Mechanistic Studies on the Prenylated- Flavin-Dependent Phenazine-1-Carboxylic Acid Decarboxylase

by

Prathamesh Madhav Datar

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2024

Doctoral Committee:

Professor E. Neil G. Marsh, Chair
Assistant Professor Jennifer Bridwell-Rabb
Assistant Professor Markos Koutmos
Associate Professor Bruce A. Palfey

Prathamesh Madhav Datar

mdatar@umich.edu

ORCID iD: 0000-0003-1514-9767

Dedication

This dissertation is dedicated to Prof. Dexter B. Northrop (1940 – 2023) as well as to humanity's relentless pursuit of knowledge by means of the scientific method.

Acknowledgements

A great number of people have been instrumental to my PhD and none of this was possible without their contributions. I would like to start off by thanking my advisor, Prof. E. Neil G. Marsh for his constant support. All his valuable suggestions on experiment design, data analysis and text editing have helped shape this dissertation in its current form. Thank you Prof. Marsh for being patient with me and guiding me on this journey.

I am extremely grateful to Prof. Bruce A. Palfey for entertaining me and providing key insights every time I stormed into his office with seemingly inexplicable data. I will miss our discussions on the philosophy of science, Steenbock symposia and the use of Wolfram Mathematica to write rate equations, including webZyme.

I would like to thank Dr. Markos Koutmos and Dr. Jennifer Bridwell-Rabb for their advice during my candidacy, data meeting and also providing me access to lab instruments/DNA plasmids from time to time. During my rotation project in Dr. Koutmos's lab, I gained firsthand experience in X-ray crystallography and I am grateful to him as well as my mentor, Dr. Kazuhiro Yamada, for the opportunity.

A friendly lab environment is perhaps the most crucial element conducive to constructive research. I am grateful to past and current members of the Marsh lab for making CHEM 4500 my second home. I thank Dr. Kyle Ferguson and Dr. Nattapol Arunrattanamook for laying the foundations of the prFMN project. I am grateful to Dr. April Kaneshiro, Dr. Pronay Roy, Dr. Anushree Mondal, Luis Vazquez-Rivera, Daniel DiRocco and Ryan McGinnis for being such

fantastic collaborators. Dr. Ajitha Christie-David, Dr. Kelsey Diffley, Dr. Karl Koebke, Dr. Marie Hoarau, Dr. Hannah Chia, Dr. Arti Dumbrepatil, Dr. Timothy Grunkemeyer, Dr. Ayesha Patel, Dr. Soumi Ghosh have all mentored me at some point in my PhD and I am indebted to them. Victor Rivera-Santana, Harsha Gouda, Yulduz Rakibova, Jiying Liu, Katherine Hunter, Yue Xin, Srijoni Majhi and Makayla Brunt, thank you for being wonderful friends and lunch/coffee/cycling buddies.

I take this opportunity to thank my departmental collaborators, Prof. Zhan Chen and Dr. Tieyi Lu as well as external collaborators, Dr. Sanket Deshmukh and Soumil Joshi (Virginia Tech). Soumil and I have been friends for the past 23 years and ever since middle school, we would work on whacky science projects. We always dreamt of becoming scientists and collaborating officially. Therefore, I am elated that our dream has come true after so many years.

I also thank my high school friends Kaunil Dhruv and Swapneel Datta for being absolute nerds and playing video games with me during 2020, which helped me push through the pandemic.

I will be forever indebted to my parents, Madhav and Mohita Datar as well as my grandparents, Manorama Datar, Ujwala Joshi and Prakash Joshi for raising me in an environment that fosters critical thinking, perseverance, mindfulness, meditation and a sense of pride for one's culture and values. These ideals have been the bedrock of my PhD journey and will continue to be for the rest of my life. I thank my younger sister Charuta for cheering me with her witty one-liners and for being the guinea pig of my dad jokes. I am also thankful to my new family, Sunita, Susheel, Chetna and Naman Virmani for encouraging me on this journey.

Finally, I want to thank Mishika Virmani, my friend, partner and wife. Of the 8 years that we are together, 5 were spent with us 13,000 km apart, bent over our aging 15"-laptop screens, squinting to catch a glimpse of each other through low picture quality video calls with sometimes

sketchy internet connections. Our respective PhD's have been a journey like no other and I am extremely grateful to have you as a fellow traveler. Thank you for putting up with me all this time over morning coffee (via Skype), weekend dates (via Skype) and those precious 60 minutes of screen time that we got with each other during weekdays, because of the grueling time difference. Thank you for being a good listener, a great advisor, a compassionate counselor, and a strong critic. Your constant support has made this thesis a reality.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Decarboxylases are chemically versatile enzymes capable of manipulating C–C bonds by reversibly converting carboxylic acids to their corresponding hydrocarbons. Thus, they are being considered as viable biocatalysts for the sustainable production of commodity chemicals. The formation of a carbanion intermediate poses a significant kinetic barrier to decarboxylation and therefore, nature has evolved cofactors such as thiamine pyrophosphate (TPP), pyridoxal-5'-phosphate (PLP) as well as metal ions, to facilitate the reaction.

Recently, a modified flavin cofactor was discovered that contains an extra 6 membered ring between the N5 and C6 positions of the isoalloxazine moiety. Named as prenylated flavin mononucleotide (prFMN), this cofactor features a unique azomethine ylide that is essential for catalysis. The UbiD-family of decarboxylases, named after the archetypical enzyme found in bacterial ubiquinone biosynthesis, utilizes prFMN to (de)carboxylate a number of α,β-unsaturated, (hetero)aromatic and phenolic carboxylic acids. In the well-studied enzyme ferulic acid decarboxylase (FDC), the reaction proceeds through the formation of a 1,3-dipolar cycloadduct between prFMN and the substrate, trans-cinnamic acid. On the other hand, in the protocatechuic acid decarboxylase AroY, an electrophilic mechanism is suggested. Overall, a detailed characterization of different UbiD-like enzymes can uncover novel mechanisms and benefit their development as biocatalysts. While FDC has been studied extensively, a kinetic evaluation of other UbiD-like enzymes is lacking. These enzymes are known to crystallize in distinct 'open' and 'closed' conformers but their relevance to catalysis also remains to be discovered. Lastly, the

biggest hurdle in studying UbiD-like enzymes is that oxidative maturation of prFMN is poorly understood. My work addresses some of these problems in the field of UbiD-catalyzed reactions.

Initially, I characterized a recently discovered prFMN dependent enzyme from *Mycolicibacterium fortuitum*. Named PhdA, this enzyme decarboxylates phenazine-1-carboxylic acid, providing *M. fortuitum* a competitive advantage over phenazine producers in soil. I developed an optimal method for reconstituting PhdA that doesn't require the use of reducing agents described previously. Moreover, I showed that PhdA can decarboxylate a number of (hetero)aromatic carboxylic acids, including anthracene-1-carboxylic acid. It also catalyzes the much slower exchange of solvent deuterium in phenazine. Finally I proposed a 1,3-dipolar cycloaddition mechanism for PhdA.

For a detailed analysis of PhdA's mechanism, I studied solvent isotope and viscosity effects. Surprisingly, I discovered that D_2O-associated changes in protein conformations significantly improved reaction rates. Molecular dynamics (MD) simulations performed in collaboration with Soumil Joshi and Dr. Sanket Deshmukh from Virginia Tech suggest that D_2O leads to domain closure, akin to the 'closed' conformer observed in crystal structures of several UbiD-like enzymes. Given that many UbiD-like enzymes crystallize in the 'open' form, these results show that optimizing solvent systems and/or engineering to adapt a 'closed' conformer might improve the efficiency of UbiD-catalyzed reactions.

Finally, I studied the biosynthesis and maturation of prFMN in detail to shed light on this process. I showed that the *in vitro* prenylation of FMN catalyzed by UbiX is inefficient and several products are formed that affect prFMN maturation. These species (collectively called as $prFMN^{ox}$) subsequently undergo solvolysis to re-form FMN as well as other degradation products.

Overall, my work shed light on the yet poorly understood prFMN maturation, expanded the substrate scope of UbiD-like enzymes and identified a novel way to engineer these proteins. This research would benefit future studies, improving scientific scholarship and building towards a more sustainable way for synthesizing commodity chemicals.

Chapter 1 Introduction

1.1 Decarboxylases – chemically versatile enzymes

Decarboxylation – the evolution of CO_2 from organic acids is one of the most fundamental chemical reactions in biochemistry, accounting for almost all of the CO_2 released in metabolism.[1] Therefore, it is hardly surprising that decarboxylases are prevalent in both primary and secondary metabolic processes. Even though decarboxylation is thermodynamically favorable owing to the release of CO_2, the formation of a subsequent carbanion intermediate presents a high-energy kinetic barrier (Figure 1.1A).[2] Hence, barring a few reactions where the negative charge can be stabilized by the substrate itself, most decarboxylases employ organic or metal-ion cofactors. A number of strategies, both oxidative and non-oxidative, have evolved for decarboxylation. In some cases, β-hydroxy acids are oxidized by $NAD(P)^+$ to their corresponding β-keto acids which can then stabilize the negative charge (Figure 1.1B). Malic enzymes and isocitrate dehydrogenase are prime examples of such oxidative decarboxylases. In other enzymes, decarboxylation itself is coupled to oxidation (e.g. ferrous-dependent oxidative decarboxylases, Figure 1.1C).[1]

Most non-oxidative decarboxylases convert carboxylic acids to their corresponding hydrocarbons by simply replacing the –COOH group with a proton. They utilize cofactors that interact with unique functional groups (FG) present on the substrate, and act as 'electron sinks' to stabilize the negative charge (Figure 1.1D).[3] Prominent examples of such 'electron sinks' are thiamine pyrophosphate (TPP) and pyridoxal 5'-phosphate (PLP) (Figure 1.1E). In the case of TPP, the FG is a keto group α to –COOH whereas for PLP, the FG is an α–NH_2. The most prominent TPP utilizing enzyme is pyruvate dehydrogenase whereas PLP features in amino acid

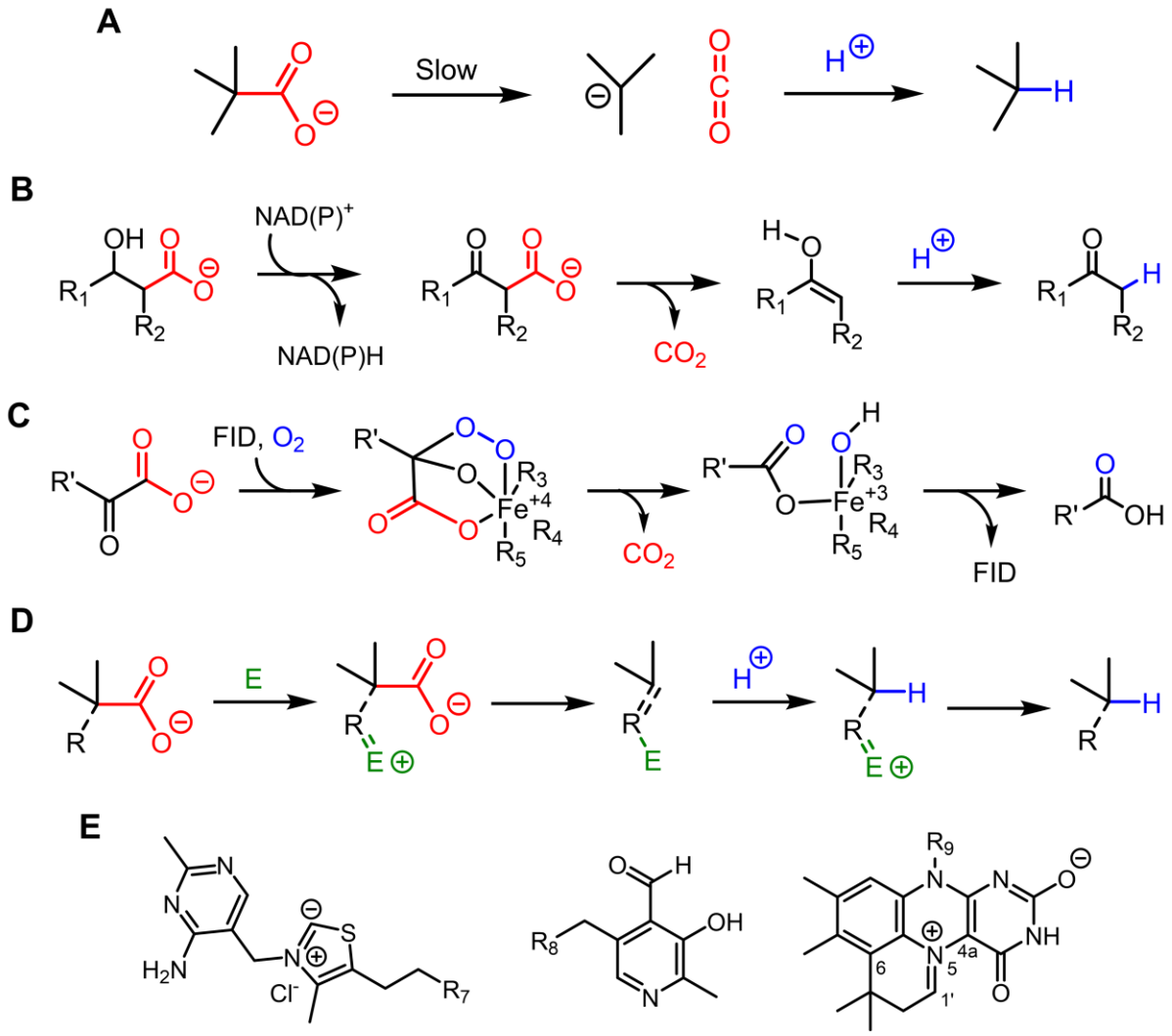


Figure 1.1: Decarboxylations in biochemistry. **A:** The carbanion intermediate poses a kinetic barrier to decarboxylation. **B:** Oxidative decarboxylation of β-hydroxy acids catalyzed by $NAD(P)^+$ dependent enzymes. **C:** Reactions catalyzed by ferrous-ion dependent oxidative decarboxylases (FID). **D:** Non-oxidative decarboxylation catalyzed by cofactors (E) that act as electron sinks and interact with unique functional groups (R) on the substrate. **E:** Examples of electron sinks. *Left* – thiamine pyrophosphate (R_7 = pyrophosphate), *middle* – pyridoxal 5'-phosphate (R8 = phosphate), *right* – Prenylated flavin mononucleotide (prFMN, R_9 = ribose phosphate). Representative mechanisms are adapted from ref.1 and ref.3.

decarboxylases. Recently, a new modified flavin cofactor was discovered that features a fourth ring added to the isoalloxazine moiety by virtue of a prenylation at the N5-position and subsequent cyclization with the C6 position (Figure 1.1E).[4] This modification imparts a unique azomethine ylide character to the parent flavin and allows interaction with α, β-unsaturated carboxylic acids. Named as prenylated flavin mononucleotide (prFMN), this cofactor is synthesized by bacterial prenyl transferases (UbiX) or their fungal homologs (PAD1). The widely distributed UbiD-family

of enzymes utilizes prFMN to facilitate decarboxylation reactions on a variety of unsaturated substrates.

1.2 Decarboxylases in biocatalysis

Apart from being important in metabolism, several reversible decarboxylases have sparked interest in the field of biocatalysis as sustainable alternatives to synthesizing commodity chemicals. In this regard both decarboxylation and carboxylation reactions performed by these enzymes have been exploited.[5-9]

1.2.1 Production of 'bioplastics' from biomass

The monomeric units of most conventional plastics originate from petroleum products, making their use unsustainable.[10] Therefore, several efforts are being undertaken to engineer microbes for the renewable production of these 'biomonomers'. Many metabolic pathways involved in biomass derived synthesis of commodity chemicals rely on decarboxylases.

For example, polyamides (nylons) are obtained through co-polymerization of diamines and diacids. Putrescine and cadaverine are two such diamines obtained from engineered microbes through PLP dependent ornithine and lysine decarboxylases (Figure 1.2A). Styrene, another valuable monomer used for making polystyrenes, is commercially synthesized from ethylbenzene, a petrochemical.[10] However, recently, efforts are being made to obtain styrene from biomass through the action of two enzymes, phenylalanine ammonia lyase (PAL) and the prFMN dependent ferulic acid decarboxylase (FDC) (Figure 1.2B).[5] Conjugated alkenes such as polyenes are used for manufacturing rubber as well as pharmaceutical compounds. Recently, 1,3,5-heptatriene was synthesized by the coupled reaction of FDC and a polyketide synthase as a proof

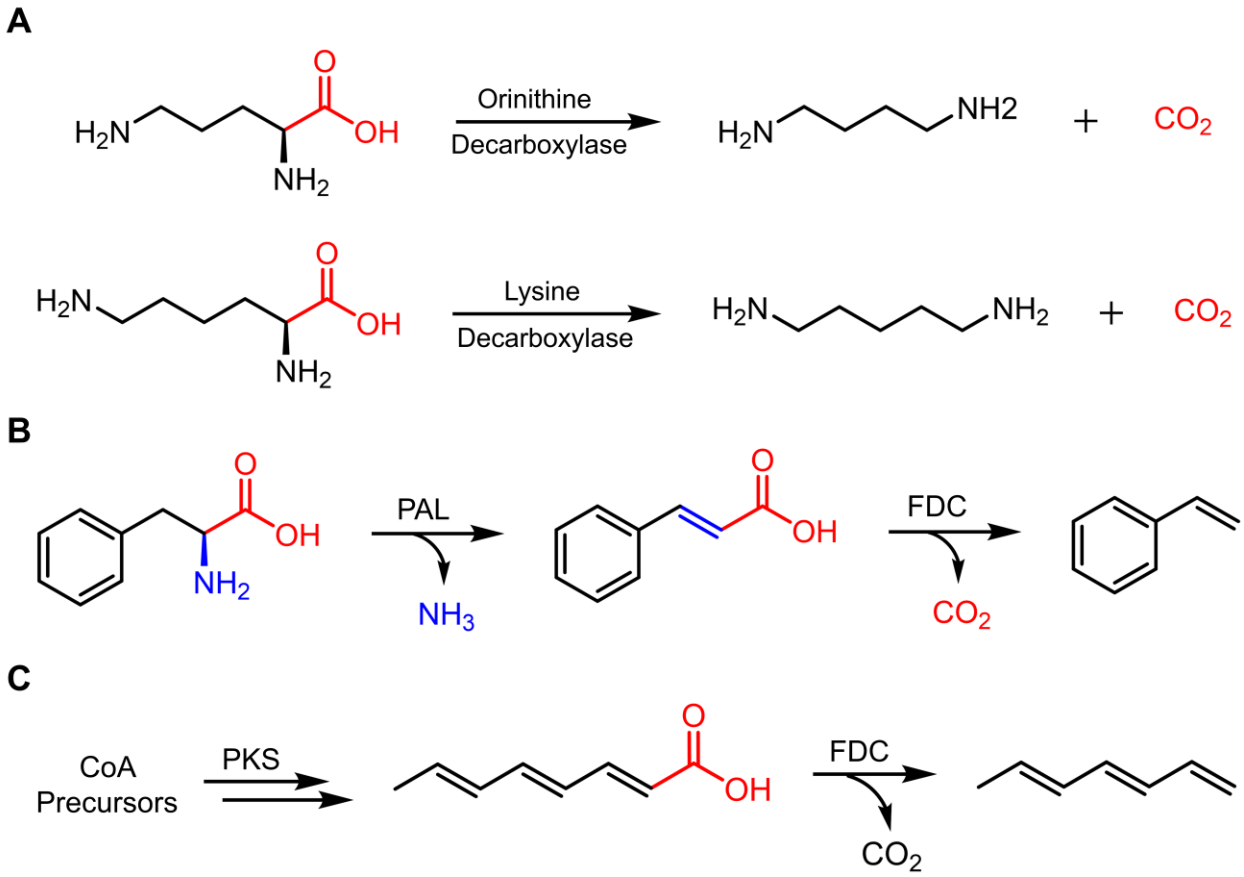


Figure 1.2: Synthesis of 'biomonomers' through reactions catalyzed by decarboxylases. **A:** Putrescine (top) and cadaverine (bottom) synthesized from ornithine and lysine through their respective PLP dependent decarboxylases. **B:** Phenylalanine is converted to styrene through trans-cinnamic acid by the enzymes phenylalanine ammonia lyase (PAL) and ferulic acid decarboxylase (FDC). **C:** Synthesis of 2,4,6-octatrienoic acid by a polyketide synthase and subsequent FDC catalyzed decarboxylation yields 1,3,5-heptatriene.

of principle demonstrating how decarboxylases can synthesize polyenes from natural products (Figure 1.2C).[6]

1.2.2 Enzymatic CO_2 fixation

CO_2 is an abundant naturally occurring C1 building block that can be utilized to build complex molecules from simple starting materials.[11] Enzymatic CO_2 fixation is an efficient method of atmospheric CO_2 capture making it an important aspect of reversible decarboxylases. Unfortunately, owing to its low free energy of formation, CO_2 fixation is thermodynamically uphill. Moreover, CO_2 is a gas under most conditions and equilibrates with H_2CO_3, HCO_3^- and

CO_3^{2-} in a pH-dependent manner. These species interconvert rapidly and the identity of the exact carboxylation cosubstrate is often ambiguous. Two general strategies exist to improve the yield of carboxylation: (i) Adding external energy (e.g. through phosphorylation) to make 'high energy' substrates or 'low energy' products; (ii) Modifying the reaction quotient by increasing the initial concentration of substrates or 'removing' products to shift the equilibrium concentrations of all reaction components (Figure 1.3A).[12]

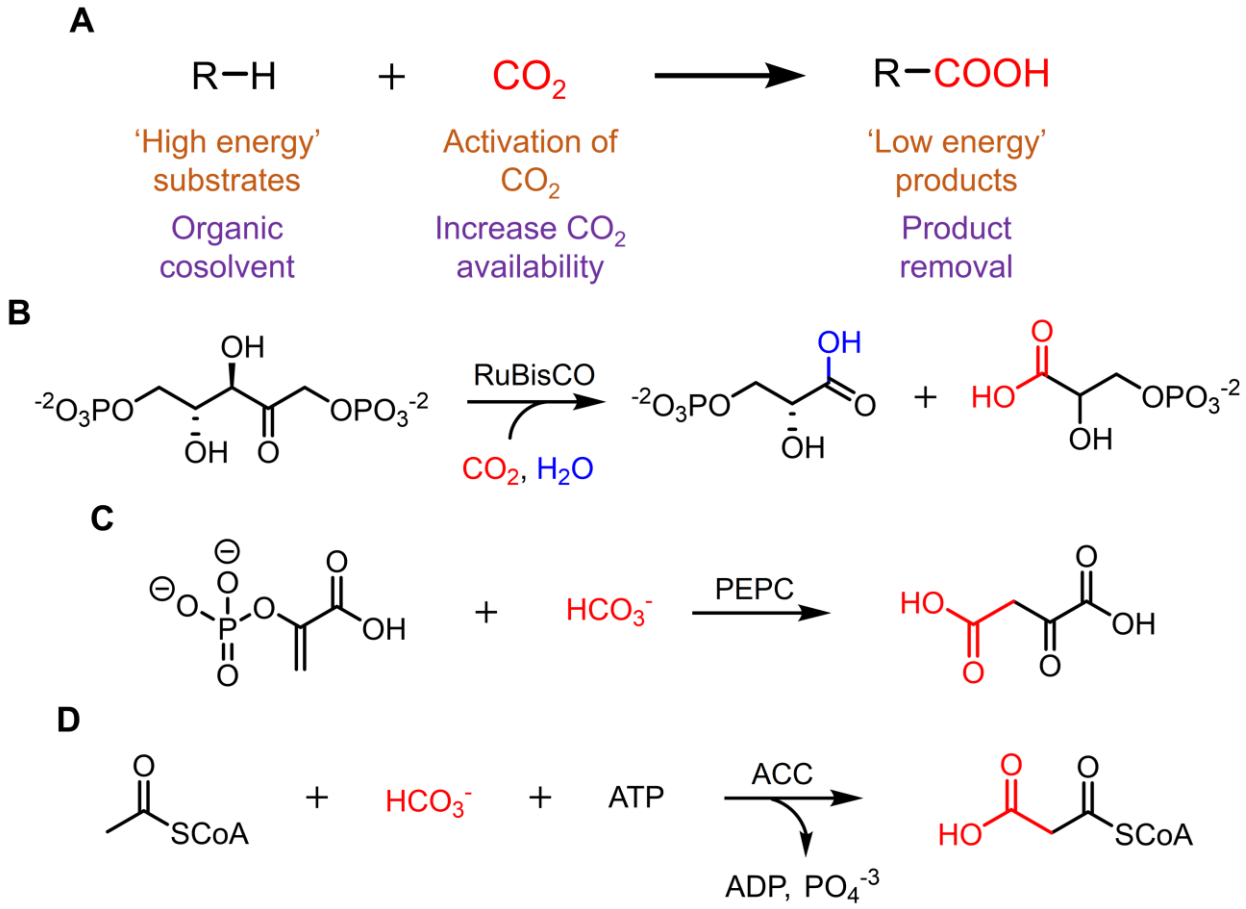


Figure 1.3: Reactions catalyzed by natural carboxylases. **A:** Different strategies to facilitate carboxylation – improving reaction energetics (orange), shifting equilibrium concentrations (purple). **B:** Synthesis of 3-phosphoglycerate catalyzed by ribulose-1,5-bisphosphate carboxylase (RuBisCO). **C:** Phosphoenolpyruvate carboxylase (PEPC) converts phosphoenolpyruvate to 2-oxosuccinic acid. **D:** Acetyl-CoA is carboxylated to malonyl-CoA by Acetyl-CoA carboxylase (ACC).

Many natural carboxylases such as ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and phosphoenolpyruvate carboxylase (PEPC) utilize the first strategy (Figure 1.3B

and C). The substrates used by these enzymes are activated through phosphorylation, increasing their free energy, thus, carboxylation is in fact thermodynamically favorable (ΔG° = –32 kJ/mol for RuBisCO and –32.2 kJ/mol for PEPC). Other enzymes such as biotin-dependent acetyl-CoA carboxylase (ACC) activate CO_2 through phosphorylation, facilitating carboxylation (Figure 1.3D).[12]

The biocatalytic utility of naturally occurring carboxylases is limited by strict substrate specificity.[13] On the other hand, the robustness and substrate promiscuity exhibited by reversible decarboxylases make them attractive biocatalysts. Some examples include TPP dependent, prFMN dependent, metal dependent and cofactor independent decarboxylases.[12] However, the reaction equilibrium for many of these enzymes favors decarboxylation and several strategies have been devised to overcome this limitation.

The simplest alteration involves increasing the availability of CO_2 through higher concentrations of bicarbonate salts or using pressurized and even supercritical CO_2.[13] For example, optimizing pH and bicarbonate concentrations led to ~80% conversion of acetaldehyde to pyruvate by the TPP dependent pyruvate decarboxylase (PyDC).[14] Similarly, using organic cosolvents or whole cells to increase the concentration of substrate is a great strategy. Adding ~20% (v/v) of water-miscible cosolvents improved the 2,6-dihydroxybenzoic acid decarboxylase (2,6-DHBD) catalyzed carboxylation of resorcinol by ~50%.[15] The yield of prFMN-dependent HmfF catalyzed carboxylation of 2-furoic acid to 2,5-furandicarboxylic acid was improved by using *E. coli* whole cells expressing HmfF and adding millimolar quantities of substrate (Figure 1.4A).[9]

Other strategies include removing the carboxylic acid produced by coupling carboxylation to a subsequent, favorable reaction. Adding tetrabutylammonium bromide to 'salt out' 2,6-dihydroxybenzoic acid produced in 2,6-DHBD catalyzed carboxylation of resorcinol improved

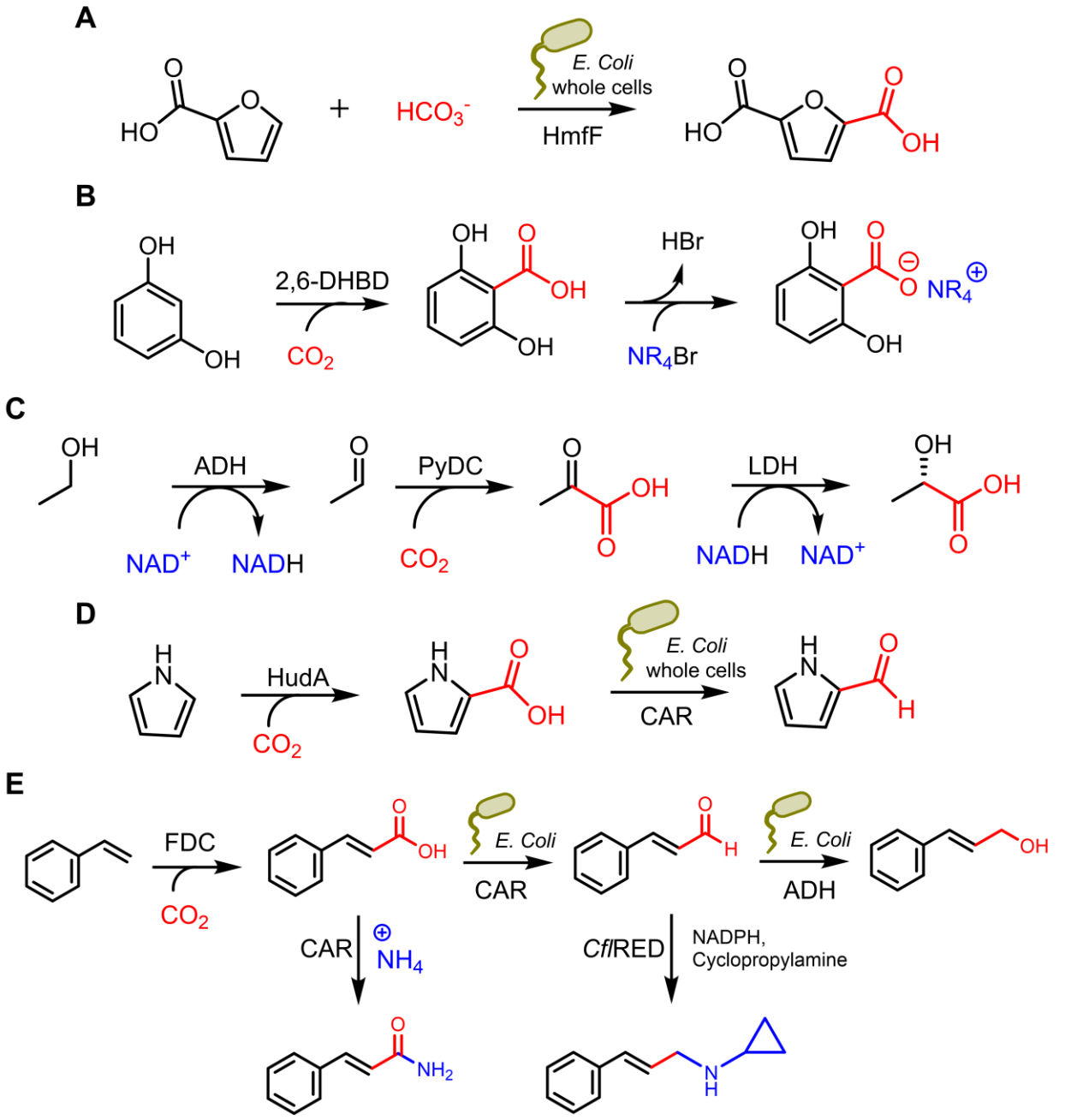


Figure 1.4: Carboxylations performed by reversible decarboxylases. **A:** *E. coli* cells expressing HmfF convert 2-furoic acid to 2,5-furandicarboxylic acid. **B:** 2,6-dihyrdoxybenzoic acid decarboxylase (2,6-DHBD) catalyzed carboxylation of resorcinol is improved by adding tetrabutylammonium bromide (R = butyl) to 'salt out' the carboxylic acid product. **C:** Alcohol dehydrogenase (ADH), pyruvate decarboxylase (PyDC) and lactate dehydrogenase (LDH) are combined to efficiently convert ethanol to lactic acid. **D:** HudA is coupled to carboxylic acid reductase (CAR) expressed in *E. coli* cells to improve carboxylation of pyrrole. **E:** CO_2 fixation and subsequent functionalization of styrene is achieved through coupled reactions of ferulic acid decarboxylase (FDC), CAR, ADH and imine reductase (*Cfl*RED).

yield from 37% to 80–97% (Figure 1.4B).[7] An ingenious method of converting ethanol to lactic acid was developed by Tong et al.[8] Here, acetaldehyde synthesized from ethanol and NAD^+ by

alcohol dehydrogenase (ADH) was utilized by the TPP-dependent PyDC to form pyruvate. Furthermore, lactate dehydrogenase (LDH) catalyzed the reduction of pyruvate using NADH synthesized by ADH, thus, preventing decarboxylation of pyruvate at equilibrium (Figure 1.4C).

The Leys group has demonstrated the ability of prFMN-dependent UbiD-like enzymes to functionalize C–H bonds in hydrocarbons by coupling carboxylation to reductions performed by carboxylic acid reductases (CARs).[16, 17] The pyrrole-2-carboxylic acid decarboxylase, HudA, was coupled to CAR to yield pyrrole-2-carboxyldehyde (Figure 1.4D). Similarly, coupling FDC with *E. coli* whole cells expressing CAR and endogenous ADH led to 95% conversion of styrene to cinnamyl alcohol (Figure 1.4E). Moreover, excluding NADPH from the above reaction yielded cinnamamide, and coupling the reaction to an imine reductase (*Cfl*RED) afforded further functionalization. The above discussion shows the versatility of reversible decarboxylases in the production of value added chemicals from biomass as well as through CO_2 fixation pathways.

1.3 The UbiX/UbiD enzyme system

The *ubiX*/*ubiD* genes were first identified in the ubiquinone biosynthesis pathway[18, 19] and were proposed to convert 4-hydroxy-3-octaprenylbenzoic acid to 2-octaprenylphenol, a precursor to ubiquinone.[2] Also known as coenzyme Q, ubiquinone is a redox cofactor that is ubiquitously found in electron transport.[20] Therefore, UbiX/UbiD enzymes are prevalent in many bacterial species. Additionally, hydroxybenzoic acid decarboxylases typified by *shdB*/*vdcB* and *shdC*/*vdcC* from *Sedimentibacter hydroxybenzoicus*/*Streptomyces* sp. D7 and present in all three microbial domains, were found to share sequence homology with UbiX (>50% similar to ShdB/VdcB) and UbiD (~50% similar to ShdC/VdcD) respectively.[21] Furthermore, the fungal enzymes FDC and PAD1, involved in production of styrene, are homologs of UbiD and UbiX respectively[2]

suggesting that the larger UbiX/UbiD enzyme family is not only found in bacteria, but also exists in archaea, fungi and unicellular algae.[22]

Initially, UbiD/UbiX as well as the corresponding FDC/PAD1 were mischaracterized as isofunctional enzymes. However, no decarboxylase activity could be detected for purified UbiX.[22] Instead, FDC, upon expression and purification from *E. coli*, could decarboxylate cinnamic acid to styrene.[2] Interestingly, FDC purified from an *E. coli* strain lacking *ubiX* (*ΔubiX*) was inactive, but its activity could be restored upon addition of cell lysates from wild type *E. coli* or *E. coli ΔubiX* overexpressing PAD1. Moreover, upon co-expression with PAD1 or UbiX, FDC's activity improved significantly. These results showed that the homologous proteins PAD1 and UbiX synthesized a cofactor that was responsible for FDC activity. Eventually, a modified flavin cofactor, later named prFMN, was discovered in the active site of the FDC crystal structure.[4] Subsequently, it was shown that UbiX/PAD1 prenylate reduced flavin mononucleotide ($FMNH_2$) to synthesize prFMN.[23, 24]

1.3.1 Biosynthesis and maturation of prFMN

Dimethylallylphosphate (DMAP) or dimethylallylpyrophosphate (DMAPP) are used as prenyl donors by UbiX/PAD1 enzymes for prFMN synthesis.[23, 24] Interestingly, no correlation exists between the choice of prenyl donor and whether the enzyme is prokaryotic or eukaryotic.[25] While some homologs can use both DMAP and DMAPP, others are more specific. Biosynthesis of DMAP is suggested to occur through the phosphorylation of prenol by a ThiM-like kinase or by dephosphorylation of DMAPP using a Nudix-based hydrolase.[26] In fact, addition of prenol to *E. coli* cultures expressing UbiX was shown to improve bound prFMN in the purified enzyme.

The UbiX homolog from *Pseudomonas aeruginosa* (*Pa*UbiX) was the first one to be characterized.[23] Stopped flow spectroscopy revealed the formation of a long lived intermediate,

with observed rate constants for its formation and decay as 177 ± 7 s^{-1} and 0.316 ± 0.002 s^{-1} respectively. Kinetic crystallography suggests that this stable intermediate is an FMN-N5-prenyl-C1' adduct, based on which the following mechanism was proposed (Figure 1.5).[25] Here, the N5 of $FMNH_2$ attacks the dimethylallyl moiety in an S_N1 type fashion to form the N5-C1' intermediate (**Int III**, Figure 1.5). The prenyl group then accepts a proton from PO_4^{3-} to become an electrophile that undergoes a Friedel-Crafts type alkylation, forming the fourth ring. The slow rate constant observed for the decay of the N5-C1' adduct (0.316 ± 0.002 s^{-1}) would indicate that Friedel-Crafts alkylation is rate-limiting. Eventually, S15 and E49 are proposed to remove the flavin C6 proton to facilitate rearomatization. While mutagenesis reveals the importance of S15 and E49 in abstracting the initial proton from N5, their role in converting **Int V** to **Int VI** (referenced in Figure 1.5) has not be determined. Studies on other UbiX homologs speculate that if DMAP is used as a substrate for enzymes specific to DMAPP, the reaction stalls at the N5-C1' adduct, which can be rescued by supplementing the reaction with (pyro)phosphate.[25]

UbiX catalyzed reaction yields reduced prFMN ($prFMNH_2$), whereas mass-spectrometry and crystallography studies on FDC reveal that an oxidized form ($prFMN^{iminium}$) is catalytically active.[4] While some UbiD-like enzymes can be obtained in their active form by co-expression with UbiX, the exact oxidative maturation process is poorly understood. O_2 exposure of UbiX bound $prFMNH_2$ following *in vitro* anaerobic synthesis leads to the formation of a purple $prFMN^{radical}$ species and no further oxidation occurs.[23] The same species is observed upon reduction of FDC with $NaBH_3CN$ and re-oxidation.[4] Moreover, the presence of this species in FDC is negatively correlated to decarboxylation activity.[27] Extraction of free $prFMNH_2$ from UbiX reaction mixture and subsequent oxidation leads to the transient formation of a $prFMN^{radical}$ like species that eventually degrades to a molecule exhibiting a flavin-like UV-Vis spectrum and an m/z of 559.183

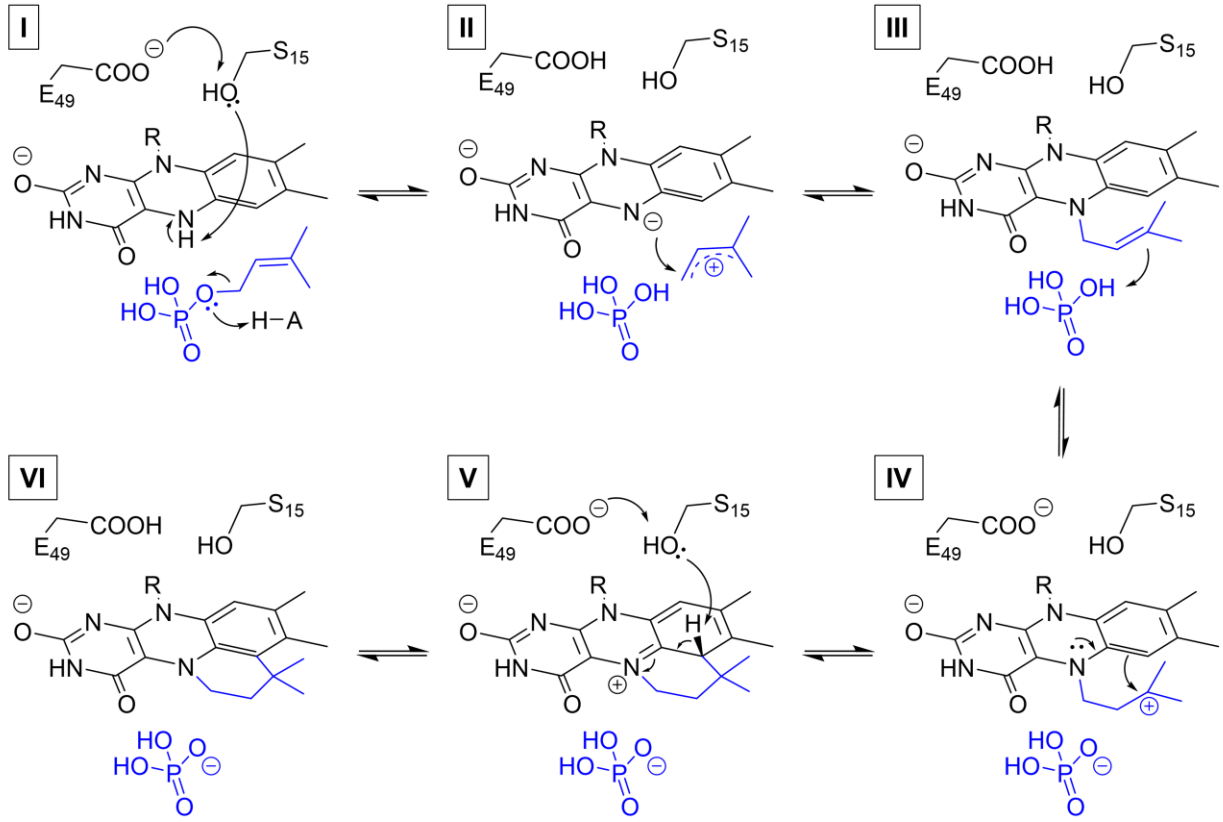


Figure 1.5: Proposed mechanism for UbiX catalyzed synthesis of prFMNH$_2$. Different reaction intermediates are denoted by roman numerals.

units. More importantly, this molecule, proposed to be prFMN-C4a-OOH, is unable to reconstitute apo-FDC. In fact, adding this putative prFMN hydroperoxide to holo-FDC obtained from *E. coli* leads to loss of activity.[27]

Active UbiD-like enzymes can only be obtained by incubation with prFMNH$_2$ under anaerobic conditions followed by oxidation. Furthermore, O_2 appears to be more effective than other oxidizing agents, such as $K_3[Fe(CN)_6]$.[27] This suggests that UbiD-like enzymes play an essential role in prFMN maturation, although the exact mechanism is unknown. Mutagenesis has revealed that the highly conserved catalytic residues R173, E277 and E282 (*Aspergillus niger* FDC – *An*FDC numbering) play an important role in this process.[28] The following mechanism is proposed for UbiD-mediated prFMNH$_2$ maturation (Figure 1.6).[27] Although experimental evidence argues that FDC-bound prFMN^{radical} or prFMN-C4a-OOH don't progress to

prFMN^{iminium}, they are proposed as intermediates in the maturation process. This is rationalized through the formation of a transient prFMN-C4a-OO^{-} species during maturation, perhaps assisted by R173.

O_2 mediated prFMN oxidation is a double edged sword as prolonged aerial exposure is detrimental to decarboxylation activity.[4] Older crystals of FDC reveal an inactive prFMN C1' hydroxy adduct (prFMN^{OH}).[4] Furthermore, in the presence of UV light, prFMN^{iminium} isomerizes to prFMN^{ketiminie}, leading to loss of catalytic activity.[28]

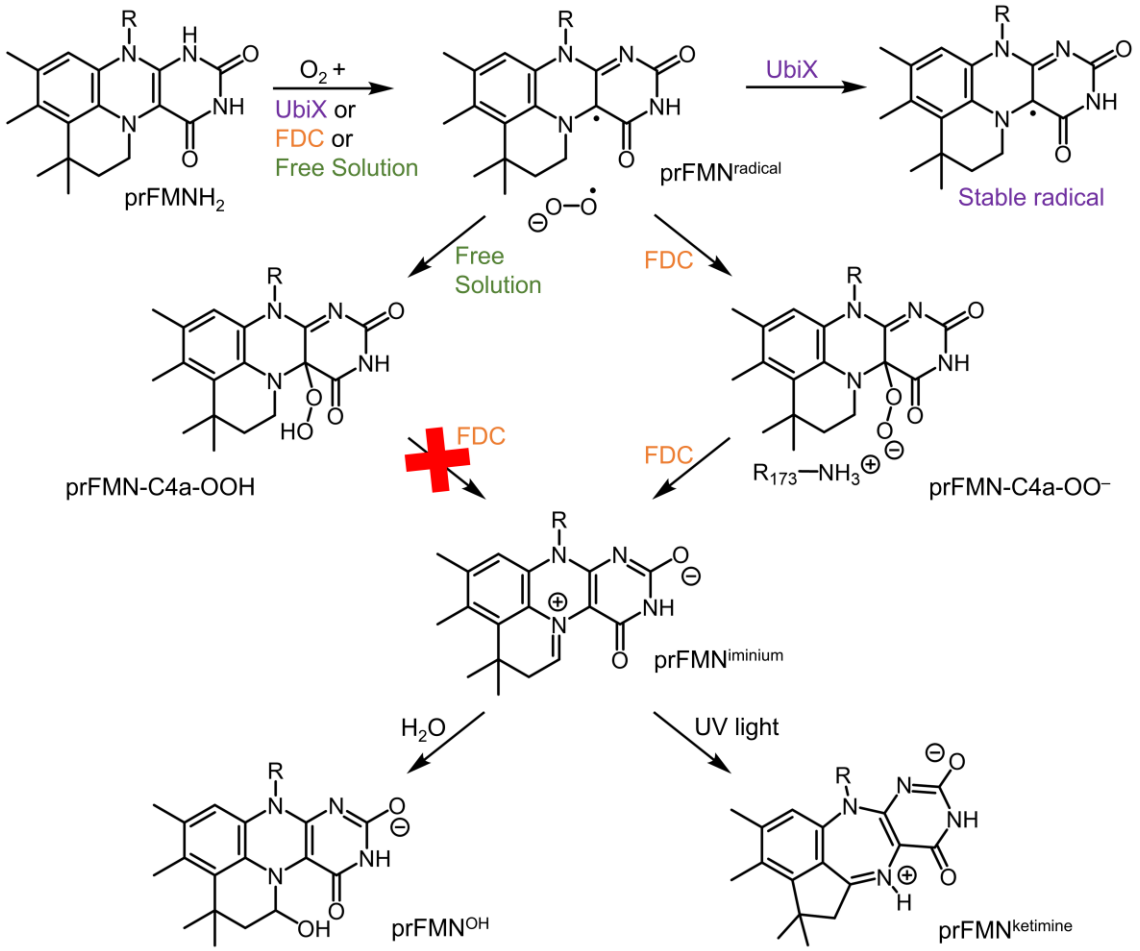


Figure 1.6: Maturation of prFMN. Upon exposure to O_2, prFMNH$_2$ forms prFMN^{radical} under all conditions. When bound to UbiX, the prFMN^{radical} is stable. Formation of active prFMN^{iminium} occurs only in the presence of FDC. Adding prFMN-C4a-OOH to FDC doesn't lead to active cofactor. With time, prFMN^{iminium} can degrade to prFMN^{OH} or prFMN^{ketimine}.

