A MECHANISTIC STUDY OF COPROHEME DECARBOXYLASE

by

SHELBY R. PARROTT

(Under the Direction of William Lanzilotta)

ABSTRACT

Heme is an essential cofactor required for numerous biological reactions in the vast majority of organisms, and its biosynthesis is a complex process. Three heme biosynthetic pathways have been identified, with the most recent being the coproporphyrin dependent (CPD) pathway. In the CPD pathway, coproporphyrinogen III is oxidized to coproporphyrin, followed by iron insertion into the tetrapyrrole, forming coproheme. The final step of the CPD pathway is the decarboxylation of coproheme to heme by the enzyme, coproheme decarboxylase (ChdC). The CPD pathway is exclusively found in firmicutes and actinobacteria, both of which are pathogenic and highly antibiotic resistant bacteria that contribute to human disease. The investigation into the mechanism of the penultimate enzyme in the CPD pathway, ChdC, will assist in the development of new antimicrobial compounds.

INDEX WORDS: Coproheme Decarboxylase, Heme, Heme Biosynthesis, Iron, Porphyrin

A MECHANISTIC STUDY OF COPROHEME DECARBOXYLASE

by

SHELBY R. PARROTT

B.S., Mars Hill University, 2019

A Thesis Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2022

A MECHANISTIC STUDY OF COPROHEME DECARBOXYLASE

by

SHELBY R. PARROTT

Major Professor: William N. Lanzilotta

Committee: Amy E. Medlock
Todd C. Harrop

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2022

DEDICATION

This work is dedicated to those I love most in this world:

Mom, thank you for sharing with me your empathy and creativity.

Dad, thank you for imparting me with your sense of humor and aptitude for science.

Jake, thank you for teaching me to be who I am above all else.

Sam, thank you for the years of love and support, here’s to many, many more.

ACKNOWLEDGEMENTS

I would like to thank my advisor, William Lanzilotta, for his patience and open-mindedness. After a difficult start to my time as a chemistry graduate student, I found my home in the Lanzilotta Lab in the biochemistry department. Bill supported me fully through my unconventional cross-departmental graduate program and through my shift from a doctorate to a master's degree. I am forever grateful for his seasoned guidance and continuous support.

I wish to thank my lab mates, Clayton Prichett, Marley Brimberry, and Liju Mathew for showing me the ropes and for being constant and reliable resources throughout my graduate journey. I would also like to thank Mika Baltes for all of his hard work assisting me with sample prep and data collection and analysis.

I would also like to express my gratitude towards Sophia Weerth and Harry Dailey both for their guidance and collaboration on the research presented in this thesis. They were invaluable resources and graciously allowed me to utilize some of their instrumentation for my research.

Lastly, I would like to thank my other committee members, Todd Harrop, and Amy Medlock, for their meaningful feedback and support throughout my time as a graduate student.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

1.1 The Pigments of Life

Tetrapyrroles are chemical compounds used in essential biological processes throughout nature. This important family of molecules is classified by the presence of four pyrrole rings arranged in either a linear or cyclic fashion. While the majority of tetrapyrroles found in nature are metallated macrocycles, there are also biologically relevant linear tetrapyrroles, known as bilins. Bilins play important roles in the photopigment and photoreceptor proteins, as well as a role in the biosynthesis of tetrapyrroles (Dailey et al., 2017). For example, hydroxymethylbilane (HMB) is a linear tetrapyrrole in the biosynthesis of uroporphyrinogen III, a cyclic tetrapyrrole that serves as the last shared ancestor in the core biosynthetic pathway, from which, all tetrapyrrole macrocycles are derived (Scheme 1.1) (Bryant et al., 2020). Other bilins are synthesized following the oxidative cleavage of heme to biliverdin. In plants, algae, and cyanobacteria, biliverdin acts as a base tetrapyrrole for the production of additional bilins that play important roles in light-harvesting enzyme mechanisms (Tu and Lagarias, 2005).

In contrast to linear tetrapyrroles, the family of cyclic tetrapyrroles is far more diverse and consists of three parent structures: chlorins, corrins, and porphyrins. Each parent structure differs slightly in structure and metal specificity (Figure 1.1). The porphyrin structure consists of four pyrrole rings connected via one-carbon bridges in a

cyclic fashion with alternating single and double bonds. Porphyrins often coordinate with iron, producing a strong red color upon chelation. Chlorins have been shown to bind both magnesium and iron, resulting in green pigmentation after metal binding. Chlorin molecules have a similar structure to porphyrins but lack a double bond on one of the four pyrrole rings. Corrins coordinate cobalt, giving them a pink color. Their structure differs from porphyrins, with a direct bond between two of the pyrrole rings instead of a one-carbon bridge. This family of macrocycles has become known as ‘the pigments of life’ due to their widespread distribution in nature and the range of colors produced upon their coordination with select transition metals.

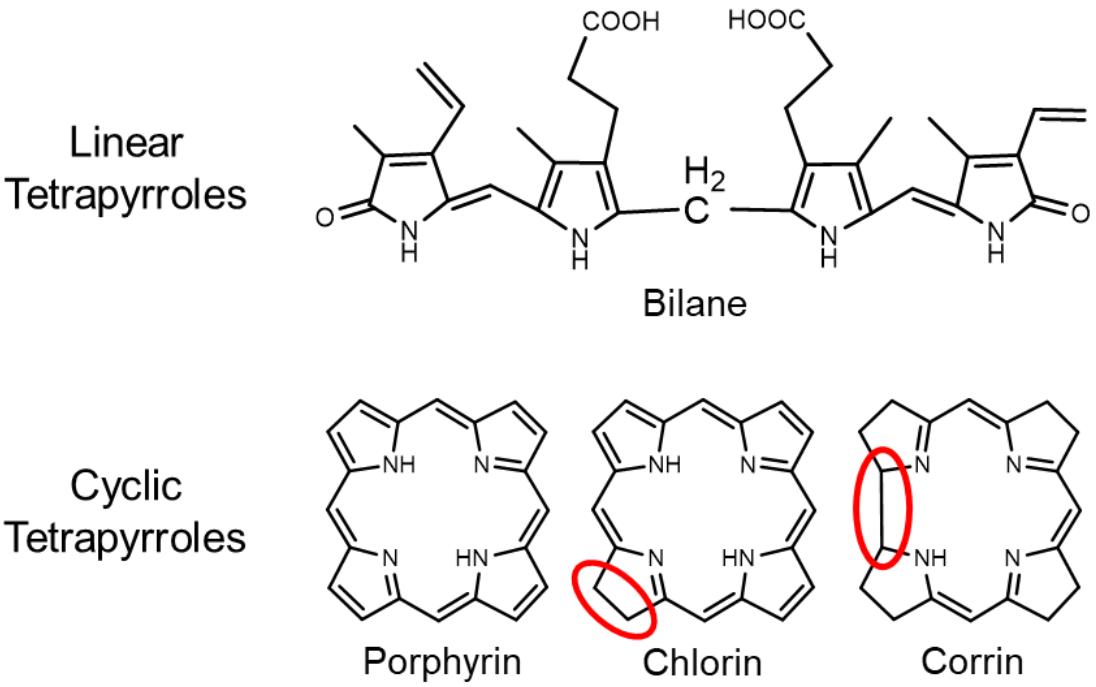


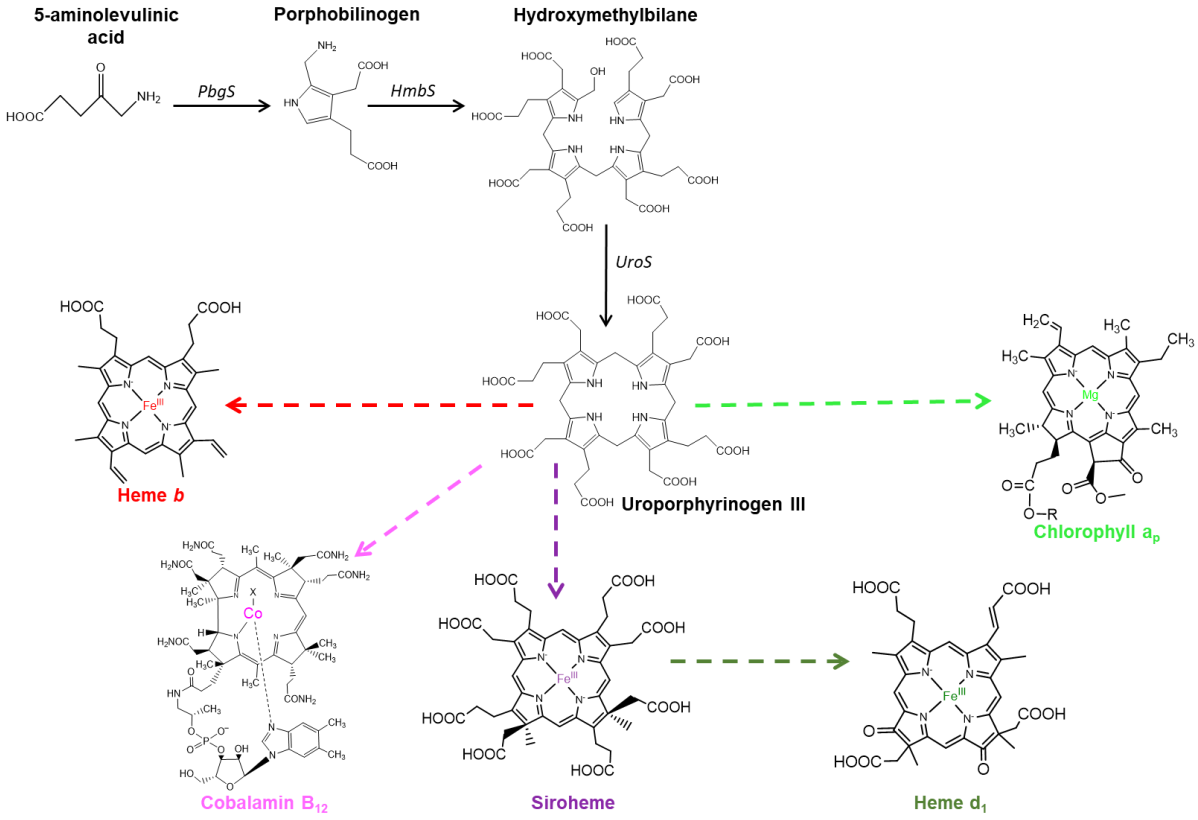
Figure 1.1

Parent structures of linear and cyclic tetrapyrroles. Structural differences among the cyclic tetrapyrroles are circled in red.

Numerous metallated macrocycles, that are essential cofactors for life, are contained within this vast superfamily. This includes, heme, chlorophyll, bacteriochlorophyll, cobalamin, siroheme, and heme d_1 (Scheme 1.1). Iron containing porphyrins, also known as heme, have been found to act as a cofactor in a variety of crucial reactions due to the several oxidation states iron can exist in. Hemes are known to play important roles in electron transport, gas sensing and transport, signaling, transcription, and catalysis (Zhang, 2011). Similar to hemes, cobalamins or vitamin B_{12} are also functionally diverse. Cobalamins are cobalt bound corrin molecules that play essential roles in light sensing, cell metabolism, DNA replication and repair, and detoxification via dehalogenation (Bridwell-Rabb and Drennan, 2017; Banerjee and Ragsdale, 2003). The magnesium bound chlorin molecules, chlorophylls and bacteriochlorophylls, are light-absorbing pigments found in plants and select bacteria that play a role in the photosynthetic transition of solar energy to chemical energy (Tymoczko et al., 2015). Siroheme and heme d1 are both iron-containing chlorin molecules that catalyze the reduction of sulfite or nitrite in the respective reducing bacteria (Bali et al., 2011).

The biosynthesis of tetrapyrroles begins with a common set of reactions that produce uroporphyrinogen III (Scheme 1.1). This reaction is initiated with the synthesis of 5-aminolevulinic acid (ALA), a process that occurs by one of two routes: the C4 or C5 pathway. Found in most metazoan and α-proteobacteria, the C4 pathway converts glycine and succinyl-CoA to ALA via 5-aminolevulinic acid synthase (AlaS) (Kikuchi et al., 1958). The C5 pathway, found in plants and most bacteria, produces ALA through a multistep enzymatic process (Dailey et al., 2017). In the C5 pathway, glutamyl-tRNA is

converted to glutamate semialdehyde (GSA) by glutamyl-tRNA reductase (GtrR). GSA then binds to GSA aminomutase (GsaM), forming ALA. ALA is transformed into uroporphyrinogen III via a conserved protein network including porphobilinogen synthase (PbgS), hydroxymethylbilane synthase (HmbS), and uroporphyrinogen III synthase (UroS) (Bryant et al., 2020). Uroporphyrinogen III is the last common intermediate among all tetrapyrroles, and following the synthesis of this molecule, a complex biosynthetic network of diverse tetrapyrroles begins to emerge (Scheme 1.1).



Scheme 1.1

Synthesis of uroporphyrinogen III and the tetrapyrroles derived from this intermediate. Solid arrows indicate individual enzymatic steps. Dashed arrows represent a multistep enzymatic reaction. The color of the dashed arrows and metal cores correlate to the unique chromophore each molecule produces.

1.2 Significance of Heme

Heme is an iron-containing porphyrin which is a cofactor in a variety of proteins that perform diverse and essential functions. Within the macrocycle, an iron ion is tetradentally coordinated in the equatorial plane of protoporphyrin IX, leaving the proximal and distal axial positions open for ligation. The ligand may come from protein interaction via amino acids (Figure 1.2) or chemical substrates. The conjugation of the macrocycle and type of iron ligation contribute to the distinguishing red color of hemes (Bryant et al., 2020). Heme binds to the protein active site through favorable interactions with the hydrophobic active site residues, with the heme substituent groups occasionally forming additional links to the protein through hydrogen bonds, salt bridges, or covalent interactions (Zhang, 2011). The iron in the heme macrocycle is also redox active and may perform key catalytic functions. One important function of heme, related to the electronic structure of iron, is its ability to activate molecular oxygen (O_2). Heme activates oxygen through partial reduction (forming one molecule of water) and formation of an iron(IV)-oxo, or ferryl, intermediate from the ferric oxidation state (Moody and Raven, 2018). Two of these key catalytic intermediates have been widely documented, Compound I and Compound II. Compound I is a ferryl porphyrin π-cation radical and Compound II is a one-electron reduced form of Compound I (Figure 1.2) (Moody and Raven, 2018).

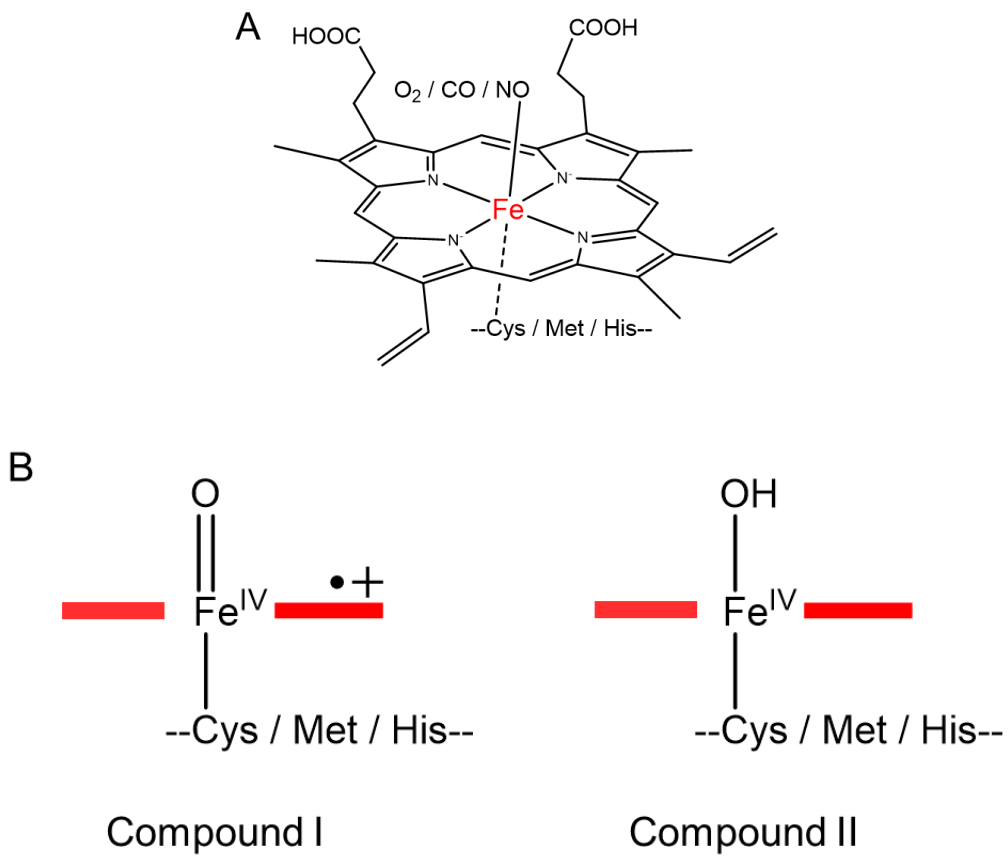


Figure 1.2

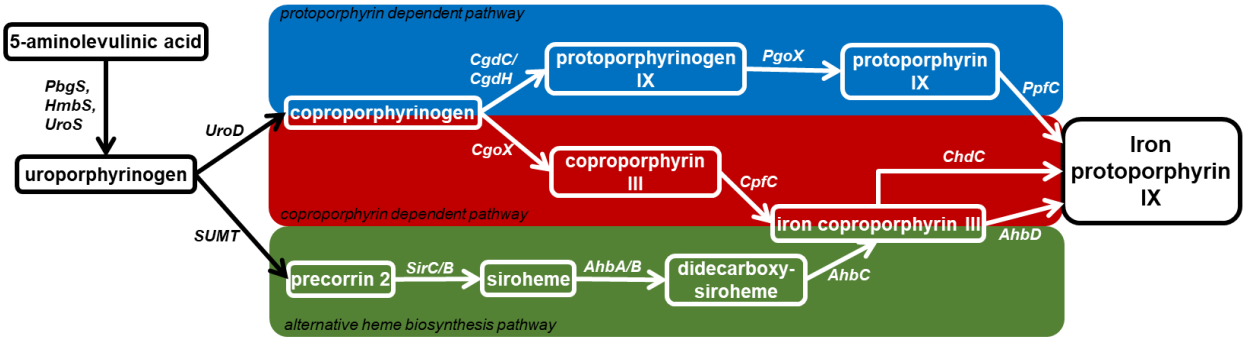
(A) Overall structure of heme and potential proximal and distal ligands. The heme macrocycle is formed by four pyrrole rings connected via methylene bridges and the following substituents: four methyl groups, two vinyl groups, and two propionate groups. The iron ion is equatorially coordinated by the nitrogens of the four pyrrole rings. The two axial positions of the iron ion are coordinated by small molecules such as oxygen, carbon monoxide, or nitric oxide, or by amino acids, cysteine, methionine, or histidine, when bound to a protein. (B) Schematic illustrations of Compound I and Compound II.

Heme is necessary for the survival of most living organisms. Some organisms acquire heme from their environment, but the vast majority have the necessary machinery to synthesize heme *de novo*. Whether it is synthesized or acquired, heme acts as a cofactor in numerous reactions that are vital to life, particularly for aerobic organisms. The most familiar example of heme-binding proteins are hemoglobin and

myoglobin, the proteins responsible for the transport and storage of oxygen, respectively. Heme also plays a role in respiratory cytochromes as an electron carrier and key catalytic cofactor. Moreover, heme has roles as a molecular sensor and transport in bacteria, redox catalysis, transcription, and signaling molecule (Zhang, 2011).

1.3 Heme Biosynthesis – Three Pathways to Heme

Because heme is vital in sustaining the vast majority of cellular life, and iron is an essential growth-limiting nutrient, iron homeostasis and heme biosynthesis is a strictly regulated process. Currently, three known pathways to heme biosynthesis exist: the 'Alternate heme biosynthesis' (Ahb) pathway, the Protoporphyrin Dependent (PPD) pathway, and the Coproporphyrin Dependent (CPD) pathway (Dailey et al., 2017). All three pathways begin with the synthesis of uroporphyrinogen III from 5-aminolevulinic acid (ALA) via PbgS, HmbS, and UroS (Scheme 1.1). Upon the synthesis of uroporphyrinogen III, the three pathways begin to branch off (Scheme 1.2). The Ahb pathway transforms uroporphyrinogen III to siroheme, which is converted to heme by AhbA, AhbB, AhbC, and AhbD. The CPD and PPD pathways both form coproporphyrinogen III through the decarboxylation of uroporphyrinogen III by uroporphyrinogen decarboxylase (UroD). After the formation of coproporphyrinogen III, the CPD and PPD pathways diverge.

