OXYGEN-DEPENDENT ENZYMES INVOLVED IN PHOSPHONATE BIOSYNTHESIS

BY

MICHELLE NICOLE GOETTGE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Microbiology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Professor William W. Metcalf, Chair
Professor Wilfred A. van der Donk
Professor James A. Imlay
Associate Professor Carin K. Vanderpool

Abstract

Microorganisms are an invaluable source of natural products. As humans, we have taken advantage of these small molecules to improve overall human health and quality of life. However, the rise in the number of pathogens that have become resistant to modern medicines requires action to deliver new compounds to treat disease. Natural products are a rich source of potential drug leads, as approximately 65% of our current antimicrobials are natural products or inspired by natural products. With the advent of genome sequencing, we now have the ability to quickly identify new biosynthetic gene clusters with the capacity to produce new natural products and quickly identify the genes required for the production of a natural product. Characterization of these biosynthetic gene clusters allow us to identify the enzymes required for the biosynthesis and functionalization of these compounds and the enzymes that may be used for inactivation of toxic compounds. They also inform us on strategies that can be used to manipulate these organisms to produce derivatives of potentially useful natural products.

Among the most useful natural products are the phosphonic acids, which contain a carbon-phosphorus bond. Over the years, these small molecules have shown great success as drugs and herbicidal compounds, thus it is prudent to facilitate the rapid discovery and characterization of new phosphonate natural products. The work in this thesis focuses on the characterization of the enzymes needed for formation and functionalization of many different phosphonic acid biosynthetic pathways.

Non-heme iron-dependent oxygenases are important enzymes in all domains of life that catalyze various chemical reactions involving the functionalization of an un-activated carbon atom. These enzymes are present in many phosphonate biosynthetic

pathways. **Chapter 2** will focus on the *in vitro* biochemical characterization of the TauD-family dioxygenases that are involved in the biosynthesis of dehydrophos, a compound with proven antibiotic activity, fosfazinomycin, a compound containing a unique phosphono-hydrazide linkage, and hydroxyphosphonocystoximate, a phosphonate containing an unusual thiohydroximate moiety. Here, I show that these enzymes share little identity with one another, which drive the substrate specificity of these oxygenases.

Oxime-containing natural products are uncommon in nature, with even less being understood about how organisms make these functional groups. Phosphonocystoximic acid and its hydroxylated congener, hydroxyphosphonocystoximic, are two phosphonate natural products containing a free oxime moiety. **Chapter 3** will focus on the enzymes required for N-oxidation and oxime formation during the biosynthesis of these two natural products. We have identified an NADPH- and FAD-dependent enzyme responsible oxidation of a primary amine to the oxime that is catalyzed early in the biosynthesis of these natural products.

Phosphonothrixin is a potent herbicidal natural product that is produced by *Saccharothrix* sp. ST-888. Until recently, the biosynthetic gene cluster responsible for production of this compound remained elusive. I will describe *in vitro* characterization of three enzymes encoded within the phosphonothrixin biosynthetic gene cluster in **Chapter 4**. Additionally, the isolation of dihydroxypropylphosphonic acid from culture extracts of *Saccharothrix* sp. ST-888 have allowed us to formulate a hypothesis for the biosynthesis of the herbicidal compound.

Acknowledgements

First and foremost, I would like to thank Bill Metcalf for all his intellectual guidance and support over the years. Your continued patience and encouragement has helped me to develop into the independent scientist I am today. I would also like to thank many professors who have provided invaluable feedback on my research over the years. John Cronan, Brenda Wilson, Carin Vanderpool, James Imlay, and Wilfred van der Donk have been involved in critiquing my research proposals and have provided continued support as part of my thesis committee. Finally, I have to thank the many faculty members who are part of the Mining Microbial Genomes theme at the Carl. R. Woese Institute for Genomic Biology: Paul Jensen, John Gerlt, Doug Mitchell, Satish Nair, and Wilfred van der Donk. You have all provided exceptional feedback of my research over the years at our theme meetings. I am privileged to have the opportunity to work and collaborate with this knowledgeable group of natural product experts.

I would also like to thank my collaborators who I have had the privilege to work with over the years. First, thank you to Kou-San Ju. You welcomed me into the MMG lab in IGB, showed me the ropes during my first few years, and provided a constant source of intellectual guidance. Next, thank you to Joel Cioni and Katharina Pallitsch. Your contributions to the *N*-oxidase project have been invaluable. I would like to acknowledge Ryuichi Hirota for being an open and supportive collaborator while working on the phosphonothrixin project. Thank you to Nektaria Petronikolou, Shihui Dong, and Satish Nair for being great collaborators on the oxygenase paper.

My Ph.D. would have been much more stressful without the continued support and friendship of so many different people over the years. To my friends who have been with me since day one in graduate school: Nayab Abidi, Amruta Bhate, and Kirsten

Eckstrum. We have been on so many adventures over the years and I miss all our late-night coffee and pie dates. Alisa King, Dr. Nick Hess, Dr. Katie Hess, Dr. Melissa Ryerson, Dr. Daniel Ryerson, and Dr. Chelsea Loyd: you all have been an amazing group to watch movies with, play board games with, and conquer escape rooms with.

I cannot thank my IGB lab mates enough for providing a stimulating and always entertaining work environment over the years: Dr. Jason Bouvier, Dr. Michael Steven Carter, Dr. Brian San Francisco, Dr. Tyler Stack, Dr. Betsy Parkinson, Dr. Manuel Ortega, Dr. Courtney Cox, Dr. Joel Melby, Dr. Caitlin Deane, Dr. Brandon Burkhart, Adam Dicaprio, Andi Liu, Arash Firouzbakht, Dr. Zedu Huang, Dr. Despina Bougioukou, Abraham Wang, and Martin McLaughlin. I especially must thank those friends and lab mates who have supported me both in and out of the lab during my Ph.D.: Dr. Tucker Maxson, Dr. Kyle Dunbar, Dr. Evelyn Malloy, Dr. Spencer Peck, Dr. Jonathan Tietz, Dr. Emily Ulrich, Dr. Yelena Ilin, Dr. Christopher Schwalen, and Xiao Rui Guo. Our game nights, potlucks, and endless shenanigans will never be forgotten.