1.3.2 The UbiD-family of enzymes

While the substrates of UbiX/PAD1 enzymes are limited to FMN and DMAP/DMAPP, the UbiD-family of enzymes demonstrates a broad substrate scope. The basic requirement for decarboxylation appears to be the presence of a double bond α, β- to the carboxylic acid. A phylogenetic analysis reveals that UbiD-like enzymes with similar substrates group together. Thus, three distinct 'clades' can be identified: α, β-unsaturated acids, aromatic acids and phenolic acids (Figure 1.7).

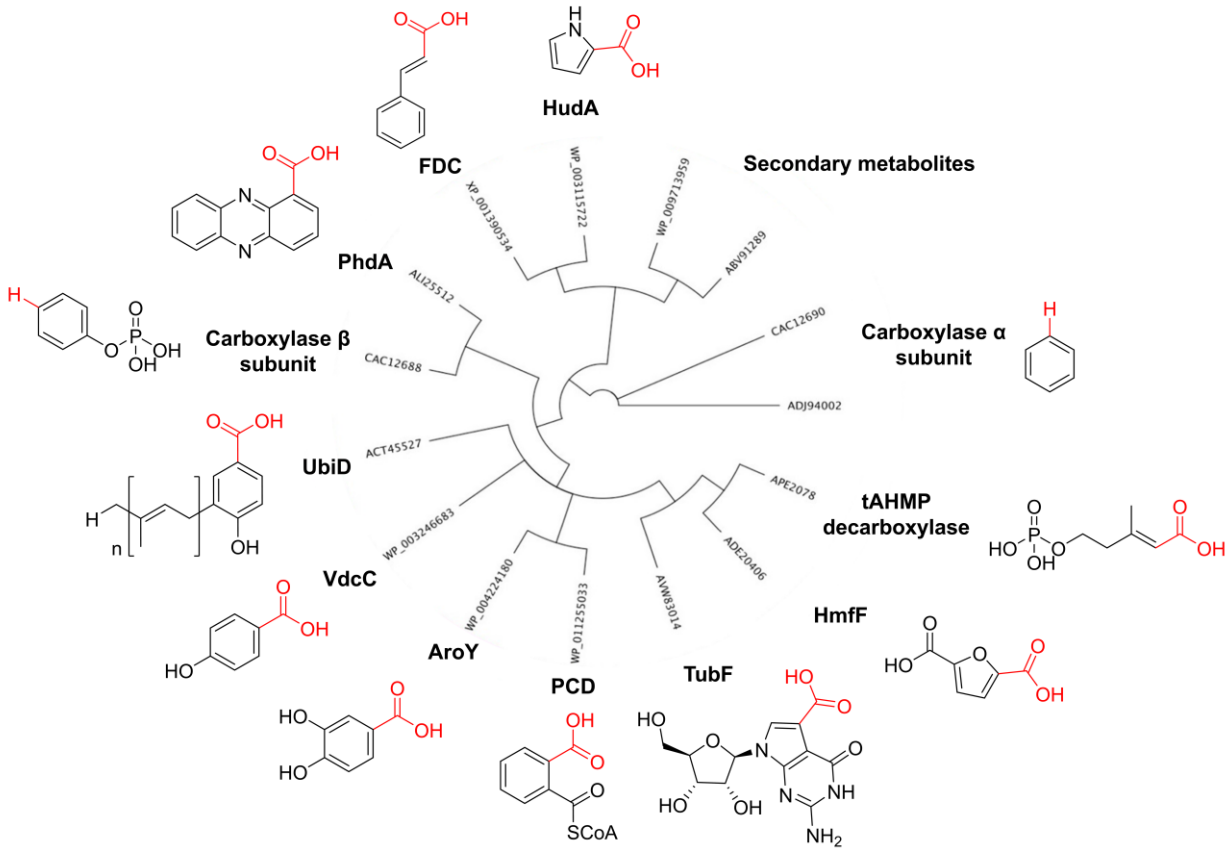


Figure 1.7: Phylogeny tree for UbiD-like enzymes. Each annotated enzyme name is written against its accession code. Chemical structures of their respective substrates are also included (reproduced from ref.32). CAC12688 and CAC12690 are annotated as carboxylases and therefore, their substrates are depicted as hydrocarbons instead of carboxylic acids.

The enzymes that decarboxylate α, β-unsaturated acids are typified by FDC, which is the most well studied enzyme in the entire UbiD-family. TtnD, involved in tautomycetin biosynthesis

is another characterized member of this group[29] whereas putative members include 9-methylstreptimidone biosynthesis enzyme (SmdK) and trans-anhydromevalonate-5-phosphate (tAHMP) decarboxylase.[22] Protocatechuic acid decarboxylase (AroY) was the first prFMN-based phenolic acid decarboxylase to be investigated,[11] followed by vanillic acid decarboxylase (VdcC).[30] The archetypical enzyme UbiD also belongs to this group, however, purified UbiD couldn't be reconstituted and showed no activity when incubated with its substrate.[31] Yet to be studied prFMN dependent phenolic acid decarboxylases include gallic acid decarboxylase (LpdC) and phenylphosphate carboxylase (PpC).[22] Furandicarboxylic acid decarboxylase (HmfF),[32] pyrrole-2-carboxylic acid decarboxylase (HudA),[33] indole-3-carboxylic acid decarboxylase (InD),[34] phthaloyl-CoA decarboxylase (PCD)[35] and phenazine-1-carboxylic acid decarboxylase (PhdA)[36] are some prFMN dependent aromatic acid decarboxylases that have been characterized. Putative benzene and naphthalene carboxylases are the most sought after enzymes in this category, because of their implications in biocatalysis and aromatic C–H functionalization.[22]

1.4 Reaction mechanisms for UbiD-like enzymes

The chemical mechanism of prFMN dependent enzymes has been a subject of interest ever since the discovery of this peculiar cofactor. Historically, flavins are known to catalyze both one- and two-electron redox reactions by virtue of their isoalloxazine ring, which can exist in quinone, semiquinone and hydroquinone forms. The heavily modified isoalloxazine core in prFMN allows the formation of an N5-prenyl C1' iminium cation coupled to an N1-O2 anion, giving the cofactor an azomethine-ylide type character. This dipolar character is central to prFMN catalysis and facilitates two distinct types of reactions: 1, 3-dipolar cycloaddition and electrophilic addition.

1.4.1 1,3-dipolar cycloaddition

As the name suggests, 1,3-dipolar cycloaddition utilizes the dipolar nature of prFMN. First observed in FDC, this mechanism is proposed for most prFMN based α,β-unsaturated acid decarboxylases. The mechanism for the decarboxylation of trans-cinnamic acid is shown in Figure 1.8.[37] Here, the substrate binds FDC to form a Michaelis complex (**E.S**). This is followed by the formation of a 1, 3-dipolar cycloadduct between the enzyme and substrate (**I1**), decarboxylation (**I2**), proton abstraction from a catalytic glutamate residue to form enzyme-product cycloadduct (**I3**) and subsequent cycloelimination (**E.P**) to yield styrene. Akin to other 'electron sink' cofactors, prFMN stabilizes the negative charge formed during decarboxylation by delocalizing it on the isoalloxazine ring.

The mechanism was initially proposed based on co-crystallization studies with FDC and phenylpyruvate which revealed an enzyme-inhibitor complex between prFMN C1' and a phenylacetaldehyde like species.[4] This complex can be formed through 1,3-dipolar cycloaddition between prFMN and the phenylpyruvate tautomer α-hydroxycinnamic acid, decarboxylation to yield a prFMN C1'-α-hydroxystyrene adduct and subsequent tautomerization. Linear free energy relationships (LFER),[38] native mass spectrometry with inhibitiors[39] and kinetic crystallography with phenylpropiolic acid[40] provides further evidence in support of a cycloaddition mechanism. While the equilibrium favors decarboxylation, incubating FDC with styrene in D_2O revealed that it will regioselectively exchange solvent D into styrene, but only in the presence of CO_2.[38] Moreover, the forward reaction exhibits a large *normal* solvent isotope effect on the proton transfer (3.33 ± 0.09) and substrate secondary isotope effects indicate that cycloelimination of **I3** is the rate-limiting step. This is further corroborated by UV-Vis and native mass spectrometry which

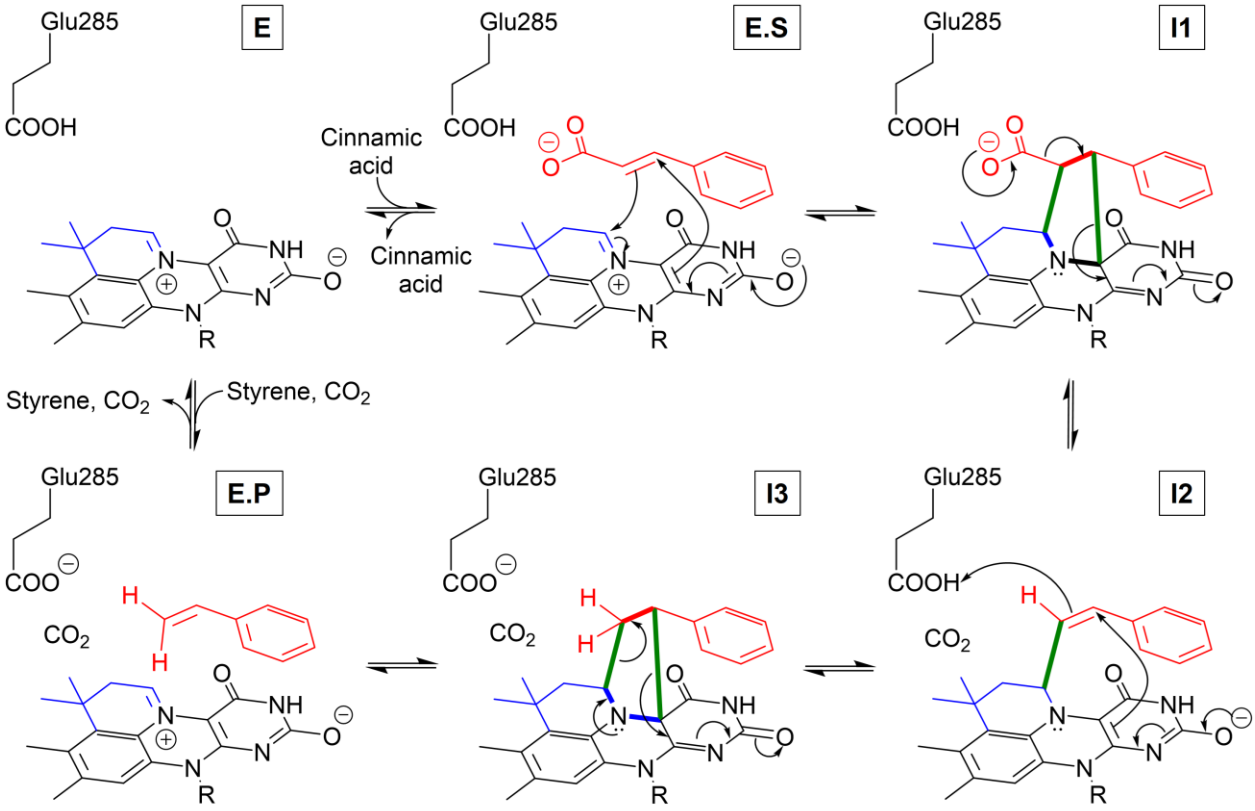


Figure 1.8: 1, 3-dipolar cycloaddition mechanism proposed for FDC. The catalytic Glu residue (E285) is numbered as per *Saccharomyces cerevisiae* FDC (*Sc*FDC).The various reaction intermediates are labelled as **E, E.S, I1, I2, I3** and **E.P** (adapted from ref. 37).

show the accumulation of **I3** at steady state, irrespective of whether cinnamic acid or styrene were used as substrates.[37]

1.4.2 Electrophilic addition

This mechanism exploits the electrophilic nature of the N5-C1' bond in prFMN. First reported in AroY,[11] the electrophilic mechanism is proposed for most prFMN based phenolic and heteroaromatic acid decarboxylases since the heteroatom can activate the aromatic ring for a nucleophilic attack (Figure 1.9). This is followed by decarboxylation, proton abstraction from a Glu/Asp residue and elimination of prFMN. Unlike the cycloaddition mechanism, negative charge stabilization is afforded by the substrate heteroatom.

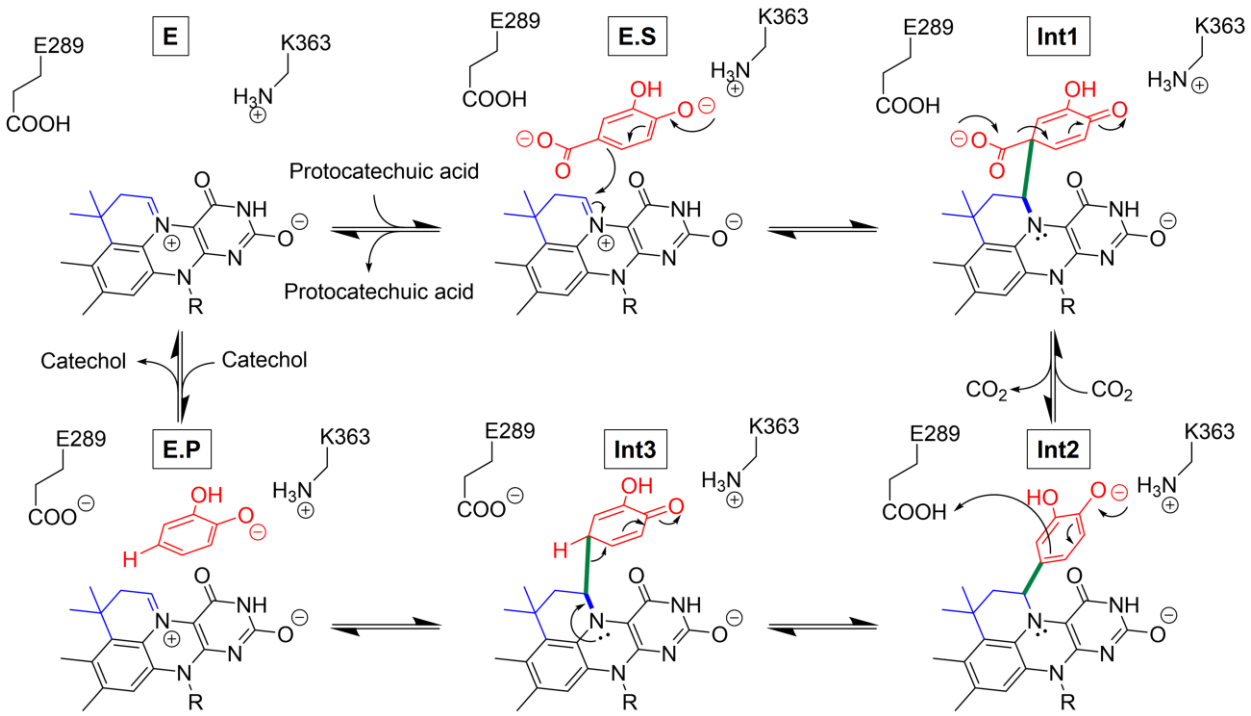


Figure 1.9: Electrophilic mechanim for AroY. The catalytic residues are numbered as per *Enterobacter cloacae* AroY (*Ec*AroY).The various reaction intermediates are labelled as **E, E.S, Int1, Int2, Int3** and **E.P** (adapted from ref. 11)

Experimental evidence for the electrophilic mechanism is limited. This is mainly because of difficulties in obtaining active holo-enzymes, problems in cofactor maturation and lack of co-crystallization with substrates (refer to section 1.7 for a detailed discussion). For AroY, the substrate protocatechauic acid was modelled in the active site and density functional theory (DFT) calculations revealed that an electrophilic mechanism was more feasible than cycloaddition.[11] Mechanistic studies on furan dicarboxylic acid decarboxylase (HmfF) were inconclusive and suggested that both cycloaddition and electrophilic addition through activation of the furan oxygen were possible routes.[32] Pyrrole-2-carboxylic acid decarboxylase (HudA) was the first enzyme after FDC to be crystallized with a covalent prFMN-inhibitor (imidazole) adduct. While DFT calculations suggested the formation of a non-concerted cycloadduct species; reactivity trends with different pyrrole, indole and imidazole carboxylic acids were in accordance with an electrophilic

aromatic substitution type mechanism.[33] Similar experiments support the electrophilic mechanism for vanillic acid decarboxylase (VdcC) and indole-3-carboxylic acid decarboxylase (InD).[30, 34]

1.4.3 Alternate mechanisms

A Michael-type addition mechanism based on the nucleophilic nature of prFMN was initially proposed for FDC (Figure 1.10). Here, the C4a from prFMN attacks the position β to the carboxylic group. However, crystallographic studies suggested a bond between the prFMN C1' and the position α to the carboxylic acid. Moreover, the m/z of a prFMN-inhibitor adduct detected through native mass spectrometry was inconsistent with Michael-type addition. Therefore, this mechanism was discarded for FDC, but theoretically, can be envisioned for other UbiD-like enzymes.

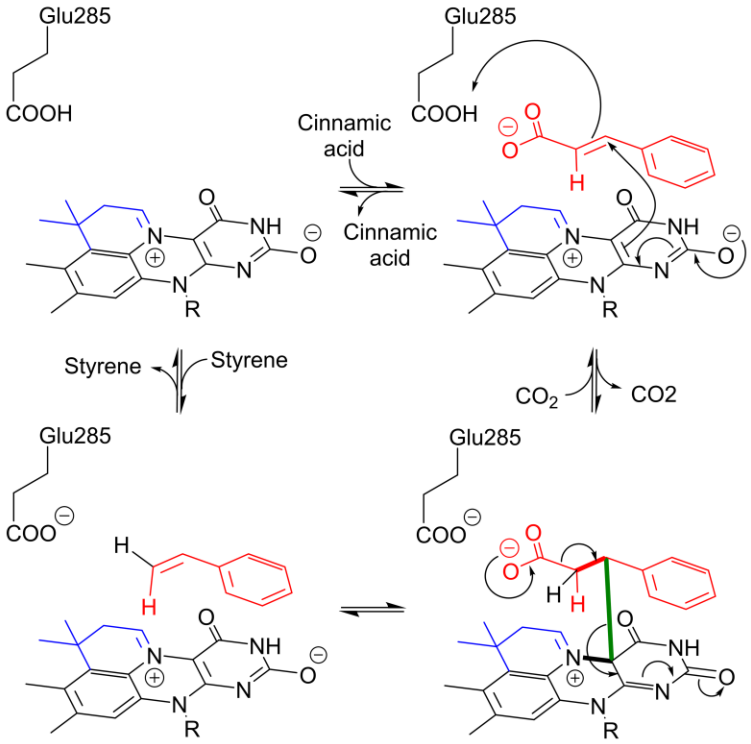


Figure 1.10: Alternate mechanisms for prFMN catalyzed decarboxylation. Michael-type addition reaction proposed for FDC.

As discussed in section 1.3.1, prFMN can exist in different redox states, including $prFMN^{iminium}$, $prFMNH_2$, $prFMN^{radical}$ and $prFMN^{ketimine}$. So far, there is no evidence to suggest that any of these species apart from $prFMN^{iminium}$ are catalytically active. While the existence of radical intermediates, possibly involving $prFMN^{radical}$ has been hinted in the recently discovered prFMN dependent phenzine-1-carboxylic acid decarboxylase (PhdA, refer to section 1.6), additional experiments are needed to verify this proposal. Since prFMN does exist in different redox forms, it is possible that a yet undiscovered enzyme family might utilize other forms of prFMN for catalysis.

1.5 Structural features and domain dynamics of UbiD-like enzymes

Quaternary structures of UbiD-like enzymes are quite diverse and comprise homodimers, homotetramers, homohexamers and even heterododecamers.[41] Interestingly, the tertiary structure of the monomeric unit (Figure 1.11A) is conserved and consists of: N-terminal prFMN binding domain (green), central α-helix (orange), oligomerization domain (magenta) and a C-terminal α-helix (grey). The active site is located at the interface of the N-terminal and oligomerization domains (Figure 1.11A, prFMN displayed in spheres). The prFMN binding region consists of a split barrel fold, akin to some other FMN binding proteins. Two metal ions, Mn^{2+} and K^+ (Fe^{2+} and Na^+ have also been observed) are involved in tethering prFMN via its phosphate group (Figure 1.11B). Except for E233 (*An*FDC numbering), the metal ion coordinating residues are not conserved, displaying the potential diversity in the choice of the metal ion.

Based on FDC's crystal structure, the substrate stacks on top of prFMN C4a-N5-C1' bonds to facilitate catalysis (Figure 1.11C). The conserved residues E282, R173 and E277 (*An*FDC) appear essential not only for activity but also for cofactor maturation, as discussed previously.[28] E282 (replaced with D in some enzymes) is the catalytic residue involved in proton transfer. Other

active site residues vary widely between enzymes and appear to facilitate proper binding and orientation of the substrate. For example, H296 in HmfF and N318 in HudA appear to hydrogen bond with the heteroatom and modifying these residues leads to an altered substrate scope.[32, 33] Moreover, modifying certain bulky residues expands the active site volume, further modifying reactivity, for example, *An*FDC I327S can decarboxylate naphthalene-2-carboxylic acid.[16]

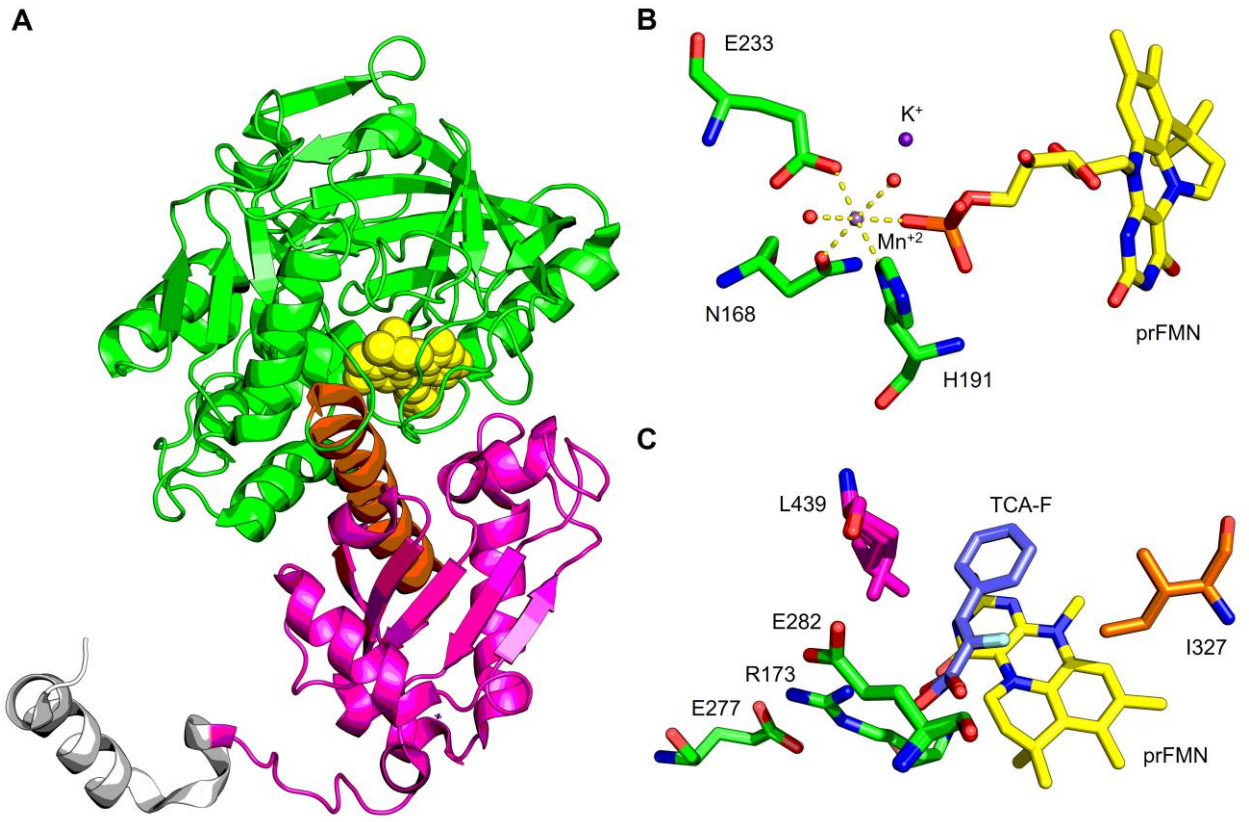
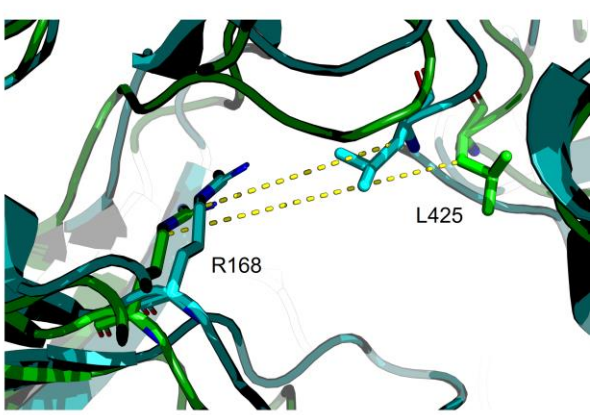


Figure 1.11: Structural insights into UbiD-like enzymes. **A:** Crystal structure of *An*FDC monomer (PDB:4ZA4). *Green* – N-terminal prFMN binding domain, *orange* – central α-helix, *magenta* – oligomerization domain, *grey* – C-terminal α-helix, *yellow* – prFMN. **B:** prFMN is bound to the enzyme through Mn^{+2} and K^{+} ions in *An*FDC (PDB:4ZA4). **C:** Active site of *An*FDC (PDB:6R2R). *TCA-F* – α-fluorocinnamic acid.

Certain members of the hydroxybenzoic acid decarboxylase category, such as vanillic acid decarboxylase (VdcC) require the presence of an auxiliary protein (VdcD) to be active.[30] VdcD is an allosteric modulator and doesn't directly participate in catalysis. It consists of a tetra-cysteine motif binding a central Zn ion and resembles Zn ribbon proteins.



C

Enzyme	PDB	Oligomer	R–L Distance (Å)	Conformer
FDC	4ZA4	Dimer	9.8	Closed
HmfF	6H6V	Hexamer	12.0 ± 0.1	Open
UbiD	5M1B	Hexamer	13.5 ± 1.1	Open
HudA	7ABN	Dimer	10.0 ± 0.1	Closed
	4IWS	Dimer	10.9 ± 0.4	Open
AroY	5O3N	Hexamer	11.9 ± 0.1	Open
VdcCD	7AE4	Hexamer	9.3 ± 0.1	Closed
	7AE5	Hexamer	14.4 ± 0.2	Open
InD	7P9Q	Hexamer	12.7 ± 0.2	Open

Figure 1.12: Distinct conformations of UbiD-like enzymes. Superimposition of 'open' (PDB:7AE5) and 'closed' (PDB:7AE4) VdcC monomers **(A)** as well as visualization of the R168–L425 distance **(B)**. *Green* – open, *cyan* – closed. **C:** Table summarizing available crystal structures for UbiD-like enzymes, oligomeric states, R-L distances and conformational classification (table adapted from ref. 41).

As more crystal structures were solved, large conformational heterogeneity was observed between the enzymes, as well as for different monomers of the same enzyme. In both UbiD and AroY, the N-terminal and the oligomerization domains were further away from each other than in FDC and appeared to be related via a hinge motion across the central α-helix.[11, 31] Differences were also observed in the active site structure, which is at the interface of the two domains. For UbiD and AroY, the active site can be classified as more 'open' compared to the 'closed' conformer in FDC. Based on the 'open' structure of apo-HudA and the 'closed' structure obtained for holo-HudA, it was suggested that domain motion is associated with cofactor binding.[33] However, for the VdcCD complex, both 'open' and 'closed' conformers were observed for the apo-enzyme (Figure 1.12A).[30] The two conformers were formally characterized based on the distance between

centers-of-masses of R168/173 on the N-terminal domain and L425/439 on the oligomerization domain (VdcCD/*An*FDC nomenclature) (Figure 1.12B). An R–L distance of ~10Å or less was considered as 'closed' whereas >11 Å was determined to be 'open'. A meta-analysis based on this classification discovered that most UbiD-enzymes occupy and 'open' conformation (Figure 1.12C).[41] An important consideration for this analysis is that the average β-factor varied between the 'open' and 'closed' crystal structures (56.9 and 130.56 $Å^2$ respectively) and therefore, significant errors could be associated with the R–L distance measurements.

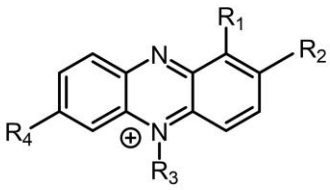
Domain motions appear to play an important role in catalysis. Recently, stopped-flow spectroscopy revealed that the dimeric FDC exhibited half-of-sites reactivity owing to negative cooperativity between the two subunits.[37] The negative cooperativity was proposed to be a result of domain motions, where substrate binding can lead to conformational changes interconverting the 'fast, tight' and the 'slow, loose' active sites.[42] Binding simulations of reaction intermediates for various R–L distances in VdcCD showed that while **Int2** (refer to Figure 1.9) prefers a more 'open' conformer, **Int3** and **E.S** complexes preferred smaller R–L distances.[30] Thus, it was proposed that domain closure following the formation of **Int2** might provide stability for **Int3** and help overcome the barrier associated with dearomatization of **Int2**.

1.6 PhdA – a novel UbiD-like enzyme

Recently, a novel UbiD-like enzyme was discovered in the soil bacterium *Mycolicibacterium fortuitum*. Named as PhdA, this enzyme decarboxylates phenazine-1-carboxylic acid (PCA) to phenazine.[36] Phenazines are redox active secondary metabolites secreted by *Pseudomonas* spp. in addition to other species, and consist of a central heteroaromatic pyrazine core with benzene rings appended on each side (Figure 1.13A).[43] Phenazine secretors functionalize the central aromatic core to produce a range of compounds that not only exhibit a variety of colors

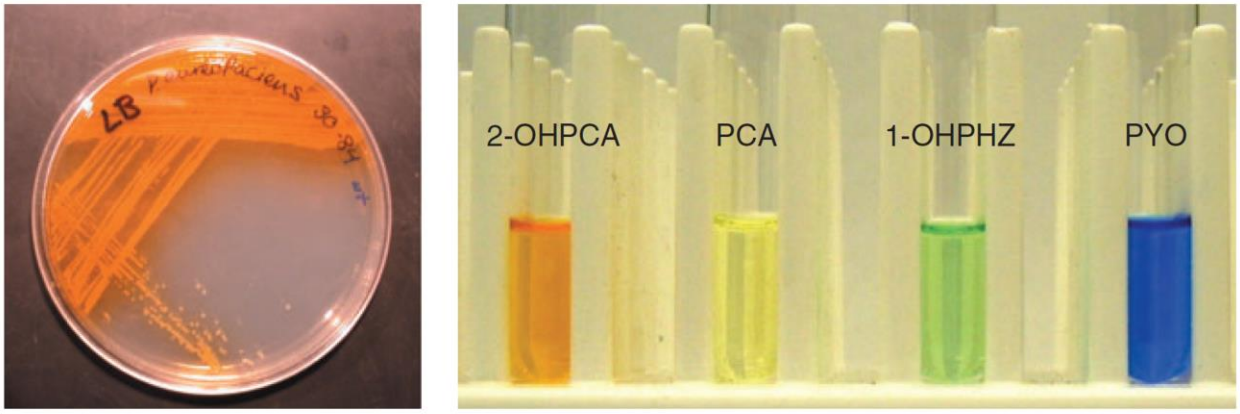
but also differ in their physical and redox properties (Figure 1.13B). These carefully modulated metabolites play an important role in the survival of pseudomonads under different conditions.

A



Name	R_1	R_2	R_3	R_4
Aeruginosia A	-COOH		$-CH_3$	$-NH_2$
2-Hydroxyphenazine-1-carboxylic acid (2-OHPCA)	-COOH	-OH		
Phenazine-1-carboxylic acid (PCA)	-COOH			
1-Hydroxyphenazine (1-OHPHZ)	-OH			
Pyocyanin (PYO)	-OH		$-CH_3$	

B



C

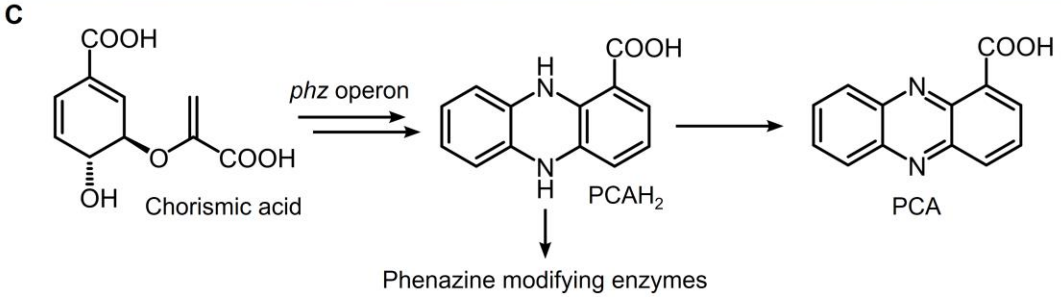


Figure 1.13: Biosynthesis and diversity of phenazines. **A:** Summary of different phenazines produced by pseudomonads. **B:** *Left* – *Pseudomonas aureofaciens* secreting orange colored 2-OHPCA, *right* – different colors exhibited by phenazines described in **A** (reproduced from ref. 44). **C:** PCA biosynthesis begins from chorismic acid. Reduced PCA ($PCAH_2$) acts as a substrate for phenazine modifying enzymes, making it the precursor to most phenazines.

The redox active nature of phenazines allows them to be terminal electron acceptors in metabolism and maintain redox homeostasis.[43] Phenazines also act as signaling molecules for quorum sensing and biofilm formation.[44] Moreover, they appear to play an important role in the virulence of *Pseudomonas aeruginosa*-based lung infections and *P. aeruginosa* strains unable to produce phenazines are severely attenuated.[43] On the other hand, pseudomonads found in the

rhizosphere secrete phenazines as antibiotics and protect crops from infections.[36] In fact, degradation of phenazines appears to make plants more prone to disease.

The biosynthesis of phenazines branches from the shikimate pathway, where the gene products of the *phz* operon convert chorismate to phenazine-1-carboxylic acid (PCA) (Figure 1.13C). Diversification of phenazines is proposed to start from PCA since its reduced form ($PCAH_2$) is readily used by many phenazine modifying enzymes as a substrate. Therefore, PCA is the target of phenazine degrading bacteria that compete with pseudomonads, such as *M. fortuitum*.

While *M. fortuitum* can thrive on PCA as the sole source of carbon,[45] the enzymes involved in its degradation were not known. In 2018, Costa and co-workers cultured *M. fortuitum* cells anaerobically with PCA as the only carbon source and observed the build-up of phenazine, confirming that the latter is an intermediate in PCA degradation.[36] This suggests that the first enzyme in the pathway is most likely a decarboxylase. When all the annotated decarboxylases from *M. fortuitum* genome were individually transformed in *E. coli*, only cells expressing the gene product of XA26_16650 could decarboxylate PCA. Moreover, a *M. fortuitum* ΔXA26_16650 strain couldn't grow on PCA as the sole carbon source, but could be rescued with an external vector expressing XA26_16650. XA26_16650 is annotated as a UbiD-like decarboxylase and analyzing its locus revealed an operon with two additional genes, XA26_16670, annotated as a UbiX-like prenyltransferase and XA26_16660, whose identity was unknown. Since mycobacteria don't synthesize ubiquinone, XA26_16650 was named as phenazine-degrading decarboxylase (PhdA) and XA26_16670 as PhdB. Bioinformatic analysis showed that the '*phd* operon' is present in two other PCA degrading *Actinobacteria*, *Nocardia* sp. strain LAM0056 and *Rhodococcus* sp. strain JVH1 but not in members incapable of PCA-degradation, further confirming the role of this operon.

Characterization of purified PhdA showed that upon reconstitution with prFMN, storage at an alkaline pH of 9.2 was required. When the protein was stored at pH 7.2, it lost activity over time. Surprisingly, neither oxidized nor reduced PCA appeared to be substrates, but incubating the enzyme with one electron reducing agents such as sodium dithionite or radical mediators such as paraquat led to significantly enhanced activity.

1.7 Goals and scope

The crucial role played by decarboxylases in metabolism and their potential as biocatalysts for sustainable synthesis of commodity chemicals emphasizes the need to study these chemically versatile enzymes. The newly discovered family of UbiD-like decarboxylases is rapidly expanding as unknown enzymes are discovered and previously misannotated enzymes are characterized. FDC has been the subject of intense mechanistic and biocatalytic studies, including development of novel assays to perform directed evolution,[46] protein engineering to expand substrate scope, as well as cascade reactions to diversify products.[16] However, if we focus all our efforts on researching a single enzyme in a diverse, unexplored family, we are missing out on the vast chemical space available to us. UbiD-like enzymes decarboxylate α,β-unsaturated, phenolic, heteroaromatic and even polyaromatic acids. Moreover, this enzyme family utilizes a previously unknown modified flavin cofactor (prFMN), whose chemical reactivity is not yet completely explored. Thus, studying novel enzymes in this family is exciting not only from a biocatalytic but also from a scholarly and academic standpoint.

PhdA is a newly discovered member of UbiD-like enzymes from *M. fortuitum* that decarboxylates phenazine-1-carboxylic acid (PCA).[36] As discussed in section 1.6, phenazines are redox active metabolites produced by pseudomonads that serve diverse functions. PhdA-mediated PCA degradation is proposed to provide *M. fortuitum* a competitive advantage over pseudomonads

in soil. Given the structure of PCA and relative substrate promiscuity exhibited by UbiD-like enzymes, it would be interesting to see if PhdA can utilize other polyaromatic and heteroaromatic carboxylic acids as substrates, including perhaps anthracenes and acridines. Moreover, PCA and prFMN can both exist in different redox states, including electron rich hydroquinone, radical semiquinone and electron deficient oxidized states.[27,47] Therefore, the reaction catalyzed by PhdA could occur through 1,3-dipolar cycloaddition if oxidized PCA is the substrate, or alternatively through electrophilic addition if reduced PCA is the substrate. Moreover, the initial characterization of PhdA revealed a dependence on one-electron reductants and radical mediators. This is particularly interesting and suggests that PCA semiquinone and/or prFMN^{radical} could exist as reaction intermediates.

Perhaps the biggest hurdle to extensive mechanistic and biocatalytic studies on UbiD-like enzymes is their recalcitrant nature. Apart from FDC and HudA, most enzymes cannot be purified in holo-form by simply co-expressing with UbiX, therefore a cumbersome *in vitro* reconstitution process is necessary. Even then, co-crystallization with substrate and/or inhibitors may be unsuccessful. Structural studies reveal the existence of distinct 'open' and 'closed' conformers. While binding simulations suggest that the 'closed' conformer is catalytically important, most UbiD-like enzymes seem to adapt the 'open' conformer. Moreover, the existence of an 'open' conformer in crystal structures is correlated to poor cofactor uptake and unsuccessful co-crystallization with substrates. There is a need to develop alternate methods to study UbiD-like enzymes.

Fortunately, many of these enzymes have a diverse substrate scope and robust activity assays. In fact, mechanistic proposals of several enzymes are purely based on substrate scope and computational studies.[11, 33, 34] Activity assays and substrate diversity can provide an in-depth

understanding of the steady state parameters, V_{max} and V_{max}/K_M. Investigating how these parameters vary as a function of the solvent isotope, substrate isotope, solvent viscosity and temperature can provide information on the nature and location of the rate limiting step(s).[48-50] Hence, another goal of my studies is to develop assays for studying the kinetic mechanisms of UbiD-like enzymes.

The peculiar cofactor, prFMN, is at the heart of UbiD-chemistry. However, oxidative maturation of this cofactor is still not well understood. While we know that incubation of UbiD-like enzymes with $prFMNH_2$ and subsequent oxidation leads to formation of $prFMN^{iminium}$, the exact mechanism is not known. A prFMN-C4a-OOH moiety is proposed as an intermediate in the FDC mediated maturation process. The same species is found to exist if free $prFMNH_2$ is oxidized in the absence of any enzyme. However, adding oxidized free $prFMNH_2$ to apo-FDC doesn't lead to activity, which contradicts prFMN-C4a-OOH being an intermediate in the maturation process. This is rationalized by suggesting that FDC-mediated transient formation of a prFMN-C4a-OO^- is necessary for successful maturation, which won't occur if free prFMN-C4a-OOH is added to apo-FDC. Further experiments are required to validate this claim and an in-depth investigation is essential to gain a better understanding of the process. Therefore, the final chapter of this dissertation discusses the biosynthesis and maturation of prFMN in detail.

Through these studies, I hope to expand our knowledge of the UbiD-enzyme family and the prFMN cofactor. This research will support future scholarship in the field of biocatalysis, decarboxylase chemistry and UbiD-like enzymes.

Chapter 2 Decarboxylation of Aromatic Carboxylic Acids by the Prenylated-FMN-Dependent Enzyme Phenazine-1-Carboxylic Acid Decarboxylase[1]

2.1 Introduction

Phenazine-degrading decarboxylase (PhdA) is a recently discovered prFMN-dependent enzyme that catalyzes the decarboxylation of phenazine-1-carboxylic acid (PCA) (Figure 2.1A).[36] Phenazines are redox-active metabolites that are secreted by a wide variety of bacteria,[43] for which PCA serves as the precursor. For organisms such as *Pseudomonas* spp, phenazine secretion confers a competitive advantage by facilitating anoxic survival and biofilm formation, thereby inhibiting other microbes. Phenazine-producing microbes are found in clinical, environmental, and agricultural contexts. Clinically, infection with *P. aeruginosa* poses a serious health risk, in part because the biofilms it forms, render the bacterium resistant to antibiotic treatment.[51] In agriculture, phenazines secreted by *Pseudomonas* spp. are important in biocontrol, where they protect cereal crops from a variety of fungal and parasitic diseases.[44] Degradation of phenazines by competing bacteria renders the plants more susceptible to infection.

Being redox active, phenazines can undergo two step-wise single electron reductions or a single step two-electron reduction, depending on the pH.[47] Thus, most phenazines, including PCA, can exist in oxidized, reduced and radical semiquinone states (Figure 2.1B). Hence, the reaction catalyzed by PhdA is intriguing because, *a priori*, the decarboxylation reaction could occur by

[1] The work presented in Chapter 2 is partially adapted from: Datar, P. M. and Marsh, E. N. G; Decarboxylation of Aromatic Carboxylic Acids by the Prenylated-FMN-dependent Enzyme Phenazine-1-carboxylic Acid Decarboxylase, *ACS Catal.* **2021**, 11, 11723−11732

either the cycloaddition mechanism used by FDC (if oxidized PCA is the substrate) or by the electrophilic addition mechanism used by AroY (if reduced PCA is the substrate). A recent report[36] described the initial characterization of PhdA from *Mycobacterium fortuitum*. The authors noticed that upon reconstitution of apo-PhdA with prFMN at pH 7, the enzyme was inactive within 2 hrs. However, if stored at pH 9.2, the activity was retained for several days. Furthermore, it appeared that both the fully reduced and the fully oxidized forms of PCA were poor substrates for PhdA. Instead, maximum decarboxylation was observed when PhdA and PCA were incubated with sub-stoichiometric amounts (w.r.t. substrate) of the one-electron reducing agent sodium dithionite. It was also noticed that the activity was enhanced by the presence of paraquat radical, a common electron mediator.[36] These results are intriguing and suggest that PhdA catalyzed PCA degradation might occur through formation of radicals, with the semiquinone form of PCA being the substrate.

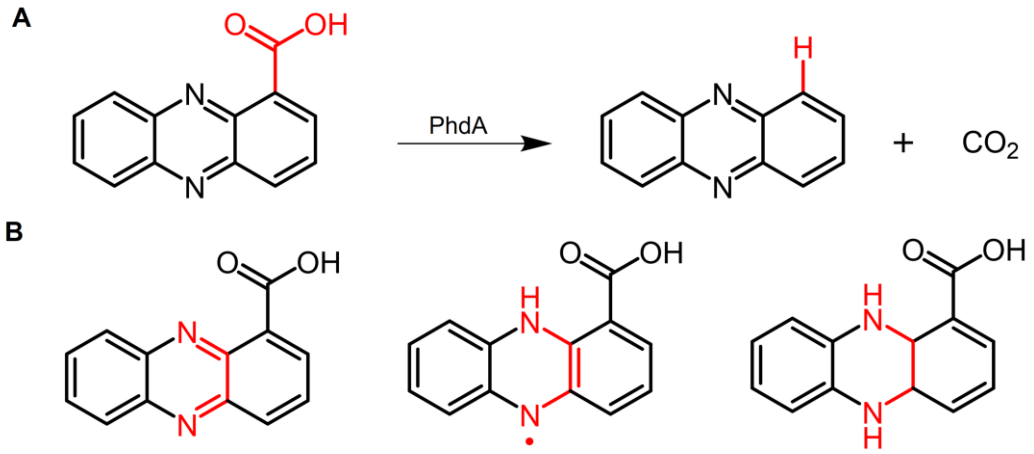


Figure 2.1: PhdA – a novel prFMN based decarboxylase. **A:** PhdA catalyzes the conversion of phenazine-1-carboxylic acid (PCA) to phenazine and CO_2. **B:** Possible redox states of PCA. *Left* – oxidized, *middle* – radical semiquinone, *right* – reduced.