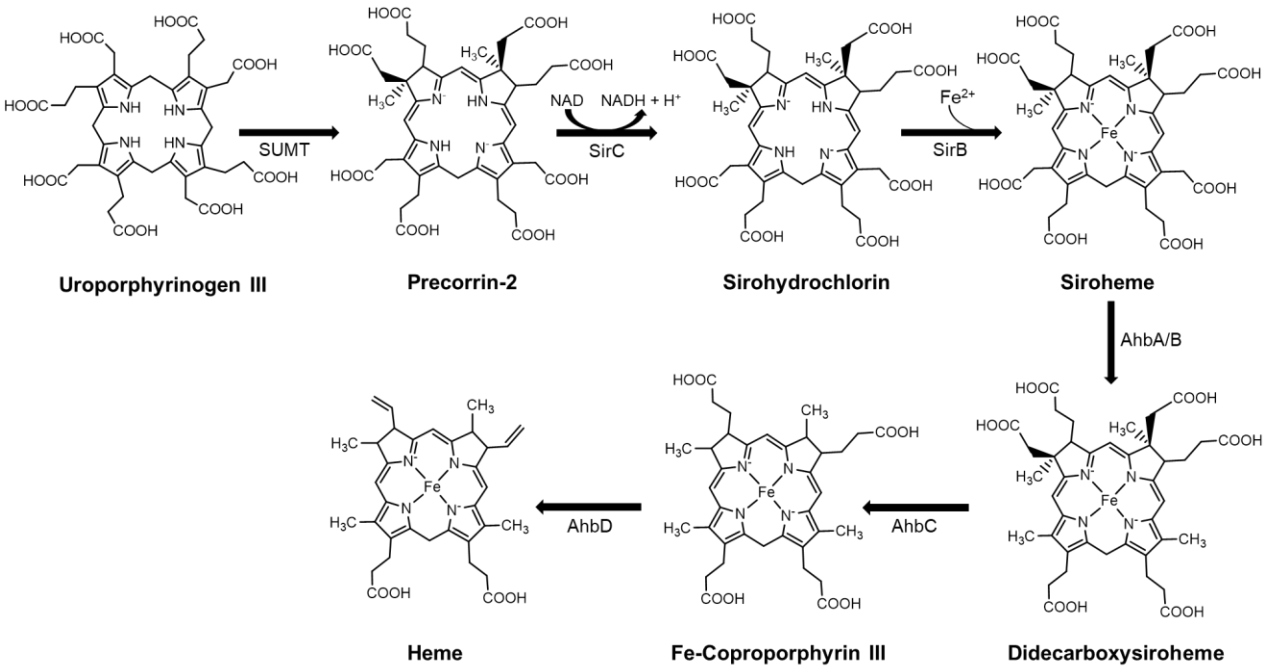


Scheme 1.2

Known pathways to heme (iron protoporphyrin IX). Schematic representation of the PPD, CPD, and AHB pathways are shown within the blue, red, and green boxes, respectively. Arrows indicate individual enzymatic steps. Intermediates and enzymes common to other biosynthetic pathways are shown in black.

1.4 The Alternative Heme Biosynthesis Pathway

The Ahb pathway is an oxygen independent route to heme used by archaea and sulfate-reducing/denitrifying bacteria (Buchenau et al., 2006). Following the synthesis of uroporphyrinogen III, *S*-adenosyl-L-methionine uroporphyrinogen III methyltransferase (SUMT) converts uroporphyrinogen III to precorrin-2 (Scheme 1.3) (Blanche et al., 1989). Precorrin-2 dehydrogenase (SirC) catalyzes the oxidation of precorrin-2, producing sirohydrochlorin (Raux et al., 2003). Iron is then inserted into sirohydrochlorin by sirohydrochlorin ferrochelatase (SirB), forming siroheme (Raux et al., 2003). The enzymes AhbA and AhbB decarboxylate two acetic acid groups of siroheme to produce 12,18-didecarboxysiroheme (Bali et al., 2011). Two additional acetic acid groups of 12,18-didecarboxysiroheme are decarboxylated by the radical *S*-adenosyl-L-methionine (SAM) enzyme, AhbC, forming iron coproporphyrin III (Bali et al., 2011). Lastly, the radical SAM enzyme, AhbD, performs the oxidative decarboxylation of the propionate groups, transforming iron coproporphyrin III to heme (Bali et al., 2011).



Scheme 1.3

Multi-enzyme reaction scheme of the siroheme dependent Abh pathway.

1.5 The Protoporphyrin Dependent Pathway

Known as the classic heme biosynthesis pathway, the PPD pathway was the first heme biosynthetic route to be extensively characterized. Found in eukaryotes and gram-negative bacteria, the PPD pathway can be further broken down into two branches following the formation of coproporphyrinogen III via UroD: the O_2 -dependent PPD pathway and the O_2-independent PPD pathway.

Following its formation, the two propionate groups of coproporphyrinogen III are subsequently oxidatively decarboxylated. The decarboxylation of the first and second propionate group result in the formation of a monovinyl monopropionate deuteroporphyrin (MMD) intermediate and protoporphyrinogen IX, respectively. In the

O_2 -dependent branch of the PPD pathway, the decarboxylation of coproporphyrinogen III to protoporphyrinogen IX is catalyzed by the dimeric membrane bound enzyme, coproporphyrinogen decarboxylase (CgdC). In the CgdC mechanism, an active site conserved aspartic acid residue is proposed to act as a base, initializing the decarboxylation reaction by facilitating the deprotonation of one the pyrrole nitrogen groups (Stephenson et al., 2007). Two molecules of O_2 are required to convert two of the propionate groups of coproporphyrinogen III to vinyl groups, using one molecule of O_2 for each decarboxylation (Lash, 2005). It is proposed that two conserved arginine residues form hydrogen bonds with the propionate groups of the metallated macrocycle which orient the substrate in the active site (Stephenson et al., 2007). Each decarboxylation yields one molecule of CO_2 (Lash, 2005).

The O_2 dependent branch continues with protoporphyrinogen oxidase (PgoX), a homodimeric membrane bound enzyme that catalyzes the conversion of protoporphyrinogen IX to protoporphyrin IX (Poulson and Polglase, 1974). PgoX utilizes three molecules of O_2 and a non-covalently bound molecule of flavin adenine dinucleotide (FAD). The cofactor acts as a two-electron cofactor to conduct the six-electron oxidation of the protoporphyrinogen IX ring (Dailey and Dailey, 1998). Evidence has shown that PgoX has one molecule of FAD bound per reaction. Therefore, it is widely accepted that this six-electron oxidation occurs in three two-electron steps, where the first two steps produce partially oxidized intermediates and the final step yields the fully oxidized product, protoporphyrin IX. In this reaction, O_2 acts as the terminal electron acceptor which generates three molecules of H_2O_2 per turnover (Dailey and Dailey, 1996).

Another homodimeric membrane associated enzyme, protoporphyrin IX ferrochelatase (PpfC) catalyzes the final reaction of the O_2 dependent branch. Interestingly, a [2Fe-2S] cluster has been found to be conserved among all metazoan and a few bacterial ferrochelatases, however, no [2Fe-2S] cluster has been reported in ferrochelatases found in plants (Weerth et al., 2021). The role of the [2Fe-2S] cluster has yet to be elucidated, but it does not appear to participate in iron chelation (Dailey et al., 2000). Structural studies of PpfC helped to illuminate the iron insertion mechanism behind this enzyme. The structure of PpfC contains an “active site lip” that retains an open conformation prior to substrate binding (Figure 1.3) (Medlock et al., *JMB*, 2007). The binding of protoporphyrin IX initiates rearrangement of an active site hydrogen bond network and the “active site lip” shifts to a closed conformation (Dailey et al., 2007). In the closed conformation, the porphyrin macrocycle is distorted, allowing for the deprotonation of one of the pyrrole nitrogens and subsequent insertion of ferrous iron, forming iron protoporphyrin IX (heme *b*) (Medlock et al., *PNAS*, 2007). Following Fe(II) insertion, a hydrogen from another pyrrole ring is abstracted by a conserved active site histidine, triggering the opening of the “active site lip” for product release (Dailey et al., 2007).

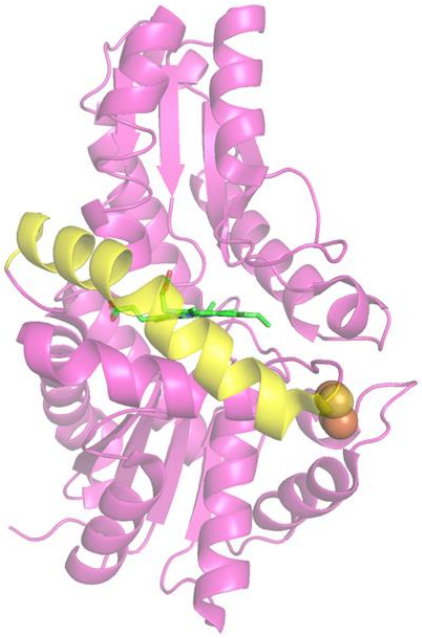


Figure 1.3

Cartoon representation of the PpfC model (PDB: 2HRE). The porphyrin substrate is shown as green sticks and iron sulfur cluster is shown as spheres. The “active site lip” is in the closed conformation and colored yellow.

The O_2-independent PPD pathway follows the same steps leading to the formation of heme *b*, but the enzymes that catalyze the O_2-independent branch are structurally and mechanistically distinct from those of the O_2-dependent branch (Sano and Granick, 1961; Ehteshamuddin, 1968). Decarboxylation of coproporphyrinogen III is catalyzed by coproporphyrinogen dehydrogenase (CgdH, formally "HemN"), a radical *S*-adenosylmethionine (SAM) enzyme containing a [4Fe-4S] cluster. The reaction is initiated by reduction of the [4Fe-4S] cluster and homolytic cleavage of SAM to methionine and a 5'-deoxyadenosyl radical (Frey and Magnusson, 2003). The radical abstracts a hydrogen atom from the beta carbon of a propionate group of coproporphyrinogen III, setting into motion the decarboxylation of the first propionate group to a vinyl group (Layer et al., 2006). This enzymatic step results in the generation of one molecule of protoporphyrinogen IX and two molecules of CO_2.

Oxidation of protoporphyrinogen IX in the O_2-independent branch of the PPD pathway is catalyzed by either protoporphyrinogen dehydrogenase 1 (PgdH1) or protoporphyrinogen dehydrogenase 2 (PgdH2). PgdH1 is a membrane bound enzyme mainly found in *γ*-proteobacteria that belongs to the long-chain flavodoxin family. PgdH1 is thought to use FMN and electron carriers from the cell respiratory chain to oxidize protoporphyrin IX (Boynton et al., 2009). PgdH2 is also a membrane bound protein found in all heme-synthesizing Gram-negative bacteria that do not possess PgoX or PgdH1 (Boynton et al., 2011). Though PgdH2 is the most common form of protoporphyrinogen oxidase, not much is known about its reaction mechanism. Regardless of which enzyme catalyzes the oxidation of protoporphyrinogen IX the O_2-

independent branch is concluded in the same manner as the O_2-dependent PPD pathway: PpfC catalyzes the insertion of iron into protoporphyrin IX producing heme *b*.

1.6 The Coproporphyrin Dependent Pathway

Although only recently discovered, the CPD pathway is thought to be the product of an evolutionary transition between the AHB pathway and the PPD pathway (Scheme 1.2) (Dailey et al., 2017). Found mostly in Gram-positive bacteria with no known mammalian homolog, this pathway reveals an alternative strategy for heme biosynthesis that requires formation of a coproporphyrin intermediate. The CPD pathway begins with the conversion of uroporphyrinogen III to coproporphyrinogen III by UroD (Scheme 1.2). However, unlike the PPD pathway, the first committed step of the CPD branch is oxidation of coproporphyrinogen III to coproporphyrin III. This oxidation is catalyzed by the enzyme coproporphyrinogen oxidase (CgoX). One FAD molecule bound to CgoX facilitates the six-electron oxidation of coproporphyrinogen III to coproporphyrin III, generating three molecules of H_2O_2 and one molecule of coproporphyrin III per turnover (Dailey et al., 1994). While similar to PgoX in sequence, structure, and function, CgoX uses coproporphyrinogen III as a substrate instead of protoporphyrinogen IX (Figure 1.4). One reason that the CPD pathway remained unidentified for so long was because CgoX from *Bacillus subtilis* can utilize protoporphyrinogen III as a substrate *in vitro*. Since the PPD pathway was understood first, a scientific bias caused the enzyme to be initially assayed with protoporphyrinogen III, leading to the mischaracterization of CgoX as a protoporphyrinogen oxidase enzyme.

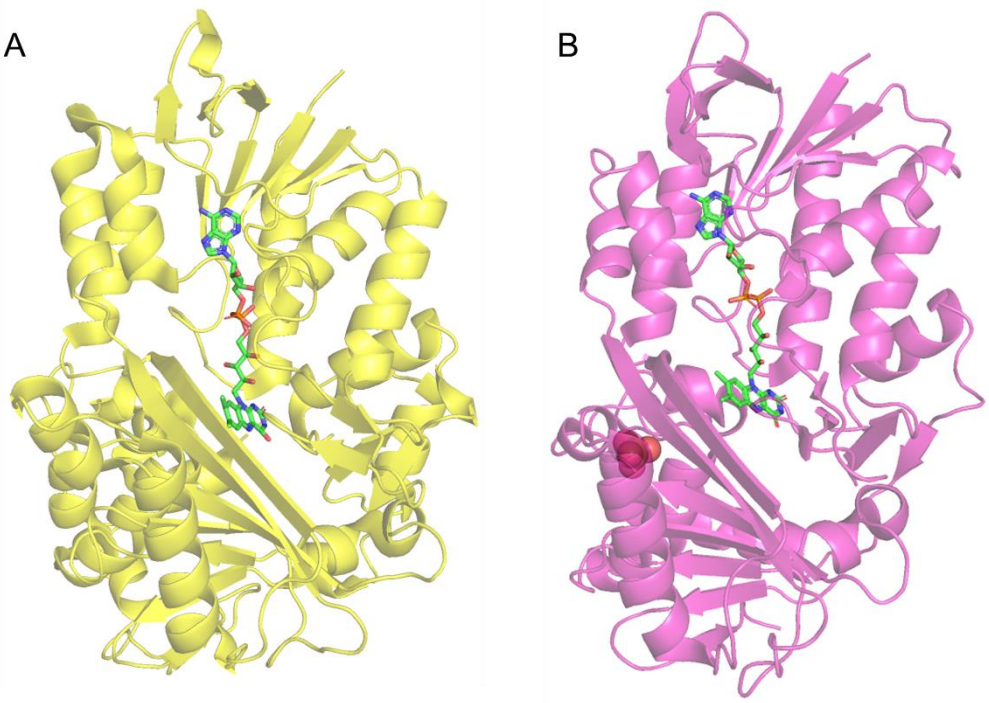


Figure 1.4

Cartoon representation of PgoX (A; PDB: 3NKS) and CgoX (B; PDB:3I6D). FAD molecule shown as green sticks.

The next enzyme in the CPD pathway is coproporphyrin ferrochelatase (CpfC), which catalyzes the insertion of ferrous iron into coproporphyrin III. Despite its similarities to PpfC in function, CpfC has low sequence identity with PpfC, and CpfC lacks certain structural elements retained by PpfC (Figure 1.5) (Celis and DuBois, 2019). For example, it is thought that the absence of the “active site lip” in CpfC allows additional space for the larger substrate, coproporphyrin III (Al-Karadaghi et al., 1997).

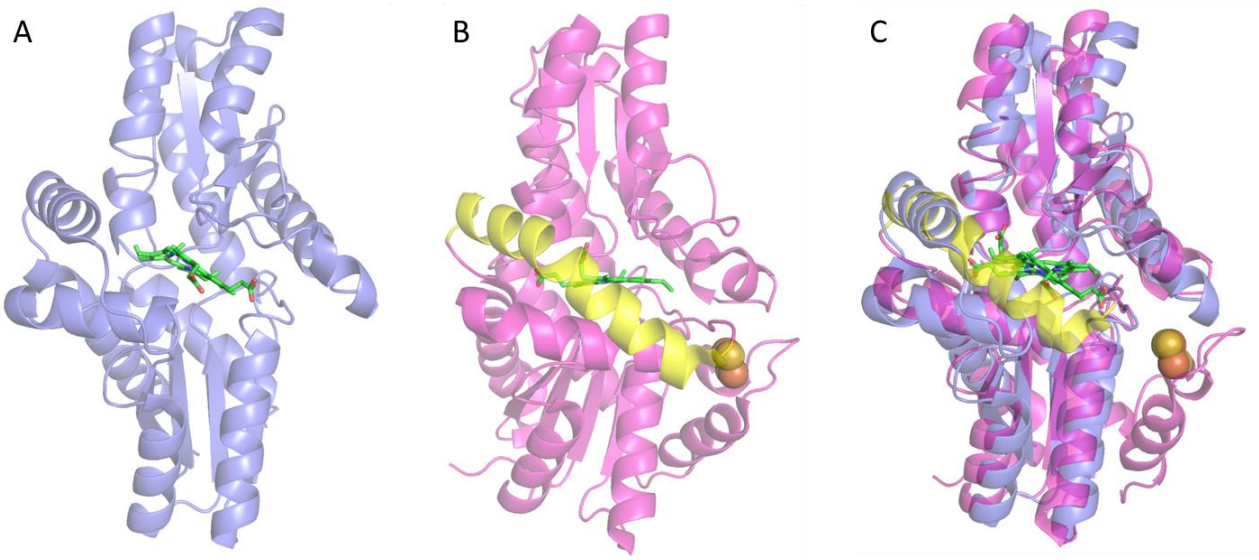


Figure 1.5

Cartoon representation CpfC (A; PDB: 1C1H) and PpfC (B; PDB: 2HRE). (**C**) Structural alignment of PpfC and CpfC. Porphyrin substrate is shown in green and the iron sulfur cluster in PpfC is shown as spheres. The 'active site lip" found in PpfC is shown in the closed conformation in yellow.

In contrast to the PPD pathway, the final step of the CPD pathway is decarboxylation of two propionate groups to form heme *b* via the enzyme coproheme decarboxylase (ChdC) (Dailey and Gerdes, 2015). Structurally, this enzyme belongs to the chlorite dismutase protein family, and retains a homopentameric structure, where each of the 5 domains bind coproheme III at the active site. Upon binding coproheme III, the enzyme performs two sequential and clockwise decarboxylations (Figure 1.6A). One of the axial positions of iron is coordinated by a conserved histidine residue, H168, while the other is left open for oxygen binding (Figure 1.6B) (Hofbauer et al., 2016). Another histidine residue (H128), part of an active site loop conserved among actinobacteria, acts as a distal base, and assists with the binding of H_2O_2 to iron (Figure 1.6B) (Michlits et al., 2020). Lastly, a conserved tyrosine residue (Y145) forms a tyrosyl

radical that is essential to the decarboxylation of each propionate group. (Figure 1.6B) (Streit et al., 2018). All amino acid numbering above and throughout this thesis adheres to that of *Streptomyces coelicolor* ChdC (*Sc*ChdC).

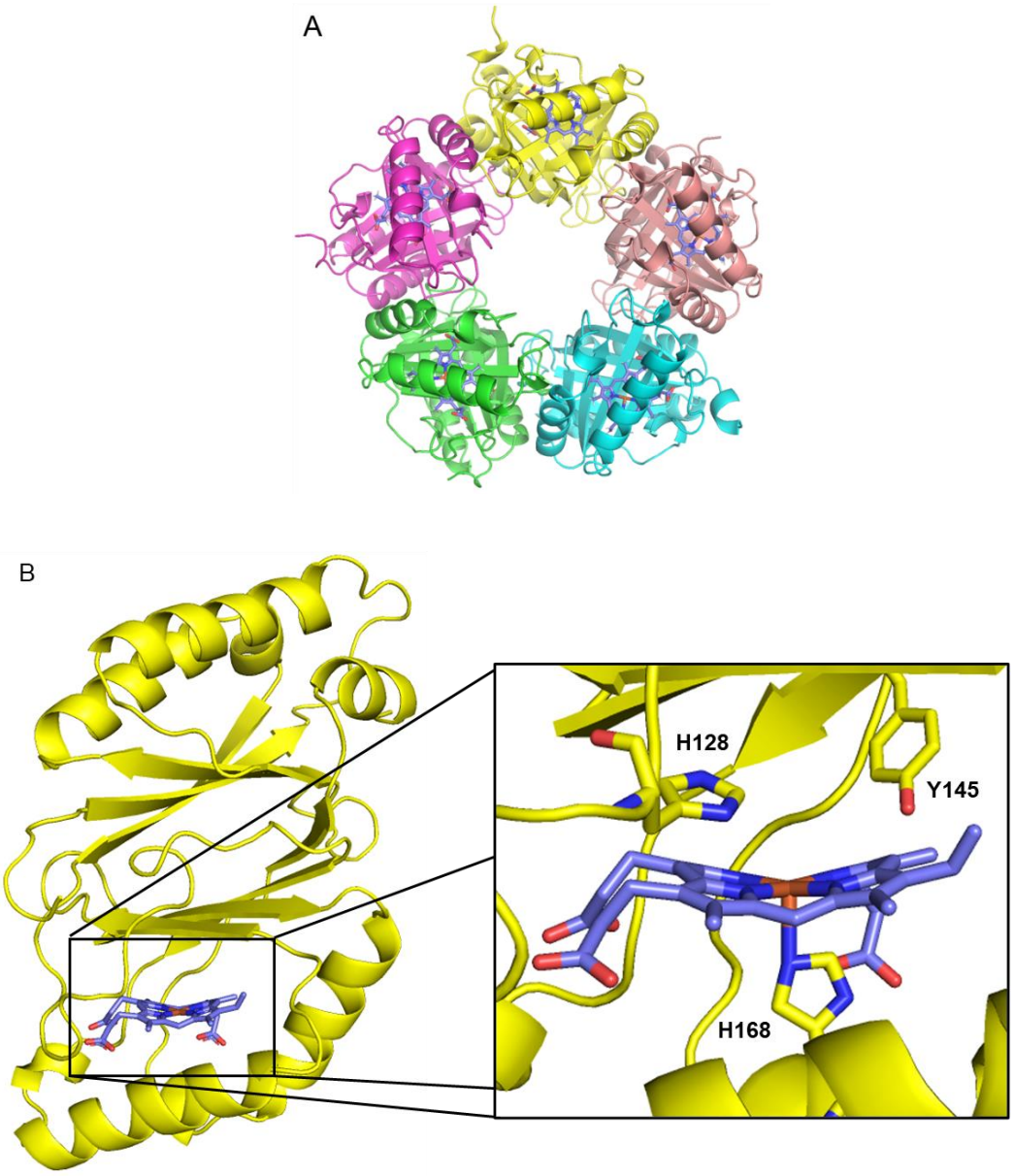


Figure 1.6

(A) The overall homopentameric structure of actinobacterial from *Sc*ChdC, where each domain has a single molecule of MMD bound (dark blue). (B) A single monomer of *Sc*ChdC with MMD bound in the active site. Inset shows conserved catalytic active site residues shown as sticks.