I have a wonderful group of friends who have supported me since we were in grade school together: Katrina Kalcic, Nelson Schneider, Steve Labedz, Michael Jonas, and Megan Nigh. Even with all the late work nights, missed phone calls, and last-minute change of plans, I can always count on you to still love and support me to this day. I look forward to our continued adventures over the years.

Finally, to my family: my mom and dad, Pat and Craig Goettge, as well as my brother, Curtis Goettge, sister-in-law, Cara Beck-Goettge, and nephew, Jacoby Goettge. Thank you for always supporting my interests and hobbies, even if I did not know what I wanted at the time. Your patience and love have supported me through the tough times

during my Ph.D. and I am eternally grateful for your endless support. This journey would not have been possible without you.

"You certainly usually find something, if you look, but it is not always quite the something you were after."

- J.R.R. Tolkien

Table of contents

Chapter 1: Introduction to phosphonic acid natural products

1.1 INTRODUCTION

1.1.1 Importance of natural products in medicine

Human lives have been influenced through the use of natural products for centuries. There is documentation from cultures around the world that date back as early as 2600 BC supporting the use of natural products to benefit human life (24, 54). From ancient herbal remedies to modern medicines, our lives would look vastly different without the influence of compounds isolated from natural sources. Currently, nearly two-thirds of drugs used therapeutically are inspired by natural products, with upwards of 80% of antibacterials being natural products or derivatives of natural compounds (102). Discovery of novel drugs with new targets is prudent as the number of antimicrobial-resistant pathogens is continuously rising, presenting a severe burden to healthcare around the world (95, 114, 134).

1.1.2 Natural product discovery in the genomics era

Development of new technologies for genome sequencing has driven the cost of a genome down and increased the quality of data received within the past decade (35). This has helped drive the exponentially growing number of sequences that have been deposited into GenBank (8). Genome sequencing of *Streptomyces coelicolor* A3(2) was instrumental to understanding actinobacterial potential for natural product production (9, 15). We now understand that most actinobacteria can produce over 25 different natural products each (15). With the growing body of knowledge about how different natural

product classes are biosynthesized, we can rapidly dereplicate and prioritize organisms for the potential to produce novel natural products.

1.2 PHOSPHONIC ACID NATURAL PRODUCTS

Phosphonic acids are reduced phosphorus small molecules containing direct carbon-phosphorus (C-P) bond, while phosphinic acids are defined as having a C-P-C or C-P-H bond. These molecules are known to be potent enzyme inhibitors due to molecular mimicry to phosphate esters, carboxylic acids, and tetrahedral intermediates in numerous metabolic pathways (93).

1.2.1 Phosphonic acids in agriculture and medicine

The agriculture industry would look vastly different without the success of phosphonates. Glyphosate, which goes by the trade names RoundUp®, Rodeo®, Accord®, Touchdown®, and Shackle®, is a widely used herbicide. This synthetic phosphonate targets enolpyruvylshikimate-3-phosphate (EPSP) synthase, which is part of the shikimate pathway of aromatic amino acid biosynthesis (127). A natural product phosphinate that has found renowned success as an herbicide is phosphinothricin, trademarked under the names Finale®, Basta®, and Liberty®. This compound mimics the tetrahedral intermediate of glutamine synthase (32). The last phosphonate that has shown success is ethephon. This compound is used as a ripening agent, as ethephon is metabolized into ethylene within the plant (12). Fig 1.1 shows the different phosphonate natural products used as herbicidal compounds.

Many phosphonate compounds have importance as medicines as well. The synthetic bisphosphonate, alendronic acid (trade named Fosamax®), is one of the most

widely prescribed drugs used to treat osteoporosis (90). Many different phosphonates are nucleotide mimics used as antiviral compounds, such as tenofovir (antiretroviral used for HIV treatment) (7, 30), adefovir (reverse transcriptase inhibitor used for hepatitis B treatment) (22, 23, 110), and cidofovir (DNA polymerase inhibitor of cytomegalovirus) (22, 142). Fosinopril is the only clinically-approved phosphonate for the use as an angiotensin converting enzyme (ACE) inhibitor to treat hypertension (40, 137). A small number of phosphonate natural products also have important uses clinically. Fosfomycin, known by the trade name Monurol®, targets cell wall biosynthesis of Gram negative pathogens and is widely used for the treatment of urinary tract infections in women (28, 29, 144, 149). Fosmidomycin on the other hand has shown promise in clinical trials for use in the treatment of malaria. This compound target DXP reductoisomerase involved in the non-mevalonate isoprenoid biosynthetic pathway of *Plasmodium* spp. (2, 76, 133). Fig 1.1 shows the different clinically used phosphonic acid natural products.

