ABSTRACT

MOSLEY, DONTAE ALLEN. Di-metal Enzymes Responsible for Para-aminobenzoic Acid Biosynthesis (Under the direction of Dr. Thomas Makris)

Di-metal enzymes are key to several biological functions and can perform various types of chemistry. It was recently revealed that two unique di-metal enzymes are involved in para-amino benzoic acid (pABA) synthesis. pABA is a building block of folic acid which is needed for microorganisms to survive. *Chlamydia trachomatis* lacks the *pabABC* genes required for pABA synthesis. The Chlamydia Protein Associating with Death Domains (CADD) is an enzyme found in *Chlamydia trachomatis*. In this study we show that the CADD enzyme can undergo a "self-sacrificial" reaction to synthesize pABA. The NE1434 enzyme is an ortholog of CADD and is found in *Nitrosomonas europaea*. We believe this enzyme is essential to the survival of the *Nitrosomonas europaea* species. This study utilizes various mass spectrometry techniques to show CADD undergoes a "self-sacrificing" reaction which results in pABA synthesis. This study also provides insight into how NE1434 orchestrates the synthesis of pABA.

© Copyright 2022 by Dontae Allen Mosley

All Rights Reserved

Di-metal Enzymes Responsible for Para-aminobenzoic Acid Biosynthesis

by

Dontae Allen Mosley

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Biochemistry

Raleigh, North Carolina

2022

APPROVED BY:

_____	_____
Dr. Thomas Makris	Dr. Robert Rose
Committee Chair	
_____	
Dr. Paul Swartz	

DEDICATION

This is dedicated to my family and loved ones. This is especially for May your grandson is still staying focused

BIOGRAPHY

Dontae Mosley is from Rocky Mount North Carolina. Growing up he was intrigued by disease and infections. He knew he wanted to pursue a career in which he would help make the world a better place by conducting research. After Graduating from Nash Central High School, he decided to pursue his bachelor's degree in Biotechnology at The University of North Carolina at Pembroke (UNCP). While at UNCP he gained the research experience needed to be successful in graduate school. Upon earning his bachelor's degree, he decided to pursue his Masters in Biochemistry at North Carolina State University (NCSU). Over the past two years he has been learning about cellular metabolism and protein biochemistry. After graduating from NCSU, he looks forward to starting his professional career and hopes to impact human health through his research.

ACKNOWLEDGMENTS

First and foremost, I'd like to thank God for always putting me in position to be successful. I would also like to thank my esteemed professor Dr. Thomas Makris. He has provided invaluable support and guidance during my graduate career. I also appreciate Dr. Bob Rose, Dr. Paul Swartz, Dr. Peter Thompson, and Dr. Greg Buhrman; you guys have also helped me be successful during my graduate career. I am extremely grateful for Dr. Rachel Smith, Dr. Conner Sandefur, Dr. Courtney Carroll Alexander, Dr. Meredith Storms, and the rest of my BRAVE family from The University of North Carolina at Pembroke.

To my colleagues from the Makris lab I love you guys and I am thankful for the part you all played in my success. Thank you, Dr. Olivia Manley, for teaching me the ropes and coaching me through my project. I appreciate Brian McMillan and Hannah Gering, the good times we had in the lab were amazing you guys have become lifelong friends. Suman Das you are amazing, and your personality kept me smiling. Han Phan you are also amazing thank you for fixing the HPLC. Leticia Rade and Mayara Chagas de Ávila you guys are wonderful thank you for bringing homemade cakes to our lab group meetings. Priscilla Baum, you have a great personality I enjoyed our conversations. I also appreciate Jo-Yann Tremolet De Villers, Will Shumate, and Pierson Rucker the energy you guys brought to the lab was amazing.

Lastly, I would like to thank my parents Kream John Mosley and Chantelle Brake Mosley thank you guys for always supporting me. I am also thankful for my Godparents, my beautiful girlfriend, and the rest of my family. None of this would've been possible without the support and love you all gave throughout my graduate career.

TABLE OF CONTENTS

LIST OF TABLES vii
LIST OF FIGURES viii

Chapter 1: Introduction 1
1.1 References 6

Chapter 2: The Unique Pathway of pABA Biosynthesis in C. trachomatis 10
2.1 Introduction 10
2.2 Methods 12
2.2.1 Expression of CADD 12
2.2.2 Expression of CADD with incorporation of isotopic labels 13
2.2.3 Purification of CADD 13
2.2.4 Anaerobic purification of CADD 14
2.2.5 Site-directed mutagenesis 15
2.2.6 Metal analysis 15
2.2.7 Note on isolation of metal-loaded and apo CADD 16
2.2.8 Activity assays 16
2.2.9 Targeted mass spectrometry data acquisition of pABA quantification 17
2.2.10 Targeted data processing 18
2.2.11 Untargeted mass spectrometry data acquisition 18
2.2.12 Untargeted data processing 19
2.2.13 Filter-aided sample preparation for proteomics 19
2.2.14 NanoLC-MS/MS for protein analysis 20
2.3 Results 22
2.3.1 Analysis of di-iron form of CADD 22
2.3.2 Metal dependence of pABA formation 23
2.3.3 Requirement of oxygen for CADD reactivity 24
2.3.4 CADD oxygen incorporation into pABA structure 25
2.3.5 CADD utilizes an endogenous “substrate” tyrosine 26
2.3.6 “Substrate” tyrosine at position 27 cleaved by CADD 28
2.3.7 CADD could undergo a proton coupled electron transfer (PCET) reaction 31
2.4 Discussion 32
2.4.1 CADD utilizes a Mn/Fe cofactor to synthesize pABA 32
2.4.2 Oxygen dependence of CADD reactivity 32
2.4.3 CADD utilizes Y27 as a precursor to pABA 33
2.4.4 Proposed CADD mechanism 33
2.5 Conclusion 35
2.6 References 37

Chapter 3: Assay Development for pABA Detection & Quantification 39
3.1 Introduction 39
3.2 Methods 42

3.2.1 Expression of CADD 42
3.2.2 Purification of CADD 42

3.2.3 CADD activity assay 42
3.2.4 Expression of NE1434 42
3.2.5 Purification of NE1434 43
3.2.6 Structural alignment of NE1434 and CADD 44
3.2.7 Ne1434 activity assay 44
3.2.8 Colorimetric pABA assay 44
3.2.9 pABA quantification by OPA reagent 45
3.3 Results45
3.3.1 Quantification of NE1434 45
3.3.1 pABA detection via colorimetric assay 46
3.3.2 Metal dependence studies with NE1434 48
3.4 Discussion 49
3.4.1 pABA detection by colorimetric assay 49
3.5 Conclusion 50
3.6 References 52

Chapter 4: Future Directions 53
4.1 The CADD project 53
4.2 The NE1434 project 54
4.3 Conclusion 55

LIST OF TABLES

Table 2.1 ICP results of CADD expressed in *E. coli* grown in M9 media 23

LIST OF FIGURES

Figure 1.1 Biosynthetic gene cluster for folate in *C. trachomatis* 3

Figure 1.2 Structural view of CADD diiron core 4

Figure 2.1 pABA biosynthesis in most bacteria 11

Figure 2.2 Metal dependence of CADD reactivity 24

Figure 2.3 Requirement of oxygen for CADD reactivity 25

Figure 2.4 CADD oxygen incorporation into pABA structure 26

Figure 2.5 CADD utilizes an endogenous "substrate" tyrosine 28

Figure 2.6 "Substrate" tyrosine at position 27 is cleaved by CADD 30

Figure 2.7 CADD could undergo a proton coupled electron transfer (PCET) reaction 32

Figure 2.8 Active site of CADD showing the diiron cluster and aromatic residues 34

Figure 2.9 Proposed CADD mechanism 35

Figure 3.1 Structural alignment of CADD and NE1434 40

Figure 3.2 Structure of para-amino benzoic acid 41

Figure 3.3 NE1434 SDS-PAGE gel 46

Figure 3.4 pABA detection via colorimetric assay 48

Figure 3.5 NE1434 metal dependence studies via colorimetric assay 49

Figure 3.6 Structural alignment of PCET aromatic residues in CADD and NE1434 50

Chapter 1: Introduction

Chlamydia trachomatis is the bacterial species that causes the commonly spread sexually transmitted infection known as chlamydia. The Center for Disease Control and Prevention reported more than one million cases of chlamydia in the United States in 2019.[1] There are also many cases of chlamydia reported in developing countries that lack treatment and diagnostic services. Chlamydia is a major health concern because of its adverse effects on reproduction. In women it can cause pelvic inflammatory disease, ectopic pregnancy, infertility, and chronic pelvic pain.[2] Also, untreated chlamydia can lead to abnormal discharge, pain or tenderness in the testicles, and blood in semen for men.[3]

Chlamydia an example of a disease caused by a pathogenic microbe. Pathogenic microbes are infectious when associated with a host organism.[4] There are two classes of pathogenic microbes - extracellular and intracellular. Both classes cause disease but utilize different pathways to infect a host. Extracellular pathogens rely on specific places of entry to infect a host.[4] Examples of extracellular pathogens include *Staphylococcus aureus* and *Escherichia coli*. These bacteria can cause infection through open wounds. Intracellular pathogens must make contact with the appropriate type of cell host in order to cause infection.[4] Examples of intracellular pathogens include *Brucella abortus* and *Chlamydia trachomatis*. These bacteria use a wide range of host cells to cause infection.[4]

C. trachomatis normally infects the single-cell columnar layer of the epithelium in the endocervix of women and the urethra of men.[2] Inside epithelial cells, *C. trachomatis* undergoes a unique developmental cycle that produces infective forms.[2] These forms are known as elementary bodies which then infect neighboring epithelial cells. *C. trachomatis* causes inflammation and redness at the sight of infection. However, despite causing inflammation, *C.*

trachomatis infection can remain undetected. Untreated *C. trachomatis* infection can cause irreversible damage to reproductive organs and facilitate the transmission of human immunodeficiency virus.[2] *C. trachomatis* infections are treatable with antibiotics.

A variety of antibiotics have been utilized to treat *C. trachomatis* infections. Examples of these antibiotics are doxycycline and azithromycin. Doxycycline belongs to the tetracycline class of antibiotics.[5] Tetracyclines interfere with bacterial protein synthesis by preventing the association of aminoacyl-tRNA with the bacterial ribosome.[5,6] Doxycycline has been prescribed to treat chlamydia for many years[5] and is the most prescribed medication for *C. trachomatis* infections. Azithromycin belongs to the macrolide class of antibiotics.[5] Macrolides inhibit bacterial protein synthesis by binding to and interfering with the assembly of the 50s large ribosomal subunit.[7] Azithromycin was introduced as an alternative to erythromycin and is prescribed to men and women who are not pregnant.[5] However, due to the rising threat of antimicrobial resistance researchers are searching for new means of treating disease causing microbes such as *C. trachomatis* infections.

Analysis of the *C. trachomatis* genome[8] revealed it lacks genes required for para-amino benzoic acid (pABA) biosynthesis which is an essential building block of folic acid.[8,9] Folic acid is a tripartite molecule which consist of a molecule of pABA with its amino end attached to a pteridine, and its carboxyl group attached to the α-amino group of glutamic acid.[10] Folic acid is used by microorganisms for the biosynthesis of DNA and amino acids.[10] Without the synthesis or acquisition of folic acid, microbes cannot survive. The genes *C. trachomatis* lacks are the *pabABC* genes which encodes the pabABC enzymes which are required to synthesize pABA.[8] C. trachomatis has been shown to contain folates different from those of host cells.[11] This suggest that *C. trachomatis* has its own unique pathway for folate biosynthesis. *C. trachomatis* contains

the ct610 gene in place of the *pabABC* genes that are typically required for pABA synthesis (Figure 1.1). The ct610 gene encodes a unique metalloenzyme known as the Chlamydia Protein Associating with Death Domains (CADD).

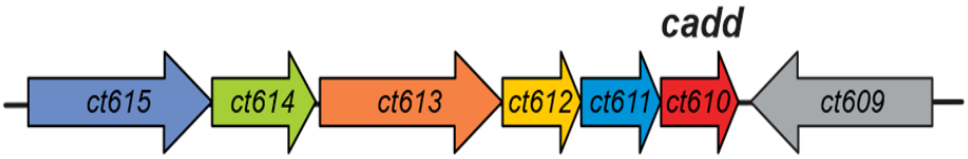

Figure 1.1: Biosynthetic gene cluster for folate in *C. trachomatis*. The C160 gene found in an operon that contains genes required by folate assembly.

CADD is part of a group of enzymes capable of catalyzing different types of chemistry. These enzymes are known as heme oxygenase-like diiron oxidases (HDOs). CADD was the first HDO to be structurally characterized.[19] However, the enzymatic mechanism of CADD remains elusive. Other examples of HDOs include the fatty acid decarboxylase UndA[12-14], the terminal alkene-amino acid-forming BesC[15-17] and the *N*-oxygenase SznF.[18-19] Interestingly, all the enzymes in this family have similar coordination motif of the diiron active site. Their active sites contain two histidine residues ligated to one iron, and one histidine ligand and two carboxylates ligated to the second iron.[17] Also, a carboxylate and a glutamate bridges the two irons (Figure 1.2). Even though these enzymes contain similar active sites, the chemistry they perform are diverse.