Prior to the discovery of the CPD pathway, it was thought that protoporphyrin IX was a required intermediate for heme biosynthesis in all bacteria, with the exception of sulfate-reducing and denitrifying bacteria (Dailey et al., 2015). However, recent identification of the CPD pathway showed that heme can be synthesized using coproporphyrin III as an intermediate through the combined action of CgoX, CpfC, and ChdC. The CPD pathway is used exclusively by the members of actinobacteria and firmicutes, many of which are pathogenic, such as *Mycobacterium tuberculosis* and *Listeria monocytogenes,* which contribute to human disease and have relatively high rates of antibiotic resistance (Cornaglia, 2009). In this context, ChdC is an important pharmaceutical target, and therefore, illumination of the ChdC mechanism will give us key insights for inhibiting the CPD heme biosynthetic pathway in gram-positive pathogens. This study utilizes a variety of biochemical assays and X-ray crystallography to further elucidate the reaction mechanism of ChdC.

1.7 References

1. Dailey HA, Dailey TA, Gerdes S, Jahn D, Jahn M, O'Brian MR. **2017**. Prokaryotic heme biosynthesis: multiple pathways to a common essential product. *Micro Molec Bio Rev.* 81(1): 1-62.

2. Bryant DA, Hunter CN, Warren MJ. **2020**. Biosynthesis of the modified tetrapyrroles—the pigments of life. *J Biol Chem.* 295(20): 6888-6925.

3. Tu S, Lagarias JC. **2005**. The Pytochromes. In W. R. Briggs & J. L. Spudich (Eds.), *Handbook of Photosensory Receptors.* (pp. 121-149). Wiley-VCH Verlag GmbH & Co. KGaA. https://doi.org/10.1002/352760510X.ch6

4. Zhang L. **2011**. Heme biology: the secret life of heme in regulating diverse biological processes. *World Scientific*. https://doi.org/10.1142/7484

5. Bridwell-Rabb J, Drennan CL. **2017**.Vitamin B_{12} in the spotlight again. *Curr Opin Chem Bio.* 37: 63-70.

6. Banerjee R, Ragsdale SW. **2003**. The many faces of vitamin B_{12}: catalysis by cobalamin-dependent enzymes. *Annu Rev Biochem.* 72: 209-247.

7. Tymoczko JL, Berg JM, Stryer L. **2015**. Biochemistry: a short course. *W.H. Freeman Company*. 3rd Edition. (p.410-411).

8. Bali S, Lawrence AD, Lobo SA, Saraiva LM, Golding BT, Palmer DJ, Howard MJ, Ferguson SJ, Warren MJ. **2011**. Molecular hijacking of siroheme for the synthesis of heme and d_1 heme. *PNAS*. 108(45): 18260-18265.

9. Kikuchi G, Kumar A, Talmage P, Shermin D. **1958**. The enzymatic synthesis of δ-aminolevulinic acid. *J Bio Chem.* 233(5): 1214-1219.

10. Moody PCE, Raven EL. **2018**. The nature and reactivity of ferryl heme in compounds I and II. *Acc Chem Res*. 51: 427-435.

11. Buchenau B, Kahnt J, Heinemann IU, Jahn D, Thauer RK. **2006**. Heme biosynthesis in methanosarcina bakeri via a pathway involving two methylation reactions. *Am Soc Microbio.* 188(24): 8666-8668.

12. Blanche F, Debussche L, Thibaut D, Crouzet J, Cameron B. **1989**. Purification and characterization of S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase from pseudomonas denitrificans. *J Bacterio.* 171(8): 4222-4231.

13. Raux E, Leech HK, Beck R, Schubert HL, Santander PJ, Roessner CA, Scott AI, Martens JH, Jahn D, Thermes C, Rambach A, Warren MJ. **2003**. Identification and functional analysis of enzymes required for precorrin-2 dehydrogenase and metal ion insertion in the biosynthesis of sirohaem and cobalamin in bacillus megaterium. *Bichem J*. 370(2): 505-516.

14. Stephenson JR, Stacey JA, Morgenthaler JB, Friesen JA, Lash TD, Jones MA. **2007**. Role of aspartate 400, arginine 262, and arginine 401 in the catalytic mechanism of human coproporphyrinogen oxidase. *Protein Sci*. 16:401–410.

15. Lash TD. **2005**. The enigma of coproporphyrinogen oxidase: how does this unusual enzyme carry out oxidative decarboxylations to afford vinyl groups? *Bioorg Med Chem Lett*. 15:4506–4509.

16. Poulson R, Polglase WJ. **1975**. The enzymatic conversion of protoporphyrinogen IX to protoporphyrin IX. *J Bio Chem.* 250(4): 1269-1274.

17. Dailey TA, Dailey HA. **1998**. Identification of an FAD superfamily containing protoporphyrinogen oxidases, monoamine oxidases, and phytoene desaturase. *J Bio Chem.* 273(22): 13658-13662.

18. Dailey HA, Dailey TA. **1996**. Protoporphyrinogen oxidase of myxcoccus xanthus. *J Bio Chem.* 271(15): 8714-8718.

19. Weerth RS, Medlock AE, Dailey HA. **2021**. Ironing out the distribution of [2Fe-2S] motifs in ferrochelatases. *J Biol Chem.* 297(5): 101017.

20. Dailey HA, Dailey TA, Wu CK, Medlock AE, Wang KF, Rose JP, Wang BC. **2000**. Ferrochelatase at the millennium: structures, mechanisms, and [2Fe-2S] clusters. *CMLS*. 57:1909-1926.

21. Medlock AE, Dailey TA, Ross TA, Dailey HA. Lanzilotta WN. **2007**. A π-helix switch selective for porphyrin deprotonation and product release in human ferrochelatase. *J Molec Bio.* 373(4): 1006-1016.

22. Dailey HA, Wu C-K, Horanyi P. Medlock AE, Najahi-Missaoui W, Burden AE, Dailey TA, Rose J. **2007**. Altered orientation of active site residues in variants of human ferrochelatase: evidence for a hydrogen bond network involved in catalysis. *Bichemistry*. 46: 7973-7979.

23. Medlock AE, Swartz L, Dailey TA, Dailey HA, Lanzilotta WN. **2007**. Substrate interactions with human ferrochelatase. *PNAS*. 104(6): 1789-1793.

24. Sano S, Granick S. **1961**. Mitochondrial coproporphyrinogen oxidase and protoporphyrin formation. *J Bio. Chem*. 236:1173–1180.

25. Ehteshamuddin AF. **1968**. Anaerobic formation of protoporphyrin IX from coproporphyrinogen III by bacterial preparations. *Biochem J*. 107: 446–447.

26. Frey PA, Magnusson OT. **2003**. S-adenosylmethionine: a wolf in sheep's clothing, or a rich man's adenosylcobalamin? *Chem Rev.* 103: 2129-2148.

27. Layer G, Pierik AJ, Trost M, Rigby SE, Leech HK, Grage K, Breckau D, Astner I, Jänsch L, Heathcote P, Warren MJ, Heinz DW, Jahn D. **2006**. The substrate radical of Escherichia coli oxygen-independent coproporphyrinogen III oxidase HemN. *J Bio Chem.* 281(23): 15727-15734.

28. Boynton TO, Daugherty LE, Dailey TA, Dailey HA. **2009**. Identification of *Escherichia coli* HemG as a novel, menadione-dependent flavodoxin with protoporphyrinogen oxidase activity. *Biochemistry*. 49: 6705-6711.

29. Boynton TO, Gerdes S, Craven SH, Neidle EL, Phillips JD, Dailey HA. **2011**. Discovery of a gene involved in a third bacterial protoporphyrinogen oxidase activity through comparative genomic analysis and functional complementation. *App Enviro Microbio*. 77(14): 4795-4801.

30. Dailey TA, Meissner P, Dailey HA. **1994**. Expression of a cloned protoporphyrinogen oxidase. *J Bio Chem.* 269(2): 813-815.

31. Celis AI, DuBois JL. **2019**. Making and Breaking Heme. *Curr Opin Struc Bio*. 59: 19-28.

32. Al-Karadaghi S, Hansson M, Nikonov S, Jönsson B, Hederstedt L. **1997**. Crystal structure of ferrochelatase: the terminal enzyme in heme biosynthesis. *Structure*. 5(11): 1501-1510.

33. Dailey HA and Gerdes S. **2015**. HemQ: An iron-coproporphyrin oxidative decarboxylase for protoheme synthesis in firmicutes and actinobacteria. *Arch Biochem Biophys.* 574: 27-35.

34. Hofbauer S, Mlynek G, Milazzo L, Pühringer D, Maresch D, Schaffner I, Furtmüller PG, Smulevich G, Djinović-Carugo K, Obinger C. **2016**. Hydrogen peroxide-mediated conversion of coproheme to heme b by HemQ--lessons from the first crystal structure and kinetic studies. *The FEBS Journal*. 283: 4386-4401.

35. Michlits H, Lier B, Pfanzagl V, Djinović-Carugo K, Furtmüller PG, Oostenbrink C, Obinger C, Hofbauer S. **2020**. Actinobacterial coproheme decarboxylases use histidine as a distal base to promote compound I formation. *ACS Catalysis*. 10: 5405-5418.

36. Streit BR, Celis AI, Moraski GC, Shisler KA, Shepard EM, Rodgers KR, Lukat-Rodgers GS, Dubois JL. **2018**. Decarboxylation involving a ferryl, propionate, and a tyrosyl group in a radical relay yields heme b. *J Biol Chem*. 293(11): 3989-3999.

37. Dailey HA, Gerdes S, Dailey TA, Burch JS, Phillips JD. **2015**. Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not us protoporphyrin. *PNAS*. 112(7): 2210-2215.

38. Cornaglia G. **2009**. Fighting infections due to multidrug-resistant gram-positive pathogens. *Clin Microbiol Infect*. 15(3): 209-211.

CHAPTER 2

THE MECHANISM OF COPROHEME DECARBOXYLASE

2.1 Introduction

Identification of the CPD pathway and the correct functional assignments of CgoX and ChdC brings a new promise and prospective for development of new antimicrobial compounds targeting heme biosynthesis in pathogenic bacteria that utilize the CPD pathway. The prevalence of the CPD pathway, specifically in multi-drug resistant Gram-positive bacteria, makes it a compelling therapeutic target. Of the three enzymes unique to the CPD pathway, the mechanism of the penultimate enzyme and subject of this work, ChdC, is the least understood.

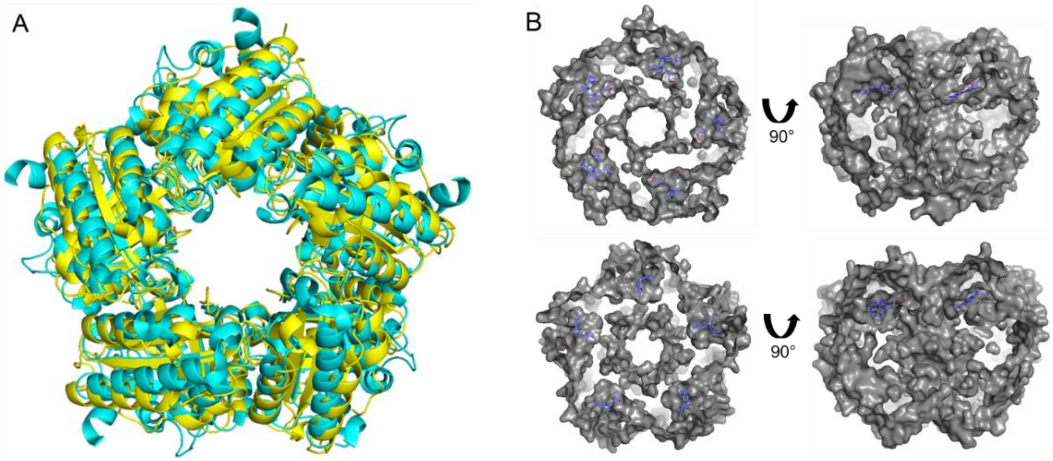


Figure 2.1

(A) Structural alignment of the cartoon representations of actinobacterial ChdC from *Streptomyces coelicolor* (*Sc*ChdC) in yellow and chlorite dismutase from *Candidatus* Nitrospira defluvii (NdCld) in cyan (PDB: 3NN2). (B) Surface representation of the overall homopentameric structure of actinobacterial *Sc*ChdC (top) and NdCld (bottom). The heme binding pocket of each domain is indicated by a bound molecule of MMD (*Sc*ChdC) or heme *b* (NdCld) shown as dark blue sticks.

Classified as a member of the heme binding chlorite dismutase family, ChdC was originally thought to be a chlorite dismutase. Like this class of proteins, ChdC retains a homopentameric structure with a molecular mass ~30 kDa (Figure 2.1A). The five monomers of ChdC and chlorite dismutases are arranged in a circular manner, resembling a donut, and each monomer contains a hydrophobic heme binding pocket (Figure 2.1A, B).

Despite the structural similarity between ChdC and chlorite dismutases, the knowledge that most organisms containing the *chdc* gene are unable to metabolize chlorite, along with the elucidation of the CPD pathway, led to the classification of ChdC as an oxidative decarboxylase (Dailey and Gerdes, 2015). Correct functional assignment and structural studies have begun to elucidate the mechanism of ChdC. Prior to introducing this mechanism, the prevalence of the CPD pathway in actinobacteria and firmicutes, necessitates highlighting key structural differences of ChdCs in each phylum. Both actinobacteria and firmicutes retain a flexible “active site loop,” that sits at the mouth of the active site; however, the length and architecture of each loop is significantly different. Recent data has shown that the active site loop actinobacteria ChdC is shorter and located closer to the active site, relative to a similar loop found in the firmicute ChdC (Figure 2.2) (Michlits et al., 2020). It has been suggested that the loop in firmicutes acts as an active site gate, regulating substrate entry and product release; however, the function of the loop has yet to be confirmed (Celis et al., 2017). In contrast, the active site loop of actinobacterial ChdCs have recently been shown to play a role in catalysis (Michlits et al., 2020). Specifically, a conserved histidine residue within the loop contributes to properly orienting the

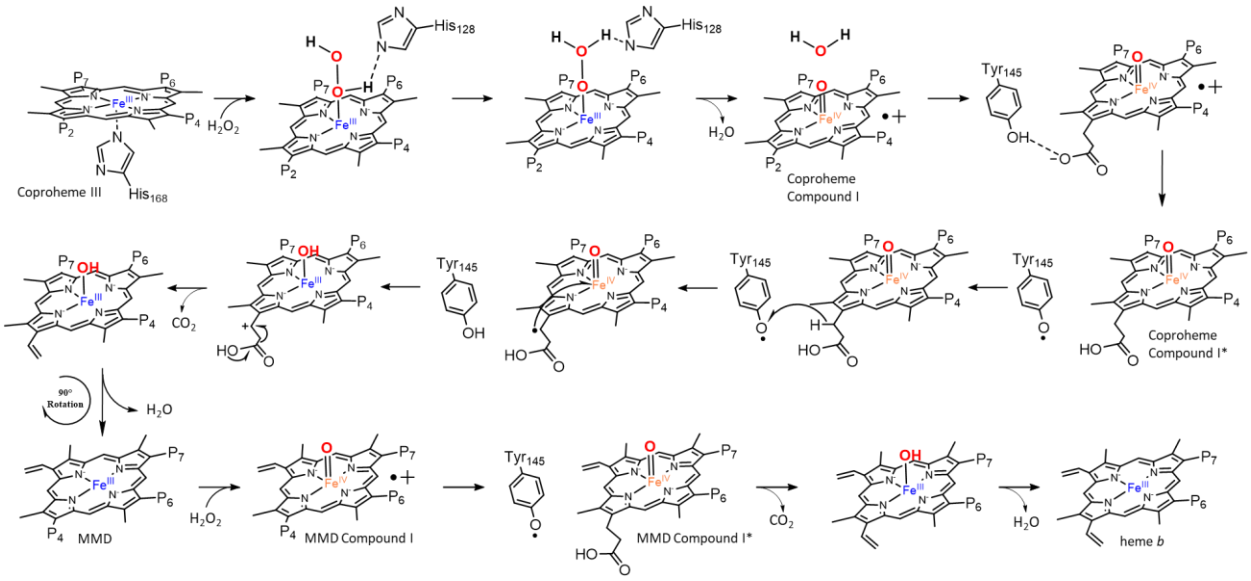
substrate by hydrogen bonding with one of the propionates of coproheme. Additionally, this histidine acts as a distal base, assisting with the deprotonation of H_2O_2 (Michlits et al., 2020). An electron acceptor is required for decarboxylation and, although neither the precise mechanism of electron transfer, nor the binding site for the acceptor, have been described, studies have shown that FMN and H_2O_2 will both mediate the reaction *in vitro* (Dailey et al., 2015). H_2O_2 is believed to be the biologically relevant electron acceptor, as 3 molecules of H_2O_2 are produced upstream in the CDP pathway**.**

Figure 2.2

Cartoon representation showing a structural alignment for the single monomer from actinobacterial *Sc*ChdC (yellow) and firmicute *Listeria monocytogenes* ChdC (*Lm*ChdC, pink, PDB: 6FXQ). The active site loops of both species are circled in black.

Previous work established a mechanism for decarboxylation of coproheme in actinobacterial ChdC (Scheme 2.1). In general, ChdC utilizes a peroxide-dependent mechanism where the substrate also functions as a coenzyme (Dailey and Gerdes, 2015). Specifically, H_2O_2 binds to the ferric iron of a penta-coordinate coproheme bound to the enzyme through a conserved histidine (H168). A conserved distal histidine (H128) acts as a general base, promoting the heterolytic cleavage of H_2O_2, resulting in the release of H_2O and the formation of the Compound I, a porphyrin cation radical intermediate containing an

oxyferryl (Michlits et al., 2020). This state is known as Compound I, and abstracts a hydrogen atom from another conserved active site residue, Y145, to yield a tyrosyl radical. As is depicted in Scheme 2.1, the tyrosyl radical abstracts a hydrogen atom from the beta carbon of the propionate group, p2. This triggers decarboxylation of the propionate group and reduction of the porphyrin ring and iron, leading to formation of a vinyl group and production of CO_2, and water (Streit et al., 2018). Conversion of the first propionate group to a vinyl group yields the intermediate monovinyl monopropionate deuteroporphyrin (MMD) (Dailey et al., 2015). A key aspect of the proposed mechanism is that MMD must rotate 90 degrees in order to position propionate group p4 near the catalytic tyrosine residue (Milazzo et al., 2019). A second molecule of H_2O_2 binds to the ferric iron and the mechanism repeats, completing the reaction and production of heme *b*.



Scheme 2.1

The proposed reaction mechanism for the decarboxylation of coproheme to form heme *b* via ChdC. The color of the central iron correlates to oxidation state changes throughout the mechanism.

To further investigate specific aspects of the proposed mechanism, site-specific mutations were made and subjected to rigorous characterization. A total of seven point mutations were made, including the following; T128C, H128Y, H168C, L208F, G169V, G169D, and G169L.The rational for these variants is summarized in Table 2.1. Each variant was then investigated by biochemical assays, spectroscopic evaluations, and crystallization to further assess the effect of the mutation and if the observations were consistent with its proposed role. In addition to the proposed role, Table 2.1 also summarizes predicted outcomes for each of the altered residues.

Table 2.1

Summary of intended effects and possible outcomes of mutated *Sc*ChdC.

Variant	Proposed Mechanistic Role	Possible Outcomes
T182C	Affect the coordination of coproheme by blocking H_2O_2 binding site.	Prevent substrate turnover; influence rate of reaction.
G169V G169L G169D	Block substrate entry or slow down the rotation of the porphyrin. Side-chain interactions may prevent substrate binding or Compound I formation.	Prevent substrate binding; capture active site intermediates
H168C	Prevent the porphyrin from exiting the active site.	Prevent turnover; capture product-bound structure
H128Y	Cause π stacking interactions, reducing ruffling of the porphyrin ring and preventing MMD from rotating. Possible coordination of tyrosine to iron may block H_2O_2 binding site.	Prevent turnover; capture MMD prior to rotation
L208F	A larger residue within hydrophobic channel may block rotation of the porphyrin.	Capture MMD prior to rotation

The T182C variant was made to prevent substrate turnover and facilitate the formation of a substrate-bound protein complex. This threonine is strictly conserved and sits directly above MMD in the active site of our current model (Figure 2.3A, B). We had several hypotheses in respect to the result of this mutation. First, we hypothesized that by mutating T182 to a larger residue, with metal-binding properties, such as cysteine,

would disrupt the water or hydrogen bonding network required for deprotonation of peroxide, or potentially serve as a ligand for coproheme. Alternatively, peroxide may still react, but lead to different chemistry, potentially an oxidation of the porphyrin structure. The variant could also simply influence the rate of reaction and turnover to be much slower. The latter would be interesting as it would result in changing the kinetics and potentially facilitate the characterization of new intermediates.