Here we report a detailed characterization of PhdA, where we address some of the puzzling questions posed by the previous report. We have established conditions for reconstituting the enzyme in a highly active and stable form, which does not require reducing agents, and have investigated the substrate range for the enzyme. We show that PhdA will decarboxylate a number of polyaromatic compounds, including unfunctionalized anthracene carboxylic acids.

Additionally, we study the reversibility of PhdA by observing H/D exchange of the product, phenazine in buffered D_2O. We also attempt to improve the reverse reaction, carboxylation of phenazine, through various means.

2.2 Materials and Methods

2.2.1 Reagents and chemicals

Aromatic carboxylic acids and the corresponding compounds lacking carboxyl groups were purchased from Apollo Scientific Co., Sigma Aldrich Co., ChemScene, TCI Co., Thermo Fischer Scientific Co., 1 ClickChemistry Inc., Enamine Ltd., and used without further purification. All other reagents were purchased from Sigma Aldrich Co. or Thermo Fischer Co. Deuterated solvents were purchased from Cambridge Isotopes Laboratories Inc. or Thermo Fischer Co.

Reduced PCA was synthesized under anaerobic conditions by mixing 9 – 10 mM PCA with 230 mM sodium dithionite in H_2O. The resulting bright orange precipitate was collected by centrifugation and washed with H_2O to remove excess sodium dithionite. The presence of reduced PCA was confirmed by UV-Vis spectroscopy.[52]

2.2.2 Strains and plasmids

E. coli Rosetta strains with pET20b(+) vector individually containing either *phdA* or *phdB* were kindly provided by Prof. Dianne Newman (Caltech). The plasmids were purified using standard methods and transformed into *E. coli* BL21DE3 (Invitrogen) to facilitate protein expression. Codon-optimized genes encoding *ubiX* from *Pseudomonas aeruginosa* (*PaubiX*) or *E. Coli* (*EcubiX*) were designed by Dr. Nattapol Arunrattanamook, synthesized commercially and subcloned individually in pET28b(+) (GenScript Biotech Co.). pET28b(+) encoding a truncated

version of *S. cerevisiae PAD1* (*SctPAD1*) designed by Dr. Fengming Lin[2] was used without further modification.

For co-expression of PhdA and *Pa*UbiX from a single plasmid, their respective genes were PCR amplified using primers mentioned in Table 2.1 (PhdA F, PhdA R, UbiX F and UbiX R). Empty pMCSG7 expression vector kindly provided by Dr. Markos Koutmos was linearized with SspI and stitched to the PCR amplified genes through Gibson Assembly (New England Biolabs).

Alternatively, co-expression of PhdA with other UbiX homologs was achieved by co-transformation of pET20b(+)(AmpR) containing *phdA* and pET28b(+)(KanR) containing *PaubiX*, *EcubiX* or *SctPAD1* in *E. coli* BL21DE3 (Invitrogen). To facilitate co-transformation of *phdA* and *phdB*, *phdB* was PCR amplified from pET20b(+) vector (PhdB F and R primers, see Table 2.1) and cloned into pET28b(+) between NcoI and BamHI sites via Gibson Assembly. The resulting pET28b(+) vector containing *phdB* (without His tag) was co-transformed with pET20b(+) vector containing *phdA*.

Primer Name	Sequence (5' to 3')
PhdA F	TGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCAATATGCGGCATTACATCGACAC
PhdA R	CATATCTATATCTCCTTCTTAAAGTTAAACAAAGCTCGAGCTTCAGGCGAGCGGCAATG
UbiX F	TGAAGCTCGAGCTTTGTTTAACTTTAAGAAGGAGATATAGATATGAGCGGTCCGGAACG
UbiX R	CGACGGAGCTCGAATTCGGATCCGTTATCCACTTCCAATTTACTCGTCGCTGACAAGG
PhdB F	TTTTGTTTAACTTTAAGAAGGAGATATAC
PhdB R	CTTGTCGACGGAGCTCGAATTCGGATCCTCAGCTCTGGAAGTACAGGTTTTC

Table 2.1: PCR amplification primers used to construct plasmids for co-expression of PhdA with various UbiX homologs

The gene for carboxylic acid reductase (*CAR*) from *Tsukamurella paurometabola* (*TpCAR*) was synthesized commercially (GenScript) and cloned between NheI and EcoRI restriction sites of pET28b(+). Similarly, *sfp* from *Bacillus subtilis* (*Bssfp*) was cloned between NdeI and EcoRI

sites of pET20b(+). To obtain active *Tp*CAR enzyme, both expression vectors were co-transformed in *E. coli* BL21DE3.

2.2.3 Protein expression and purification

All *E. coli* BL21 DE3 strains were cultivated at 37 ˚C with shaking at 200 rpm in LB broth supplemented with 50 µg/ml Ampicillin [for pET20b(+)] and/or 50 µg/ml Kanamycin [for pET28b(+)]. After reaching an OD_{600} of 0.6 - 0.8, protein expression was induced by adding 0.1 mM IPTG. Wherever indicated, the medium was supplemented with 0.1 - 1 mL of prenol[53] and 1 mM $MnCl_2$. For anaerobic protein induction, the cultures were transferred to reagent bottles and capped tightly. Following an overnight incubation (20 ˚C, 170 rpm), cells were harvested by centrifugation (4˚C, 5000 rpm, and 15 min) and stored at -80˚C.

Proteins were purified by Ni-affinity chromatography. For anaerobic purification, all steps were performed in an anaerobic chamber (Coy) using N_2 purged buffers. Cells were resuspended in Buffer A (20 mM Tris-Cl pH 7.2 or pH 9.2, 500 mM KCl, 1 mM $MnCl_2$, 10 mM imidazole, 5% glycerol) supplemented with complete EDTA-free protease inhibitor cocktail (Roche) and sonicated using 3 sec pulses separated by 5 sec for a total time of 18 - 20 min. The lysate was clarified by centrifugation at (4 ˚C, 12000 rpm, and 45 min) and applied to a pre-equilibrated 5 mL HisTrap (GE Healthcare) column at a flow rate of 1 mL/min. Initially, the column was washed with 3-5 column volumes of Buffer A at 1 – 2 mL/min. Subsequently, a linear gradient of 0.1 M to 1.0 M imidazole was applied over 40-45 mL and the eluent was collected in 1 – 2 mL fractions. After SDS-PAGE analysis on a 10% polyacrylamide gel (Bio-Rad), relevant fractions of PhdA were combined, desalted (Bio-Rad 10-DG column) into Buffer C (20 mM Tris pH 7.2 or pH 9.2,

500 mM KCl, 1 mM $MnCl_2$, 5% glycerol) and stored. *Pa*UbiX was stored directly without desalting. All purified proteins were stored at -80 ˚C.

Co-expression and purification of *Bs*Sfp and *Tp*CAR was performed as previously described.[54]

2.2.4 Enzymatic synthesis of prFMN and reconstitution of PhdA

- All materials such as vials, desalting columns etc were transferred to an anaerobic chamber (Coy chamber with N_2 atmosphere containing 2.3% - 2.5% H_2 gas) and allowed to equilibrate overnight. All chemical reagents were deoxygenated by purging with N_2 overnight before transferring to the Coy chamber
- Sodium dithionite (DT) was freshly prepared under anaerobic conditions in H_2O. 300 – 400 μM FMN was reduced by titrating with DT (0.5 – 1 mM final concentration) and mixed with 2 mM DMAP in the presence of 20 – 30 μM *Pa*UbiX in 20 mM Tris-Cl, pH 7.2, 100 mM KCl and 5% glycerol. Reactions were conducted in O-ring capped vials (such as cryovials) to avoid oxidation from trace amounts of O_2 present in the Coy chamber
- Interestingly, *Pa*UbiX precipitated immediately but still catalyzed almost 100% conversion of FMN to prFMN as determined by HPLC analysis
- The UbiX reaction was incubated at room temperature in the dark for 4-5 hrs following which *Pa*UbiX was separated by centrifugation at 10,000 rpm for 15 min. The resulting supernatant was used to reconstitute PhdA
- 100 μL of 200 – 250 μM PhdA was reconstituted with 240 μL of supernatant from the UbiX reaction in 20 mM Tris-Cl, pH 7.2, 10 mM $MnCl_2$, 100 mM KCl and 5% glycerol. The final reaction volume was 400 μL and final concentrations of PhdA and prFMN were ~50 μM and

~240 μM respectively. The reconstitution was performed in cryovials and incubated at room temperature in the dark for 10-15 min

- Meanwhile, 4x100 μL Zeba spin desalting columns (Thermo Fischer Co.) were scrubbed with 300 μL of 1 mM DT and buffer exchanged with 3x300 μL of storage buffer (20 mM Tris-Cl, pH 7.2, 1 mM $MnCl_2$, 500 mM KCl and 5% glycerol) following manufacturers protocol
- 4x100 μL reconstituted PhdA was applied to each desalting column and centrifuged at 1500 rcf for 2 min. All 4 protein fractions were pooled into a single vial, brought out of the anaerobic chamber with the cap open and incubated at 4°C in the dark
- A Bradford assay was performed to measure protein concentration. After being exposed to O_2 for ~40 min, reconstituted, desalted PhdA was directly injected on HPLC to observe bound prFMN
- To verify the presence of prFMN, PhdA and holo-FDC purified from *E. coli* were also directly injected on HPLC. The protein and cofactor separated in line and provided distinct peaks
- Reconstituted, desalted PhdA was stored at -80°C after ~1.5 hrs following O_2 exposure in 10-20 μL aliquots

2.2.5 Enzymatic assays

All assays to test PhdA activity were performed at room temperature with 0.2 μM enzyme and 100 μM PCA in 20 mM Bis-Tris-Cl buffer (pH 6.5). Reactions were quenched at various time points by adding 100 mM NaOH (final concentration) and analysed using HPLC. Most activity assays were performed under aerobic conditions.

The following assays were conducted in an anaerobic chamber (Coy) with degassed buffers: (a) For verifying if reduced PCA is a substrate of PhdA, 0.2 μM of enzyme was incubated with 100 μM reduced PCA. (b) The effect of sodium dithionite on PhdA's activity was tested by

incubating different concentrations of the reducing agent with 0.2 μM enzyme in 20 mM Bis-Tris-Cl buffer (pH 6.5) for 5 – 10 min following which reactions were initiated with 100 μM PCA. (c) The influence of paraquat radical were also studied in a similar fashion. 1 mM paraquat dichloride was mixed with 0.5 mM sodium dithionite under anaerobic conditions. The presence of paraquat radical was evident by the rapid formation a dark blue solution. 0.2 μM PhdA was incubated with 100 μM freshly prepared paraquat radical in 20 mM Bis-Tris-Cl buffer (pH 6.5) for 10 min after which 100 μM PCA was added to the reaction.

To study the substrate scope of PhdA, 10 μM enzyme was incubated with 500 μM of the aromatic carboxylic acid under investigation in 20 mM Bis-Tris-Cl buffer, pH 6.5. Control reactions without PhdA were performed in parallel. Reactions were quenched after 17 hrs by adding an equal volume of acetonitrile, centrifuged at 14,000 rpm for 20 min and the supernatant was analyzed using HPLC.

To study the reverse reaction, 2-10 μM PhdA was incubated with 0.5-1 mM phenazine in 1-3 M $NaHCO_3$, NH_4HCO_3 or $KHCO_3$. The bicarbonates were titrated to pH ~ 7 and worked as both a source of CO_2 and a buffer. Reactions were quenched with equal volume of ACN and injected on HPLC for analysis. For biphasic reactions, 10 mM phenazine was dissolved in an organic solvent and carefully placed over the H_2O layer. After overnight incubation, phenazine was extracted in the organic layer, dried, and re-dissolved in 1:1 H_2O:ACN before injection on HPLC.

Standard reactions to measure the activity of *Tp*CAR were performed as previously described.[54] Reductions of various aromatic acids with *Tp*CAR were performed under similar conditions, but with 10x more enzyme compared to the standard assay and ~1mM substrate.

2.2.6 HPLC analysis

All HPLC analysis was performed on a Shimadzu Prominence LC-20AT series chromatography system equipped with a diode array detector. A phenomenex kinetex C18 column (5 μm particle size, 250 x 4.6 mm) was used to obtain separation at a flow rate of 0.5 $mL.min^{-1}$ and a detection wavelength of 360 nm (unless specified otherwise). The mobile phase consisted of 10 mM trifluoroacetic acid in water (buffer A) and 10 mM trifluoroacetic acid in acetonitrile (buffer B). Different methods were employed for different analytes.

For studying various flavins, the system was used at a flow rate of 0.4 ml/min. The gradient consisted of 5% buffer B for 5 min, 5% - 100% B over 25 min and 100% B for 5 min. This was followed by a re-equilibration at 5% B for 5 min.

For separating PCA and phenazine the column was subjected to 5% B over 1 min, 5 – 55% B over 1 min, held at 55% B for 3 min, a slow gradient from 55 – 60% B over 10 min, 60 – 95% B over 1 min, 95% B for 3 min and re-equilibration at 5% B for 6 min.

For separating compound **4** (refer to Table 2.2 for structures of all compounds) from the product, 2,3-dimethylquinoxaline, a similar gradient was employed with initial equilibration at 5% B for 1 min, 5 – 45% B over 9 min, held at 45% B for 5 min, 45 – 95% B over 1 min, held at 95% B over 4 min and re-equilibration at 5% B for 5 min. Detection was carried out at 310 nm. For studying conversion of compounds **5**, **6**, **7**, and **8** to their respective hydrocarbons, the system was equilibrated at 50% B for 5 min followed by a gradient from 50 – 100% B over 10 min, holding 100% B for 5 min and re-equilibrating to 50% B for 7 min.

For separating **1**, **2** and **3** from their respective products, the following modified solvent system was used: Buffer A consisted of 10 mM Tris-Cl pH 7.2 in water and buffer B was 10 mM Tris-Cl pH 7.2 in 60 % ACN:40 % H_2O mixture. The detection wavelength was 310 nm. Initially,

column was equilibrated with 10 % B for 5 min followed by a gradient from 10 – 100 % B over 7 min, holding 100 % B for 5 min and re-equilibration to 10% B over 5 min.

Aromatic acids were separated from their aldehydes by the same methods described above for the respective acids. Putative aldehyde peaks were verified by injecting a standard on HPLC or by obtaining their mass through LC-MS.

2.2.7 LC-MS analysis

An Agilent 1290 series LC system equipped with an Agilent 6545 quadrupole-TOF mass spectrometer was used for LC-MS analysis. Analytes were eluted with 0.1% formic acid in water (buffer A) and 0.1% formic acid in 95% acetonitrile, 5% water (buffer B). Unless specified otherwise, the type of column and the elution method used were identical to the HPLC conditions. Mass acquisition was carried out in positive ion mode from 50 – 1200 m/z.

To separate different hydrocarbons for H/D exchange reactions, an Agilent Zorbax Eclipse Plus C18 column (1.8 µm particle size, 50 x 2.1 mm) was used at a flow rate of 0.4 ml/min. The elution method consisted of 5% B for 1 min followed by a gradient from 5% - 95% B over 3 min, 95% B for 1 min and a final re-equilibration to 5% B over 1 min.

2.2.8 H/D exchange assays and NMR analysis

Reactions were performed in D_2O with 0.5 – 10% DMSO-d_6, 50 – 500 µM phenazine and 5 µM reconstituted PhdA. Control reactions without the enzyme were also performed. After overnight incubation, the reactions were quenched and analyzed by either LC-MS or NMR. To study deuterium exchange under low CO_2 levels, CO_2 was removed by bubbling argon gas for 6 - 8 hrs through the buffers, with gaseous CO_2 removed by an in-line 5 M KOH trap.

For NMR analysis, the reactions were quenched by adding an equal volume of $CDCl_3$, vortexed to extract the phenazine in $CDCl_3$ and centrifuged at 14,000 rpm for 10 min. The $CDCl_3$ layer was collected separately, dried with Na_2SO_4 and analyzed on a Varian MR 400 MHz spectrometer.

H/D exchange assays for substrate analogs were performed as follows: PhdA was buffer exchanged into 20 mM Bis-Tris-Cl pD 6.5 in D_2O. 0.5 – 1 mM substrate and 38 µM PhdA were incubated in 20 µL D_2O buffer overnight. Reactions were quenched with 3M HCl, diluted 10x in H_2O and purified using Pierce C18 tips (Thermo Fisher Scientific Co.) following manufacturers protocol. To separate ^{13}C and D isotopes, 5 – 10 µL of the purified sample was injected on an Orbitrap Lumos mass analyser (Thermo Fisher Scientific Co.) through a nano-ESI source in positive ion mode. Data was analyzed using Freestyle software (Thermo Fisher Scientific Co.)

2.3 Results

2.3.1 Initial purification attempts

Initially, pET20b(+) containing *phdA* + C-terminal His tag was expressed in *E. coli* BL21DE3 cells and purified through Ni-affinity chromatography (Figure 2.2A). Upon purification, PhdA was analysed by UV-Vis spectroscopy to observe bound cofactor, however, no spectra corresponding to prFMN could be detected. Furthermore, no evidence for PCA decarboxylation was found by HPLC, suggesting that PhdA was purified in the inactive apo form. Although *E. coli* expresses UbiX, the endogenous levels of the cofactor appear to be insufficient for obtaining holo-PhdA.

For some prFMN dependent enzymes like FDC, co-expression with the prFMN synthase, UbiX (or its yeast homolog, PAD1) was shown to be beneficial for obtaining holo-enzyme. Therefore, we designed a construct for co-expression of PhdA and a homolog of UbiX from

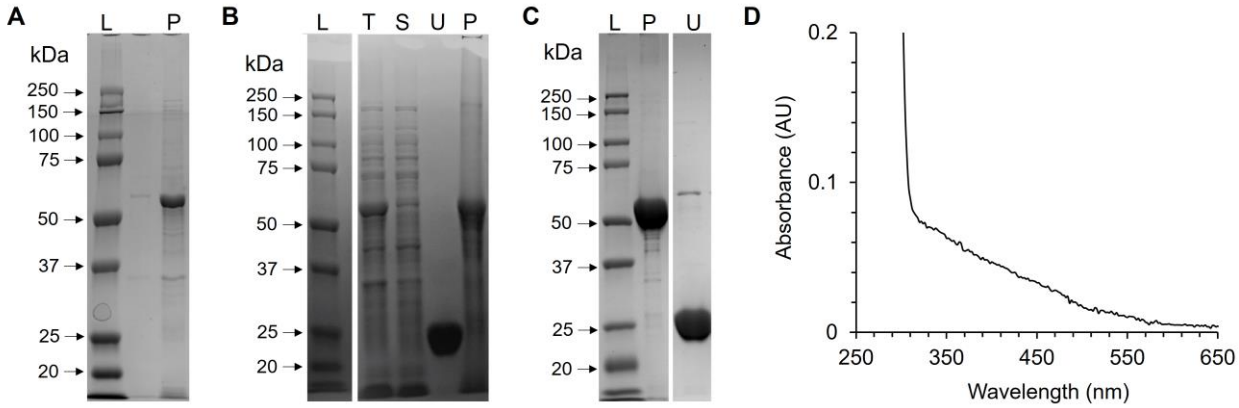


Figure 2.2: Purification of PhdA from *E. coli* BL21DE3. **A:** SDS-PAGE for apo-PhdA (53.2 kDa) with C-terminal His tag, purified by Ni-affinity chromatography. *L* – protein ladder, *P* – PhdA elution fraction (200 – 300 mM Imidazole). **B:** SDS-PAGE to study the co-expression of PhdA and *Pa*UbiX (26.4 kDa) cloned in pMCSG7 vector. *L* – protein ladder, *T* – total cell lysate, *S* – clarified lysate, *U* – *Pa*UbiX std., *P* – PhdA std. **C:** SDS-PAGE for Ni-affinity purification of PhdA co-transformed with *Pa*UbiX. *L* – protein ladder, *P* – PhdA elution fraction (200 – 300 mM Imidazole), *U* – *Pa*UbiX elution fraction (700 – 1000 mM Imidazole). **D:** UV-Vis spectrum for PhdA co-transformed with *Pa*UbiX.

Pseudomonas aeruginosa (*Pa*UbiX). The genes *phdA* and *PaubiX*, each with their respective ribosome binding sites, were cloned in pMCSG7 vector between the SspI site by Gibson assembly. *phdA* was cloned with a N-terminal His tag whereas no affinity tags were added for *PaubiX*. Unfortunately, no expression was detected for *Pa*UbiX and PhdA expressed in the insoluble fraction (Figure 2.2B).

Eventually, co-expression was achieved by co-transformation of pET20b(+)(AmpR) containing *phdA* and pET28b(+)(KanR) containing *PaubiX* in *E. coli* BL21DE3. To facilitate efficient synthesis of prFMN, the cultures were supplemented with prenol after inducing protein expression.[26] While both proteins were expressed with His-tags, an appreciable separation was obtained during Ni-affinity purification, with PhdA eluting between 200 – 300 mM imidazole and *Pa*UbiX eluting after 700mM imidazole (Figure 2.2C). Moreover, purified PhdA exhibited a broad UV-Vis absorption shoulder extending from ~ 300 – 550 nm (Figure 2.2D) and also catalyzed the decarboxylation of PCA. Therefore, all future purifications of PhdA were performed by co-transformation with *Pa*UbiX unless specified otherwise.

2.3.2 Dependence of PhdA activity on reducing agents

Initially, purification of PhdA was conducted at pH 9.2, following the protocol described by Costa and co-workers.[36] Under these conditions, the enzyme required incubation with sodium dithionite (> 50 µM) at pH 6.5 for activity, as previously reported (Figure 2.3A). However, in our hands, it proved hard to obtain consistent values for PhdA activity, and large variations in enzyme activity between individual assays were apparent. We also investigated whether, as previously reported, the electron mediator, paraquat, was necessary for activity and found only a marginal improvement in the presence of the paraquat radical (Figure 2.3B).

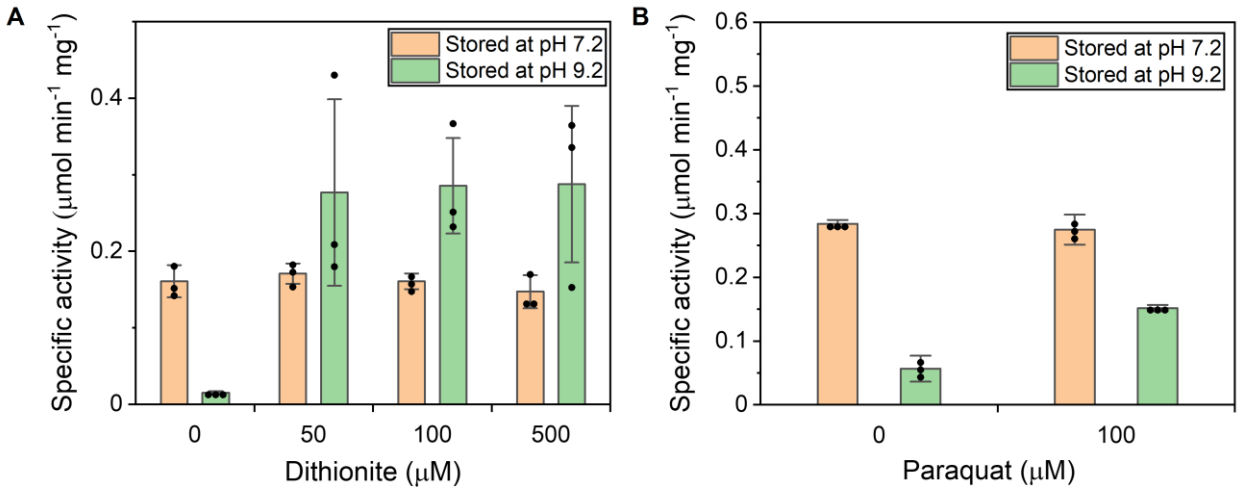


Figure 2.3: Dependence of PhdA activity on reducing agents. **A:** Specific activity of PhdA catalyzed PCA decarboxylation with varying concentrations of sodium dithionite, under anaerobic conditions. **B:** Specific activity of PCA decarboxylation in the presence of the paraquat radical in an anaerobic environment.

Because prFMN is sensitive to hydrolysis and is known to isomerize between the active iminium and inactive enamine and ketimine forms, [4, 28, 31]we suspected that the relatively high pH at which the enzyme was purified and stored might result in some degradation of the cofactor. Therefore, we repeated the purification of PhdA at pH 7.2. At this pH, the activity of PhdA was not affected by the presence of dithionite or paraquat radical in the assay, and was similar to that exhibited by the dithionite-treated enzyme purified at pH 9.2. Therefore, for further experiments,

the enzyme was purified and stored at pH 7.2 and activity assays were performed on oxidized PCA without adding any reducing agents.

The above analysis indicates that the presence of reducing agents or electron mediators is not necessary for PhdA catalyzed PCA decarboxylation, providing evidence against a mechanism involving radical intermediates.

2.3.3 Optimizing prFMN incorporation and in vitro reconstitution

Regardless of whether PhdA was purified at pH 7.2 or 9.2, the enzyme showed quite low levels of activity, suggesting that not all the active site contained the cofactor. Maturation and installation of prFMN is still poorly understood and has been reported to be problematic for many UbiD-like enzymes.[11, 30, 31] Moreover, in addition to the presence of FMN, multiple prFMN species are observed during maturation[27] that precludes a definitive analysis of bound cofactor by UV-Vis. Therefore, we developed an HPLC based assay to analyze the flavin content of PhdA. As a reference, we also measured the prFMN bound to FDC, for which the cofactor is efficiently installed upon co-expression with a prFMN synthase.[2] Reference chromatographs (at 360 nm absorbance) for this analysis are shown in Figure 2.4B. The active form of prFMN can be seen eluting at a retention time of ~18.9 min whereas the peak at ~ 17.5 min is for FMN, based on a standard. The identities of these peaks were further confirmed by LC-MS. Compared to FDC, PhdA has significantly lower prFMN bound, even though equal amounts of both enzymes (20 µL of 50 µM) are injected. Additionally, we observe that other chromophores, possibly degraded forms of prFMN, are present.

Amongst several attempts to improve the fraction of bound prFMN to PhdA, we: (i) Supplemented PhdA + *Pa*UbiX cultures with different concentrations of prenol to boost prFMN biosynthesis; (ii) Expressed and/or purified PhdA anaerobically owing to the air sensitivity of

prFMN (iii) Co-expressed PhdA with other UbiX homologs, such as *E. coli* UbiX (*Ec*UbiX), *S. cerevisiae* tPAD1 (*Sc*tPAD1) and even PhdB, the native prFMN synthase from *M. fortuitum*; (iv) Purified PhdA via ion exchange chromatography (IEX) to prevent the possibility of Ni^{+2} from the HisTrap column interfering with the Mn^{+2}-mediated binding of prFMN to PhdA. For each case, the prFMN content was analyzed by HPLC. Unfortunately, none of the perturbations resulted in improving the amount of bound cofactor. We therefore turned to reconstituting holo-PhdA *in vitro* using enzymatically synthesized prFMN.

prFMN was synthesized from reduced FMN and dimethylallyl phosphate (DMAP) using *Pa*UbiX,[26] as described in section 2.2.4. The presence of prFMN was confirmed by HPLC (Figure 2.4A). After removing UbiX, PhdA (50 μM, final concentration) was added and incubated under anaerobic conditions for 10-15 min at room temperature. Subsequently, PhdA was purified from the reaction mixture by desalting into fresh buffer and the reduced cofactor was allowed to oxidize in air to produce the active form of prFMN. HPLC analysis of reconstituted PhdA (Figure 2.4B) revealed a significant increase in the peak at ~18.9 min for active prFMN (peak P), along with two other peaks at ~20.2 min (peak Q) and ~21.8 min (peak R) respectively. The mass (m/z = 525.17) and UV-Vis spectrum of peak P match with the sample from FDC (Figure 2.4D); the mass (m/z = 526.18) and UV-Vis spectrum of peak Q (Figure 2.4E) suggests that is the stable prFMN radical which is known to be formed as an off-pathway byproduct of prFMN maturation;[31] the mass (m/z = 527.19) of and UV-Vis spectrum of peak R (Figure 2.4F) indicates it is a form of prFMN, the structure of which remains unknown.

The reconstituted enzyme obtained by this method was ~8-fold more active than the "holo"-PhdA initially purified from *E. coli* (Figure 2.5A). Interestingly, the final desalting of PhdA

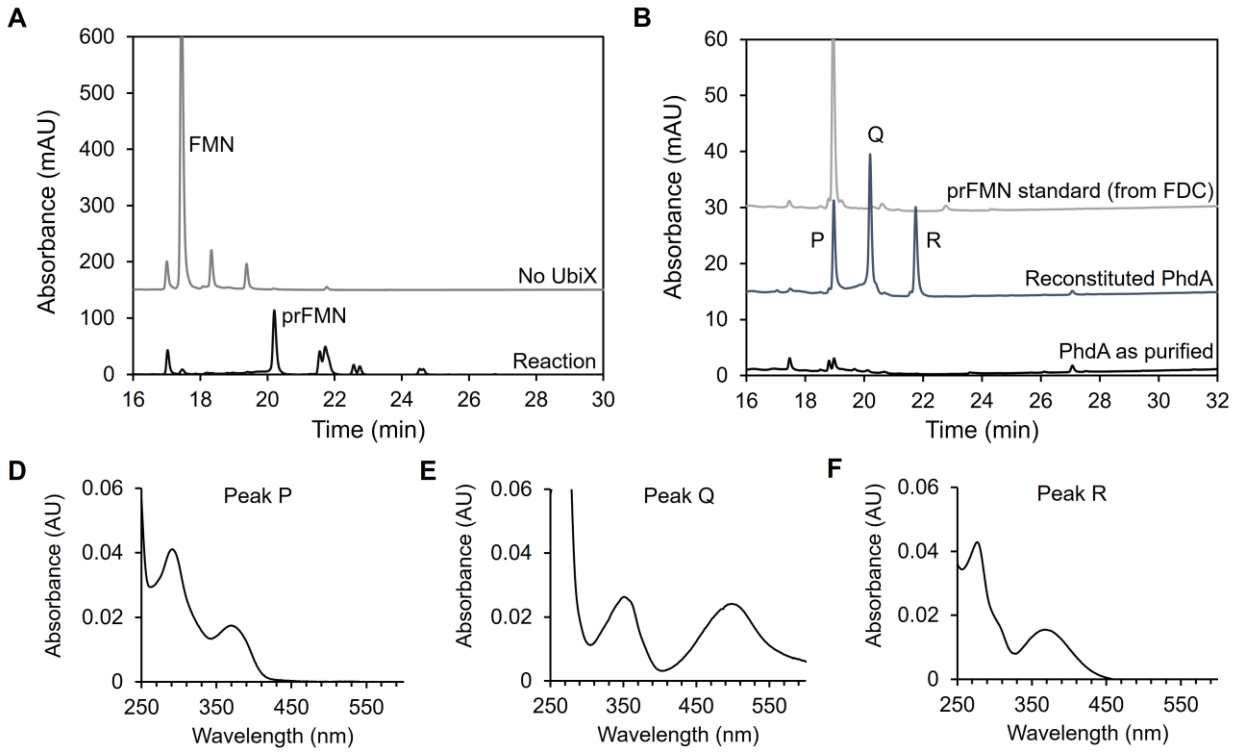


Figure 2.4: Reconstitution of PhdA. **A:** HPLC chromatograph for UbiX catalyzed synthesis of prFMN. *Top* – HPLC trace without UbiX, *bottom* – trace for the complete reaction. **B:** Chromatographs for PhdA reconstitution. *Top* – holo-FDC injected to display prFMN standard, *middle* – reconstituted PhdA with different prFMN forms at 18.9 (peak P), 20.2 (peak Q) and 21.8 min (peak R), *bottom* – 'holo' PhdA co-transformed with UbiX in *E. coli*. **D, E and F:** UV-Vis spectra for peaks P, Q and R respectively.

prior to oxidation proved to be essential to produce highly active enzyme; if this step was omitted, reconstituted PhdA showed no increase in activity.

2.3.4 Steady state kinetics and reactivity with reduced PCA

Having established the conditions for reconstituting PhdA we examined the reaction in more detail. Using the reconstituted enzyme, we measured the steady state kinetic parameters for the decarboxylation of PCA (Figure 2.5B); we determined the apparent k_{cat} to be 2.6 ± 0.1 s^{-1} (155 ± 4 min^{-1}) and the K_M for PCA to be 53 ± 2 µM. Given that our cofactor analysis showed that prFMN is present in only a fraction of the active sites, the true k_{cat} is likely several fold higher.

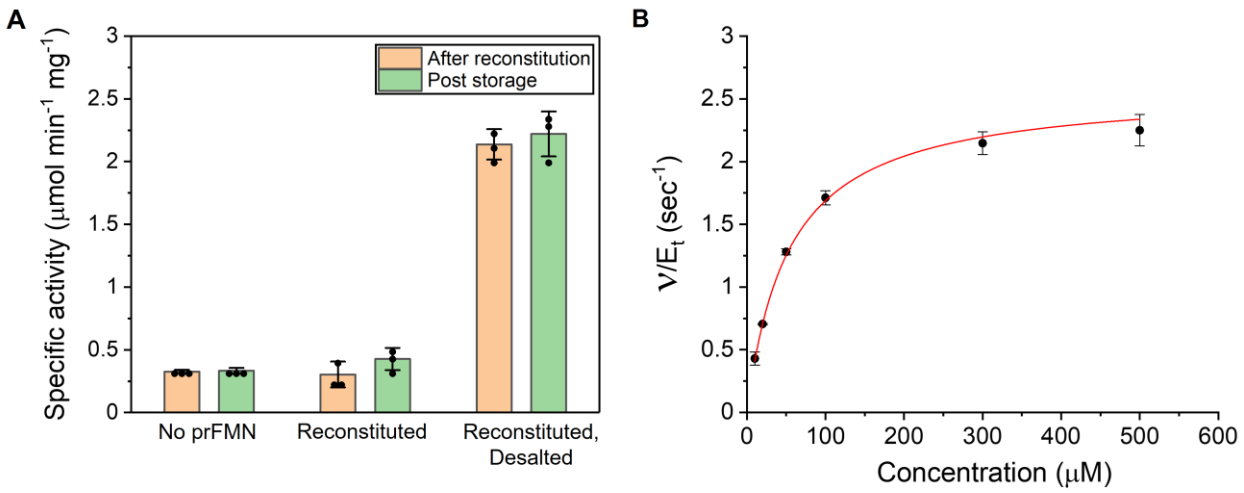


Figure 2.5: Kinetics of reconstituted PhdA. **A:** Specific activity for different enzyme fractions of PhdA with oxidized PCA. **B:** Steady state kinetics for PhdA catalyzed PCA decarboxylation.

The initial report describing the characterization of PhdA left unresolved which oxidation state(s) of PCA are substrates for decarboxylation.[36] As discussed earlier, the reduced form of PCA is an electron rich system which can, in principle, undergo an electrophilic addition reaction similar to the mechanism proposed for AroY. To resolve the ambiguity, we prepared the fully reduced form of PCA and evaluated it as a substrate for PhdA under anaerobic conditions, as the compound is readily oxidized in air. Under these conditions, only a trace amount of phenazine was detected, which we consider was most likely due to the slow oxidation of reduced PCA under micro-aerobic conditions followed by decarboxylation. Therefore, we consider that only the oxidized form of PCA is a substrate for decarboxylation.

2.3.5 Substrate scope of PhdA

We next examined the substrate range of PhdA, with the objective of determining what features of the phenazine ring system were important for substrate recognition and reactivity. The various potential aromatic carboxylic acid substrates listed in Table 2.2 (500 μM final concentration) were incubated with 10 μM PhdA in reaction buffer at room temperature for 17 hours. The fraction of substrate decarboxylated was then determined by HPLC analysis and also

compared with % conversion for PCA. To obtain accurate % conversion values, HPLC peak calibrations were performed for all hydrocarbon products.

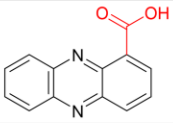
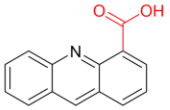
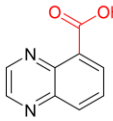
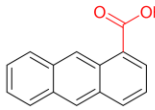
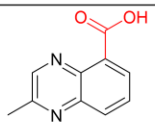
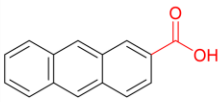
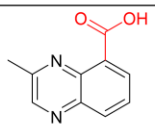
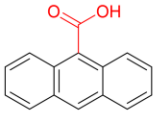
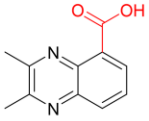
Compound	Substrate	% Conversion	Compound	Substrate	% Conversion
PCA		100	**5**		55.5 ± 3.9
1		0.14 ± 0.01	**6**		15.0 ± 1.7
2		10.2 ± 0.2	**7**		0
3		9.7 ± 0.4	**8**		0.54 ± 0.06
4		76.5 ± 6.6			

Table 2.2: Substrate scope of PhdA catalyzed decarboxylation. % conversions for each compound are measured under identical conditions and are averages of two independent measurements.

Quinoxaline-5-carboxylic acid (**1**), which lacks the distal phenyl ring, proved to be a very poor substrate, with only 0.1 % of the compound undergoing decarboxylation under the conditions of the reaction. However, addition of a methyl group at either the 2 or 3 positions, markedly improved reactivity. Both 2-methyl-quinoxaline-5-carboxylic (**2**) and 3-methyl-quinoxaline-5-carboxylic acid (**3**) underwent 10 % decarboxylation to 2-methylquinoxaline. 2,3-dimethylquinoxaline-5-carboxylic acid (**4**) proved an even better substrate, undergoing 76% decarboxylation to 2, 3-dimethylquinoxaline. These results demonstrate that the extended aromatic system of the phenazine nucleus is not required for the reaction; rather, the distal ring likely contributes more to substrate recognition, as its steric bulk can be substituted with methyl groups.

Furthermore, we examined the contribution of the heterocyclic nitrogen atoms to the reactivity of the substrate. The isosteric compound, acridine-4-carboxylic acid (**5**) (which contains

only one nitrogen atom) proved to be substrate for PhdA and was decarboxylated to a moderate extent (56 %), thereby demonstrating that the quinoxaline functionality is not required for activity. We extended this line of investigation to examine the reactivity of PhdA with un-activated polyaromatic compounds. No reaction was observed with naphthalene-1-carboxylic acid or naphthalene-2-carboxylic acid. However, anthracene-1-carboxylic acid (**6**) proved surprisingly reactive and underwent 15 % decarboxylation under the conditions of the reaction. We further investigated the regioselectivity of PhdA decarboxylation using anthracene carboxylic acids. No reaction was observed with anthracene-2-carboxylic acid (**7**), however, interestingly, a small amount of decarboxylation (0.5 %) was observed for anthracene-9-carboxylic acid (**8**), which suggests an alternate mode of substrate binding. Other acids tested in this study that didn't show any reactivity with PhdA include: fluorene-4-carboxylic acid, fluorene-1-carboxylic acid and benzoic acid.

2.3.6 Optimizing the carboxylation reaction

Initially, phenazine was incubated overnight with PhdA and a large excess of bicarbonate as a CO_2 source, following which the reaction was analyzed by HPLC. A very small amount of PCA corresponding to < 0.1% conversion was formed in the reaction (Figure 2.6A), suggesting that the equilibrium constant for the reaction heavily favors decarboxylation. Additionally, the lower solubility of phenazine (0.5 – 1 mM) further affected the outcome.

To favor the equilibrium towards carboxylation by Le Chatelier's principle, we attempted to improve the solubility of phenazine through addition of co-solvents and detergents, biphasic reactions or supplementing with cyclodextrins. However these changes were either detrimental to the stability of PhdA or effectively reduced the amount of available phenazine, providing no improvement in carboxylation. Similarly, efforts to increase the concentration of CO_2 by using

different bicarbonate salts or pressurized CO_2 (through a septum and a CO_2 balloon) were also unsuccessful.

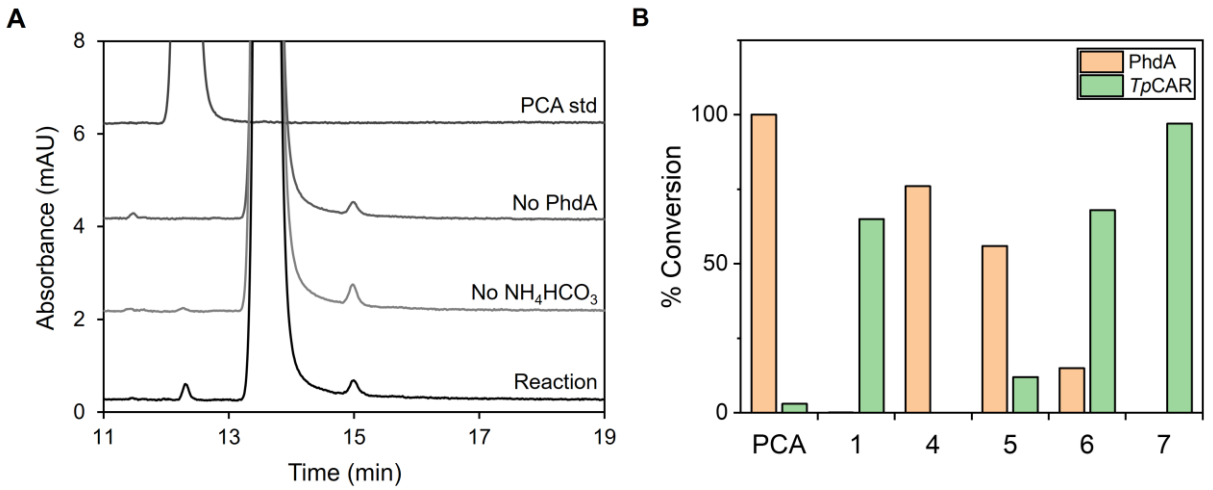


Figure 2.6: Optimizing PhdA catalyzed carboxylation. **A:** Representative chromatograph for PhdA catalyzed carboxylation of phenazine. **B:** Incompatible substrate scope of PhdA and *Tp*CAR. % conversions for PhdA catalyzed decarboxylation and *Tp*CAR catalyzed reduction are plotted for PCA as well as compounds 1, 4, 5, 6 and 7 (refer to Table 2.2)

In section 1.2.2, we described how the reversibility of UbiD-like enzymes has been improved by coupling carboxylation to thermodynamically favorable reductions performed by a carboxylic acid reductase (CAR).[16] Since reduction is essentially irreversible under the reaction conditions, according to Le Chatelier's principle, it shifts the reaction towards carboxylation by removing the product (carboxylic acid) from the chemical equilibrium. Therefore, CAR from *Tsukamurella paurometabola* (*TpCAR*) was co-expressed with *Bacillus subtilis* Sfp (*Bs*Sfp) and purified as previously described.[54] The turnover number for purified *Tp*CAR with benzoic acid (102 min^{-1}) was comparable to the value reported previously (~140 min^{-1}). However, upon studying *Tp*CAR-catalyzed reduction of aromatic acids described in Table 2.2, we discovered that the two enzymes have incompatible substrate scope (Figure 2.6B). In accordance with this observation, PhdA + *Tp*CAR coupled assays led to no detectable products in the reverse reaction.

2.3.7 Studying PhdA catalyzed H/D exchange

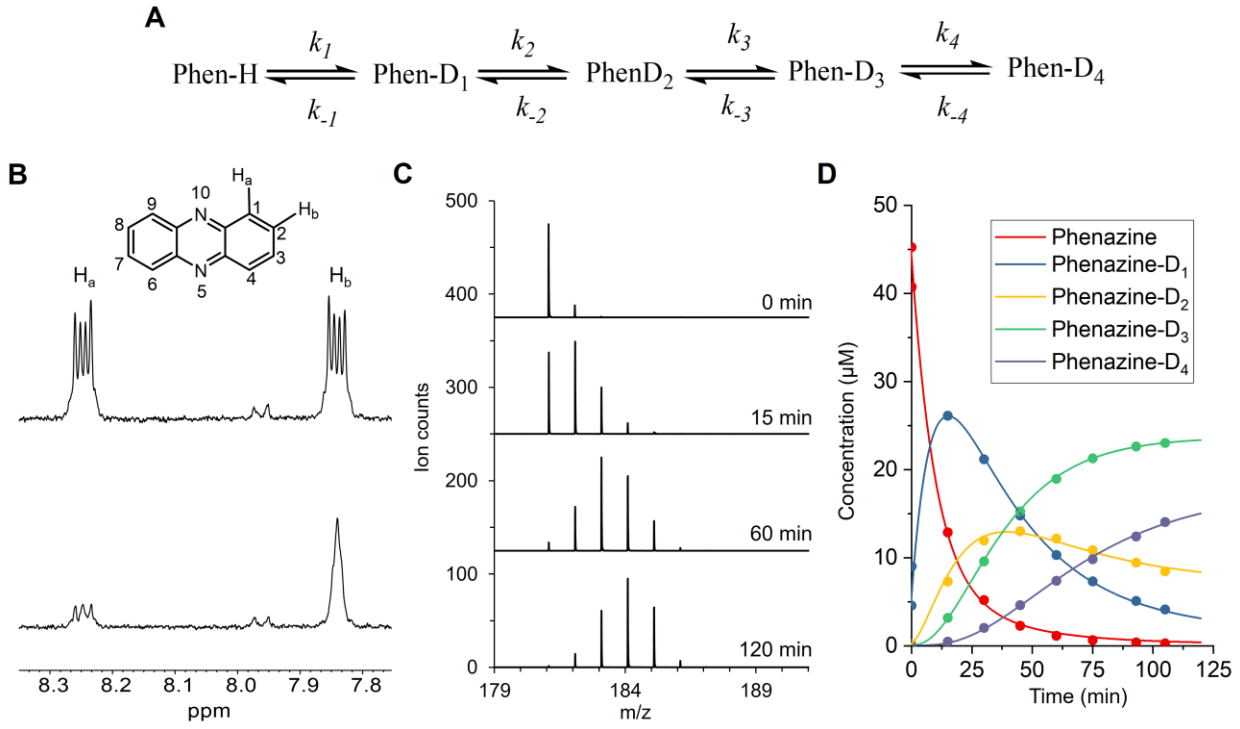


Figure 2.7: Deuterium exchange into phenazine catalyzed by PhdA. **A:** Scheme depicting the 4-step sequential model used to fit the H/D exchange data. **B:** ^{1}H-NMR spectrum for phenazine before (*top*) and after (*bottom*) incubating with PhdA overnight in buffered D_2O. **C:** Representative mass spectra monitoring the time course for deuterium exchange in phenazine. **D:** Time course for deuterium exchange into phenazine determined by peak integration of MS data; the data are fitted to the model depicted in **A** using KinTek Explorer. The forward (odd numbered) and reverse (even numbered) rate constants obtained for the individual steps are as follows (in min^{-1}): k_1 = 0.0921 ± 0.0063, k_{-1} = 0.0084 ± 0.0035, k_2 = 0.0381 ± 0.0023, k_{-2} = 0.008 ± 0.0036, k_3 = 0.0624 ± 0.0058, k_{-3} = 0.0188 ± 0.0038, k_4 = 0.0208 ± 0.0042, k_{-4} = 0.028 ± 0.0095.