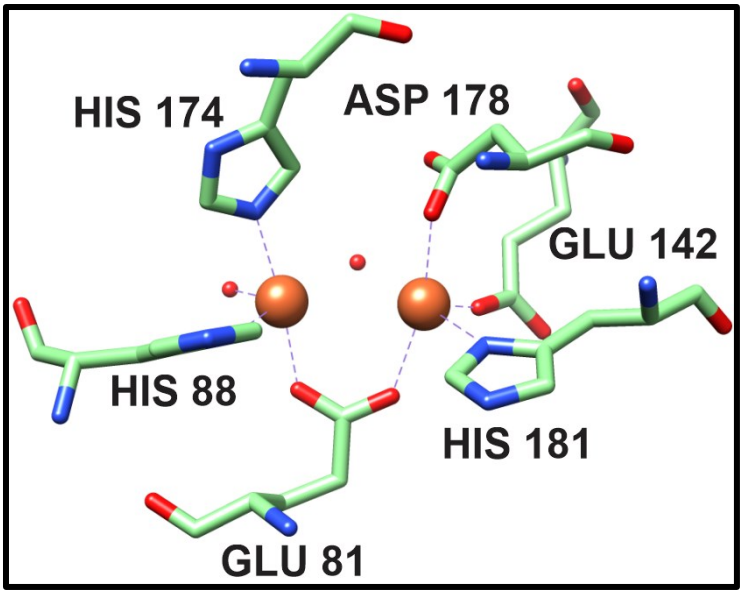

Figure 1.2: Structural view of CADD diiron core (PDB ID: 1RCW)[21]

Previous studies revealed CADD can interact with cell death domains to cause cell apoptosis.[21] This function of CADD is essential to the spread of *C. trachomatis* once it infects a host. Previous research also showed that CADD is able to rescue mutant *Escherichia coli* (*E. coli*) cells which lack the *pabABC* genes.[8] These findings suggest CADD has the ability to provide cells which lack the *pabABC* genes with the pABA needed for folate biosynthesis. Due to CADD being involved in the obligatory biosynthesis of folate, it may also serve as a therapeutic targe. An effective inhibitor could then lead to a drug that specifically targets infections caused by *C. trachomatis*.

There are antifolate drugs capable of inhibit folate biosynthesis. Examples of antifolate drugs include sulfamethoxazole and trimethoprim. Sulfamethoxazole is a mimic of pABA[20] while trimethoprim inhibits reduction of dihydrofolic acid to tetrahydrofolic acid.[20] However, these drugs are also known to cause adverse effects. These drugs can have neurologic effects, hypersensitivity reactions, affect pregnant women by causing reproductive toxicity, and they can

cause hyperkalemia.[22,23] Understanding the mechanism of CADD may lead to the development of treatments specific to *C. trachomatis*.

Previous analysis revealed that there are several homologs of CADD. These orthologs are distributed between prokaryotes and intracellular bacteria.[25] One of the orthologs was found in the *Nitrosomonas europaea* species which similarly lacks the *pabABC* genes like *C. trachomatis*.[23] *N. europaea* is an obligate chemolithotroph that derives all the reductant required for energy and biosynthesis from the oxidation of NH_3 to nitrite.[24] *N. europaea* is also used in the treatment of industrial and sewage waste.[24] This bacterium is also capable of degrading a variety of organic compounds.

In place of the *pabABC* genes, *N. europaea* contains the *NE1434* gene which encodes a metalloenzyme homologous to CADD. Previous research has suggested NE1434 could fulfill the same role as CADD.[25] NE1434 is the only ortholog of CADD that has been shown to be responsible to pABA biosynthesis. However, the mechanism CADD and NE1434 utilize to perform this unique chemistry has yet to be identified. The goal of this study is to study provides insight into the mechanisms of CADD and NE1434.

1.1 References

1. National Overview of STD's, 2019
2. Brunham, R. C.; Rey-Ladino, J., Immunology of Chlamydia infection: implications for a Chlamydia trachomatis vaccine. *Nature reviews immunology* **2005,** *5* (2), 149-161.
3. Miller, N., 5 Possible Long-Term Effects of Chlamydia Everlywell: Home Health Testing Made Easy 2021.
4. Silva, M. T., Classical labeling of bacterial pathogens according to their lifestyle in the host: inconsistencies and alternatives. *Frontiers in microbiology* **2012,** *3*, 71.
5. Guaschino, S.; Ricci, G., How, and how efficiently, can we treat Chlamydia trachomatis infections in women? *Best Practice & Research Clinical Obstetrics & Gynaecology* **2002,** *16* (6), 875-888.
6. Chopra, I.; Roberts, M., Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and molecular biology reviews* **2001,** *65* (2), 232-260.
7. Parnham, M. J.; Haber, V. E.; Giamarellos-Bourboulis, E. J.; Perletti, G.; Verleden, G. M.; Vos, R., Azithromycin: mechanisms of action and their relevance for clinical applications. *Pharmacology & therapeutics* **2014,** *143* (2), 225-245.
8. Adams, N. E.; Thiaville, J. J.; Proestos, J.; Juárez-Vázquez, A. L.; McCoy, A. J.; Barona-Gómez, F.; Iwata-Reuyl, D.; de Crécy-Lagard, V.; Maurelli, A. T., Promiscuous and adaptable enzymes fill "holes" in the tetrahydrofolate pathway in Chlamydia species. *MBio* **2014,** *5* (4), e01378-14.

9. Green, J. M.; Nichols, B. P., p-Aminobenzoate biosynthesis in Escherichia coli. Purification of aminodeoxychorismate lyase and cloning of pabC. *Journal of Biological Chemistry* **1991,** *266* (20), 12971-12975
10. Maynard, C.; Cummins, I.; Green, J.; Weinkove, D., A bacterial route for folic acid supplementation. *BMC biology* **2018,** *16* (1), 1-10.
11. Fan, H.; Brunham, R. C.; McClarty, G., Acquisition and synthesis of folates by obligate intracellular bacteria of the genus Chlamydia. *The Journal of clinical investigation* **1992,** *90* (5), 1803-1811.
12. Rui, Z.; Li, X.; Zhu, X.; Liu, J.; Domigan, B.; Barr, I.; Cate, J. H.; Zhang, W., Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proceedings of the National Academy of Sciences* **2014,** *111* (51), 18237-18242.
13. Manley, O. M.; Fan, R.; Guo, Y.; Makris, T. M., Oxidative decarboxylase UndA utilizes a dinuclear iron cofactor. *Journal of the American Chemical Society* **2019,** *141* (22), 8684-8688.
14. Zhang, B.; Rajakovich, L. J.; Van Cura, D.; Blaesi, E. J.; Mitchell, A. J.; Tysoe, C. R.; Zhu, X.; Streit, B. R.; Rui, Z.; Zhang, W., Substrate-triggered formation of a peroxo-Fe2 (III/III) intermediate during fatty acid decarboxylation by UndA. *Journal of the American Chemical Society* **2019,** *141* (37), 14510-14514.
15. Marchand, J.; Neugebauer, M.; Ing, M.; Lin, C.-I.; Pelton, J.; Chang, M., Discovery of a pathway for terminal-alkyne amino acid biosynthesis. Nature **2019**, 567 (7748), 420-424.
16. Neugebauer, M. E.; Sumida, K. H.; Pelton, J. G.; McMurry, J. L.; Marchand, J. A.; Chang, M. C., A family of radical halogenases for the engineering of amino-acid-based products. Nature chemical biology **2019**, 15 (10), 1009-1016.

17. Manley, O. M.; Tang, H.; Xue, S.; Guo, Y.; Chang, W.-c.; Makris, T. M., BesC Initiates C–C Cleavage through a Substrate-Triggered and Reactive Diferric-Peroxo Intermediate. *Journal of the American Chemical Society* **2021,** *143* (50), 21416-21424.
18. McBride, M. J.; Sil, D.; Ng, T. L.; Crooke, A. M.; Kenney, G. E.; Tysoe, C. R.; Zhang, B.; Balskus, E. P.; Boal, A. K.; Krebs, C., A peroxodiiron (III/III) intermediate mediating both N-hydroxylation steps in biosynthesis of the N-nitrosourea pharmacophore of streptozotocin by the multi-domain metalloenzyme SznF. *Journal of the American Chemical Society* **2020,** *142* (27), 11818-11828.
19. McBride, M. J.; Pope, S. R.; Hu, K.; Okafor, C. D.; Balskus, E. P.; Bollinger, J. M.; Boal, A. K., Structure and assembly of the diiron cofactor in the heme-oxygenase–like domain of the N-nitrosourea–producing enzyme SznF. *Proceedings of the National Academy of Sciences* **2021,** *118* (4).
20. Majewsky, M.; Wagner, D.; Delay, M.; Bräse, S.; Yargeau, V.; Horn, H., Antibacterial activity of sulfamethoxazole transformation products (TPs): general relevance for sulfonamide TPs modified at the para position. *Chemical research in toxicology* **2014,** *27* (10), 1821-1828.
21. Schwarzenbacher, R.; Stenner-Liewen, F.; Liewen, H.; Robinson, H.; Yuan, H.; Bossy-Wetzel, E.; Reed, J. C.; Liddington, R. C., Structure of the Chlamydia protein CADD reveals a redox enzyme that modulates host cell apoptosis. *Journal of Biological Chemistry* **2004,** *279* (28), 29320-29324.
22. Huovinen, P., Trimethoprim resistance. *Antimicrobial agents and chemotherapy* **1987,** *31* (10), 1451-1456.

23. Ho, J. M.-W.; Juurlink, D. N., Considerations when prescribing trimethoprim–sulfamethoxazole. *Cmaj* **2011,** *183* (16), 1851-1858.

24. Arp, D. J.; Sayavedra-Soto, L. A.; Hommes, N. G., Molecular biology and biochemistry of ammonia oxidation by Nitrosomonas europaea. *Archives of microbiology* **2002,** *178* (4), 250-255.

25. Satoh, Y.; Kuratsu, M.; Kobayashi, D.; Dairi, T., New gene responsible for para-aminobenzoate biosynthesis. *Journal of bioscience and bioengineering* **2014,** *117* (2), 178-183.

Chapter 2: The Unique Pathway of pABA Biosynthesis in *C. trachomatis*

Abstract

The Chlamydia Protein Associating with Death Domains (CADD) is found in *C. trachomatis* and is essential to the biosynthesis of folic acid. *C. trachomatis* lacks the *pabABC* genes, which are required by most bacteria that synthesize *para*-aminobenzoic acid (pABA) de novo. Knockout studies showed CADD can rescue mutant *E. coli* cells that lack the *pabABC* genes.[1] This suggests that CADD can rescue this phenotype and supplant the role of pabABC enzymes. We believe CADD provides *C. trachomatis* with pABA needed for survival. However, previous enzyme activity assays performed with CADD produced little product.[2] We performed activity assays with the diiron form of CADD and observed very little pABA formation. We concluded that diiron form of CADD may not be its active form. Our study utilizes LC-MS and protein MS to analyze CADD enzyme activity assays. We show that CADD utilizes a Fe/Mn cofactor to synthesize pABA. Using these techniques, we also show the oxygen dependence of CADD reactivity. We propose that CADD undergoes a proton coupled electron transfer reaction similar to ribonucleotide reductases (RNR). This radical transfer reaction results in the cleavage of a Y27, which undergoes further processing by CADD to form pABA. Our results provide insight into the CADD mechanism which has been elusive thus far.

2.1 Introduction

Metalloenzymes are a large family of enzymes that use a metal cation as a cofactor and are essential to a wide range of essential biological activities.[3] CADD is a metalloenzyme found in *C. trachomatis* which we believe to be essential to the survival of the species. *C. trachomatis* is an intracellular pathogen that causes a wide range of sexually transmitted infections.[4-5] Diseases caused by *C. trachomatis* are usually treatable with antibiotics. However due to the rise

of antimicrobial resistance a new way of treating microbial disease is needed.[6-7] It was previously discovered that *C. trachomatis* lacks the *pabABC* genes which are used by most bacteria for pABA synthesis.

pABA is a growth factor needed by all microorganisms.[8] In bacteria, pABA is used for the synthesis of folic acid. Folic acid is required for the synthesis of bacterial DNA and amino acids.[1] Folic acid is a tripartite molecule, and its structure consists of a pterin ring, pABA and glutamate. pABA is an essential building block of folate and its synthesis in most bacteria involve, the *pabABC* enzymes (Figure 2.1). The *pabABC* enzymes are involved in the pABA synthesis from chorismate. However, instead of containing the *pabABC* genes *C. trachomatis* contains the *ct610 gene* which encodes CADD (Figure 1.1).

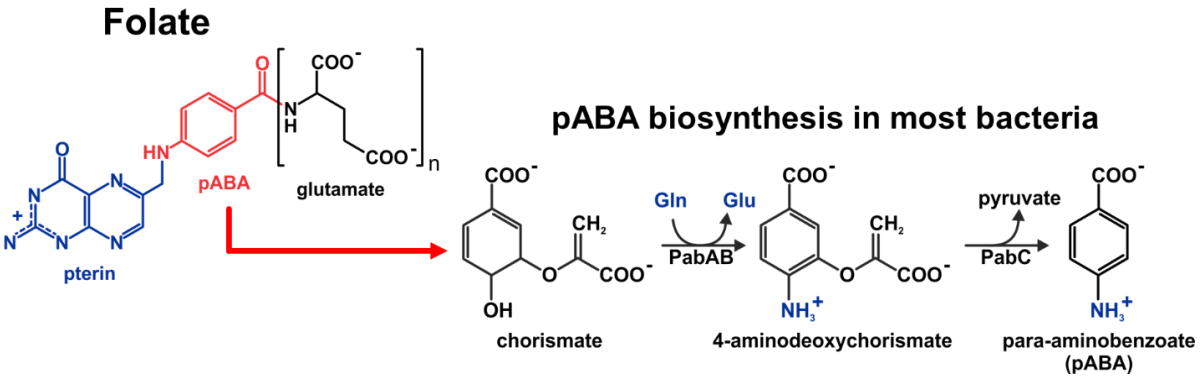

Figure 2.1: pABA is an essential building block of folate and pABA biosynthesis in most bacteria involve the *pabABC* enzymes.

Previous research demonstrated the ability of CADD to rescue *E. coli* cells which lacked the *pabABC* genes.[1-9] Due to not having the *pabABC* genes the *E. coli* cells were not able to proliferate. However, when the ct610 gene was transformed into the cells, cell growth improved exponentially, and the introduced gene rescued pABA auxotrophy. A previously solved structure of CADD revealed it contained a di-iron center,[10] and this center was believed to be essential to catalytic activity. Recent research has also demonstrated that L-tyrosine is the metabolic

precursor pABA.[2] This previous research performed activity assays with CADD. They lost enzymatic all activity when specific tyrosine residues on the protein's surface was mutated. CADD was deemed a "Self-Sacrificing" enzyme cleaving its own Tyr to synthesize pABA. However, enzyme activity assays yielded 0.004 molar equivalents of pABA. This suggests even though CADD can form a diiron core, it may not be the active form of the enzyme.