H168 is conserved and serves as the proximal ligand to the iron (Figure 2.3A, C). This histidine residue was also mutated to a cysteine to either disrupt a key proton transfer event, preventing the porphyrin from exiting the active site post-turnover or alter the redox couple of the iron atom.

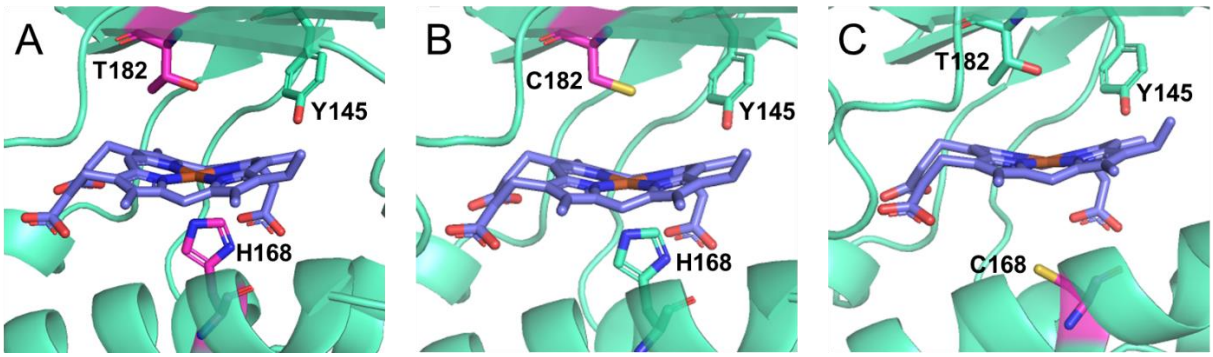


Figure 2.3

Cartoon representation of wt *Sc*ChdC active site (A) and the variants T182C (B) and H168C (C) obtained by in silico mutations using PyMOL. Important residues are shown as pink sticks and MMD is shown as dark blue sticks.

G169 is located in the conserved α-helix at the opening of the active site and was mutated to valine, leucine, and aspartic acid (Figure 2.4). All three amino acids (V, L, and D) favor helix formation. However, given G169 sits at the entrance to the active site, the variant side chains could sterically constrain this loop, directly addressing loop

flexibility importance. Another possibility would be to influence rotation of porphyrin within the active site, trapping intermediates. In addition, at higher pH values, an aspartic acid may form a salt bridge between G169D and H168, preventing either substrate binding or formation of Compound I.

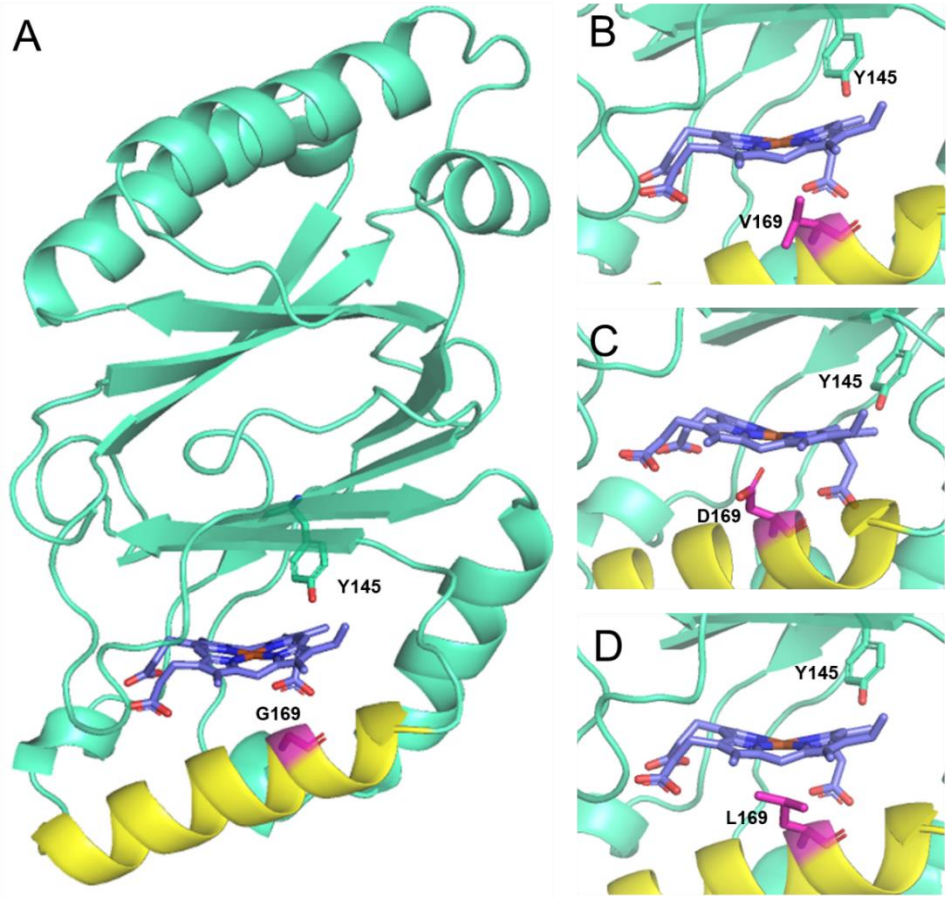
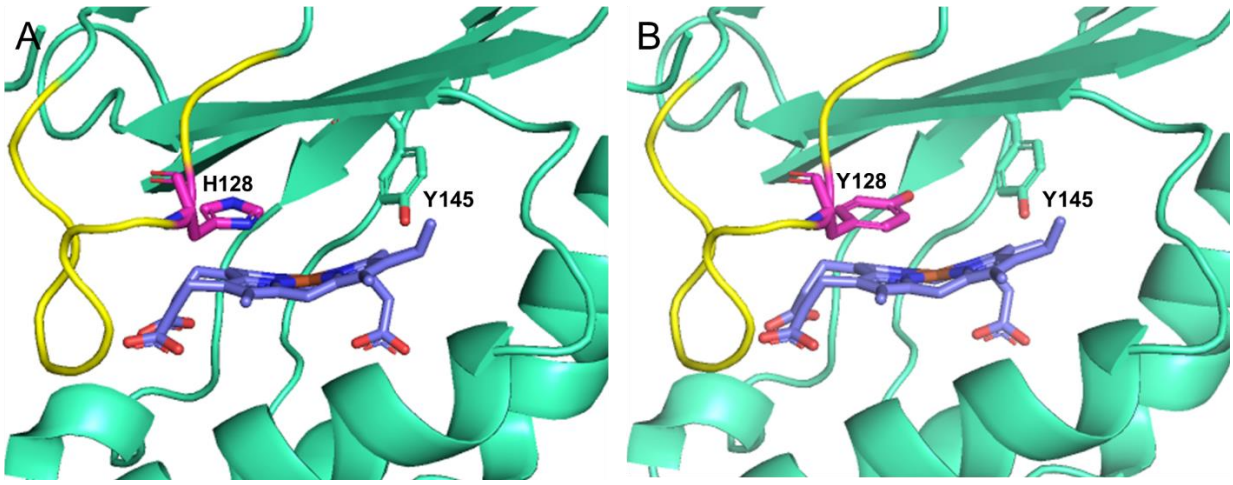


Figure 2.4

Cartoon representation of a single monomer of wt *Sc*ChdC (A) and the variants G169V *Sc*ChdC (B), G169D *Sc*ChdC (C), and G169L *Sc*ChdC (D) obtained by in silico mutations using PyMOL. The conserved α-helix containing G169 is highlighted in yellow. Important residues are shown as pink sticks and MMD is shown as dark blue sticks.

Two additional important features of the active site are a loop that could function to cover the entrance, and a hydrophobic channel that we hypothesize aids in rotation of the intermediate (Figure 2.5). Recent studies indicated a conserved histidine within the active site loop of a ChdC from actinobacteria acts as a general base in the deprotonation and heterolytic cleavage of H_2O_2, which inspired the H128Y variant (Michlits et al., 2020; Figure 2.5A). The intention of this mutation was to create π stacking interactions between the tyrosine and coproheme, reducing ruffling of the porphyrin ring, and preventing MMD rotation. In addition, coordination in tyrosine to iron of coproheme would prevent H_2O_2 binding and subsequent turnover. The hydrophobic channel, required to rotate the MMD intermediate, is made up of F197, M209, L208, and V178 (Figure 2.5B). Following oxidative decarboxylation of the first propionate group, the resulting vinyl group passes through the hydrophobic channel during intermediate rotation. We chose to mutate L208 to a phenylalanine with the hypothesis that a larger residue at this position could block rotation of the porphyrin, or stack on the porphyrin ring structure allowing us to capture an intermediate prior to rotation.



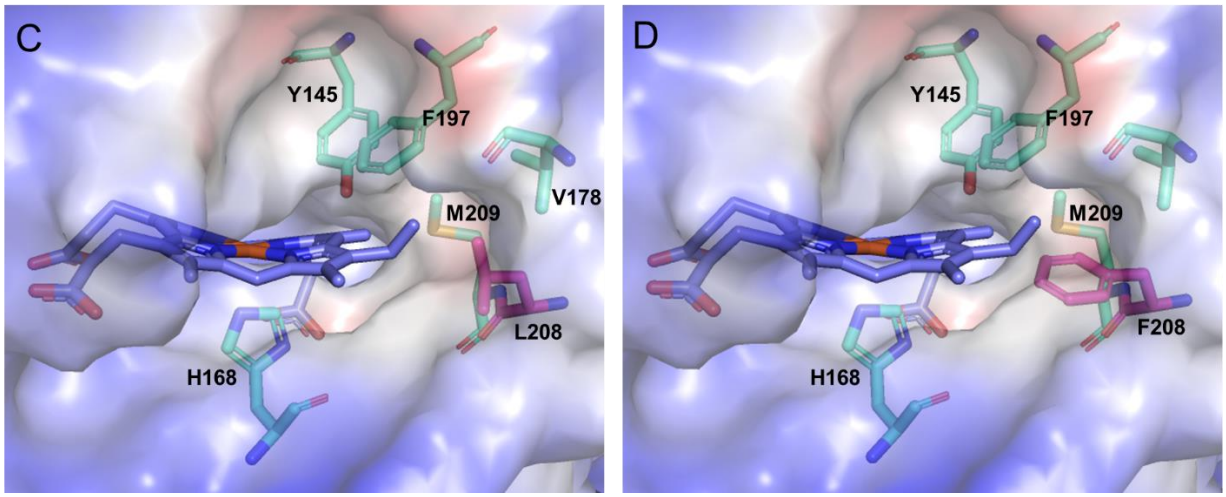


Figure 2.5

Cartoon representations of wt *Sc*ChdC (A) and the H128Y *Sc*ChdC variant (B) and surface representations of wt *Sc*ChdC (C) and the L208F *Sc*ChdC variant (D). Models of *Sc*ChdC variants were obtained by in silico mutations using PyMOL. Important residues are shown as pink sticks and MMD is shown as dark blue sticks. The "active site loop" containing H128 is highlighted in yellow. The surface representation in panel C and D is colored by electrostatic potential, depicting the active site hydrophobic channel in red

While previous work has answered many of the questions surrounding the ChdC mechanism, unanswered questions regarding the mechanics behind rotation of the porphyrin intermediate and product release remain. It has recently been accepted that the rotation of the intermediate occurs within the active site, rather than MMD exiting and reentering the active site in the correct orientation (Sebastiani et al., 2021). Past work had shown that ChdC can reversibly bind and release product, suggesting that this enzyme may fulfill a regulatory role in the CPD pathway (Hofbauer et al., 2015). Therefore, a primary focus of this research is to further characterize the rotation mechanism of the intermediate, MMD. Specifically, we propose that this is the rate limiting step, and that movement is initially facilitated by repulsion forces between the

newly formed vinyl group and surrounding residues. Based on structural data presented in this work, we hypothesize that due to the loss of the negatively charged propionate group after the first decarboxylation, the vinyl group is excluded from the polar region next to the catalytic Y145 residue and slides 90 degrees through a hydrophobic channel composed of F197, M209, L208, and V178. Additionally, we propose that as MMD rotates, R149 helps guide the second propionate to its post-rotation position, next to the catalytic Y145. To investigate this hypothesis, we have taken a multifaceted approach involving spectroscopic and structural characterization of ChdC enzymes.

2.2 Methods

Protein Expression and Site-Directed Mutagenesis

All proteins were expressed using a pTrcHisA constitutive expression vector. *Streptomyces coelicolor* ChdC was cloned by Tammy Dailey; the detailed method can be found in Dailey et al. 2010 (Dailey et al., 2010). Variant *Sc*ChdCs were synthesized using the NEB Q5 site directed mutagenesis protocol. The primers and correlating melting temperatures for each variant are listed below (Table 2.2). Wild type (wt) and variant *Sc*ChdC DNA was subsequently transformed into BL21 (DE3) Competent *E. coli* competent cells.

Table 2.2

Froward and reverse primers and correlating melting temperatures (T_m) used for the site directed mutagenesis of *Sc*ChdC.

Variant	Primer		T_m
T182C	Forward	GCGCGCCAAC**tgc**GTCGCCTCCTTCTCGCTGGG	72
	Reverse	ACGTCCGGGTAGCCGCGG	
	Forward	CAACCGCTCG**tat**ATCCCGGCCTTCC	

H128Y	Reverse	AACTCGGCGGGGCGGTGC	72
H168C	Forward	GCTCGCCGAC**tgc**GGCAAGATGGCCC	72
	Reverse	ATGCGGCGGCGGTCCTCG	
G169V	Forward	CGCCGACCAC**gtg**AAGATGGCCCGCG	72
	Reverse	AGCATGCGGCGGCGGTCC	
G169L	Forward	CGCCGACCAC**ctg**AAGATGGCCCGCGG	72
	Reverse	AGCATGCGGCGGCGGTCC	
G169D	Forward	CGCCGACCAC**gat**AAGATGGCCCGCG	72
	Reverse	AGCATGCGGCGGCGGTCC	
L208F	Forward	CATCGTCGAC**ttt**ATGCGCCACCTGC	71
	Reverse	CGGTGCAGCTCGTCGGCC	

Protein Purification

Transformed cells were grown in 6 x 1L Circle Grow media in 2L baffled flasks at 30°C for 20-24 hours. Cell cultures produced were spun down and harvested cell pellets were resuspended in 100mM HEPES/KCl pH7.5. Resolubilized cells were lysed using a French Press and lysates were cleared via ultra-centrifugation. Resulting lysates were administered onto an equilibrated Cobalt Chelation Resin (G-Biosciences) and the column was subsequently washed with 3 column volumes of 100mM HEPES/KCl pH7.5. Purified protein was eluted with a series of imidazole solutions (50, 100, 150, 200mM) and eluted fractions were analyzed using SDS gel electrophoresis to confirm presence and quality of purified protein. Clean fractions were then concentrated down to 1-2 mL using Amicon Ultra-15 centrifuge filters. Concentrated protein was dialyzed overnight in 100mM HEPES/KCl pH 7.5.

Reagents and Stocks

Ferric coproheme was purchased from Frontier Scientific, Inc. (Logan, Utah, USA) as lyophilized powder. A coproheme solution was prepared by first dissolving the coproheme powder in a 0.5 M NaOH solution. Subsequently, a concentrated aliquot of this solution was diluted to an appropriate volume with a 100mM HEPES/KCl buffer, pH 7.5. A hemin solution, used for product binding and crystallization experiments, was prepared by first resolubilizing solid hemin in 0.5M NaOH and DMSO. This solution was then diluted to an appropriate volume with 100mM HEPES/KCl buffer, pH 7.5.

Crystallization and Structure Determination

Five crystallization screens were used for initial screening to determine optimal crystallization conditions for each of the variant *Sc*ChdC proteins: PEG/Rx, Crystal Screen, Index, and PEG/Ion by Hampton Research and Wizard Classic by Rigaku. Initial screens were set up for substrate and product bound *Sc*ChdC variants using a sitting drop vapor diffusion method, where the drop consisted of a 1:1 ratio of crystal screen condition and protein. After the initial screening, matrix screens of suitable conditions were developed, optimizing pH, salt, additives, or precipitation reagents. A batch diffusion method using capillaries was also utilized to improve crystal size and diffraction quality for identified conditions. An in-house diffractometer was used to assess diffraction quality and determine the space group before sending samples to SER-CAT. Subsequent rounds of model building, and refinements were performed using the programs COOT (Emsley et al., 2010) and PHENIX (Afonine et al., 2018).

UV-Visible Spectroscopy

Protein concentrations, substrate and product binding, and substrate turnover were all measured using a HP 8453 UV/Vis spectrophotometer. Absorbance data from 250-900 nM was collected and analyzed using OlisWorks and GlobalWorks software, respectively (Olis, Inc., Athens, GA). Substrate and product binding assays were performed at 10 μM protein in 100 mM HEPES/100 mM KCl pH7.5 buffer. To begin, sample was blanked, and equimolar amounts of substrate or product were added to the cuvette. The assay was monitored for 5 minutes with constant mixing and scanning every 3 seconds. Substrate turnover assays were performed by mixing 10 μM protein with 10-60 μM H_2O_2 in 100 mM HEPES/100 mM KCl pH7.5 buffer. Prior to start, the sample was blanked, followed by addition of 10 μM of coproheme. The assay was monitored for 15 minutes with constant mixing and scans every 9 seconds.

Ultra-High Performance Liquid Chromatography (UHPLC) Assays

ChdC turnover assays were performed by mixing equimolar amounts of wild type or variant ChdC with substrate. Turnover was initiated by the addition of 0.5-3 eq H_2O_2. After five minutes, the reaction was terminated with 10 μL concentrated HCl and iron porphyrins were extracted with ethyl acetate phase separation. Dried porphyrins were resolubilized in 0.5% triethylamine and evaluated with an Agilent UHPLC system. Each sample was injected into a C18 Poroshell column at a flow rate of 0.5 mL/min at 50°C. Solvent A was 0.1% formic acid in Milli-Q H_2O and solvent B was 0.1% formic acid in 100% acetonitrile. Analytes of interest were eluted following a linear gradient from 0 to 90% B over 6 min.

2.3 Results

Crystal structure of MMD bound *Sc*ChdC

The crystal structure of wild type (wt) *Sc*ChdC in complex with MMD was determined to a resolution of 2.7 Å (Table 2.3). The MMD bound structure was obtained through sitting drop optimizations performed by Clayton Pritchett. The protein solution was prepared by mixing equimolar concentrations of wt *Sc*ChdC and coproheme (15 mg/mL) in 100 mM HEPES/100 mM KCl pH7.5 buffer. 10 mM of KCN^- was added to the protein solution to prevent substrate turnover during crystallization. After 9 weeks, large, red, square bipyramidal crystals formed and were collected for X-ray diffraction. These crystals were initially screened on an in-house diffractometer to confirm diffraction prior to being sent to SER-CAT for data collection. Several rounds of model building and data refinement were performed, and despite the addition of cyanide, the refined electron density indicated the presence of MMD rather than coproheme. Likely small amounts of H_2O_2 present in aqueous solution promoted the first decarboxylation of ChdC.

Table 2.3

Data Collection and Refinement Statistics of wt *Sc*ChdC. Statistics for the highest-resolution shell are shown in parentheses.

	wt *Sc*ChdC
Wavelength (Å)	1.0
Resolution range (Å)	39.1 - 2.698 (2.794 - 2.698)
Space group	P 41 21 2
Unit cell (Å)	184.842 184.842 241.463 90 90 90
Unique reflections	114784 (11324)
Completeness (%)	99.82 (99.91)

Wilson B-factor (Å2)	62.19
Reflections used in refinement	114767 (11324)
Reflections used for R-free	2000 (197)
R-work	0.1898 (0.2725)
R-free	0.2366 (0.3278)
Number of non-hydrogen atoms	18835
macromolecules	18319
ligands	460
solvent	56
Protein residues	2215
RMS (bonds) (Å)	0.009
RMS (angles) (°)	1.05
Ramachandran favored (%)	92.03
Ramachandran allowed (%)	7.33
Ramachandran outliers (%)	0.64
Rotamer outliers (%)	3.57
Clashscore	15.56
Average B-factor (Å2)	60.17
Macromolecules (Å2)	60.08
Ligands (Å2)	65.49
Solvent (Å2)	46.93

The overall fold of the MMD-bound *Sc*ChdC structure, presented in this work, is consistent with previously published structural data for actinobacterial ChdC from *Corynebacterium diphtheriae (Cd*ChdC) (PDB: 6XUB; Michlits et al., 2020). Two

homopentamers are seen in the asymmetric unit of *Sc*ChdC, with a single homopentamer being the physiologically relevant form of the enzyme. The monomers are arranged in a circular manner, resembling a donut, with each monomer containing an N-terminal active site to bind the redox-active substrate (Figure 2.6).