Mechanistically, carboxylation involves breaking a C–H bond. Thus, we sought an alternate route to study the reverse reaction – to observed H/D exchange of phenazine with PhdA in buffered D_2O. The enzyme readily exchanged up to 4 deuterium atoms into phenazine, as determined by ^{1}H-NMR. Furthermore, our analysis confirmed that, as expected, the chemically equivalent protons at positions 1, 4, 6 and 9 were exchanged, as is evident from the loss of the doublet-of-doublets at 8.25 ppm and simplification of the doublet-of-doublets at 7.85 ppm to a singlet (Figure 2.7B). In contrast to FDC, where CO_2 appears to be required for deuterium exchange into styrene,[38] removing dissolved CO_2 from the buffer did not appear to affect

deuterium exchange into phenazine by PhdA. The time course for deuterium exchange was monitored using LC-MS, which allowed the formation mono-, di-, tri-, and tetra-deuterated phenazine to be followed (Figure 2.7C). These data were well fitted to a 4-step sequential kinetic model (Figure 2.7A, D) yielding rate constants for every step, based on which the rate of exchange of the first deuterium was calculated (4.13 ± 0.27 μM/min) and normalized with the enzyme concentration to yield an apparent rate constant of 0.83 ± 0.06 min^{-1}.

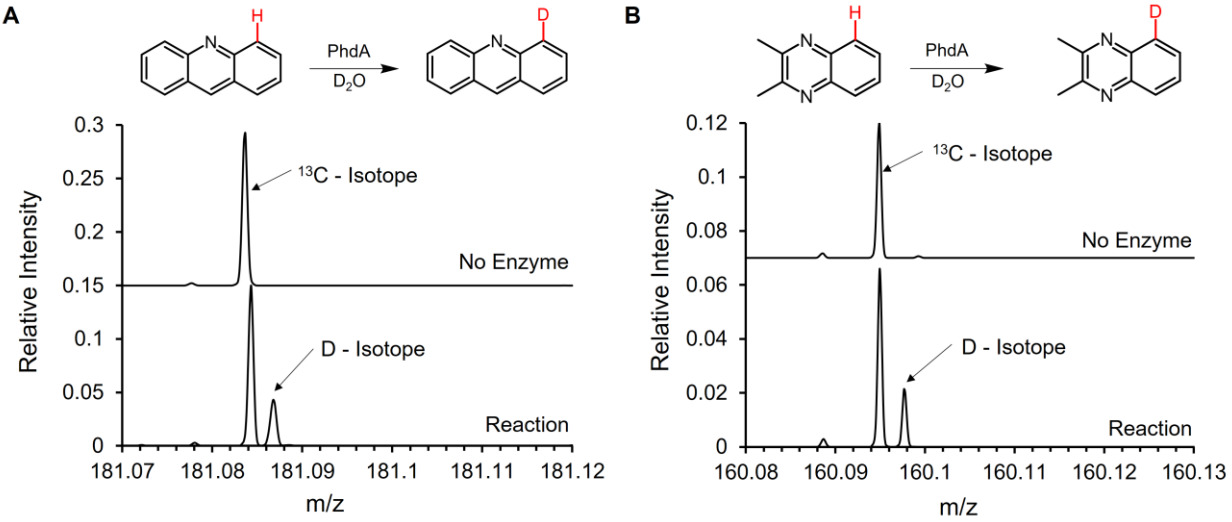


Figure 2.8: H/D exchange of substrate analogs. Orbitrap mass spectra, zoomed into the [M+H+1] region are shown for acridine (**A**) and DQ (**B**). The ^{13}C and D isotopes are labelled.

Encouraged by these results, we sought to study PhdA catalyzed H/D exchange of substrate analogs such as acridine and 2,3-dimethylquinoxaline (DQ). Unfortunately, we were unable to detect H/D exchange through ^{1}H NMR and owing to the large ^{13}C abundance (~10%) in the [M+H+1] m/z, the LC-MS analysis was inconclusive. In order to separate the ^{13}C and D isomers, we sought the ultra-high-m/z-resolution provided by an orbitrap mass analyser (Orbitrap Lumos, Thermo Fisher Scientific Co.). The [M+H+1] region of the orbitrap mass spectrum showed formation of a distinct peak ~0.003 units higher in m/z than the ^{13}C isotope, which matches with the expected m/z for the D-isotope (Figure 2.8). Relative to the [M+H] peak, this peak corresponds

to about ~5% conversion, thus explaining why we couldn't observe it through a traditional mass analyser.

2.4 Discussion

The phylogeny tree described in Figure 1.7 places PhdA in the aromatic acid decarboxylase subgroup. However, based on the initial characterization of PhdA, the ambiguity surrounding the oxidation state of the substrate meant that *a priori* the reaction could plausibly proceed by either an electrophilic mechanism if reduced PCA was the substrate, a cycloaddition mechanism if oxidized PCA was the substrate or through a radical intermediate considering the need for 1 electron reducing agents.[36] In this study, we aimed to get a better understanding of the reaction catalyzed by PhdA as well as gauge its utility as a biocatalyst.

Our inability to obtain holo-PhdA from *E. coli* despite employing various methods highlights the primary difficulty associated with studying UbiD-like enzymes. Barring a few exceptions, these enzymes purify in their inactive apo-form even after co-expression with a prFMN synthase. The distinct 'open' and 'closed' conformers observed in crystal structures might be related to this phenomenon and we hypothesize that the 'open' conformer found in most of these enzymes has poor affinity for prFMN. Nevertheless, *in vitro* synthesis and reconstitution of prFMN is the only reliable way to obtain active enzyme fractions.

The oxidative maturation of prFMN is still not well understood. It appears that maturation requires the UbiD enzyme in question to bind reduced prFMN prior to its oxidation for efficient reconstitution of the holo-enzyme. Studies on the maturation of prFMN in FDC, the best understood system, found that incubating the enzyme with oxidized prFMN actually led to loss of activity.[27] Also, an inverse relationship was observed between the amount of oxidized prFMN present in the reconstitution reaction and the final activity of the reconstituted FDC. This is in

accord with our observation that to efficiently reconstitute PhdA it was necessary to remove inhibitory prFMN species prior to oxidation. Even then, subsequent on-enzyme oxidation led to 3 different prFMN species bound to PhdA (Figure 2.4), whereas, in contrast, only the active iminium form of prFMN is found in FDC. This hints at the possibility that other protein components might be needed as chaperones to efficiently reconstitute some UbiD-like enzymes with prFMN.

Upon obtaining active enzyme, we observed that for the enzyme stored at pH 9.2, the presence of one-electron reducing agents was required to obtain turnover whereas no such pre-incubation was necessary for the enzyme stored at pH 7.2. Since prFMN is prone to hydrolysis and oxidation, we hypothesized that storing PhdA at alkaline pH leads to degradation of prFMN which is partly reversed by adding reducing agents. As such, this disproved the possibility of any radical reaction intermediates and indicates that PCA semiquinone is not a substrate. Furthermore, our experiments clearly establish that oxidized PCA is the substrate for PhdA, helping us resolve the ambiguity surrounding the oxidation state of the substrate.

Following this, we performed a broad substrate screen of PhdA that provided us with useful structure-activity relationships.The trends for the reaction of compounds **1**, **2**, **3** and **4** with PhdA suggest that the extended π-system afforded by the distal aromatic ring of PCA is not important for reactivity, but rather contributes to substrate binding. Similarly, the interesting and somewhat surprising observation that compounds **5** and **6** are substrates demonstrates that the nitrogen atoms of phenazine are not required for activity. Overall, the structure-activity relationships show that PhdA prefers electron-poor aromatic substrates and thus is unlikely to react by an electrophilic mechanism. An electrophilic mechanism is also inconsistent with the decarboxylation of

anthracene carboxylic acids. If we consider an electrophilic intermediate, the activity at C9 should be much higher than C1 position,[55] but we observe that is not the case.

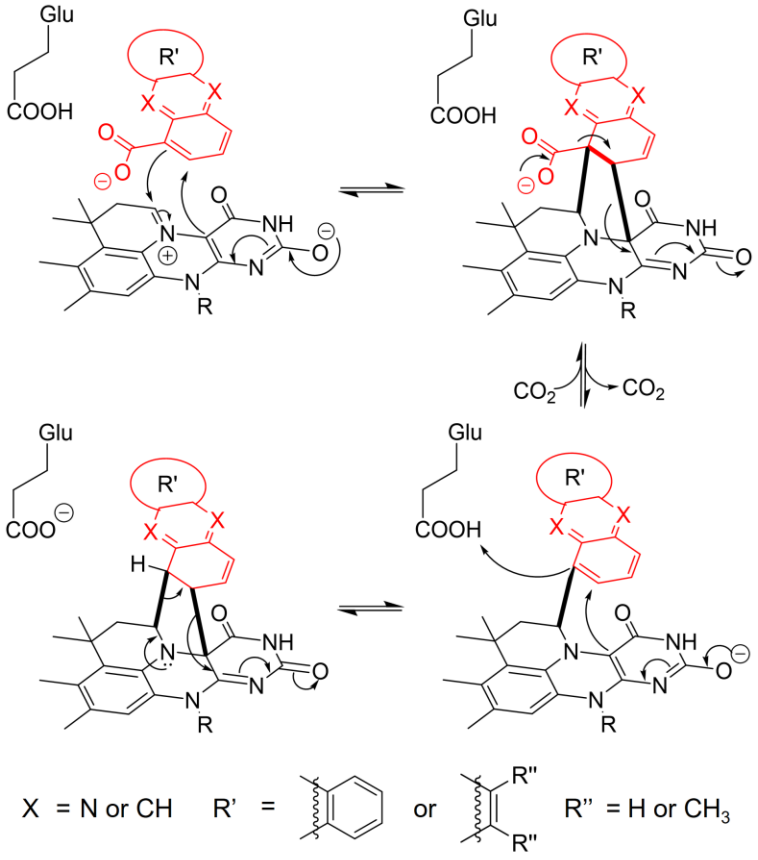


Figure 2.9: Proposed mechanism for PhdA based on substrate scope and deuterium exchange assays.

On the other hand, non-enzymatic reactions of anthracene and acridine with strong dipoles such as nitrile oxides form 1,3-dipolar cycloadducts.[56-58] The monocycloadducts are regioselective, with the nucleophilic O atom of nitrile oxides attacking only the C_2 position on both molecules. If we consider that a similar cycloadduct is formed for PhdA, the implied regioselectivity suggests a C4a – C2 bond between prFMN and phenazine/anthracene (Figure 2.9). For this cycloadduct, prFMN can activate only the C1 position of the substrate, whereas a reasonable mechanism cannot be proposed for activating the C2 position. This is consistent with the observed regioselectivity of deuterium exchange in phenazine and the decarboxylation of anthracene carboxylic acids (compounds **6**, **7** and **8**). Therefore, while the 1,3-dipolar

cycloaddition mechanism for prFMN reacting with PCA remains to be rigorously established for PhdA, we consider this mechanism is most likely.

Perhaps the most promising aspect of prFMN based decarboxylases is their potential ability to functionalize C–H bonds via carboxylation. Carboxylation reactions are useful because they not only trap CO_2 and use it as a C_1 building block but also provide a facile way to functionalize hydrocarbons.[11] Unfortunately, for such enzymes, the reaction equilibrium favours decarboxylation.[16] Our attempts to carboxylate phenazine provided similar results. Deuterium exchange experiments, on the other hand, offer evidence for C–H functionalization as they show the enzyme's ability to abstract a proton from the hydrocarbon. They also provide quantitative information about the site-selectivity, rate and extent of enzymatic C–H functionalization. In the case of PhdA, deuterium exchange experiments on phenazine provide the first evidence (to our knowledge) of the activation of electron-deficient aromatic molecules by prFMN. The observed site-selectivity in deuterium exchange of phenazine provides valuable insight into the reaction mechanism as discussed earlier. The turnover number for exchanging the first deuterium (~ 0.826 min^{-1}) is much lower than the k_{cat}^{app} for decarboxylation (~ 156 min^{-1}). This suggests that deprotonation of phenazine (or a step preceding deprotonation) is rate-limiting in the exchange reaction. Moreover, we demonstrated that PhdA can functionalize C–H bonds in other aromatic hydrocarbons, albeit to a much lesser extent.

In summary, we showed that PhdA is a potentially lucrative biocatalyst that can regioselectively decarboxylate otherwise unreactive aromatic molecules under mild conditions. Our experiments also show that the enzyme is capable of carboxylating hydrocarbons, if alternative strategies can be developed for removing the carboxylic acid product from the reaction mixture.

Chapter 3 Probing the Role of Protein Conformational Changes in the Mechanism of Prenylated-FMN-Dependent Phenazine-1-Carboxylic Acid Decarboxylase[2]

3.1 Introduction

Prenylated flavin mononucleotide (prFMN) is the cofactor for a recently discovered class of (de)carboxylase enzymes that remove or attach carboxylate groups at sp^2-hybridized carbon atoms.[4, 23] prFMN-dependent enzymes are also referred to as 'UbiD-like' enzymes, after the eponymous enzyme involved in bacterial ubiquinone biosynthesis.[18] Although, so far, few of these enzymes have been characterized in detail, the UbiD family of decarboxylases are widely distributed among microbes where many appear to be involved in the metabolism of aromatic compounds. Because of their potential to catalyze (de)carboxylation reactions at otherwise unreactive carbon centers, UbiD-like enzymes have attracted interest as selective and environmentally benign catalysts for organic synthesis.[16, 59, 60]

In prFMN, the isoalloxazine moiety of the flavin is modified by the addition of an isoprene-derived 6-membered ring that spans N5 and C6 of the flavin. This modification occurs on reduced FMN and is catalyzed by a specialized prenyl transferase, with either dimethylallyl phosphate or dimethylallyl pyrophosphate as the prenyl donor. Upon re-oxidization, prFMN forms a nitrogen

[2] The work presented in this chapter is adapted from: "Datar, P.M., Joshi, S.Y., Deshmukh, S.A., Marsh, E.N.G., Probing the Role of Protein Conformational Changes in the Mechanism of Prenylated-FMN-dependent Phenazine-1-carboxylic Acid Decarboxylase; *J. Biol. Chem.*, **2024**, *In press*"

P.M.D and E.N.G.M. conceptualized the idea, P.M.D. planned and conducted the experiments and analyzed the data. S.Y.J. and S.A.D. conducted computational studies and analyzed the data. All authors contributed to writing the manuscript.

ylide and it is this unusual modification that converts this ubiquitous redox cofactor into one that facilitates (de)carboxylation reactions at sp^2-hybridized carbon atoms.[60-63]

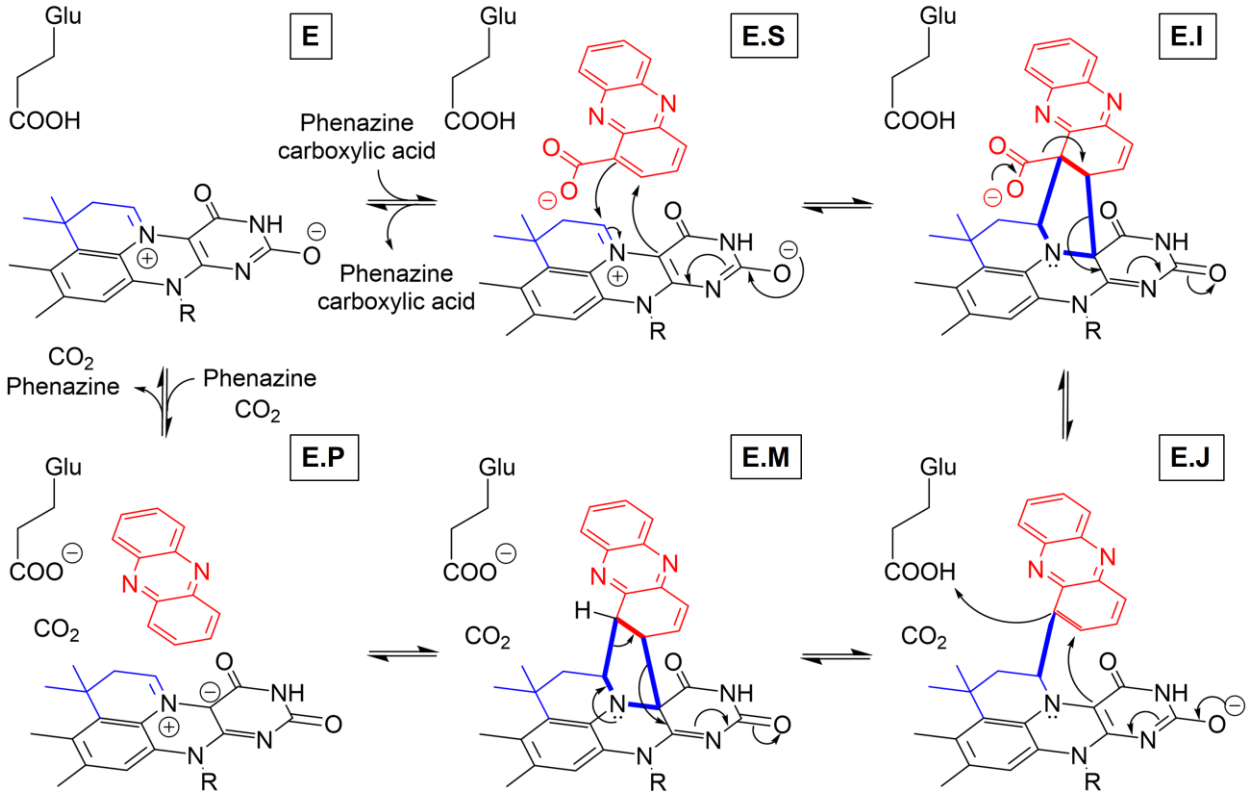


Figure 3.1: Proposed dipolar cycloaddition reaction mechanism for PhdA-catalysed decarboxylation of PCA. E – free enzyme; E.S – Michaelis complex; E.I, E.J and E.M – covalent intermediates; E.P – Enzyme-product complex.

The mechanism by which prFMN-dependent enzymes catalyze decarboxylation reactions hinges upon the reactivity of the nitrogen ylide towards electron-rich unsaturated C–C double bonds. For ferulic acid decarboxylase (FDC), which is the best understood enzyme, experimental evidence[2, 4, 28, 37-40] points to the reaction being initiated through a 1,3-dipolar cycloaddition between the nitrogen ylide with the double bond adjacent to the carboxyl-group of the substrate.[39, 40] This allows the flavin nucleus to act as an electron sink in the subsequent decarboxylation step.[4] However, for enzymes such as AroY[11] that decarboxylate electron-rich aromatic carboxylic acids, the mechanism is more likely to involve electrophilic addition of the substrate to prFMN.[11]

Recently, a novel prFMN-dependent enzyme, PhdA, was discovered which catalyzes the decarboxylation of the redox-active metabolite phenazine-1-carboxylic acid (PCA) to phenazine.[36] Phenazines are secreted by a wide variety of bacteria and are integral to biofilm formation and anoxic survival.[43] Biofilm formation in a clinical setting poses a serious health risk because the biofilm renders pathogens such as *P. aeruginosa* resistant to antibiotic treatment.[51] In contrast, in an agricultural setting, phenazines secreted by *Pseudomonas* spp. are beneficial as they protect cereal crops from various parasitic and fungal diseases.[44]

In our previous work we established conditions for reconstituting PhdA with prFMN and surveyed its substrate scope. We established that the enzyme decarboxylates the oxidized form of phenazine, which earlier studies had left open to question. We showed that, in addition to PCA, PhdA will catalyze the decarboxylation of a wide range of aromatic compounds including such unreactive compounds as anthracene-1-carboxylic acid, albeit at slow rates. We also showed that PhdA catalyzes the exchange of deuterium into phenazine and measured the kinetics of this reaction. Our kinetic analysis suggested that deprotonation of phenazine would likely be the rate determining step for the reverse carboxylation reaction. Based on these results, we proposed a 1,3-dipolar cycloaddition mechanism for PhdA, akin to the reaction catalyzed by FDC (Figure 3.1).[64]

Here we have analyzed the kinetics of PhdA-catalyzed decarboxylation in more detail. We have compared the reaction of the physiological substrate, PCA, with the slow substrate 2,3-dimethylquinoxaline-5-carboxylic acid (DQCA). The kinetics of both substrates display an unusual dependence on D_2O, with the observed isotope effects being more prominent for DQCA. The enzymatic reaction discriminates against the heavy isotope in the transfer of lyonium to the product, as expected, resulting in a *normal* kinetic isotope effect (KIE) on V_{max}/K_M. However, in 98% D_2O buffer under V_{max}/K_M conditions, the reaction exhibits an *inverse* solvent KIE on the

rate. Our studies suggest that the unusual inverse KIE can be explained by a medium effect related to protein conformational changes. We have investigated the nature of this conformation change by performing all-atom molecular dynamics (MD) simulations using the crystal structure of PhdA (PDB ID: 7PDA)[65] as a starting point. We have also developed a kinetic model of the reaction that accounts for these apparently contradictory isotope effects.

3.2 Materials and methods

3.2.1 Reagents and chemicals

Phenazine-1-carboxylic acid (PCA), phenazine, 2,3-dimethylquinoxaline-5-carboxylic acid (DQCA) and 2,3-dimethylquinoxaline (DQ) were purchased from Apollo Scientific Co., Sigma Aldrich Co., 1 ClickChemistry Inc. or Thermo Fischer Scientific Co. and used without further purification. Deuterium oxide (99.8% atom D) was purchased from Thermo Fischer Co. All other reagents were purchased from Sigma Aldrich Co. or Thermo Fischer Co.

3.2.2 Purification and reconstitution of PhdA

E. coli BL21DE3 cells (Invitrogen) were co-transformed with pET20b(+) containing *phdA* and pET28b(+) containing *ubiX* from *Pseudomonas aeruginosa* (*paubiX*). Expression, purification and *in vitro* reconstitution of PhdA was carried out as previously reported (sections 2.2.2 to 2.2.4).[64] Reconstituted PhdA was quantified via Bradford Assay and tested for PCA decarboxylation under standard conditions. The activity didn't vary significantly between different batches of the reconstituted enzyme.

3.2.3 HPLC and LC-MS analysis

The decarboxylation of substrates was monitored using the discontinuous HPLC-based assay previously described (section 2.2.6).[64] The incorporation of solvent deuterium into product was monitored by LC-MS as previously described.[64]

3.2.4 pL-rate profiles

The following buffers were used, all at 0.1 M concentrations: sodium citrate (pL 5.5 – 6), Bis-Tris-Cl (pL 6 – 6.5), potassium phosphate (pL 6.5 – 8) and Tris-Cl (pL 8 – 8.5). Buffers were prepared in H_2O/D_2O and titrated with HCl/DCl or NaOH/NaOD to the desired pH. For phosphate buffers, KH_2PO_4 and K_2HPO_4 stocks were prepared in H_2O/D_2O and titrated. For D_2O buffers, atom fraction of D (χ) was re-calculated based on protium added from the buffer components (in most cases $\chi > 0.99$). Corrections were applied to the pH-meter readings using the equation[38]:

$$pD = pH_{meter} + 0.076\chi^2 + 0.3314\,\chi + 0.00009$$

Activity assays were performed at room temperature (20 - 22°C) and consisted of 0.1 M buffer, 0.1 – 0.5 µM reconstituted PhdA and different concentrations of substrates. For PCA, 10-15 µM substrate was added under V_{max}/K_M conditions whereas for V_{max}, 500-1000 µM PCA was used. Similarly, for DQCA, 30 – 50 µM of the acid was used for V_{max}/K_M and 5 – 10 mM for V_{max}. Reactions were quenched by adding 500 mM NaOH (final concentration) and analyzed by HPLC. Reaction rates, normalized by enzyme concentration (v/E_t), were plotted as a function of pL and fit to the following equation[66]:

$$v/E_t = \frac{(v/E_t)_{max}}{1 + 10^{pK_{a1} - pL} + 10^{pL - pK_{a2}}}$$

Where it is assumed that the pL-rate behavior of PhdA arises from the titration of 2 ionizable groups, each with a single pK_a. $(v/E_t)_{max}$ is the pL-independent rate, pK_{a1} corresponds to

the group that needs to be deprotonated for activity and pK_{a2} is for the residue that needs to be protonated.

3.2.5 Solvent Viscosity Studies

All activity assays were setup in potassium phosphate buffer (pH = 7.0) at 22 ^{o}C with the addition of either glucose or sucrose (between 8 – 40% w/w) as viscosogens. The relative viscosity, η_{rel} of the resulting solutions was calculated based on the % w/w of the viscosogen.[67] Reaction rates were measured at 'low' substrate concentrations (V_{max}/K_M conditions, 10 - 15 μM for PCA, 50 - 100 μM for DQCA) and 'high' substrate concentrations (V_{max} conditions, 0.5 - 1 mM for PCA, 5 -10 mM for DQCA). The ratio of the rate without (ν_0) and with (ν_η) viscosogen was plotted against $\eta_{rel} - 1$ and wherever applicable, the plots were fit to the following equations[49]:

$$\nu_0/\nu_\eta = m(\eta_{rel} - 1) + 1 \qquad \textit{(For normal viscosity effect)}$$

$$\nu_0/\nu_\eta = \frac{1}{1+A\left[\frac{(\eta_{rel}-1)}{(\eta_{rel}-1)+B}\right]} \qquad \textit{(For inverse viscosity effect)}$$

Here, m, A and B are parameters of the fit and describe the extent to which the rate depends on solvent viscosity.

3.2.6 Solvent isotope effects

Potassium phosphate buffers (pH or pD = 7) in H_2O and D_2O were made as described above and added volumetrically to obtain mixed isotopic buffers. The atom fraction of D (χ) was adjusted by applying the necessary corrections.[66] All subsequent reactions were performed in these buffers.

For measuring $^{D_2O}V/K_P$, reactions containing PhdA (0.2 μM) and PCA (15 μM) were performed in mixed isotopic buffers and quenched after 20 sec. Alternatively, 1 μM PhdA was reacted with 50 μM DQCA and quenched after 15 min. Samples were analyzed by LC-MS in

positive ion mode. Due to the substantial presence (~10%) of naturally occurring ^{13}C-isotopes that also incorporate D, the LC-MS peaks for [M+H], [M+H+1] and [M+H+2] changed with χ. Thus, the mole fraction of the deuterated product ($\chi_{D,product}$) was calculated according to the following equation and plotted against χ:

$$\chi_{D,product} = \frac{[M+H+1]+[M+H+2]}{[M+H]+[M+H+1]+[M+H+2]}$$

The isotope effect associated with transfer of deuterium to the product is written as (Here, $[P_H]/[P_D]$ is the ratio of proteated to deuterated product, ${}^{D_2O}V/K_P$ is the isotope effect and χ is the atom fraction of D in solvent)[38]:

$$\frac{[P_H]}{[P_D]} = {}^{D_2O}V/K_P\left(\frac{1-\chi}{\chi}\right)$$

From this, the expression relating $\chi_{D,product}$ to χ can be derived as follows:

$$\frac{[P_H]}{[P_D]}+1 = {}^{D_2O}V/K_P\left(\frac{1-\chi}{\chi}\right)+1;$$

$$\frac{[P_D]}{[P_D]+[P_H]} = \left(\frac{\chi}{{}^{D_2O}V/K_P(1-\chi)+\chi}\right);$$

$$\chi_{D,product} = \frac{1}{{}^{D_2O}V/K_P}\left(\frac{\chi}{\left(\frac{1}{{}^{D_2O}V/K_P}-1\right)\chi+1}\right);$$

Now, in order to account for the m/z signal from ^{13}C-isotopes, a correction B (ratio of ^{13}C-only isotopes) needs to be applied. As mentioned previously, the value of B decreases with χ by the factor $\left(\frac{1-\chi}{\left(\frac{1}{{}^{D_2O}V/K_P}-1\right)\chi+1}\right)$ as ^{13}C-isotopes also incorporate D (here we assume that B changes only because of the H/D isotope effect and ^{12}C/^{13}C isotope effect is negligible). Applying the above correction and further simplification yields:

$$\chi_{D,product} = \frac{B + \left(\frac{1}{D_2O V/K_P} - B\right)\chi}{\left(\frac{1}{D_2O V/K_P} - 1\right)\chi + 1}$$

Proton inventories were obtained under either V_{max} (5 – 10 mM DQCA) or V_{max}/K_M (50 – 100 µM DQCA) conditions. The ratio of the rate in mixed isotopic water to H_2O (ν_χ/ν_0) was plotted as a function of the D-atom fraction, χ and fitted to equations 2 – 5 (see 'Results' section) as applicable.[68]

For midpoint SIE,[48] the rate of DQCA decarboxylation (ν) was monitored at low substrate concentrations (V_{max}/K_M conditions) in H_2O ($\chi = 0$), D_2O ($\chi = 0.99$) and $\chi = 0.5$, following which $\nu_{0.5}/\nu_0$ and $\nu_{0.99}/\nu_0$ were calculated. Later, Eq.2, Eq.3 and Eq.4 were solved for $\chi = 0.99$ to determine the parameters ϕ_T, Z and ϕ_R. The experimental value of $\nu_{0.5}/\nu_0$ was then compared to the theoretical values calculated for the different mechanisms represented by Eq.2, Eq.3 and Eq.4 at $\chi = 0.5$.

3.2.7 Protein Unfolding

All reactions were performed at 20˚C. Stocks of 9 M urea in H_2O or D_2O were added to 0.1 µM PhdA in 0.1 M potassium phosphate buffer (pL = 7) to achieve different concentrations of urea. The samples were incubated for 60 - 90 min after which protein fluorescence emission spectra were recorded. Excitation wavelength = 295 nm; emission spectrum recorded between 310 – 470 nm. The average wavelength of emission (λ_{avg}) at each concentration of urea was calculated by[69]:

$$\lambda_{avg} = \sum_{i=1}^{N}(I_i\lambda_i) \Big/ \sum_{i=1}^{N} I_i$$

Where N is the total number of data points and λ_i, I_i are the wavelength and intensity of the i^{th} data point. The normalized λ_{avg} values were plotted against urea concentration and fitted to the following equation[70]:

$$f_U = \frac{F + U.e^{-m(K_{1/2}-x)}}{1 + e^{-m(K_{1/2}-x)}}$$

Here, f_U is the unfolded fraction as a function of urea (x), F, U, m are the fitting parameters and $K_{1/2}$ is the urea concentration for $f_U = 0.5$. The free energy of unfolding was calculated as $\Delta G_U = m.R.T.K_{1/2}$.[70] Here, R is the gas constant and T = 293.15 K.

3.2.8 Inhibition Studies

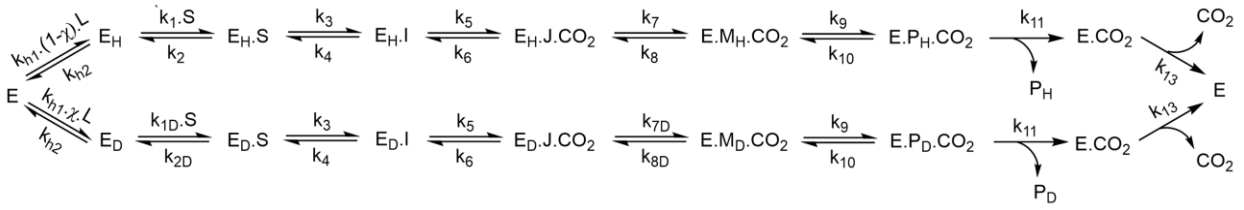
Reactions were setup in potassium phosphate buffer (pH or pD = 7) and consisted of 0.2 μM PhdA, 50 μM PCA and varying concentrations of DQCA. The residual normalized rate (v/E_t) was plotted against DQCA concentration and fit to the following equation, assuming competitive inhibition[71]:

$$v/E_t = \frac{k_{cat}.S}{S + \left(1 + \frac{I}{K_I}\right)K_M}$$

Here, I is the independent variable (DQCA concentration), S = 50 μM is the concentration of PCA, k_{cat} and K_M are the steady-state parameters for PCA and K_I is the apparent inhibition constant for DQCA.

3.2.9 Modelling Kinetic Mechanisms for Isotope Effects

The *normal* isotope effect measured for the transfer of deuterium to the product and the *inverse* isotope effect measured by comparison of reaction rates indicate that multiple steps are isotopically sensitive. Therefore, it is necessary to develop a kinetic scheme that can explain both isotope effects. The simplest mechanism conceptually (Mechanism 1) is shown below:



Mechanism 1

Here, the catalytic Glu269 residue exchanges L (either proton or deuteron) with solvent (indicated as E_H or E_D) before substrate binding and both products are released after all the chemical steps. *Inverse* medium effects are usually general effects, which basically means that they would be applicable to multiple, if not all, kinetic steps.[72] However, for simplicity, we consider that only the first step (k_1) expresses the medium effect (k_{1D}) and k_7 is the proton transfer step with the *normal* isotope effect (k_{7D}). Expressions for ${}^{D_2O}V/K_S$ and ${}^{D_2O}V/K_P$ in terms of the elementary rate constants were derived using Cleland's method of net rate constants.[73] Briefly, for any step *i*, the net rate constant $k_i' = k_i \frac{k_{i+2}}{(k_{i+1} + k_{i+2})}$. Here, all forward rate constants are odd numbers and all reverse rate constants are even numbers. Expressions were manipulated in Wolfram Mathematica. For Mechanism 1 the net rate constants, k_i' are given by:

$$k_9' = k_9 \frac{k_{11}}{(k_{11} + k_{10})};\ k_7' = k_7 \frac{k_9'}{(k_9' + k_8)};\ k_5' = k_5 \frac{k_7'}{(k_7' + k_6)};\ k_3' = k_3 \frac{k_5'}{(k_5' + k_4)};\ k_1' = k_1.S \frac{k_3'}{(k_3' + k_2)}$$

$$k_{h1}' = k_{h1}.(1-\chi).L \frac{k_1'}{k_{h2}} \qquad \textit{(Assuming rapid exchange, } k_{h2} \gg k_1'.\ \chi \textit{ is atom fraction of D)}$$

$$k_{7D}' = k_{7D} \frac{k_9'}{(k_9' + k_{8D})};\ k_{5D}' = k_5 \frac{k_{7D}'}{(k_{7D}' + k_6)};\ k_{3D}' = k_3 \frac{k_{5D}'}{(k_{5D}' + k_4)};\ k_{1D}' = k_{1D}.S \frac{k_{3D}'}{(k_{3D}' + k_{2D})};$$

$$k_{h1D}' = k_{h1}.\chi.L.\frac{k_{1D}'}{k_{h2}}$$

The expressions for ${}^{D_2O}V/K_S$ can be written as:

$${}^{D_2O}V/K_S = \frac{\left(\frac{V/K}{E_t}\right)_{H_2O}}{\left(\frac{V/K}{E_t}\right)_{D_2O}} = \frac{k_1'/S}{k_{1D}'/S}$$

Thus we get (Equation S1),

$$^{D_2O}V/K_S = \frac{k_1k_7(k_{2D}k_4k_6k_{8D}k_{10}+k_{2D}k_4k_6k_{8D}k_{11}+k_{2D}k_4k_6k_9k_{11}+k_{2D}k_4k_7{}_Dk_9k_{11}+k_{2D}k_5k_7{}_Dk_9k_{11}+k_3k_5k_7{}_Dk_9k_{11})}{k_{1D}k_7{}_D(k_2k_4k_6k_8k_{10}+k_2k_4k_6k_8k_{11}+k_2k_4k_6k_9k_{11}+k_2k_4k_7k_9k_{11}+k_2k_5k_7k_9k_{11}+k_3k_5k_7k_9k_{11})}$$

While Eq. S1 is rather intimidating, it simply states that $^{D_2O}V/K_S$ depends on the isotopically sensitive steps, k_1 and k_7, 'modulated' by the other steps. Two approaches exist to reduce such complex equations into simpler terms – the 'Commitment Factor' concept developed by O'Leary, Cleland and Northrop[74] and the 'Virtual Transition State' theory conceived by Schowen, Stein and Quinn.[48, 75-77] While both methods yield identical results algebraically, the former approach assumes that only a single step is isotopically sensitive whereas no such limitation exists for the latter. Therefore, we considered the virtual transition state theory for our analyses. Briefly, it treats the steady state parameters as weighted averages of the individual kinetic steps. Therefore, $^{D_2O}V/K_S$ can be expressed as:

$$^{D_2O}V/K_S = Z\left(\sum_i^m f_i \,.\, ^{D_2O}k_i\right)$$

Here, Z is the general medium effect and $^{D_2O}k_i$ is the isotope effect on the i^{th} step that has a fractional contribution f_i to $^{D_2O}V/K_S$. Thus, by definition, $\sum f_i = 1$. Based on the above treatment, the equation for $^{D_2O}V/K_S$ in Mechanism 1 can be reduced to (Equation S2):

$$^{D_2O}V/K_S = f_1 .\, ^{D_2O}k_1 + f_7 .\, ^{D_2O}k_7 + (1 - f_1 - f_7)$$

Here, $^{D_2O}k_1 = k_1/k_{1D}$, $^{D_2O}k_7 = k_7/k_{7D}$ and f_1, f_7 are their respective fractional contributions to $^{D_2O}V/K_S$. The above equation assumes that the equilibrium isotope effect on each step is unity, i.e. $(k_1/k_2)/(k_{1D}/k_{2D})$ and $(k_7/k_8)/(k_{7D}/k_{8D}) \sim 1$. Now, if χ is the solvent D-atom fraction, Z is the value of the medium effect and ϕ_T is the transition state fractionation factor for proton transfer, then $\frac{k_{1,\chi}}{k_{1,0}} = Z^{\chi}$ and $\frac{k_{7,\chi}}{k_{7,0}} = 1 - \chi + \chi.\phi_T$. The proton inventory expression form of Eq. S2 can be written as:

$$\left[\frac{\left(V/K_S\right)_\chi}{\left(V/K_S\right)_0}\right]^{-1} = f_1.Z^{-\chi} + f_7(1-\chi+\chi.\phi_T)^{-1} + (1-f_1-f_7)$$

Therefore, Mechanism 1 is shown to be incompatible with the experimental data as it predicts a 'dome' shaped proton inventory for ${}^{D_2O}V/K_S$ with competing *normal* and *inverse* isotope effects, which is not observed.

Furthermore, consideration of the expression for the product isotope derived for Mechanism 1 also shows it to be incompatible with the experimental data. If one considers the upper (H_2O) and lower (D_2O) pathways as competing reactions in mixed solvent isotopes, then the net rate constants for the upper (k_{upper}) and lower (k_{lower}) pathways are given by:

$$k_{upper} = \frac{k'_{h1}}{1+k'_{h1}\left(\frac{1}{k'_1}+\frac{1}{k'_3}+\frac{1}{k'_5}+\frac{1}{k'_7}+\frac{1}{k'_9}+\frac{1}{k_{11}}+\frac{1}{k_{13}}\right)+k'_{h1D}\left(\frac{1}{k'_{1D}}+\frac{1}{k'_{3D}}+\frac{1}{k'_{5D}}+\frac{1}{k'_{7D}}+\frac{1}{k'_9}+\frac{1}{k_{11}}+\frac{1}{k_{13}}\right)}$$

$$k_{lower} = \frac{k'_{h1D}}{1+k'_{h1}\left(\frac{1}{k'_1}+\frac{1}{k'_3}+\frac{1}{k'_5}+\frac{1}{k'_7}+\frac{1}{k'_9}+\frac{1}{k_{11}}+\frac{1}{k_{13}}\right)+k'_{h1D}\left(\frac{1}{k'_{1D}}+\frac{1}{k'_{3D}}+\frac{1}{k'_{5D}}+\frac{1}{k'_{7D}}+\frac{1}{k'_9}+\frac{1}{k_{11}}+\frac{1}{k_{13}}\right)}$$

At any time *t* (under steady state), $[P_H] = k_{upper}.E.t$ and $[P_D] = k_{lower}.E.t$

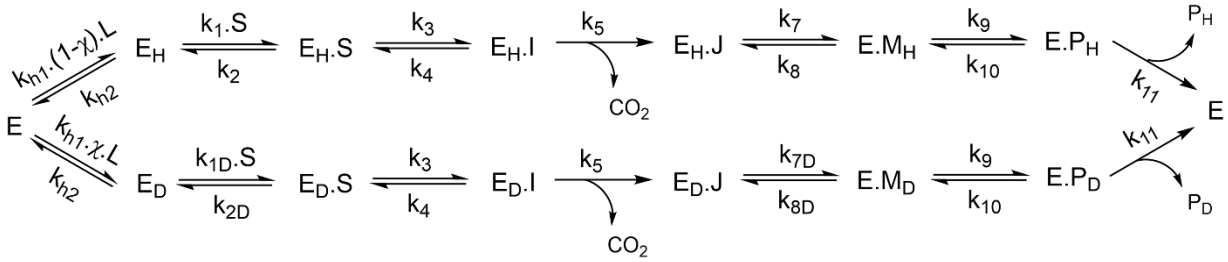
Therefore, $\frac{[P_H]}{[P_D]} = \frac{k_{upper}}{k_{lower}} = \frac{k'_{h1}}{k'_{h1D}}$ and simplifying this we get:

$$\frac{[P_H]}{[P_D]} = {}^{D_2O}V/K_S\frac{(1-\chi)}{\chi}$$

Therefore, for this mechanism, the product isotope effect should reflect the rate isotope effect, which is not the case, further confirming that Mechanism 1 is incorrect.

We next consider Mechanism 2, in which decarboxylation is functionally irreversible under the conditions of the experiment. This assumption is plausible if CO_2 is released immediately following the decarboxylation step.

Here k_6, the reverse of step k_5, is 0 so that the expression for ${}^{D_2O}V/K_S$ and $[P_H]/[P_D]$ are given by, respectively:



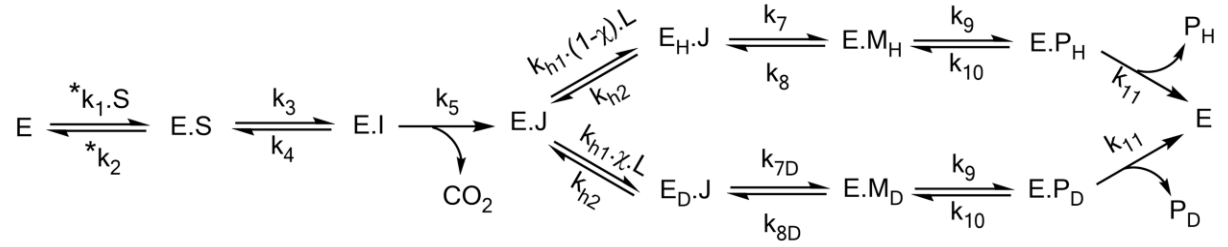
Mechanism 2

$$^{D_2O}V/K_S = \frac{k_1(k_{2D}k_4 + k_{2D}k_5 + k_3k_5)}{k_{1D}(k_2k_4 + k_2k_5 + k_3k_5)} = f_1.\,^{D_2O}k_1 + (1 - f_1)$$

$$\frac{[P_H]}{[P_D]} = \frac{k_1(k_{2D}k_4 + k_{2D}k_5 + k_3k_5)(1-\chi)}{k_{1D}(k_2k_4 + k_2k_5 + k_3k_5)\chi} = {}^{D_2O}V/K_S\frac{(1-\chi)}{\chi}$$

Thus, Mechanism 2 predicts that only $^{D_2O}k_1$ is expressed in the rate and product distribution, whereas $^{D_2O}k_7$ is not. This is because the enzyme can exchange H/D only in the beginning of the reaction and cannot partition back after decarboxylation. Therefore, the flux through the upper and lower pathways is not sensitive to k_7, the step in which proton transfer occurs.[78]

Mechanism 3 represents a modification of Mechanism 2 and allows H/D exchange to occur on the enzyme after the decarboxylation step:



Mechanism 3

Here, the equation for $^{D_2O}V/K_S$ is the same as Mechanism 2:

$$^{D_2O}V/K = \frac{k_1(k_{2D}k_4 + k_{2D}k_5 + k_3k_5)}{k_{1D}(k_2k_4 + k_2k_5 + k_3k_5)} = f_1.\,^{D_2O}k_1 + (1 - f_1)$$

However, for $[P_H]/[P_D]$, the net rate constants for the upper and lower pathway have to be re-written as follows:

$$k_{upper} = \frac{k'_{h1}}{1 + k'_{h1}\left(\frac{1}{k'_7} + \frac{1}{k'_9} + \frac{1}{k_{11}}\right) + k'_{h1D}\left(\frac{1}{k'_{7D}} + \frac{1}{k'_9} + \frac{1}{k_{11}}\right)}$$

$$k_{lower} = \frac{k'_{h1}}{1 + k'_{h1}\left(\frac{1}{k'_7} + \frac{1}{k'_9} + \frac{1}{k_{11}}\right) + k'_{h1D}\left(\frac{1}{k'_{7D}} + \frac{1}{k'_9} + \frac{1}{k_{11}}\right)}$$

And

$$\frac{[P_H]}{[P_D]} = \frac{k_{upper}}{k_{lower}}$$

$$= \frac{k_7(k_{8D}k_{10} + k_{8D}k_{11} + k_9k_{11})}{k_{7D}(k_8k_{10} + k_8k_{11} + k_9k_{11})} \cdot \frac{(1-\chi)}{\chi}$$

$$= [f_7 .\, {}^{D_2O}k_7 + (1 - f_7)]\frac{(1-\chi)}{\chi}$$

$$= {}^{D_2O}V/K_P \frac{(1-\chi)}{\chi}$$

Therefore, Mechanism 3 allows for the *inverse* isotope effect to be expressed in ${}^{D_2O}V/K_S$ and the *normal* isotope effect on the product distribution to be expressed as ${}^{D_2O}V/K_P$.