A recently discovered RNR found in *C. trachomatis* that utilizes a Mn/Fe cofactor for radical initiation.[11] The RNR found in *C. trachomatis* lacks a conserved Tyr residue that is found in class 1 RNR. Due to there being a Mn:Fe RNR in *C. trachomatis* it was speculated that CADD could also utilizes a Mn:Fe cofactor to synthesize pABA. This study provides insight into the unique pathway of pABA biosynthesis in *C. trachomatis*.

2.2 Methods

2.2.1 Expression of CADD

The *cadd* gene (*ct610*) was codon-optimized and synthesized by Bio Basic 2.0. The gene was subcloned into pET28b with NdeI and XhoI restriction enzymes for expression with a C and N terminal hexahistidine tag. The final coding sequence follows with restriction sites underlined.

CATATGATGAACTTCCTGGATCAGCTGGATCTGATCATCCAGAACAAACACATGCTG
GAACACACCTTCTACGTTAAATGGTCTAAAGGGGAACTGACCAAAGAACAGCTGCA
GGCGTACGCGAAAGATTACTACCTGCACATCAAAGCGTTCCCGAAATACCTGAGCG
CGATCCACTCTCGTTGCGATGATCTGGAAGCGCGTAAACTGCTGCTGGATAACCTGA
TGGATGAAGAAAACGGCTACCCGAACCACATCGATCTGTGGAAACAGTTCGTTTTCG
CGCTGGGTGTTACCCCGGAAGAACTGGAAGCGCACGAACCGTCTGAAGCGGCGAAA
GCGAAAGTTGCGACCTTCATGCGTTGGTGCACCGGCGATAGCCTGGCGGCGGGCGTT
GCGGCGCTGTACTCTTACGAAAGCCAGATCCCGCGTATCGCGCGTGAAAAAATCCG

TGGCCTGACCGAATATTTCGGCTTTAGTAACCCTGAAGATTATGCATACTTTACTGA
ACATGAAGAAGCTGACGTTAGACACGCACGTGAAGAGAAAGCGCTGATTGAAATGC
TGCTGAAAGATGACGCGGATAAAGTCTTAGAAGCTAGTCAGGAAGTGACCCAATCC
CTGTACGGGTTTCTAGATTCTTTTCTTGATCCGGGTACCTGCTGTAGCTGCCACCAGT
CTTACCTCGAG

Protein expression was carried out in *E. coli* BL21(DE3) cultures (1 L each) grown in M9 media supplemented with 2 g/L casamino acids, 50 μg/mL kanamycin, and 50 μM of the desired metal ($FeCl_3$, $MnCl_2$, or both). M9 media was used for precise control of the metal content of the cultures. The cells were grown with shaking at 37 °C to an OD_{600} of 0.6-0.8, at which point the cultures were cooled to 25 °C and protein expression was induced with the addition of 200 μM IPTG. The cell cultures were allowed to incubate while shaking overnight. The cells were harvested by centrifugation roughly 16 hours later, and the cell pellets were stored at -20°C until further use.

2.2.2 Expression of CADD with incorporation of isotopic labels

Modifications to the above expression conditions were made for isotopic label incorporation as needed. To isotopically label the protein Tyr residues, casamino acids were omitted from the culture medium, and the cultures were supplemented instead with 10 mg/L L-Tyr-(phenyl-3,5-d_2).

2.2.3 Purification of CADD

Thawed cell paste was resuspended in buffer comprised of 25 mM HEPES pH 7.5, 100 mM KCl, 10 mM imidazole, and 5% glycerol (~10 mL buffer per g cell paste) and chilled on ice. Cells were lysed by sonication and then centrifuged to pellet cell debris. The cleared lysate was loaded onto a Ni-NTA column while maintained at 4 °C (as was the remainder of the

purification). The resin was washed with 5 column volumes (CV) of the above buffer containing 20 mM imidazole, and protein was eluted with 5 CV of buffer with 500 mM imidazole. The fractions containing CADD were identified by an absorbance at 280 nm and by SDS-PAGE. The concentration of CADD was determined by dividing the absorbance at 280 nm by the extension coefficient (36 $M^{-1}cm^{-1}$). Fractions containing CADD were pooled, concentrated, and flash frozen in liquid nitrogen and stored at -70 °C until further use. Typical yields of CADD were ≥200 mg purified protein per L bacterial culture regardless of the type of medium used.

2.2.4 Anaerobic purification of CADD

Due to initial difficulty in obtaining pABA activity using aerobically purified, diiron-loaded CADD, the enzyme was expressed from Fe-containing M9 cultures was purified anaerobically in a Coy anaerobic chamber. Purification buffers were made anaerobic by sequential evacuations and refills with argon gas via Schlenk line prior to transferring to the anaerobic chamber, where they sat uncapped and stirring overnight prior to usage the following day. Buffers used were identical to those of the aerobic purification except that 0.5 mM sodium dithionite was added to the buffers immediately before use. Cell paste was thawed and resuspended in the anaerobic chamber, and the sonication was kept cold using sealed ice packs. The lysate was sealed in centrifuge tubes and removed from the chamber for centrifugation. The tubes were then returned to the anaerobic chamber and the remainder of the purification was performed as described above. Purified protein was sealed in gas-tight vials prior to removal from the anaerobic chamber to be frozen and stored for later use.

2.2.5 Site-directed mutagenesis

Tyrosine- and tryptophan-to-phenylalanine mutations were made using the semi-overlapping primers listed below. The mutations were incorporated by a whole-plasmid

amplification PCR strategy using Q5 DNA polymerase. The PCR product was treated with DpnI to digest the parent plasmid and then transformed into chemically competent *E. coli* DH5α cells. Colonies were screened for the mutation by culturing, miniprepping, and sequencing. Plasmid confirmed to contain the desired mutation was used for protein expression as described for WT CADD.

Y27F Primers:

Forward 5'-CCTTCTTCGTTAAATGGTCTAAAGGGGAACTGACC-3'

Reverse 5'- CAAACACATGCTGGAACACACCTTCTTCGTTAAATGG-3'

W92F Primers:

Forward 5'- CACATCGATCTGTTCAAACAGTTCGTTTTCGCGCTGGGTGTTACCC-3'

Reverse 5'- CGAACTGTTTCAACAGATCGATGTGGTTCGGGTAGCCGTTTTCTTCATCC-3'

2.2.6 Metal analysis

To determine the metal content of purified protein, 100 μM CADD was acidified in 2 M HCl. The denatured protein was pelleted by centrifugation. The supernatant was given a Ge/Y/In/ Pr internal standard solution (50 ppb, final concentration), diluted appropriately, and measured for Fe, Mn, Zn, Co, and Ni, content using an Thermo Scientific iCAP RQ inductively coupled plasma mass spectrometer equipped with an Elemental Scientific SC 2DX Autosampler via the METRIC facility at North Carolina State University. Quantitation was achieved by comparison to a standard curve of the appropriate metal. Alternatively, the supernatant of the digested protein sample was also analyzed by ferrozine assay[12] for Fe content when appropriate.

2.2.7 Note on isolation of metal-loaded and apo CADD

CADD isolated from cultures grown with Fe supplementation yielded 50% loaded diiron CADD as determined by metal quantification (Table S1). This was supplemented with an additional equivalent of Fe (from an anaerobic solution ferrous ammonium sulfate or ^{57}Fe(II) made in 25 mM HEPES pH 7.5, 100 mM KCl, and 5% glycerol) for diiron-loaded CADD. By contrast, CADD from cultures grown with Mn supplementation purified with very little bound metal (Table S1), thus we consider this virtually metal-free form to be *apo* CADD. This preparation of *apo* CADD was given Fe(II) and/or Mn(II) for turnover reactions as indicated.

2.2.8 Activity assays

Single-turnover CADD reactions were carried out in Ni-NTA elution buffer (25 mM HEPES pH 7.5, 100 mM KCl, 500 mM imidazole, 5% glycerol) and were generally comprised of 100 μM diiron-loaded CADD or *apo* CADD (the preparations of which is described above) and 0-2 molar equivalents of metal (via anaerobic stock solutions of Fe(II) and/or Mn(II)). Experiments were carried out using different ratios of Fe:Mn to determine the optimal ratio for pABA production. Multiple turnovers were achieved with the inclusion of 1.25 mM ascorbic acid along with 1:1 Fe:Mn. The reactions were incubated at 37°C for ~12 hours under an aerobic atmosphere and then quenched with an equal volume of Acetonitrile (ACN). The protein aggregate was pelleted by centrifugation, and the supernatant was stored at -20°C before analysis for pABA quantification.

To probe the identity of the cosubstrate, multiple turnover reactions were performed with the addition of 80 μM catalase or 10 μM bovine superoxide dismutase using the optimal metal ratio of 1:1 Fe:Mn. Anaerobic reactions were performed by degassing all buffer, protein, and reaction components in sealed vials via Schlenk line and then transferring into an anaerobic

chamber to carry out the reaction. Experiments using isotopically labeled 99% $^{18}O_2$ were carried out by the same procedure as the anaerobic experiments but instead of maintaining anaerobicity, the reaction samples were evacuated and refilled with $^{18}O_2$ gas while stirring. The product was analyzed by untargeted high-resolution mass spectrometry as described below.

To determine the origin of the substrate Tyr, multiple turnover reactions were performed as described above using the optimal metal ratio of 1:1 Fe:Mn, with or without the addition of 200 μM exogenous L-tyrosine or labeled L-tyrosine-(phenyl-3,5-d_2). Reactions were also carried out with isotopically enriched CADD (as described above) with or without the addition of exogenous unlabeled L-tyrosine. The samples were analyzed by untargeted high-resolution mass spectrometry for pABA quantification and proteomics mass spectrometry for examination of peptide modifications, as described below.

2.2.9 Targeted mass spectrometry data acquisition for pABA quantification

Targeted analysis of CADD reaction samples was performed using a Thermo Vanquish LC instrument (Thermo Fisher Scientific, San Jose, CA) coupled to a Thermo TSQ Altis triple quadrupole mass spectrometer with a heated electrospray ionization (HESI) source. Chromatographic separation was achieved on a Thermo Accucore aQ C18 column (2.1 x 100 mm, 2.6 mM) maintained at 40°C. The following linear gradient of mobile phase A (H_2O + 0.1% FA) and mobile phase B (MeOH + 0.1% FA) was used: 0-1 min (5%B, 0.4 mL/min), 1-2 min (5-10%B, 0.4 mL/min), 2-6 min (10-15%B, 0.4 mL/min), 6-9.5 min (15-50%B, 0.4 mL/min), 9.51-12 min (50-95%B, 0.4 mL/min, 12.01-12.02 min 95-5%B, 12.02-15 min (5%B). For quantification, samples were compared to a pABA standard calibration curve in the linear range of 125 nM to 100 mM. Both samples and standards were analyzed (5 mL injections) in positive ion mode (spray voltage 3.5 kV, ion transfer tube temperature 300 °C, vaporizer temperature

350°C, sheath gas 50 a.u., aux gas 10 a.u., sweep gas 1 a.u.) using a Q1 resolution of 0.7 m/z, a Q3 resolution of 1.2 m/z. The following multiple reaction monitoring (MRM) transitions and collision energies were used: 138.1 → 120.1 (pABA quantifier, CE 13), 138.1 →94.1 (pABA qualifier, CE 16), 138.1 → 77.0 (pABA qualifier, CE 22), 140.1 → 122.1 (d_2-pABA quantifier, CE 13), 140.1 →96.1 (d_2-pABA qualifier, CE 16), 140.1 → 79.0 (d_2-pABA qualifier, CE 22), 140.1 → 122.1 (^{18}O-pABA (1) quantifier, CE 13), 140.1 →120.1 (^{18}O-pABA (1) or (2) qualifier, CE 16), 140.1 → 77.0 (^{18}O-pABA (1) or (2) qualifier, CE 22), 142.1 → 122.1 ($^{18}O_2$-pABA quantifier, CE 13), 142.1 →94.1 ($^{18}O_2$-pABA qualifier, CE 16), 142.1 → 77.0 ($^{18}O_2$-pABA qualifier, CE 22).

2.2.10 Targeted data processing

Peak integration and quantification were performed in Skyline.[13] Standard curves were constructed using peak areas from the quantifier transitions listed above for commercially bought pABA standards of known concentration. The concentrations of each analyte in the unknown samples were calculated in an identical manner relative to the regression line and the natural abundance. The peak area of the qualifier transition was compared with the peak area of the quantifier transitions to generate an ion ratio used for compound validation. Comparison of the ion ratios for the standards to the ion ratio of the unknowns' samples was carried out with a threshold of 30%.

2.2.11 Untargeted mass spectrometry data acquisition

Samples that were processed with stable isotope labeled protein or cosubstrate (3,5-d_2-tyrosine or $^{18}O_2$, respectively) were subjected to further confirmatory analysis using untargeted high resolution mass spectrometry. These measurements were performed using a Thermo Vanquish LC instrument (Thermo Fisher Scientific, San Jose, CA) coupled to a Thermo Orbitrap

Exploris 480 mass spectrometer (Thermo Fisher Scientific) with a heated electrospray ionization (HESI) source. Chromatographic separation was achieved using identical conditions to the targeted methods. Samples were analyzed (5 mL injections) in positive ion mode (spray voltage 3.5 kV, ion transfer tube temperature 300°C, vaporizer temperature 350°C, sheath gas 50 a.u., aux gas 10 a.u., sweep gas 1 a.u.) using a top5 ddMS2 method (RF Lens 50, scan range m/z 50-500, MS1 120K RP, MS2 30K RP, CE 50).

2.2.12 Untargeted data processing

High resolution tandem mass spectrometry data was visualized using Thermo Freestyle to confirm the molecular formulas of the d_2- and ^{18}O-containing pABA products. Spectral intensities were exported from FreeStyle for figure generation.