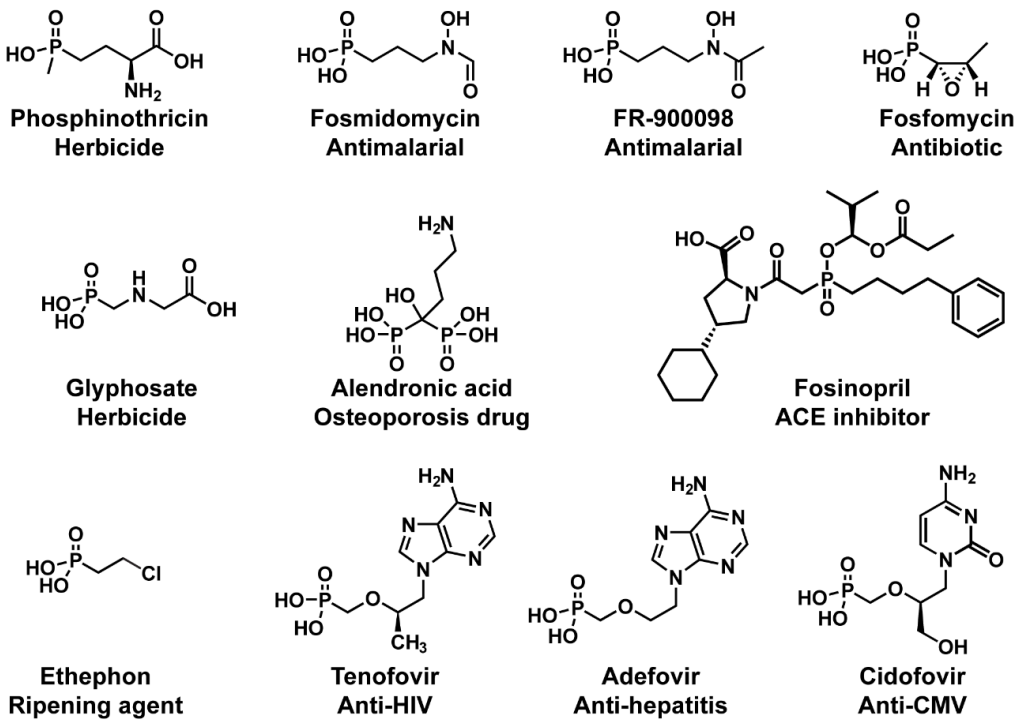


Figure 1.1. Phosphonic acids that are used clinically and agriculturally. The name of the phosphonate is listed below the structure, with the function of each compound listed below the name.

1.2.2 Genome mining for novel phosphonate natural products

Given the clinical and agricultural applicability of phosphonates, it is prudent that we continue to invest time and resources to discover novel phosphonate natural products with new targets and functions. Discovery of novel phosphonate natural products is facilitated by the fact that all known phosphonate biosynthetic pathways, except for the structurally similar ACE inhibitors K-4 (72), K-26 (143), and I5B2 (65), start with the same biotransformation catalyzed by a single enzyme, phosphoenolpyruvate phosphomutase (PepM) (14, 44, 67, 120). PepM funnels the central metabolic intermediate, phosphoenolpyruvate (PEP), into secondary metabolism through the rearrangement of PEP to phosphonopyruvate, which is the first biosynthetic step catalyzing carbon-phosphorus bond formation.

Using *pepM* as a genetic handle, we are able to mine the genomes of sequenced organisms to identify those with the potential to produce novel phosphonate compounds (58, 59). We can distinguish bona fide PepM proteins from the related isocitrate lyase proteins through identification of a conserved active site motif, EDKxxxxxNS (16, 146). This same study also discovered that the degree of similarity between PepM amino acid sequences predicts the divergence of the phosphonate gene clusters, thus correlating the probability of discovering a new, undescribed phosphonate natural product with the sequence identity of PepM (146). Using this knowledge, a large study was done to screen for actinobacteria that have the genetic machinery to produce new phosphonate natural products (59). Close to 10,000 actinobacteria were analyzed for the presence of the *pepM* gene. Those *pepM*$^{+}$ isolates were genome sequenced because the genes required for most bacterial natural products are clustered together on the genome (92, 109, 130). Thus, this quickly allowed for the identification of the putative phosphonate gene clusters in those organisms. Through this genome-mining effort, the number of phosphonate natural products reported in the literature nearly doubled, from 20 to 39 (59).

This chapter will describe the discovery of different phosphonic acid natural products that have identified gene clusters and discuss the current understanding of how each of these natural products are made. This chapter is organized into sections based on common biosynthetic intermediates during biosynthesis: 2-aminoethylphosphonate, 2-hydroxyethylphosphonate, phosphonoacetate, and dihydroxypropylphosphonic acid.

1.3 2AEP: THE FIRST DESCRIBED PHOSPHONATE AND A COMMON CELL-WALL ASSOCIATED COMPOUND

1.3.1 Physiological incorporation of 2AEP

The first identified phosphonate natural product, 2-aminoethylphosphonic acid (2AEP), was discovered in 1959 and was isolated from a sheep rumen protozoan, *Tetrahymena pyriformis* (46). This natural product is arguably the most common phosphonate on Earth, as many different bacteria (5, 53, 136), fungi (135), protozoa (26, 46, 64), and even invertebrates (61, 87, 88, 94, 112, 145) produce 2AEP and incorporate it into lipid head groups (70) as glycerophosphonolipids and sphingnolipids (53, 61, 136). Interestingly, 2AEP has been isolated from bovine livers (39), while rat hepatocytes readily incorporate 2AEP as 1,2-diacylglyceryl-aminoethylphosphonate. Fig 1.2 shows the different phosphonic acids that have been isolated over the years from the cell membranes of different organisms. The exact physiological role of phosphonolipids in different organisms remains unknown, although the most logical explanation involves the stability of the C-P bond over the more labile carbon-oxygen-phosphorus (C-O-P) bond. In densely populated communities where phospholipases may be abundantly produced and excreted, the stability of the C-P bond may give a producing organism a competitive advantage over those organisms that do not produce C-P bonds (62, 139).

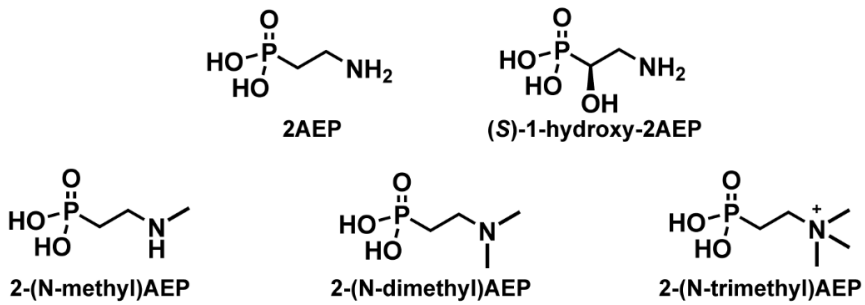


Figure 1.2. Phosphonic acids commonly found associated with the cell wall of different microorganisms.

1.3.2 The biosynthesis of 2AEP

The biosynthesis of 2AEP is carried out through three enzymatic transformations outlined in Fig 3. Prior to isolation of the enzymes required for this biotransformation, the Dunaway-Mariano group elucidated the biosynthesis of 2AEP by incubating radiolabeled substrates with *Tetrahymena pyriformis* extracts (4). Isolation and characterization of each individual enzyme over the years has confirmed these initial labeling studies. As is common with almost all phosphonate biosynthetic pathways, the isomerization of phosphoenolpyruvate (PEP) to phosphonopyruvate (PnPy) through the action of PEP mutase (PepM) is the first conserved biosynthetic step (48, 67, 100, 115, 119, 121). This highly unfavorable reaction is driven forward by phosphonopyruvate (PnPy) decarboxylase to form phosphonoacetaldehyde (PnAA) (55, 99, 101, 119, 150). The enzyme, 2AEP aminotransferase, then catalyzes the formation of 2AEP from PnAA (17, 25, 66, 83) (Fig 1.3).