It is worth noting that while the current analysis cannot provide an exact value for the intrinsic isotope effect (${}^{D_2O}k_7$), the virtual transition state theory allows us to make reasonable estimates of its limiting values. Assuming no quantum tunneling, the maximum value of ${}^{D_2O}k_7$ is 10.[68] Thus, in the case of DQCA, for ${}^{D_2O}V/K_P$ = 4.15 and ${}^{D_2O}k_7$ = 10, $f_7 \sim 0.35$ i.e the proton transfer step will be ~35% rate limiting. Now, if the proton transfer is completely rate limiting for ${}^{D_2O}V/K_P$, $f_7 = 1$ and under these conditions, ${}^{D_2O}k_7 = {}^{D_2O}V/K_P = 4.15$, which is the minimum possible value of ${}^{D_2O}k_7$. If ϕ_T is the transition state fractionation factor for proton transfer, then ${}^{D_2O}k_7 = \frac{1}{\phi_T}$ and hence, the upper and lower limits on ϕ_T are 0.24 and 0.1 respectively.

Now, we can also write an equation for V_{max}, ${}^{D_2O}V$. For this, we consider the medium effect to be a general effect rather than limited to a single step. The expression for proton inventory as per the virtual transition state theory is:

$$\left(\frac{\nu_\chi}{\nu_0}\right)^{-1} = Z^{-\chi}\left[w_7.\left(\frac{k_{7,\chi}}{k_{7,0}}\right)^{-1} + (1 - w_7)\right]$$

$$= Z^{-\chi}\left[\frac{w_7}{1 - \chi + \chi.\phi_T} + (1 - w_7)\right]$$

Here, ϕ_T is the transition state fractionation factor for proton transfer and w_7 is the contribution of k_7 to V_{max}. Simplifying the above expression gives:

$$\nu_\chi/\nu_0 = \frac{Z^\chi(1 - \chi + \chi.\phi_T)}{1 - \chi + \chi.w_7 + \chi.\phi_T - \chi.w_7.\phi_T}$$

All the equations described here and in 'Results' section were fit to experimental data in Origin 2022 graphing software.

3.2.10 Molecular Dynamics simulations

Molecular Dynamics (MD) simulations were performed using the GROMACS package.[79] Crystal structure of PhdA monomer (PDB:7PDA)[65] was solvated in a 100 Å x 100 Å x 100 Å cubic box with ~30,000 solvent (H_2O or D_2O) molecules. The protein was parameterized using the CHARMM36 force-field[80], whereas the cofactor was parameterized using the CHARMM general force-field.[81] The TIP3P water model was used to represent H_2O[82], whereas D_2O was parameterized using the TIP3P-HW model.[83] Additionally, 150 mM KCl ions were used for creating proper physiological conditions and extra neutralizing ions were added to maintain an overall charge neutral system. With a 12 Å cutoff, the Particle-Mesh Ewald (PME) approach was utilized for long-range electrostatics.[84] Similarly, the Van der Waals cutoff was also set to 12 Å. Initially, the model systems were energy minimized using the steepest descent minimization

algorithm.[85] The MD simulations were performed at a temperature of 298 K, with periodic boundaries, using the NPT thermodynamic ensemble, where temperature and pressure control was achieved using Nose-Hoover thermostat and Parrinello-Rahman barostat.[86-88] The trajectory was recorded every 1 ps, giving a total of 1,000,000 frames for each of the 1 μs simulations. The entire trajectory was analyzed for the root mean squared deviation (RMSD), free energy landscape (FEL) and the evolution of R159–I416 and prFMN–E269 distances. Allowing the system to equilibrate during the initial 500 ns, the trajectories from 500 – 1000 ns were used to analyze the root mean squared fluctuation (RMSF) as well as protein-solvent and intraprotein hydrogen bonding using GROMACS in-built packages and Visual Molecular Dynamics.[89]

3.2.11 Hydrogen bonding in MD simulations

Hydrogen bond autocorrelations, calculated for both protein simulations were fitted to a sum of exponentials using the following equation:

$$C(t) = \sum_{i=1}^{N} A_i e^{-t/T_i}$$

Where T_i is the time constant, A_i is the amplitude of the *i*th decay process and *N* is the number of exponentials. For the best fit, we used five exponentials which are presented in Table 3.1. This table can be interpreted as A_i % of h-bonds break in T_i ps time. The high R^2 values for curve fitting and similar amplitude and time constant values for both solvents shows that the equation was appropriately fit to the autocorrelation data obtained from the MD simulations. Generally, higher percentages of hydrogen bonds persisted for longer when solvated in D_2O as compared to H_2O. Furthermore, intra-protein hydrogen bonds were also significantly longer in the presence of D_2O which agrees with our earlier analyses and previous literature.[90] A detailed list of unique intra-

protein hydrogen bonds found in the final protein conformers in H_2O and D_2O are presented in Appendix B.

Protein-solvent hydrogen bonding											
	R^2	**A1**	**A2**	**A3**	**A4**	**A5**	**T1**	**T2**	**T3**	**T4**	**T5**
Water	0.9997	27.60%	37.06%	21.08%	7.90%	6.19%	0.77	14.51	82.86	600.19	12780.06
Heavy water	0.9997	24.65%	35.60%	22.39%	7.37%	9.82%	0.77	15.00	81.68	603.24	15580.13
Intraprotein hydrogen bonding											
	R^2	**A1**	**A2**	**A3**	**A4**	**A5**	**T1**	**T2**	**T3**	**T4**	**T5**
Water	0.9995	12.79%	2.60%	2.78%	4.25%	77.59%	0.01	8.76	140.72	1188.10	215049.55
Heavy water	0.9996	13.35%	2.49%	2.72%	4.35%	77.09%	0.39	24.61	291.22	2987.92	3904080.73

Table 3.1: Amplitudes and time constants obtained by fitting the hydrogen bond autocorrelation function to the sum of five exponentials.

3.3 Results

3.3.1 Solvent isotope and viscosity effects for PhdA reacting with PCA

Decarboxylation reactions catalyzed by UbiD-like enzymes involve the transfer of a solvent proton to carbon, often mediated by an active site glutamate residue (Figure 3.1), that is subject to an isotope effect.[38] Therefore, we reasoned that investigating the kinetic behavior of PhdA in buffered D_2O might be mechanistically informative. Solvent deuterium content affects the pK_a's of most acids including enzymatic functional groups,[48] which necessitates measuring isotope effects in a pL-independent region (L = H or D) of the pL-rate profile. Furthermore, comparing isotope effects under V_{max} (high substrate) and V_{max}/K_M (low substrate) provides information about different regions of the kinetic mechanism. Previously, we determined k_{cat} = 155 ± 4 min^{-1} and K_M = 53 ± 2 µM for the PhdA catalyzed decarboxylation of PCA.[64] Based on these parameters, we measured pL-rate profiles under V_{max} ([PCA] = 500 µM; ~ 10.K_M) or V_{max}/K_M ([PCA] = 15 µM; ~ K_M/3) conditions (Figure 3.2 B and C).

Under both conditions the reaction exhibited a classical bell-shaped activity profile. Under V_{max}/K_M conditions in H_2O, the acidic and basic limbs have apparent pK_as = 6.7 ± 0.3 and 7.0 ± 0.3 respectively. The values under V_{max} conditions are shifted further apart: pK_as = 6.2 ± 0.2 and 7.6 ± 0.2. The apparent pK_as measured under V_{max} and V_{max}/K_M conditions reflect the pK_as of the enzyme-substrate complex and free enzyme respectively. The pK_a of N1 proton in reduced FMN is ~ 6.2 to 6.9, based on the local environment.[91] Therefore, we hypothesize that pK_{a1} of PhdA's pH-rate profile reflects the N1 proton on prFMN, which needs to be deprotonated for activity. On the other hand, pK_{a2} most likely reflects the active site glutamate (Glu269 in PhdA), which needs to be protonated for activity. Repeating the measurements in D_2O resulted in a small but significant upward shift of ~ 0.3 pH units (after correcting for the difference in the activity between protium and deuterium ions).

Having established the pL profile for PhdA reacting with PCA under V_{max} and V_{max}/K_M conditions, we measured the corresponding SIEs, ^{D_2O}V and $^{D_2O}V/K_S$, by direct comparison of the rates of reaction at the pL maxima for each condition. These measurements yielded ^{D_2O}V = 0.93 ± 0.12 (n = 6) and $^{D_2O}V/K_S$ = 0.75 ± 0.17 (n = 6). It is unremarkable that ^{D_2O}V is unity within error, but it is surprising that $^{D_2O}V/K_S$ appears slightly inverse, although the error is quite large. Apparent inverse SIEs observed in some enzymes e.g. NAD-Malic enzyme,[92] have been attributed to the increased viscosity of D_2O. However, when we examined the effect of sucrose or glucose on V_{max} or V_{max}/K_M for PCA (Figure 3.2D, E), the rates of reaction were either slightly decreased or unaffected by increasing concentrations of viscosogen. These observations indicate that viscosity differences are not responsible for the apparent *inverse* SIE.

To gain further insight into the mechanism, we measured the solvent isotope effect for the transfer of deuterium to the product, phenazine, ($^{D_2O}V/K_P$), by internal competition (Figure 3.2F).

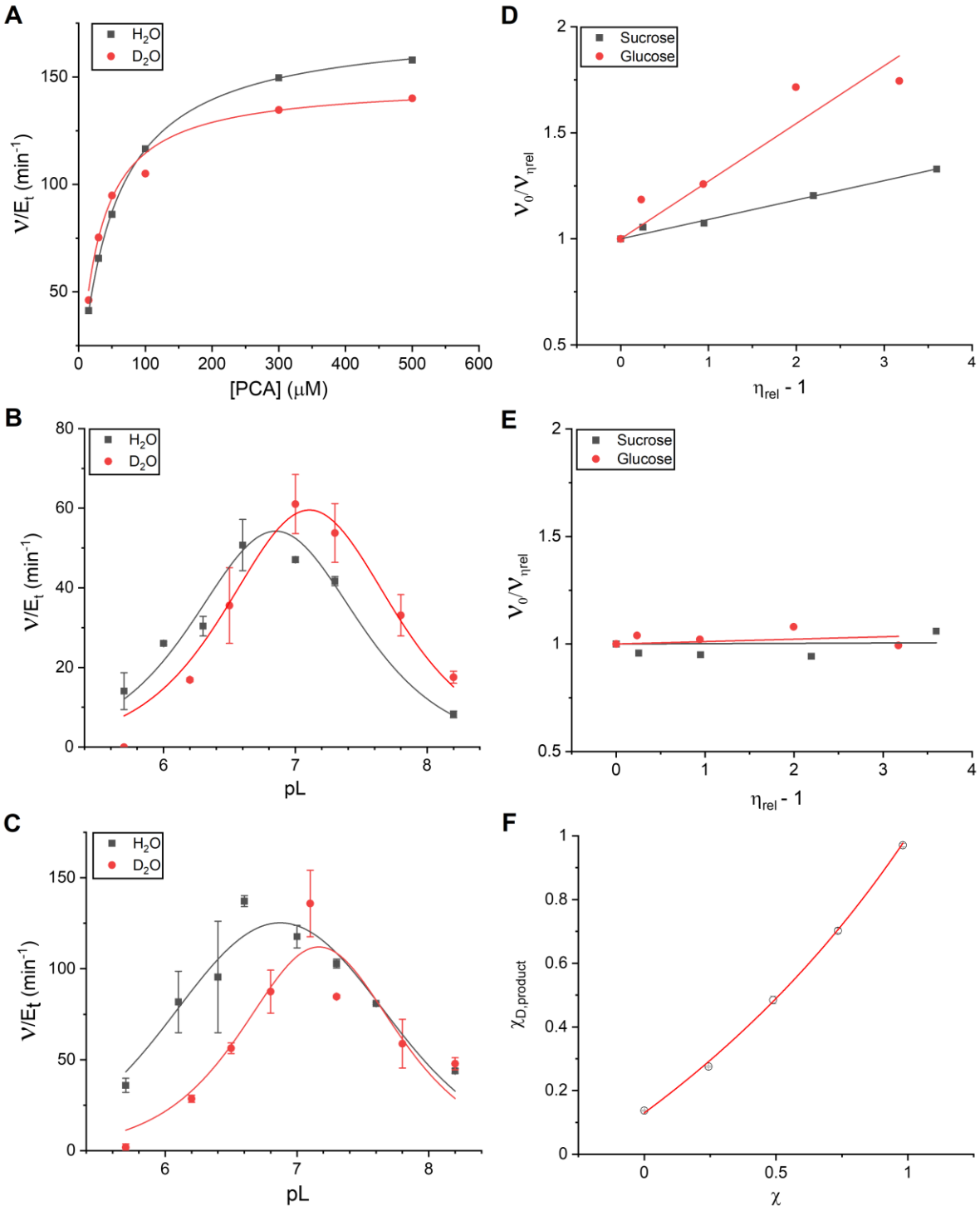


Figure 3.2: Solvent isotope and viscosity effects on the decarboxylation of PCA by PhdA. **A:** Steady state kinetics for PhdA catalyzed PCA decarboxylation at pL = 7 in H_2O and D_2O**. B and C:** pL-rate dependence for the reaction monitored under V_{max}/K_M and V_{max} conditions respectively. Data are fitted assuming 2 ionizable groups contribute to the rate profile. **D and E:** Solvent viscosity studies under V_{max}/K_M and V_{max} respectively. Data are fitted as per the equations described in section 3.2.5. **F:** Determination of $^{D_2O}V/K_P$ for the transfer of deuterium to phenazine. The mole fraction of deuterated phenazine ($\chi_{D,product}$) is plotted as a function of solvent D-atom fraction (χ) and the isotope effect calculated by fitting the data to Eq. 1.

The PhdA-catalyzed decarboxylation of PCA was monitored at pL=7 in buffers containing increasing D-atom fraction (χ), and the phenazine produced analyzed by LC-MS to determine the mole fraction of deuterium appearing in the product, ($\chi_{D,product}$). ${}^{D_2O}V/K_P$ was calculated from these data by fitting them to Eq. 1 (Refer to section 3.2.6 for derivation):

$$\chi_{D,product} = \frac{B+\left(\frac{1}{{}^{D_2O}V/K_P} - B\right)\chi}{\left(\frac{1}{{}^{D_2O}V/K_P} - 1\right)\chi+1} \quad (1)$$

Here, B is the ^{13}C isotopic abundance in product. The solvent KIE on V_{max}/K_M calculated from product distribution (${}^{D_2O}V/K_P$ = 1.43 ± 0.06) is small and, as expected, normal. This is in contrast to the KIE on V_{max}/K_M calculated from the reaction rate (${}^{D_2O}V/K_S$). The difference in the SIE obtained between the two methods suggests that the two effects arise from different steps in the mechanism.

3.3.2 Solvent isotope and viscosity effects for PhdA reacting with DQCA

The reactions of enzymes with "slow" substrates can often be mechanistically informative because they may uncover steps that are kinetically masked in reactions with the physiological substrates. Previously, we identified 2,3-dimethylquinoxaline-5-carboxylic acid (DQCA) as a substrate for PhdA; therefore, we reasoned that a detailed examination of the kinetics of DQCA decarboxylation might be informative. We first determined k_{cat} and K_M for DQCA reacting with PhdA at pH 7.0 and 22 ºC (Figure 3.3A). k_{cat} = 6.1 ± 0.1 min^{-1} is approximately 25-fold slower than k_{cat} for PCA, whereas K_M = 509 ± 42 μM is ~ 10-fold higher than K_M for PCA. The catalytic efficiency, k_{cat}/K_M, for DQCA is therefore ~250-fold lower than PCA, making it a significantly poorer substrate.

Next, we examined the pL-rate profile for DQCA reacting with PhdA in more detail (Figure 3.3 B and C). Unsurprisingly, in H_2O the bell-shaped pH profile is similar to that observed for the

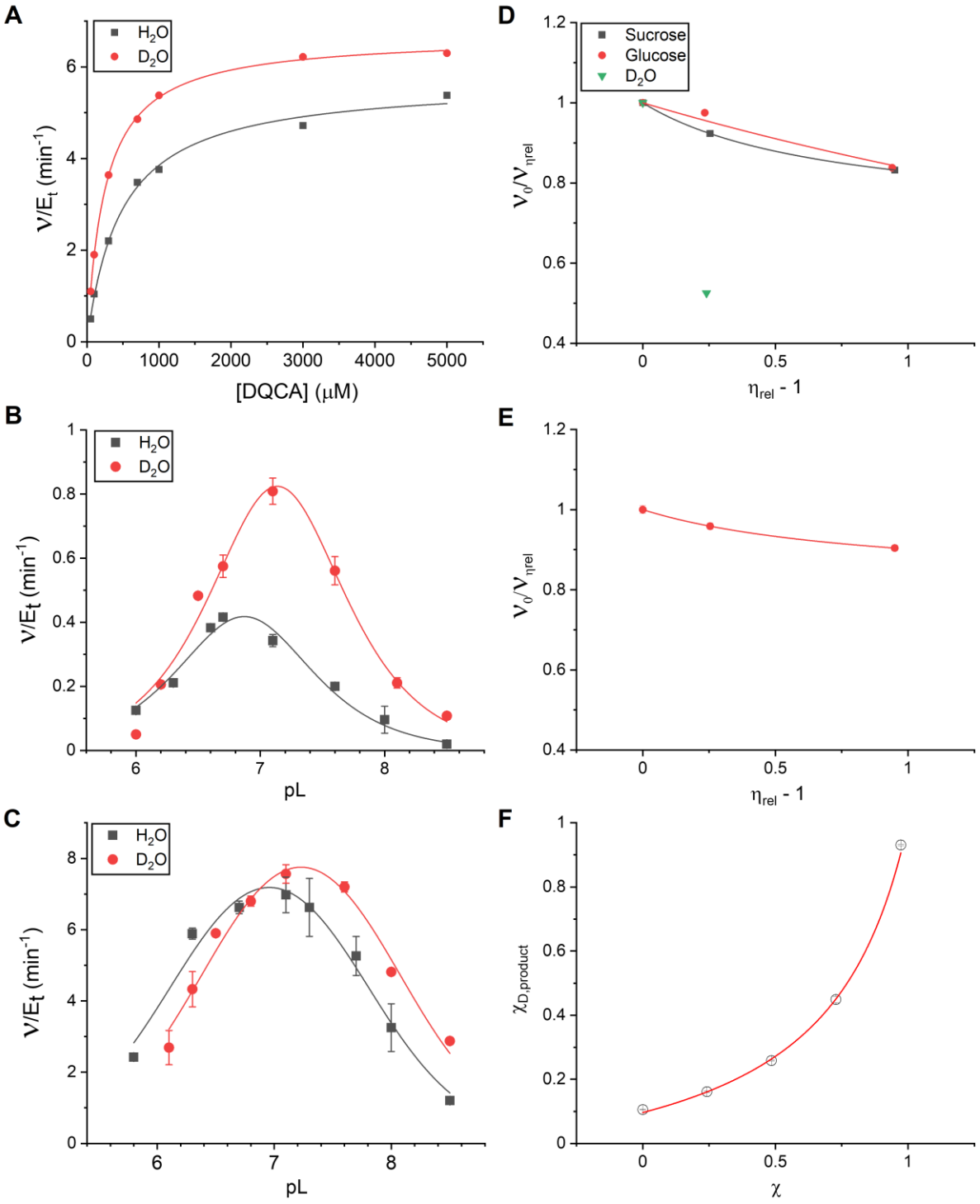


Figure 3.3: Solvent isotope and viscosity effects on the decarboxylation of DQCA by PhdA. **A:** Steady state kinetics for PhdA catalyzed DQCA decarboxylation at pL = 7 in H_2O and D_2O. **B and C:** pL-rate dependence for the reaction monitored under V_{max}/K_M and V_{max} conditions respectively. Data are fitted assuming 2 ionizable groups contribute to the rate profile. **D and E:** Solvent viscosity studies under V_{max}/K_M and V_{max} respectively. Data are fitted as per the equations described in 'Materials and methods'. **F:** Determination of ${}^{D_2O}V/K_P$ for the transfer of deuterium to DQ. The mole fraction of deuterated phenazine ($\chi_{D,product}$) is plotted as a function of solvent D-atom fraction (χ) and the isotope effect calculated by fitting the data to Eq. 1.

reaction of PCA. Under V_{max} conditions ([DQCA] = 10 mM; ~ 20.K_M), the acidic and basic limbs have apparent pK_as of 6.2 ± 0.1 and 7.7 ± 0.1, whereas under V_{max}/K_M conditions ([DQCA] = 30 μM; K_M/16) the apparent pK_as are 6.7 ± 0.3 and 7.1 ± 0.4 respectively. We then repeated the measurements in D_2O and determined the apparent pK_as as 6.4 ± 0.1, 8.0 ± 0.1 under V_{max} conditions and 6.9 ± 0.4, 7.4 ± 0.4 under V_{max}/K_M conditions. Like PCA, the pK_as are shifted ~ 0.3 pH units towards higher pD in D_2O. Under V_{max}/K_M conditions a substantial *inverse* SIE is evident, although under V_{max} conditions the solvent isotope effect is close to unity. We measured $^{D_2O}V/K_S$ (0.53 ± 0.01, n = 6) and ^{D_2O}V (0.9 ± 0.04, n = 5) at pL = 7. We also measured the values at the respective pL maxima and observed no significant differences ($^{D_2O}V/K_S$ ~ 0.5 to 0.6 for both measurements). Therefore, for simplicity, all future isotope effect studies were performed at pL=7. To verify that the *inverse* isotope effects did not arise from changes in viscosity, we examined the PhdA catalyzed decarboxylation of DQCA in the presence of various viscosogens (Figure 3.3 D and E). Although the rate of the reaction increased slightly with increasing viscosogen concentration, it cannot explain the significantly higher reaction rate in D_2O.

We also measured $^{D_2O}V/K_P$ for the transfer of deuterium into the product, 2,3-dimethylquinoxaline (DQ), by internal competition, using similar methodology and analysis as described for PCA (Figure 3.3F). In this case, the solvent KIE is much larger and *normal* ($^{D_2O}V/K_P$ = 4.15 ± 0.22). This value is quite typical for the transfer of a deuteron from a solvent-exchangeable residue to carbon.

In summary, the ^{D_2O}V, $^{D_2O}V/K_S$ and $^{D_2O}V/K_P$ measured for the decarboxylation of DQCA follow a similar trend to those observed with PCA, but with the corresponding SIEs becoming more pronounced. It appears that, because DQCA is a poor substrate for PhdA, the isotopically

sensitive steps in the mechanism become more rate limiting when compared to the physiological substrate, PCA.

3.3.3 Proton inventory analysis

The seemingly contradictory values of the $^{D_2O}V/K_S$ and $^{D_2O}V/K_P$ suggest the presence of more than one isotopically sensitive steps. While the *normal* $^{D_2O}V/K_P$ is most likely associated with proton transfer to product, the source of the *inverse* $^{D_2O}V/K_S$ is less clear. Proton inventory is a powerful tool that can identify exchangeable protons (or deuterons) that actively participate in a reaction.[68] This technique involves measuring reaction rates for the steady state parameters in mixed isotopic waters. The resulting curve can be analyzed according to the Gross-Butler equation. The virtual transition state theory could be applied wherever multiple steps are rate limiting and/or isotopically sensitive (see section 3.2.9 for a detailed explanation). The ability to differentiate between one proton and multi-proton mechanisms depends on the precision of the proton inventory data which in turn increases with increasing magnitude of the isotope effect.[66] Since DQCA provided a higher value for the *inverse* isotope effect (~2 as opposed to ~1.33 for PCA), proton inventory analyses were performed on V_{max}/K_M and V_{max} for DQCA (Figure 3.4).

For V_{max}/K_M conditions, the rate increases gradually with increasing D-atom fraction (Figure 3.4A). The data were fitted to the three simplest mechanisms: one transition state proton in flight (linear proton inventory, Eq.2); one reactant state proton (hyperbolic proton inventory, Eq. 3) and a medium effect (exponential proton inventory, Eq. 4).[68]

(2) $$\nu_\chi/\nu_0 = 1 - \chi + \chi.\phi_T$$

(3) $$\nu_\chi/\nu_0 = \frac{1}{1-\chi+\chi.\phi_R}$$

(4) $$\nu_\chi/\nu_0 = Z^\chi$$

Where, ϕ_T is the fractionation factor of the proton involved in the transition state, ϕ_R is the fractionation factor for a reactant state proton and Z is the value of the overall medium effect in D_2O. Although the data are plausibly fitted by a simple linear function (Eq. 2), for the isotope effect to arise from a single transition state proton its fractionation factor would need to be extremely high: $\phi_T = 1.92$. Such a high value for ϕ_T is unprecedented,[66] and hence unlikely. A hyperbolic proton inventory (Eq. 3) is attributed to the deprotonation of a reactant state proton with $\phi_R < 1$. Whereas Cys residues frequently display $\phi_R \sim 0.55$,[66] they are not implicated in the mechanism of PhdA and its crystal structure confirms that there are no active site Cys residues. Statistically too, a hyperbolic function fits the data least well: the midpoint isotope effect[48] calculated at $\chi = 0.5$ for a hyperbolic proton inventory is 1.31, whereas the experimentally determined value is 1.42 ± 0.04 (n=6). The *inverse* SIE is therefore unlikely to arise from a single reactant state proton. This leaves a medium effect (Eq. 4) as the most plausible interpretation of the data: i.e. the SIE arises from the combination of many small fractionation factors at many protonic sites involved in overall solvent reorganization or conformational change.

Under V_{max} conditions, a dome-shaped proton inventory is obtained (Figure 3.4B) which is diagnostic of competing *normal* and *inverse* isotope effects contributing to ^{D_2O}V. [75-77]The data may be fitted to Eq. 5 which describes this situation (see section 3.2.9 for derivation):

$$(5) \qquad \nu_\chi/\nu_0 = \frac{Z^\chi(1-\chi+\chi.\phi_T)}{1-\chi+\chi.w_7+\chi.\phi_T-\chi.w_7.\phi_T}$$

Here, Z is the value of the general medium effect in D_2O, ϕ_T is the transition state fractionation factor for the proton transfer and w_7 is its fractional contribution to V_{max}. Eq. 5 contains mutual dependency between parameters and therefore satisfactory fits cannot be obtained without 'fixing' a parameter. From the value of $^{D_2O}V/K_P$, the upper limit of ϕ_T is 0.24 (refer to section 3.2.9 for

further explanation). This can be used to estimate Z ~ 1.4 and w_7 ~ 0.1. Thus, ^{D_2O}V arises from a general medium effect (Z) offset by a single transition state proton (ϕ_T) contributing ~10% to V_{max}.

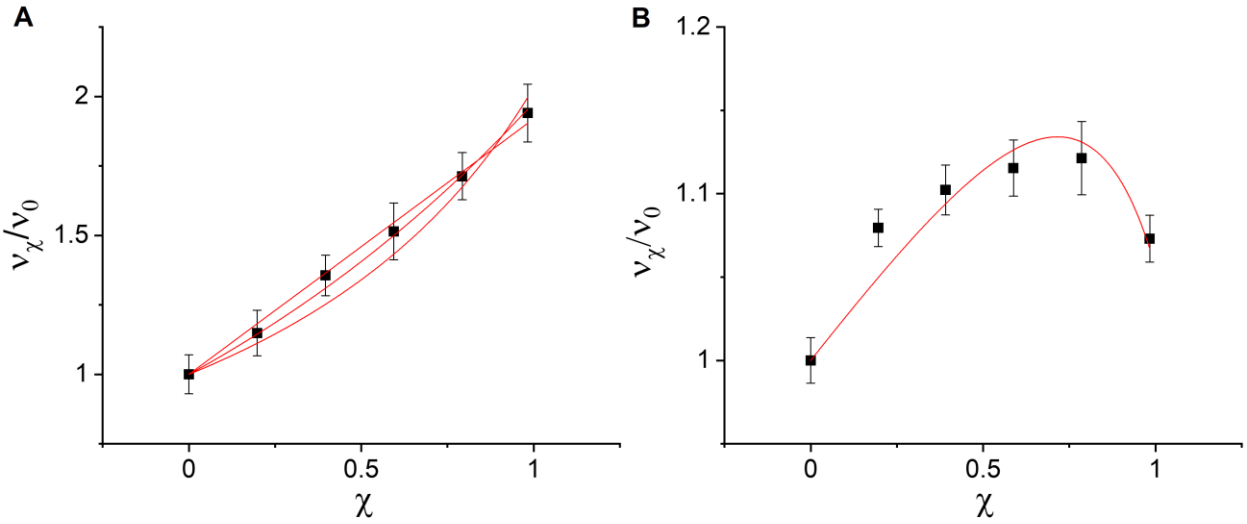


Figure 3.4: Proton inventory analysis of PhdA catalysed decarboxylation of DQCA. **A:** The reaction rate relative to 100 % H_2O plotted as a function of solvent D-atom fraction (χ) under low [DQCA] (V_{max}/K_M conditions). The data are fitted to Eq. 2 (top), Eq. 3 (bottom) and Eq. 4 (middle) described in the main text as indicated **B:** The reaction velocity relative to 100 % H_2O plotted as a function of χ under high [DQCA] (V_{max} conditions). The data are fitted to Eq. 5 described in the main text.

3.3.4 Origin of the medium effect in PhdA

While a proton inventory analysis predicts the type of isotope effect, it cannot provide information on the cause of the medium effect. However, it is well documented that D_2O alters the stability of proteins, which leads many proteins to unfold more slowly and at higher denaturant concentration in D_2O.[93-96] To examine if PhdA is more stable in D_2O, we compared the urea-induced unfolding of PhdA in buffered H_2O and D_2O by following the red-shift in intrinsic protein fluorescence (Figure 3.5 A, B). The normalized unfolding curves were plotted as a function of urea concentration (Figure 3.5C). We observe that PhdA unfolds at higher urea concentrations in D_2O than in H_2O. Fitting the unfolding curves to a simple two-state model of protein unfolding (see section 3.2.7) gave $K_{1/2}$ for unfolding in H_2O = 1.54 ± 0.03 M and ΔG_U = 18 ± 4 kJ.mol^{-1}. In D_2O

$K_{1/2}$ = 2.27 ± 0.04 M and ΔG_U = 23 ± 6 kJ.mol^{-1}. Thus, D_2O appears to stabilize PhdA by ~ 5 kJ.mol^{-1}, although there is significant error associated with this estimate.

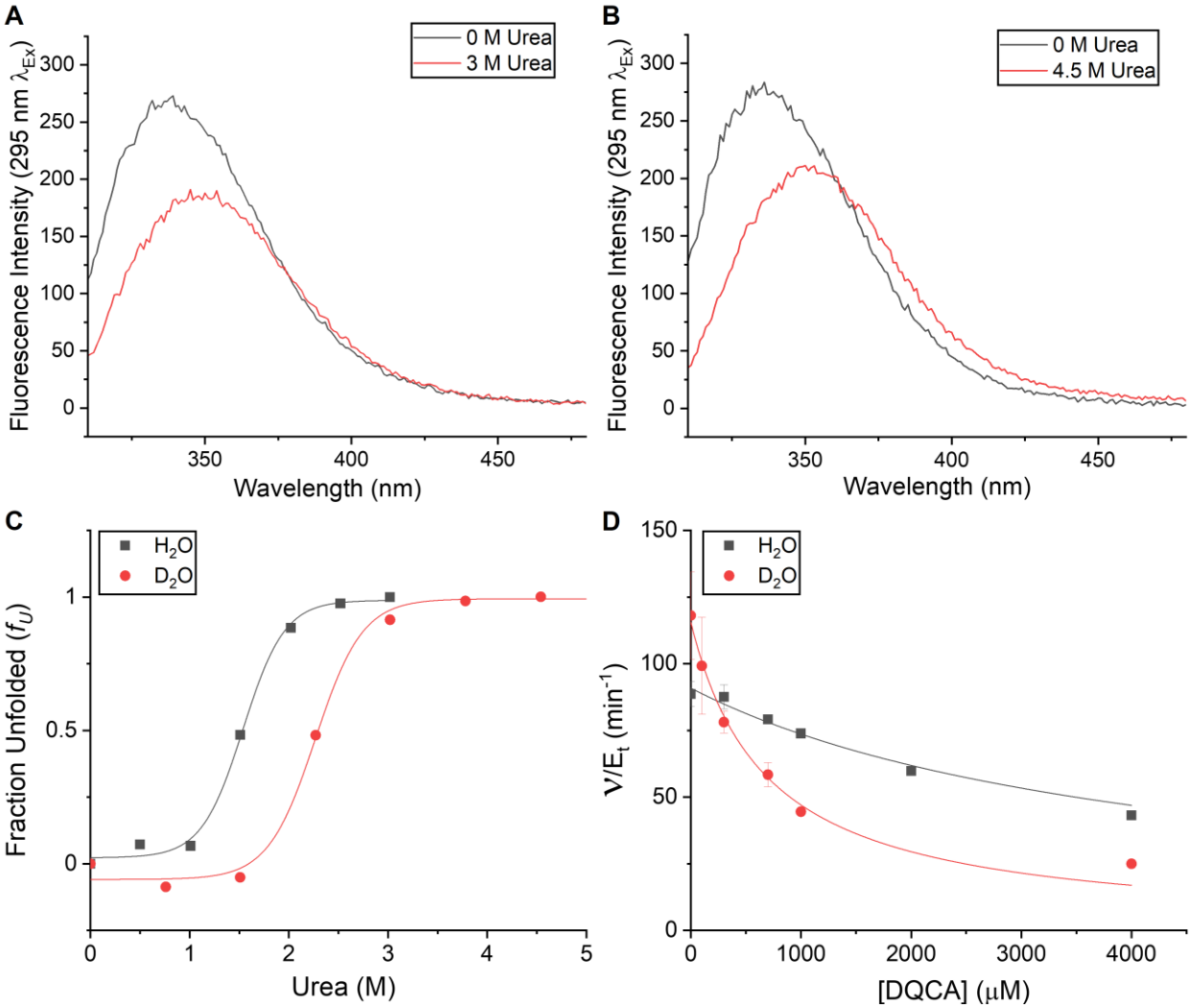


Figure 3.5: Origin of the medium effect in PhdA. **A and B:** Representative fluorescence emission spectra (excitation at 295 nm) for folded and unfolded PhdA (with urea) in phosphate-buffered H_2O and D_2O respectively. **C:** Normalized fraction of unfolded protein is plotted as a function of urea concentration. **D:** Competitive inhibition of PCA decarboxylation by the slow substrate DQCA in buffered H_2O and D_2O. The normalized residual rate (v/E_t) of PCA decarboxylation is plotted as a function of DQCA concentration.

3.3.5 Conformational stability affects reaction kinetics.

With evidence that D_2O affects the stability of the free enzyme, we sought to identify how this effect can manifest in ${}^{D_2O}V/K_S$. V_{max}/K_M provides information on all the steps from free enzyme up to and including the first irreversible step.[74] Thus, we hypothesized that this stable conformer

in D_2O favors formation of the Michaelis complex and subsequent steps in the reaction, leading to an *inverse* $^{D_2O}V/K_S$. To verify this, we first measured the K_M of PCA and DQCA in H_2O and D_2O at pL=7 (Figures 3.2A and 3.3A). We observed that in both cases, while k_{cat} doesn't change significantly, the K_M is appreciably lower in D_2O (PCA: K_M = 49.9 ± 0.7 μM in H_2O and 28.6 ± 1.7 μM in D_2O; DQCA: K_M = 509 ± 42 μM in H_2O and 252 ± 6 μM in D_2O). The lower K_M value suggests that D_2O favors the capture of substrate to form an effective complex that is poised for turnover.[97]

Although changes in K_M are often used as a proxy for K_d, it is well known that this assumption is unreliable. We therefore exploited the fact that the slow substrate DQCA acts as a competitive inhibitor of the physiological substrate, PCA, to measure apparent inhibition constants K_I^{app} for DQCA in H_2O and D_2O. The inhibition curves obtained with [PCA] = 50 μM are shown in Figure 3.5D. The change in inhibition is quite striking: in D_2O DQCA behaves as a much more potent inhibitor, K_I^{app} = 250 ± 30 μM, than in H_2O, K_I^{app} = 2100 ± 300 μM. The drastic lowering in K_I^{app} indicates that the decrease in K_M could be related to formation of a more productive Michaelis complex that drives subsequent steps.

3.3.6 Molecular dynamics simulations show conformational differences in the two solvents.

To understand the microscopic origins of the conformational stability exhibited by PhdA in the two solvents, we performed all-atom MD simulations using its crystal structure (PDB:7PDA), over a period of 1 μs in the presence of explicit H_2O or D_2O molecules. In common with other UbiD-like enzymes, the tertiary structure of PhdA comprises an N-terminal prFMN binding domain, a central α-helix, an oligomerization domain and a C-terminal α-helix.[41] Although PhdA is hexameric, the simulations were performed using a monomer, a simplification that greatly

expedited the calculations, but was not expected to affect the behavior of active site residues which are located away from the protein-protein interfaces (Figure 3.6).[98, 99]

Figure 3.6: Structure of the PhdA hexamer. For one of the subunits, the domains are color-coded as follows: N-terminal prFMN binding domain (green), central α-helix (orange), oligomerization domain (magenta), C-terminal α-helix (white), active site loop (cyan). The cofactor (yellow) and relevant active site residues are displayed in spheres.

Analysis of the root mean squared deviation (RMSD) of the protein backbone from the starting structure showed a higher value for the protein in H_2O than in D_2O during the initial equilibration period (Figure 3.7A). The RMSD was stable for the remainder of the trajectories,

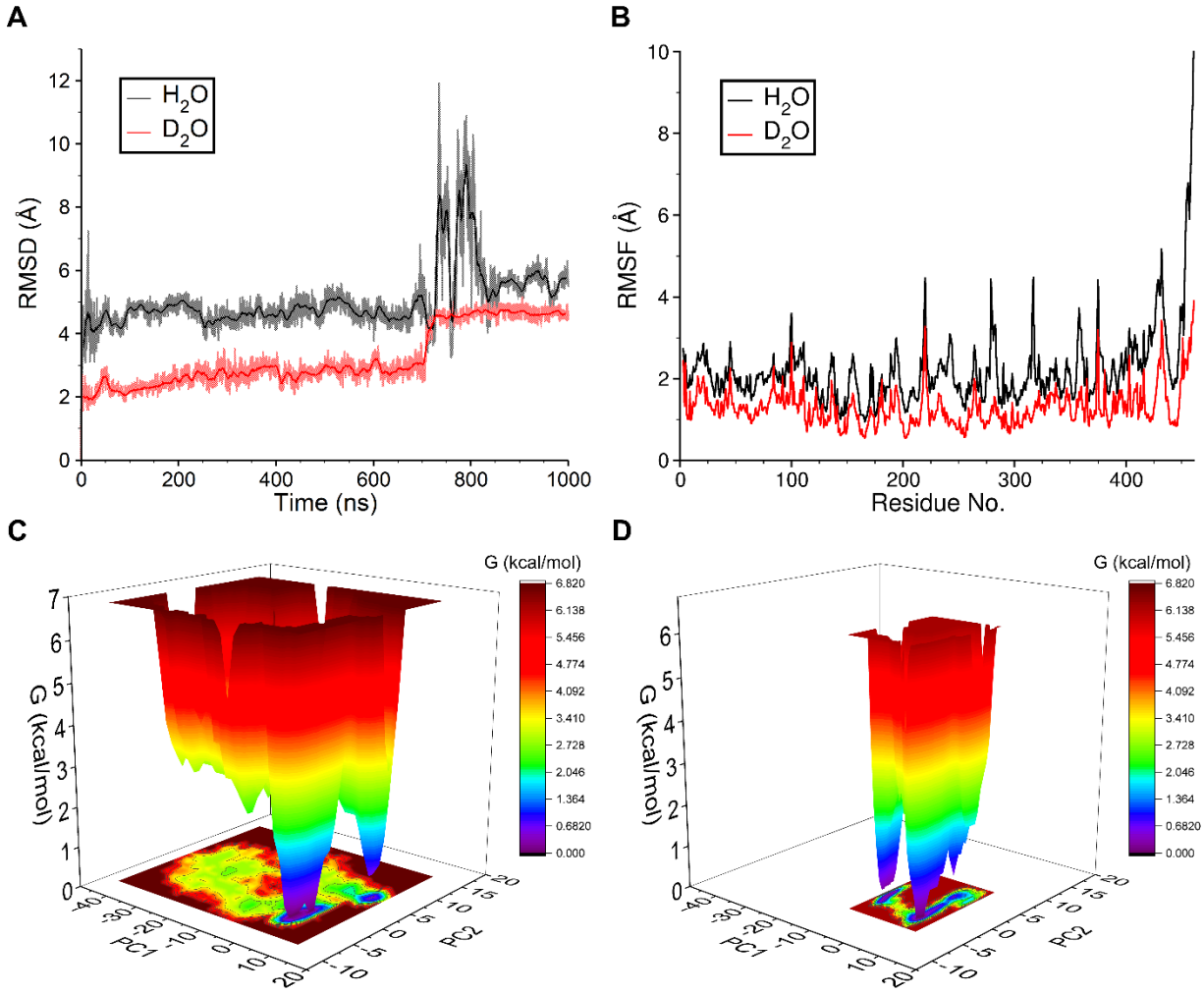


Figure 3.7: MD simulations of PhdA in H_2O and D_2O. **A:** RMSD for the protein backbone throughout the simulation run. **B:** RMSF for protein residues. Free energy landscapes for PhdA monomer in **C:** H_2O, and **D:** D_2O. Colorbar includes the energy values. The x and y axes represent two principal components (PCs) of the proteins with the highest variations. PCs are extracted using a multivariate statistical technique called Principal Component Analysis and represent maximum protein dynamics in lower dimensions. For **A**, translucent background lines display data points from the entire simulation trajectory (1,000,000 points) while the solid foreground lines display smoother running averages considering every 2,000 points. Analysis in **B** was carried out for the last 500 ns of the simulation trajectories

except for the data at ~750 ns, at which time, larger RMSD values were observed. This perturbation arises from the movement of the C-terminal helix which can be observed in the visual simulation. However, this is likely an artifact because in the quaternary structures of UbiD-like enzymes, the C-terminal helix normally makes extensive contacts with other protein subunits.[41] Excluding the C-terminal helix from the RMSD analysis removed the discontinuity. The residue-wise root mean squared fluctuation (RMSF) was calculated for the final 500 ns of the MD trajectories. Consistent

with the RMSD analysis, residues in the H_2O-solvated system showed greater fluctuations than those in D_2O, even after excluding the C-terminal helix (Figure 3.7B). These data indicate that PhdA is less conformationally mobile when solvated in D_2O, compared to H_2O.[100]

For validation, protein free energy landscapes (FELs), which provide a statistical description of the various possible states explored by a protein in MD simulations, were obtained through the Boltzmann inversion of the joint probability density function of the principle components (PCs).[101] The protein in H_2O (Figure 3.7C) displayed a significantly higher exploration of the conformational space as compared to the protein in D_2O (Figure 3.7D). In agreement with our experimental results, the D_2O-solvated protein FEL exhibited prominent narrow troughs indicating highly stable conformations that dominated throughout the simulations, as opposed to broader troughs and multiple conformers separated by low energy barriers in the H_2O-solvated system. Consistent with this analysis, 80 unique protein conformers were identified for the H_2O-solvated protein whereas only 7 were found in D_2O. These conformers were recognized by calculating the RMSDs between all structures throughout the trajectory and segregating them based on a RMSD cutoff of 2 Å.[102]

Overlaying the dominant conformers identified in H_2O and D_2O established significant differences between the two, with an RMSD of 3.82 Å (Figure 3.8). Based on their crystal structures, UbiD-like enzymes are known to exhibit distinct ‘open’ and ‘closed’ conformers. The distance between the centers-of-masses (COM) of residues R159 and I416 (residues that are broadly conserved in UbiD-like enzymes) serves as a convenient metric to measure the openness of the active site.[30] In the crystal structure of PhdA, solved without substrate bound, the R–I distance is ~14.3 Å, corresponding to an ‘open’ conformer. However, in H_2O this distance increased to ~15.7 Å in the dominant conformer, whereas in D_2O it decreased to ~11.2 Å. Surprised