2.2.13 Filter-aided sample preparation for proteomics

A modified filter-assisted-sample-preparation (FASP) procedure was adopted for proteomics workflow. Total protein amount in each sample (described above in *Activity Assays*) was determined using a BCA kit (Thermo), and a volume equal to 200 µg protein was pipetted out and reconstituted to 200 µL with 50 mM ABC, 1% SDC solution. A 15 µL volume of 50 mM DTT solution in 0.1 M Tris pH 8.0 was added to each sample, and the protein solution was then incubated at 56 ºC for 30 min to reduce disulfide bonds. After adding 200 µL of 8 M urea in 0.1 M Tris pH 8.0, the solution was placed on a 30 kDa molecular weight cutoff filter and centrifuged at 12,000 *g* for 15 min at 21 °C. To the filter was added 200 µL of 8 M urea and 64 µL of 200 mM IAA in 8M urea containing 0.1 M Tris pH 8.0. The samples were placed in the dark at ambient temperature for 1 hr. Solvent was removed by centrifugation as before. The sample was washed 3 consecutive times with 2 M urea, 10 mM calcium chloride in 0.1 M Tris pH 8.0 with centrifugation at 125,000 *g* for 10 min at 21 °C between washes. This was followed by 3 consecutive washes

with 0.1 M Tris pH 7.5, once again with centrifugation at 130,000 *g* for 15 min at 21 °C between washes. The sample was digested on the filter in a 37 °C oven overnight with trypsin added in 0.1 M Tris pH 7.5 at an trypsin/protein ratio of 1:50. The digestion process was quenched using 50 µL 0.001% Zwitttergent 3-16 in 1% formic acid (quench buffer) followed by centrifugation at 13,500 *g* for 30 min. This step was repeated using 400 µL quench buffer followed by centrifuge at 14,000 *g* for 10 min. Eluents were collected, lyophilized, and stored in -20 °C freezer until nanoLC-MS/MS analysis was performed.

2.2.14 NanoLC-MS/MS for protein analysis

Analysis was performed on a Thermo EASY nano-LC 1200 coupled to an Orbitrap Exploris 480 with EASY-spray ion source (Thermo Scientific, Bremen, Germany). A trap column - Acclaim PepMap™ 100 (75 um x 2 cm, C18, 3 um, 100 Å) - was set up in-line with an analytical column - PepMap™ RSLC C18 (75 um x 25 cm, 2 um, 100 Å, held at 45 °C) - both from Thermo Fisher Scientific (San Jose, CA). This “trap-and-elute” configuration facilitated an additional cleaning step for the samples prior to nanoLC separation of tryptic peptides. Sample protein digests were reconstituted in 200 µL of 98% water, 2% ACN containing 0.1% formic acid and 2 µL of each sample were injected on-column. Mobile phases were 98% water, 2% ACN containing 0.1% formic acid (Mobile Phase A or MPA) and 80% ACN, 20% water containing 0.1% formic acid (Mobile Phase B or MPB). Samples were loaded onto the trap column at 1 µL/min for a total loading volume of 6 µL. The following solvent nanoLC gradient was used for the protein digests: initial condition of 5% MPB was held for 2 min, increased to 25% MPB over 100 min, increased to 40% MPB over 15 min, increased to 95% MPB over 1 min, and held at 95% MPB for 17 min. Before each injection, the trap column was equilibrated with 3 µL mobile phase and the analytical column equilibrated at 600 bar with 6 µL mobile

phase at initial composition. The autosampler was programmed with a custom wash sequence to rinse the needle and components with 80% isopropanol, MPB, and MPA between injections. Standard BSA digest and blanks (MPA) are run in between every 3 samples and several standard HeLa digest runs are conducted at the beginning and end of experiments to monitor the instrument performance.

The mass spectrometer was operated in positive ion mode using full MS scan and data dependent acquisition (DDA) MS/MS using the following settings: spray voltage of 1900 V, capillary temperature of 275°C, funnel RF level of 40 V, *m/z* 375–1600 for precursor scan at 120k resolving power, 300% normalized AGC target, maximum injection time of 120 ms, precursor dynamic exclusion of 20 s, MS^2 resolving power at 15k, cycle time of 1.5 s, 100% normalized AGC target for MS^2, 21 ms MS/MS maximum injection time, isolation window of *m/z* 1.5, and 30% normalized collision energy.

Raw nanoLC-MS/MS data were analyzed using Proteome Discoverer 2.4.0.305 (Thermo Scientific, San Jose, CA) to identify the CADD protein and verify the presence and location of phenylalanine, deuterated tyrosine, or the conversion of tyrosine to glycine. 2 FASTA files were used in database searching, one for the *Escherichia coli* (4,201 sequences) proteome and one for the *Chlamydia trachomatis* (889 sequences) proteome. Sequences corresponding to human keratins and porcine trypsin were also searched as possible contaminants. A maximum of 2 missed cleavage sites were allowed with a minimum peptide length of 6 residues and a maximum length of 144. The precursor mass tolerance was set to 5 ppm and the fragment mass tolerance was 0.02 Da. The weights of *b* and *y* ions were set to 1 and that of *a* and *x* ions set to 0.1. Oxidation of methionine (+15.995 Da), conversion of tyrosine to phenylalanine (-15.995 Da), deuteration of tyrosine (+2.013 Da), and the conversion of tyrosine or d_2-tyrosine to glycine (-106.042 or -

108.042 Da, respectively) were included as dynamic modifications, and a maximum of 3 equal modifications per peptide were allowed. Protein *N*-terminus modifications allowed were addition of acetyl group, loss of methionine, and loss of methionine with addition of acetyl group. Carbamidomethylation of cysteine was set as a static modification. The Percolator node was used with the following settings: concatenated target/decoy selection, validation based on *q*-value, and FDR targets of 0.01 for strict and 0.05 for relaxed false discovery rates. Data were imported into Skyline v.21.1.0.278 to extract ion chromatograms of targeted peptides and integrate the corresponding chromatographic peaks. Semiquantification was done by comparing the relative peak area of the modified ion from reactions of Fe:Mn CADD to that of a presumably unmodified *apo* CADD. Freestyle 1.6 software (Thermo Scientific) was also used in data analysis.

2.3 Results

2.3.1 Analysis of di-iron form of CADD

CADD was heterologously expressed in *E. coli* grown in minimal media supplemented with iron or manganese. After the CADD was purified it was analyzed via ICP-MS for quantification of metal content. CADD grown in the presence was 50% loaded with iron (Table 2.1). CADD grown in the presence of Mn had low concentrations of Mn. This suggested Mn is labile during purification. Interestingly CADD grown in the presence of Fe and Mn was 60% loaded with Fe (Table 2.1). This suggested CADD has a preference to bind Fe. This is consistent with the previously solved structure which revealed, CADD contains a diiron core.[10]

Table 2.1: ICP results of CADD expressed in *E. coli* grown in M9 media supplemented with iron or manganese.

	Metal loading (mol metal/mol protein)			
Media	Fe	Mn	Zn	Ni
M9 + Fe	1.00	0.00	0.11	0.16
M9 + Mn	0.01	0.01	0.04	0.06
M9 + Fe & Mn	0.61	0.04	0.01	0.23

2.3.2 Metal dependence of pABA formation

In vitro reactions carried out with the diiron form of CADD yielded much less than 0.1 molar equivalents of pABA (Figure 2.2 A). However, *in vitro* reactions carried out with the dimanganese form of CADD displayed a small product peak consistent with pABA (Figure 2.2 A). Due to there being a previously discovered RNR in *C. trachomatis* that utilized a Fe/Mn cofactor[11], we decided to carry out reactions with varying rations of Fe/Mn. Maximum activity was observed was observed with a Fe/Mn molar ratio of 1:1 (Figure 2.2 B). Stoichiometric pABA formation was observed when reactions were carried out in the presence of excess reductant (Figure 2.3 A). Also, when reactions were performed with excess Tyr pABA formation did not change significantly (Figure 2.3A). This suggested that CADD does not require an exogenous substrate for catalytic activity. Overall, these findings show that CADD requires an Fe/Mn cofactor for activity.

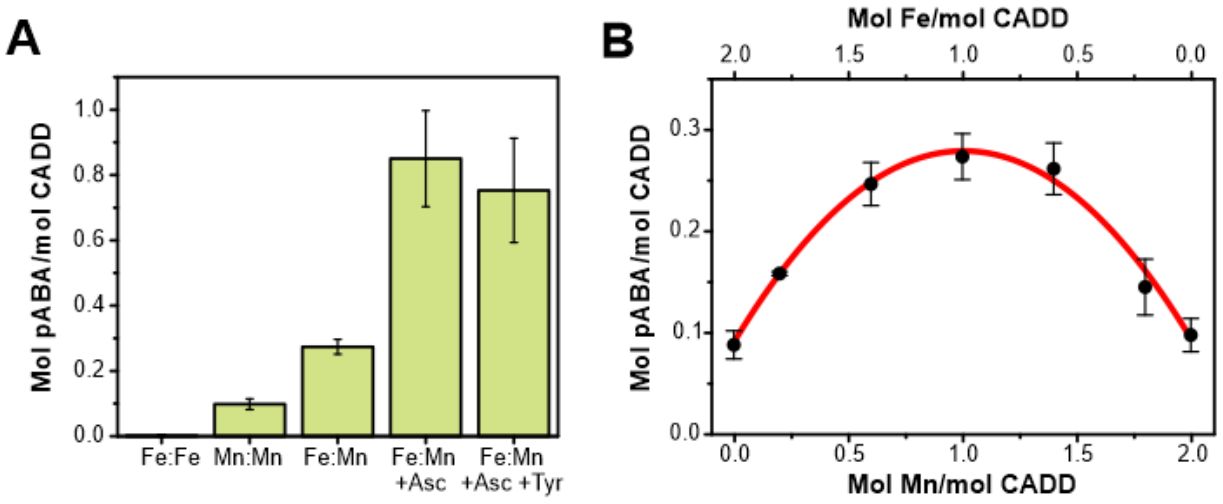

Figure 2.2: A) *In vitro* reactions of CADD expressed with Mn, reconstituted with Mn or Fe:Mn, Fe:Mn reactions were also provided ascorbic acid and exogenous Tyr. pABA production is reported as molar equivalents of pABA produced per protein B) pABA formation from the reaction of CADD reconstituted with varying ratios of Fe:Mn without excess reductant. Error bars represent the standard deviations of three different reactions.

2.3.3 Requirement of oxygen for CADD reactivity

After determining that CADD requires a Fe/Mn cofactor to synthesize pABA, then we asked if the CADD reaction required oxygen as a cosubstrate. CADD reactions were performed in the absence of O_2. These reactions yielded less than 0.1 molar equivalents of pABA (Figure 2.3). These findings suggested that CADD does in fact require oxygen for catalysis. To rule out alternative oxidants such as superoxide or hydrogen peroxide reactions were carried out using superoxide dismutase (SOD) and catalase. SOD is a scavenger of superoxide and catalase is a decomposer of hydrogen peroxide. These scavengers did not abolish pABA formation which confirmed that neither superoxide nor hydrogen peroxide is utilized by CADD as a cosubstrate (Figure 2.3). These findings allowed us to conclude that CADD utilizes O_2 as a cosubtrate.

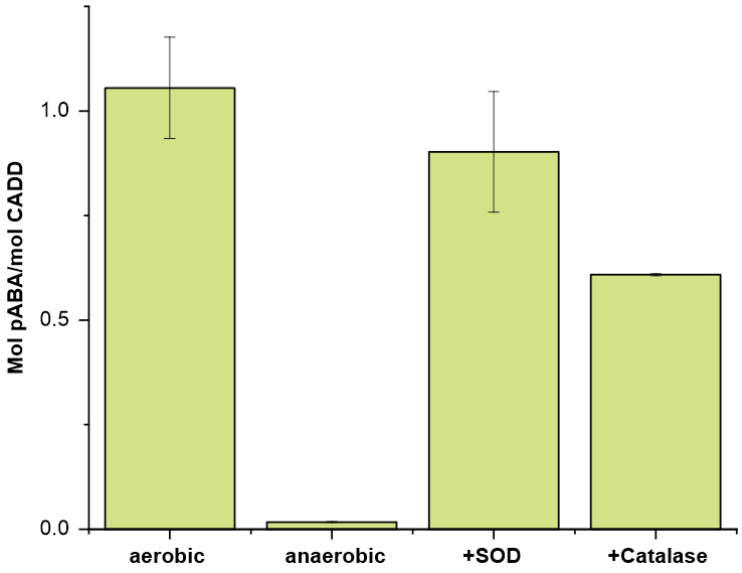

Figure 2.3: *In vitro* reaction results of Fe:Mn CADD with excess ascorbic acid under aerobic or anaerobic conditions and with the inclusion of 10 μM SOD or 10 mg/mL catalase. pABA production is reported as molar Equiv. of pABA produced per μM protein. Error bars represent the standard deviations of three different reactions.

2.3.4 CADD oxygen incorporation into pABA structure

To form a complete pABA, molecule the reaction of CADD would also necessitate the oxygenation of the Cβ of the substrate Tyr to form a carboxylate. To determine whether the oxygen atoms derived from O_2 isotopic tracer studies were performed with $^{18}O_2$. Native pABA has a product peak with m/z = 138. Incorporation of two $^{18}O_2$ molecules would result in a +2 increase per oxygen molecule of native product peak. Untargeted MS analysis of product peaks revealed products with masses of 140.1 and 142.1 (Figure 2.4 B, C). This is consistent with the incorporation of one and two $^{18}O_2$ atoms. The major product of the reactions was the incorporation of two integrated ^{18}O atoms ($^{18}O_2$-pABA) into the structure of pABA. However, a small percentage of products consisted with the incorporation of both $^{16}O_2$ and $^{18}O_2$ into the structure of pABA (Figure. 2.4 A). This could be due to the oxygen atoms exchanging with water molecules or it could be due small amounts of O^{16} contamination. Ultimately, these results

allow us to conclude that the carboxylate oxygens of pABA come from oxygen insertion. These oxygens are inserted by the Fe/Mn cofactor of CADD during multiple rounds of O_2 activation.