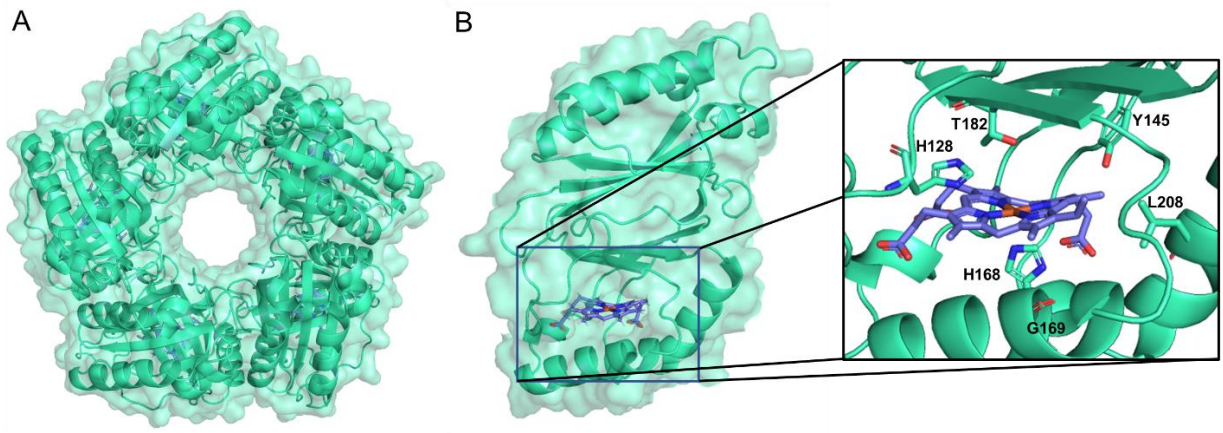


Figure 2.6

Cartoon and surface representation showing (A) the overall homopentameric structure and (B) single monomer of *Sc*ChdC. The inset outlined in black depicts the active site with residues relevant to this work shown as sticks. A single molecule of MMD is bound to the N-terminus of each monomer and is shown as dark blue sticks.

MMD is axially coordinated by a conserved proximal histidine (H168), and the distal position is open for peroxide binding. Residues Y145 and H128 were shown to be catalytically relevant in actinobacterial ChdC. EPR signatures collected by Streit et al. (2018) revealed the formation of a tyrosyl radical within 30 s of initiating the reaction. The same experiment using MMD rather than coproheme showed similar EPR signatures, indicating that formation of the radical T145 residue is involved in oxidative decarboxylation of the substrate and intermediate (Streit et al., 2018). It was suggested that the reaction between coproheme-ChdC and H_2O_2 results in the formation a porphyrin cation radical intermediate, Compound I, which oxidizes Y145 to a tyrosyl

radical (Streit et al., 2018). This radical goes on to abstracts a hydrogen atom from the β-carbon of the porphyrin propionate group, initiating decarboxylation of the propionate group to a vinyl. Using site-directed mutagenesis and stopped-flow spectroscopy, Michlits et al. (2020) showed a conserved residue, H118 (H128 in *Sc*ChdC), in the active site loop acts as a distal base, assisting with Compound I formation. Turnover reactions of wt *Cd*ChdC and *Cd*ChdC variants Y135A, and H118A/Y135A were monitored via stopped-flow spectroscopy and Compound I formation was traced using a single exponential. Data indicated Compound I formation was 2 orders of magnitude slower in H118A *Cd*ChdC compared to that of the wt and Y135A (Michlits 2020). The significant decrease in the rate of Compound I formation seen in the absence of H118, the higher catalytic efficiency of actinobacteria, and the proximity of H118 to the H_2O_2 binding site, led to the conclusion that H118 acts as a distal base to assist with Compound I formation (Michlits 2020).

Crystal structure of coproheme bound *Sc*ChdC

The coproheme bound structure of *Sc*ChdC was obtained through sitting drop optimizations of the T182C variant following crystal screening of the wt and *Sc*ChdC variants (Table 2.4). Following initial screening, suitable conditions for crystallization were optimized through the capillaries, kosmotrope additives, and microseeding. During the optimization trials, we conducted a kosmotrope screen, adding 1-200mM betaine to select conditions from initial screening. Addition of 200mM betaine to one of the parent conditions (0.2 ammonium citrate tribasic pH 7.0, 20% w/v polyethylene glycol 3,350) produced large red crystals (Figure 2.7). The T182C variant crystals diffracted to 3.8Å and were subsequently used for microseeding.

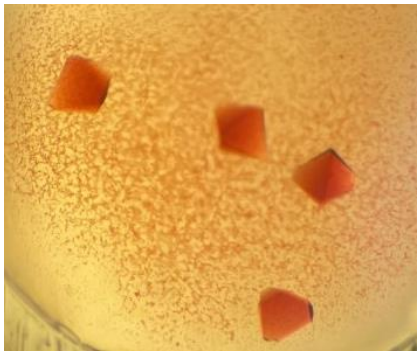


Figure 2.7

Substrate bound T182C *Sc*ChdC crystals. Crystals were grown in 0.2 ammonium citrate tribasic pH 7.0, 20% w/v polyethylene glycol 3,350, and 200mM betaine.

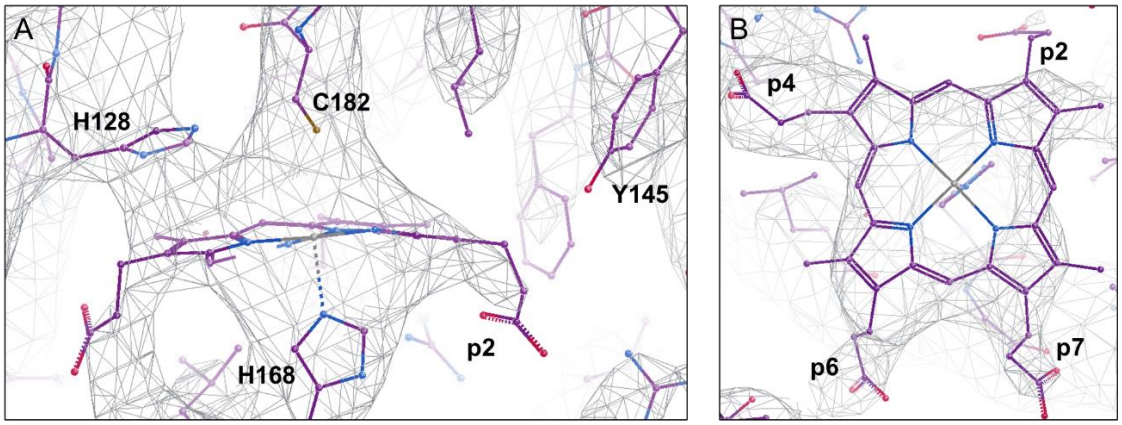
Table 2.4

Data Collection and Refinement Statistics of T182C *Sc*ChdC. Statistics for the highest-resolution shell are shown in parentheses.

	T182C *Sc*ChdC
Wavelength (Å)	1.0
Resolution range (Å)	39.03 - 3.808 (3.944 - 3.808)
Space group	P 41 21 2
Unit cell (Å)	185.374 185.374 245.277 90 90 90
Unique reflections	42122 (3999)
Completeness (%)	99.04 (94.63)
Wilson B-factor (Å2)	119.49
Reflections used in refinement	42100 (3998)
Reflections used for R-free	2000 (190)
R-work	0.2288 (0.2777)
R-free	0.2686 (0.3561)
Number of non-hydrogen atoms	18730
macromolecules	18270
ligands	460
Protein residues	2221

RMS (bonds) (Å)	0.008
RMS (angles) (°)	1.12
Ramachandran favored (%)	95.12
Ramachandran allowed (%)	4.20
Ramachandran outliers (%)	0.68
Rotamer outliers (%)	0.00
Clashscore	8.68
Average B-factor (Å2)	129.23
Macromolecules (Å2)	129.27
Ligands (Å2)	127.87

Despite low resolution, analysis of T182C variant electron density and active site structure does support our hypothesis that the threonine to cysteine mutation would capture substrate (Figure 2.8). Specifically, we observe continuous density from the porphyrin iron to the cysteine, as well as density to support the presence of substrate, coproheme, not the MMD intermediate (Figure 2.8).



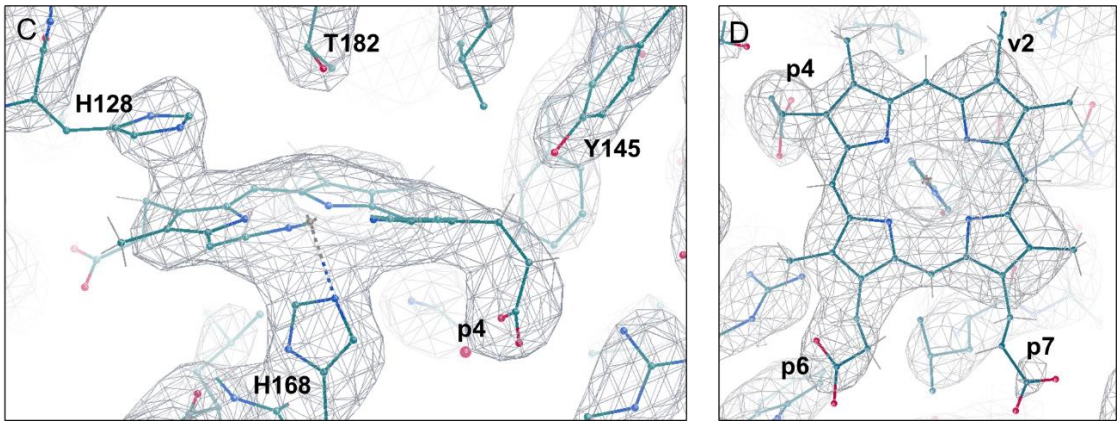


Figure 2.8

Ball and stick model for the T182C variant (A,B) and wt *Sc*ChdC (C,D) with the $2F_o$-F_c composite omit map (grey cage). C182 blocks H_2O_2 from binding, trapping substrate in the active site of T182C *Sc*ChdC. T182 in wt *Sc*ChdC allows space for H_2O_2 binding, resulting in partial turnover to MMD captured in the active site. Composite omit 2|Fo| − |Fc| electron density maps are shown at σ = 1.5.

With structural data for both coproheme and MMD bound to *Sc*ChdC, it is possible to analyze the architecture of each active site for structural differences. The most significant difference is the porphyrin interactions with an active site arginine residue, R149. In the coproheme-*Sc*ChdC structure, R149 is located near, and appears to interact with, propionate 4 (p4) (Figure 2.9). However, in the MMD-*Sc*ChdC structure, where substrate has undergone a 90-degree rotation, R149 does not appear to interact with the propionate (p6) that took the position of p4 after the rotation, as expected. Rather, the post-rotation orientation of MMD suggests continued interaction between R149 and p4, in its post-rotation position (Figure 2.9). This observation supports the hypothesis that R149 facilitates proper substrate orientation upon binding and rotation of the porphyrin intermediate.

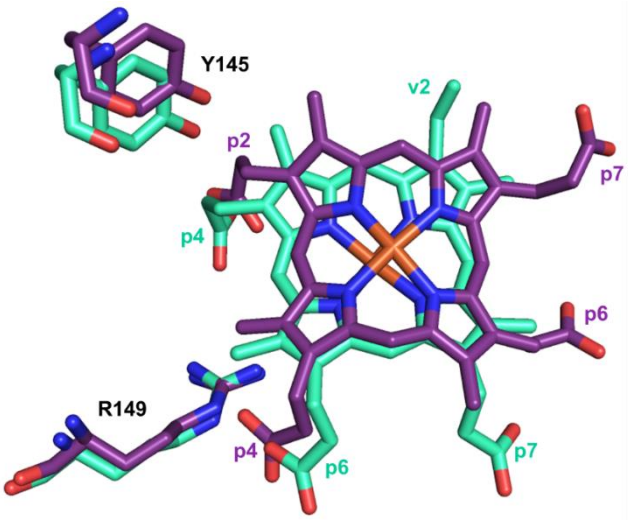


Figure 2.9

Overlay of the MMD bound wt *Sc*ChdC active site (green) and the coproheme bound T182C *Sc*ChdC active site (purple), shown as sticks. Active site residue R149 shifts positions with the rotation of the substrate.

Substrate and Product Binding

To assess the active site accessibility and binding of the wt and variant *Sc*ChdC, substrate and product binding assays were performed using UV-Visible spectroscopy. This provided a simple way to compare porphyrin binding modes by comparing absorption bands for the wt and variant enzyme. Spectral data for substrate binding to the wt and variants is summarized in Tables 2.5.

Table 2.5

Summary of absorption peaks, spin state assignments, and proposed porphyrin-protein complex following the addition of coproheme.

Protein	Soret maximum (nm)	Prominent bands (nm)	Spin State	Proposed Complex
*Sc*ChdC	393	495, 530, 625	5cQS	coproheme-ChdC
T182C	391	493, 530, 610	5cQS/5cHS	coproheme-ChdC; free coproheme
H128Y	392	493, 530, 625	5cQS	coproheme-ChdC
H168C	394	500, 530, 585, 625	5cQS/6cLS/ 5cHS	coproheme-ChdC; compound I-ChdC; MMD-ChdC
L208F	398	535, 575, 615	5cQS/6cLS/ 5cHS	coproheme-ChdC; compound I-ChdC; MMD-ChdC

G169V	393	492, 530, 575, 615	5cQS/6cLS/ 5cHS	coproheme-ChdC; compound I-ChdC; MMD-ChdC
G169D	391	490, 525, 610	5cQS/5cHS	coproheme-ChdC; free coproheme
G169L	393	493, 530, 575, 615	5cQS/6cLS/ 5cHS	coproheme-ChdC; compound I-ChdC; MMD-ChdC

The spectral features of wt *Sc*ChdC in complex with coproheme are in agreement with published data for actinobacterial ChdC (Pfanzagl et al., 2018). The wt coproheme-ChdC complex has a Soret maxima at 393 nm, Q-band features at 495 and 530, and a charge-transfer (CT) at 625 (Figure 2.10; Table 2.5). Though uncommon in biological systems, recent work established ferric ChdC exits in a five-coordinate quantum mechanically mixed spin state (5cQS) (Milazzo et al., 2018). In agreement with this finding, the CT band at 625 nm, in the wt coproheme-ChdC complex, indicates this 5cQS species (Smulevich et al., 2005). Of the seven *Sc*ChdC variants analyzed in this dataset, the H128Y variant, in complex with coproheme, displays absorbance features most similar to the wt complex. Specifically, H128Y had a Soret maxima at 392 nm, Q-bands at 493 and 530 nm, as well as a CT band at 625 nm (Figure 2.10; Table 2.5).

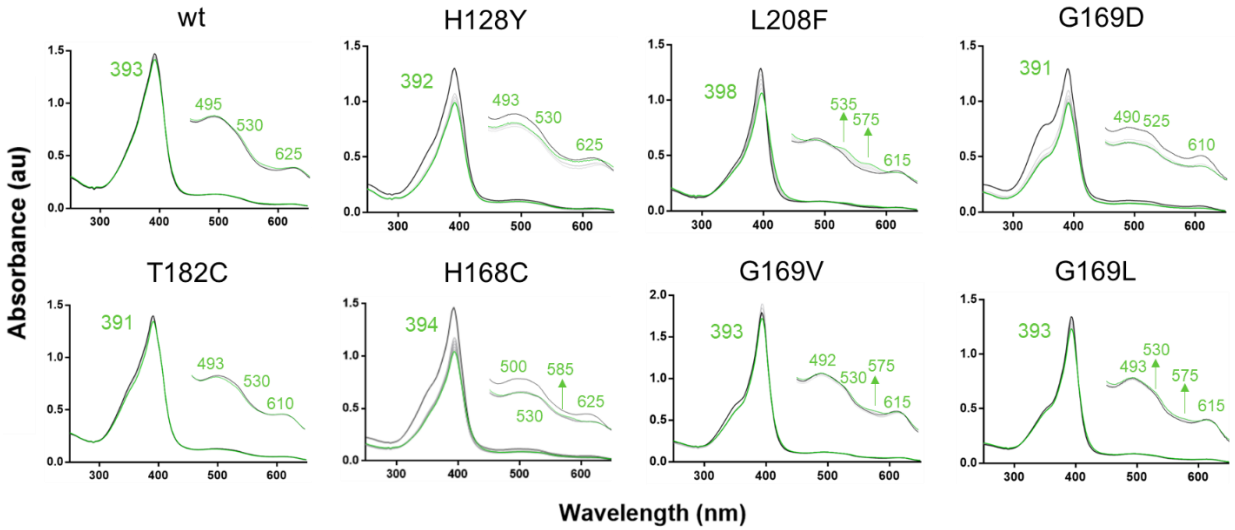


Figure 2.10

UV-Visible spectra following coproheme binding to wt and variant *Sc*ChdCs. Each spectrum follows the spectral transition upon mixing 10 μM coproheme with 10 μM *Sc*ChdC. The initial scan, taken after 3s of mixing, is shown in black. The final scan, taken after 300s and colored green, depicts the coproheme-*Sc*ChdC complex. Intermediate spectral scans are shown in gray.

The T182C and G169D variants have similar spectral features, with the Soret maxima and CT bands that are slightly blue shifted compared to the wt enzyme at 391 and 610 nm, respectively (Figure 2.10; Table 2.5). They each also display Q-band features around 490 and 530 nm, like the wt-coproheme complex (Figure 2.10; Table 2.5). The spectral features of T182C and G169D are indicative of 5cHS free coproheme in solution, suggesting a reduced affinity for substrate (Milazzo et al., 2018).

G169V, H168C, G169L, and L208F in complex with coproheme display red-shifted Soret maxima of 393, 394, 393, and 398 nm, respectively (Figure 2.10; Table 2.5). The Q-band features and CT bands are wt-like, with the exception of an additional peak near 580 nm for all four variants and a blue-shifted CT band at 615 nm for G169V, G169L and L208F (Figure 2.10; Table 2.5). The blue-shifted CT band is proposed to be the MMD-ChdC complex (Streit et al., 2017) and the peak near 580 nm was recently established as spectral marker for MMD-bound ChdC (Sebastiani et al., 2022). Sebastiani et al. (2022) also recently established the MMD-ChdC complex as a 5cHS system. The peak at 575 nm was assigned to ChdC in complex with Compound I (Michlits et al., 2020) and is characteristic of a six-coordinate low-spin (6cLS) species, the typical spin state of the ferric porphyrin cation radical, Compound I (Pfanzagl et al., 2018). These findings suggest G169V, H168C, G169L, and L208F are prone to partial turnover using residual H_2O_2 already in solution. For the H168C variant, this is

surprising considering H168 is a ligand to the iron of coproheme and the substitution to a cysteine ligand would significantly decrease the reduction potential of the iron atom (Hosseinzadeh and Lu, 2016).

The wt *Sc*ChdC protein with heme reflects spectral features of a six-coordinate high-spin (6cHS) system, with a Soret maxima at 407 and Q-bands at 535 and 580 (Sebastiani et al., 2022). H128Y, G169V, and G169D all display wt-like features for the heme-ChdC complex, with Soret maxima at 406, 407, and 407 nm, respectively (Figure 2.11; Table 2.6). H128Y and G169V heme complexes both display Q-bands around 530 and 580, similar to the wt (Figure 2.11; Table 2.6). However, H128Y, G169V, and G169D have additional bands around 490 and 630, indicating the presence of a 6cLS species in addition to the 6cHS system (Pfanzagl et al., 2018). This may imply that these variants increase flexibility of the active site, allowing substrate to adopt several different orientations.

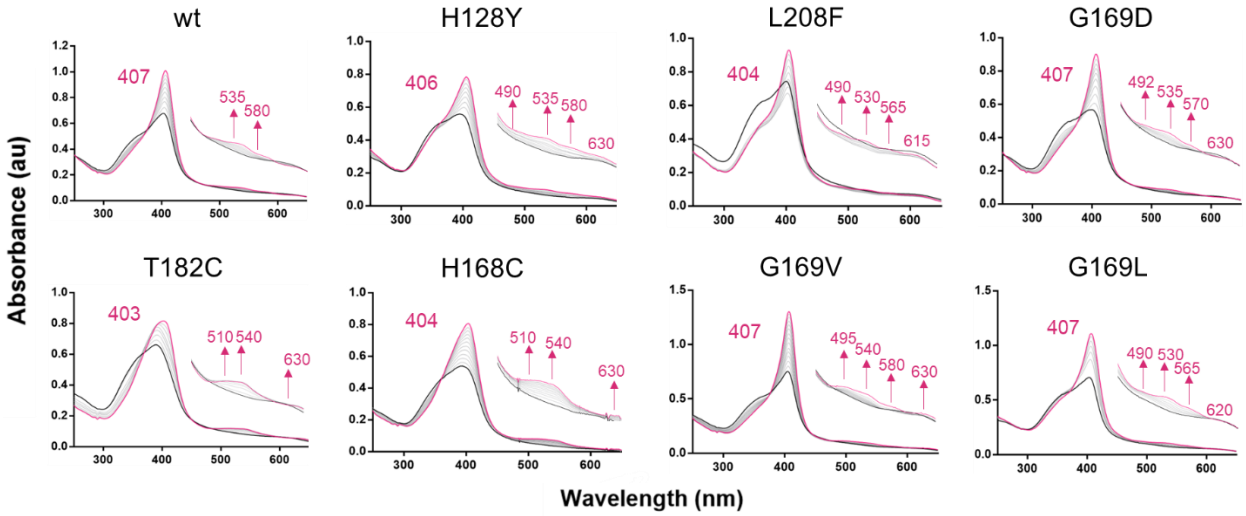


Figure 2.11

UV-Visible spectra following heme binding to wt and variant *Sc*ChdCs. Each spectrum follows the spectral transition upon mixing 10 μM hemin with 10 μM *Sc*ChdC. The initial

scan, taken after 3s of mixing, is shown in black. The final scan, taken after 300s and colored pink, depicts the heme-*Sc*ChdC complex. Intermediate spectral scans are shown in gray.