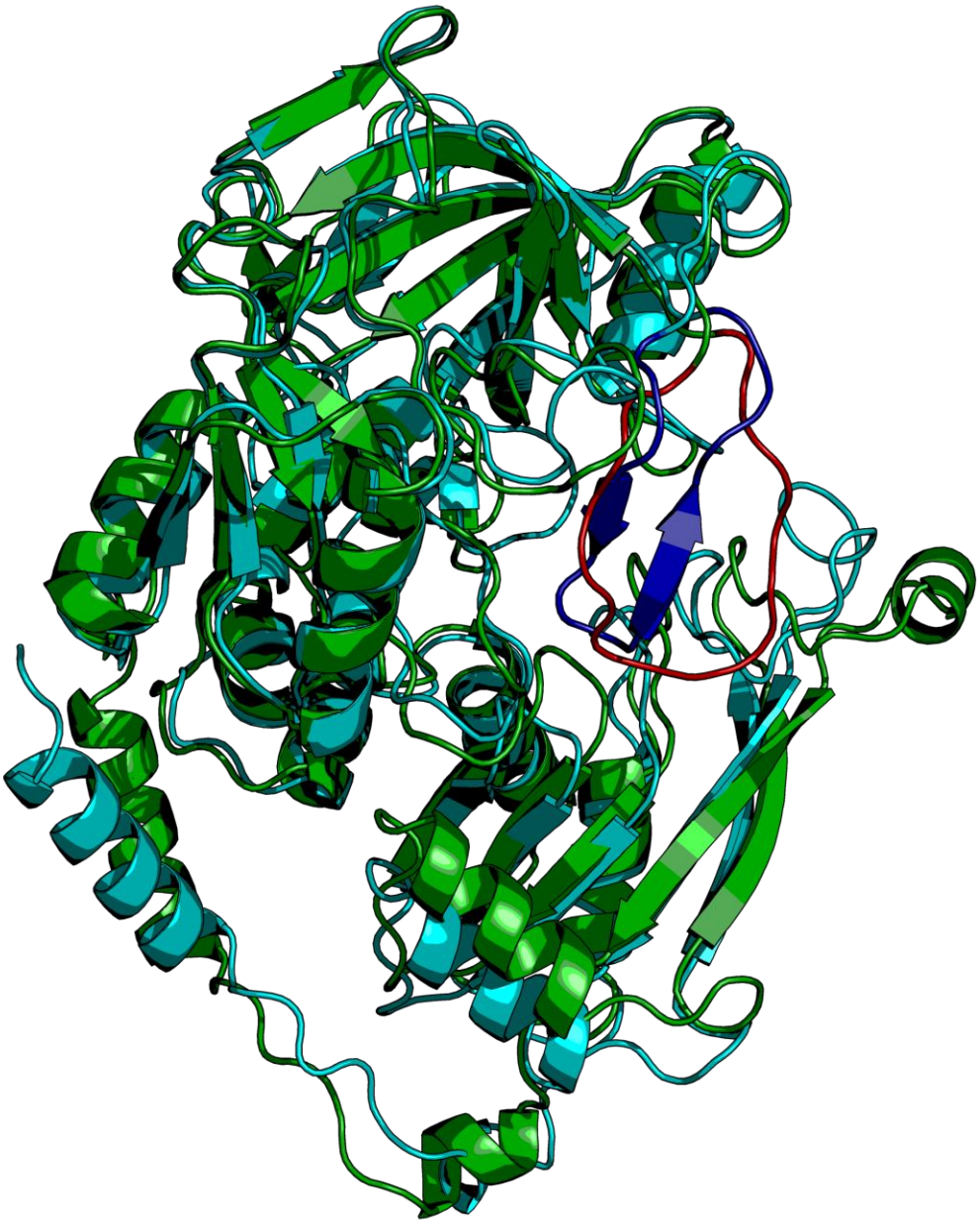


Figure 3.8: Overlay of the dominant conformers in H_2O (green) and D_2O (cyan). The active site loop (residues 264-279) is colored in red for the H_2O conformer and blue for the D_2O conformer.

by this result, we calculated R–I distances over the entire simulation trajectory. In H_2O, the R–I distance progressively increased from ~14.3 Å to ~17.8 Å (Figure 3.9A) suggesting significant further domain opening. In contrast, in D_2O, this distance reduced to ~11.5 Å (Figure 3.9B) indicating its evolution to a more 'closed' conformer (Figure 3.9C). The RMSF data coupled to a visual inspection of the simulation showed that the active site loop, comprising residues 264-279,

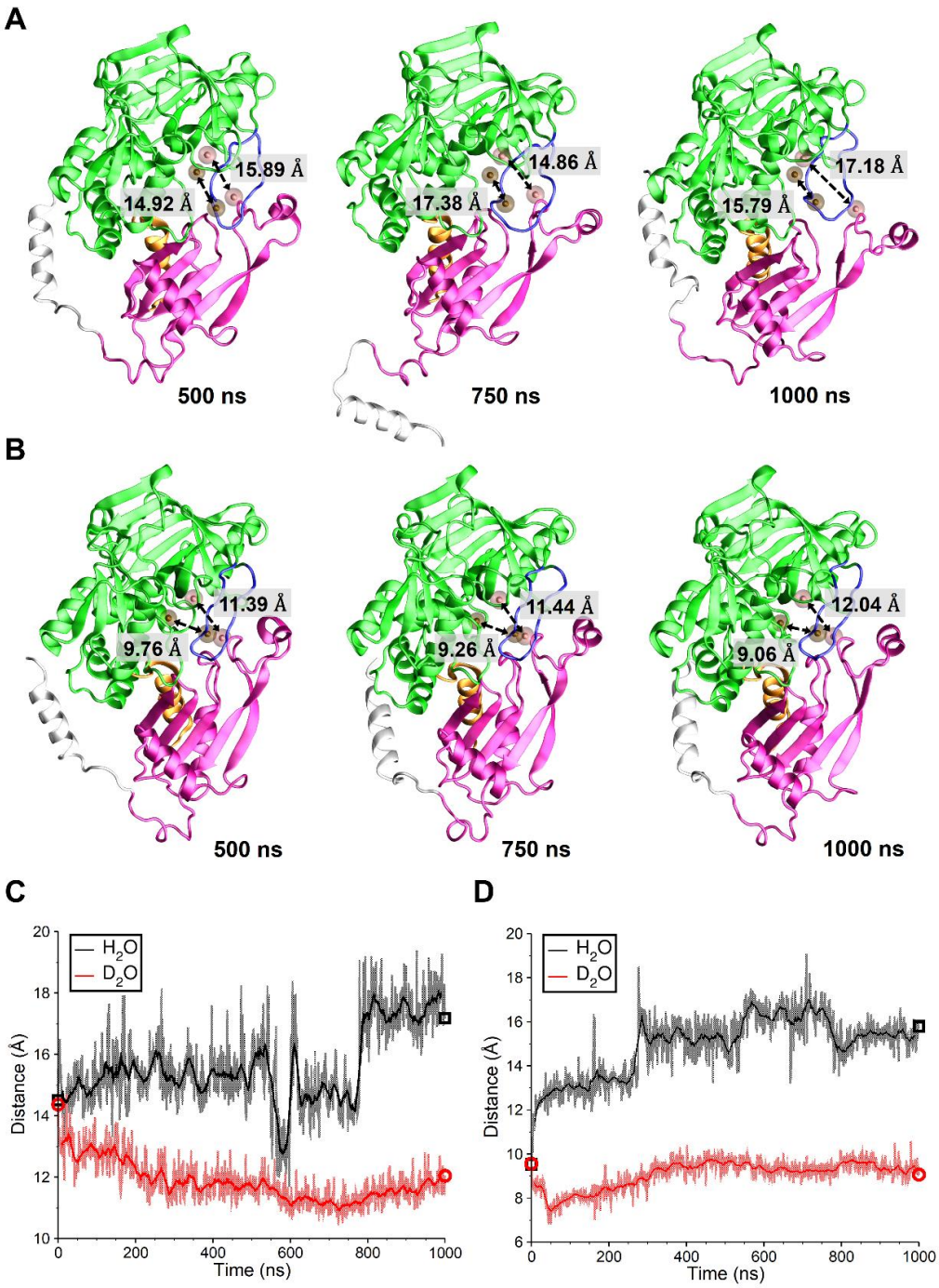


Figure 3.9: Representative snapshots at 500 ns, 750 ns and 1000 ns for PhdA in **A:** H_2O, and **B:** D_2O. The domains are color-coded as follows: N-terminal prFMN binding domain (green), central α-helix (orange), oligomerization domain (magenta), C-terminal α-helix (white), active site loop (blue). The COMs of R159 and I416 are presented in pink whereas COMs of prFMN and E269 in brown. For each snapshot, the respective COM distances are given in boxes. Temporal evolution of COM distances for **C:** R159–I416 and **D:** prFMN–E269.

that contains the catalytic residue E269 was more structured and compact in D_2O. In line with this observation, the distance between the centers-of-masses of the cofactor and E269 is short in D_2O

than H_2O (Figure 3.9D). Our simulations suggest that D_2O leads to significant structural changes in the active site which can affect catalysis.

3.3.7 D_2O promotes intra-protein hydrogen bonds resulting in a more compact structure.

To ascertain why these solvent-specific differences exist in protein conformations, the protein-solvent and intra-protein hydrogen bonds were calculated over the last 500 ns of the trajectories. In D_2O the protein exhibited fewer protein-solvent hydrogen bonds (Figure 3.10A) and more intra-protein hydrogen bonds (Figure 3.10B). Additionally, the hydrogen bond autocorrelation functions were calculated to analyze their lifetimes in both simulations. Interestingly, it was found that the probability of persistence of hydrogen bonds decayed slower in D_2O for both protein-solvent and intra-protein hydrogen bonding. The intra-protein hydrogen bonds were also generally found to persist significantly longer than hydrogen bonds with the solvent. These observations agree with the work of Sheu et al., who found that D_2O did not facilitate as significant a decrease in the activation energy for hydrogen bonding as H_2O resulting in a decreased decay rate.[103] Additional analyses regarding the hydrogen bonding behavior in both simulations are presented in section 3.2.11.

Previous studies suggest that proteins compress in D_2O, leading to a more compact shape.[96, 103] To validate this behavior, the solvent accessible surface area (SASA) and volume of the protein was also calculated for the simulations. It was observed that the D_2O-solvated protein showed lower values of SASA (Figure 3.10E) and volume (Figure 3.10F) throughout the time course, indicating that D_2O induced a more compact protein conformation. This analysis shows that the lower R-I distances and a more compact active site loop can be rationalized through the differences in protein-solvent and intra-protein hydrogen bonding between the two solvents.

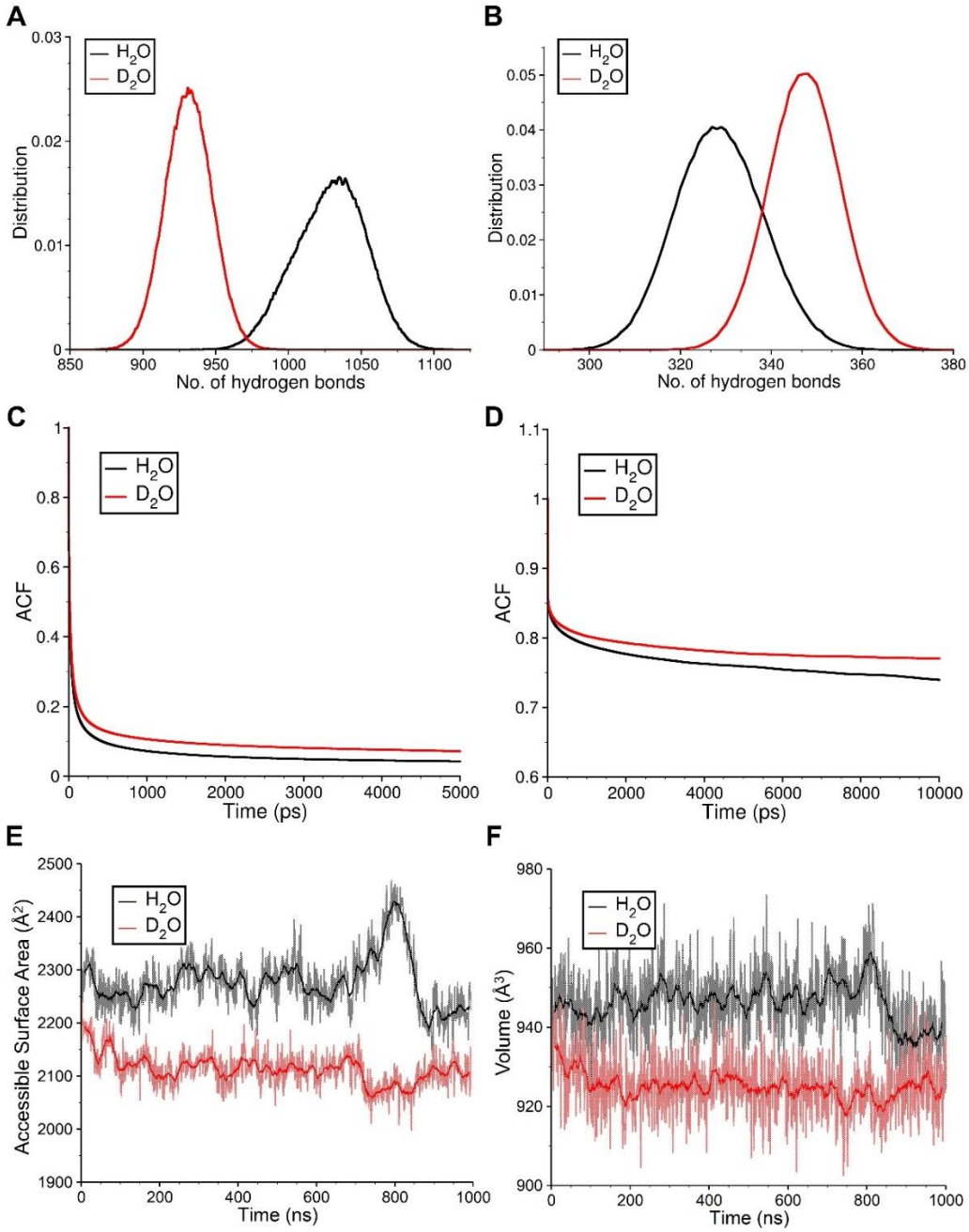


Figure 3.10: Differences in hydrogen bonding and surface accessible area in PhdA in H_2O and D_2O. Distribution of number of **A:** protein-solvent, and **B:** intra-protein hydrogen bonds. Autocorrelation functions (ACF) for **C:** protein-solvent, and **D:** intra-protein hydrogen bonding. Analyses in A-D were performed for the final 500 ns of the simulation trajectory. Evolution of protein **E:** solvent accessible surface area (SASA), and **F:** volume, throughout the simulation. For **E**, and **F**, translucent background lines display data points from the entire simulation trajectory (1,000,000 points) while the solid foreground lines display smoother running averages considering every 2,000 points.

3.4 Discussion

Recent studies on UbiD-like enzymes have suggested the importance of domain motions in catalysis. Stopped flow spectroscopy indicates that the dimeric FDC exhibits negative cooperativity between the two subunits arising from a conformational change that facilitates inter-conversion between the 'fast, tight' and the 'slow, loose' active sites.[42] Docking studies on the 'open' and 'closed' conformers of vanillic acid decarboxylase (VdcCD complex) suggest that lower binding energies for substrate (**E.S** complex from Figure 1.9), enzyme-substrate adducts (**Int1**) and enzyme-product adduct (**Int3**) are associated with domain closure whereas the decarboxylated intermediate (**Int2**) is stabilized by a more 'open' conformer.[30] Moreover, it is proposed that the favorable conformational change from the 'open' to 'closed' involved in the conversion of **Int2** to **Int3** would help to overcome the presumably unfavorable dearomatization of **Int2**. For the best studied enzyme, FDC, the mechanism of decarboxylation is quite well defined, however the kinetics and mechanisms of most other UbiD-like enzymes remain poorly understood.

We previously showed that the recently discovered UbiD-like enzyme, phenazine-1-carboxylate decarboxylase, (PhdA) could decarboxylate a range of polyaromatic carboxylic acids. Here we have focused on defining the kinetics and mechanism of the enzyme by using the slow substrate, DQCA to probe the nature of the rate-determining step. Our studies suggest that domain motions are important in the mechanism of PhdA and that these may, in part, account for the unusual kinetic behavior observed in D_2O, which is accentuated by using the slow substrate DQCA. The experimental observations are consistent with, and to some extent rationalized by, our MD simulations of the response of the protein structure to D_2O.

MD simulations on PhdA in H_2O and D_2O suggest that D_2O increases intra-protein H-bonding, which is consistent with the increase in ΔG_{unfold} observed in D_2O. The simulations also indicate that D_2O affords a more compact structure to the protein, including stabilizing the active site loop as well as promoting domain closure towards a more 'closed'-like conformer. Comparing to previous studies, these features are expected to stabilize the Michaelis complex and subsequent reaction intermediates. The lower K_M values and the significantly tighter K_I^{app} measured for DQCA in D_2O are consistent with the MD simulations. Moreover, the D_2O-induced conformational shift towards the catalytically active form of the enzyme rationalizes the inverse ${}^{D_2O}V/K_S$ effect.

Our study indicates that while a *normal* SIE effect is observed for deuterium incorporation into product, it doesn't affect ${}^{D_2O}V/K_S$. This means that the proton transfer step is not rate limiting for V_{max}/K_M and/or doesn't feature in its mathematical expression. We propose that 'Mechanism 3' described earlier (section 3.2.9) best explains the KIE data. Here, k_7 is the rate constant for proton transfer and displays a *normal* intrinsic isotope effect (${}^{D_2O}k_7 = k_7/k_{7D}$). We assume that CO_2 release is functionally irreversible, effectively dividing the reaction into two halves. Consequently, ${}^{D_2O}V/K_S$ is governed only by steps k_1 through k_5 and therefore, doesn't include ${}^{D_2O}k_7$.

CO_2 release and subsequent re-aromatization might lead to domain opening and stabilization of the **E.J** complex.[30] At this point E269 has to exchange a proton with the solvent in order for ${}^{D_2O}k_7$ to be expressed in ${}^{D_2O}V/K_{product}$.[78] Since subsequent intermediates (**E.M**) are also stabilized on the trajectory of domain closure, D_2O should drive the reaction forward. This is reflective in the proton inventory of V_{max}, where the rate initially increases with increasing solvent D-atom fraction. But, at high D-atom fractions, proton transfer becomes significantly rate limiting and the advantage provided by the *inverse* medium effect diminishes, leading to a dome shaped

proton inventory. This also implies that the medium effect is more likely a 'general' effect encompassing multiple kinetic steps. Moreover, previous studies show that microviscosogens stabilize protein conformers[104] and *inverse* solvent viscosity effects have been reported in conjunction with *inverse* solvent isotope effects.[50, 105] The fact that we observe slightly *inverse* viscosity effects for DQCA decarboxylation suggests that microviscosogens also stabilize domain closure in PhdA, albeit not to the same extent as D_2O.

Overall, the kinetics of PhdA decarboxylation indicate that conformational changes are significantly rate limiting for both V_{max}/K_M and V_{max}. Even for the poor substrate DQCA, under V_{max} conditions, proton inventory analysis shows that proton transfer is only ~10% rate limiting. Conformational switching between 'open' and 'closed' forms appears to be a general feature of UbiD-like enzymes, although whether such motions are kinetically significant in the decarboxylation reactions catalyzed by other members of this enzyme family remains to be determined. We note that there is intense interest in using UbiD-like decarboxylases for biocatalytic applications. In this context, our results suggest that engineering these enzymes (or optimizing solvent systems) to increase the stability of the 'closed' form may have the added benefit of improving their catalytic efficiency.

Chapter 4 Biosynthesis and Maturation of prFMN

4.1 Introduction

The biosynthesis and maturation of prFMN is central to the activity of UbiD-like (de)carboxylases. prFMN is synthesized through UbiX-catalyzed prenylation of reduced flavin mononucleotide ($FMNH_2$) using dimethylallyl phosphate (DMAP) or dimethylallyl pyrophosphate (DMAPP) as prenyl donors. Crystallography reveals that a nucleophilic attack of the FMN-N5 on the prenyl moiety occurs in the first step to form a FMN-N5-prenyl-C1' bond.[23] The presence of a dimethylallyl group is crucial for this initial bond formation, suggesting an S_N1 type mechanism.[25] Next, the prenyl group of this long-lived N5-C1' intermediate (observed decay constant = 0.316 ± 0.002 s^{-1}) converts to an electrophile by accepting a proton and subsequently cyclizes with FMN C6 through a Freidel-Crafts alkylation. Finally, the enzymatic residues S15 and E49 (*Pseudomonas aeruginosa* UbiX – *Pa*UbiX numbering) are proposed to facilitate re-aromatization of the flavin ring.

While the UbiX-catalyzed reaction forms reduced prFMN ($prFMNH_2$), an iminium ($prFMN^{iminium}$) form of the cofactor is catalytically active, showing the need for oxidative maturation.[4] Analogous to the oxidation of flavins, incubating $prFMNH_2$ with a UbiD-like enzyme followed by oxidation is proposed to form a purple radical semiquinone ($prFMN^{radical}$) and superoxide. Loss of the superoxide stalls this maturation at the stable $prFMN^{radical}$, which can be observed in UbiX via EPR spectroscopy.[27] The UbiD-like enzyme active site presumably traps this superoxide and promotes the formation of a putative prFMN-C4a-peroxide (prFMN-C4a-OO^-

) which abstracts a proton from C1' of the prenyl moiety and is eliminated, yielding prFMN^{iminium} (Figure 1.6).[27, 28] The conserved residues R173, E277 and E282 (*An*FDC numbering) are proposed to assist in the maturation.[28]

Recently, Balaikaite *et al* studied the oxidation of prFMNH$_2$ in free solution, without the addition of any UbiD-like enzyme (hereafter, prFMNH$_2$ oxidized in the absence of any enzyme would be referred to as prFMN^{ox}). Upon O_2 exposure, they observed the rapid formation (k_{obs1} = 0.5 ± 0.3 min^{-1}) of a new species (species B) which decayed slowly (k_{obs2} = 0.078 ± 0.015 min^{-1}) to another species (species C). The authors noted significant discrepancy in the observed rate constants depending upon sample aeration and under micro-aerobic conditions, species B was stable for hours. Based on UV-Vis and EPR spectroscopy, species B was identified as prFMN^{radical}. The UV-Vis spectrum of species C appeared similar to that of FMN but its mass spectrum showed a peak at 559.18 m/z. Ferrous-oxidized xylenol-orange (FOX) assay and surface enhanced raman spectroscopy (SERS) revealed the presence of peroxides in prFMN^{ox} and therefore, species C was assigned as a prFMN-hydroperoxide (prFMN-C4a-OOH).[27]

Thus, as per the proposed mechanism, the oxidation of prFMNH$_2$ in the presence or absence of a UbiD-like enzyme proceeds via nearly identical intermediates (Figure 1.6). This suggests that adding prFMN^{ox} (containing prFMN^{radical} and/or prFMN-C4a-OOH) to a UbiD-like enzyme should lead to cofactor maturation. However, the available data on prFMN maturation unequivocally shows that decarboxylation activity is only observed when prFMNH$_2$ is incubated with UbiD-like enzymes prior to oxidation.[4, 11, 32, 64] In fact, the addition of prFMN^{ox} to holo-FDC led to a reduction of enzyme activity.[27] This discrepancy is explained by suggesting that either the transient formation of prFMN-C4a-OO$^-$ is crucial for proper cofactor maturation or that prFMN maturation doesn't proceed via peroxy-adducts. Overall, the oxidative maturation of prFMN is a

complex process, and alternative mechanisms need to be proposed for explaining the apparently contradictory data.

Here, we studied the oxidation of $prFMNH_2$ through LC-MS and discovered multiple peaks in the reaction (peaks Q and R). Upon prolonged oxidation, we observed the formation of the 559.18 m/z species reported previously and, surprisingly, FMN. Through solvent deuterium labelling, we assigned unique chemical structures to peaks Q and R. Finally, we examined the role of these peaks in the oxidative maturation of prFMN and developed an alternate mechanism that reasonably explains both the present and the previous data.

4.2 Materials and Methods

4.2.1 Enzymatic synthesis of $prFMNH_2$ and prRiboflavin

UbiX-catalyzed $prFMNH_2$ synthesis with FMN (100 – 400 µM) and DMAP (usually 5 equivalents w.r.t. FMN) was performed as previously described (Section 2.2.4). For solvent deuterium labelling experiments, 0.2 M Tris base was titrated with DCl in D_2O. The solvent D-atom fraction was calculated by accounting for the H-atoms added and appropriate corrections were applied to the pH meter measurement to obtain a pD of 7.2 (refer to section 3.2.4). Similarly, 2.5 M KCl was dissolved in D_2O. All buffers and D_2O were purged with N_2 gas overnight and transferred to an anaerobic chamber (Coy). Concentrations of all reaction components were identical to the H_2O reactions. The final solvent D-atom fraction in the reaction was ~ 0.8.

For reactions with riboflavin, 10 mM riboflavin was solubilized in H_2O with 50 mM NaOH. The solutions were prepared fresh each time and used immediately to avoid decomposition of riboflavin. Reactions consisted of 100 – 400 µM riboflavin and DMAP (usually 5 equivalents w.r.t. riboflavin). The concentrations of all other components were identical to those of the FMN reactions.

4.2.2 Reconstitution of PhdA with different prFMN fractions

Reconstitution with PhdA was performed as described previously (section 2.2.4). For 'aerobic' reconstitution, free prFMNH_2 was oxidized by exposure to O_2 prior to incubating it with PhdA and subsequently desalted.

4.2.3 PhdA activity assays

Decarboxylation assays for reconstituted PhdA were performed under standard conditions as previously described (section 2.2.5). Briefly, 100 µM PCA was reacted with 0.2 µM PhdA in 20 mM Bis-Tris-Cl (pH 6.5). The reactions were quenched at different time-points and analyzed by HPLC. Reaction rates normalized by the enzyme concentration (min^{-1}) were compared for various enzyme fractions to determine reconstitution efficiency.

4.2.4 HPLC and LC-MS analysis

HPLC and LC-MS solvents, columns and elution methods were as previously described (sections 2.2.6 and 2.2.7). Precipitated UbiX was separated from the reaction mixture through centrifugation. To study prFMNH_2 oxidation under micro-aerobic conditions, the reduced reaction mixtures were removed from the anaerobic chamber and immediately transferred to HPLC vials with spring-bottom inserts, capped tightly and injected on a Shimadzu LC-20 system. This afforded a 'slower' oxidation, owing to the slow diffusion of O_2 into the sample. For vigorous oxidation, the reduced reaction mixtures were exposed to O_2 by vigorously tapping the vials or by pipetting air into the sample before HPLC analysis.

4.3 Results

4.3.1 UbiX reaction produces multiple products

In order to study the oxidation of prFMNH$_2$ in free solution, Balaikaite *et al* primarily utilized UV-Vis spectroscopy.[27] However, Wang *et al* studied the same process through HPLC and observed multiple compounds.[26] In section 2.3.3, we too described the presence of various species while studying the *in vitro* reconstitution of PhdA through HPLC. Therefore, we argued that a thorough HPLC-based analysis of prFMN^{ox} would be beneficial. Since our HPLC method is ~40 min long, an obvious limitation of this assay is the lack of accurate information regarding the temporal evolution of prFMNH$_2$ to species C via species B. However, given that species B was stable for hours under micro-aerobic conditions[27], prFMNH$_2$ was transferred to HPLC vials with spring-bottom inserts and capped tightly before HPLC analysis. This allowed the oxidation to proceed under micro-aerobic conditions, leading to the detection of species B. Moreover, the slow degradation of species B could be studied through subsequent HPLC injections. Similarly, species C could be studied by oxidizing prFMNH$_2$ overnight under micro-aerobic conditions or by vigorously tapping the vial/pipetting air into the sample before HPLC analysis.

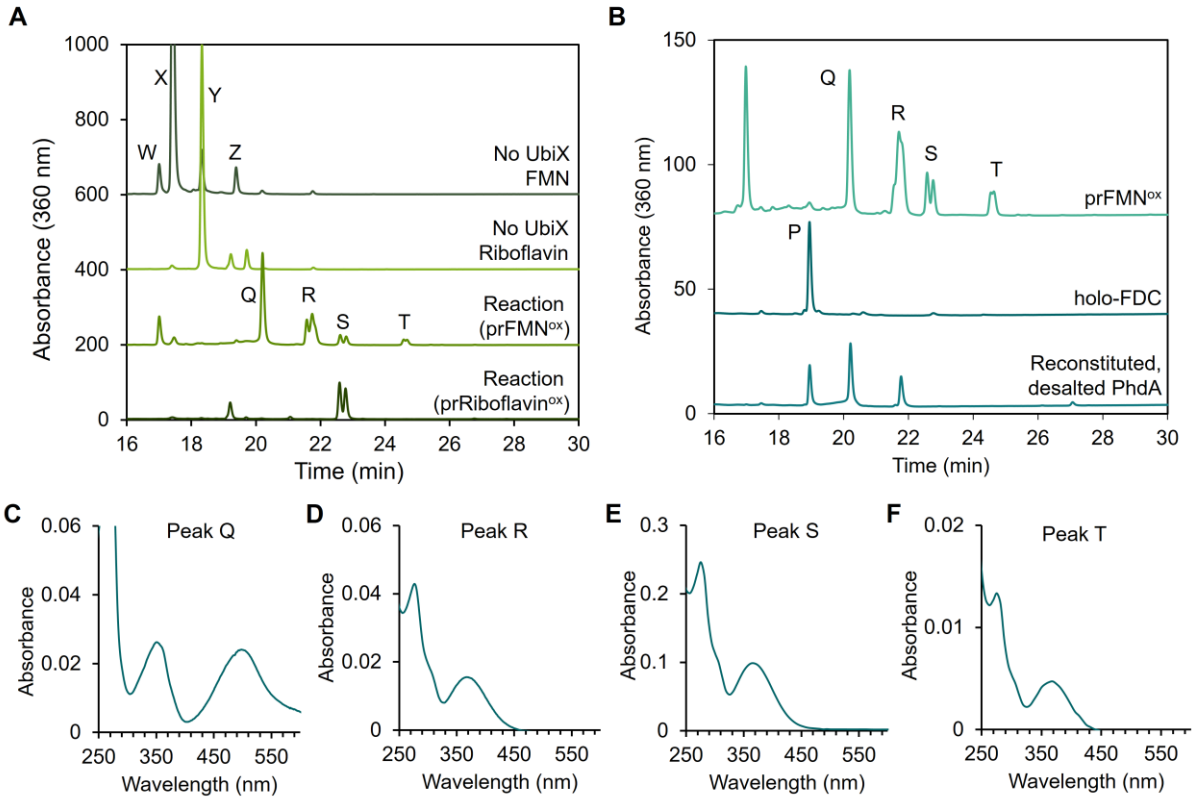
Initially, we assessed the purity of FMN through LC-MS and detected 4 distinct peaks with flavin-like UV-Vis spectra (peaks W, X, Y and Z, Figure 4.1A). Based on their mass spectra, we identified peaks W and X as nucleotide isomers of FMN (both 457.11 m/z) and peak Y as riboflavin (377.14 m/z), all of which are known to exist in commercially available FMN.[106] Similarly, peak Z (359.13 m/z) is suggestive of a loss of H_2O from riboflavin.

Upon studying the UbiX-catalyzed synthesis of prFMNH$_2$ and subsequent oxidation under micro-aerobic conditions, we made an interesting observation. With the exception of peak W (likely an inactive nucleotide isomer of FMN), all other peaks are consumed in the reaction with

UbiX, leading to the formation of four new peaks (peaks Q, R, S and T, Figure 4.1A). Based on their mass spectra, peaks Q (525.17 m/z) and R (527.19 m/z) represent prenylation of FMN, peak S (447.22 m/z) is prenylated riboflavin and peak T (429.21 m/z) is prenylated peak Z (Figure 4.1G). The UV-Vis spectra of peaks R, S and T are identical (Figure 4.1D, E and F), suggesting similar modifications to the isoalloxazine moiety whereas their mass spectra indicate that the flavin core is reduced. On the other hand, the UV-Vis (Figure 4.1C) and mass spectra of peak Q resemble an oxidized flavin core with distinct chemical modifications. Thus, 'species B' (Balaikaite *et al* nomenclature) found in $prFMN^{ox}$ is in fact a mixture of multiple compounds. Even if we consider a pure FMN sample, at least two species (Q and R) are produced in the UbiX-catalyzed prenylation and subsequent oxidation.

To confirm that UbiX prenylates riboflavin, 200 µM reduced riboflavin was reacted with 1 mM DMAP in the presence of UbiX. Subsequent HPLC analysis under micro-aerobic conditions revealed a new peak at ~22.6 min whose UV-Vis and mass spectra are identical to peak S (hereafter referred to as prRiboflavin) (Figure 4.1A). Interestingly, a species analogous to peak Q (expected m/z = 445.21 units) is not observed. Therefore, UbiX-catalyzed prenylation of riboflavin affords a simpler reaction than that of FMN, which may be beneficial for future analysis.

When the chromatograph of $prFMN^{ox}$ is compared to that of reconstituted PhdA, we observe that while peaks Q and R are found in both samples (Figure 4.1B), the peak corresponding to the active form of prFMN (peak P) is not observed in $prFMN^{ox}$. This is in line with the previous reports that UbiX alone cannot synthesize $prFMN^{iminium}$ and incubation with a UbiD-like enzyme prior to oxidation is necessary.[4, 11, 32, 64] Moreover, peaks Q and R, which arise from the enzyme-free oxidation of $prFMNH_2$ under micro-aerobic conditions, can also bind PhdA during *in vitro*



G

Reactant peak	m/z	Identity	Product peak	m/z	Identity
W	457.11	FMN isomer	—	—	—
X	457.11	FMN	Q	525.17	Oxidized prFMN
			R	527.19	Reduced prFMN
Y	377.14	Riboflavin	S	447.22	Reduced prRiboflavin
Z	359.13	Dehydrated Riboflavin	T	429.21	Dehydrated prRiboflavin

Figure 4.1: Different species observed in UbiX catalyzed reactions. **A:** Prenylation of FMN (prFMN^{ox}) and riboflavin (prRiboflavin^{ox}), with their corresponding no UbiX controls. Peaks W, X (both 457.11 m/z), Y (377.14 m/z) and Z (359.13 m/z) are observed in the commercial FMN sample. Peaks Q (525.17 m/z), R (527.19 m/z), S (447.22 m/z) and T (429.21 m/z) are the prenylated products formed in the reaction. **B:** Comparing prFMN^{ox} peaks to reconstituted PhdA and holo-FDC. **C, D, E and F:** UV-Vis spectra for peaks Q, R, S and U respectively. **G:** Table summarizing the data displayed in A. All reactant and their corresponding product peaks are mentioned along with their proposed identities.

reconstitution. If present in excess, they might eventually out-compete the binding of prFMN^{iminium} to PhdA, leading to loss of activity and thus, demonstrating the need for desalting.

4.3.2 Prolonged oxidation of prFMN^{ox} leads to loss of prenylation

To observe the evolution of peaks Q and R with time, free prFMNH_2 was transferred to HPLC vials and consecutive injections were performed under micro-aerobic conditions. Surprisingly, subsequent analyses showed the occurrence of a peak at 17.5 min, which corresponds to FMN (peak X). In fact, all the species consumed in the reaction (peaks X, Y and Z) were slowly being re-generated (Figure 4.2A), presumably by hydrolysis of the prenyl moiety. To confirm this observation, prFMNH_2 was oxidized overnight under micro-aerobic conditions and analyzed by HPLC (Figure 4.2B). Indeed, the chromatograph shows the re-occurrence of peaks X, Y and Z as evident from their retention times and UV-Vis spectra (Figure 4.2C, D and E). Interestingly, when the chromatograph is viewed at 280nm, a new species (peak U) is detected at ~20.2 min, whose UV-Vis spectrum lacks any features in the 300 – 600 nm region (Figure 4.2F). To verify that FMN and peak U are formed due to oxidative degradation, prFMNH_2 was stored overnight under either anaerobic or aerobic conditions, before analysis by HPLC (Figure 4.2B). The corresponding chromatographs show no degradation in the anaerobically stored prFMNH_2.

To further characterize peak U, prFMN^{ox} was analyzed by LC-MS (Figure 4.2G). While peaks Q and R appear immediately following exposure to O_2, further oxidation leads to the emergence of FMN and peak U, whose m/z (559.18 units) matches that of the putative prFMN-C4a-OOH reported previously.[27] Based on the current analysis, it is unclear which of the two peaks, Q or R, degrade to FMN.

Since UbiX catalyzed prenylation of riboflavin affords a single product, we studied the oxidation of prRiboflavin in detail. Peak S is observed immediately upon exposure to O_2 whereas prolonged oxidation leads to the loss of peak S and concomitant formation of riboflavin (Figure 4.2H). On the other hand, a species resembling peak U without the PO_4^{3-} group (expected m/z =

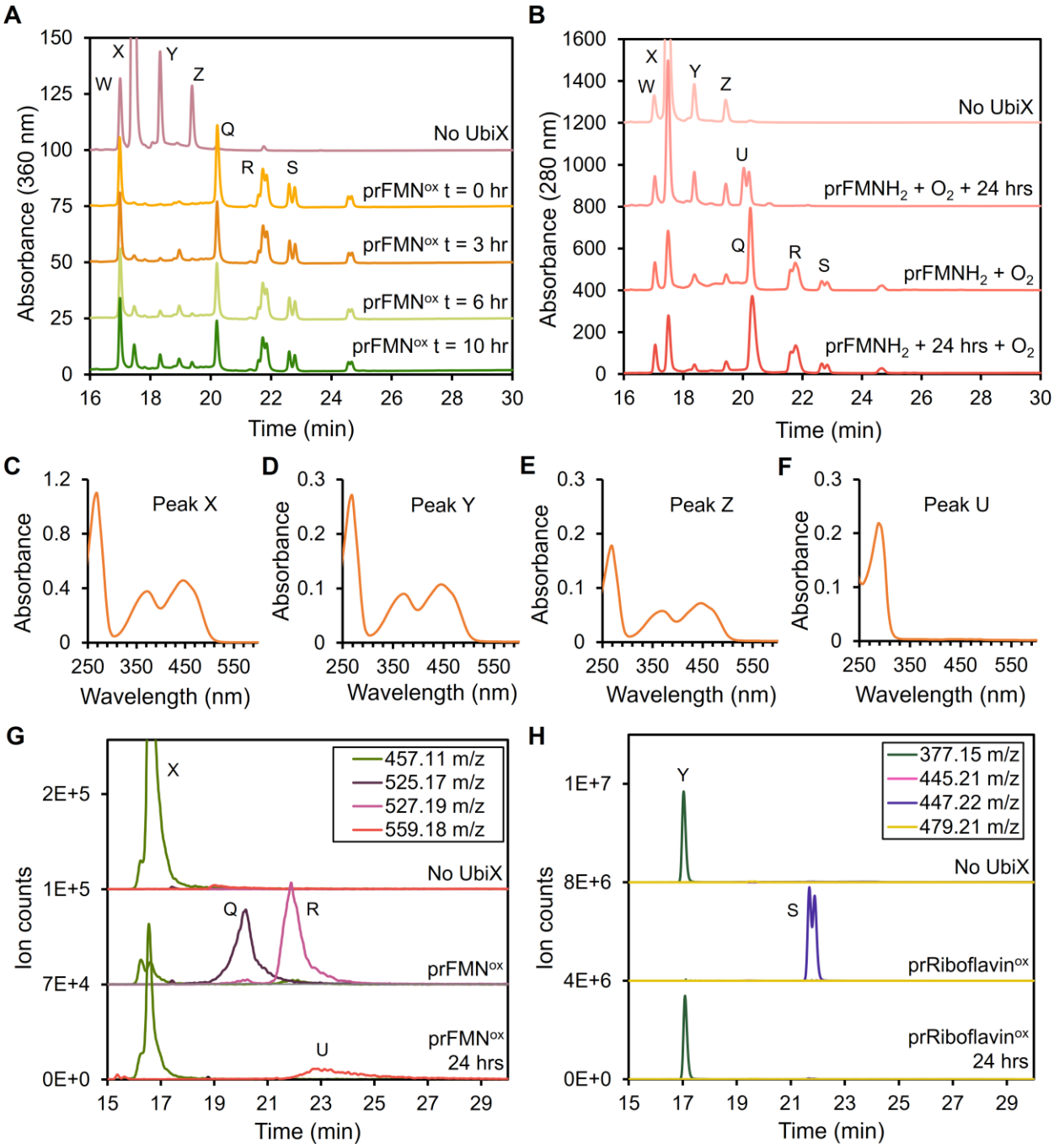


Figure 4.2: Prolonged oxidation of prFMN^{ox} and prRiboflavin^{ox}. **A:** Time-dependent analysis of prFMNH$_2$ oxidation under micro-aerobic conditions. **B:** Chromatographs showing no UbiX control (No UbiX), free prFMNH$_2$ injected under micro-aerobic conditions (prFMNH$_2$ + O$_2$), prFMNH$_2$ oxidized for 24 hrs (prFMNH$_2$ + O$_2$ + 24 hrs) and prFMNH$_2$ stored anaerobically for 24 hrs before HPLC analysis (prFMNH$_2$ + 24 hrs + O$_2$). **C, D, E and F:** UV-Vis spectra for peaks X, Y, Z and U respectively. **G:** LC-MS analysis of prFMN^{ox} immediately following O$_2$ exposure and after overnight oxidation. **H:** LC-MS analysis of prRiboflavin^{ox} immediately following O$_2$ exposure and after overnight oxidation. The retention times for the same species vary slightly between the LC-MS and HPLC chromatographs owing to differences in mobile phase (0.1% formic acid vs 10 mM trifluoroacetic acid) and LC systems (Agilent 1290 vs Shimadzu LC-20).

479.21 units) is not observed. By analogy to prFMN^{ox}, this suggests that peak R likely degrades to FMN and peak Q converts to peak U. We note that the retention times for all peaks are slightly shifted in the LC-MS chromatographs, owing to the different elution buffers (0.1% formic acid vs 10 mM trifluoroacetic acid) and LC systems (Agilent Series 1290 vs Shimadzu LC-20).

4.3.3 Solvent deuterium labelling

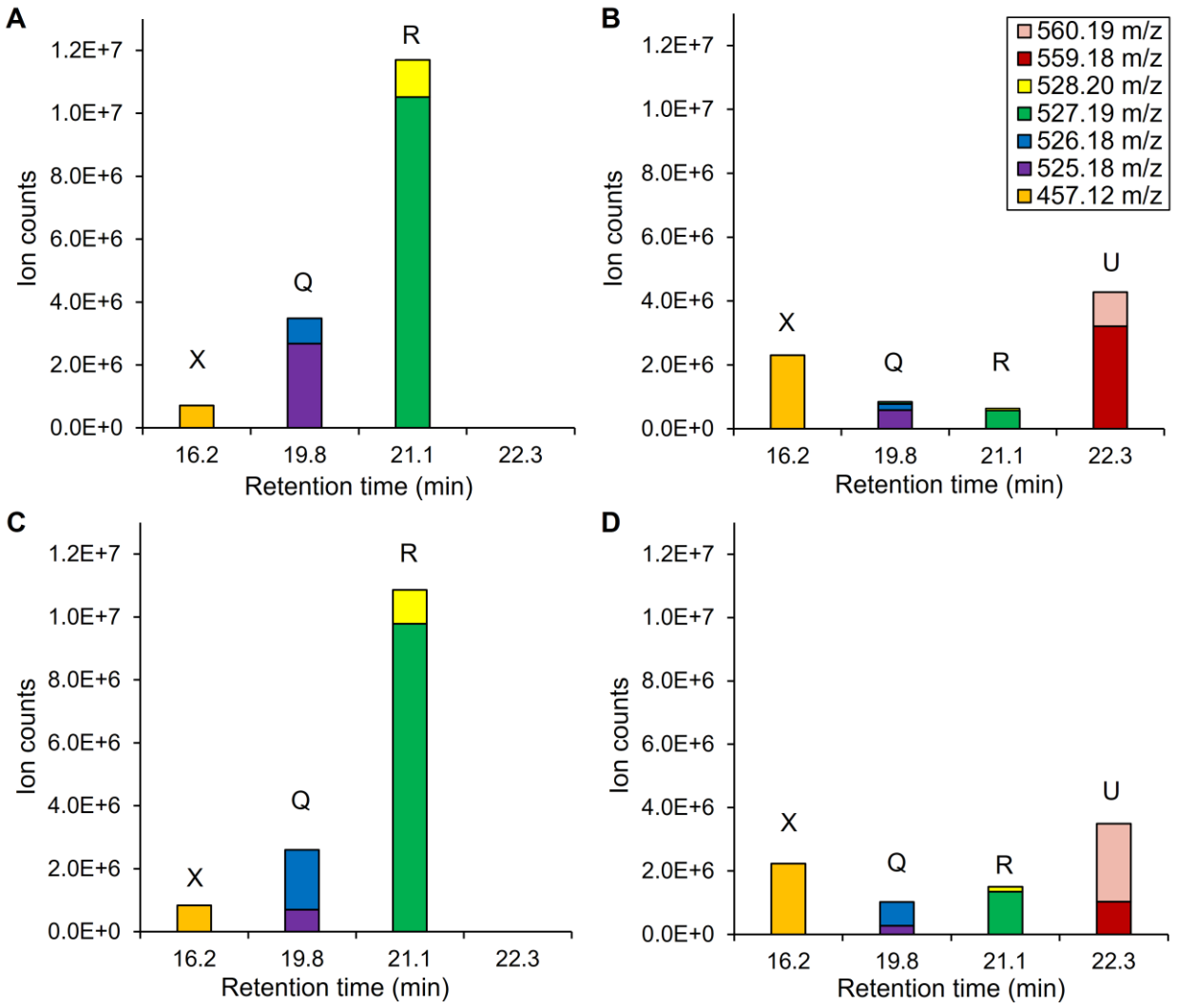


Figure 4.3: Solvent D-atom labelling of prFMN. Ion counts of all m/z signals detected for each peak in the chromatograph are displayed for prFMN^{ox} synthesized in H_2O (**A**) and after overnight oxidation (**B**) as well as for prFMN^{ox} synthesized in D_2O (**C**) and oxidized overnight (**D**). Each peak is labelled as mentioned in the text.