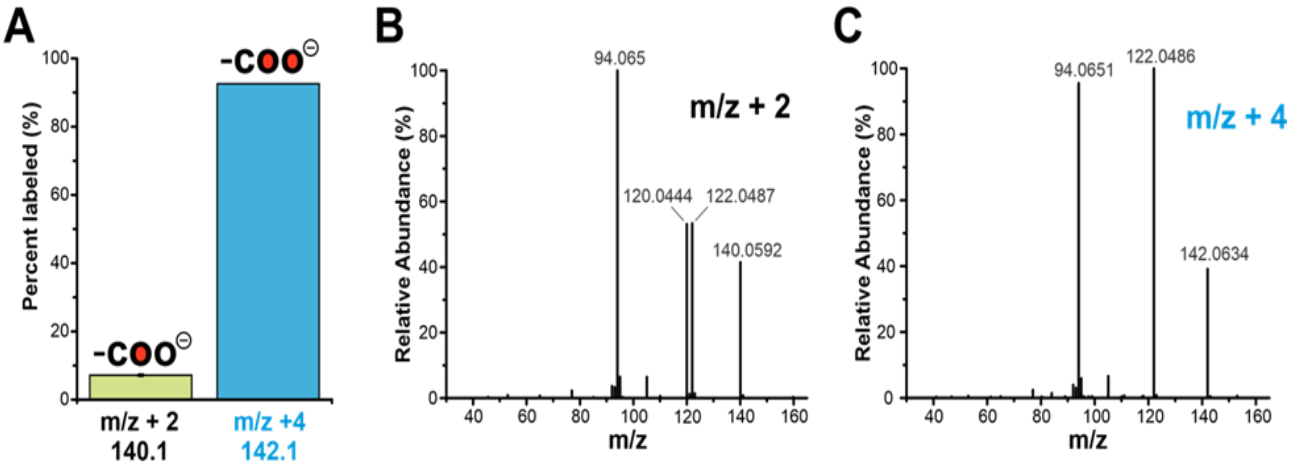

Figure 2.4: A) Relative extent of $^{18}O_2$ isotope incorporation into the product pABA and the corresponding mass spectra (**B** and **C**)

2.3.5 CADD utilizes an endogenous "substrate" tyrosine

Our previous results from the metal dependence reactions showed that the addition of free Tyr does not enhance pABA formation (Figure. 2.2 A). This implies the substrate may come from a protein derived Tyr as was hypothesized by previous research.[2] To confirm this hypothesis, CADD was expressed in M9 media supplemented with 3,5-d_2-L-tyrosine. CADD from this expression was purified and utilized in subsequent turnover reactions. This expression condition labeled majority of the Tyr residues on CADD with two deuterium atoms (Figure 2.5 A). If CADD utilizes a protein derived Tyr as a "substrate" there would be a shift in pABA product peak (m/z= 138.1) of +2. Reactions were carried out with deuterium labeled CADD and WT CADD. These reactions either contained exogenous deuterium labeled Tyr, exogenous unlabeled Tyr, or no Tyr. We observed that deuterium labeled CADD produced pABA with a product peak of m/z= 140.1 (Figure 2.5 C). This is consistent with the incorporation of two deuterium atoms into the structure of pABA. However, deuterium labeled CADD also produced

small amounts of unlabeled pABA (Figure 2.5 B). This is likely due to incomplete incorporation of the deuterium label during expression. This could be due to not using a high enough concentration of deuterated tyrosine during expression. On the other hand, WT CADD only produced unlabeled pABA. This was observed even when WT CADD was incubated with exogenous deuterium labeled Tyr (Figure 2.5 B). Also, when deuterium labeled CADD was incubated with exogenous unlabeled Tyr the ratio of labeled to unlabeled pABA did not change (Figure 2.5 B). These findings suggest CADD does not require and exogenous Tyr source for activity.

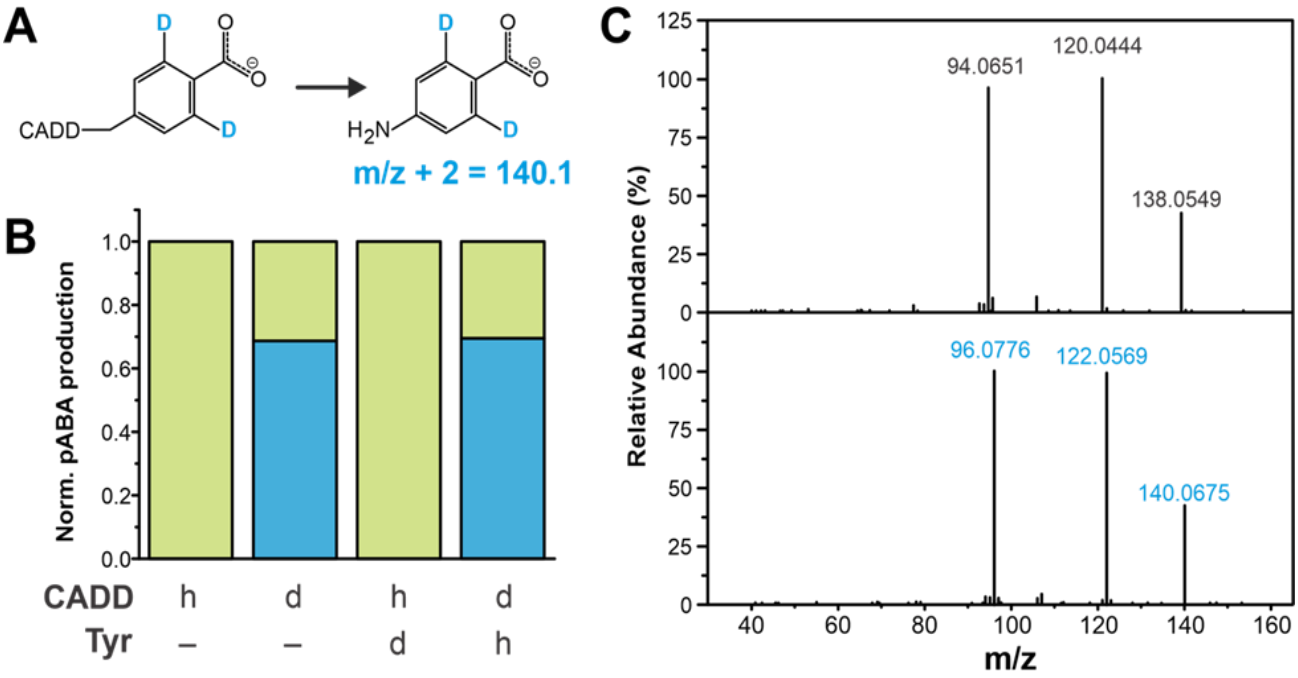

Figure 2.5: A. Incorporation of isotopically labeled 3,5-d_2-Tyr derived from CADD into the product pABA. **B.** Reactions were performed using wild type CADD or deuterated CADD, with the addition of two equivalents of exogenous labeled (d) or unlabeled (h) Tyr. **C.** Mass spectra of the products from reactions carried out with unlabeled protein (top) and labeled protein (bottom) indicate incorporation of the isotope labels in the latter.

2.3.6 "Substrate" tyrosine at position 27 is cleaved by CADD

To determine the position of the substrate, Tyr, 3,5-d_2-L-tyrosine CADD was analyzed via protein MS. Protein MS analysis examined for the loss of a deuterium labeled Tyr side chain (m/z = -108). Reactions were performed with deuterium labeled CADD and contained optimal conditions for reactivity. After protein MS analysis, a mass difference was observed for fragment y_3^+ which corresponds to Tyr27 of CADD (Figure 2.6 A). The cleavage of Y27 left a glycine in its place. Reactions were also carried out under no metal conditions. We wouldn't expect to see any missing Tyr. This is due to the apo reactions not containing the proper metal cofactor required for reactivity. The apo reactions contained a y_3^+ fragment (m/z = 411.25710) consistent with deuterium labeled Tyr at Y27 (Figure 2.6 A). This finding suggested that apo CADD was unreactive. We also observed that Fe/Mn was required for Tyr cleavage, when compared to the apo reactions carried out in the absence of metal ions (Figure 2.6 B). This confirms that a Fe/Mn cofactor is in fact required for reactivity. To confirm that Y27 is the precursor to pABA, reactions were carried out with Y27F CADD mutant (Figure 2.6 C). Indeed, this mutation significantly reduced CADD reactivity yielding less than 0.1 molar equivalents of pABA. Ultimately, these findings confirm that Y27 is cleaved and further processed into pABA.

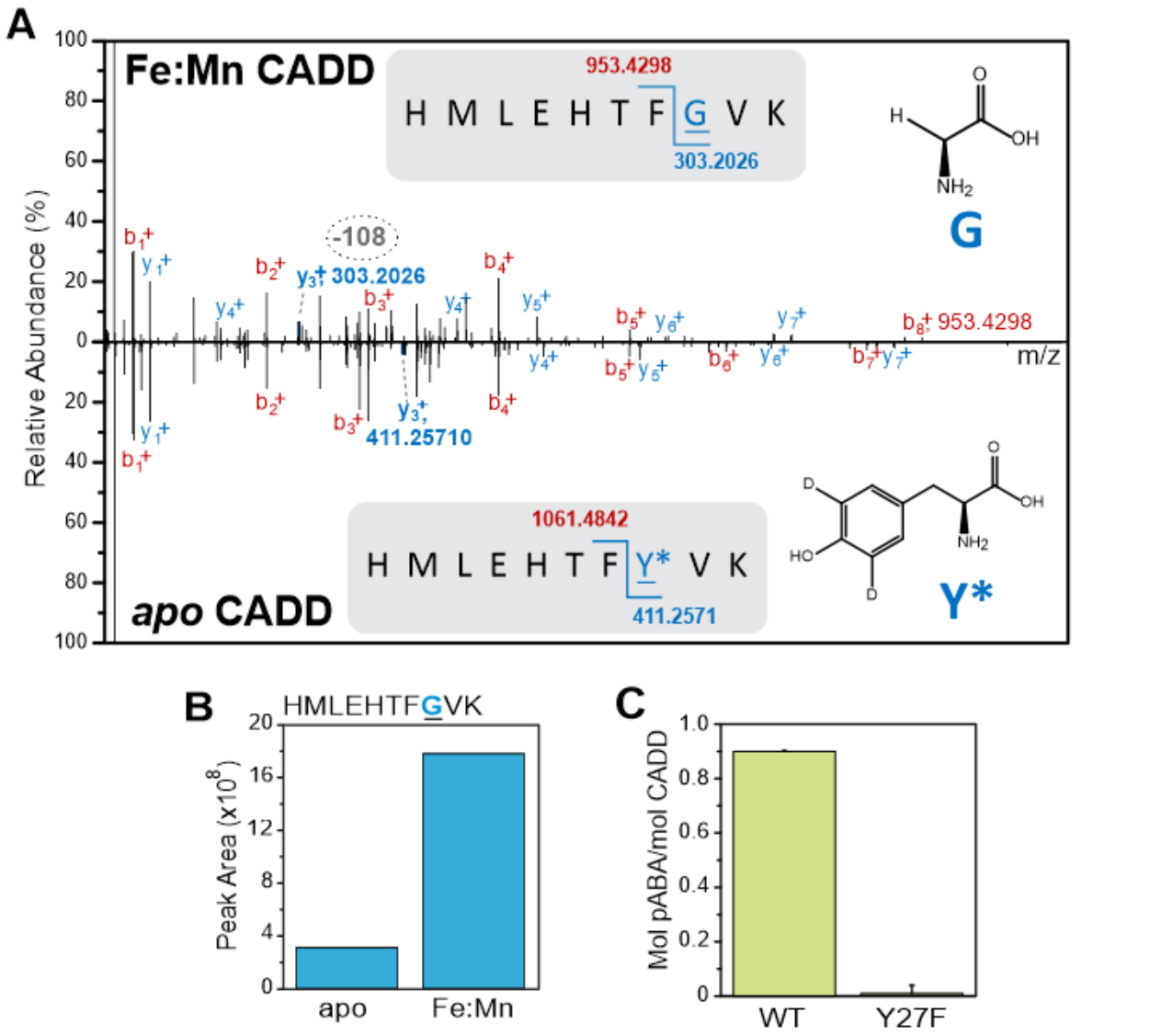

Figure 2.6: A. Mass spectra of the modified peptide from a tryptic digestion of CADD post-reaction with an Fe:Mn cluster (top) or metal-free (bottom). A -108 was observed with Fe:Mn CADD reactions which correlates with the loss Y27. The loss of this tyrosine leaves a glycine in its place. **B.** Abundance of the modified peptide which contains a Glycine at position 27 in apo and Fe:Mn CADD reactions. **C.** Production of pABA by WT and Y27F CADD

2.3.7 CADD could undergo a proton coupled electron transfer (PCET) reaction

The crystal structure of CADD indicates that Tyr27 is 14 Å from the bimetal cluster (Figure 2.7 A). This indicates that a direct reaction with the Fe/Mn cluster is not a possibility. We considered that CADD could under a PCET reaction similar to RNR.[14-15] There exist numerous Tyr and Trp residues that could be used to pass and electron and proton to Tyr27 (Figure 2.8). -Site-directed mutagenesis of the aromatic residue in close proximity to the metal-binding site (Trp92) to Phe indeed decreased pABA yield (Figure 2.7). However, pABA formation was not abolished. This suggest that CADD may use other aromatic residues for reactivity. CADD contains an abundance of oxidizable, aromatic residues. This suggest that the mutation of a single position results in merely re-routing of the electron transfer pathway. However, the decrease in pABA formation in the mutants suggests this chain of oxidizable amino acids play a role in cleavage of the distant "substrate" Tyr.

