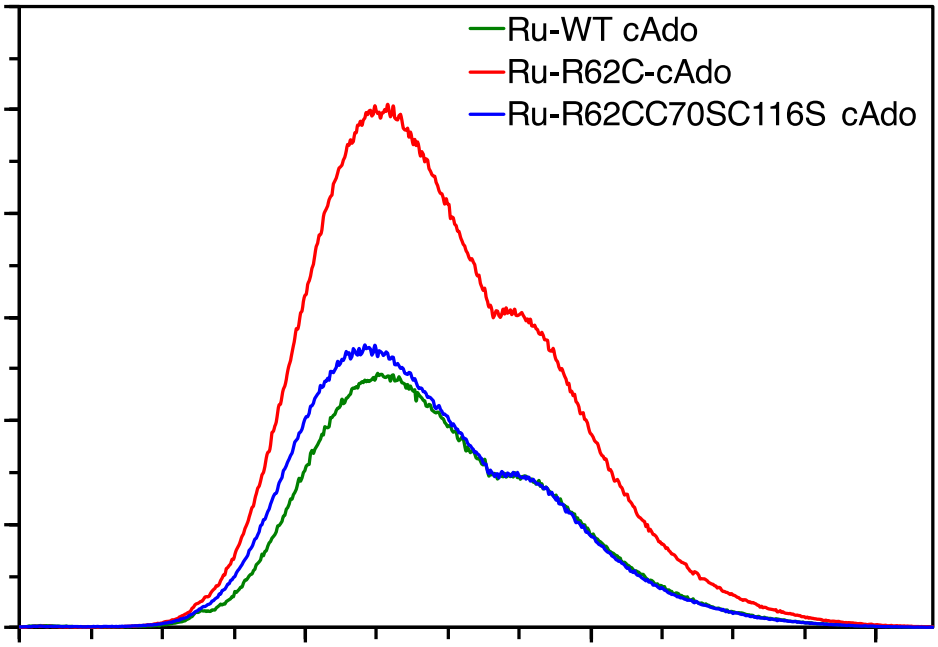

Figure 3-9. **Steady-state fluorescence spectra were generated with an excitation wavelength of 455 nm. The Ru-cAdo models were deoxygenated in 50 mM sodium phosphate and 100 mM NaCl buffer (pH 7.5). Ru-WT-cAdo (green trace), Ru-R62C-cAdo (red trace), and Ru-R62C C70S-C116S cAdo (blue trace).**

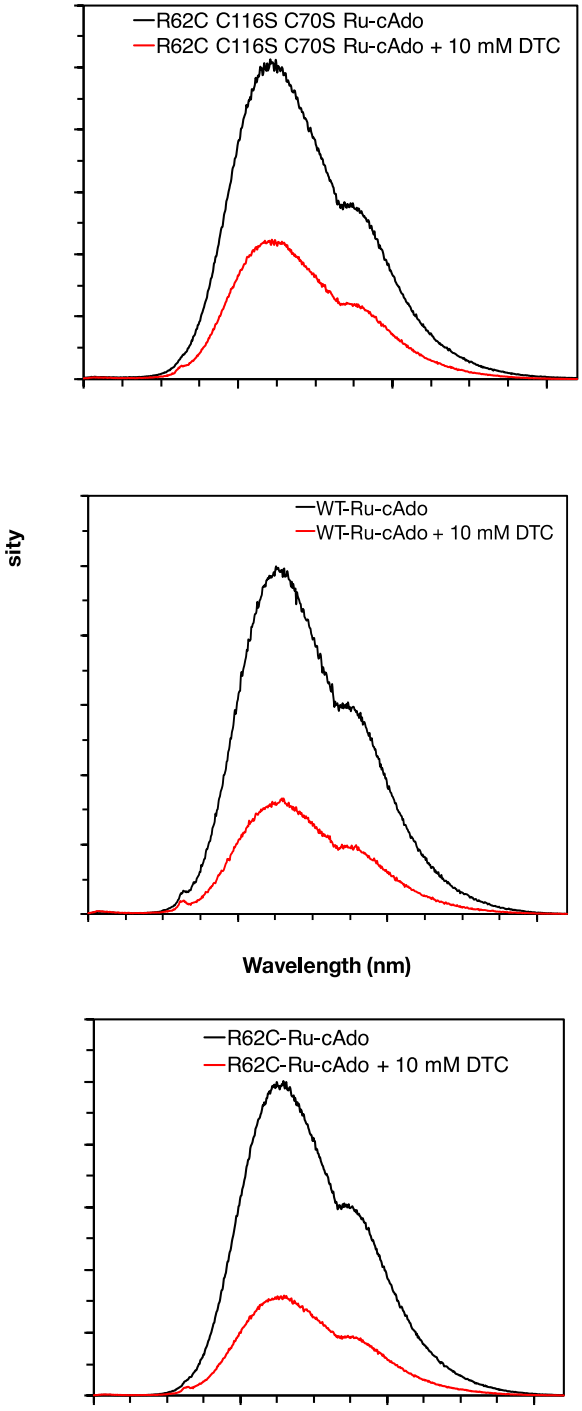

Figure 3-10. Steady state emission spectra of Ru-cAdo models in deoxygenated 50 mM sodium phosphate and 100 mM NaCl, pH 7.5 (black traces). Subsequent emission recorded under identical conditions in the presence of 10 mM DTC quencher (red traces).

Trypsin digestion

In order to confirm Ru-labelling at the desired Cys residues, proteins were digested using trypsin and the peptide fragments were analysed using mass spectrometry. Attachment of Ru-photosensitizers can be probed by comparing the peptide fragments of unlabeled cAdo against the labeled WT protein. The structure of the protein (PDB ID 4RC5) shows two solvent exposed Cys residues, C116 and C70, which are potential sites capable of reacting with the brominated Ru-label. The digested WT-cAdo, and Ru-WT-cAdo clearly identify the labeled residue as Cys70, due to the presence of a distinct peak at 1309 Da. **Table 3-2** shows that the peptide fragment containing Cys70 (713 Da) and the Ru-photosensitizer (597 Da) give rise to a peak at 1310 Da, confirming that Cys70 is the labeled residue. The same approach can be used to confirm that for C62 and C70 are the labeled residues for Ru-R62C-cAdo.

Table 3-2. Selected fragments of digested proteins from the Trypsin solution digestion

Name	Fragment position	Peptide fragment (Da)	Ru-label (Da)	Theoretical mass (Da)	Actual mass (Da)
Ru-WT-cAdo	65-71 GFMACGK	713	597	1310	1309
‘ Ru-R62C-cAdo	65-71 GFMACGK	713	597	1310	1308
	58-64 MEQCHMK	907	597	1504	1501

3.2.1. Photochemical reactions

Proton NMR

Proton NMR was used to evaluate the deformylation activity of our Ru-cAdo models. **Figure 3-11** sets out typical NMR spectra observed for the time course of the photochemical reactions. A peak at 8.3 ppm grows over the course of 40 minutes, corresponding to the production of formate. The assignment was confirmed by spiking with authentic sample. Additionally, the peaks at 1.1 and 3.9 ppm decrease over time. The loss of these signals is consistent with quenching of DTC and production of the sparingly soluble oxidized DTC dimer**.**

The absolute integral value of the DTC methyl or ethyl resonances provides a measure of DTC consumption. Comparison of light and dark samples suggests that ca. 6 mM DTC remains after 60 minutes of illumination. Comparison with the integral value for formate produced suggests that 50 ± 10 µM formate is produced during the reactions. Thus, much more DTC is consumed then is needed to satisfy the $4e^-$ stoichiometry required by cAdo.

In many cases, a side product with a resonance at 9.04 ppm is also observed. Such downfield shifts are most commonly observed for aldehyde C*H*O and for peroxide protons. Production of aliphatic peroxide side products are known for cAdo [24]. To test for this possibility, reaction mixtures were extracted with 2 mM triphenylphosphine (PPh_3) in $CDCl_3$ and analyzed using ^{31}P NMR. Dark samples showed a single resonance at -5.31 ppm, which is assigned to PPh_3. In the light samples, a second resonance is observed at 29.4 ppm, consistent with production of triphenylphosphine oxide ($OPPh_3$). The proton NMR of the aqueous layer was taken after extraction with 2 mM PPh_3 in $CDCl_3$. The peak at approximately 9.04 ppm disappeared indicating that it had been extracted with $CDCl_3$. This result strongly suggests that the production of peroxide (H_2O_2 or ROOH) side products can at least partially account for the discrepancy between DTC consumption and formate production.

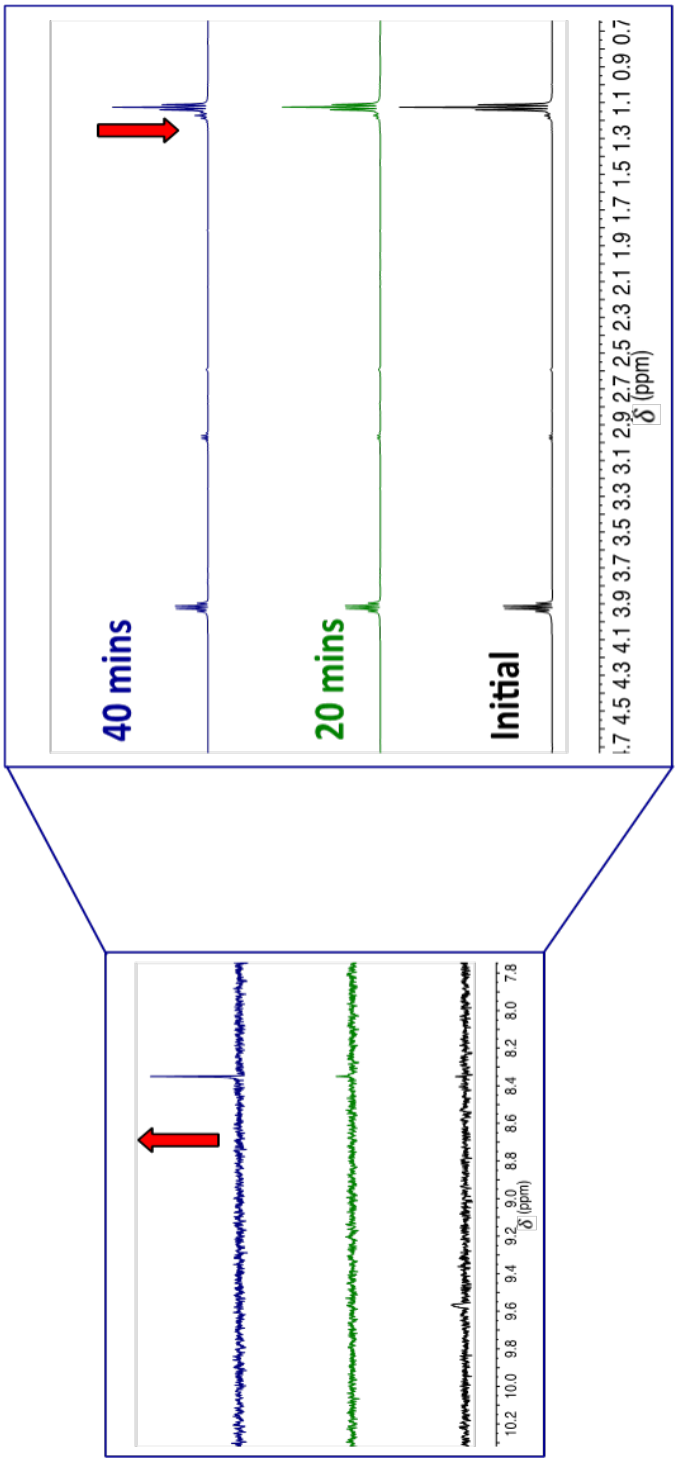

Figure 3-11. **^{1}H NMR light reaction with undecanal using Ru-WT-cAdo over 40 minute's exposure. Left insert: the growing peak at 8.3 ppm shows the formation of formate over time. Right: insert: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time.**

GC-MS analyses

The low water solubility of alkane deformylation products makes them difficult to detect using ^{1}H NMR. Consequently, GC-MS was used to identify the alkane produced from photocatalytic reactions. Initially, short chain alkanes produced from C_5-C_7 aldehydes were explored. However, developing a reliable headspace analysis method for short chain alkanes proved challenging and only qualitative analyses were possible under the best circumstances. Detection of hydrocarbon production for longer chain alkanes was more straightforward and undecanal was chosen as the aldehyde substrate. Hydrocarbon products (decane) were extracted with ethyl acetate, as described above. Using this method, decane was reproducibly detected, but comparison to yields of formate suggest that the extraction method could be improved. Several controls were also run in parallel and suggest ~ 15% recovery of decane. A possible explanation for this can be due to polypropylene microcentrifuge tubes. One potential strategy to minimize the loss could involve using glass centrifuge tubes.