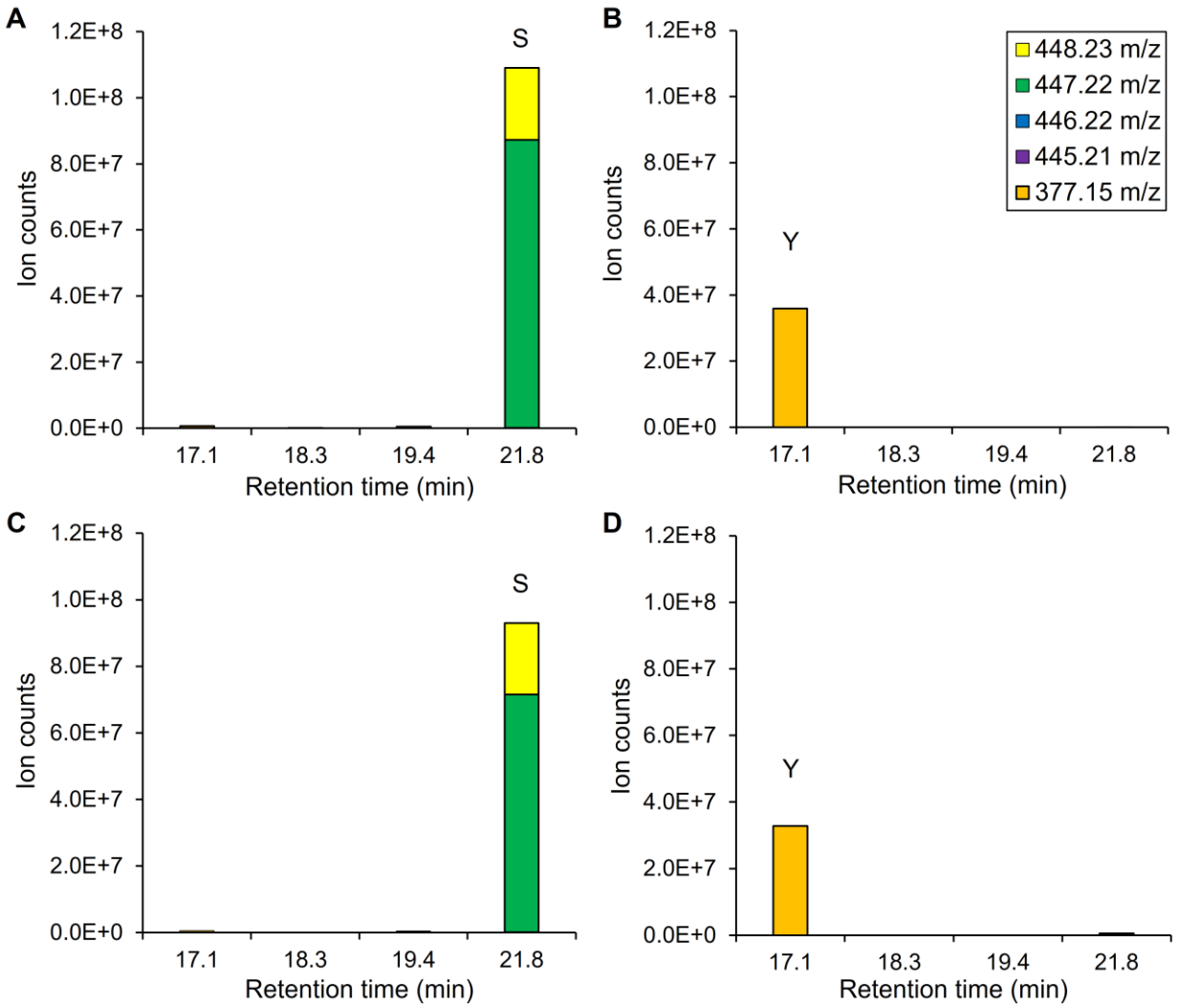


Figure 4.4: Solvent D-atom labelling of prRiboflavin. Ion counts of all m/z signals detected for each peak in the chromatograph are displayed for prRiboflavin^{ox} synthesized in H_2O (**A**) and after overnight oxidation (**B**) as well as for prRiboflavin^{ox} synthesized in D_2O (**C**) and oxidized overnight (**D**). Each peak is labelled as mentioned in the text.

Loss of prenylation can occur if the reaction stalls after formation of the N5-C1' bond, preventing cyclization. Oxidation and subsequent imine-hydrolysis will then form the corresponding flavin starting material and prenal (Figure 4.7). The proposed mechanism for UbiX suggests that the prenyl moiety of N5-C1' accepts a PO_4^{3-}-mediated solvent proton, prior to cyclization. Therefore, we considered that studying prFMN synthesis in buffered D_2O might be informative. We followed the D-atom label in prFMN^{ox} by LC-MS and observed that for the reaction performed in ~80% D_2O, peak Q displayed an enhancement of the [M+1] signal (526.18

m/z), whereas peak R was unchanged (Figure 4.3). This confirms that peak Q represents a cyclized form of prFMN that has incorporated solvent deuterium, while peak R is the N5-C1' adduct and hence, doesn't incorporate deuterium. Moreover, peak U also shows a higher proportion of the [M+1] signal (560.19 m/z), indicating solvent D-atom incorporation.

Upon studying the prenylation of riboflavin in D_2O, we observe that peak S doesn't incorporate any deuterium and converts to riboflavin over time (Figure 4.4). Overall, our analysis confirms that peak R (and the analogous peak S) is an N5-C1' adduct that loses the prenyl group upon oxidation and subsequent hydrolysis whereas peak Q is a cyclized form of prFMN that eventually degrades to peak U.

4.3.4 The role of different prFMN forms in oxidative maturation

To understand the role of peaks Q and R in oxidative maturation, PhdA was reconstituted *in vitro* under three different conditions. Condition 1 ($-O_2$, $-O_2$, $+O_2$) represents the standard reconstitution protocol described in section 2.2.4, where free $prFMNH_2$ was incubated with PhdA in the absence of O_2, followed by an anaerobic desalting and subsequent oxidation. For condition 2 ($-O_2$, $+O_2$, $+O_2$), $prFMNH_2$ was incubated with PhdA under anaerobic conditions, oxidized and then desalted under aerobic conditions. For condition 3 ($+O_2$, $+O_2$, $+O_2$), $prFMNH_2$ was oxidized, incubated with PhdA and subsequently desalted, all in the presence of O_2. Under each condition, the different PhdA fractions were analyzed for prFMN content (Figure 4.5A) and their decarboxylation activity was measured with PCA (Figure 4.5B).

Our data shows distinct variation in the proportion of bound cofactors, including the active $prFMN^{iminium}$ (peak P), under each condition. The worst reconstitution is for condition 3 ($+O_2$, $+O_2$, $+O_2$), where negligible quantities of peak P are seen. This suggests that once $prFMNH_2$ has oxidized to peak Q, PhdA cannot convert it to peak P. This is in line with the previous report,

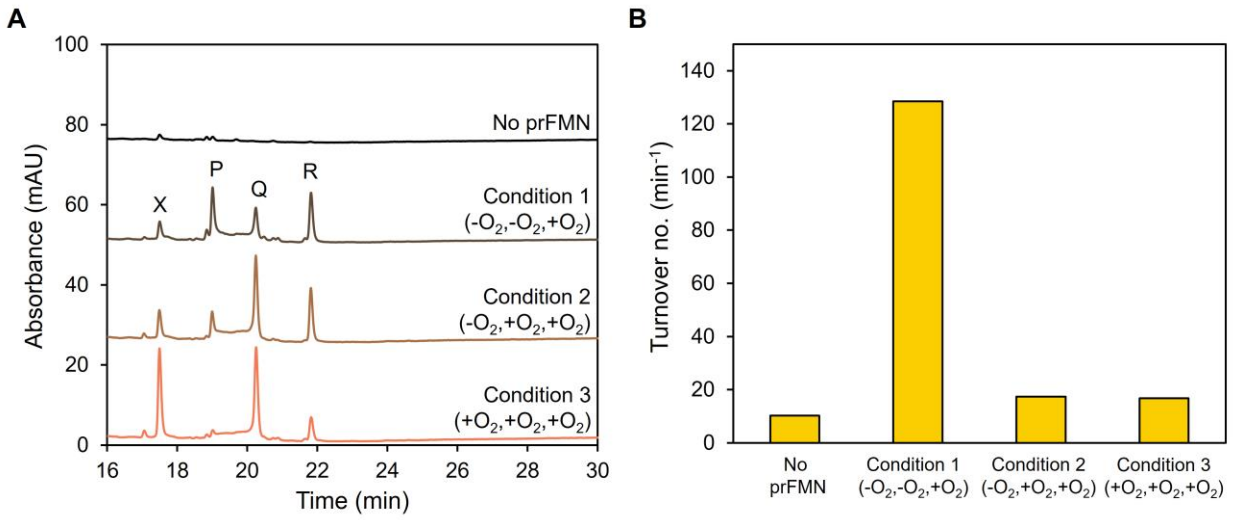


Figure 4.5: Reconstitution of PhdA under different conditions. **A:** HPLC chromatographs of reconstituted PhdA fractions showing different prFMN forms. **B:** Normalized reaction rate (min^{-1}) measured for PCA decarboxylation by different PhdA fractions.

where $prFMN^{ox}$ couldn't activate apo-FDC.[27] Moreover, it is also evident that for this condition, peak R has converted to FMN (peak X). While some amount of peak P is observed for condition 2 ($-O_2$, $+O_2$, $+O_2$), the proportion of peak Q is significantly higher. Since ~5 equivalents (w.r.t. PhdA) of free $prFMNH_2$ are added for *in vitro* reconstitution, this suggests that PhdA binds a fraction of $prFMNH_2$, converting it to the active form, but the unbound cofactor oxidizes to peak Q and eventually out-competes binding of peak P to PhdA. Finally, the highest proportion of active prFMN (peak P) is observed only for the standard reconstitution (condition 1, $-O_2$, $-O_2$, $-O_2$). The corresponding enzymatic activities are in agreement, with the highest rate observed for condition 1 and all other PhdA fractions being similar to the un-reconstituted enzyme.

4.4 Discussion

Previous studies on the oxidation of $prFMNH_2$ in free solution detected the formation of a $prFMN^{radical}$ 'species B' which slowly degrades in an oxygen dependent manner to a prFMN-C4a-OOH 'species C' that appears to be remarkably stable (>5 days). These species are nearly identical to the proposed intermediates in prFMN maturation. Thus, one would expect that adding $prFMN^{ox}$

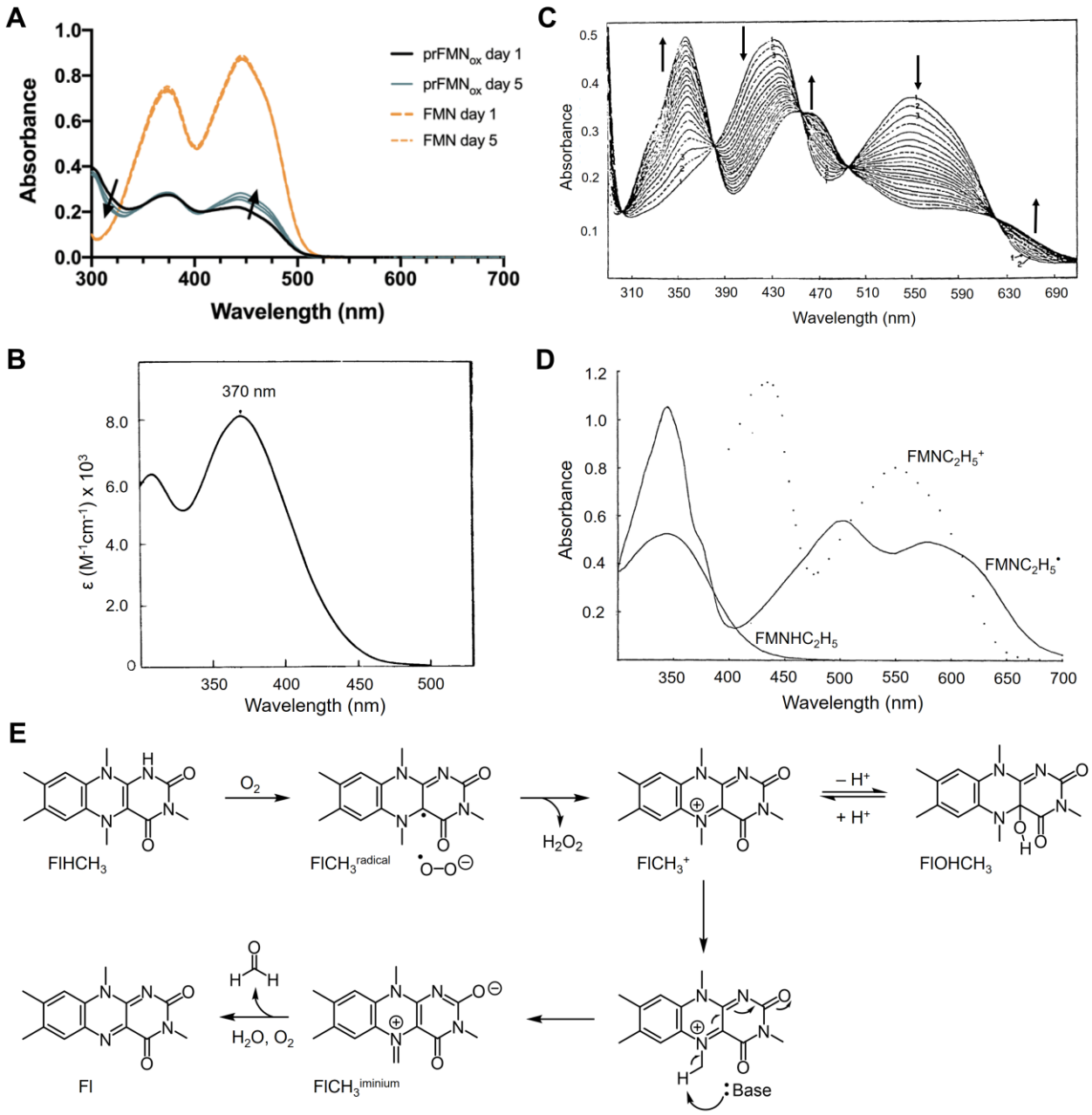


Figure 4.6: Oxidation and solvolysis of N5-alkyl flavins. **A:** UV-Vis spectrum of prFMN^{ox} containing the putative prFMN-C4a-OOH species (reproduced from ref.27) **B:** UV-Vis spectrum for 4a-hydroperoxy-5-ethyl-3-methyllumiflavine (reproduced from ref.107). **C:** Changes in the UV-Vis spectrum of $FlCH_3^+$ during solvolysis. Arrows indicate spectral changes with time (reproduced from ref.111). **D:** UV-Vis spectra of reduced ($FMNHC_2H_5$), radical semiquinone ($FMNC_2H_5^•$) and oxidized ($FMNC_2H_5^+$) N5-ethyl-FMN (reproduced from ref.111). **E:** Scheme depicting the solvolysis of $FlCH_3$.

to apo-FDC should provide active enzyme. However, the observed data contradicts this prediction.[27] In order to explain the discrepancy, two alternatives were provided – either the

transient formation of a prFMN-C4a-OO^- is necessary for proper prFMN maturation, or cofactor maturation doesn't proceed through peroxy-intermediates.

Here, we propose a third possibility – the assignment of 'species C' as prFMN-C4a-OOH is incorrect. The evidence that led to this assignment is as follows: (i) UV-Vis spectrum akin to flavins (Figure 4.6A), (ii) a 559.18 m/z peak which could correspond to the [M+H] of a prFMN-C4a-OOH and (iii) ferrous-oxidized xylenol−orange (FOX) assay as well as surface enhanced Raman spectroscopy (SERS), both of which detected peroxide species.

Our data clearly shows that prolonged oxidation of $prFMNH_2$ leads to the re-formation of FMN, which explains the flavin-like UV-Vis spectrum reported by Balaikate *et al*. Additionally, the spectrum for the 559.18 m/z species (peak U, Figure 4.2F) looks quite unlike the previously reported UV-Vis spectra for N5-alkyl-C4a-hydroperoxy flavins (Figure 4.6B).[107] Moreover, while N5-alkyl-C4a-OOH flavins can be formed by reacting N5-alkyl flavins with H_2O_2 and are stable in dioxane for days, they are highly labile in H_2O and collapse to N5-alkyl-C4a-OH adducts.[107] Finally, oxidation of flavins with O_2 is known to generate H_2O_2[108] which invalidates the results obtained by FOX and SERS, especially since the authors didn't show any control experiments with an oxidized $FMNH_2$ sample.

The solvolysis of N5-alkyl flavins, such as of 3,5-dimethyllumiflavin ($FlCH_3$) has been studied extensively[107, 109, 110] and appears to be kinetically complex (Figure 4.6C). When reduced $FlCH_3$ ($FlHCH_3$) is mixed with O_2, an $FlCH_3$-superoxide radical pair is formed which collapses to oxidized $FlCH_3^+$. $FlCH_3^+$ exists in equilibrium with the C4a-OH species $FlOHCH_3$ as a function of pH. A general-base catalyzed deprotonation converts $FlCH_3^+$ to the corresponding imine, $FlCH_3^{iminium}$ which undergoes rapid hydrolysis to form reduced 3-methyllumiflavin (FlH_2) and formaldehyde. Under anaerobic conditions, FIH_2 comproportionates with $FlCH_3^+$ to give further

complex products and under aerobic conditions, all $FlCH_3^+$ eventually converts to oxidized 3-methyllumiflavin (Fl) and formaldehyde (Figure 4.6E). The deprotonation of $FlCH_3^+$ to form $FlCH_3^{iminium}$ is the rate limiting step in the solvolysis. This step also appears to be irreversible as H/D exchange of the methyl protons with solvent was not observed. Furthermore, the observed bimolecular rate constant for proton abstraction in H_2O was 6.4 x 10^{-6} $M^{-1}s^{-1}$, with a Bronsted β value of 0.58.[109] This suggests that in bulk H_2O, proton abstraction occurs slowly and the presence of a strong base is required for efficient deprotonation.

Studies on solvolysis of $FlCH_3^+$ could act as a model to understand prFMN oxidation (Figure 4.7). Our data indicates that UbiX catalyzed prenylation of FMN is inefficient and forms N5-prenyl-$FMNH_2$ in addition to $prFMNH_2$. The m/z of peak R (527.19 units) matches the [M+H] mass of N5-prenyl-$FMNH_2$ and its UV-Vis spectrum (λ_{max} ~ 365 nm) resembles that of N5-ethyl-FMN (λ_{max} ~ 345 nm, ~20 nm shift) studied previously[111] (Figure 4.6D). Analogous to $FlHCH_3$,[110] $prFMNH_2$ and N5-prenyl-$FMNH_2$ can react with O_2 to form the corresponding radical pairs ($prFMN^{radical}$, N5-prenyl-$FMN^{radical}$) with superoxide and eventually collapse to their oxidized forms ($prFMN^+$ and N5-prenyl-FMN^+). The m/z of peak Q (525.17 units) is identical to the [M^+] mass of $prFMN^+$ and its UV-Vis spectrum (λ_{max} ~ 360 and 500 nm) looks similar to $FlCH_3^+$ (λ_{max} ~ 420 and 550 nm, ~50-60 nm shift), suggesting that peak Q could be $prFMN^+$. Further, $prFMN^+$ and N5-prenyl-FMN^+ can convert to $prFMN^{iminium}$ and N5-prenyl-$FMN^{iminium}$ through buffer catalyzed deprotonation. However, these species wouldn't be stable in free solution, N5-prenyl-$FMN^{iminium}$ would hydrolyze to yield FMN and prenal. The solvolysis of $prFMN^+$ is likely more complicated and we propose the following alternative structure for species U (Figure 4.7). This species has the expected [M+H] of 559.18 m/z but it represents a degraded form of prFMN that cannot activate PhdA.

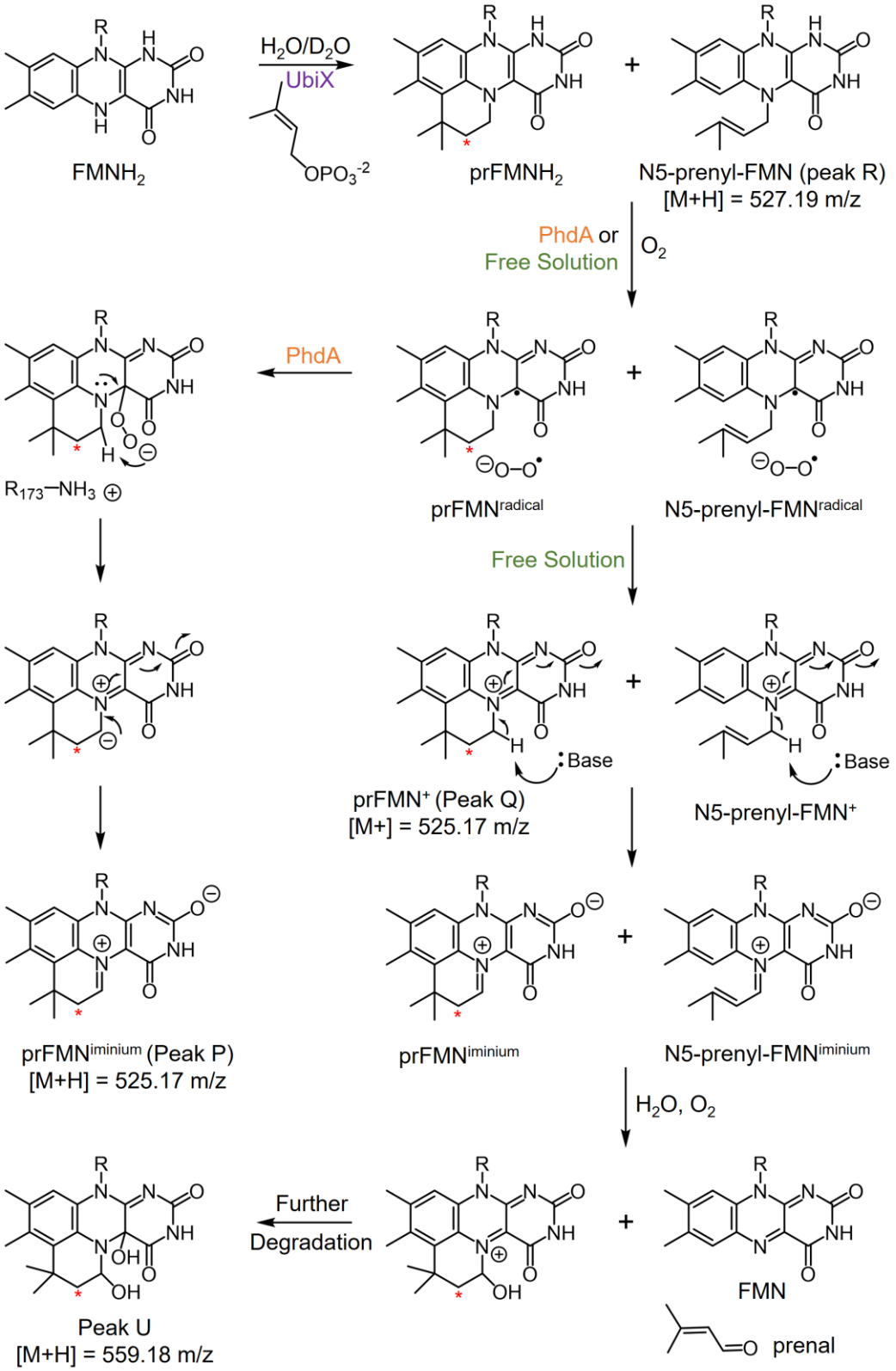


Figure 4.7: Proposed pathway for the oxidative maturation of $prFMNH_2$ in the presence of UbiD-like enzymes or its solvolysis in free solution under aerobic conditions. The position where a solvent D-atom could be incorporated during cyclization of the N5-C1' adduct is marked with (*)

Interestingly, even though prFMN^{iminium} can be produced through a general base mediated deprotonation of prFMN^{+} (Figure 4.7), our data suggests that peak Q is not converted to peak P when incubated with PhdA (Figure 4.5A). It is possible that the PhdA active site doesn't provide a sufficiently basic residue in close proximity to prFMN that can bring about this conversion. It would thus be interesting to investigate if nature has evolved maturases to efficiently convert prFMN^{+} to prFMN^{iminium} and protect the cofactor from subsequent solvolysis.

As discussed previously, an important limitation of our study is the inability to detect precise oxidation states of the different prFMN forms. Solvolysis of N5-alkyl flavins involves numerous redox states, each with distinct protonated/deprotonated forms that comproportionate together.[109] LC-MS analysis cannot detect all these species and only the prominent redox forms that exist at steady-state can be identified. However, our data provides an initial blueprint that can direct future studies on understanding prFMN oxidative maturation.

Chapter 5 Conclusions and Future Directions

5.1 PhdA is a prFMN dependent (hetero)aromatic acid decarboxylase

Recently, UbiD-like enzymes have garnered significant interest in the field of biocatalysis owing to their ability to perform reversible decarboxylations on a variety of industrially and environmentally relevant substrates. While only about 10 enzymes have been characterized so far, they (de)carboxylate a range of α,β-unsaturated acids, (hetero)aromatic acids as well as phenolic acids.[41] Characterizing previously unknown enzymes expands the available chemical space for biocatalysis.

Moreover, the peculiar cofactor, prFMN, is central to UbiD-based catalysis and chemically interesting. The reactivity of prFMN hinges on its nitrogen ylide character, which has been shown to perform both 1,3-dipolar cycloaddition as well as electrophilic addition reactions. Studying novel enzymes from the UbiD-family might uncover other mechanisms, leading to an overall better understanding of prFMN-based catalysis.

Therefore, we investigated the reactivity and mechanism of the recently discovered prFMN dependent phenazine-1-carboxylic acid decarboxylase (PhdA) from *M. fortuitum*. Phenazines are redox active secondary metabolites produced by pseudomonads that act as antibiotics and are also important in signaling.[36] Phenazine-1-carboxylic acid (PCA) is the precursor to most phenazines. Therefore, PhdA catalyzed decarboxylation of phenazine which acts as the first step in PCA degradation, provides a competitive advantage to *M. fortuitum* against pseudomonads in soil.[36] Chemically too, PhdA is interesting not only because it decarboxylates electron deficient

(hetero)aromatic rings, but also because initial studies revealed the possibility of a novel mechanism involving radical intermediates.

5.1.1 PhdA decarboxylates oxidized PCA

Studies by Costa *et al.* suggested that PhdA reconstituted at pH 7.2 lost activity within a few hours.[36] However, the enzyme stored at pH 9.2 was active for several days when pre-incubated with a one electron reducing agent such as sodium dithionite or a radical mediator such as paraquat. This suggested the existence of a novel decarboxylation mechanism, perhaps involving radical intermediates. Moreover, since phenazines can exist in oxidized, reduced and radical semiquinone forms, the redox state of the active substrate form couldn't be determined.

Our initial efforts to optimize the expression and purification of holo-PhdA from *E. coli* failed and *in vitro* reconstitution was deemed necessary. Here, we noticed that while the enzyme reconstituted at pH 7.2 lost activity as previously reported, the addition of a desalting step prior to oxidation led to a stable enzyme that remained active for several weeks. This suggests that the presence of free, incorrectly oxidized prFMN ($prFMN^{ox}$) is detrimental to proper maturation of the bound cofactor. Similarly, PhdA reconstituted by our method was active irrespective of the presence of reducing agents, indicating that their dependence is an artifact of the storage buffer's pH. Moreover, we showed that only oxidized PCA is efficiently utilized by PhdA dissipating the previous ambiguity regarding the correct redox state of the substrate.

5.1.2 PhdA activates a variety of (hetero)aromatic substrates

Investigating the substrate scope of PhdA revealed that while the third ring of phenazine helps in proper orientation of substrate in the active site, it can be replaced with methyl groups. Similarly, the presence of the ring nitrogens is not required for activity as acridine and even

anthracene carboxylic acids are substrates. Additionally, PhdA decarboxylation is regioselective and prefers the carboxylic acid group on C1 over C2 or C9. Apart from providing valuable insights into the mechanism of PhdA, these studies also predict that several other (hetero)aromatic carboxylic acids may be potential substrates of PhdA. This includes regioisomers of acridine carboxylic acids, methylquinoline carboxylic acids and methylnaphthalene carboxylic acids amongst several others.

Furthermore, we showed that the PhdA-catalyzed reaction is partially reversible by demonstrating the incorporation of solvent D-atoms into phenazine in a regioselective manner. Kinetic studies reveal that proton abstraction from phenazine is likely rate-limiting in the reverse direction. PhdA also functionalizes C–H bonds in acridine and DQ, albeit at much slower rates that can only be detected using a sensitive mass spectrometer.

The substrate scope and H/D exchange studies indicate that the PhdA-catalyzed decarboxylation proceeds through 1,3-dipolar cycloaddition, akin to FDC. While there is no definitive evidence supporting this proposal, acridines and anthracenes are known to form 1,3-dipolar cycloadducts with strong dipoles such as nitrile oxides and nitrogen ylides. Moreover, the substrates of PhdA are electron deficient, making an electrophilic mechanism less likely.

5.1.3 High-throughput assays to study substrate scope of UbiD-like enzymes

An important hurdle to elucidating the substrate scope of UbiD-like enzymes is the difficulty in assembling a sufficiently diverse substrate library. Commercially available carboxylic acids are relatively scarce and expensive. On the other hand, it is easy to curate a library of the corresponding hydrocarbons for studying H/D exchange. However, our experiments revealed that H/D exchange reactions are significantly slower than decarboxylation, leading to difficulties in detecting products.

We showed that an Orbitrap mass analyser that can identify even minute D-atom incorporation because it has a high enough resolution to distinguish between ^{13}C and D isotopes. Moreover, since each hydrocarbon has a unique m/z signature, multiple samples can be combined for injection, increasing throughput. Therefore, optimizing an Orbitrap mass analyzer based H/D exchange assay may be a viable alternative for studying the substrate scope of UbiD-like enzymes.

5.1.4 Strategies to optimize carboxylation

Owing to the large thermodynamic barrier associated with CO_2 addition, we were unable to optimize conditions for PhdA catalyzed carboxylation of phenazine. Our efforts to couple carboxylation to the thermodynamically favorable reduction catalyzed by carboxylic acid reductase (CAR) were also futile, due to the incompatible substrate scope of the two enzymes. However, we do see promise in this approach.

Several CAR enzymes have been characterized so far with diverse substrates[54] and future studies can explore the reactivity of different CAR homologs for the various substrates of PhdA. Furthermore, directed evolution and protein engineering can help improve the substrate scope of both enzymes leading to better compatibility. In section 1.2.2 we identified several other strategies for product removal to improve carboxylation yield. Their efficacy in enhancing PhdA catalyzed carboxylation could also be evaluated.

5.2 Protein conformational changes affect catalysis

While the mechanism of FDC has been studied extensively, thorough investigations into other UbiD-like enzymes are lacking. This is largely due to their recalcitrant nature and inability to co-crystallize with substrate and/or prFMN.[41] Fortunately, these enzymes display a diverse

substrate scope, allowing for a detailed examination of structure-activity relationships and steady state kinetics.

In search of alternate methods to study UbiD-like enzymes, we turned to measuring the effects of solvent isotope and viscosity on the steady state parameters of the PhdA-catalyzed decarboxylation of PCA and DQCA. Coupled to MD simulations, our results not only exposed a surprising relation between reaction rate and protein conformation, but also provided a blueprint to study similar effects in other enzymes.

5.2.1 Solvent isotope effects alter protein conformational mobility

Initially, we employed solvent isotope effects to identify rate-limiting chemical steps in the mechanism. However, we soon discovered that multiple steps were isotopically sensitive and a large *inverse* medium effect was apparent, indicating protein conformational changes. For some time, D_2O has been known to affect protein conformations,[93-96] and we showed that for PhdA, these conformational changes led to rate enhancement under V_{max}/K_M conditions by favoring the formation of the Michaelis complex and subsequent steps.

To identify the nature of these conformational changes, we collaborated with Soumil Joshi and Dr. Sanket Deshmukh from Virginia Tech to perform MD simulations on the PhdA crystal structure (PDB:7PDA). It was observed that D_2O promotes domain closure akin to the catalytically active 'closed' conformer observed in several UbiD-like crystal structures. Moreover, D_2O affords stability to the active site loop and promotes an overall more compact structure to the protein. These results rationalize the *inverse* solvent isotope effect by showing how D_2O might stabilize the Michaelis complex and subsequent steps.

In conclusion, our studies showed that protein conformational changes are significantly more rate limiting than chemical steps such as proton transfer in the mechanism of PhdA. DQCA,

being a poorer substrate, has a further enhancement of the solvent isotope effects than for the native substrate, PCA.

5.2.2 The importance of protein conformations in the larger UbiD-like family

In contrast to the variety observed in their substrate scope, the UbiD-family of enzymes share several structural features. The tertiary structure of all enzymes studied so far consists of an N-terminal prFMN binding domain, a central α-helix, an oligomerization domain and a C-terminal α-helix. Distinct 'open' and 'closed' conformers are observed in several crystal structures.[41] These conformers display significant changes in the active site. Moreover, they appear to play an important role in catalysis as binding simulations reveal that several reaction intermediates are stabilized on the trajectory of domain closure. Despite this, many UbiD-like enzymes can only be crystallized in the 'open' conformer.

Therefore, our results with PhdA have the potential to impact broader understanding of the UbiD-family as a whole. Indeed, the existence of an 'open' conformer in crystal structures is associated with poor cofactor binding, lack of substrate co-crystallization and an overall recalcitrant nature of many UbiD-like enzymes. It therefore, would be interesting to see if some of these enzymes display similar *inverse* solvent isotope effects, owing to the D_2O afforded stabilization of the 'closed' conformer. It would also be worth investigating if switching to D_2O based buffers would provide a higher success in obtaining co-crystals with substrate/cofactor.

Moving forward, further computational modeling can guide rational design approaches for stabilizing the 'closed' conformer in H_2O. In Appendix B, we have identified several unique intramolecular H-bonds observed in the D_2O conformer for PhdA. A detailed investigation into how these bonds stabilize the 'closed' conformer will be informative. Based on our studies, we predict that the 'closed' conformer would not only improve the V_{max}/K_M but also the V_{max} rate,

since many late intermediates (such as **E.M** or **Int3**) will also be stabilized on the trajectory of domain closure. This might provide a tremendous push in utilizing UbiD-like enzymes for biocatalysis.

5.2.3 Intrinsic isotope effect for proton transfer

The two mechanisms proposed for UbiD-like enzymes have very similar reaction intermediates. In fact, the m/z of these putative intermediates would be identical, making it difficult to design mechanistic probes to differentiate between the two. Therefore, mechanistic proposals for several UbiD-like enzymes rely on substrate scope and/or computational simulations.

Here we propose a method to differentiate between the two mechanisms using solvent isotope effects. The proton transfer step occurs between **I2/Int2** and **I3/Int3** (refer to Figure 1.8 and 1.9). While the 'reactant' for this step is identical, the 'product' structure differs slightly for both mechanisms. Therefore, depending on whether the transition state is an 'early' or a 'late' transition state, the intrinsic isotope effect on proton transfer could be mechanistically informative. We propose that measuring the V/K_P isotope effect in proteated, deuterated and tritiated water followed by using the Swain-Schaad relationship can provide us a value of the intrinsic isotope effect for proton transfer.[112] This can then be compared to theoretical values calculated from computational simulations for both mechanisms. In principle, such an analysis can be performed for all UbiD-like enzymes.

5.3 Biosynthesis and maturation of prFMN

Despite being central to the catalysis of UbiD-like enzymes, the biosynthesis and maturation of prFMN is poorly understood. Many inactive forms of the cofactor are observed as intermediates, byproducts and side products of the maturation process. A prFMN radical

semiquinone (prFMN^{radical}) and a prFMN-C4a-OOH are proposed intermediates in the maturation process. Interestingly, the freely oxidized prFMN (prFMN^{ox}) is also proposed to contain prFMN-C4a-OOH, but adding this to apo-FDC doesn't lead to enzyme activation.

Hence, we studied the oxidation of free prFMN in detail to identify the different species present. We also studied how prFMN^{ox} inhibits oxidative maturation of PhdA. Based on the results, we proposed an alternate scheme for the maturation of prFMNH_2 or its solvolysis following oxidation.

5.3.1 Understanding the solvolysis of prFMN

Careful analysis of the UbiX-catalyzed prenylation of FMNH_2 revealed that several other flavin impurities present in the FMN sample, such as riboflavin, were also prenylated. Two prominent peaks were observed in the LC-MS chromatograph of prFMN^{ox}, one of them (peak Q) had the same mass as prFMN^{iminium}, but a different retention time, while the other (peak R) was 2 m/z higher. Solvent D-atom labelling showed that peak Q has incorporated a D atom whereas peak R hadn't. Furthermore, upon prolonged oxidation, FMN was re-formed, along with a new species (peak T) that had an m/z of 559.18 units and incorporated solvent D.

Similar experiments on oxidation of prenylated riboflavin (prRiboflavin^{ox}) showed that only one species predominanted (peak S). This peak (447.22 m/z) resembled peak R, but lacking a PO_4^{-3} group and upon prolonged oxidation, led to the re-formation of riboflavin. By drawing analogy to prFMN^{ox}, this suggests that peak Q converts to peak T whereas peak R converts to FMN.

Comparing these results to previous studies on other N5-alkyl flavins, such as 3,5-dimethyllumiflavin ($Fl_{ox}^{+}CH_3$), we hypothesized that peak Q is a cyclized form of prFMN that has the same m/z as prFMN^{iminium} (prFMN^{+}) whereas peak R is a N5-C1' intermediate that hasn't

cyclized (N5-prenyl-FMN). Further, we proposed that these species undergo a general base mediated deprotonation at C1', followed by solvolysis which leads to the formation of FMN and peak T.

We acknowledge that solvolysis of N5-alkyl flavins are kinetically complex and several experiments need to be performed to validate our hypothesis. Future studies could include performing prFMN synthesis with D-labeling at the C1' of DMAP. If our hypothesis is correct, peaks Q and R should retain the label whereas peak P ($prFMN^{iminium}$) would lose a D-atom. Upon prolonged oxidation, peak T should also lose the label. Further, LC-MS analysis could be expanded to detect formation of prenal and tandem mass spectrometry might provide unique fragmentation patters of peaks Q and R, providing a better understanding of their structure.

5.3.2 Is there a prFMN maturase?

While the transiently formed prFMN-C4a-OO^{-} during oxidative maturation might deprotonate C1' to form $prFMN^{iminium}$,[28] PhdA is unable to convert $prFMN^{+}$ to $prFMN^{iminium}$ when incubated with $prFMN^{ox}$. The deprotonation of N5-alkyl flavins to form the corresponding imines is rate limiting, irreversible and requires a strong base. Buffer catalyzed deprotonation is very slow and leads to immediate hydrolysis of the imine that subsequently forms.

Hence, we propose the existence of a putative prFMN maturase that may be able to perform this deprotonation sufficiently rapidly and protect $prFMN^{iminium}$ from hydrolysis. Earlier, Costa et al. identified a previously unknown gene in the '*phd* operon'.[36] This gene (labelled XA26_16660) is annotated as a 'hypothetical protein'. Recent experiments from our lab show that this small ~18kDa protein (called as PhdC) is capable of binding $prFMNH_2$ and matures it to the correct $prFMN^{iminium}$, apparently in a more efficient manner than PhdA. It also has a lower affinity for prFMN and incubation of 'holo-PhdC' with apo-PhdA leads to transfer of $prFMN^{iminium}$, thus

activating PhdA. Hence, we propose that PhdC is a prFMN maturase, though it remains to be seen whether it can convert $prFMN^{+}$ to $prFMN^{iminium}$ when incubated with $prFMN^{ox}$.

5.4 Conluding remarks

The UbiD-enzyme family and the peculiar prFMN cofactor were discovered only about eight years ago and while a lot remains to be unearthed, there is tremendous interest in utilizing these enzymes for biocatalysis. This thesis aims to address some of the important questions in prFMN biosynthesis and maturation in addition to characterizing a novel (hetero)aromatic decarboxylase from the UbiD-family. Here, I also develop alternate methods to study the mechanism of recalcitrant UbiD-like enzymes using solvent isotope effects and discover the importance of protein conformational changes in catalysis.

Overall, this thesis provides a better understanding of certain aspects of UbiD-chemistry and is expected to support future research and scholarship in the field.