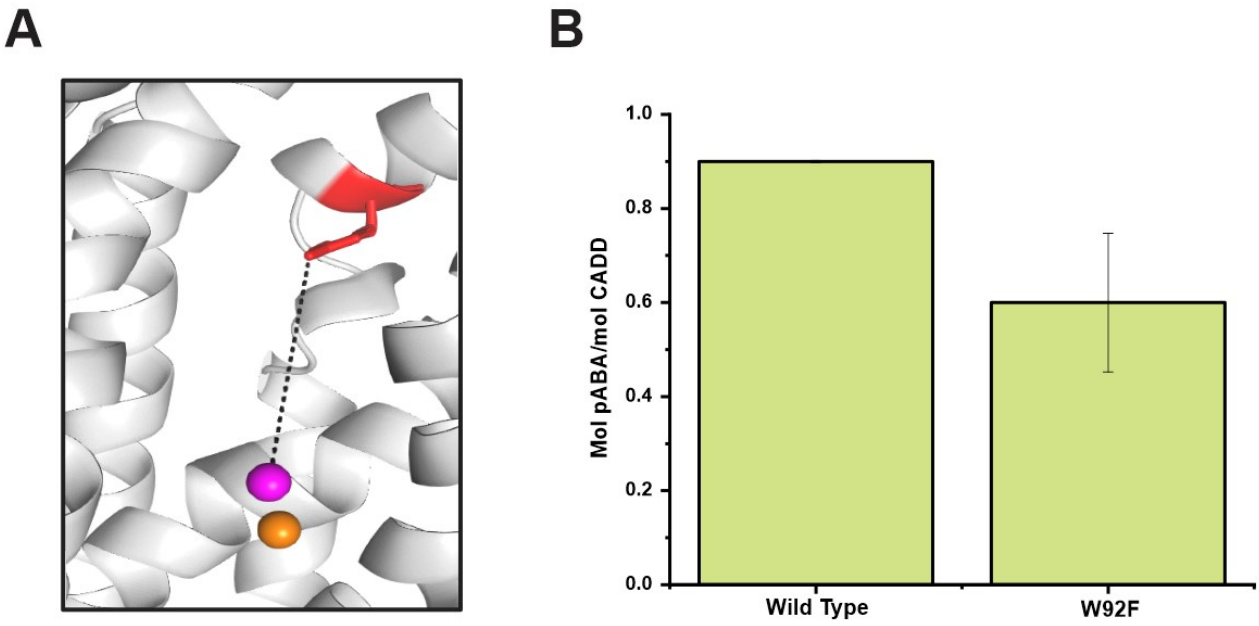

Figure 2.7: A) The metal cluster of CADD is 14 Å away from Y27. Distance is depicted by dashed line. B) pABA production by wild type CADD and the W92F mutant. pABA production is reported as mol. Equiv. of pABA produced per protein. Error bars represent the standard deviations of three different reactions.

2.4 Discussion

2.4.1 CADD utilizes a Mn/Fe to synthesize pABA

The results in this study show that the diiron form of CADD is not the active form of the enzyme. However, the diiron form could be essential to other functions since its structure revealed it contains a diiron core.[10] Metal dependence studies revealed that when CADD is given 1 equivalent of Mn and iron pABA formation is observed. Stoichiometric pABA synthesis is observed when reactions were carried out with in the presence of excess ascorbate. These reactions could be mimicking cellular environments. CADD is found in an intracellular pathogen and cells tend to contain a variety of reducing agents which could enhance CADD reactivity. However, it is surprising that CADD utilizes an Mn/Fe cofactor because it belongs to the HDO family. All the other HDOs us a diiron core to catalyze reactions CADD is the first HDO that has been shown to use a Mn/Fe cofactor.

2.4.2 Oxygen dependence of CADD reactivity

CADD utilizes a Mn/Fe cofactor to activate O_2. Reactions carried out in anerobic environments yielded no pABA. However, reactions caried out with scavengers superoxide dismutase and catalase had no effect of pABA formation. Results also, show that CADD incorporates two oxygens into the structure of pABA. Oxygen tracer studies performed with $^{18}O_2$ revealed that two labeled oxygens were incorporated into the structure of pABA. Even though the major product of the tracer studies was two oxygen incorporations a minor product one oxygen incorporation was also observed. This could be due to small amounts of $^{16}O_2$ contamination in the reactions. This could also be due to labeled oxygens exchanging with water molecules. Overall, our study shows that oxygen is essential to CADD pABA catalysis.

2.4.3 CADD utilizes Y27 as a precursor to pABA

Our findings shows that CADD is in fact a "self-sacrificial" enzyme. Previous studies suggested that Y27 could be the precursor to pABA.[2] We found that Y27 is in fact the precursor. The protein MS analysis revealed that Y27 was missing after the CADD reaction takes place. To confirm Y27 is the precursor to pABA, CADD mutant strain (Y27F) was made. Reactions carried out with the Y27F mutant yielded no pABA formation. Also, Reactions carried out with deuterium labeled CADD yielded labeled pABA. Even when labeled CADD reactions were carried out in the presence of excess unlabeled tyrosine, only labeled pABA was observed. To confirm these findings, deuterium-labeled CADD reactions were analyzed via protein MS. Ultimately, our study proves that CADD undergoes a self-sacrificial reaction which results in the specific and long-range cleavage of Y27.

2.4.4 Proposed CADD mechanism

We hypothesize that the mechanism CADD uses to synthesize pABA is similar to the reaction of ribonucleotide reductases (RNR). RNR are key enzymes that mediates the synthesis of deoxyribonucleotides (DNTPs) which are essential to DNA synthesis in all living cells.[14] RNR utilizes radical chemistry to catalyze the reduction of ribonucleotides into DNTPs.[15] RNRs. A conventional class 1 RNR utilizes a diiron cofactor to activate oxygen to produce an intermediate which generates a tyrosyl radical which is transported among aromatic amino acid side chain residues to reduce ribonucleotides and produce DNTPs.[14,15] We purpose that like RNR CADD undergoes a radical transfer reaction which results in the self-cleavage of Y27. CADD contains multiple aromatic residues that could be utilized to transfer a radical to Y27 (figure 2.8).

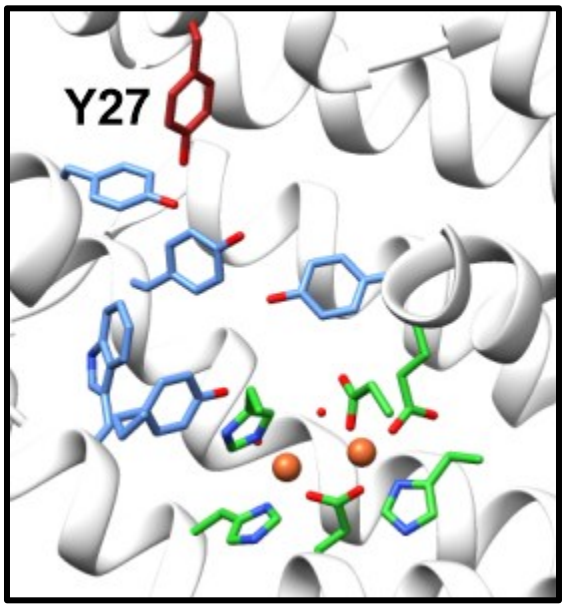

Figure 2.8*:* Active site of CADD from the crystal structure (PDB ID = 1RCW) showing the diiron cluster (ligands shown in green), putative substrate Tyr27 (red), and intervening aromatic residues (blue).

The data obtained in this study has allowed us to develop presumed mechanism (figure 2.9). The substrate tyrosine Y27 is 14 Å away from the Fe/Mn core. We believe CADD undergoes a proton coupled electron transfer reaction which results in the cleavage of Y27. This cleavage leaves a glycine in place of Y27. The substrate Tyr is further processed by CADD with

two oxygen molecules are incorporated into its structure. There also must be the incorporation of an amine to form a complete pABA structure. There is a lysine at position 152 of CADD that could be the source of the amine. However, more experiments must be done to confirm that the lysine at position 152 is the source of the amine. Overall, our mechanism provides insight into how CADD can synthesize pABA.

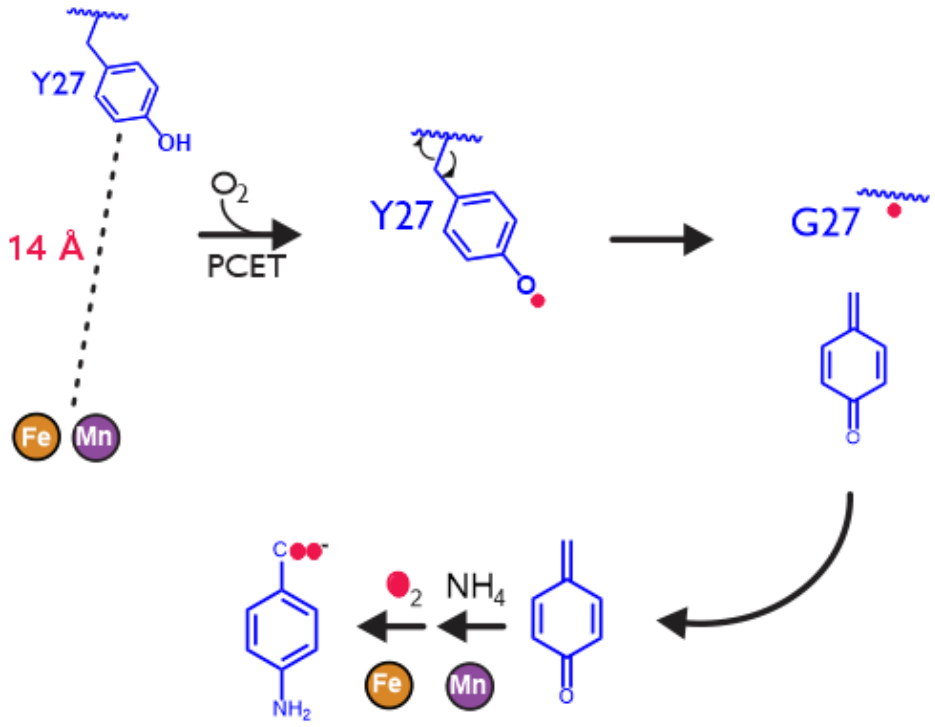

Figure 2.9: Proposed CADD mechanism which PCET which results in the cleavage of Y27. Two oxygen molecules and an amine are also incorporated into the substrate Tyr structure to form pABA.

2.5 Conclusion

Metal dependent reactions revealed CADD utilizes a Fe/Mn cofactor to synthesize pABA. CADD is the first characterized HDO to use a Fe/Mn cofactor to catalyze reactions. After determining the proper metal cofactor, it was shown that CAAD reactivity is oxygen dependent. Protein mass spec analysis of CADD reactions revealed that Y27 on the surface of CADD is the substrate Tyr that is synthesized into pABA. These findings confirm that CADD undergoes a

self-cleavage reaction. This reaction results in the synthesis of pABA. More experiments must be performed to determine if CADD has radical activity and to determine the source of the amine that is incorporated into the structure of pABA.

2.6 References

1. Adams, N. E.; Thiaville, J. J.; Proestos, J.; Juárez-Vázquez, A. L.; McCoy, A. J.; Barona-Gómez, F.; Iwata-Reuyl, D.; de Crécy-Lagard, V.; Maurelli, A. T., Promiscuous and adaptable enzymes fill "holes" in the tetrahydrofolate pathway in Chlamydia species. *MBio* **2014,** *5* (4), e01378-14.
2. Macias-Orihuela, Y.; Cast, T.; Crawford, I.; Brandecker, K. J.; Thiaville, J. J.; Murzin, A. G.; de Crécy-Lagard, V.; White, R. H.; Allen, K. D., An unusual route for p-aminobenzoate biosynthesis in Chlamydia trachomatis involves a probable self-sacrificing diiron oxygenase. *Journal of bacteriology* **2020,** *202* (20), e00319-20.
3. Hoppert, M., Metalloenzymes. In *Encyclopedia of Geobiology*, 2011.
4. Silva, M. T., Classical labeling of bacterial pathogens according to their lifestyle in the host: inconsistencies and alternatives. *Frontiers in microbiology* **2012,** *3*, 71.
5. Brunham, R. C.; Rey-Ladino, J., Immunology of Chlamydia infection: implications for a Chlamydia trachomatis vaccine. *Nature reviews immunology* **2005,** *5* (2), 149-161.
6. Meštrović, T.; Ljubin-Sternak, S., Molecular mechanisms of Chlamydia trachomatis resistance to antimicrobial drugs. *Frontiers in bioscience* **2018,** *23*, 656-670.
7. Sandoz, K. M.; Rockey, D. D., Antibiotic resistance in Chlamydiae. *Future microbiology* **2010,** *5* (9), 1427-1442.
8. Britannica, T. E. o. E. para-aminobenzic acid https://www.britannica.com/science/para-aminobenzoic-acid (accessed Januarty 2022).
9. Satoh, Y.; Kuratsu, M.; Kobayashi, D.; Dairi, T., New gene responsible for para-aminobenzoate biosynthesis. *Journal of bioscience and bioengineering* **2014,** *117* (2), 178-183.

10. Schwarzenbacher, R.; Stenner-Liewen, F.; Liewen, H.; Robinson, H.; Yuan, H.; Bossy-Wetzel, E.; Reed, J. C.; Liddington, R. C., Structure of the Chlamydia protein CADD reveals a redox enzyme that modulates host cell apoptosis. *Journal of Biological Chemistry* **2004,** *279* (28), 29320-29324.
11. Jiang, W.; Yun, D.; Saleh, L.; Barr, E. W.; Xing, G.; Hoffart, L. M.; Maslak, M.-A.; Krebs, C.; Bollinger Jr, J. M., A manganese (IV)/iron (III) cofactor in Chlamydia trachomatis ribonucleotide reductase. *Science* **2007,** *316* (5828), 1188-1191.
12. Riemer, J.; Hoepken, H. H.; Czerwinska, H.; Robinson, S. R.; Dringen, R., Colorimetric ferrozine-based assay for the quantitation of iron in cultured cells. *Analytical biochemistry* **2004,** *331* (2), 370-375.
13. Adams, K. J.; Pratt, B.; Bose, N.; Dubois, L. G.; St. John-Williams, L.; Perrott, K. M.; Ky, K.; Kapahi, P.; Sharma, V.; MacCoss, M. J., Skyline for small molecules: A unifying software package for quantitative metabolomics. *Journal of proteome research* **2020,** *19* (4), 1447-1458.
14. Torrents, E., Ribonucleotide reductases: essential enzymes for bacterial life. *Frontiers in cellular and infection microbiology* **2014,** *4*, 52.
15. Stubbe, J.; Nocera, D. G.; Yee, C. S.; Chang, M. C., Radical initiation in the class I ribonucleotide reductase: long-range proton-coupled electron transfer? *Chemical reviews* **2003,** *103* (6), 2167-2202.

Chapter 3: Assay Development for pABA Detection & Quantification

Abstract

An enzyme found in Nitrosomonas europaea is suggested to fulfill the same role as CADD. The Ne1434 enzyme is homologous to CADD.[1] The two enzymes share structure similarities, and we believe NE1434 undergoes a similar reaction to CADD. To measure if NE1434 can synthesize pABA two assays were developed to observed and quantify pABA formation. A colorimetric assay was used to observe pABA formation after enzyme activity reactions were carried out. Also, a reagent capable of reacting with primary amines to form fluorescent compounds was utilized to quantify pABA formation. In this study we present a new means of detecting pABA and experiments to quantify pABA is ongoing.