Despite the low overall recovery of decane, the reproducibility of the extraction method allowed for comparison of the relative catalytic activity of our Ru-cAdo models. Accounting for this low extraction efficiency, and comparing the resulting GC-MS yields with the above NMR integrations, suggests that ca. 70 µM decane is produced. This value is consistent with the NMR integrations (~50 µM formate) and the expected 1:1 stoichiometry between formate and decane.

The activity of our Ru-R62C/C116S/C70S-cAdo model is significantly lower than our other Ru-cAdo models. Recently, Bollinger and colleagues implicated a radical on C70 as a potential radical species that could participate in the unproductive free-radical mechanism (**Figure 3-13,** red trace) when cAdo is exposed to non-native aldehyde substrates**.** By introducing a mutation at C70A they observed a 73% reduction in activity, which is similar to the approximately 70% reduction of decane production that was observed in comparison to Ru-R62C-cAdo and Ru-WT-cAdo models [24]. It also is notable that the addition of Ru at C70 does not affect catalysis. This result suggests that Ru-C70 can play a role analogous to C70 in WT-cAdo.

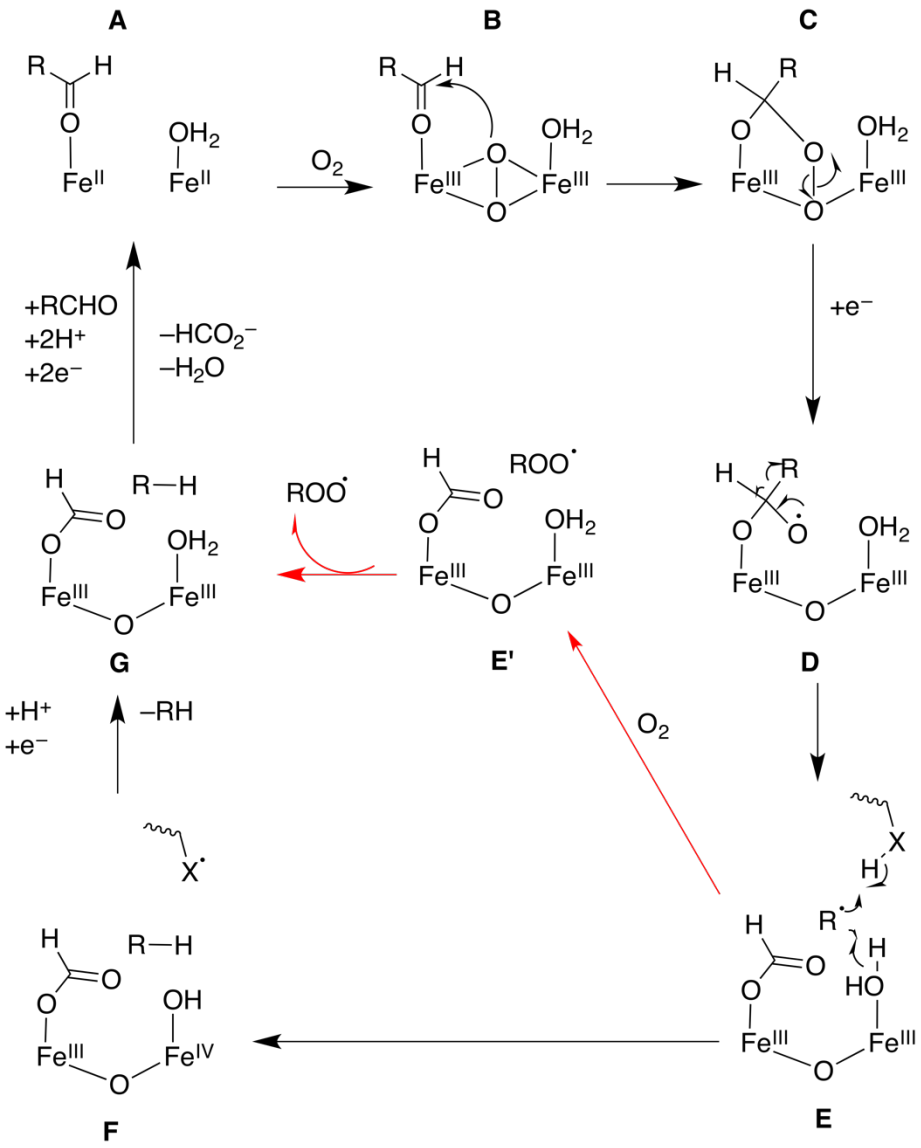

Figure 3-12. The putative free radical mechanism for cAdo; the black arrows represent the productive pathway; red arrows represent the unproductive pathway. [24]

Control reactions

To assess the efficacy of the photochemical reactions several control reactions were carried out. Activity of heterogeneously expressed protein was confirmed using a standard activity assay (see Experimental section). The ^{1}H NMR spectra were collected shortly after indicating deformylation of octanal. The NMR data confirmed enzymatic activity of WT and mutant proteins with chemical reductants. Moreover, in the literature there have been several examples of photoreduction via non-covalently bound $Ru(bpy)_3^{2+}$

[25]. To confirm that covalent attachment is necessary for our Ru-cAdo systems to undergo photocatalytic turnover, control experiments using $Ru(bpy)_3^{2+}$ as the photosensitizer were carried out. After exposure to visible light, DTC consumption was observed, but formate was not detected by ^1H NMR. This result strongly suggests that the bimolecular approach to active site reduction is not effective, even with the powerful $Ru(bpy)_3^{+}$ reducing agent. Hence, through our controls were able to demonstrate the need for DTC, and Ru-labeled cAdo for photocatalytic deformylation of aliphatic aldehydes.

3.3. Discussion

The work described in this chapter included the design and photocatalytic activity of several Ru-modified cAdo proteins. Three different Ru-cAdo models were prepared and fully characterized using several techniques that confirm the attachment of a Ru-photosensitizer to the surface of the protein. Following characterization of these labeled proteins, a photocatalytic system was developed that was used to decarbonylate aliphatic aldehydes to the corresponding alkanes. ^1H NMR and GC-MS measurements were used to show the production of formate and alkane products, respectively.

The native substrates for cAdo are C_n-aldehydes (n = 13 to 17). However, investigating catalysis with these native substrates is problematic due to their very low water solubility. Therefore, the work described in this chapter employed non-native C_n-aldehydes (n = 5 to 11) aldehyde substrates, which have significantly higher water solubility. It is important to note that this may provide misleading information about natural ET of cAdo. However, these data are still quite useful in the context of chemical synthesis or biofuel production. The modest increase in catalytic activity suggest that ET to the active site is not the limiting factor for cAdo catalysis. This is fully supported by production of peroxide side products, which likely arise from quenching of the alkyl radical intermediate by O_2, or by production and disproportionation of superoxide by Ru. These data suggest that other factors may be limiting the activity of the enzyme. A possible explanation for low catalytic activity may be attributed to slow release of the hydrocarbon product. The highly hydrophobic binding pocket may inhibit the release of products, therefore preventing other substrates from binding in the pocket. Future work should include exploring strategies to promote faster product release, including the use of mixed solvent systems where

aldehydes/alkanes have better solubility. Use of an appropriate hydrophobic co-solvent could promote faster product release, and result in a higher turnover. One potential solvent to start with is ethyl acetate for the extraction of alkanes. Alternatively, miscible solvents, such as DMSO, could play a similar role, and it is known that cAdo can tolerate low levels of DMSO in buffered media.

Finally, though dioxygen is an essential substrate for cAdo, it has also been reported to inhibit catalysis [26,27]. The molecular basis for this inefficiency was only recently recognized (**Figure 3-12**) [24] by detection of a peroxyl radical that forms from the reaction a proposed alkyl radical intermediate with O_2 (**Step E'**). It was found that the formation of the peroxyl radical was positively correlated to O_2 concentrations, and negatively correlated to the alkane production. Though we did not attempt a thorough investigation of microaerobic conditions (low oxygen conditions), it would be interesting to investigate our Ru-cAdo models under these conditions, as reported in the literature [24,26,28].

3.4. Conclusion

Three different Ru-cAdo models were characterized, and evaluated for catalytic activity. The order of increasing catalytic reactivity for our labeled proteins are as follows: Ru-R62C-cAdo > Ru-WT-cAdo > Ru-R62C/C116S/C70S-cAdo. The Ru-R62C/C116S/C70S-cAdo model has significantly (~70%) lower yields of alkane in comparison to the other two Ru-cAdo models. This is in agreement with literature reports which also report that the C70A mutation decreases the activity by more than 70% compared to WT activity. Our results reinforce the conclusion that C70 is an important residue involved in the catalytic cycle. For the remaining two Ru-cAdo models we see similar levels of activity, with slightly greater activity for Ru-R62C-cAdo. This difference in activity is attributed to the presence of two photosensitizers on the surface. These two redox-active groups likely enable more efficient delivery of the four electrons required to complete one turnover. Our work suggests that ET could still be optimized in engineered cAdo systems, but it is not the limiting factor. There likely are other underlying factors that are limiting cAdo activity, such as partitioning of aliphatic substrates and products between aqueous media and the enzyme active site. As such, controlling and optimizing the interplay between ET and substrate/product partitioning to the active site is an important goal.