L208F, T182C, H168C, and G169L in complex with heme all display a blue-shifted Soret maxima ranging from 403 to 404 nm (Figure 2.11; Table 2.6). Though these four variants have similar Soret maxima, their Q-bands and CT bands vary a great deal. G169L, T182C, and H168C all have a Q-band at 540 nm, however, T182C and H168C have an additional Q-band at 510 nm and a CT band at 630 nm (Figure 2.11; Table 2.6). The spectral features seen for T182C and H168C are similar to other heme binding proteins, where heme is coordinated by cysteine (Vetter et al., 2009). L208F displays Q-bands features at 490, 530, and 565 nm and a CT band at 615 nm. While G169L has a wt-like Soret maximum (407 nm), the Q-bands collected for this variant are similar to that of L208F. The peaks at 490 and 615 nm are indicative of 5cHS free hemin, suggesting poor binding of the product (Neugebauer et al., 2012). HPLC and UV-Visible spectroscopic analysis of coproheme decarboxylase activity suggests L208F has a higher affinity for heme than the other variants (Figure 2.12 & 2.13).

Table 2.6

Summary of absorption peaks, spin state assignments, and proposed porphyrin-protein complex resulting from the mixing of wt and variant *Sc*ChdC proteins with hemin.

Protein	Soret maximum (nm)	Prominent bands (nm)	Spin State	Proposed Complex
*Sc*ChdC	407	535, 580	6cHS	heme-ChdC
T182C	403	510, 540, 630	6cHS/6cLS	heme-ChdC
H128Y	406	490, 535, 580, 630	6cHS/6cLS	Heme-ChdC; free heme
H168C	404	510, 540, 630	6cHS/6cLS	heme-ChdC
L208F	404	490, 530, 565, 615	6cHS/5cHS	Heme-ChdC; free heme

G169V	407	495, 540, 580, 630	6cHS/6cLS	Heme-ChdC
G169D	407	492, 535, 570, 630	6cHS/6cLS	heme-ChdC
G169L	407	490, 530, 565, 620	6cHS/5cHS	Heme-ChdC; free heme

Coproheme Decarboxylase Activity

To assess the coproheme decarboxylase activity, turnover experiments were performed and analyzed via UV-Visible spectroscopy and HPLC. Equimolar amounts of protein and substrate were mixed followed by addition of 3 equivalents H_2O_2 to initiate the reaction. UV-Visible spectral data was collected for each protein over 15 minutes. Turnover assays for HPLC analysis were allowed to react for 5 min, before termination with concentrated HCl. Porphyrins were extracted using ethyl acetate, resolubilized in TEA and applied to a C18 reverse phase column, allowing for the visualization of the relative amounts of coproheme, MMD, or product present after each 5-minute turnover reaction. Spectral data from turnover assays are summarized in Table 2.7.

Table 2.7

Summary of absorption peaks, spin state assignments, and proposed porphyrin-protein complex resulting from the mixing of wt and variant *Sc*ChdC proteins with coproheme and H_2O_2.

Protein	Soret Maximum (nm)	Prominent bands (nm)	Spin State	Proposed Complex
*Sc*ChdC	405	535, 580, 630	5cHS/6cHS	MMD-ChdC; heme *b*-ChdC
T182C	396	495, 530, 610	5cQS/5cHS/ 6cHS	Coproheme-ChdC; MMD-ChdC; heme *b*-ChdC
H128Y	393	495, 535, 630	5cQS/5cHS	Coproheme-ChdC; MMD-ChdC
H168C	398	510, 540, 580	5cQS/5cHS/ 6cHS	Coproheme-ChdC; MMD-ChdC; heme *b*-ChdC
L208F	397	490, 535, 578, 620	5cQS/5cHS/ 6cHS	Coproheme-ChdC; MMD-ChdC; heme *b*-ChdC

G169V	399	490, 535, 580, 625	5cQS/5cHS/ 6cHS	Coproheme-ChdC; MMD-ChdC; heme *b*-ChdC
G169D	393	495, 540, 585	5cQS/5cHS/ 6cHS	Coproheme-ChdC; MMD-ChdC; heme *b*-ChdC
G169L	394	493, 530, 580, 615	5cQS/5cHS/ 6cHS	Coproheme-ChdC; MMD-ChdC; heme *b*-ChdC

The wt enzyme showed coproheme decarboxylase activity similar to that of other wt actinobacterial ChdC and 6cHS systems, with a Soret maxima of 405 nm, Q-bands at 535 and 580 nm, and a CT band at 630 nm (Figure 2.12; Table 2.7). A key spectral feature indicative of coproheme turnover to heme *b* is a Soret shift from 390 nm to 405-407 nm (Sebastiani et al., 2021). Additionally, the peak near 580 nm was recently determined to be an indicator for the presence of 5cHS MMD (Sebastiani et al., 2022). The H128Y variant was the only variant that did not reveal any evidence for production of heme *b*. The spectral data following turnover for the wt *Sc*ChdC has been previously published, the only difference being an additional CT band at 630 nm, suggesting the presence of some 5cQS species in addition to the 6cHS product bound ChdC (Smulevich et al., 2005). The Soret maxima at the end of the 15-minute reaction for each of the variants ranged from 393 nm to 399 nm, indicating none of the variant ChdCs accomplished a complete turnover within the reaction time (Figure 2.12; Table 2.7).

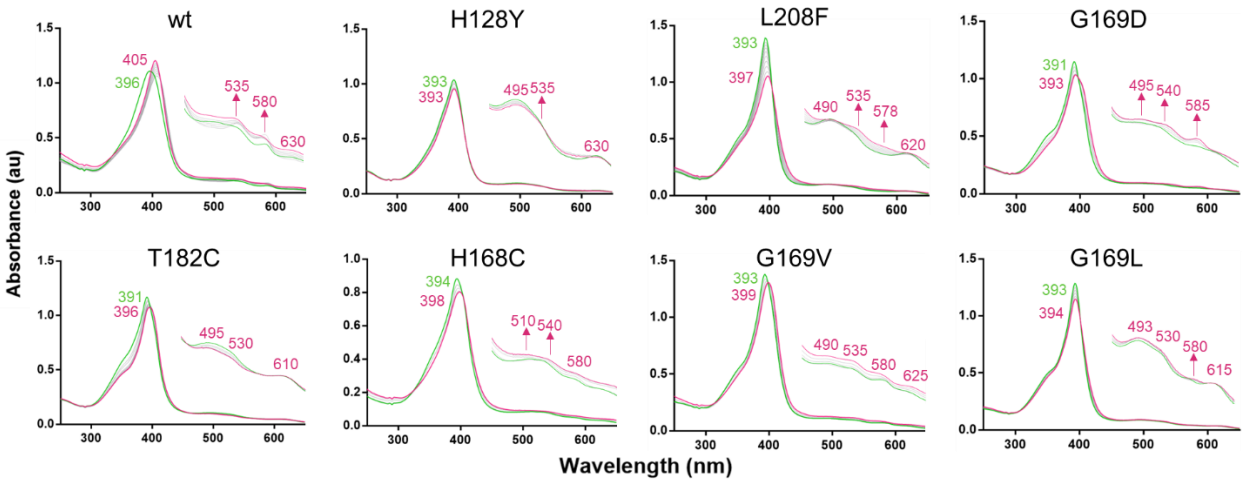


Figure 2.12

UV-Visible spectra following coproheme decarboxylase reactivity of wt and variant *Sc*ChdCs. Each spectrum follows the spectral transition upon mixing 3 equivalents H_2O_2 with 10 µM coproheme-*Sc*ChdC complex. The initial scan, taken after 9s of mixing, is shown in green. After 900s of mixing, the resulting porphyrin-*Sc*ChdC complex is represented in pink. Intermediate spectral scans are shown in gray.

The H128Y and T182C variants have Soret maxima at 393 nm and 396 nm, respectively, and both variants lack a defined peak at 580 nm (Figure 2.12; Table 2.7), indicative of no turnover. This, along with HPLC data, suggest that the T182C variant carries out partial turnover of substrate, yielding small amounts of MMD and heme *b* (Figure 2.12, 2.13). H128Y only accumulates MMD, as HPLC and UV-Visible absorbance data do not show any heme *b* formation (Figure 2.12, 2.13). Sebastiani et al. produced and analyzed a similar variant, H118F (numbering for *Cd*ChdC), where the active site loop histidine was substituted for a phenylalanine in place of a tyrosine. Mass spectroscopy and UV-Visible analysis similarly indicated an accumulation of MMD and no detection of heme *b* (Sebastiani et al., 2021).

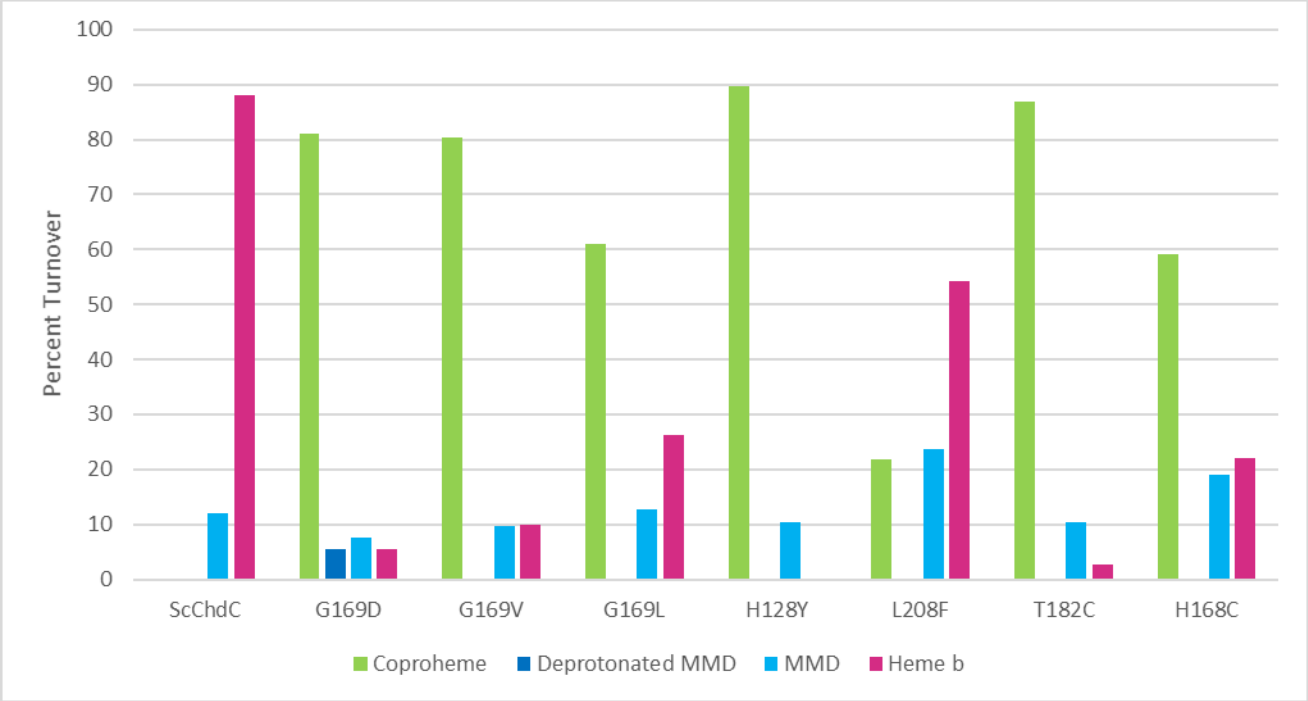


Figure 2.13

HPLC analysis following a five-minute reaction between ChdC (10 μM), coproheme (10 μM) and H_2O_2 (30 μM). The data is summarized by the ratio of heme species (coproheme (green), deprotonated MMD (dark blue), MMD (light blue), and heme *b* (pink)) remaining upon terminating the reaction.

Spectral and HPLC data for G169D and G169V indicate partial substrate turnover, with the detection of small amounts of MMD and heme *b*. G169D and G169V display Soret maxima at 393 and 394 nm, respectively, indicating low substrate turnover to either intermediate or product. Both variants have Q-band features near 490, 530, and 580 nm, suggesting the presence of the 5cQS substrate, 6cHS product, and 5cHS intermediate, respectively. G169D and G169V also have a CT band near 630 nm, suggesting the presence of the 5cQS substrate. The HPLC traces for G169D and G169V also indicate poor substrate turnover, where, the 5-minute reaction of coproheme, 3 equivalents H_2O_2, and either G169V or G169D, show coproheme and

small traces of MMD and heme *b*. Interestingly, the G169D variant displays a split peak around the expected elution time for the intermediate, MMD. We hypothesize that the peak splitting seen for this variant is most likely due to MMD being present in two protonation states (Sullivan et al., 2015).

Post turnover HPLC and spectral analysis of H168C, L208F, and G169L indicate that these variants generate MMD and heme *b* following the addition of H_2O_2 (Figure 2.12, 2.13). This is somewhat surprising considering the critical roles assigned to each amino acid. Regardless, compared to the wt, H168C, L208F, and G169L each have blue shifted Soret maxima at 398, 397, and 394 nm, respectively, aligning with that expected for partial turnover to MMD and heme *b*. These three variants also have a peak at 580 nm, indicative of 5cHS MMD (Sebastiani et al., 2022). H168C has additional Q-band features at 510 and 540 nm, reflecting spectral features seen for H168C heme binding (Figure 2.12, 2.11). HPLC traces of H168C turnover activity further confirm presence of MMD and heme *b* (Figure 2.13). In addition to the peak at 580 nm, L208F and G169L have Q-bands around 490 and 530 nm and a blue-shifted CT band near 615 nm (Figure 2.12; Table 2.7). The bands near 490, 535, and 580 nm again suggest the presence of the 5cQS coproheme substrate, the 6cHS heme *b* product, and the 5cHS MMD intermediate, respectively. Blue-shifted CT bands for L208F and G169L could be indicative of accumulation of the MMD-ChdC complex, or free hemin in solution (Streit et al., 2017; Neugebauer et al., 2012). Overall, H168C, L208F, and G169L displayed the highest substrate turnover of the seven variants.

2.4 Discussion and Conclusion

The crystal structure of MMD-bound wt *Sc*ChdC enabled us to rationalize functions of amino acids within the active site. Several point mutations were constructed and used to address the ChdC mechanism. However, many questions remain including if oxidative decarboxylation of substrate or rotation of the MMD intermediate is the rate limiting step. Regardless, characterization of each point mutation, coupled with structural data, provides important mechanistic information.

At the present time, only one of the variants, T182C, produced diffraction-quality crystals (3.8Å). Despite low resolution diffraction, electron density revealed two key changes in the active site architecture when compared to the MMD bound wt *Sc*ChdC structure. First, the T182C variant captured a pre-turnover, substrate bound state (Figure 2.8, 2.9). In this state, the porphyrin is oriented differently in relation to the active site residue R149 when compared with the MMD bound structure (Figure 2.8, 2.9). Our structural data clearly show the T182C variant binds coproheme. Specifically, the threonine to cysteine substitution resulted in a six-coordinate iron atom and which blocks H_2O_2 from binding. Interactions between the cysteine and the iron trap the substrate in its binding position, preventing turnover. This conclusion is further supported by HPLC traces and electronic absorption data of biochemical binding and turnover assays (Figure 2.12, 2.13). While the enzyme is inactive, the structure provides an opportunity to see the substrate-bound conformation and how positioning of all four propionates influences active site structure prior to turnover. In the coproheme bound T182C structure, R149 participates in close interactions with p4, suggesting that R149 assists in orienting substrate in the active site after binding. Comparing with the MMD

bound wt *Sc*ChdC structure, where substrate has undergone the first decarboxylation and 90-degree rotation, R149 sits closer to p4 to its new, post-rotation position. This new structural observation provides evidence for an additional role for R149 in facilitating rotation and re-orientation of MMD in the active site.

Furthermore, comparison between coproheme bound structures of actinobacterial *Sc*ChdC and firmicute ChdC from *Geobacillus stearothermophilus* (*Gs*ChdC) further supports the importance of R149 to substrate orientation. In the T182C *Sc*ChdC structure, p2 of coproheme is coordinated by R218 and Y145, and p4 is coordinated by R149 and W153. The coproheme bound *Gs*ChdC structure (PDB: 5T2K; Milazzo et al., 2018) revealed a catalytically relevant lysine residue (K148), in the place of R149, that interacts with p4 similarly. In strong contrast to actinobacterial, p2 of the coproheme bound *Gs*ChdC is coordinated by Y144, S222, and M146. While S222 in *Gs*ChdC and R218 in *Sc*ChdC are both located near p2, S222 sits much closer to p2 (2.6Å) compared to R218 (3.0Å). M146 of *Gs*ChdC corresponds to a nonpolar phenylalanine residue in *Sc*ChdC (F147) and has been shown to interact with p2 of coproheme (Milazzo et al., 2018). The lack of this active site methionine residue in actinobacteria, and the differences in the interactions of R149 in the coproheme and MMD bound *Sc*ChdC structures suggests R149 fulfills the roles of M146 and K148 found in firmicutes. In the substrate bound state, R149 coordinates with p4, similar to K148 in firmicutes. However, in the post-rotation MMD bound state, R149 is closer to p4 in its post rotation position (4.2Å) compared to p2 in original binding position (5.2Å). This observation provides an explanation for how actinobacteria reorient the MMD intermediate without the conserved active site methionine.

Spectral and HPLC data for the H128Y variant is in agreement with recently published data for the variant H118F *Cd*ChdC (Figure 2.12 & 2.13; Sebastiani et al., 2021). While the tyrosine and phenylalanine substitutions from both studies can deprotonate H_2O_2 prior to the first decarboxylation, the results are consistent with the larger side chain inhibiting intermediate rotation within the active site, prohibiting the second oxidative decarboxylation. A reasonable conclusion is that H128 is required for the deprotonation of the oxidant and, a resulting change in active site electrostatics, results in the rotation of the intermediate. In addition, the newly generated vinyl group must move through a hydrophobic channel including L208. Interestingly, the leucine to phenylalanine (L208F) variant did not hinder rotation of the intermediate as intended. We predict the hydrophobicity of the phenylalanine still allowed the newly formed vinyl group to slide through the hydrophobic channel. In fact, the L208F results in the highest accumulation of heme *b* among the variants, second only to the wt (Figure 2.13). We now propose placing an amino acid that is aliphatic and more hydrophilic in this position may prevent rotation of MMD.

The HPLC traces of G169D indicate the presence of two protonation states for MMD (Figure 2.13). We have considered two possible explanations for the peak splitting seen in the HPLC traces of G169D ChdC. The first is that D169 may act as a general base, resulting in an additional deprotonation event, and trapping of the intermediate in the pre-rotation position. An alternative hypothesis is that D169 could lead to a partial deprotonation of p6 and p7 of the tetrapyrrole, resulting in a low barrier hydrogen bond between the D169 side chain and the propionate group. The pKa for each carboxylate group would be similar, therefore, either propionate would be favorable. The fact that

there is some turnover, suggests that the rotation still takes place, however, rotation could be significantly slower. The application of stopped-flow spectroscopy will allow us to address this directly. Following the rotation, interaction between D169 and the tetrapyrrole is not as likely because of the new position of the propionate groups. Another possibility is that electrostatic repulsion from D169 forces the substrate to bind in two different ways. A portion of the substrate binds normally, and the enzyme performs the two decarboxylations to form heme *b* and the remaining substrate binds in the post 90-degree rotation pose, causing p4 to be decarboxylated first instead of p2. This would result in an isomeric form of MMD unable to rotate within the active site due to steric restraints caused by p7 and R149.

Spectral scans and HPLC traces of H168C, G169L, and G169V turnover assays revealed that coproheme binds and turns over to heme *b,* though at a much slower rate than the wt ChdC (Figure 2.12 & 2.13). Electronic absorbance data for the H168C variant reflect features typical of a cysteine coordinated heme group (Figure 2.12; Vetter et al., 2009). Further analysis of the H168C variant using the pyridine hemochromagen assay and kinetic analysis may shed light on whether the histidine to cysteine mutation results in a tighter coordination of heme to the active site, resulting in a product bound form of the protein. Additional spectroscopic analysis using resonance Raman and EPR may also provide insight into the coordination environment. Likewise, a similar analysis of G169L and G169V may indicate whether L169 or V169 is participating in side chain interactions that inhibit substrate binding, intermediate rotation, or product release from the active site as the degree of porphyrin “ruffling” would be affected by the larger side chains.

The research presented in this work provide a more detailed understanding of and evidence for the mechanism of rotation of the intermediate within the active site of actinobacterial ChdC. We propose that once the first vinyl group is formed, repulsive forces caused by nearby polar residues propel the vinyl group through a hydrophobic channel composed of F197, M209, L208, and V178. At the same time, the side chain of R149 guides p4 into position where it will be coordinated by Y145 and R218 prior to the second oxidative decarboxylation. Following decarboxylation/oxidation of the second propionate group, we propose the same repulsive forces that initiate the intermediate rotation also push the second vinyl group through the hydrophobic channel, guiding the heme group out of the active site. Both the rotation, and the ability of intermediate or product to exit the active site is influenced by the flexibility of the active site loop (Residues R120-P130). A multi-faceted approach including additional variants and kinetic analysis using stopped-flow spectroscopy will help us to further describe the atomic details of the ChdC mechanism. Combined with a steadily growing base of knowledge, our structure/function-based approach to designing new inhibitors may be utilized to develop antimicrobial compounds.