3.1 Introduction

Previously it was revealed that there is an enzyme found in *Nitrosomonas europaea (N. europaea) that* could have a similar function as CADD.[1] The *ne1434* gene is found in *N. europaea* and encodes the NE1434 enzyme. This enzyme is believed to be essential to the biosynthesis of pABA. Previous studies with NE134 displayed its ability to rescue mutant *E. coli* cells which lacked the *pabABC* genes.[1] Predicted structures of NE1434 that have been compared to CADD reveal similarities between the proteins (Figure 3.1). The predicted structure of NE1434 has small loop variations and differences in alpha helix extensions between the H2 and H3 helixes of NE1434.[2] However, these differences are not anticipated to impact protein function which suggest NE1434 can undergo a similar reaction as CADD. Previous knockout/knockin studies performed with NE1434 and its structural similarity to CADD suggest NE1434 has the same function as CADD. However, there has not been research studies on NE1434 that measure its ability to synthesize pABA.

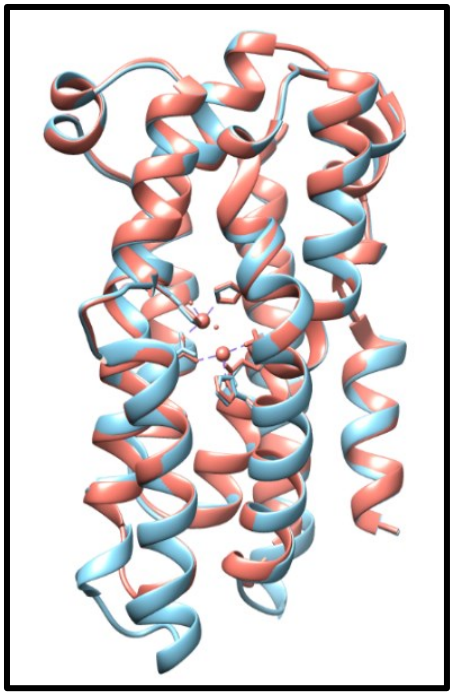

Figure 3.1: Structural alignment of CADD (PDB ID: 1RCW) in red and predicted structure of NE1434 in light blue.

Instead of taking NE1434 through a similar workflow as CADD, we decided to develop an assay that would allow pABA synthesis to be measured by utilizing o-phthaldialdehyde (OPA). OPA is a compound in conjunction with reduced sulfhydryl groups reacts with primary amines to form fluorescent compounds. After reacting with OPA the fluorescence of the compound can be monitored by using a fluorescent plate reader. The fluorescent compound's absorbance can also be measured at 350nm using a spectrophotometer. The structure of pABA contains a primary amine (Figure 3.2) which makes it a likely candidate for OPA reactions.

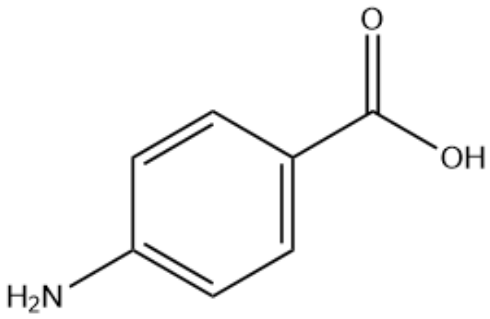

Figure 3.2: Structure of para-amino benzoic acid which contains a primary amine.

We also explored the possibility of using other reagents for colorimetric pABA detection. Previous research utilized a combination of sodium nitrite, ammonium sulfonate, and N-naphthalene dihydrochloride to measure pABA concentrations in urine samples.[3] After, the reagents are added the amount of pABA in each sample was quantified by measuring the absorbance at 540nm.[3] A purple color was also observed when the reagents were added to the samples. We decided to pursue this colorimetric assay because like OPA it is a simple and inexpensive means detect and quantify pABA.

The assays developed in this study will be used to determine if NE1434 is able to synthesize pABA in a similar manner as CADD. CADD was used as a model to determine if the assays are a good way to detect and accurately quantify pABA concentrations. These assays will also be used to perform future studies with CADD. We hope to perform time trial activity assays using the colorimetric and OPA assay to quantify pABA. This will determine the rate in which pABA is synthesized. Understanding the time required for pABA synthesize could lead to the capturing of intermediate products. However, the main goal of this assay is to reveal if NE1434 can synthesize pABA in similar manner as CADD.

3.2 Methods

3.2.1 Expression of CADD

Expression of CADD was done as described in Chapter 2.

3.2.2 Purification of CADD

Purification of CADD was done as described in Chapter 2.

3.2.3 CADD activity assay

CADD reactions were carried out in Ni-NTA elution buffer (25 mM HEPES pH 7.5, 100 mM KCl, 500 mM imidazole, 5% glycerol) and were comprised of 100 μM CADD (the preparations of which is described above)1.25mM ascorbic acid, 80 μM Fe(II) and 120 μM Mn(II). The reactions were incubated at 37°C for ~12 hours under aerobic atmosphere and then quenched with 20uL of HCL. The protein aggregate was pelleted by centrifugation, and the supernatant was stored at -20°C before analysis for pABA quantification.

3.2.4 Expression of NE1434

The NE1434 gene was codon optimized and synthesized by Bio Basic 2.0. The gene was subcloned into pET-21B with NcoI and BamHI restriction enzymes for expression with a C-terminal hexahistidine tag. The sequence follows.

CCATGGCCACCAACACCTTCAAACAGCAGGTTGATTCTATCATCCAGTCTCGTCACC
TGCTGCAGCACCCGTTCTACATCGCTTGGACCGAAGGTAAACTGACCCGTGAACAGC
TGCGTCACTACGCTGAACAGTACTTCTACAACGTTCTGGCTGAACCGACCTACCTGT
CTGCTGTTCACTTCAACACCCCGCACTTCCACAACGTTGAAAACTCTGGTGACATCT
CTATCCGTCAGGAAGTTCTGAAAAACCTGATCGATGAAGAACACGGTGAAAAGAAC
CACCCGGCTCTGTGGAAAGCGTTCGCGTTCGCTCTGGGTGCTGATGACGCTTCTCTG
ACCCAGGCTGATGCTCTGCCGGAAACCGAAAACCTGGTTGCTACCTTCCGTGATATT

TGTATCAACGAACCGTTCTACGCTGGTCTGGCTGCTCTGCACGCTTTCGAATCTCAG
GTTCCGGATATTGCTGCTGTTAAAATCGATGGTCTGGCTAAATTCTACGGTATGAAA
GATCCGGATTCTTACGAATTTTTCTCTGTTCACCAGACTGCAGACATCTTCCACTCTC
AGGCTGAATGGGCGATCATCGAAAAATTCGCTGACACCCCGGAAAAACAGGCTGAA
GTTCTGGCTGCTACCCGTCGTGCATGTGATGCTCTGTGGAAATTCCTGGATGGTATCC
ACGAAAACTACTGTGCTAACCTGATCTGTGAAGAAAAAACCGCTGCTACCCTGCAC
GGATCC

The plasmid was expressed in *Escherichia coli* BL21 (DE3) cultures grown in M9 media supplemented with 50 μg/mL kanamycin and 50 μM Mn. The cells were grown to and OD_{600} of 0.6-0.8. Protein expression was induced with the addition of 200 μM IPTG. The cell cultures were cooled to 24 °C upon induction and incubated overnight. The cells were harvested and stored at -20°C until further use.

3.2.5 Purification of NE1434

Thawed cell paste was resuspended in 25 mM HEPES pH 7.5, 100 mM KCl, 10 mM imidazole, and 5% glycerol (~10 mL buffer per g cell paste) and chilled on ice. Cells were lysed by sonication, and the cleared lysate was loaded onto a Ni-NTA column while maintained at 4 °C (as was the remainder of the purification). The resin was washed with 5 CV of the above buffer with 15 mM imidazole, and protein was eluted with 5 CV of buffer with 500 mM imidazole. The fractions containing NE1434 were identified by an absorbance at 280 nm and by SDS-PAGE. Fractions containing NE1434 were pooled, concentrated, and flash frozen in liquid nitrogen and stored at -70 °C until further use.

3.2.6 Structural alignment of Ne1434 and CADD

All structural alignments were performed using Chimera 1.14 and the prior align function within it.[2] First the software aligns the sequence and then fits the α-carbons of residues in the same columns of the sequence alignment. The NE1434 protein structure was predicted using Expasy Swiss.[4,5]

3.2.7 NE1434 activity assay

NE1434 reactions were carried out in Ni-NTA elution buffer (25 mM HEPES pH 7.5, 100 mM KCl, 500 mM imidazole, 5% glycerol) and were comprised of 100 μM NE1434 (the preparations of which is described above) 1.25mM ascorbic acid and 0-2 molar equivalents of metal (via anaerobic stock solutions of Fe(II) and/or Mn(II)). Experiments were carried out using Fe, Mn, and different ratios of Fe:Mn to determine the optimal ratio for pABA production. The reactions were incubated at 37°C for ~12 hours under aerobic atmosphere and then quenched with 20uL of HCL. The protein aggregate was pelleted by centrifugation, and the supernatant was stored at -20°C before analysis for pABA quantification (described below).

3.2.8 Colorimetric pABA Assay

The activity assays (described above) were diluted with 500 μL of buffer (25 mM HEPES pH 7.5, 100 mM KCl, 500 mM imidazole, 5% glycerol). Stock solutions of sodium nitrite (14mM), ammonium sulfonate (43mM), and N-naphthalene dihydrochloride (3.8mM) were made prior to the colorimetric assay was performed. 100 μL of sodium nitrite stock solution was added to the activity assays and allowed to sit for two minutes. Next, 100 μL of the ammonium sulfonate stock solution was added to the activity assays and allowed to sit for two minutes. Finally, 100 μL of the N-naphthalene dihydrochloride was added to the activity assays.

Following the addition of the N-naphthalene dihydrochloride pABA was detected by a light/dark purple color.

3.2.9 pABA quantification by OPA reagent

The OPA reagent was equilibrated to room temperature before use. After the reagent reached room temperature, 50 μL of BME was added to 20mL of the OPA. Stock solutions of 10 mM pABA and 10mM lysine were made to establish the sensitivity of the OPA. The OPA was aliquoted into 1mL tubes. To establish that the reagent worked varying concentrations of lysine was added to the 1mL tubes and the absorbance was via uV spectrophotometer. To determine if OPA could be used to quantify pABA, varying concentrations of pABA were added to 1mL tubes and the absorbance was measured via uV spectrophotometer.

3.3 Results

3.3.1 Quantification of NE1434

After NE1434 was expressed and purified, it was analyzed by SDS-PAGE electrophoresis to ensure the protein had no impurities and to confirm its molecular weight. Based on how far NE1434 traveled down the SDS-PAGE gel it has a molecular weight of ~26kDa (Figure 3.3). The SDS-PAGE gel also showed that NE1434 had no significant impurities. Thus, we concluded this protein could be used for future activity assays.

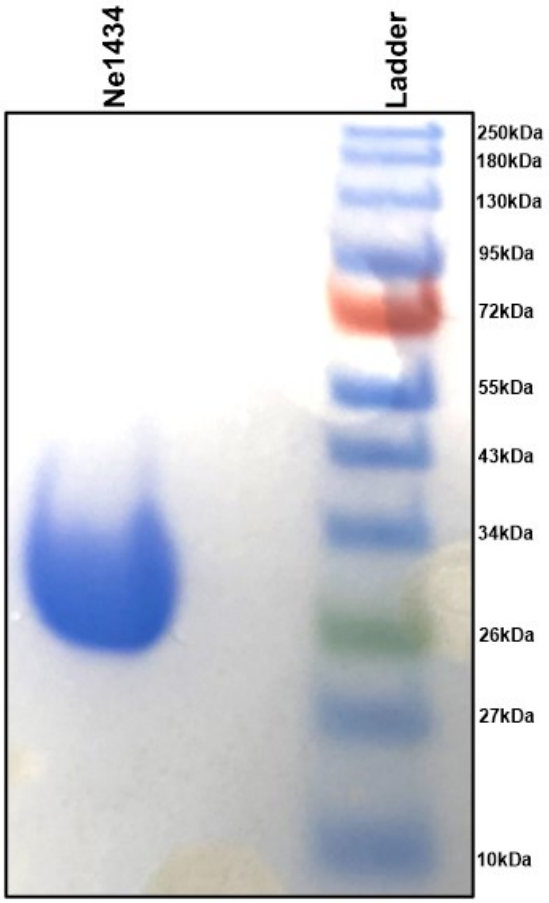

Figure 3.3: Representative SDS-PAGE of purified NE1434. Predicted molecular weight is 26kDa

3.3.2 pABA detection via colorimetric assay

pABA standards and CADD reactions were used as a model to determine if the colorimetric assay could be utilized to detect pABA. CADD reactions were carried out under optimal conditions to ensure pABA formation was observed. After, the CADD reactions were quenched and diluted the reagents were added. A purple color was observed after the addition of the reagents (tube 1 Figure 3.3 A). However, the color was not observed when the reagents were added to water (tube 2 Figure 3.3 A). pABA standards ranging from 0 to 80μM were made to confirm the color was consistent with pABA. After the reagents were added to the standards the absorbance of the standards was measured using a UV/Vis spectrophotometer. The absorbance at 550nm of the standards increased with increasing concentrations of pABA (Figure 3.3 B). A

standard curve of the pABA solutions was made and used to quantify pABA concentrations in CADD reactions. The pABA concentration observed via colorimetric assay was compared to previous LC-MS results (Figure 3.3 B). The colorimetric assay yielded data that was far too large and was not consistent with previously gathered data. We determined that the colorimetric assay may not be the most efficient to accurately quantity pABA concentrations. However, the colorimetric assay can likely be used as a primary verification method for pABA production.