3.5. References

1. Schirmer, A.; Rude, M. A.; Li, X.; Popova, E.; Cardayre, S. B. D. *Science*. **2010**, 559–562.
2. Cheesbrough, T. M.; Kolattukudy, P. E. *J. Biol. Chem*. **1988**, 263, 2738−2743.
3. Howard, R. W.; Blomquist, G. J. *Annu. Rev. Entymol*. **2005**, 50, 371−393
4. Fengming, L.; Das, D.; Xiaoxia, L.N.; Marsh, N.G.; *FEBS J*. **2013**, 280, 4773
5. Li, N.; Chang, W.C.; Warui, D.M.; Booker, S.J.; Krebs, C.; Bollinger, J.M.; *Biochemistry*, **2012**, 51, 7908.
6. Howard, T. P.; Middelhaufe, S.; Moore, K.; Edner, C.; Kolak, D.M.; Taylor, G. N.; Parker, D. A.; Lee, R.; Smirnoff, N.; Aves, S.J.; Love, J.; *Proc. Natl. Acad. Sci. U.S.A.*, **2013**, 110, 7636.
7. Kallio, P.; Pàsztor,A.; Thiel, K.; Akhtar, M.K.; Jones, P.R.; *Nat. Commun.* **2014**, 5, 4731.
8. C. Andre, S. W. Kim, X.-H. Yu, and J. Shanklin, *Proc. Natl. Acad. Sci. U.S.A.,* **2013**, 110, 3191.
9. Li, N.; Norgaard, H.; Warui, D. M.; Booker, S. J.; Krebs, C.; Bollinger, J. M. *J. Am. Chem. Soc.* **2011**, 133, 7148−7152.
10. Das, D.; Eser, B. E.; Han, J.; Sciore, A.; Marsh, E. N. G. *Angew. Chem.* **2011**, 50, 7148−7152.
11. L.J. Rajakovich, H. Nørgaard, D. Warui, W. C. Chang, S.J.Booker, C. Krebs, J.M Bollinger, M.E. Pandelia, *J. Am. Chem. Soc.*, **2015**, 3, 2515.
12. Berglund, J.; Pascher, T.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **1997**, 119, 2464–2469.
13. Bjerrum, M. J.; Casimiro, D. R.; Chang, I.-J.; Bilio, A. J. D.; Gray, H. B.; Hill, M. G.; Langen, R.; Mines, G. A.; Skov, L. K.; Winkler, J. R.; Wuttke, D. S. *J Bioenerg Biomembr.* **1995**, 27, 295–302.
14. Tran, N.-H.; Nguyen, D.; Dwaraknath, S.; Mahadevan, S.; Chavez, G.; Nguyen, A.; Dao, T.; Mullen, S.; Nguyen, T.-A.; Cheruzel, L. E. *J. Am. Chem. Soc.* **2013**, 135, 14484–14487.
15. F. N. Castellano, J. D. Dattelbaum, and J. R. Lakowicz, *Anal. Biochem.*, **1998**, 255, 165.
16. Deronzier A, Meyer TJ. *Inorg Chem.* **1980**,19, 2912–2917.
17. Tran, N., Nguyen, D., Dwaraknath, S., Mahadevan, S., Chavez, G., Nguyen, A.. . Cheruzel, L. *J. Am. Chem. Soc.*, **2013**, 135, 14484-14487.
18. Tran, N.-H.; Huynh, N.; Bui, T.; Nguyen, Y.; Huynh, P.; Cooper, M. E.; Cheruzel, L. E. *Chem. Commun.* **2011**, 47, 11936.
19. Compute Pi/MW tool (website: http://web.expasy.org/compute_pi/) (accessed May 20th 2016)

20. Peptidecutter http://web.expasy.org/peptide_cutter/ (accessed May 20th 2016)
21. Pandelia, M. E.; Li, N.; Nørgaard, H.; Warui, D. M.; Rajakovich, L. J.; Chang, W.-C.; Booker, S. J.; Krebs, C.; Bollinger, J. M. *J. Am. Chem. Soc.* **2013**, 135, 15801–15812.
22. Greenfield, N. *J. Nat. Protoc.* **2007**, 1, 2876–2890.
23. Ener, M. E. PhD. Dissertation, California Institute of Technology, 2014.
24. Rajakovich, L. J.; Nørgaard, H.; Warui, D. M.; Chang, W.-C.; Li, N.; Booker, S. J.; Krebs, C.; Bollinger, J. M.; Pandelia, M.-E. *J. Am. Chem. Soc.* **2015**, 137, 11695–11709.
25. Nilsson, T. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, 89, 6497–6501
26. Das, D.; Ellington, B.; Paul, B.; Marsh, E.N.G.; *ACS Chem. Biol.* **2014**, 9 570
27. Andre, C.; Kim, S. W.; Yu, X.-H.; Shanklin, *J. Proc. Natl. Acad. Sci. U.S.A.* **2013**, 110, 3191–3196.
28. Waugh, M. W.; Marsh, E. N. G. *Biochemistry*, **2014**, 53, 5537–5543.

Chapter 4.

Summary, Conclusions, and Future Directions

Electron transfer (ET) in biological systems has attracted much interest from chemists, biochemists, and biologists. The work described in this thesis focused on application of a great body of fundamental work that already existed in the literature. We describe how we can use this knowledge to introduce new redox functionality into metalloproteins by the addition of artificial electron wires. I used two different dioxygen-activating metalloproteins as models to produce artificial electron wires. First, I demonstrated a novel approach for creating amino acid ET 'wires' from the heme to the protein surface of yeast cytochrome *c* peroxidase (CcP) (**Chapter 2**). Incorporation of two mutations, Y229W and A193W, results in a properly folded CcP protein that supports formation of high-valent heme intermediates. Addition of H_2O_2 to these proteins promoted oxidation of stable organic compounds slightly faster than WT CcP. Second, I designed several Ru-modified cAdo systems where flash-quench generated Ru(I) shuttled electrons to the embedded diiron active site (**Chapter 3**). The Ru-modified cAdo systems were shown to catalyze the deformylation of aldehydes and produce formate and alkanes.

Chapter 2: Mobilizing electrons in cytochrome c peroxidase: designing and testing an electron transfer wire

An electron transfer wire was introduced into CcP by introducing two AAs in close proximity to Trp191, the well-known radical in the proximal pocket of CcP [1]. The ET wire is most efficient under alkaline pH conditions. This implies that under more acidic conditions the net driving force for ET from the heme to the surface is above 200 meV for at least one ET step. It is known that this makes electron hopping unfavourable [2]. It would be desirable to engineer a mutant that performed under more acidic conditions, as the stability of CcP Cmpd I significantly drops above pH 8 [3]. Although I showed that the ET behaviour can be modulated by varying the pH, one can also modify the protein environment to influence changes in reduction potential. For example, the remarkable stability of Cmpd I can be attributed to the protein environment of Trp191. Early work on

CcP revealed a large area of negative electrostatic potential contained in the proximal pocket of the active site [4]. Future work can look at constructing similar environments for our newly introduced AAs to generate wires at lower pH values.

Crystallographic information for the double mutant (CcP-WW) would provide additional insight into the nature of ET. With crystallographic data one can generate hopping maps, which use semiclassical ET theory to display of heat map for multi-step ET systems [5]. Furthermore, one can investigate how the orientation of the AA residues affects catalysis. Aromatic residues typically orient themselves in either parallel-displaced (PD) or T-shapes orientation. PD orientations are known to favour ET due to the increased overlap of the molecular orbitals [6]. As such, the orientation of the aromatic residues could be analyzed by inspecting crystal structures for our CcP-WW mutant. This would allow us to gain a better understanding on the nature of ET in our mutant and guide us in the developing second-generation electron wires.

An alternative approach could include using existing Trp/Tyr chains in biological systems. Potential candidates could include dioxygen-activating enzymes such as, cytochromes P450 (CYPs). P450 monooxygenases have Tyr/Trp chains to relieve oxidative stress [7]. Targeting mutations that modulate current reactivity, and make electron hoping through the wire a more favourable pathway would be ideal. In addition to introducing such chains, recent work demonstrates that redox reactivity at the heme site can be modified with only a few mutations [8]. Generation of a strong oxidant, as implicated in peroxidases, is critical for the activation of P450 wires.

Chapter 3: Photocatalytic production of alkanes using Ru-modified cyanobacterial aldehyde deformylating oxygenase (cAdo)

Since the early 1980s, ruthenium photosensitizers have been used to investigate ET and catalysis in biological systems [9,10]. The recent interest in cAdo prompted us to incorporate a Ru-photosensitizer to the surface of the protein as a strategy to increase the rate of ET to the active site. I designed and characterized three different Ru-modified cAdo models and demonstrated that electrons could be shuttled into the diiron site to carry out catalytic aldehyde deformylation. I showed that our Ru-cAdo models could successfully deformylate a range of short chain aldehydes (C_5-C_{11}), confirming the functional activity

of our Ru-cAdo models. I observed the corresponding C_{n-1} alkane through GC-MS, and also observed formate production through proton NMR.

The Ru-cAdo models described in **Chapter 3** can be used to investigate the mechanistic underpinnings of cAdo. **Figure 3-12** displays the putative catalytic cycle of cAdo, but there has not yet been a report of the direct detection of an alkyl radical that supports **Step E**. Nonetheless, this radical pathway has been favoured by many [11]. The most recent evidence supporting the radical mechanism came from the detection of the substrate derived peroxyl radical that accumulates with non-native short-chain aldehydes [12]. The peroxyl radical is consistent with radical mechanism, as opposed to the recently proposed heterolytic bond cleavage step [13,14]. With our modified Ru-cAdo models one can investigate alkyl radical formation with native aldehyde substrates (i.e., C_{15}-C_{19}). Flash-quench techniques and electron paramagnetic resonance (EPR) spectroscopy can be used to investigate the presence of a short-lived alkyl radical, which would provide conclusive evidence for the free radical mechanism.

In addition to the above mechanistic work, industrial applications could be tailored on the basis of our mechanistic work. Many industrial applications require the use of short chain aldehydes for biofuel production and the formation of additional oxidative products is not desired. Therefore, there is incentive to bioengineer cAdo to avoid additional by-products. Mutations to the substrate-binding pocket can help modify the substrate bias to allow for short chain aldehydes. Efforts at modifying the long hydrophobic substrate-binding pocket have been reported [15]. An example includes changing Ala134 to Phe residue. This specific mutation has been known to have a substrate bias towards short chain aldehydes. Future work can include introducing mutations to our Ru-cAdo models as a strategy to producing higher yields of short chain alkanes.

Finally, dioxygen is an essential substrate for cAdo, but the unproductive quenching of the alkyl radical in the substrate-binding pocket in the presence of excess dioxygen, makes dioxygen an inhibitor as well. This unproductive reaction pathway is outlined in **Figure 3-12** and shows how excess O_2 can produce the substrate-derived peroxyl radical. **Figure 3-12** (Step E) shows the hydrogen atom originates from the coordinated water ligand [16]. For the unproductive pathway the quenching of the dioxygen occurs faster than the abstraction of the hydrogen. Future work could include

exploring ways to engineer the active site such that one can avoid the formation of the peroxyl radical under ambient dioxygen concentrations.