Appendices

Appendix A: Nucleotide Sequences of All Proteins Used in This Study

A.1 pET20b(+)(AmpR) + *phdA* (between NdeI and NotI, vector sequence in lower case)

catATGCGGCATTACATCGACACTCTGACGGAGAAACTCGGAGCCGATGAGGTGCAG
ACGATCAAGGGCGCCAACTGGGATCTCGAAATCGGTTGCATCACCGAGTTGTCGGC
CGAGAAAGAGGGCCCGGCACTACTTTTCGACGACATCCCCGGCTACCCGTCGGGAC
ATCGCGTCTTCACCAACTTCATGGGCACCGTGTCGCGCTGCGCGGTCGCCCTCGGCC
TGCCTGCTGACACCTCCGCGATGGACATCATCCGCGCTTGGAAGGATCTTGGAAAGC
GCATCGAACCCATTCCCCCGGTTGAGGTTTCCGAGGGCGCGATCCTGGAAAACGTGC
TCGAGGGCGATGATGTCGACCTGGAGATGTTCCCGACGCCGCGGTGGCATGACGGC
GACGGCGGGCGCTACATCGGCACCGCGTGCATGGTCATCACACGTGATCCGGACAC
CGGTTGGGTCAACGTCGGCACCTACCGGGGATGCGTGCAGGGCAAGGACCGGCTGT
CGCTGTGGATGCTCGGGAACCGGCACGCGCTCGCGATCGCCAAGAAGTACTGGGAT
CGGGGCACGGCCTGCCCGATCGCGGTGGTTGTCGGTTGCGATCCGATTCTGACCACT
GCCGCCGCCATCGCCGCCCCATCGGGGGTGTGCGAGTACGACGTCGCGGGTGGTCT
ACGTGGCGTCGGCGTCGAGGTCATCTCCGCGCCCGGCACCGGGCTGCCGATTCCGGC
CAACGCCGAGATCGTCTTCGAGGGCGAGATGCCGCCGGTGGAAGAGGAGTCGGTGC
ACGAAGGCCCGTTCGGGGAGTGGACCGGCTACTTCACCCACGCCGGCGACGAGACC
GTCGTGCGCGTGCAGCGCATCCTGCATCGGGATTCGCCGATCATCCTCGGCGCGCCG
CCTATGATCCCCACCGTGCCCGCCGGCGACCAGGCGGTGCCGCTGTACTCGGCCTCG
GTCACCTGGGATCACCTGGAGGCCTCCGGTGTGCAGAACATCAAGGGGGTCTGGGC
CTACGCGCGTCAGCTCATGATGGTGATCTCGATCGAGCAGACGGGTGCAGGCGACG
CCATGCATGCGCTGCTCGCCGCCGCGGGCCGTAAGCGCACCGGAGGTGTGGATCGC
TATTTCGTGGTCGTGGATGAGGACATCGACATCACCGACATCAACCACGTGCTGTGG
GCGCTGTTCACCCGTGTCGATCCGGCCGAATCGATCCACGTGCTGCGGACGCCTACG
ACCGCGATCGACCCGCGCTTGTCGCCGGCCAAGCGGGAAGCCGGTGACATGTCGAT
GGGCATCGTGTTGATCGACGCGTGTAAGCCGTTCGCGTGGAAGGACTCCTACCCACG
GGCGAACCGGTTCGACGAGCCGTACCGGGCCGAGATCCGCGATCGGTGGAAGGCGA
CATTGCCGCTCGCCGAAAACCTGTACTTCCAGAGCCACCATCACCATCACCATTGAgc
ggccgc

A.2 pET20b(+)(AmpR) + *phdB* (between NdeI and NotI, vector sequence in lower case)

catATGCGCATCATCGTCGCGATCAGCGGCGCCAGCGGCGCACCGTTCGCGGTGCGCC
TACTGGAGACGCTACGCGAGATGCCCGATGTCGAAACCCATTTGGTGATGAGCACCT
GGGGTAAGTCGAACATCGAGGTGGAGACCGACCGCACGGTGTCCGAAGTTGTCGCC
TTGGCCGACGTGACGTACAAGCTCGGTGAGCAGGGCGCGGCGATCTCGTCCGGCTC
GTTCCGTACCACCGGGATGATCATCGTGCCGTGCAGCATGAGGACACTGTCCGCCAT
CCGCTATGGGATGGCGGACAATCTCATCTGCCGGGCCGCCGATGTGGTGCTCAAGG
AGGGGCGGCAGCTGGTGCTGGTGCCGCGGGAAACCCCGCTGAACACCATTCATTTG

GAGAACATGCTTGCGTTGAGCCGAATGGGCGCGCGCATCGTGCCACCGATGCCGGC
CTTCTACAACCATCCGCAGACCATCGGCGACATCGTCGACCATGTAGTGGTGCGCAT
TCTCGATCAGTTCGGGCTGGATGCCCCGCAGGCGAAACGGTGGCGCGGACTGGGCG
CGGCGCGGCGCGACCGGCCGACCCACGCCGGCCATGACCTCAGTCAAGCTGCCGGC
ACCGAGGCCGAAAACCTGTACTTCCAGAGCCACCATCACCATCACCATTGAgcggccgc

A.3 pET28b(+)(KanR) + *phdB* (between NdeI and BamHI, vector sequence in lower case)

catATGCGCATCATCGTCGCGATCAGCGGCGCCAGCGGCGCACCGTTCGCGGTGCGCC
TACTGGAGACGCTACGCGAGATGCCCGATGTCGAAACCCATTTGGTGATGAGCACCT
GGGGTAAGTCGAACATCGAGGTGGAGACCGACCGCACGGTGTCCGAAGTTGTCGCC
TTGGCCGACGTGACGTACAAGCTCGGTGAGCAGGGCGCGGCGATCTCGTCCGGCTC
GTTCCGTACCACCGGGATGATCATCGTGCCGTGCAGCATGAGGACACTGTCCGCCAT
CCGCTATGGGATGGCGGACAATCTCATCTGCCGGGCCGCCGATGTGGTGCTCAAGG
AGGGGCGGCAGCTGGTGCTGGTGCCGCGGGAAACCCCGCTGAACACCATTCATTTG
GAGAACATGCTTGCGTTGAGCCGAATGGGCGCGCGCATCGTGCCACCGATGCCGGC
CTTCTACAACCATCCGCAGACCATCGGCGACATCGTCGACCATGTAGTGGTGCGCAT
TCTCGATCAGTTCGGGCTGGATGCCCCGCAGGCGAAACGGTGGCGCGGACTGGGCG
CGGCGCGGCGCGACCGGCCGACCCACGCCGGCCATGACCTCAGTCAAGCTGCCGGC
ACCGAGGCCGAAAACCTGTACTTCCAGAGCTGAggatcc

A.4 pET28b(+)(KanR) + *PaubiX* (between NcoI and SalI, vector sequence in lower case)

ccATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCA
GCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGATCCGAATTCGAGC
TCCATGAGCGGCCCGGAGCGTATTACCCTGGCGATGACCGGTGCGAGCGGTGCGCA
ATATGGTCTGCGTCTGCTGGACTGCCTGGTTCAGGAAGAACGTGAGGTGCACTTCCT
GATCAGCAAAGCGGCGCAGCTGGTGATGGCGACCGAAACCGATGTTGCGCTGCCGG
CGAAACCGCAGGCGATGCAAGCGTTCCTGACCGAATATTGCGGTGCGGCGGCGGGT
CAGATTCGTGTTTTTGGTCAAAACGATTGGATGGCTCCGCCGGCGAGCGGCAGCAGC
GCGCCGAACGCGATGGTGATCTGCCCGTGCAGCACCGGTACCCTGAGCGCGGTTGC
GACCGGTGCGTGCAACAACCTGATTGAACGTGCGGCGGATGTGGCGCTGAAGGAAC
GTCGTCCGCTGGTGCTGGTTCCGCGTGAGGCGCCGTTTAGCAGCATCCACCTGGAAA
ACATGCTGAAACTGAGCAACCTGGGTGCGGTTATTCTGCCGGCGGCGCCGGGTTTCT
ATCACCAGCCGCAAAGCGTGGAAGACCTGGTTGATTTTGTGGTTGCGCGTATCCTGA
ACACCCTGGGTATTCCGCAAGATATGCTGCCGCGTTGGGGCGAACAACACCTGGTTA
GCGATGAGTAAgtcgac

A.5 pET28b(+)(KanR) + *EcubiX* (between Eco53kI and SalI, vector sequence in lower case)

gagctccATGAAGCGTCTGATCGTGGGTATTAGCGGTGCGAGCGGTGCGATCTATGGTG
TTCGTCTGCTGCAAGTGCTGCGTGACGTTACCGATATTGAAACCCATCTGGTGATGA
GCCAAGCGGCGCGTCAAACCCTGAGCCTGGAAACCGACTTCAGCCTGCGTGAAGTG
CAAGCGCTGGCGGATGTTACCCACGACGCGCGTGATATCGCGGCGAGCATTAGCAG
CGGCAGCTTTCAGACCCTGGGTATGGTGATCCTGCCGTGCAGCATTAAGACCCTGAG
CGGCATCGTTCACAGCTATACCGATGGTCTGCTGACCCGTGCGGCGGATGTGGTTCT

GAAAGAGCGTCGTCCGCTGGTGCTGTGCGTTCGTGAAACCCCGCTGCACCTGGGTCA
CCTGCGTCTGATGACCCAAGCGGCGGAGATCGGTGCGGTGATTATGCCGCCAGTTCC
GGCGTTCTATCACCGTCCGCAGAGCCTGGACGATGTTATCAACCAAACCGTGAACCG
TGTTCTGGACCAGTTCGCGATTACCCTGCCGGAAGACCTGTTCGCGCGTTGGCAGGG
TGCGCACCACCACCACCACCACTAAgtcgac

A.6 pET28b(+)(KanR) + *SctPAD1* (between NcoI and BamHI, vector sequence in lower case)

ccATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCA
GCCATATGAAACGTATTGTGGTTGCGATCACCGGCGCCACGGGTGTTGCACTGGGCA
TTCGCCTGCTGCAGGTCCTGAAAGAACTGAGCGTGGAAACCCATCTGGTTATCTCTA
AATGGGGTGCGGCCACCATGAAATATGAAACGGATTGGGAACCGCACGACGTTGCA
GCTCTGGCCACCAAAACGTACTCAGTTCGTGATGTCTCGGCATGCATTAGCTCTGGC
AGCTTTCAACACGACGGTATGATCGTCGTGCCGTGTAGTATGAAATCCCTGGCGGCC
ATTCGTATCGGCTTCACCGAAGATCTGATTACGCGCGCAGCTGACGTGTCTATCAAA
GAAAACCGTAAACTGCTGCTGGTTACCCGCGAAACGCCGCTGAGTTCCATTCATCTG
GAAAATATGCTGAGCCTGTGCCGCGCTGGCGTCATTATCTTTCCGCCGGTGCCGGCA
TTCTATACCCGTCCGAAAAGTCTGCACGATCTGCTGGAACAGTCCGTGGGTCGCATC
CTGGACTGTTTCGGCATTCACGCTGACACGTTTCCGCGCTGGGAAGGTATCAAATCA
AAATAAggatcc

A.7 pET28b(+)(KanR) + *TpCAR* (between NheI and EcoRI, vector sequence in lower case)

ccATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCA
GCCATATGGCTAGCTCGATTGAGACGGTGCAGAACGGCGTCCCCGCAGAGGGCTCG
GTGCCCCCCGCCGACCAGCAGACCGAGCGACTGCCGCAGGTGATCGCCAGGATCTT
CGCCCAGTTCGCGGATCGTCCGGCCTTCGCGACCCGCGAGGCGGGGCCGGGGACCC
CCTACGCCACCGTCTCCTATCGGGAGATCTGGCGGCGCGTCACCGCGCTGGTGGCCT
CCTGGCAGAGCGAAGTGGCTCCGGGAGACTTCGTCGCCATCCTCGGCTTCACCAGCT
CGGACTTCGTGACCGTCGACCTCGCGACCACACTGCTCGGCGCCCCGAACGTGCCGC
TCCAGGCCGGGGCCCCCGCCGCTCGCATCGCGACCATCCTCGATGAGACCCGGCCG
AAGATCCTCGCCGTGAGTGCCGATCAGGTCGACCTCGCCCAGGAGGCTCTGGCCGA
GTCCGCGGCTACCCCGCGGGTGGTCGTCTTCGACGGCGAACGCGACGGCTACGAGG
GCATCGAGGCGGACATCCTTTCCGGCTCCGCCCTGCCGGCACCGGAGTTCTTCGCGC
CCGAGCCCGGCACCGATCCTCTCGTCACGCTCATCTACACATCCGGCAGCACCGGTA
CCCCGAAGGGGGCCATGTACACCGAGCAGTTGGTTCGCGATGCCTGGCTCAAGGTG
GACAGCATCGTCGACATCGACATGCCGGCCGAGTCGCTCCTGCACTTCCTGCCTATG
AGCCATATGTACGGGCGCAACTGGCTGATCGCCGGCCTGGCATCGGGCGGGACCGG
GTACTTCGCCGGCGCCTCCGATATGTCCACCCTGTTCGACGACCTCGCCGCCGCCCG
GCCCACCGCCATCGGCCTGGTGCCCCGCGTGTGCGAGCTGATACACCAGCGCTATCT
GGCCGTCGAGGCGGACACTGATGCGGAGACCGCGCGCGTCGAACTGCGTGACCGGG
TACTCGGCGGTCGGCTGCAGGCCGCGATGTGCGGTAGCGCCGCCCTCTCGTCGGAGC
TGCAGACCTTCATGGAGTGGTTGCTCGGAATCGATATCCAGATCGGCTACGGATCCA
CCGAGGCCGGTGGTGTCATCCGCGACGGAGTGGTCGTTCGGCCGCCGGTCACGGAG
TACAAGCTGATCGATGTCCCCGAACTGGGCTACTTCGTCACCGACTCCCCGCATCCA

CGCGGCGAACTCCTGGTCAAGTCGACGCAGTTGATTCCCGGGTACTACAACTCCGAC
AAGCGGATCCGCGACGACGAAGGCTTCTACCGCACCGGCGATGTGATGGCCGAGCT
GGGACCCGACCGGCTCGAGTACGTCGACCGGCGGAGCAACGTGATCAAGTTGGCAC
AGGGAGAGTTCGTGCCGATCGCCCAACTCGAGGCCATCTACGCCGCCGGTCCCGAT
GTGCACCAGATCTTCCTGTACGGAACCAGCGAACGCTCCTACCTGATCGGCGTCGTC
GTGCCCGCGCCGGGACCCGACGGCGAGACCGATGCGCAGACCCGCACCCGCGTACT
CGATGGCCTGGCCGCGATCGCCCGTGAGAACGATCTCGCTGCCTACGAGGTGCCGC
GCGATGTGCTCATCGAACGTGATCCCTTCTCTCAGGAGAACGGGCTGCGGTCGGGGA
TCGGCAAGCTGGTGCGCCCGGCCCTCATCGCCCGCTACGGTGACCGGTTGCACGACC
TCTACGCCCAGGCCGACACCCGTCAACGCGAGGGCTTGCGCGCTCTCGACGCCTCGG
GCCCGATCATCGACACCGTGCTCGGGGCGGCTGCGTTGACGCTCGGCGCGGATATC
GCGGACTTCGACGCCGACACTCGATTCGGCGACCTCGGTGGCGACTCGTTGTCGGCG
CTCTCGCTCGCGACGACGCTCGAAGGCCTCTACGACGTGCCCGTCCCCGTGCAGACG
ATCGTCGGACCGACCGCCACACTCGGCGGCGTCGCCCGGCACATCGAGAAGGCTCG
ATCGGGTGGCGTCGCGGCACCGACCGCCGACTCGGTGCACGGCGTGGGTGCGAGCG
TCGCCCGGGCCACCGACCTGACGCTGGAGAAGTTCATCGACCCCGAGCTCCTCGCGC
TCGCGCCGACGCTTCCCGCGGCGACCGGTGAGCCGAACACCGTGCTGCTCACCGGA
TCCACCGGCTACCTCGGCCGCTTCCTGCTGCTGGACTGGTTGCGACGGGTCGCTCCG
CACGGCGGCACCGTGATCGCGCTGGTGCGCGGCGCCGACGCCGACGATGCGCGACG
CCGCGTCACGGCCGCGATCGGTGACTCGGATCCTGACCTGACACAAGAGTTCACGTC
ACTCGCGGAGCATCACCTCCACGTGATCGCCGGTGACTTCGGCAGCCCCGCACTCGG
ACTCGACGATGCCACCTGGAGCGATCTCGCCGGGCGAGTCGATCACGTGGTGCACT
GCGGCGCGCTCGTCAACCACGTGCTGCCCTACGACCAACTGTTCGGTCCCAATGTGG
TGGCCACCGGCGAAGTGGTGCGACTCGCACTCACCACGCGCCGCAAGTCCGTGGAT
TACGTCTCCACGGTGGCTGTGGTTCCGCAGGATGACGGCCGCGTCCTGGTCGAGGAC
GACGATGTTCGCGAGCTCGGCGCCGAACGGCGCATCGGGGCCGATGCCTACGCGAA
CGGCTACGCCGTGAGCAAATGGGCGGGCGAAGTGCTGTTGCATGAGGCAGCCGACC
TGGCGGACCTGCCGGTGCGGGTGTTCCGCTCCGATATGATCTTGGCGCACAGTCGAT
TCCACGGACAGTTCAACGAGGTCGACCAGTTCACCCGCCTGCTCCTGAGTATCGCCG
AGACCGGACTGGCGCCGGCGTCGTTCTACACGCCGGATCCGAGTGGACACCGCCCG
CACTACGACGGGCTGCCGGTGGACTTCACCGCCGAAGCGATCACCACGCTCAGCGC
CGCGGGGCGTTCGGGGTACCGGACCTTCCACGTGCTCAACGCCAACGATGACGGCG
TGAGCCTGGACAGCTTCGTCGACTGGATCGCCGCCTCGGGCCGGAGCATCGAACGG
ATCGACGACTACGACACCTGGTTCGCCCGGTTCGAGCAGGCGCTCCAGCAGCTCCCC
GATGAGGCGCGCCAGCGGTCGGTGCTGCCCCTGCTGCACGCGGTGCGCGAGCCGGC
TCCGGCCGCCGGGACCTCCGCGCTGTCGGTGGACCGGTTCCGTGGTGCGGTGCGTGA
GACCGGAGTAGGACCGGGGGACATCCCGGTGCTCGATCGCGCCCTGATCGAGAAGT
ACCTGCGCGACTTCGAGACCGCGGGCTGGCTCGCGCCCGGTGCGCGCGACTGAgaattc

A.8 pET20b(+)(AmpR) + *Bssfp* (between NdeI and EcoRI, vector sequence in lower case)

catATGAAGATTTACGGAATTTATATGGACCGCCCGCTTTCACAGGAAGAAAATGAAC
GGTTCATGACTTTCATATCACCTGAAAAACGGGAGAAATGCCGGAGATTTTATCATA
AAGAAGATGCTCACCGCACCCTGCTGGGAGATGTGCTCGTTCGCTCAGTCATAAGCA
GGCAGTATCAGTTGGACAAATCCGATATCCGCTTTAGCACGCAGGAATACGGGAAG
CCGTGCATCCCTGATCTTCCCGACGCTCATTTCAACATTTCTCACTCCGGCCGCTGGG
TCATTGGTGCGTTTGATTCACAGCCGATCGGCATAGATATCGAAAAAACGAAACCG

ATCAGCCTTGAGATCGCCAAGCGCTTCTTTTCAAAAACAGAGTACAGCGACCTTTTA
GCAAAAGACAAGGACGAGCAGACAGACTATTTTTATCATCTATGGTCAATGAAAGA
AAGCTTTATCAAACAGGAAGGCAAAGGCTTATCGCTTCCGCTTGATTCCTTTTCAGT
GCGCCTGCATCAGGACGGACAAGTATCCATTGAGCTTCCGGACAGCCATTCCCCATG
CTATATCAAAACGTATGAGGTCGATCCCGGCTACAAAATGGCTGTATGCGCCGCACA
CCCTGATTTCCCCGAGGATATCACAATGGTCTCGTACGAAGAGCTTTTATAAgaattc

Appendix B: List of Unique Intra-protein H-bonds

H_2O only	% Existence	D_2O only	% Existence
ARG5GLU237	66.632	TYR10PHE68	97.71
ARG5GLU237	66.792	ASN30GLU34	50.307
TYR10ASP6	74.536	THR39THR66	98.1
TYR10ARG74	50.697	HSD62MET128	66.782
HSD62PRO59	54.687	MET69GLY315	98.475
THR66GLY304	87.666	GLY70ASN67	68.752
ASN67THR39	63.247	THR71GLY70	98.9
ARG74GLU22	72.926	ILE91ALA87	77.591
ARG74GLU22	65.632	LYS95ASP96	98.985
LEU79CYS75	75.716	LYS99GLU332	82.056
LYS95ASP329	51.777	GLY121GLN292	61.037
LEU126LYS171	65.477	GLU127ASP125	98.715
ARG133GLY139	95.945	TYR142GLU270	57.227
TYR142PRO272	99.63	CYS147TYR164	66.727
ARG152GLY157	84.681	MET148VAL203	93.58
ARG152GLY245	77.286	TRP158GLY221	72.341
ARG173GLY170	67.887	VAL161GLU224	96.705
ARG173GLU266	59.452	ARG165GLU270	56.632

ARG173ASP172	67.017	ARG165GLU270	52.097
ARG173GLU266	71.361	LYS171ASP125	92.74
ARG173GLU266	60.172	LYS171GLU127	68.947
SER175GLU286	99.75	MET178GLU286	57.002
TRP177ARG165	52.107	ARG182LEU179	83.641
TRP177GLU270	58.752	ARG194ASP155	81.616
ARG232GLY235	88.806	ARG194ASP155	71.891
SER240VAL109	80.986	ARG194ASP155	54.142
GLY245PRO242	67.972	LEU210ASP207	58.172
ASN251GLU237	50.517	GLU224GLU224	91.915
HSD269ASP285	53.432	ARG232ALA228	86.976
GLY271GLU286	99.005	ARG232PHE4	81.836
ALA283GLU270	75.516	ALA241LEU246	99.425
ASP285GLU270	72.256	GLY243GLU257	73.106
ARG297ASP6	98.98	THR244GLU257	98.82
GLN317ALA314	64.632	THR244GLU257	99.75
ALA318GLY315	70.306	GLY245ALA241	53.637
VAL319GLY315	65.912	ILE248ILE239	96.84
LEU321VAL319	57.772	HSD269GLY284	67.817
SER323GLN348	91.465	PHE273THR281	91.32
ALA324PRO320	98.205	GLU275TYR279	94.12
SER325LEU321	73.436	TRP276GLU275	99.66
SER325LEU321	95.695	THR277TYR279	70.451

ASN338GLN357	63.712	THR277ILE422	99.42
ARG347GLN317	52.077	GLY278GLU275	87.236
GLN348GLY377	83.501	THR281PHE273	84.826
SER354ASP387	96.05	THR281ALA283	83.486
GLU356ASN338	93.85	ARG290GLU266	51.377
THR358LYS447	68.892	ARG290GLU266	54.577
ARG373ALA370	98.295	ALA318MET69	78.966
LYS374HSD330	97.84	TYR322ARG375	57.062
ARG375ASP380	92.55	THR327ARG373	74.031
ARG375ARG373	96.38	THR327ARG373	79.861
ARG375ASP380	97	LYS340GLU356	61.542
ARG381ASP380	98.895	LYS340GLU463	54.762
GLU388GLU388	85.206	GLN348ALA346	65.537
ASN396ASP394	52.377	GLN348VAL379	76.436
LYS452PHE449	67.292	THR358ASP362	78.546
LYS452ASP387	98.515	GLY359ASP362	64.747
LYS452GLU388	86.571	GLY372ALA369	65.877
ARG457ASP453	75.756	ARG373ASP380	93.4
ARG457TYR455	72.676	VAL379THR419	94.36
ARG457ASP389	83.041	GLU388ILE355	92.35
ARG470ASP96	61.342	VAL398ASP394	87.931
		SER427LYS430	52.132
		ARG431ALA421	93.08

		ARG431ASP423	66.342
		SER454TRP451	76.761
		ALA458ASP389	53.357
		GLU468PRO464	70.651
		ILE469TYR465	80.221
		TRP473ASP89	85.256
		LYS474ASP96	61.027
		LEU477TRP473	69.057

Table B.1: List of hydrogen bonded residues along with their % existence during the final 20 ns of the simulation trajectory. Only bonds with % existence greater than 50% are included. Bonds unique to the H_2O solvated protein are in column 1 whereas those observed only in D_2O are mentioned in column 3.

Bibliography

(1) Li, T.; Huo, L.; Pulley, C; Aimin, L. Decarboxylation mechanisms in biological system. *Bioorg. Chem.* **2012**, *43*, 2-14.
(2) Lin, F.; Ferguson, K. L.; Boyer, D. R.; Lin, X. N.; Marsh, E. N. G. Isofunctional Enzymes PAD1 and UbiX Catalyze Formation of a Novel Cofactor Required by Ferulic Acid Decarboxylase and 4-Hydroxy-3-polyprenylbenzoic Acid Decarboxylase. *ACS Chem. Biol.* **2015**, *10* (4), 1137-1144.
(3) Marshall, S. A.; Payne, K. A. P.; Leys, D. The UbiX-UbiD system: The biosynthesis and use of prenylated flavin (prFMN). *Arch. Biochem. Biophys.* **2017**, *632*, 209-221.
(4) Payne, K. A. P.; White, M. D.; Fisher, K.; Khara, B.; Bailey, S. S.; Parker, D.; Rattray, N. J. W.; Trivedi, D. K.; Goodacre, R.; Beveridge, R.; Barran, B.; Rigby, S. E. J.; Scrutton, N. S.; Hay, S.; Leys. L. New cofactor supports α,β-unsaturated acid decarboxylation via 1,3-dipolar cycloaddition. *Nature* **2015**, *522* (7557), 497-501.
(5) Fujiwara, R.; Noda, S.; Tanaka, T.; Kondo, A. Styrene production from a biomass-derived carbon source using a coculture system of phenylalanine ammonia lyase and phenylacrylic acid decarboxylase-expressing *Streptomyces lividans* transformants. *J. Biosci. Bioeng.* **2016**, *122* (6), 730-735.
(6) Messiha, H. L.; Payne, K. A. P.; Scrutton, N. S.; Leys, D. A Biological Route to Conjugated Alkenes: Microbial Production of Hepta-1,3,5-triene. *ACS Synth. Biol.* **2021**, *10* (2), 228-235.
(7) Ren, J.; Yao, P.; Yu, S.; Dong, W.; Chen, Q.; Feng, J.; Wu, Q.; Zhu, D. An Unprecedented Effective Enzymatic Carboxylation of Phenols. *ACS Catal.* **2016**, *6* (2), 564-567.
(8) Tong, X.; El-Zahab, B.; Zhao, X.; Liu, Y.; Wang, P. Enzymatic synthesis of L-lactic acid from carbon dioxide and ethanol with an inherent cofactor regeneration cycle. *Biotechnol. Bioeng.* **2011**, *108* (2), 465-469.
(9) Lopez-Lorenzo, X.; Asem, H.; Stamm, A.; Subramaniyan, S.; Hakkarainen, M.; Syrén, P.-O. Whole-cell Mediated Carboxylation of 2-Furoic Acid Towards the Production of Renewable Platform Chemicals and Biomaterials. *ChemCatChem* **2023**, *15* (6), e202201483.
(10) Adkins, J.; Pugh, S.; McKenna, R.; Nielsen, D. Engineering microbial chemical factories to produce renewable "biomonomers". *Front. Microbiol.* **2012**, *3*.
(11) Payer, S. E.; Marshall, S. A.; Bärland, N.; Sheng, X.; Reiter, T.; Dordic, A.; Steinkellner, G.; Wuensch, C.; Kaltwasser, S.; Fisher, K.; Rigby, S. E. J.; Macheroux, P.; Vonck, J.; Gruber, K.; Faber, K.; Himo, F.; Leys, D.; Pavkov-Keller, T.; Glueck, S. M. Regioselective para-Carboxylation of Catechols with a Prenylated Flavin Dependent Decarboxylase. *Angew. Chem.* **2017**, *56* (44), 13893-13897.
(12) Bierbaumer, S.; Nattermann, M.; Schulz, L.; Zschoche, R.; Erb, T. J.; Winkler, C. K.; Tinzl, M.; Glueck, S. M. Enzymatic Conversion of CO_2: From Natural to Artificial Utilization. *Chem. Rev.* **2023**, *123* (9), 5702-5754.
(13) Aleku, G. A.; Roberts, G. W.; Titchiner, G. R.; Leys, D. Synthetic Enzyme-Catalyzed CO_2 Fixation Reactions. *ChemSusChem* **2021**, *14* (8), 1781-1804.
(14) Miyazaki, M.; Shibue, M.; Ogino, K.; Nakamura, H.; Maeda, H. Enzymatic synthesis of pyruvic acid from acetaldehyde and carbon dioxide. *Chem. Comm.* **2001**, (18), 1800-1801.

(15) Wuensch, C.; Schmidt, N.; Gross, J.; Grischek, B.; Glueck, S. M.; Faber, K. Pushing the equilibrium of regio-complementary carboxylation of phenols and hydroxystyrene derivatives. *J. Biotech.* **2013**, *168* (3), 264-270.
(16) Aleku, G. A.; Saaret, A.; Bradshaw-Allen, R. T.; Derrington, S. R.; Titchiner, G. R.; Gostimskaya, I.; Gahloth, D.; Parker, D. A.; Hay, S.; Leys, D. Enzymatic C-H activation of aromatic compounds through CO_2 fixation. *Nat. Chem. Biol.* **2020**, *16*, 1255-1260.
(17) Titchiner, G. R.; Marshall, S. A.; Miscikas, H.; Leys, D. Biosynthesis of Pyrrole-2-carbaldehyde via Enzymatic CO_2 Fixation. *Catal.* **2022,** 12.
(18) Meganathan, R. Ubiquinone biosynthesis in microorganisms. *FEMS Microbiol. Lett.* **2001**, *203* (2), 131-139.
(19) Zhang, H.; Javor, G. T. Identification of the *ubiD* Gene on the *Escherichia coli* Chromosome. *J. Bacteriol.* **2000**, *182* (21), 6243-6246.
(20) Ernster, L.; Dallner, G. Biochemical, physiological and medical aspects of ubiquinone function. *Biochim. Biophys. Acta* **1995**, *1271* (1), 195-204.
(21) Lupa, B.; Lyon, D.; Gibbs, M. D.; Reeves, R. A.; Wiegel, J. Distribution of genes encoding the microbial non-oxidative reversible hydroxyarylic acid decarboxylases/phenol carboxylases. *Genomics* **2005**, *86* (3), 342-351.
(22) Saaret, A.; Balaikaite, A.; Leys, D. Biochemistry of prenylated-FMN enzymes. *Enzymes* **2020**, *47*, 517-549.
(23) White, M. D.; Payne, K. A. P.; Fisher, K.; Marshall, S. A.; Parker, D.; Rattray, N. J. W.; Trivedi, D. K.; Goodacre, R.; Rigby, S. E. J.; Scrutton, N. S.; Hay, S.; Leys, D. UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis. *Nature* **2015**, *522* (7557), 502-506.
(24) Arunrattanamook, N.; Marsh, E. N. G. Kinetic Characterization of Prenyl-Flavin Synthase from *Saccharomyces cerevisiae*. *Biochemistry* **2018**, *57* (5), 696-700.
(25) Marshall, S. A.; Payne, K. A. P.; Fisher, K.; White, M. D.; Ní Cheallaigh, A.; Balaikaite, A.; Rigby, S. E. J.; Leys, D. The UbiX flavin prenyltransferase reaction mechanism resembles class I terpene cyclase chemistry. *Nat. Commun.* **2019**, *10* (1), 2357.
(26) Wang, P. H.; Khusnutdinova, A. N.; Luo, F.; Xiao, J.; Nemr, K.; Flick, R.; Brown, G.; Mahadevan, R.; Edwards, E. A.; Yakunin, A. F. Biosynthesis and Activity of Prenylated FMN Cofactors. *Cell Chem. Biol.* **2018**, *25* (5), 560.
(27) Balaikaite, A.; Chisanga, M.; Fisher, K.; Heyes, D. J.; Spiess, R.; Leys, D. Ferulic Acid Decarboxylase Controls Oxidative Maturation of the Prenylated Flavin Mononucleotide Cofactor. *ACS Chem. Biol.* **2020**, *15* (9), 2466-2475.
(28) Bailey, S. S.; Payne, K. A. P.; Fisher, K.; Marshall, S. A.; Cliff, M. J.; Spiess, R.; Parker, D. A.; Rigby, S. E. J.; Leys, D. The role of conserved residues in Fdc decarboxylase in prenylated flavin mononucleotide oxidative maturation, cofactor isomerization, and catalysis. *J. Biol. Chem.* **2018**, *293* (7), 2272-2287.
(29) Annaval, T.; Han, L.; Rudolf, J. D.; Xie, G.; Yang, D.; Chang, C.-Y.; Ma, M.; Crnovcic, I.; Miller, M. D.; Soman, J.; Xu, W.; Phillips, G. N.; Shen, Ben. Biochemical and Structural Characterization of TtnD, a Prenylated FMN-Dependent Decarboxylase from the Tautomycetin Biosynthetic Pathway. *ACS Chem. Biol.* **2018**, *13* (9), 2728-2738.
(30) Marshall, S. A.; Payne, K. A. P.; Fisher, K.; Titchiner, G. R.; Levy, C.; Hay, S.; Leys, D. UbiD domain dynamics underpins aromatic decarboxylation. *Nat. Comm.* **2021**, *12* (1), 5065.
(31) Marshall, S. A.; Fisher, K.; Ní Cheallaigh, A.; White, M. D.; Payne, K. A.; Parker, D. A.; Rigby, S. E.; Leys, D. Oxidative Maturation and Structural Characterization of Prenylated FMN

Binding by UbiD, a Decarboxylase Involved in Bacterial Ubiquinone Biosynthesis. *J. Biol. Chem.* **2017**, *292* (11), 4623-4637.
(32) Payne, K. A. P.; Marshall, S. A.; Fisher, K.; Cliff, M. J.; Cannas, D. M.; Yan, C.; Heyes, D. J.; Parker, D. A.; Larrosa, I.; Leys, D. Enzymatic Carboxylation of 2-Furoic Acid Yields 2,5-Furandicarboxylic Acid (FDCA). *ACS Catal.* **2019**, *9* (4), 2854-2865.
(33) Payne, K. A. P.; Marshall, S. A.; Fisher, K.; Rigby, S. E. J.; Cliff, M. J.; Spiess, R.; Cannas, D. M.; Larrosa, I.; Hay, S.; Leys, D. Structure and Mechanism of *Pseudomonas aeruginosa* PA0254/HudA, a prFMN-Dependent Pyrrole-2-carboxylic Acid Decarboxylase Linked to Virulence. *ACS Catal.* **2021**, *11* (5), 2865-2878.
(34) Gahloth, D.; Fisher, K.; Payne, K. A. P.; Cliff, M.; Levy, C.; Leys, D. Structural and biochemical characterization of the prenylated flavin mononucleotide-dependent indole-3-carboxylic acid decarboxylase. *J. Biol. Chem.* **2022**, *298* (4).
(35) Mergelsberg, M.; Willistein, M.; Meyer, H.; Stärk, H.-J.; Bechtel, D. F.; Pierik, A. J.; Boll, M. Phthaloyl-coenzyme A decarboxylase from Thauera chlorobenzoica: the prenylated flavin-, K^+- and Fe^{2+}-dependent key enzyme of anaerobic phthalate degradation. *Environ. Microbiol.* **2017**, *19* (9), 3734-3744.
(36) Costa, K. C.; Moskatel, L. S.; Meirelles, L. A.; Newman, D. K. PhdA Catalyzes the First Step of Phenazine-1-Carboxylic Acid Degradation in *Mycobacterium fortuitum*. *J. Bacteriol.* **2018**, *200* (10), e00763-00717.
(37) Kaneshiro, A. K.; Koebke, K. J.; Zhao, C.; Ferguson, K. L.; Ballou, D. P.; Palfey, B. A.; Ruotolo, B. T.; Marsh, E. N. G. Kinetic Analysis of Transient Intermediates in the Mechanism of Prenyl-Flavin-Dependent Ferulic Acid Decarboxylase. *Biochemistry* **2021**, *60* (2), 125-134.
(38) Ferguson, K. L.; Arunrattanamook, N.; Marsh, E. N. G. Mechanism of the Novel Prenylated Flavin-Containing Enzyme Ferulic Acid Decarboxylase Probed by Isotope Effects and Linear Free-Energy Relationships. *Biochemistry* **2016**, *55* (20), 2857-2863.
(39) Ferguson, K. L.; Eschweiler, J. D.; Ruotolo, B. T.; Marsh, E. N. G. Evidence for a 1,3-Dipolar Cyclo-addition Mechanism in the Decarboxylation of Phenylacrylic Acids Catalyzed by Ferulic Acid Decarboxylase. *J. Am. Chem. Soc.* **2017**, *139* (32), 10972-10975.
(40) Bailey, S. S.; Payne, K. A. P.; Saaret, A.; Marshall, S. A.; Gostimskaya, I.; Kosov, I.; Fisher, K.; Hay, S.; Leys, D. Enzymatic control of cycloadduct conformation ensures reversible 1,3-dipolar cycloaddition in a prFMN-dependent decarboxylase. *Nat. Chem.* **2019**, *11* (11), 1049-1057.
(41) Roberts, G. W.; Leys, D. Structural insights into UbiD reversible decarboxylation. *Curr. Opin. Struct. Biol.* **2022**, *75*, 102432.
(42) Kaneshiro, A. K.; Datar, P. M.; Marsh, E. N. G. Negative Cooperativity in the Mechanism of Prenylated-Flavin-Dependent Ferulic Acid Decarboxylase: A Proposal for a "Two-Stroke" Decarboxylation Cycle. *Biochemistry* **2023**, *62* (1), 53-61.
(43) Price-Whelan, A.; Dietrich, L. E. P.; Newman, D. K. Rethinking 'secondary' metabolism: physiological roles for phenazine antibiotics. *Nat. Chem. Biol.* **2006**, *2* (2), 71-78.
(44) Mavrodi, D. V.; Parejko, J. A.; Mavrodi, O. V.; Kwak, Y. S.; Weller, D. M.; Blankenfeldt, W.; Thomashow, L. S. Recent insights into the diversity, frequency and ecological roles of phenazines in fluorescent *Pseudomonas* spp. *Environ. Microbiol.* **2013**, *15* (3), 675-686.
(45) Costa, K. C.; Bergkessel, M.; Saunders, S.; Korlach, J.; Newman, D. K. Enzymatic Degradation of Phenazines Can Generate Energy and Protect Sensitive Organisms from Toxicity. *mBio.* **2015**, *6* (6), e01520-01515.

(46) Duță, H.; Filip, A.; Nagy, L. C.; Nagy, E. Z. A.; Tőtős, R.; Bencze, L. C. Toolbox for the structure-guided evolution of ferulic acid decarboxylase (FDC). *Sci. Rep.* **2022**, *12* (1), 3347.
(47) Kaye, R. C.; Stonehill, H. I. The Polarographic Reduction of Pyridine, Quinoline, and Phenaxine. *J. Chem. Soc.* **1952**, 3240-3243.
(48) Schowen, K. B.; Schowen, R. L. Solvent isotope effects of enzyme systems. *Methods Enzymol.* **1982**, *87*, 551-606.
(49) Gadda, G.; Sobrado, P. Kinetic Solvent Viscosity Effects as Probes for Studying the Mechanisms of Enzyme Action. *Biochemistry* **2018**, *57* (25), 3445-3453.
(50) Raber, M. L.; Freeman, M. F.; Townsend, C. A. Dissection of the Stepwise Mechanism to β-Lactam Formation and Elucidation of a Rate-determining Conformational Change in β-Lactam Synthetase. *J. Biol. Chem.* **2009**, *284* (1), 207-217.
(51) Gellatly, S. L.; Hancock, R. E. W. *Pseudomonas aeruginosa*: new insights into pathogenesis and host defenses. *Pathog. Dis.* **2013**, *67* (3), 159-173.
(52) Wang, Y.; Newman, D. K. Redox Reactions of Phenazine Antibiotics with Ferric (Hydr)oxides and Molecular Oxygen. *Environ. Sci. Technol.* **2008**, *42* (7), 2380-2386.
(53) Batyrova, K. A.; Khusnutdinova, A. N.; Wang, P.-H.; Leo, R. D.; Flick, R.; Edwards, E. A.; Savchenko, A.; Yakunin, A. F. Biocatalytic in Vitro and in Vivo FMN Prenylation and (De)carboxylase Activation. *ACS Chem. Biol.* **2020**, *15*, 1874−1882.
(54) Finnigan, W.; Thomas, A.; Cromar, H.; Gough, B.; Snajdrova, R.; Adams, J. P.; Littlechild, J. A.; Harmer, N. J. Characterization of Carboxylic Acid Reductases as Enzymes in the Toolbox for Synthetic Chemistry. *ChemCatChem* **2017**, *9* (6), 1005-1017.
(55) Ansell, H. V.; Hirschler, M. M.; Taylor, R. Electrophilic aromatic substitution. Part 18. Protiodetritiation of anthracene, coronene (dibenzo[ghi, pqr]perylene), and triphenylene in anhydrous trifluoroacetic acid. *J. Chem. Soc., Perkin Trans. 2* **1977**, 353-355.
(56) Corsaro, A.; Pistarà, V.; Rescifina, A.; Piperno, A.; Chiacchio, M. A.; Romeo, G. A DFT rationalization for the observed regiochemistry in the nitrile oxide cycloaddition with anthracene and acridine. *Tetrahedron* **2004**, *60* (31), 6443-6451.
(57) Librando, V.; Chiacchio, U.; Corsaro, A.; Gumina, G. Dipolarophilic Reactivity of Polycyclic Aromatic Compounds Towards Nitrile Oxides. *Polycycl. Aromat. Compd.* **1996**, *11* (1-4), 313-316.
(58) Corsaro, A.; Librando, V.; Chiacchio, U.; Pistarà, V. 1,3-Dipolar cycloaddition reactions of polycyclic aromatic hydrocarbons with 3,5-dichloro-2,4,6-trimethyl- and 2,4,6-trimethylbenzonitrile oxide. *Tetrahedron* **1996**, *52* (40), 13027-13034.
(59) Kawanabe, K.; Aono, R.; Kino, K. 2,5-Furandicarboxylic acid production from furfural by sequential biocatalytic reactions. *J. Biosci. Bioeng.* **2021**, *123*, 18-24.
(60) Payer, S. E.; Faber, K.; Glueck, S. M. Non-Oxidative Enzymatic (De)Carboxylation of (Hetero)Aromatics and Acrylic Acid Derivatives. *Adv. Synth. Catal.* **2019**, *361* (11), 2402-2420.
(61) Leys, D. Flavin metamorphosis: cofactor transformation through prenylation. *Curr. Opin. Chem. Biol.* **2018**, *47*, 117-125.
(62) Piano, V.; Palfey, B. A.; Mattevi, A. Flavins as Covalent Catalysts: New Mechanisms Emerge. *Trends Biochem. Sci.* **2017**, *42* (6), 457-469.
(63) Leys, D.; Scrutton, N. S. Sweating the assets of flavin cofactors: new insight of chemical versatility from knowledge of structure and mechanism. *Curr. Opin. Struct. Biol.* **2016**, *41*, 19-26.

(64) Datar, P. M.; Marsh, E. N. G. Decarboxylation of Aromatic Carboxylic Acids by the Prenylated-FMN-dependent Enzyme Phenazine-1-carboxylic Acid Decarboxylase. *ACS Catal.* **2021**, *11* (18), 11723-11732.
(65) Gahloth, D.; Leys, D. Crystal structure of Phenazine 1-carboxylic acid decarboxylase from M*ycobacterium fortuitum*. 2021.
(66) Quinn, D. M.; Sutton, L. D. Theoretical basis and mechanistic utility of solvent isotope effects. In *Enzyme mechanism from Isotope effects*, Cook, P. F. Ed.; CRC Press, 1991; pp 73 - 126.
(67) Haynes, W. M. *CRC Handbook of Chemistry and Physics*; CRC press, 2014.
(68) Venkatasubban, K. S.; Schowen, R. L. The Proton Inventory Techniqu. *Crit. Rev. Biochem.* **1984**, *17* (1), 1-44.
(69) Royer, C. A.; Mann, C. J.; Matthews, C. R. Resolution of the fluorescence equilibrium unfolding profile of trp aporepressor using single tryptophan mutants. *Prot. Sci.* **1993**, *2* (11), 1844-1852.
(70) Pariso, M.; Mazzini, A.; Sorbi, R.; Ramoni, R.; Grolli, S.; Favilla, R. Unfolding and refolding of porcine odorant binding protein in guanidinium hydrochloride: equilibrium studies at neutral pH. *Biochim. Biophys. Acta,* **2003**, *1652* (2), 115-125.
(71) Fersht, A. *Enzyme Structure and Mechanism*; W.H. Freeman and Company, 1977.
(72) Stein, R. L. Transition-State Properties for the Association of α -1-Protease Inhibitor with Porcine Pancreatic Elastase. *J. Am. Chem. Soc.* **1985**, *107* (21), 6039-6042.
(73) Cleland, W. W. Partition Analysis and the Concept of Net Rate Constants as Tools in Enzyme Kinetics. *Biochemistry* **1975**, *14* (14), 3220-3224.
(74) Northrop, D. B. The expression of isotope effects on enzyme-catalyzed reactions. *Annual review of biochemistry* **1981**, *50* (1), 103-131.
(75) Schowen, R. L. Catalytic Power and Transititon State Stabilization. In *Transition States of Biochemical Processes*, Gandour, R. D., Schowen, R. L. Eds.; Vol. 1; Springer, 1978; pp 77-114.
(76) Stein, R. L. Analysis of Kinetic Isotope Effects on Complex Reactions Utilizing the Concept of the Virtual Transition State. *J. Org. Chem.* **1981**, *46*, 3328-3330.
(77) Quinn, D. M. Acetylcholinesterase: Enzyme Structure, Reaction Dynamics, and Virtual Transition States. *Chem. Rev.* **1987**, *87*, 955-979.
(78) Northrop, D. B. Limits on the Expression of Enzyme-Mediated Solvent Isotope Effects. *J. Am. Chem. Soc.* **1981**, *103*, 1208-1212.
(79) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25.
(80) Huang, J.; MacKerell Jr, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry* **2013**, *34* (25), 2135-2145.
(81) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31* (4), 671-690.
(82) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A,* **2001**, *105* (43), 9954-9960.
(83) Linse, J.-B.; Hub, J. S. Three- and four-site models for heavy water: SPC/E-HW, TIP3P-HW, and TIP4P/2005-HW. *J. Chem. Phys.* **2021**, *154* (19).

(84) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089-10092.
(85) Curry, H. B. The Method of Steepest Descent for Non-Linear Minimization Problems. *Q. Appl. Math.* **1944**, *2* (3), 258-261.
(86) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52* (2), 255-268.
(87) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A***, 1985**, *31* (3), 1695-1697.
(88) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52* (12), 7182-7190.
(89) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph* **1996**, *14* (1), 33-38, 27-38.
(90) Róg, T.; Murzyn, K.; Milhaud, J.; Karttunen, M.; Pasenkiewicz-Gierula, M. Water Isotope Effect on the Phosphatidylcholine Bilayer Properties: A Molecular Dynamics Simulation Study. *J. Phys. Chem. B* **2009**, *113* (8), 2378-2387.
(91) Robbins, J. M.; Ellis, H. R. Identification of Critical Steps Governing the Two-Component Alkanesulfonate Monooxygenase Catalytic Mechanism. *Biochemistry* **2012**, *51* (32), 6378-6387.
(92) Karsten, W. E.; Lai, C.-J.; Cook, P. F. Inverse Solvent Isotope Effects in the NAD-Malic Enzyme Reaction Are the Result of the Viscosity Difference between D_2O and H_2O: Implications for Solvent Isotope Effect Studies. *J. Am. Chem. Soc.* **1995**, *117*, 5914 - 5918.
(93) Maybury, R. H.; Katz, J. J. Protein denaturation in heavy water. *Nature* **1956**, *177*, 629 - 630.
(94) Hermans, J. J.; Sheraga, H. A. The thermally induced configurational change of ribonuclease in H_2O and D_2O. *Biochim. Biophys. Acta.* **1959**, *36*, 534–535.
(95) Parker, M. J.; Clarke, A. R. Amide backbone and water-related H/D isotope effects on the dynamics of a protein folding reaction. *Biochemistry* **1997**, *36*, 5786–5794.
(96) Cioni, P.; Strambini, G. B. Effect of Heavy Water on Protein Flexibility. *Biophys. J.* **2002**, *82* (6), 3246-3253.
(97) Northrop, D. On the Meaning of K_m and V/K in Enzyme Kinetics. *J. Chem. Educ.* **1998**, *75* (9), 1153-1157.
(98) Papaioannou, A.; Kuyucak, S.; Kuncic, Z. Molecular Dynamics Simulations of Insulin: Elucidating the Conformational Changes that Enable Its Binding. *PLoS One* **2015**, *10* (12), e0144058.
(99) Guterres, H.; Im, W. Improving Protein-Ligand Docking Results with High-Throughput Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2020**, *60* (4), 2189-2198.
(100) Martínez, L. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS One* **2015**, *10* (3), e0119264.
(101) Maisuradze, G. G.; Leitner, D. M. Free energy landscape of a biomolecule in dihedral principal component space: Sampling convergence and correspondence between structures and minima. *Proteins: Structure, Function, and Bioinformatics* **2007**, *67* (3), 569-578.
(102) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem.* **1999**, *38* (1-2), 236-240.
(103) Sheu, S.-Y.; Schlag, E. W.; Selzle, H. L.; Yang, D.-Y. Molecular Dynamics of Hydrogen Bonds in Protein−D2O: The Solvent Isotope Effect. *J. Phys. Chem. A* **2008**, *112* (5), 797-802.

(104) Gekko, K.; Timasheff, S. N. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry* **1981**, *20* (16), 4667-4676.
(105) Sobrado, P.; Daubner, S. C.; Fitzpatrick, P. F. Probing the Relative Timing of Hydrogen Abstraction Steps in the Flavocytochrome b2 Reaction with Primary and Solvent Deuterium Isotope Effects and Mutant Enzymes. *Biochemistry* **2001**, *40* (4), 994-1001.
(106) Entsch, B.; Sim, R. G. The purification and identification of flavin nucleotides by high-performance liquid chromatography. *Anal. Biochem.* **1983**, *133* (2), 401-408.
(107) Kemal, C.; Bruice, T. C. Simple synthesis of a 4a-hydroperoxy adduct of a 1,5-dihydroflavine: preliminary studies of a model for bacterial luciferase. *Proc. Nat. Acad. Sci.* **1976**, *73* (4), 995-999.
(108) Romero, E.; Gómez Castellanos, J. R.; Gadda, G.; Fraaije, M. W.; Mattevi, A. Same Substrate, Many Reactions: Oxygen Activation in Flavoenzymes. *Chem. Rev.* **2018**, *118* (4), 1742-1769.
(109) Kemal, C.; Bruice, T. C. The chemistry of an N5-methyl-1, 5-dihydroflavin and its aminium cation radical. *J. Am. Chem. Soc.* **1976**, *98* (13), 3955-3964.
(110) Kemal, C.; Chan, T. W.; Bruice, T. C. Reaction of $^3O_2$ with dihydroflavins. 1. N3,5-Dimethyl-1,5-dihydrolumiflavin and 1,5-dihydroisoalloxazines. *J. Am. Chem. Soc.* **1977**, *99* (22), 7272-7286.
(111) Kemal, C. Exploring Flavin Chemistry With N(5)-Blocked Models. Ph.D., University of California, Santa Barbara, United States - California, 1977.
(112) Northrop, D. B. Intrinsic Isotope Effects in Enzyme-Catalyzed Reactions. In *Enzyme Mechanisms from Isotope Effects*, Cook, P. F. Ed.; CRC, 1991; pp 182-202.