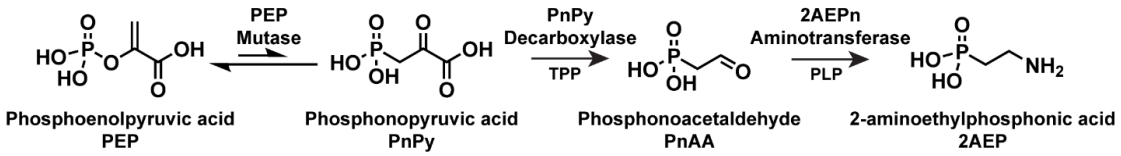


Figure 1.3. The biosynthesis of 2-aminoethylphosphonic acid involves three enzymatic transformations.

1.3.3 Beyond 2AEP as a structural phosphonic acid

As discussed at the beginning of this review, 2AEP and derivatives are common head groups incorporated into phosphonolipids and phosphonoglycans (53, 53, 61, 61, 64, 136, 145, 146). To date, 2AEP has only been characterized as part of modifications to the membranes of different organisms. This section will discuss the involvement of 2AEP in the biosynthesis of the newly described phosphonate, phosphonocystoximic acid and its hydroxylated congener along with the prospect of 2AEP being involved in the biosynthesis of the clinically relevant compound, fosmidomycin.

1.3.3.1 Phosphonocystoximic acid and hydroxynitrilaphos discovery

The phosphonic acids isolated from the culture broth of *Streptomyces regensis* NRRL WC-3744 were shown to be a collection of unique natural products. The first isolated cyanohydrin-containing phosphonate, hydroxynitrilaphos, was isolated from spent culture medium of *S. regensis* NRRL WC-3744 (19). In this same study, three other phosphonic acids were also isolated from the spent medium when the organism was grown on ISP4 medium: 2AEP, *N*-acetyl-2AEP and *N*-acetyl-1-hydroxy-2AEP. In a later study, hydroxyphosphonocystoximic acid was identified via high resolution mass spectrometry to be present in ISP4 extracts of *S. regensis* WC-3744 (59). Based on the compounds isolated from the spent media and the presence of particular genes within the

gene cluster, it is hypothesized that hydroxyphosphonocystoximic acid is biosynthesized in a manner similar to glucosinolates, a common plant natural product, starting from 2AEP (Fig 1.4).

1.3.3.2 Phosphonocystoximate biosynthesis may be reminiscent of glucosinolate biosynthesis

Glucosinolates are small, amino acid derived natural products, most commonly produced by plants (33, 37). The biosynthesis of these small molecules begins with the oxidative decarboxylation of an amino acid, catalyzed by a cytochrome P450-dependent *N*-oxygenase to yield an aldoxime intermediate (33). The hydroxyphosphonocystoximic acid gene cluster from *S. regensis* NRRL WC-3744 contains the conserved three-gene set to produce 2AEP: a PEP mutase, a phosphonopyruvate decarboxylase, and a 2AEP aminotransferase (19, 34). 2AEP could serve as an analogous compound similar to an amino acid. As discussed in **Chapter 3**, I show the subsequent FAD- and NADPH-dependent oxidation of 2AEP to yield the aldoxime, 2-hydroxyiminophosphonic acid (34), which would reflect the aldoxime formed in glucosinolate biosynthesis. In glucosinolate biosynthesis, the aldoxime is further oxidized through the action of a separate cytochrome P450-dependent enzyme to form an unstable *aci*-nitro or nitrile oxide that can react with a sulfur donor such as cysteine or glutathione(38). This S-alkylthiohydroximate structure would then be reminiscent of hydroxyphosphonocystoximic acid (Fig 1.4). *Streptomyces* do not produce glutathione (56), therefore an alternative pathway to S-alkylation needs to be proposed. Mycothiol-conjugated hydroxyphosphonocystoximic acid has been observed by high-resolution mass spectrometry, indicating that mycothiol is the likely *N*-acetylcysteine donor (59).

In contrast, cyanogenic glucosides are natural products also produced by plants in a similar fashion to glucosinolates (122). However, formation of an aci-nitro used as an acceptor for a S-alkylation, involves a cytochrome P450-dependent enzyme catalyzes oxidation of the aldoxime to the nitrile (122). This molecule would be reminiscent of hydroxynitrilaphos (Fig 1.4). There are no cytochrome P450-dependent enzymes encoded in the *S. regensis* NRRL WC-3744 phosphonate gene cluster; however, there are two oxidoreductases that have no assigned function (Fig 1.4) that could be responsible for the oxidation of 2-hydroximinophosphonic acid to yield either a *aci*-nitroethylphosphonic acid (glucosinolate biosynthesis analog) or hydroxynitrilaphos (cyanogenic glucoside biosynthesis analog) (59). The presence of hydroxynitrilaphos in culture extracts suggests that the natural product biosynthetic pathway may resemble cyanogenic glucoside biosynthesis. The presence of ATP-grasp homologs within the gene cluster further suggest that hydroxynitrilaphos may not be the final natural product, but rather a degradation product of a larger peptide (59).

1.3.3.3 Differences in the phosphonocystoximate and hydroxynitrilaphos gene cluster

Streptomyces sp. NRRL S-481 contains a similar gene cluster to *S. regensis* WC-3744(34, 59), thus it is not surprising that the natural product produced by this organism is similar to hydroxynitrilaphos produced by *S. regensis*. The *Streptomyces* sp. NRRL S-481 gene cluster is lacking a phosphonopyruvate decarboxylase homolog, therefore it is unknown how *Streptomyces* sp. NRRL S-481 produces 2AEP as an intermediate for oxime-formation. This organism is also lacking a non-heme iron and α-ketoglutarate-dependent dioxygenase homolog, which is responsible for the hydroxylation of the carbon alpha to the phosphorus atom in hydroxyphosphonocystoximate biosynthesis (34).