4.1. References

1. Miller, M. A.; Vitello, L.; Erman, J. E*. Biochemistry* **1995**, *34*, 12048–12058
2. Warren, J. J.; Ener, M. E.; Vlček, A.; Winkler, J. R.; Gray, H. B. *Coord. Chem. Rev.* **2012**, *256*, 2478–2487
3. Loo, S.; Erman, J. E. *Biochemistry* **1975**, *14*, 3467–3470.
4. Miller, M. A.; Han, G. W.; Kraut, J. *Proc. Natl. Acad. Sci. U.S.A.*. **1994**, *91*, 11118–11122.
5.. Warren, J. J.; Winkler, J. R.; Gray, H. B. *Coord. Chem. Rev.* **2013**, *257*, 165–170.
6. Polizzi, N. F.; Migliore, A.; Therien, M. J.; Beratan, D. N. *Proc. Natl. Acad. Sci. U.S.A.*. **2015**, *112*, 10821–10822
7. Winkler, J. R.; Gray, H. B. *Quart. Rev. Biophys.* **2015**, *48*, 411–420.
8. Coelho, P., Wang, Z., Ener, M., Baril, S., Kannan, A., Arnold, F., Brustad, E. *Nat. Chem. Biol.* **2013**, 9, 485-487
9. Winkler, J. R.; Gray, H. B. *Chem. Rev*. **1992**, *92*, 369–379.
10. Millett, F.; Durham, B. *Biochemistry* **2002**, *41*, 11315–11324.
11. Das, D.; Ellington, B.; Paul, B.; Marsh, E. N. G. *ACS Chem. Biol*. **2014**, *9*, 570–577.
12. Rajakovich, L. J.; Nørgaard, H.; Warui, D. M.; Chang, W.-C.; Li, N.; Booker, S. J.; Krebs, C.; Bollinger, J. M.; Pandelia, M.-E. *J. Am. Chem. Soc.* **2015**, *137*, 11695–11709.
13. Aukema, K. G.; Makris, T. M.; Stoian, S. A.; Richman, J. E.; Münck, E.; Lipscomb, J. D.; Wackett, L. P. *ACS Catal*. **2013**, *3*, 2228–2238
14. Marsh, E. N. G.; Waugh, M. W. *ACS Catal*. **2013**, *3*, 2515–2521.
15. Khara, B.; Menon, N.; Levy, C.; Mansell, D.; Das, D.; Marsh, E. N. G.; Leys, D.; Scrutton, N. S. *ChemBioChem.* **2013**, *14*, 1204–1208.
16. Waugh, M. W.; Marsh, E. N. G. *Biochemistry* **2014**, *53*, 5537–5543

Appendix A.

Supporting Information for Chapter 2

Table A1: Primers used for CcP mutants purchased from Eurofins Operon

Mutation	Sequence (5'→3')
Ala193Trp	Forward primer: GGTCCGTGGGGCTGGGCAAACAATCTGTTTACG
	Reverse primer: CGTAAACAGATTGTTTGCCCAGCCCCACGGACC
Tyr229Trp	Reverse primer: GGACAGTAAATCCGGTTGGATGATGCTGCCGACC
	Reverse primer: GGTCGGCAGCATCATCCAACCGGATTTACTGTCC

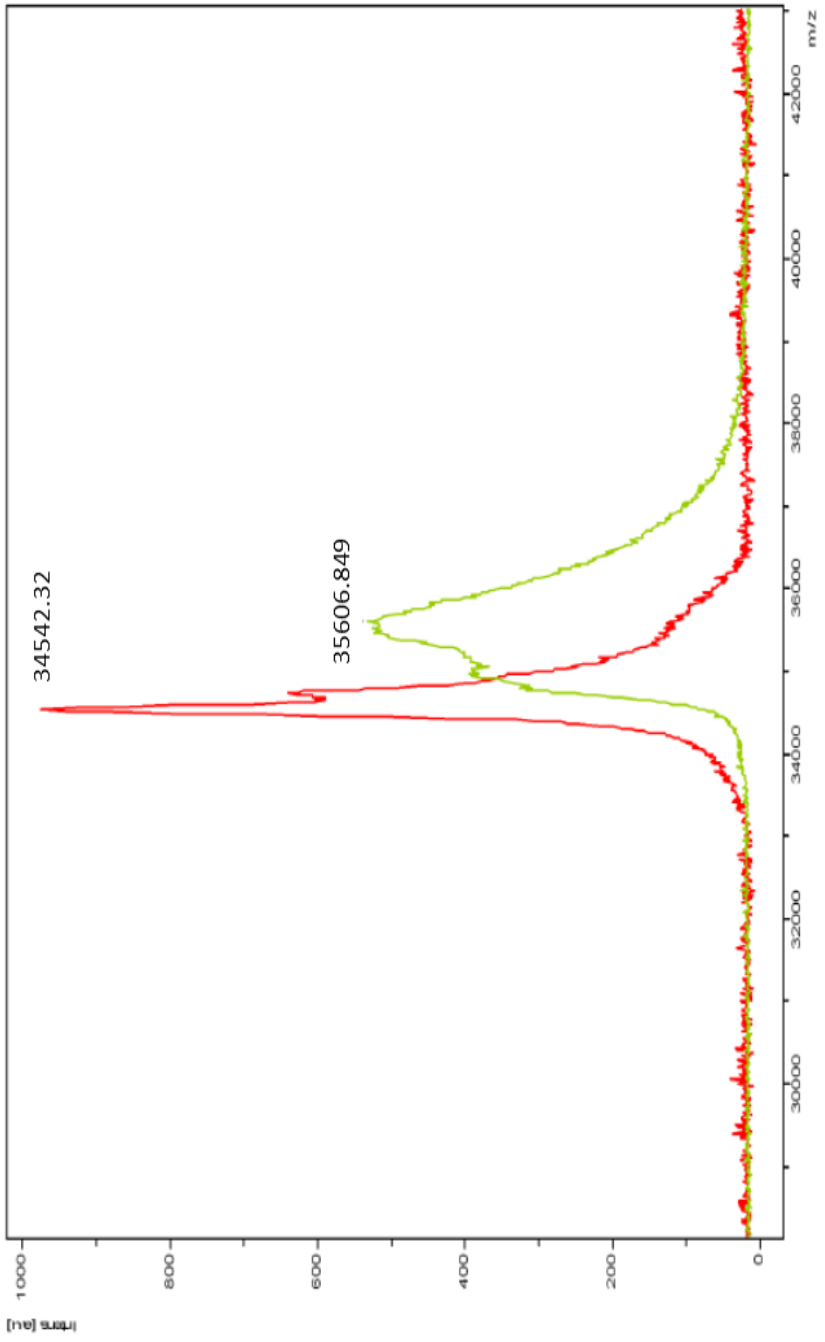

Figure A.2. Overlay of MALDI spectra for CcP-WW mutant. Red trace: CcP-WW before VA oxidation at pH 8. Green trace: CcP-WW after VA oxidation at pH 8.

Appendix B.

Supporting Information for Chapter 3

Table B1: Primers used for cAdo mutants purchased from Eurofin Operon

Mutation	Sequence (5'→3')
Arg62Cys	Forward primer: GCCAAAATGGAACAATGCCACATGAAAGGCTTTATGG
	Reverse primer: CCATAAAGCCTTTCATGTGGCATTGTTCCATTTTGGC
Cys116Ser	Forward primer: GCCTGATTATCGAAAGTTTTGCG
	Reverse primer: CGCAAAACTTTCGATAATCAGGC
Cys70Ser	Forward primer: GGCTTTATGGCAAGCGG
	Reverse primer: CCGCTTGCCATAAAGCC

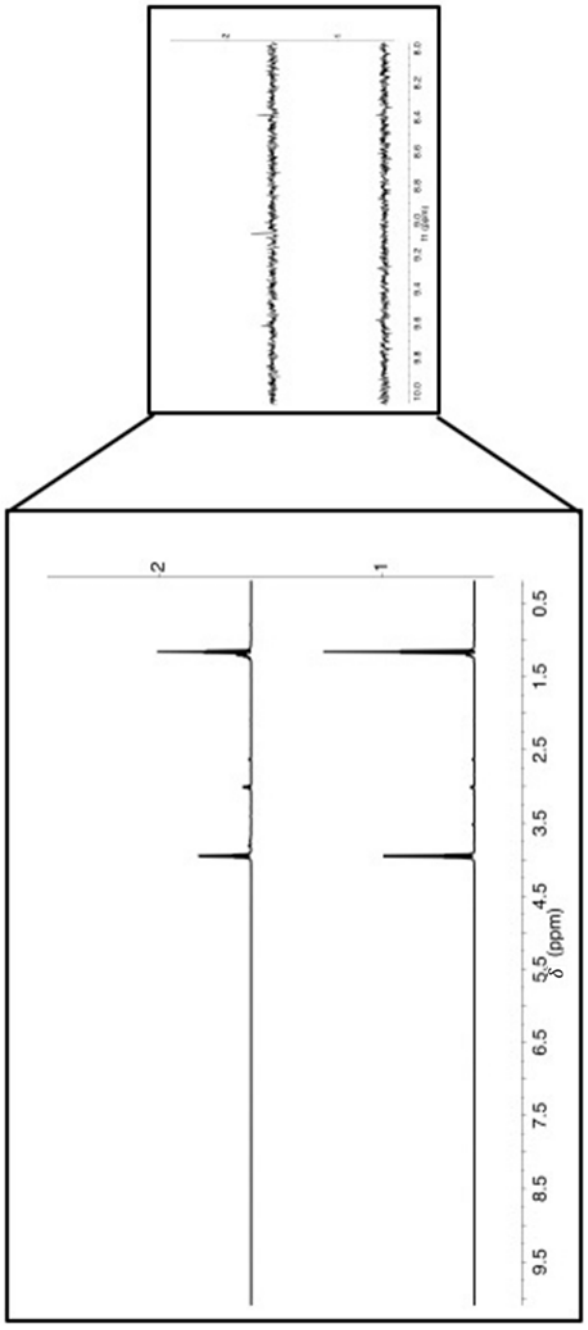

Figure B.1. ^{1}H NMR of light reaction with undecanal using Ru-R62C-C70S-C116S cAdo. Top: 60 minutes of light exposure Bottom: no light exposure. Left insert: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time. Right: the growing peak at 8.3 ppm shows the formation of formate over time; the small peak at approximately 9 ppm indicates the formation of a peroxide side product.

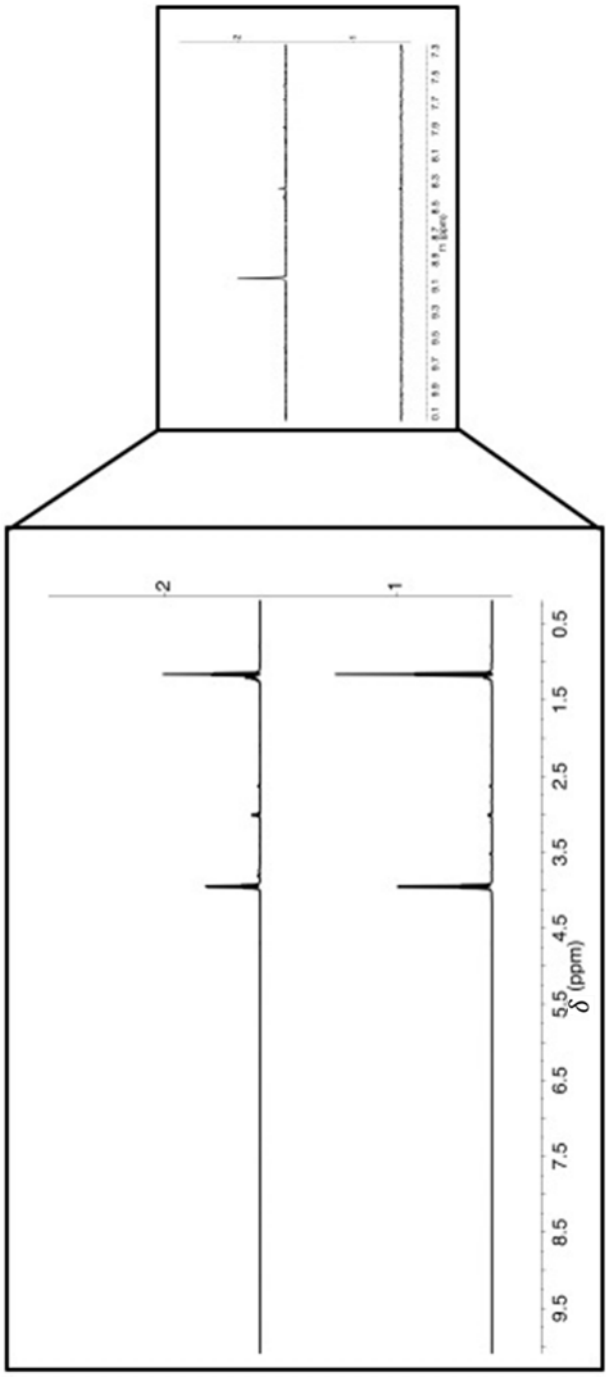

Figure B.2. ^{1}H NMR of light reaction with undecanal using Ru-R62C-cAdo. Top: 60 minutes of light exposure Bottom: no light exposure. Left insert: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time. Right: the growing peak at 8.3 ppm shows the formation of formate over time; the small peak at approximately 9 ppm indicates the formation of a peroxide side product.