2.5 References

1. Dailey HA, Gerdes S. **2015**. HemQ: An iron-coproporphyrin oxidative decarboxylase for protoheme synthesis in firmicutes and actinobacteria. *Arch Biochem Biophys.* 574: 27-35.

2. Bali S, Lawrence AD, Lobo SA, Saraiva LM, Golding BT, Palmer DJ, Howard MJ, Ferguson SJ, Warren MJ. **2011**. Molecular hijacking of siroheme for the synthesis of heme and d_1 heme. *PNAS*. 108(45): 18260-18265.

3. Michlits H, Lier B, Pfanzagl V, Djinović-Carugo K, Furtmüller PG, Oostenbrink C, Obinger C, Hofbauer S. **2020**. Actinobacterial coproheme decarboxylases use histidine as a distal base to promote compound I formation. *ACS Catalysis*. 10: 5405-5418.

4. Celis AI, Guass GH, Streit BR, Shisler K, Moraski GC, Rodgers KR, Lukat-Rodgers GS, Peters JW, DuBois JL. **2017**. Structure-based mechanism for oxidative decarboxylation reactions mediated by amino acids and heme propionates in coproheme decarboxylases (HemQ). *JACS*. 139: 1900-1911.

5. Dailey HA, Gerdes S, Dailey TA, Burch JS, Phillips JD. **2015**. Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. *PNAS*. 112(7): 2210-2215.

6. Streit BR, Celis AI, Moraski GC, Shisler KA, Shepard EM, Rodgers KR, Lukat-Rodgers GS, Dubois JL. **2018**. Decarboxylation involving a ferryl, propionate, and a tyrosyl group in a radical relay yields heme b. *J Biol Chem*. 293(11): 3989-3999.

7. Milazzo L, Gabler T, Pühringer, Jandova Z, Maresch D, Michlits H, Pfanzagl V, Djinović-Carugo K, Oostenbrink C, Furtmüller PG, Obinger C, Smulevich G, Hofbauer S. **2019**. Redox cofactor rotates during its stepwise decarboxylation: molecular mechanism of conversion of coproheme to heme b. *ACS Catalysis*. 9: 6766-6782.

8. Sebastiani F, Michlits H, Lier B, Becucci M, Furtmüller PG, Oostenbrink, Obinger C, Hofbauer S, Smulevich G. **2021**. Reaction intermediate rotation during the decarboxylation of coproheme to heme b in c. diphtheriae. *Biophys J.* 120: 3600-3614.

9. Hofbauer S, Hagmüller A, Schaffner I, Mlynek G, Krutzler M, Stadlmayr G, Pirker KF, Obinger C, Daims H, Djinović-Carugo K, Furtmüller PG. **2015**. Structure and heme-binding properties of HemQ (chlorite dismutase-like protein) from listeria monocytogenes. *Arch Biochem Biophys.* 574: 36-48.

10. Dailey TA, Boynton TO, Albetel A, Gerdes S, Johnson MK, Dailey HA. **2010**. Discovery and characterization of HemQ: An essential heme biosynthetic pathway component. *J Bio Chem.* 285(34): 25978-25986.

11. Emsley P, Lohkamp B, Scott WG, Cowtan K. **2010**. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 66: 486–501.

12. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. **2018**. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol*. 74: 531–544.

13. Pfanzagl V, Holcik L, Maresch D, Gorgone G, Michlits H, Furtmüller PG, Hofbauer S. **2018**. Coproheme decarboxylases – phylogenetic prediction versus biological experiments. *Arch Biochem Biophys.* 640: 27-36.

14. Milazzo L, Hofbauer S, Howes BD, Gabler T, Furtmüller PG, Obinger C, Smulevich G. **2018**. Insights into the active site of coproheme decarboxylase from *listeria monocytogenes. Biochemistry.* 57: 2044-2057.

15. Smulevich G, Feis A, Howes BD. **2005**. Fifteen years of raman spectroscopy of engineered heme containing peroxidases: what have we learned? *Acc Chem Res.* 38: 433-440.

16. Sebastiani F, Risorti R, Niccoli C, Michlits H, Becucci M, Hofbauer, Smulevich G. **2022**. An active site at work—the role of key residues in *C. diphteriae* coproheme decarboxylase. *J Inorg Biochem*. 229: 111718.

17. Hosseinzadeh P and Lu Y. **2016**. Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics. *Biochim Biophys Acta*. 1857: 557-581.

18. Vetter SW, Terentis AC, Osborne RL, Dawson JH, Goodin DB. **2009**. Replacement of the axial histidine heme ligand with cysteine in nitrophorin I: spectroscopic and crystallographic characterization. *J Biol Inorg Chem.* 14(2): 179-191.

19. Neugebauer U, März A, Henkel T, Schmitt M, Popp J. **2012**. Spectroscopic detection and quantification of heme and heme degradation products. *Anal Bioanal Chem*. 404: 2819-2829.

20. Hofbauer S, Mlynek G, Milazzo L, Pühringer D, Maresch D, Schaffner I, Furtmüller PG, Smulevich G, Djinović-Carugo K, Obinger C. **2016**. Hydrogen peroxide-mediated conversion of coproheme to heme b by HemQ—lessons from the first crystal structure and kinetic studies. *The FEBS Journal.* 283: 4386-4401.

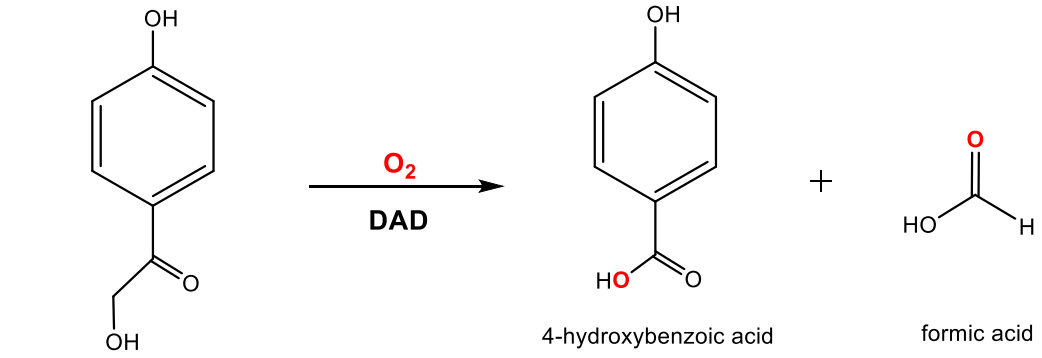
21. Sullivan SA, Streit BR, Ferguson EL, Jean PA, McNett DA, Llames LT, DuBois JL. **2015**. Mass-spectrometric profiling of porphyrins in complex biological samples with fundamental, toxicological, and pharmacological applications. *Analy Biochem.* 478: 82-89.

APPENDIX A

Crystallization of 2,4'-dihydroxyacetophenone dioxygenase

A.1 Introduction

The nonheme, iron-dependent enzyme 2,4’-dihydroxyacetophenone dioxygenase (DAD) is a homotetrameric protein made up of 20.3 kDa monomers (Hopper and Kaderbhai, 1999). DAD utilizes molecular oxygen to catalyze the breakdown of 2,4’-dihydroxyacetophenone (DHA) to 4-hydroxybenzoic acid and formic acid (Scheme A.1). The substrate, DHA, is a common intermediate aromatic compounds formed during the degradation of plant products and manmade compounds released by the chemical industry. Due to its ability to aerobically degrade aromatic compounds, DAD has become a target for removal of aromatic pollutants from groundwater and soil.



2,4'-dihydroxyacetophenone

4-hydroxybenzoic acid

formic acid

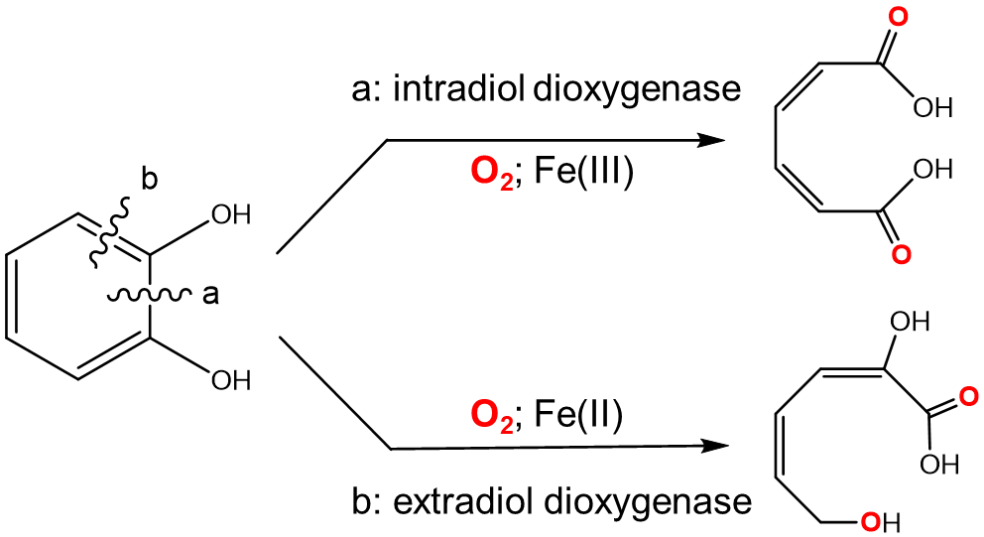
Scheme A.1

The degradation of 2,4’-dihydroxyacetophenone to 4-hydroxybenzoic acid and formic acid catalyzed by DAD in the presence of O_2.

In general, oxygenases catalyze the incorporation of molecular oxygen into a wide range of organic substrates. Early mechanistic studies investigating the products,

4-hydroxybenzoic acid and formic acid, using isotope labeling where DAD was allowed to react in the presence of either $H_2{}^{18}O$ or ${}^{18}O_2$ revealed that both products contained ${}^{18}O$ only when in the presence of ${}^{18}O_2$, confirming the designation of DAD as a dioxygenase (Hopper, 1986). Unlike the majority of characterized dioxygenases, which cleave the aromatic ring itself, DAD has been shown to cleave the C—C bond of the aliphatic substituent.

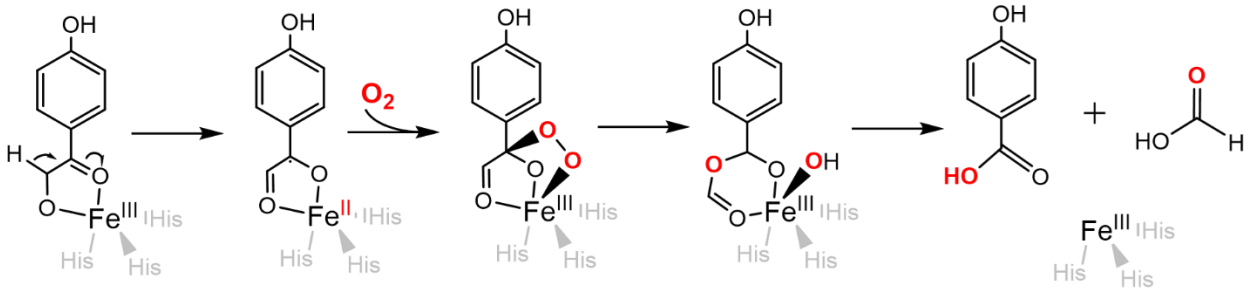
Catechol dioxygenases are another class of iron-dependent enzymes that utilize molecular oxygen to catalyze the cleavage of aromatic rings (Broderick, 1999). Ring cleavage performed by catechol dioxygenases occurs at two different locations within the aromatic ring. Catechol dioxygenases are classified based on their cleavage of the aromatic ring, where intradiol catalyzes ortho-cleavage and extradiol dioxygenases cleaves at the meta-position (Scheme A.2). Additionally, the active site metal center of intradiol and extradiol dioxygenases are Fe(III) and Fe(II), respectively, resulting in the different oxygenation chemistry between these two classes (Broderick, 1999). Two possible mechanisms for the aliphatic C—C bond cleavage catalyzed by DAD.



Scheme A.2

Ring cleavage mechanisms of catechol dioxygenases. Intradiol dioxygenases utilize O_2 and ferric iron to catalyze the ortho-cleavage of the aromatic ring, while extradiol dioxygenases cleave at the meta-position in the presence of O_2 and ferrous iron.

The first mechanism, derived from intradiol dioxygenases, was based on the observation that DAD is active in the presence of O_2 (Enya et al., 2012). This suggests that ferric iron drives catalysis rather than ferrous iron, which readily oxidizes under aerobic conditions. The mechanism is initiated with the coordination of the substrate, DHA, to the active site ferric iron, followed by the loss of a proton and formation of an intermediate reminiscent of the intradiol dioxygenases substrate, catecholic 1,2-diol. Subsequent delocalization of an electron from the substrate leads to formation of a ferrous iron benzylic radical intermediate. Molecular oxygen binds to the reduced iron and an oxygen atom is inserted into the intermediate, leading to cleavage of the aliphatic C—C bond.

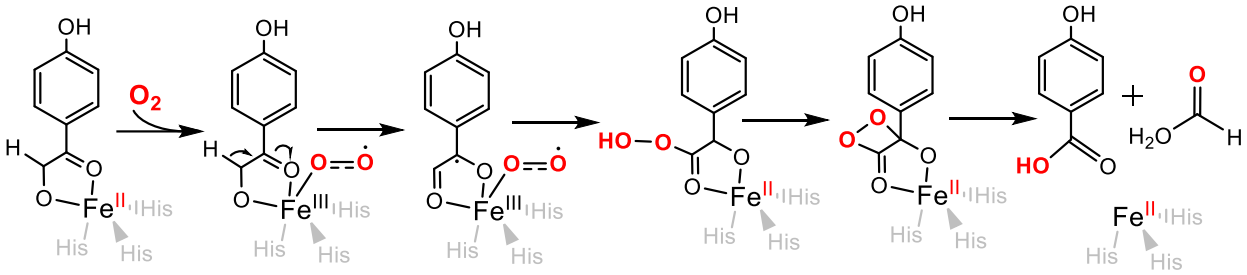


Scheme A.3

Proposed mechanism of DAD derived from intradiol dioxygenases, where the oxidative cleavage of DHA occurs in the presence of O_2 and ferric iron.

The second mechanism is based on the observation that both extradiol dioxygenases and *a*DAD produce colorless crystals, while intradiol dioxygenases

produce red crystals. The extradiol dioxygenase proposed mechanism utilizes a ferrous iron, and structural and spectroscopic data collected by Paria et al. (2012) and Keegan et al. (2014) led to an alternative DAD mechanism. Together, their mechanism begins with the binding of molecular oxygen to the nonheme iron and subsequent formation of an iron(III)-superoxide species. This radical superoxide abstracts a hydrogen atom from the hydroxyketone group, producing an iron(III)-hydroperoxide intermediate. The substrate is then hydroperoxylated via a radical rebound mechanism and the C—C bond of the resulting dioxacyclic intermediate is cleaved, yielding the reaction products, formic acid and 4-hydroxybenzoic acid.



Scheme A.4

Proposed mechanism of DAD derived from extradiol dioxygenases, where the oxidative cleavage of DHA occurs in the presence of O_2 and ferrous iron.

A recent study by Roberts et al. (2020) attempted to elucidate the mechanism of DAD by analyzing the metal dependence of the enzyme. Apo-DAD was produced and reconstituted with various divalent metal ion and the substrate, DHAP. Enzyme activity was tracked using UV-Visible spectroscopy and monitoring the decrease at 277 nm, an indicator for substrate consumption. They found most of the divalent metals had no effect on enzyme activity. In another experiment using UV-Visible spectroscopy which analyzed enzyme activity of varying ratios of apoDAD and Fe^{2+} anaerobically and

aerobically, under anaerobic conditions, ferrous iron increased enzyme activity by about 10%. After exposing each reaction to air, they saw a significant increase in enzyme activity, with a linear dependence on iron concentration. Additionally, they identified a small absorbance feature at 360 nm as a spectroscopic marker for the presence of oxidized Fe^{3+}. This peak at 360 nm is not seen with the anaerobically assayed apoDAD reconstituted with Fe^{2+} or the apoDAD. This newly identified spectral marker and the lack of DAD activity seen with reconstitution of Fe^{2+}under anaerobic conditions provide support for a mechanism that relies on ferric iron (Roberts et al., 2020).

To date, only two structures of DAD have been published. Keegan et al. published a structure of DAD from Alcaligenes sp. 4HAP (*a*DAD) in 2014 (PDB: 4P9G) and a year later, Guo et al. published another structure of DAD from the same species (PDB: 5BPX). Both structures show a conserved β-barrel fold, classifying DAD structurally as part of the cupin superfamily. This diverse superfamily consists of a wide array of metal binding proteins which utilize the β-barrel fold to facilitate a multitude of functions in nature. Iron-dependent dioxygenases are one of the largest subsets of this structure family (Dunwell et al., 2004). Structural data from *a*DAD show that the β-barrel is made up of two sets of six-stranded β-sheets and three α-helices, one at the N-terminus and two which form a helical hairpin at the C-terminus (Figure A.1). Typically, two sets of four-stranded β-sheets is the norm for the cupin superfamily, however, *a*DAD has an additional two β-sheets which makes up each side of the β-barrel. The proximity of the four additional β-sheets to one another is close enough that one might expect the opposing sheets to interact to close one end of the barrel. Unexpectantly, these β-sheets flare outward from the barrel, exposing a patch of hydrophobic residues

that are believed to play a role in substrate access to the buried active site (Figure A.1; Keegan et al., 2014).

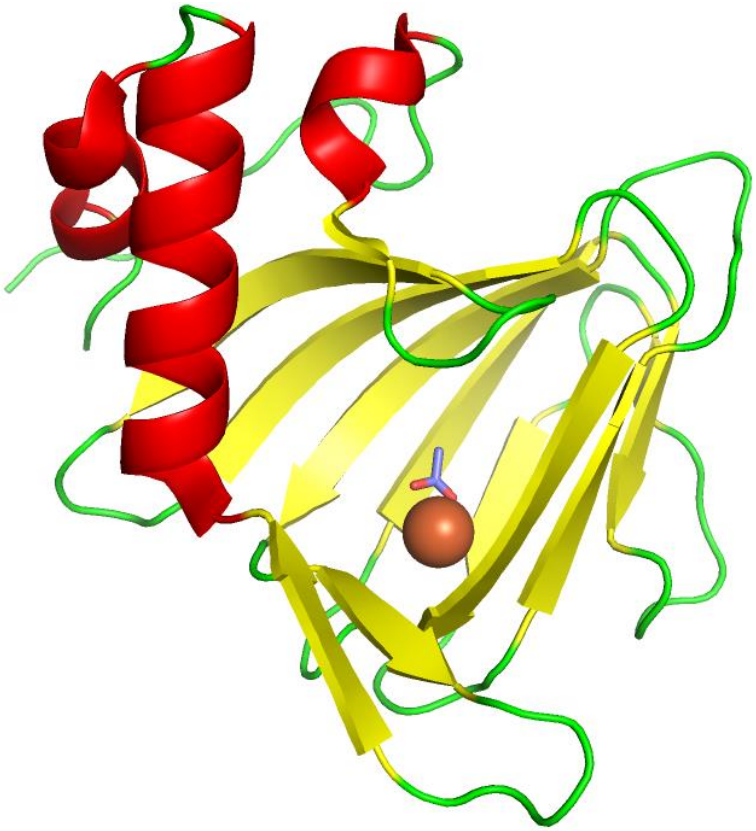


Figure A.1

Cartoon representation of a single monomer of *a*DAD (PDB:5BPX) colored by secondary structures (α-helix (red); β-sheet (yellow); loop (green)). The active site iron is shown a red sphere and an acetate ligand is shown as dark blue sticks.

The active site is located in the hollow center of the β-barrel and consists of mainly hydrophobic residues largely conserved among the DAD family (Figure A.1; Keegan et al., 2014). At the active site, the catalytic iron is coordinated by three conserved histidine residues, H78, H76, and H114 (numbering for *a*DAD), and an anionic group (Figure A.2). Initially, this anionic group was thought to be a carbonate dianion; more recent and higher resolution crystallographic data indicated the ligand is an acetate anion (Guo et al., 2015). The substrate, 2,4'-DHAP, is suspected to occupy

the location of an acetate ion and a glycerol molecule modeled into the crystal structure active site (Keegan et al., 2014).