In this regard, even though it is unknown how *Streptomyces* sp. NRRL S-481 produces 2AEP, it is speculated that the biosynthesis will proceed in a similar manner to the pathway proposed in *S. regensis* NRRL WC-3744.

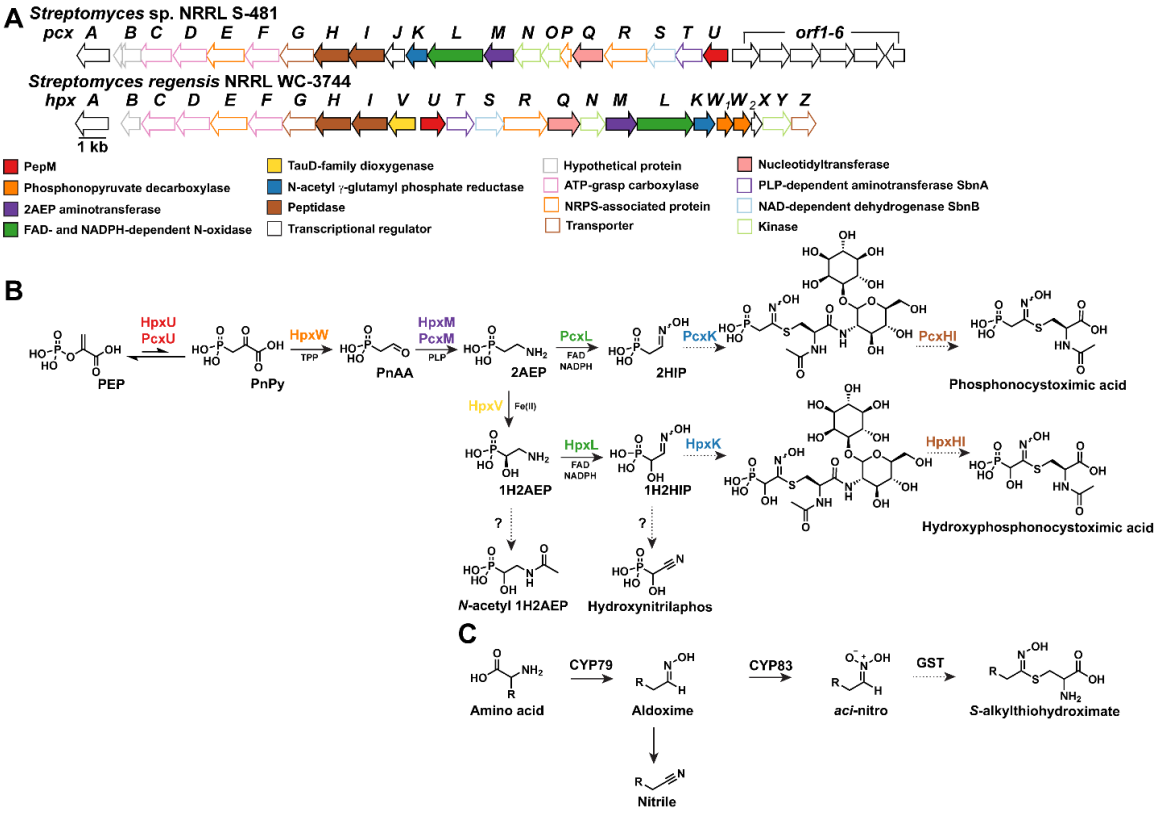


Figure 1.4. (A) Gene clusters and putative function for each protein encoded by each putative ORF from *Streptomyces* sp. NRRL S-481 for phosphonocystoximate biosynthesis and *Streptomyces regensis* NRRL WC-3744 for hydroxyphosphonocystoximate and hydroxynitrilaphos biosynthesis. (B) Putative biosynthetic pathways for phosphonocystoximate (Pcx proteins) and hydroxyphosphonocystoximate (Hpx proteins). (C) Partial biosynthetic pathway of glucosinolates ending with the S-alkylthiohydroximate intermediate resembling the phosphonocystoximic acids.

1.3.4 Fosmidomycin is a potent phosphonic acid with an unknown biosynthetic pathway

1.3.4.1 Fosmidomycin discovery and bioactivity

Fosmidomycin, first named FR-31564, was originally isolated from the fermentation broth of *Streptomyces lavendulae* No. 8006 (ATCC 31279) (52, 78, 106). A related, dehydrated compound, was isolated from the same organism and called FR-32863 (52, 78, 106). Initial studies comparing the bioactivity of fosmidomycin and fosfomycin found that fosmidomycin showed bioactivity against a wide variety of Gram-negative pathogens and only some Gram-positive organisms (97). Fosmidomycin showed equal to or greater antimicrobial activity against many different pathogens when compared to the bioactivity of gentamycin and fosfomycin, while *in vivo* efficacy of fosmidomycin was equal to fosfomycin for clearance of skin infections (97).

The original studies examining fosmidomycin mode of action noted that resistance only arose through transporter mutants, rather than physical inactivation of the molecule (73). This was different than the inactivation of fosfomycin, where phosphorylation is responsible for resistance to the antibiotic (71). However, studies investigating the mode of action of fosmidomycin discovered that the molecule inhibited carotenoid and menaquinone biosynthesis, suggesting the isoprenoid biosynthesis as the specific target (125). Indeed, fosmidomycin is a potent inhibitor of 1-deoxy-D-xylulose 5-phosphate (DXP) reductoisomerase, a key step in the non-mevalonate isoprenoid biosynthetic pathway of both bacteria (82) and plants (57). The discovery that *Plasmodium falciparum* also uses the non-mevalonate isoprenoid biosynthetic pathway was critical for launching fosmidomycin into development as an antimalarial drug (57,

86). This process was accelerated because previous Phase I clinical trials had already demonstrated tolerance for fosmidomycin in humans as a potential antibiotic (75).

1.3.4.2 Biosynthetic gene cluster of S. lavendulae No. 8006

Even though fosmidomycin has reached Phase II human clinical trials for the treatment of malaria and different Gram-negative infections, how *S. lavendulae* No. 8006 produces this bioactive molecule remains elusive. The putative phosphonate biosynthetic gene cluster was recently identified upon sequencing of the *S. lavendulae* No. 8006 genome (59). The genes within the genomic neighborhood of the PEP mutase suggest that the biosynthetic pathway may proceed through the biosynthesis of 2AEP (Fig 1.3), due to the presence of a PnPy decarboxylase and a 2AEP aminotransferase homolog (Fig 1.5). Beyond the presence of these homologs within the genomic neighborhood, little is known about the route to fosmidomycin in *S. lavendulae* No. 8006.