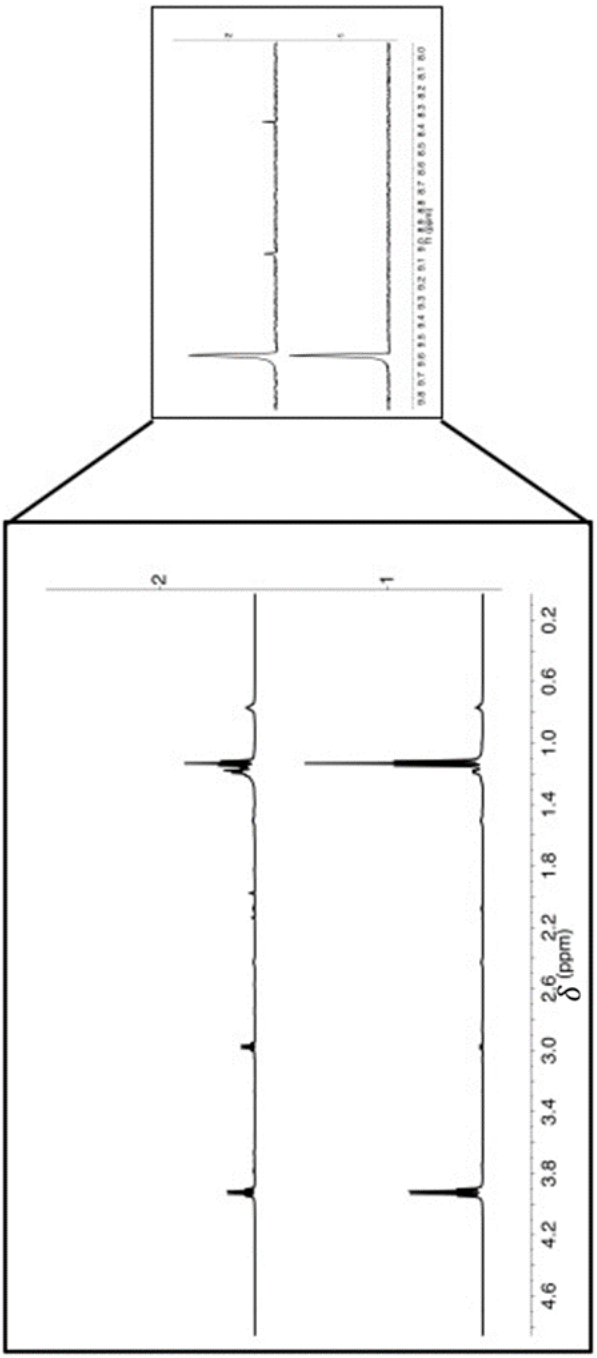

Figure B.3. ^{1}H NMR of light reaction with heptanal using Ru-WT-cAdo; Top: 30 min light exposure Bottom: no light exposure. Left inset: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time. Right: the growing peak at 8.3 ppm indicates the formation of formate over time; the small peak at approximately 9 ppm indicates the formation of a peroxide side product.

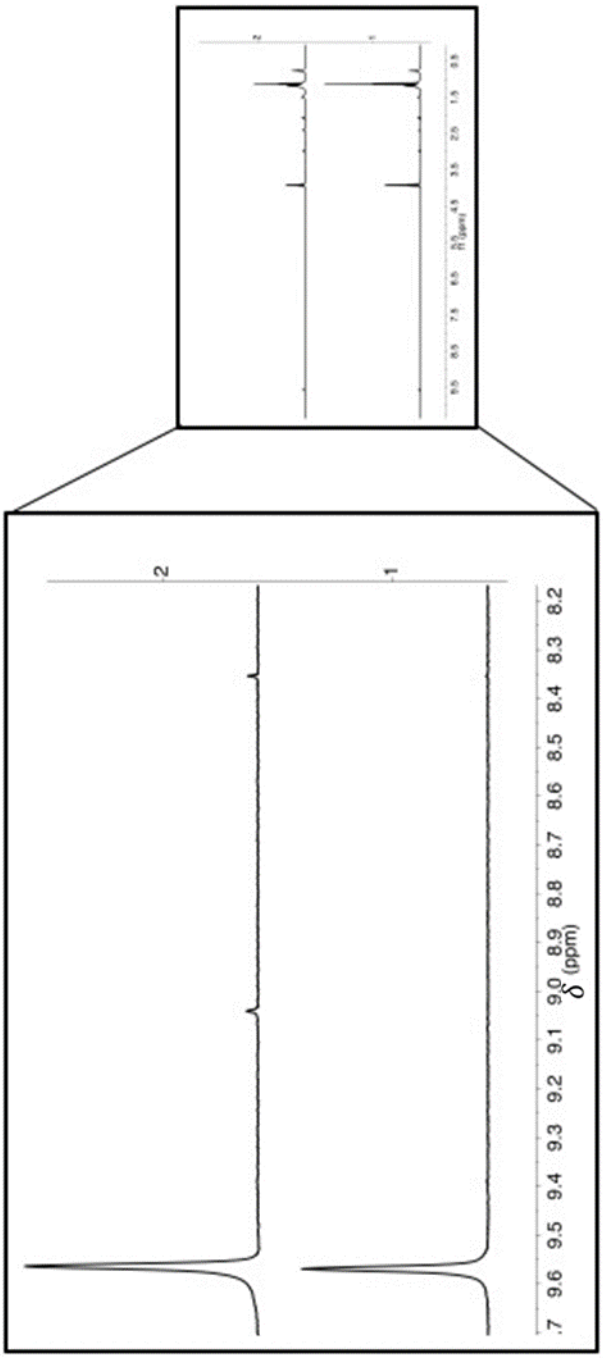

Figure B.4. ^{1}H NMR of light reaction with heptanal using Ru-R62C-C116S-C70S-cAdo; Top: 90 min light exposure Bottom: no light exposure. Right insert: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time. Left insert: the growing peak at 8.3 ppm indicates the formation of formate over time; the small peak at approximately 9 ppm indicates the formation of a peroxide side product.

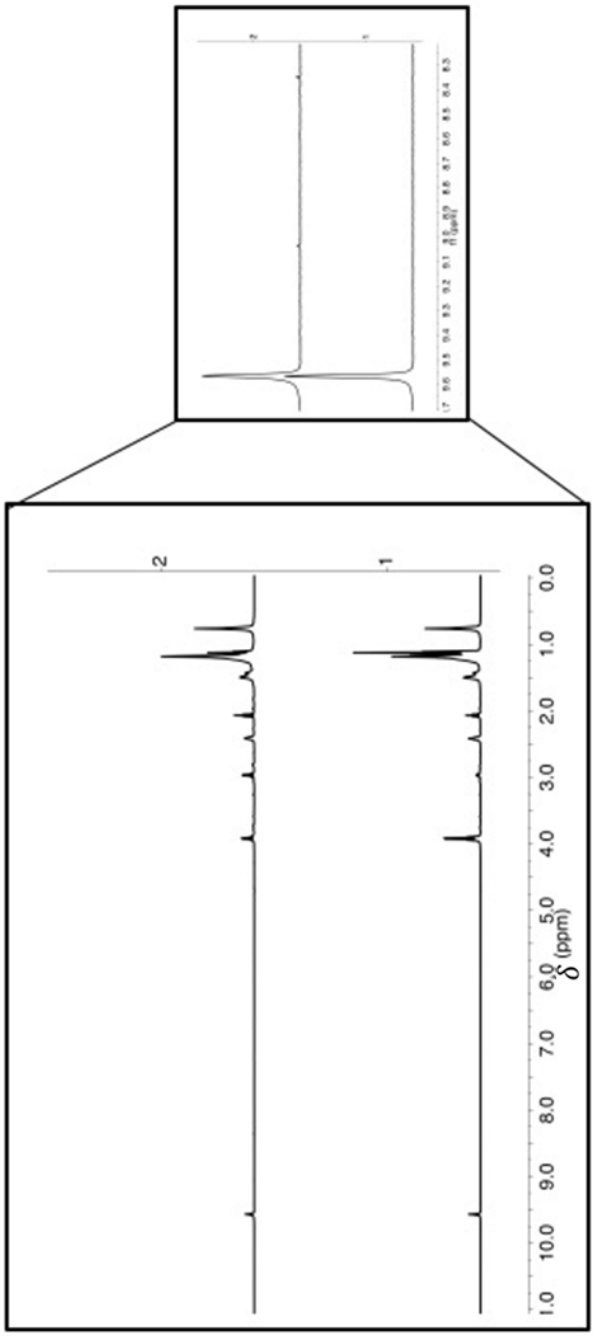

Figure B.5. ^{1}H NMR of light reaction with heptanal using Ru-R62C-cAdo; Top: 60 min light exposure Bottom: no light exposure. Left insert: the decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time. Right insert: the growing peak at 8.3 ppm indicates the formation of formate over time; the small peak at approximately 9 ppm indicates the formation of a peroxide side product.