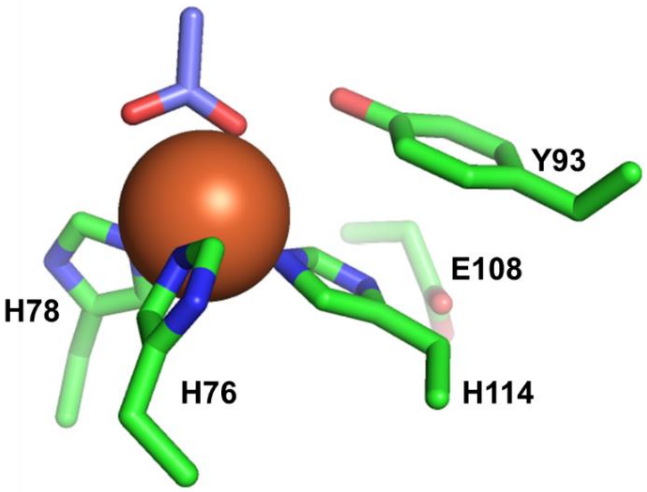


Figure A.2

Structural model of the *a*DAD active site depicting the coordination of the catalytic iron (PDB: 5BPX). The active site iron is shown as a red sphere and the acetate ligand and important active site residues are shown as dark blue and green sticks, respectively.

Keegan et al. (2014) modeled the substrate, 2,4'-DHAP, into the active site, so the α-hydroxyketone group sits near the catalytic iron. Placement of the substrate in the active site revealed hydrogen bond interactions between a conserved tyrosine residue, Tyr93, and the α-hydroxy group of the substrate (Figure A.3A). It was suggested that the hydroxy group of Tyr93 played a key mechanistic role by stabilizing reaction intermediates (Keegan et al., 2014). Initial structural studies also revealed that another active site residue, Glu108, may play a role in shuttling oxygen into the active site (Guo et al., 2015). In addition, more recent molecular dynamic (MD) simulation studies investigated a possible role for an active site water molecule (Manna et al., 2018). Theoretical calculations from this study supported the hypothesis that Tyr93 and Glu108

play an active role in catalysis. It was suggested that a hydrogen bonding network involving the active site water, Tyr93, and Glu108, help to stabilize intermediate complexes (Figure A.3B). Where the water molecule and Tyr93 act as proton carriers, and Glu108 functions as a proton reservoir (Manna et al., 2018). While the recent work from Roberts et al. (2020) provides evidence against a mechanism involving ferrous iron, the structural studies support a catalytic role for Tyr93 and Glu108.

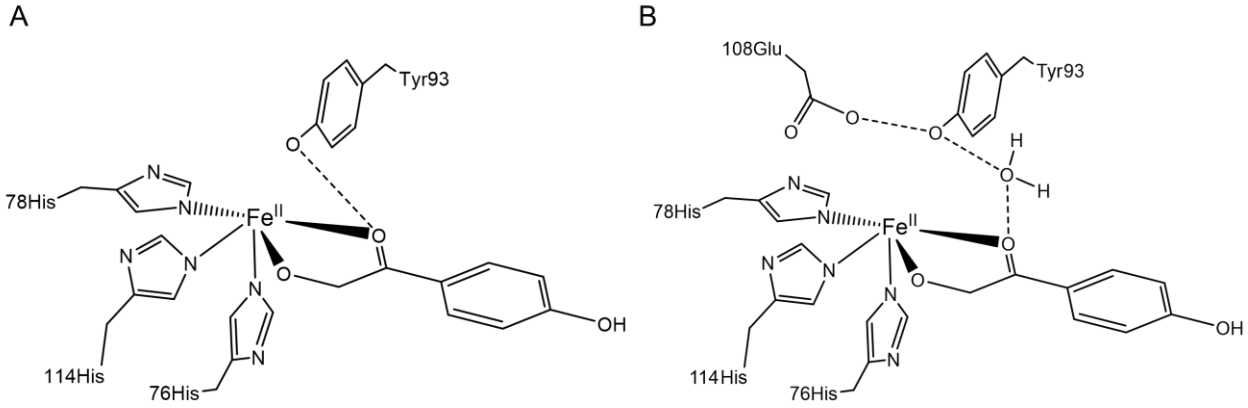


Figure A.3

Active site hydrogen bonding network of *a*DAD based on observations from X-ray crystallography data (A) and MD simulations (B). Straight dashed lines represent hydrogen bonding between molecules.

The Roberts Lab at the University of South Carolina are currently working to further elucidate the DAD using X-ray crystallography. Our collaborators tasked us with the crystallization of 2,4'-dihydroxyacetphenone dioxygenase from *Burkholderia* sp. AZ11 (*b*DAD). Protein expression and purification was performed by Gabrielle C. Connor and Kenneth M. Roberts and *b*DAD was shipped to us for crystallization,

diffraction data collection, and refinement. At the present time, a 3.3Å structure has been produced and efforts to develop higher resolution crystals are ongoing.

A.2 Methods

Protein Expression and Purification

*b*DAD was expressed and purified by Gabrielle C. Connor and Kenneth M. Roberts at the University of South Carolina. A detailed method can be found in Roberts et al., 2020.

Crystallization and Structure Determination

A set of five crystallization screens were used for initial screening to determine optimal crystallization conditions for *b*DAD: PEG/Rx, Crystal Screen, Index, and PEG/Ion by Hampton Research and Wizard Classic by Rigaku. Initial screens were prepared for the apo form of *b*DAD using sitting drop vapor diffusion method, where the drop consisted of a 1:1 mixture of 1 µL crystal screen condition to 30 mg/mL protein. Following the initial high throughput screen, matrix screens of suitable conditions were developed, optimizing pH, salt, additives, or other precipitation reagents. A batch diffusion method using gradient screens was also utilized to improve crystal size and diffraction quality. An in-house diffractometer was used to assess the diffraction quality and determine the space group parameters before sending samples to SER-CAT. Subsequent rounds of model building, and refinements were performed using the programs COOT (Emsley et al., 2010) and PHENIX (Afonine et al., 2018).

A.3 Results

Crystal structure of *b*DAD

Initial screening of *b*DAD, performed by Clayton Pritchett, revealed multiple promising conditions for crystallization. Before proceeding with optimizations, Clayton Pritchett also performed a protein concentration screen and determine the optimal *b*DAD concentration for crystallization to be 30 mg/mL. Sitting drop optimizations identified from the high throughput screen were prepared. Initial crystallization trials of *b*DAD included optimizations with the products resulting from this reaction, 4-hydroxybenzoic acid and formate, however, crystal diffraction was not improved with this addition. Additional optimizations included changes in the pH, salts, additives, or precipitation reagents included in the parent condition. Building on these optimizations, temperature trials, kosmotrope screening, microseeding, and gradient screening using capillaries were also done. Resulting crystals were screened on an in-house diffractometer to confirm they were diffraction quality crystals as well as identify the space group. Diffraction quality crystals were then sent to SER-CAT for data collection, followed by several rounds of data refinement.

Optimization ultimately led to a crystal structure of *b*DAD determined to 3.3 Å resolution obtained through microseeding (Table A.1). This was done by crushing and mixing diffraction quality crystals with the parent conditions (0.2M magnesium acetate pH 8 and 10% PEG 3,350) to make a concentrated seed stock. A serial dilution of the seed stock was performed from 1:10 to $1:10^{20}$ in a sitting drop tray where the drop consisted of a 1:1 volume of protein and the diluted seed stock. The protein solution was prepared by diluting *b*DAD to a concentration of 30mg/mL in 20mM potassium

phosphate pH 8. Colorless crystals grew after ten days and were subsequently frozen for data collection at SER-CAT. Crystals produced in the $1{:}10^{14}$ seed stock solution resulted in the 3.3 Å resolution structure presented in this work (Figure A.4A).

Table A.1

Data Collection and Refinement Statistics of *b*DAD. Statistics for the highest-resolution shell are shown in parentheses.

	***b*DAD**
Wavelength (Å)	1.0
Resolution range (Å)	40.11 - 3.305 (3.423 - 3.305)
Space group	C 2 2 21
Unit cell (Å)	216.144 219.827 214.087 90 90 90
Unique reflections	75354 (6792)
Completeness (%)	98.58 (89.56)
Wilson B-factor (Å2)	87.79
Reflections used in refinement	75202 (6792)
Reflections used for R-free	2012 (183)
R-work	0.1850 (0.2867)
R-free	0.2199 (0.3184)
Number of non-hydrogen atoms	16104
macromolecules	16044
ligands	60
Protein residues	1966
RMS (bonds) (Å)	0.015
RMS (angles) (°)	1.62
Ramachandran favored (%)	87.13

Ramachandran allowed (%)	10.35
Ramachandran outliers (%)	2.52
Rotamer outliers (%)	0.06
Clashscore	16.63
Average B-factor (Å2)	79.08
Macromolecules (Å2)	79.06
Ligands (Å2)	83.37

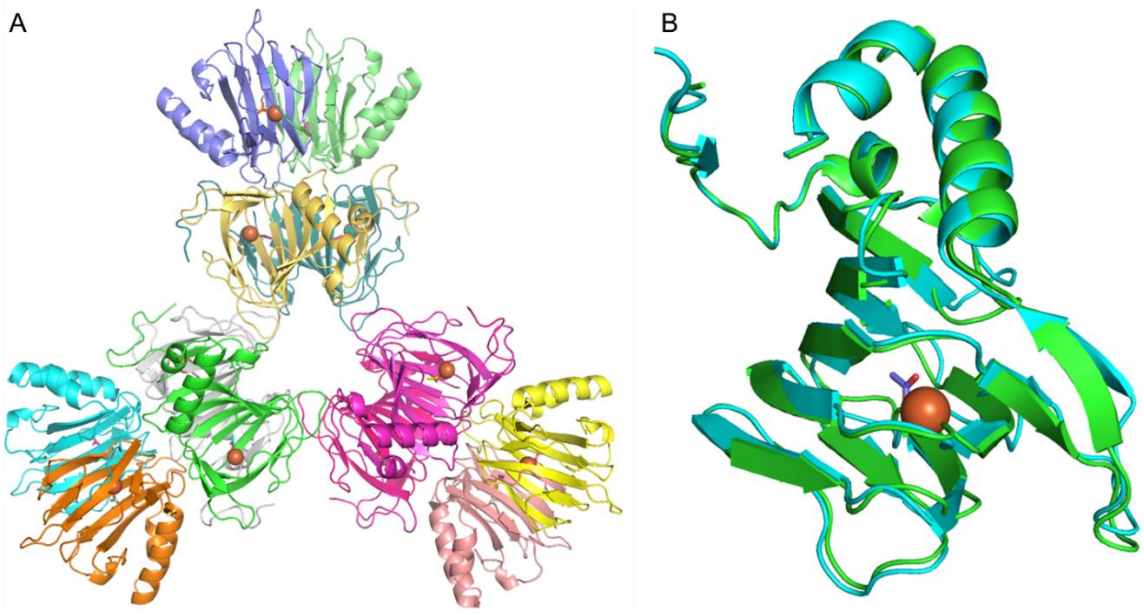


Figure A.4

(A) Cartoon representation of the overall structure of *b*DAD. The asymmetric unit consists of three homotetramers, the monomers are differentiated by color. (B) An overlay of the individual monomers from *a*DAD (PDB: 5BPX) and *b*DAD shown as cartoons in green and cyan, respectively. The active site iron is shown as red spheres and the acetate ligand is shown as dark blue sticks.

The asymmetric unit of *b*DAD contains three homotetramers. Where the homotetramer is the physiologically relevant structure (Figure A.4A). A single monomer

of *b*DAD aligns with the previously published structures of *a*DAD (Figure A.4B). Like the *a*DAD structure, the monomer is made up of a six-stranded β-barrel, a N-terminal α-helix, and a C-terminal helical hairpin comprising two α-helices. Electron density at the active site confirms the presence of the three conserved histidine residues (H78, H76, and H114) that coordinate the catalytic iron, and the active site residues, Y93 and G108 (Figure A.5). Active site density also confirms the presence of a fourth ligand coordinating the iron, and we have putatively modeled an acetate ion at this location (Figure A.5).

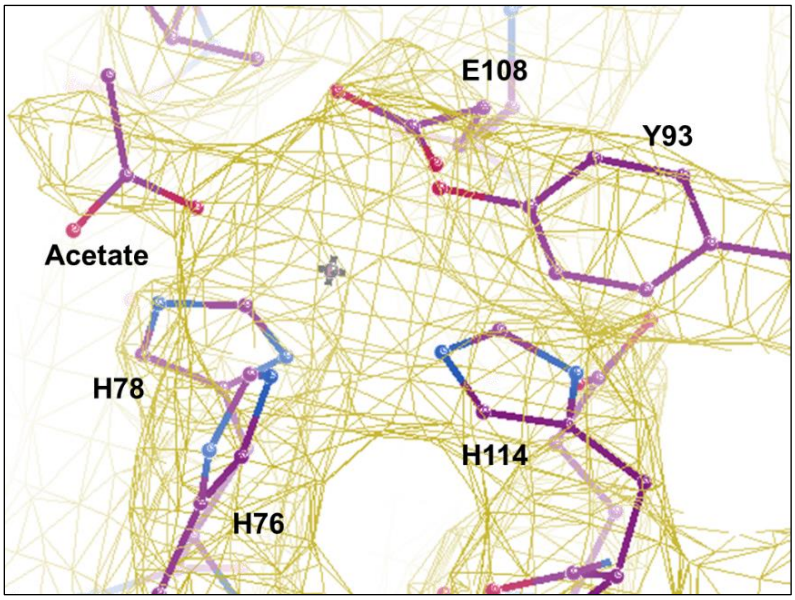


Figure A.5

Ball and stick model of *b*DAD with the $2F_o$-F_c composite omit map (yellow cage). Active site density confirms the presence of and acetate molecule and three histidine residues that coordinate with the iron. Composite omit 2|Fo| − |Fc| electron density maps are shown at σ = 1.5.

An overlay of the active site structures from aDAD and *b*DAD revealed conflicting orientations of the active site residue Glu108 (Figure A.6). In the previously published

structure of aDAD (PDB: 5BPX), Glu108 is pointed away from the iron binding site (Figure A.6A). In contrast, electron density at the active site of *b*DAD suggests a different orientation of Glu108, where the glutamate is pointing towards the catalytic iron, near the substrate binding site (Figure A.6B).

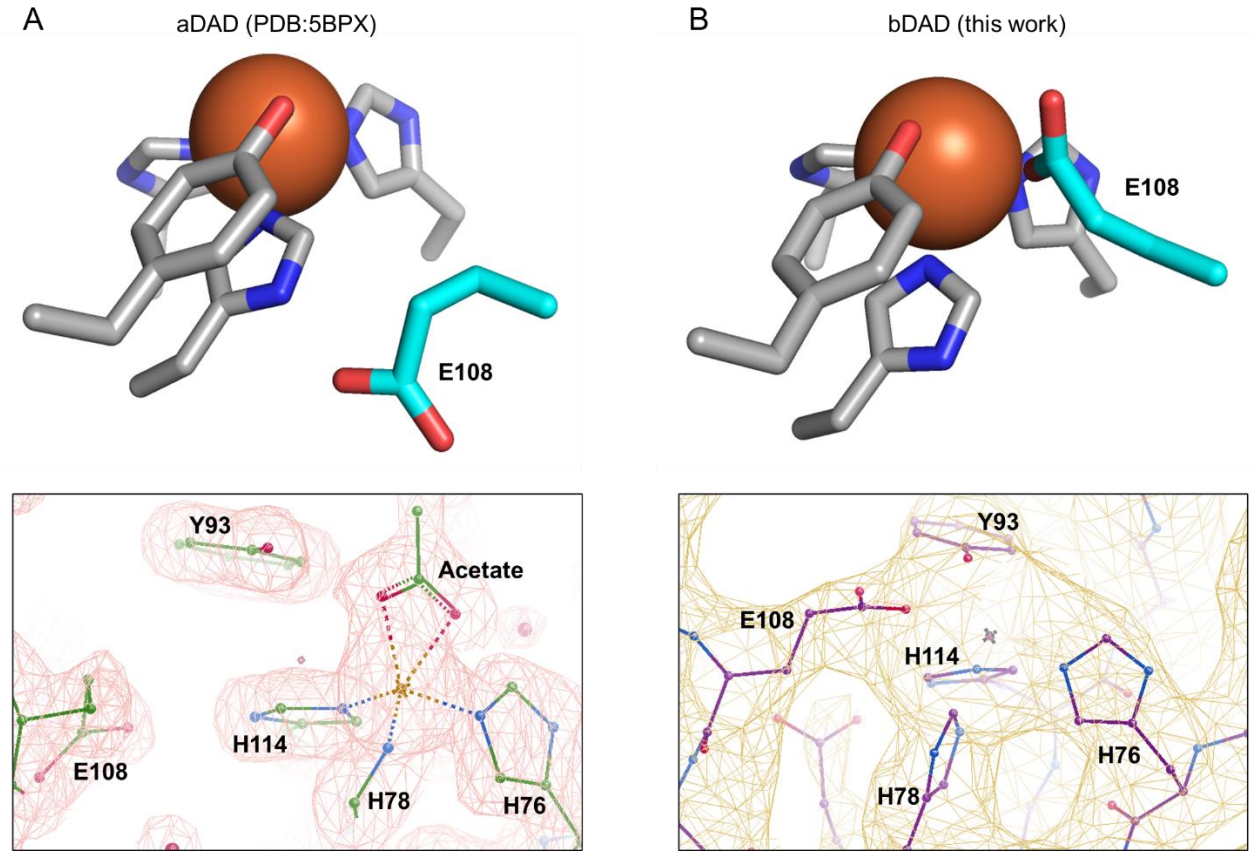


Figure A.6

Orientation of Glu108 in the active sites of aDAD (A) and *b*DAD (B). The top panel of A and B depicts a model of the active site structure where the catalytic iron, Glu108, and other active site residues are shown as red spheres, cyan sticks, and gray sticks, respectively. The bottom panel of A and B is a ball and stick representation of aDAD and *b*DAD with the associated $2F_o$-F_c composite omit map shown as red and yellow mesh, respectively. Composite omit 2|Fo| − |Fc| electron density maps are shown at σ = 1.5.

A.4 Discussion and Conclusion

Presented in this work is the first structure of DAD from *Burkholderia* sp. AZ11. At a resolution of 3.3 Å, we were able to compare our crystal structure to a previously published structure of aDAD (PDB: 5PBX). Most notably, a comparison of active site residues revealed a significant difference in the conformation of Glu108. This glutamate residue has been suggested to play a role shutting oxygen into the active site and has been implicated as part of the active site hydrogen bonding network (8: Guo 2015; 10: Manna 2018). Electron density extending from the iron to Glu108 observed in our structure is an ideal for oxygen binding. Furthermore, Glu108 has been captured at both the protein surface and within the active site, pointing to the flexible nature of this glutamate residue. This observation also provides further support for Glu108 functioning as an oxygen shuttle.

Ongoing and future crystallization experiments include optimizations of the wild type enzyme, with and without substrate and efforts to crystallize active site variants of *b*DAD. One variant of special interest is Y93F/A to gain a more detailed understanding of the role of Tyr93 in catalysis. Other variants of note are W62F and D64N/A, both of which have been implicated in substrate binding.

A.5 References

1. Hopper DJ and Kaderbhai MA. **1999**. 2,4'-dihydroxyacetophenone dioxygenase (EC 1.13.11.41) from *Alcaligenes* sp. 4HAP: a novel enzyme with an atypical dioxygenase sequence. *Biochem J.* 344: 397-402.

2. Hopper DJ. **1986**. Oxygenase properties of the (4-hydroxybenzoyl)methanol-cleavage enzyme from an *Alcaligenes* sp. *Biochem J.* 239: 469-472.

3. Broderick JB. **1999**. Catechol dioxygenases. *Essays in Biochemistry*. 34: 173-189).

4. Enya M, Aoyagi K, Hishikawa Y, Yoshimura A, Mitsukura K, Maruyama K. **2012**. Molecular and catalytic properties of 2,4'-dihydroxyacetophenone dioxygenase from *Burkholderia* sp. AZ11. *Biosci Biotech Biochem*. 76(3): 567-574.

5. Paria S, Halder P, Paine TK. **2012**. Oxidative carbon-carbon bond cleavage of a α-hydroxy ketone by a functional model of 2,4'-dihydroxyacetophenone dioxygenase. *Angew Chem Int Ed.* 51: 6195-6199.

6. Keegan R, Lebedev A, Erskine P, xGuo J, Wood SP, Hopper DJ, Rigby SEJ, Copper JB. **2014**. Structure of the 2,4'-dihydroxyacetophenone dioxygenase from *AlcaligenesI* sp. 4HAP. *Acta Cryst.* D70: 2444-2454.

7. Roberts KM, Connor GC, Cave CH, Rowe GT, Page CA. **2020**. The metal- and substrate-dependencies of 2,4'-dihydroxyacetophenone dioxygenase. *Arch Biochem Biophys.* 691: 108441.

8. Guo J, Erskine P, Coker AR, Gor J, Perkins SJ, Wood SP, Cooper JB. **2015**. Extension of resolution and oligomerization-state studies of 2,4'-dihydroxyacetophenone dioxygenase form *Alcaligenes* sp. 4HAP. *Acta Cryst.* F71: 1258-1263.

9. Dunwell JM, Purvis A, Khuri S. **2004**. Cupins: the most functionally diverse protein superfamily. *Phytochemistry*. 65: 7-17.

10. Manna RN, Malakar T, Jana B, Paul A. **2018**. Unraveling the crucial role of a single active water molecule in the oxidative cleavage of aliphatic C-C bond of 2,4'-dihydrocyacetophenone catalyzed by 2,4'-dihydroxyacetophenone dioxygenase enzyme: A quantum mechanics/molecular mechanics investigation. *ACS Catalysis.* 8: 10043-10050.

11. Emsley P, Lohkamp B, Scott WG, Cowtan K. **2010**. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 66: 486–501.

12. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. **2018**. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struc Bio*. 74: 531–544.