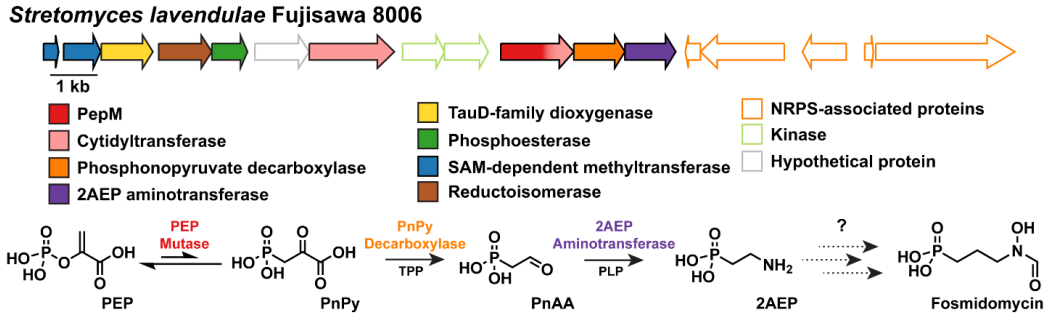


Figure 1.5. Putative fosmidomycin biosynthetic gene cluster. The gene cluster contains homologs of the genes required for production of 2AEP, although how *S. lavendulae* produces fosmidomycin from 2AEP remains elusive.

1.4 2HEP, AN INTERMEDIATE COMMON IN PHOSPHONOPEPTIDE BIOSYNTHETIC PATHWAYS

The first phosphonate biosynthetic pathways investigated after the discovery of 2AEP did not include this small molecule as a biosynthetic intermediate, but rather a

different two-carbon molecule that is hydroxylated opposed to aminated. This small molecule is derived from phosphoenolpyruvate in three enzymatic steps (Fig 1.6). PEP mutase and phosphonopyruvate decarboxylase catalyze the transformation of phosphoenolpyruvate to phosphonoacetaldehyde. A type-III iron-dependent alcohol dehydrogenase, PhpC, was originally identified in phosphinothricin biosynthesis to catalyze the reduction of phosphonoacetaldehyde to 2-hydroxyethylphosphonic acid (123, 124). Homologs of this enzyme were further identified as part of phosphonoglycan biosynthesis (147), fosfomycin biosynthesis (124, 141), dehydrophos biosynthesis (21), and phosphinothricin biosynthesis (124).

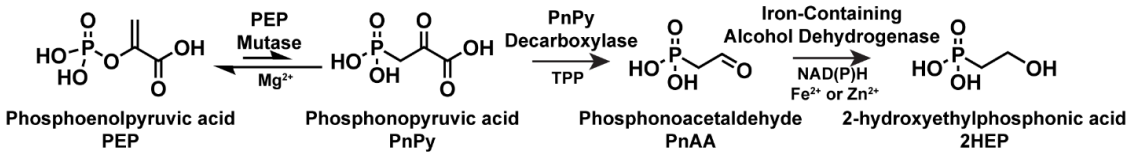


Figure 1.6. The biosynthesis of 2-hydroxyethylphosphonic acid involves three enzymatic transformations.

1.4.1 Phosphonoglycans

As described in the previous section about 2AEP biosynthesis, phosphonic acids are common modifications for lipids, proteins, and glycans. Most of these modifications involve the use of 2AEP as the phosphonic acid. In this section, I will discuss the discovery and partial biosynthetic pathway involved in production of phosphonoglycans using 2HEP as the phosphonic acid.

1.4.1.1 Production of phosphonoglycans in Glycomyces

Glycomyces sp. strain NRRL B-16210 was shown to produce phosphonic acids at a high levels through analysis by ^{31}P NMR spectroscopy (147). Degradation of the cell

walls using lysozyme increased the levels of phosphonates detected, indicating that the phosphonic acid was associated with the cell wall. After partial acid hydrolysis of the phosphonates isolated from the cells, both glycerol-attached (2-HEP mono(2,3-dihydroxypropyl)ester) and hexose-linked (likely 2-HEP galactose) 2HEP derivatives were detected using ^{31}P NMR spectroscopy and MS-MS fragmentation analysis of the compounds (147). The isolation of both glycerol-linked and hexose-linked phosphonic acids from one cell, suggested that these compounds may be part of a larger branched poly-teichoic acid cell wall present in different Gram-positive organisms, yet further experimentation is needed to demonstrate this (147).

1.4.1.2 Phosphonoglycan biosynthetic pathway

The gene cluster for phosphonoglycan biosynthesis contains PepM and PnPy decarboxylase homologs, which would have the ability to generate PnAA. In an operon-like organization with *ppm* and *ppd*, the gene cluster contains an iron-dependent alcohol dehydrogenase homolog that should form 2HEP. These are the only three genes that have homologs in other phosphonate biosynthetic pathways, thus seeing the production of 2HEP is unsurprising. In addition to this three-gene cassette, within the genomic neighborhood, there are genes homologous to those involved in glycan biosynthesis and teichoic acid biosynthesis. Therefore, the compounds isolated from *Glycomyces* sp. strain NRRL B-16210 likely are modifications to the cell membrane as phosphonoglycans (147).