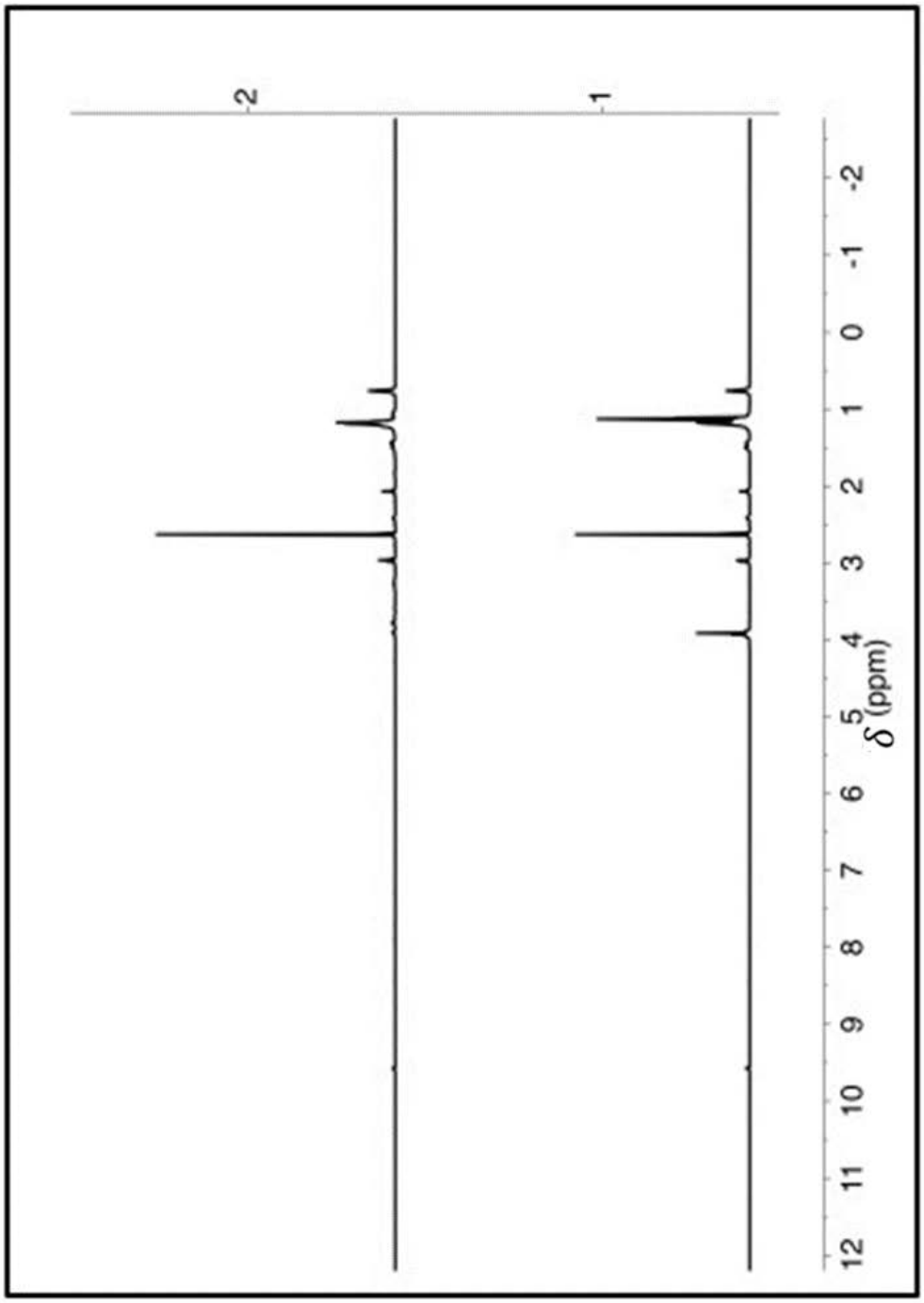

Figure B.6. ^{1}H NMR of WT-cAdo control with non-covalently attached Tris(2,2'-bipyridine)ruthenium (II) chloride. Bottom trace: no light exposure. Top trace: 180 minutes of light exposure. The decreasing peak at approximately 3.9 ppm and 1.1 ppm indicates quenching of DTC over time.

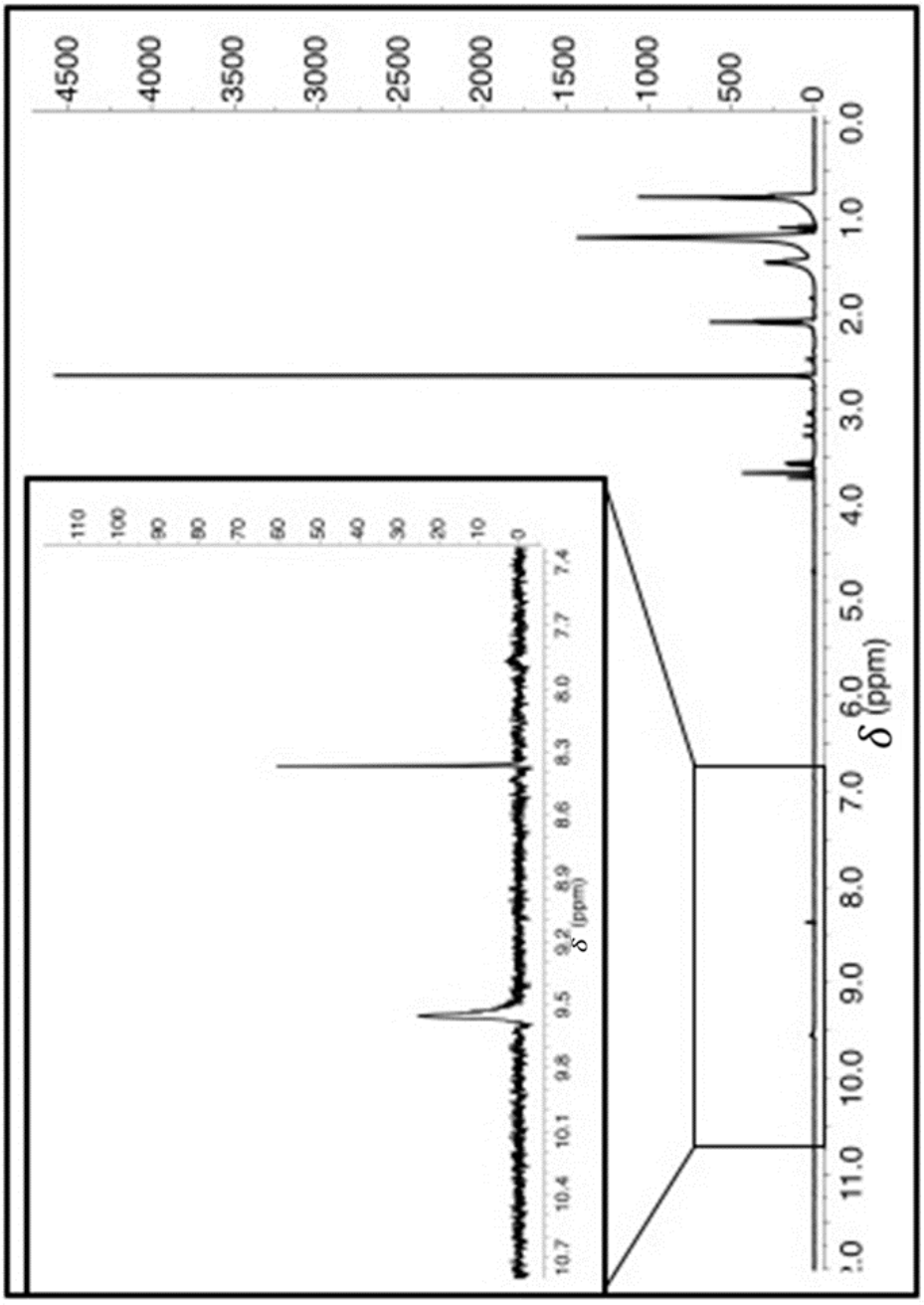

Figure B.7. ^{1}H NMR of R62C-cAdo control with reductant, Phenazine methosulfate (PMS). Addition of the reductant results in the growth of a peak at 8.3 ppm indicating the formation of formate.

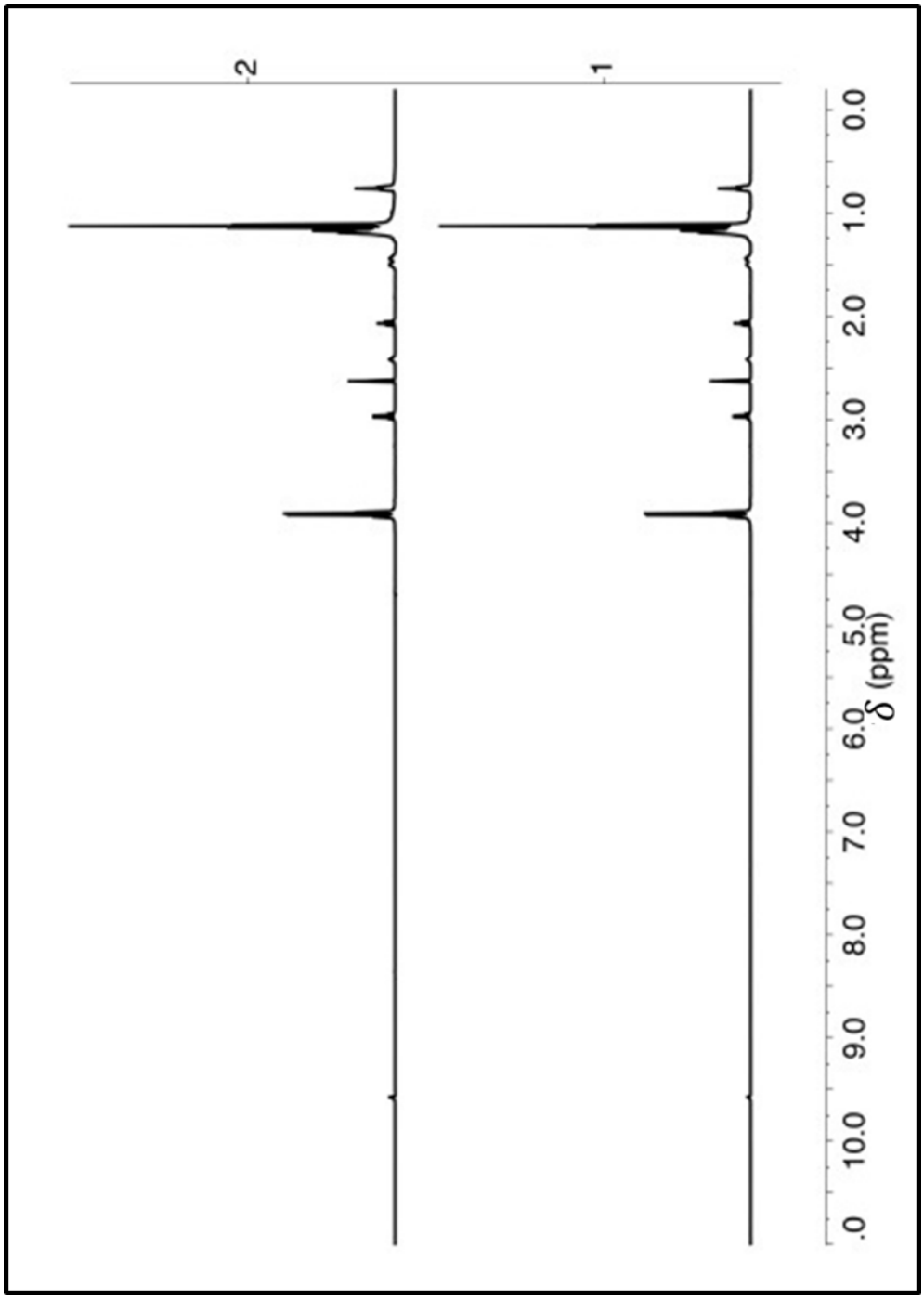

Figure B.8. ^{1}H NMR of WT-cAdo control with DTC. Bottom trace: no light exposure Top trace: 180 minutes of light exposure. No change in peaks indicating that no reaction occurs without Ru-labeled protein.