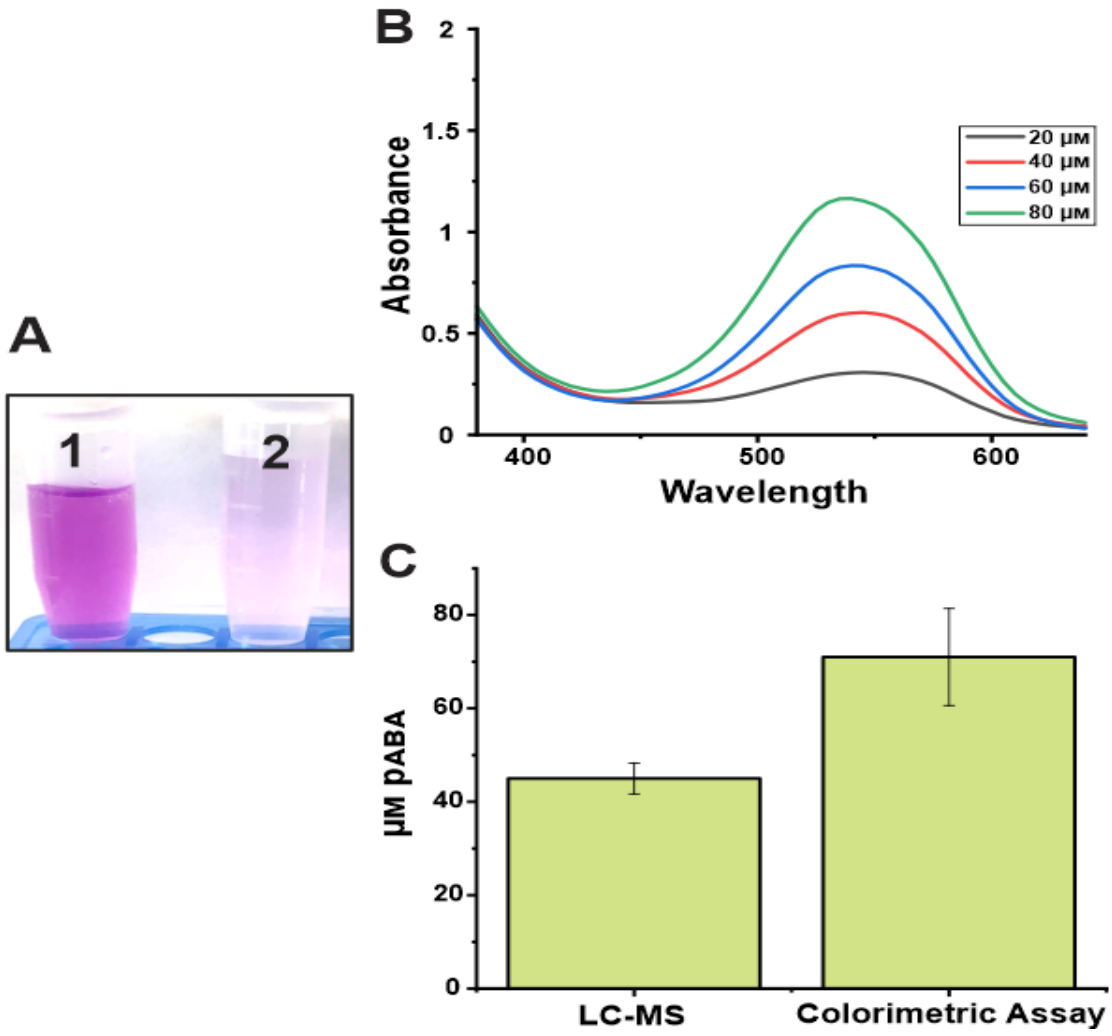

Figure 3.4: A) Tube 1 contained a CADD reaction that was carried out under optimal conditions. A purple color was observed after the addition of the reagents. Tube 2 contained water the color was not observed after the addition of the reagents. B) 20 μM-100 μM pABA standards were made and measured by uV spectroscopy. The absorbance at 550nm increased with increasing concentrations of pABA. C) CADD reactions carried out under optimal conditions were analyzed via the colorimetric assay. The results were compared to optimal CADD reactions that were analyzed via LC-MS. The optimal CADD reactions contained 100 μM CADD, 1.25mM ascorbic acid, 80 μM Fe and 120 μM Mn.

3.3.3 Metal dependence studies with NE1434

Once it was confirmed that the colorimetric assay could be used to detect pABA. We sought to determine if NE1434 can synthesize pABA like CADD. Activity assays were carried

out with Fe, Mn, and Fe/Mn ratios. The reactions also contained ascorbic acid because it allowed us to see stochiometric pABA formation with CADD. After the reagents were added to the quenched and diluted NE1434 reactions no color formation was observed (tubes 2-4 Figure 3.4). When compared to CADD there is a significant difference in color formation (Figure 3.4). From these results unable to confirm if NE1434 is able to synthesize pABA in a similar manner as CADD.

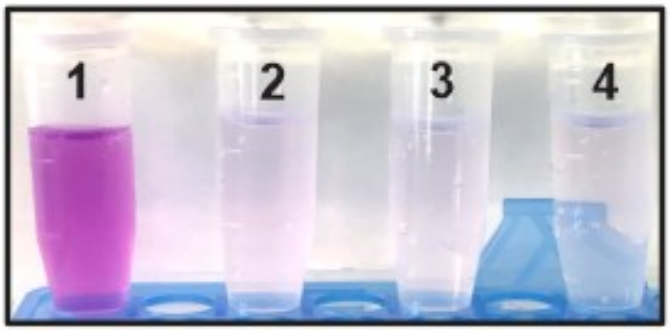

Figure 3.5: Tubes depicting the color formation after the reagents were added to CADD and NE1434 reactions. Tube 1 contained optimal conditions needed for 100 μM CADD, 1.25mM ascorbate, 80 μM Fe and 120 μM Mn. Tube 2 contained 100 μM NE1434, 1.25mM ascorbic acid and 200 μM Fe. pABA, Tube 3 contained 100 μM NE1434, 1.25mM ascorbic acid 100 μM Fe and 100 μM Mn. Tube 4 contained 100 μM Ne1434, 1.25mM ascorbic acid and 200 μM Mn.

3.4 Discussion

3.4.1 pABA Detection by colorimetric assay

We found that the colorimetric assay could be used to detect pABA. This finding agrees with published data which utilized the assay to detect pABA in urine.[3] However, more refinement must be done to the assay to determine if it can be used to quantify pABA formulation. The data gathered from the assay so far suggest NE1434 doesn't synthesize pABA in a similar manner as CADD. Comparison of the CADD structure to the predicted structure of NE1434 revealed that NE1434 does not contain the same aromatic residues we believe to be

essential to a PCET reaction. NE1434 lacks Y170, Y43, and W30 (Figure 3.6). However, NE1434 does contain the substrateTyr located at position 27. It is important to note that NE1434 also contains residues in close proximity to the metal core that could be involved in a PCET reaction (Figure 3.5). It is unclear whether NE1434 undergoes a similar reaction as CADD. Thus far, Ne1434 reactions have been carried out with Fe, Mn, and ratios of Fe/Mn. We know that CADD utilizes a Fe/Mn cofactor to synthesize pABA however this may not be the case with NE1434. Also, it is unclear if NE1434 undergoes a self-sacrificial reaction like CADD. NE1434 could require an exogenous substrate to form pABA it could also utilize a different source to cleave the substrate Tyr.

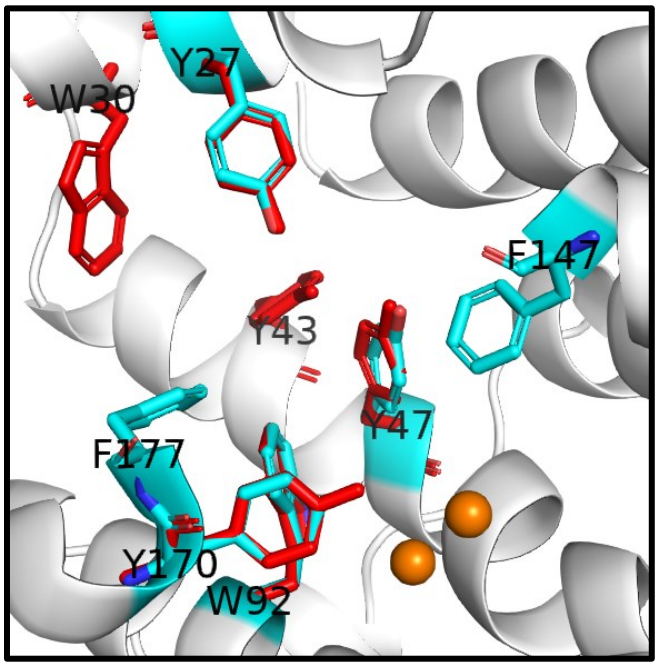

Figure 3.6: Structural alignment of CADD (PDB ID: 1RCW) in Red and predicted structure of NE1434 in blue depicting aromatic residues between metal core and substrate tyrosine.

3.5 Conclusion

We have confirmed that to colorimetric assay is a good means of observing pABA formulation. However, more experiments need to be performed to determine if the colorimetric

assay ca be used to quantify pABA. We observed no color formation after NE1434 reactions were complete. This suggest NE1434 utilizes a different cofactor or substrate for pABA biosynthesis. NE1434 could utilize an exogenous Tyr source or could require the assistance of a different enzyme for activity. Overall, more experiments most ne performed to determine if NE1434 can synthesize pABA. Lastly, we were unable to determine if OPA could be used to detect pABA. The assay needs to be refined for pABA to be detected.

3.6 References

1. Satoh, Y.; Kuratsu, M.; Kobayashi, D.; Dairi, T., New gene responsible for para-aminobenzoate biosynthesis. *Journal of bioscience and bioengineering* **2014,** *117* (2), 178-183.
2. Zambelas, J. M., A Broadly Distributed, Novel Class of Diiron Enzymes Involved in Para-Aminobenzoate Synthesis. **2020**.
3. Bingham, S.; Cummings, J., The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 h urine collections in man. *Clinical science (London, England: 1979)* **1983,** *64* (6), 629-635.
4. Chimera, U., a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem* **2004,** *25* (13), 1605-12.
5. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **2018,** *46* (W1), W296-W303.
6. Bienert, S.; Waterhouse, A.; de Beer, T. A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T., The SWISS-MODEL Repository—new features and functionality. *Nucleic acids research* **2017,** *45* (D1), D313-D319.
7. Vigo, M.; Malec, L.; Gomez, R.; Llosa, R., Spectrophotometric assay using o-phthaldialdehyde for determination of reactive lysine in dairy products. *Food Chemistry* **1992,** *44* (5), 363-365.

Chapter 4: Future Directions

4.1 The CADD project

We have been able to show that CADD utilizes a Fe/Mn cofactor to synthesize pABA. We have also shown CADD utilizes a protein derived Tyr as a precursor to pABA. Previous research set out to show CADD undergoes a self-sacrificial reaction which results in the synthesis of pABA. However, enzyme activity assays yielded poor results. Our discovery of the proper metal cofactor allowed us to observe stoichiometric pABA production and cleavage of Y27 found on CADD. However, there is more work that needs to be done to understand the mechanism CADD undergoes.

There are still parts of the CADD reaction that are left unanswered. We believe CADD undergoes a PCET reaction similar to RNR. Y27 is 14 Å away from the Fe/Mn core. Due to this long distance it is highly unlikely direct reaction of Y27, and the Fe/Mn core takes place. There are multiple aromatic residues between Y27 and the Fe/Mn core that could be utilized to pass a proton and electron. The development of mutant strains of CADD that lack single or multiple aromatic residues between Y27, and the Fe/Mn core would prove insightful. If CADD undergoes a PCET reaction enzyme activity should take a dramatic hit due to the protein lacking key aromatic residues.

Next steps of the CADD project also include crystallography. We want the crystalize the apo and Fe/Mn form of CADD. Crystallization of both forms of CADD could allow us to solve two unique structures. The apo form should provide a structure with Y27 still intact on the protein, due to it not having the proper metal cofactor required for reactivity. On the other hand, the Fe/Mn form of CADD should provide a structure in which Y27 is missing. This would be due to CADD undergoing the self-sacrificial reaction. So far, samples of CADD have been sent

to the national high- throughput crystallization center at Hauptman-Woodward Medical research institute (HWM). This research facility sets up a crystal screen with a variety of crystal screen conditions. The status of each condition is provided biweekly. This allows researchers to see what conditions are best used to crystallize their proteins of interest. HWM has provided a variety of conditions that seem to crystallize CADD well. We hope to have a CADD crystal and structure soon.

Other future directions of the CADD project include limited oxygen activity assays, time trial activity assays and studies with mutant strains of *E. coli*. By limiting oxygen, the amount of oxygen required for pABA synthesis could be determined. This can be done by performing activity assays under anaerobic conditions and providing the reactions with different concentrations of oxygen saturated buffer. By performing time trial activity assays, we could stop the reaction before pABA is formed. After stopping the reaction, LC-MS analysis could be utilized to look for intermediate products of the reaction. Lastly, we will perform studies with mutant strains of *E. coli* that lack the *pabABC* genes. It already been shown that CADD can rescue mutant *E. coli* cells which lack the *pabABC* genes. However, by performing similar studies with mutant strains of CADD we could learn more about what residues are essential to CADD activity. We could also incorporate different therapeutics into the knockout studies to see what drugs prevent CADD activity. Overall, the CADD project is far from finished and what we've discovered in this study is just the tip of the iceberg.

4.2 The NE1434 project

The NE1434 project is still in its early stages. So far, we have shown that NE1434 may not synthesize pABA in a similar manner to CADD. Future directions include performing activity assays under varying reaction conditions. This will allow us to determine what is needed

for NE1434 to synthesize pABA. Once the cofactor and substrate of NE1434 is determined knockout studies can be performed to determine what amino acid residues are essential to activity. Once these residues are determined we can compare these essential residues to the amino acid residues essential to CADD reactivity. Future directions also include solving the crystal structure of NE1434. The structure of NE1434 has yet to be solved. The predicted structure of NE1434 is similar to the structure of CADD. Solving the structure of NE1434 could allow us to perform accurate sequence alignments.

4.3 Conclusion

We have been able to make good progress with CADD. We have determined the metal cofactor and the substrate CADD utilizes to synthesize pABA. The NE1434 project is still in its early stages however, we hope to discover the cofactor and substrate required for NE1434 to synthesize pABA. Overall, this study has completed its goal of providing insight into the mechanisms of CADD and NE1434.