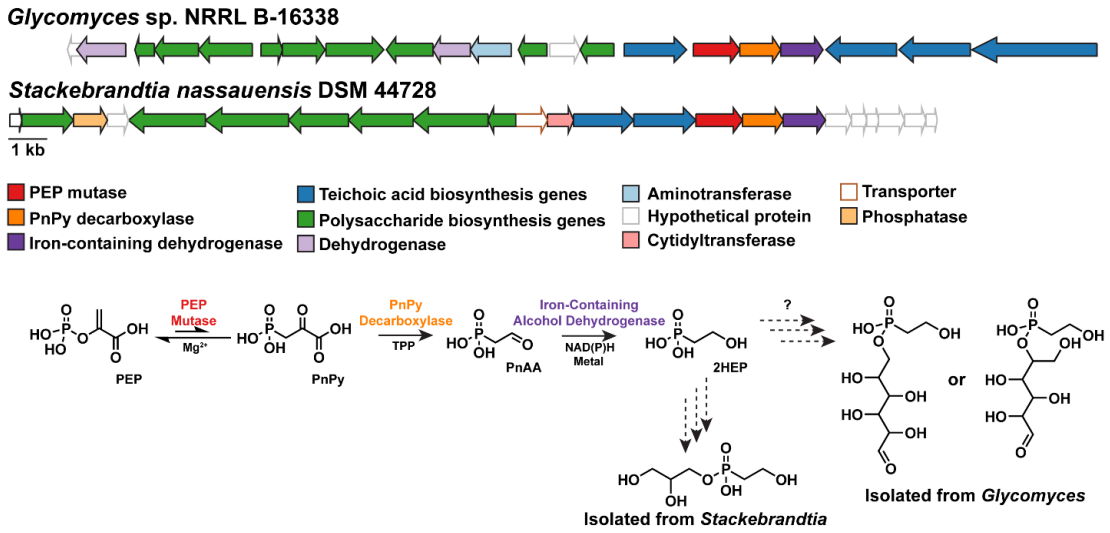


Figure 1.7. The phosphonate biosynthetic gene clusters from *Glycomyces* sp. NRRL B-16338 and *Stackebrandtia nassuensis* DSM 44728 and the natural products isolated from each strain.

1.4.2 Fosfomycin: a phosphonate so important that nature invented two biosynthetic routes to the same product?

1.4.2.1 Fosfomycin discovery and mode of action

With regards to treating disease, fosfomycin is arguably the most important natural product discovered to date. Trademarked under the name Monuril® and Monurol®, this natural product is commonly used to treat urinary tract infections in the United States (28, 29, 113, 149). Fosfomycin targets cell wall biosynthesis in both Gram-positive and Gram-negative organisms though the inhibition of phosphoenolpyruvate synthase (60, 96).

This natural product was originally discovered in 1969 to be produced by *Streptomyces wedmorensis* and *Streptomyces fradiae* (129) and showed broad spectrum antimicrobial activity (41). Seventeen years later, fosfomycin was discovered again in a search for broad spectrum antibiotics in *Pseudomonas syringae* (126). Genome

sequencing of *P. syringae* PB-5123 in 2012 revealed that this strain did not contain the same phosphonate-specific genetic machinery as used by *Streptomyces fradiae* to produce fosfomycin (68, 141), yet both strains have been shown to produce the identical final product. Investigations into this natural product resulted in a nearly complete understanding of the biosynthesis of fosfomycin in *Streptomyces*, yet the biosynthetic route to fosfomycin awaits elucidation in Pseudomonads. I will outline the biosynthetic routes to the production of this clinically used phosphonate in both organisms.

1.4.2.2 Biosynthesis of fosfomycin in Streptomyces

The fosfomycin biosynthetic pathway in *Streptomyces* has nearly been pieced together to completion, with only one enzymatic transformation likely needing to be described. The fosfomycin gene cluster in *Streptomyces* consists of eight genes, six of which are likely involved in biosynthesis (18, 42, 43, 45, 68, 141), with the two remaining proteins, FomA and FomB, providing self-resistance to fosfomycin (71). The first committed step, catalyzed by the phosphoenolpyruvate mutase, Fom1, generates phosphonopyruvate (43), which can be decarboxylated to form phosphonoacetaldehyde (PnAA) using Fom2, the phosphonopyruvate decarboxylase homolog (42). FomC catalyzes the reduction of PnAA to 2-hydroxyethylphosphonic acid (2HEP) (124). For many years, the function of a methylcobalamin-dependent radical-SAM-dependent methyltransferase, Fom3, remained elusive (1, 80) , though it was predicted to form 2-hydroxypropylphosphonic acid (2-HPP) by methylating 2HEP.

Recently, it has been shown that Fom3 methylates cytidylyl-2-hydroxyethylphosphonate (CMP-2HEP) at a much higher rate than 2HEP reported by Allen and colleagues (116). This reaction produces an equal ratio of cytidylyl-2-

hydroxypropylphosphonate (CMP-2HPP) diastereomers, suggesting that there is no stereoselectivity during the methylation of CMP-2HEP (116). Recently, Cho and colleagues have shown that like many other phosphonate biosynthetic pathways, fosfomycin biosynthesis proceeds through the cytidylation of an intermediate in the pathway (18). Fom1, the phosphoenolpyruvate phosphomutase homolog, has an extra *N*-terminal domain that shares homology to a cytidylyltransferase. When this gene is knocked out, fosfomycin production is recovered when mutants are supplemented with 2HPP, but are unable to be rescued when the mutants are supplemented with 2HEP, suggesting that this protein functions downstream of 2HEP production (18). This would suggest that Fom1 has two functions within fosfomycin biosynthesis: (1) as the carbon-phosphorus bond forming enzyme (43) and (2) as a cytidylyltransferase acting on 2HEP (18). During *in vitro* analysis of a truncated Fom1 protein containing only the cytidylyltransferase domain, Cho and colleagues discovered that the N-terminal domain of Fom1 does indeed catalyze the cytidylation of 2HEP in a cytidine triphosphate (CTP)-dependent and cobalt(II)-dependent fashion (18).

Fom4, also known as HppE, is a non-heme, ferrous iron-dependent, alpha-ketoglutarate independent oxygenase and peroxidase that catalyzes epoxidation of 2HPP (45, 91, 148). Depending on the substrate bound within the active site, HppE can either perform epoxidation of (S)-2HPP or oxidation of (R)-2HPP, yielding 2-oxopropylphosphonate (2OPP) (107).

This leaves one biosynthetic enzyme to finalize the fosfomycin biosynthetic pathway. It is still unknown how CMP-2HPP will lose its CMP moiety, however, the remaining gene (*fomD*) within the fosfomycin biosynthetic cluster likely will catalyze

this transformation. FomD is homologous to the phosphatase, SA1684, from *Staphylococcus aureus*. This enzyme cleaves nucleotide diphosphoates, thus, it is predicted that FomD may hydrolyze the CMP-2HPP bond, yielding a racemic mixture of 2HPP (116).