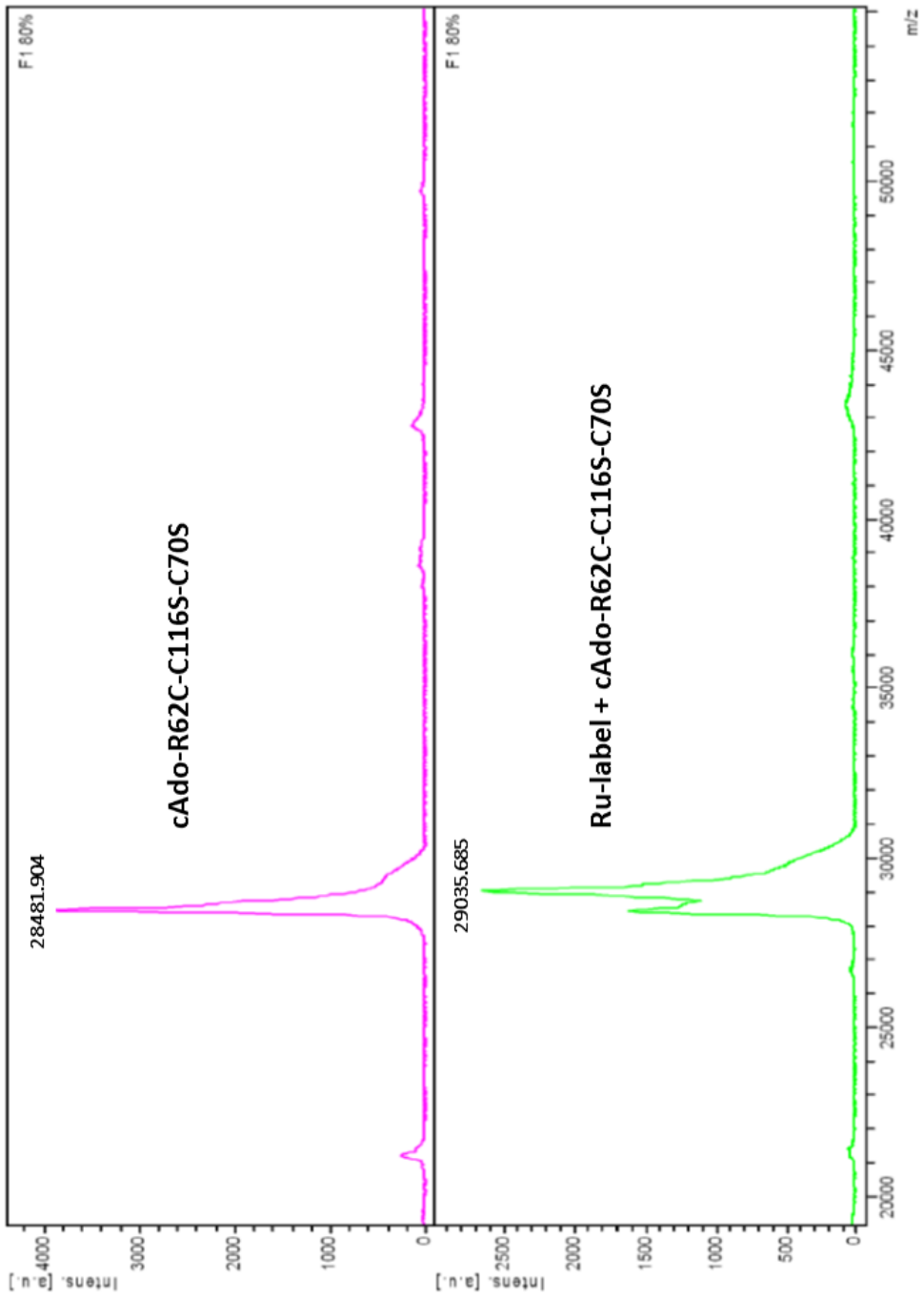

Figure B.9. Stacked MS of R62C/C116S/C70S-cAdo unlabeled (top) and Ru-R62-C116S-C70S-cAdo (bottom) collected from MALDI. The mass shift of approximately 597 Da is indicative of one Ru-label attached to the surface of the protein.

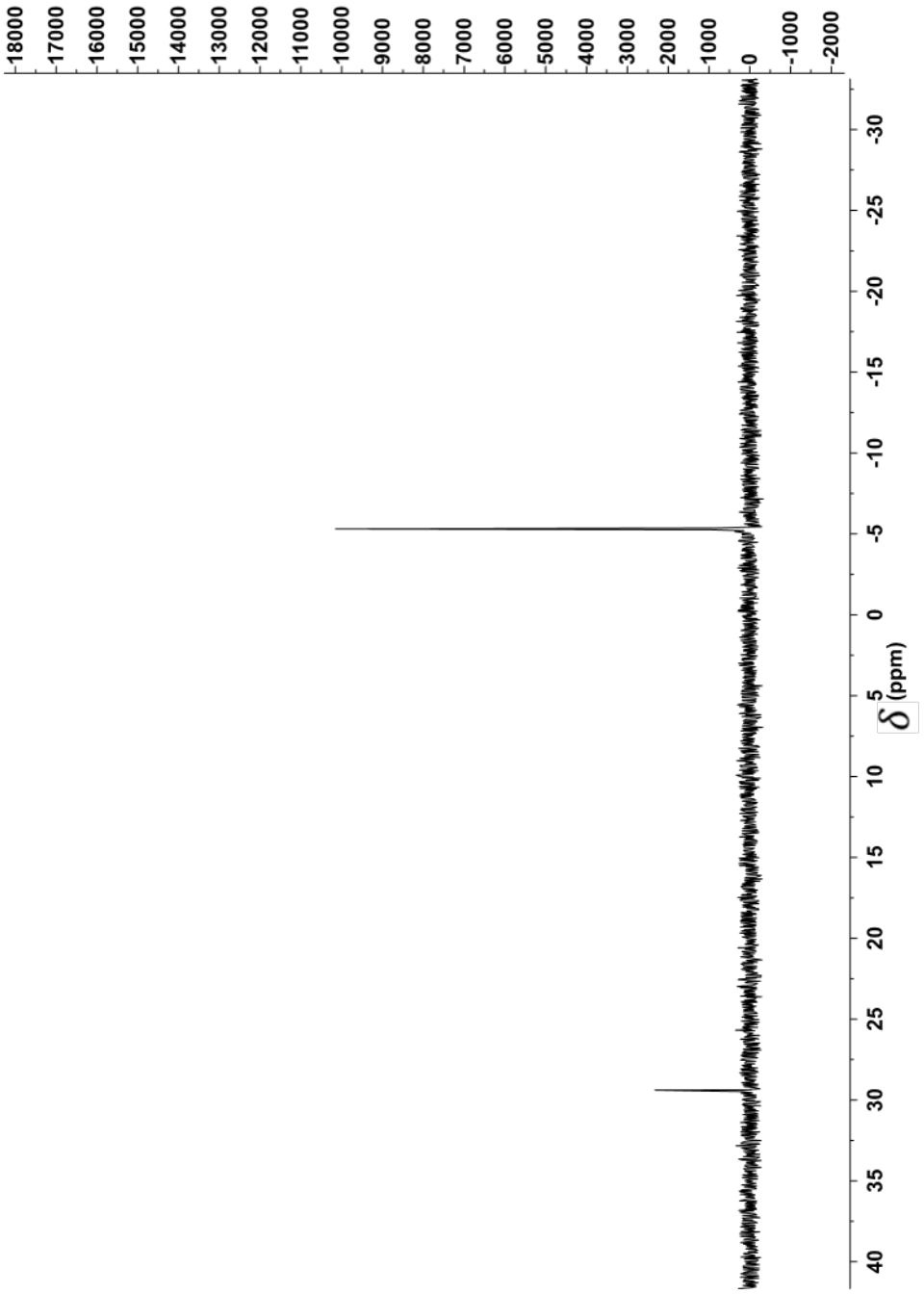

Figure B.10. ^{31}P NMR of light reaction extracted with 2 mM triphenylphosphine (PPh_3) in $CDCl_3$. Ru-R62C-cAdo. Peak at -5.3 ppm corresponds to PPh_3 peak; peak at 29.4 ppm consistent with production of triphenylphosphine oxide ($OPPh_3$).

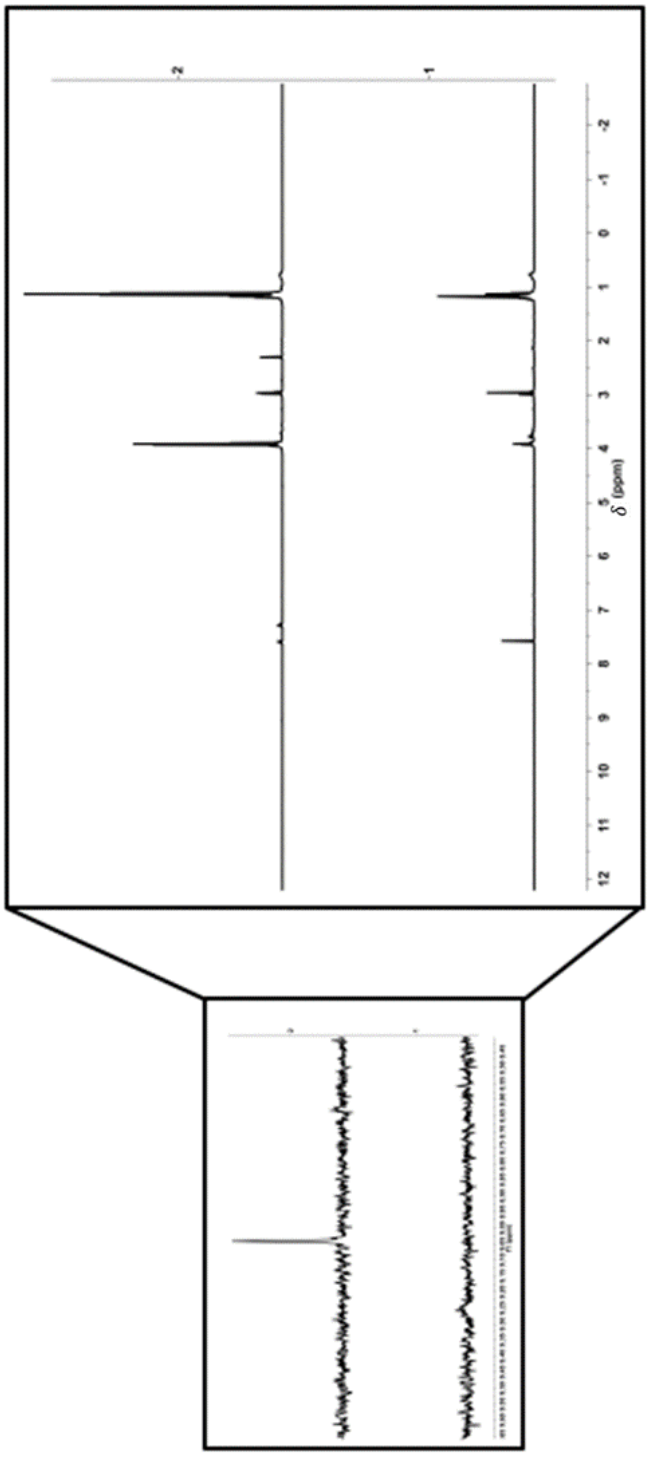

Figure B.11. ^{1}H NMR of Ru-R62C-cAdo light reaction with undecanal after 120 minutes of exposure. Top trace: ^{1}H NMR before extraction with 2 mM triphenylphosphine (PPh_3) in $CDCl_3$. Bottom trace: aqueous layer after extraction with 2 mM PPh_3 in $CDCl_3$. Loss of peak at approximately 9 ppm indicates the extraction of our byproduct upon addition of PPh_3 solution.

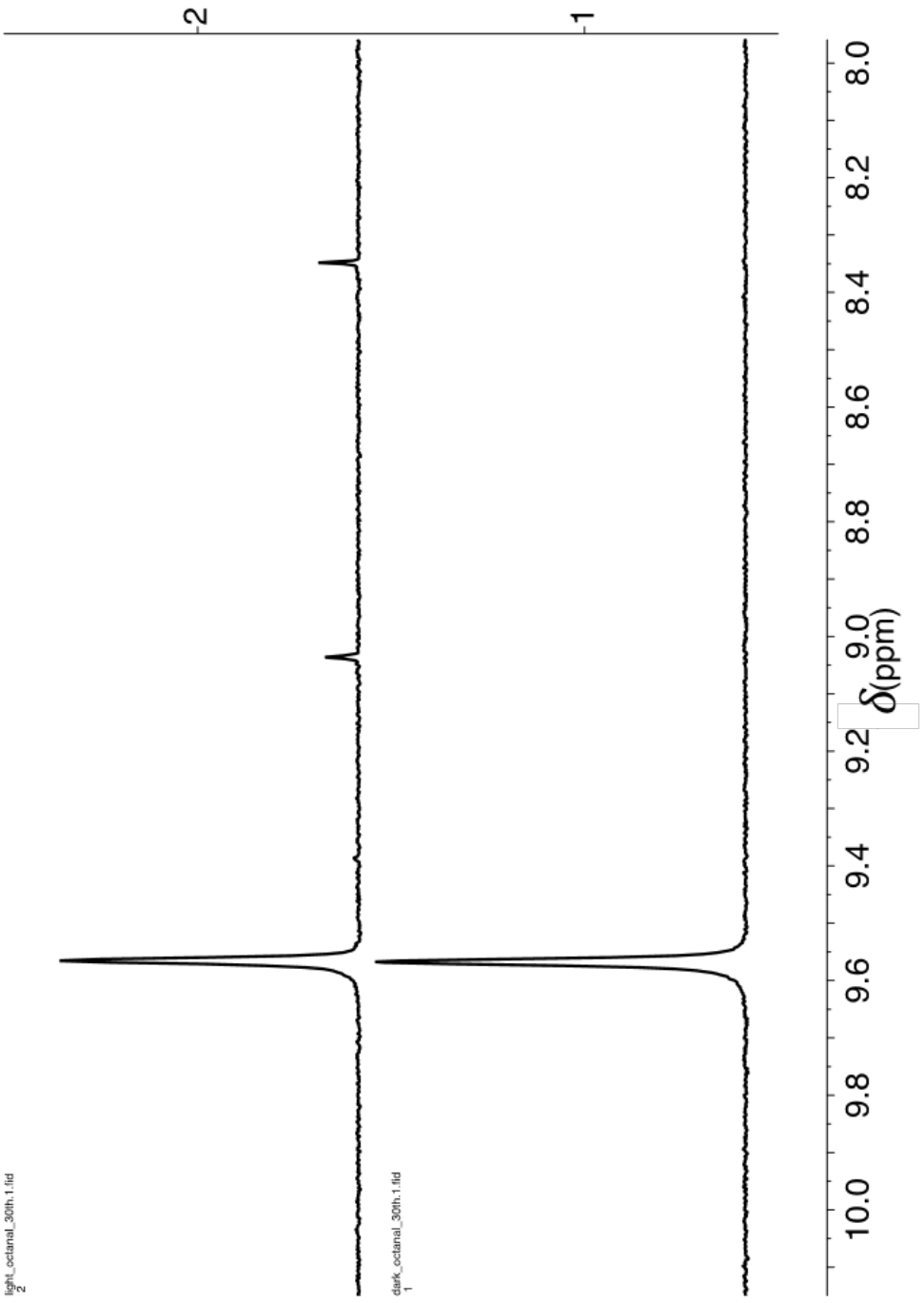

Figure B.12. ^{1}H NMR of Ru-R62C-cAdo reaction with octanal. Top trace: 90 minutes of light exposure; Bottom trace: no light exposure. Peak at approximately 9.5 ppm corresponds to the aldehyde peak. After 90 minutes of light exposure two peaks grow in. Peak at 8.35 ppm is indicative of formate and peak at approximately 9.1 ppm indicates the formation of a peroxide side product

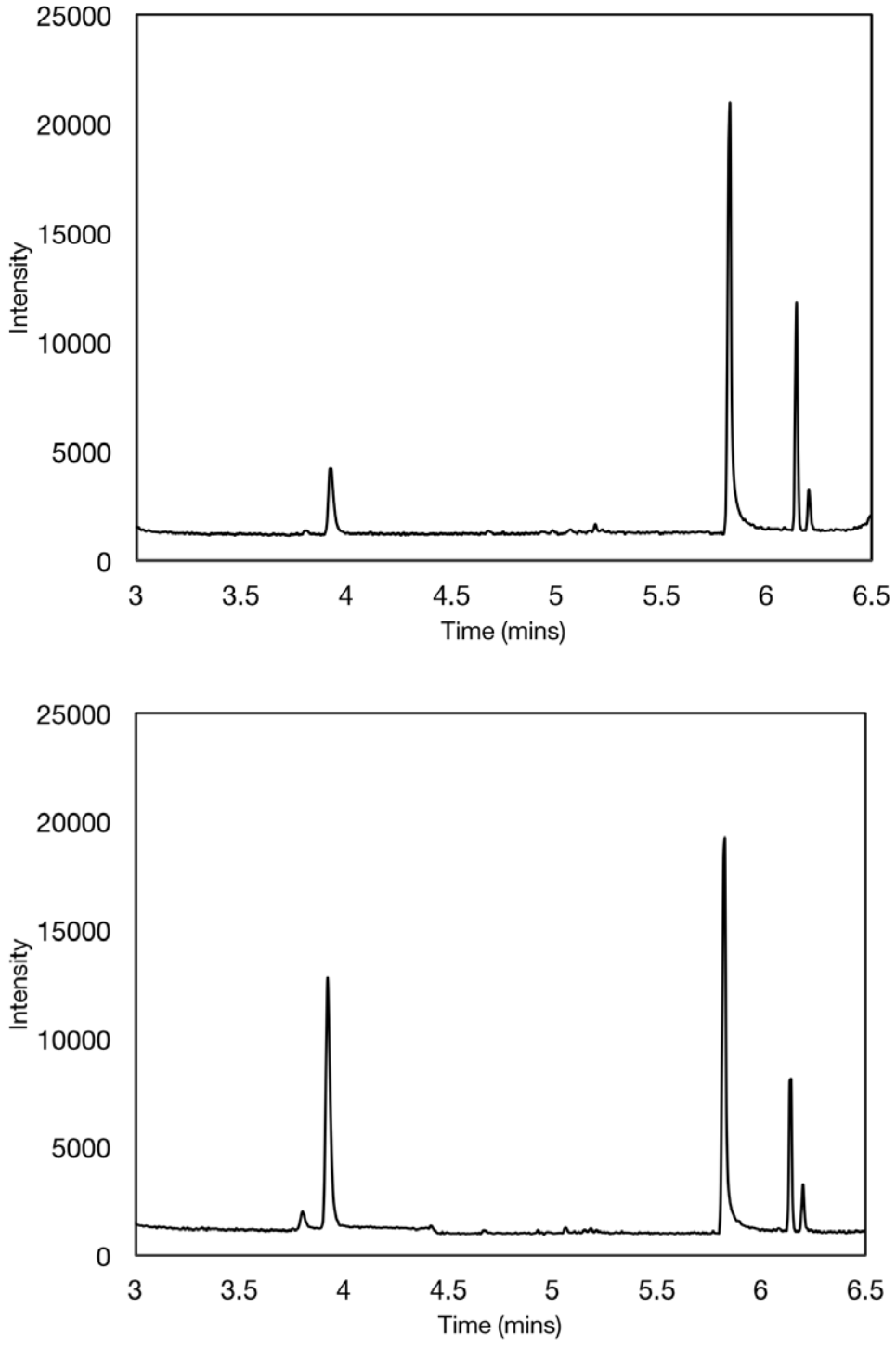

Figure B.13. GC-MS chromatograph of Ru-R62C-C116S-C70S-cAdo (**top**) and Ru-R62C- cAdo (**bottom**) after 60 minutes of light exposure with undecanal. Trace shows the production of the n-1 hydrocarbon (decane). All light reactions were extracted with ethyl acetate. Retention time: 3.9 min (decane), 5.8 min (undecanal), and 6.1 min (dodecane)

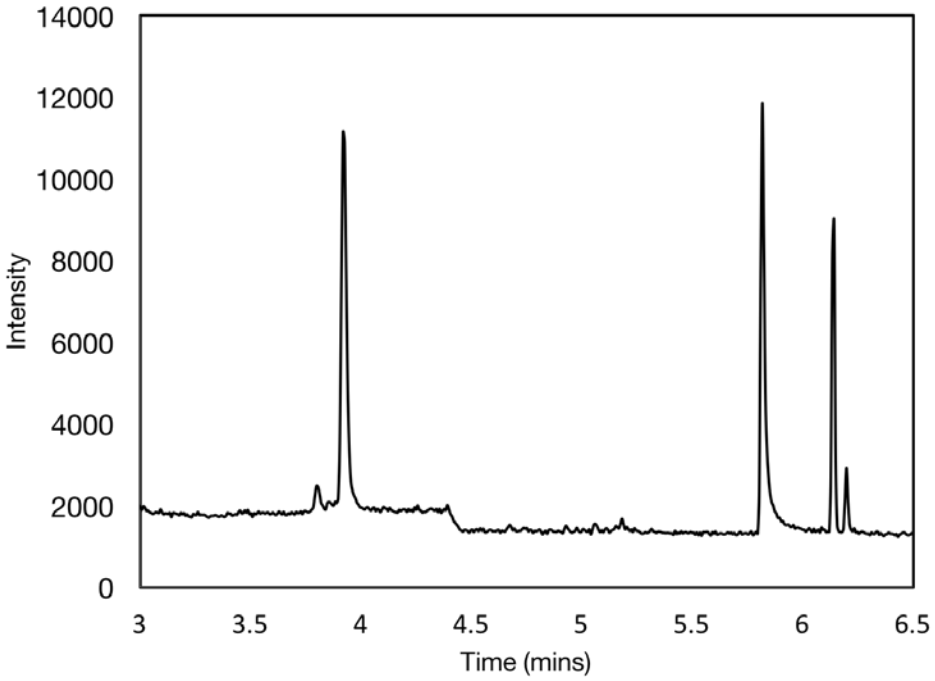

Figure B.14. GC-MS chromatograph of Ru-WT-cAdo after 60 minutes of light exposure with undecanal. Trace shows the production of the n-1 hydrocarbon (decane). All light reactions were extracted with ethyl acetate. Retention time: 3.9 min (decane), 5.8 min (undecanal), and 6.1 min (dodecane)

Table B2: GC-MS data for Ru-cAdo models.

Name	Area of decane peak	Area of dodecane peak	Decane produced (mg/L)
Ru-WT-cAdo	47185	21814	1.9mi6
	118469	103352	1.026
	167670	74468	2.016
Ru-R62-cAdo	64961	26912	2.161
	179201	71088	2.257
	47598	34130	1.249
Ru-R62C-C70S-C116S-cAdo	18236	24972	0.654
	135325	190364	0.637
	55755	92496	0.540