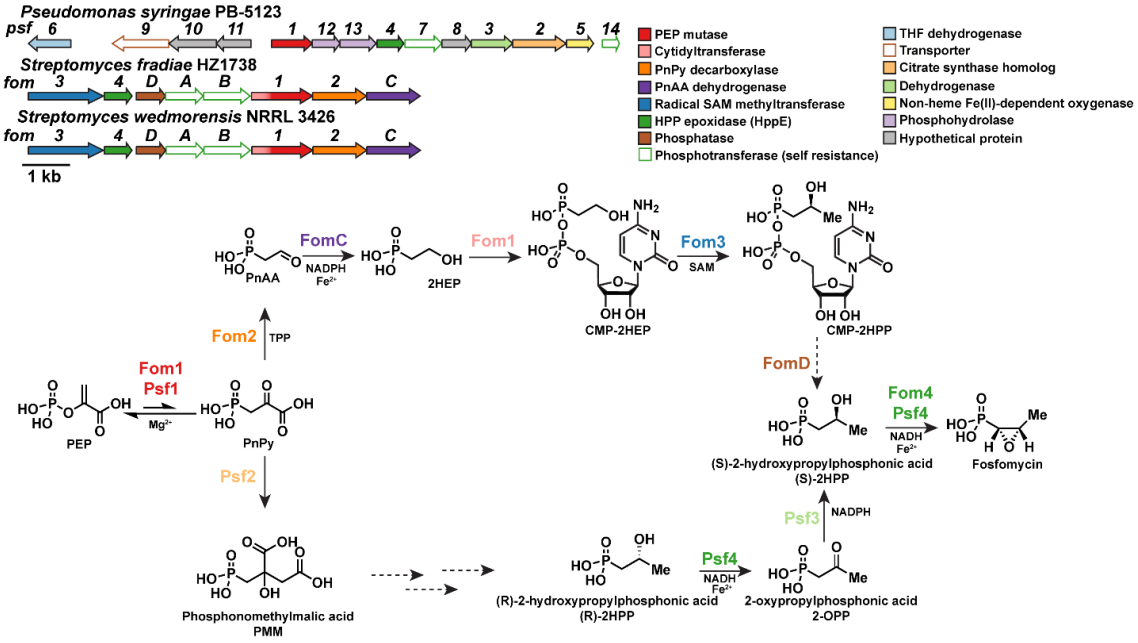


Figure 1.8. There have been two different pathways discovered to date for fosfomycin biosynthesis. (A) The biosynthetic gene cluster from Streptomyces (*fom* genes) involves the biosynthesis of 2HEP, while the biosynthetic gene cluster in Pseudomonas (*psf* genes) are shown. Homologous enzymes are highlighted with the same color. (B) The biosynthetic routes to fosfomycin. Solid arrows denote biosynthetic steps with confirmed activity, while dashed arrows denote proposed biosynthetic steps.

1.4.2.3 Biosynthesis of fosfomycin in Pseudomonas *shares three homologous enzymes with* Streptomyces

The *Pseudomonas* fosfomycin biosynthetic pathway is significantly less characterized. The fosfomycin gene cluster in *Pseudomonas* was assigned based on the heterologous expression of a fosmid containing a 22.5 kb insert with putative fosfomycin genes into a *Pseudomonas* strain unable to make phosphonates (68). Interestingly, only

three genes were identified to be homologous to the enzymes in *S. wedmorensis* and *S. fradiae*. Psf1 was identified and labeled as the PEP mutase homolog, while *psf7* encodes a kinase homologous to *fomA* that is responsible for self-resistance through phosphorylation of fosfomycin (71). The final enzymatic step of fosfomycin biosynthesis, carried out by Psf4, is also shared with the *Streptomyces* pathway. Psf4 is a homolog of HppE and is responsible for the epoxidation of (*S*)-2-hydroxypropylphosphonic acid (2HPP) to yield fosfomycin. This reaction has been reconstituted *in vitro* (81, 98).

1.4.2.4 Biosynthesis of fosfomycin in Pseudomonas *shares homologous enzymes with other phosphonate biosynthetic pathways*

Unlike the *Streptomyces* pathway, *P. syringae* does not contain a phosphonopyruvate decarboxylase homolog to drive the highly unfavorable PEP mutase reaction forward through decarboxylation (55, 99, 119, 150). However, this pathway encodes a gene to produce Psf2, a citrate synthase-like protein homologous to FrbC in FR-900098 biosynthesis (27) and Pms in phosphinothricin biosynthesis (11, 118), which is responsible for the acetylation of phosphonopyruvate or phosphinopryruvate to form phosphonomethylmalic (PMM) acid or phosphinomethylmalic acid. Indeed, when Psf1 and Psf2 are incubated with PEP in the presence of acetyl-CoA, PMM is the observed product (68).

When Psf3, Psf5, Psf6, and Psf7 were heterologously expressed and activity was tested using PMM or 2HPP as substrates *in vitro*, no product was observed in any of these reactions, suggesting that these are not the native substrates of the enzymes or that a specialized cofactor or binding protein may be missing for the reactions (68). However, more recent studies have identified a function for Psf3 and have confirmed an additional,

hypothesized function for Psf4 (107). During the biosynthesis of fosfomycin in pseudomonads, 2-HPP was hypothesized to be derived from 2-oxopropylphosphonate (2-OPP) through the reduction of 2-OPP by one of the two encoded dehydrogenases within the cluster, Psf3 or Psf6. Psf3 was found to be an NADPH-dependent reductase that forms (S)-2-HPP using 2-OPP as a substrate (107). When using (S)-2-HPP as the substrate in the presence of $NADP^+$, but not (R)-2-HPP, 2-OPP was observed in the reaction mixture, confirming that (S)-2HPP is the product formed by Psf3. As discussed above, even though fosfomycin biosynthesis is not expected to proceed through the intermediate 2-OPP in *Streptomyces*, HppE has the ability to oxidize (R)-2-HPP to 2-OPP (148). During in vitro activities, Psf4, a homolog of HppE, was shown to have identical activities, thus, when a reaction mixture containing both Psf3 and Psf4 were incubated with a racemic mixture of 2-HPP, all of the product was converted to fosfomycin (107). The biosynthetic steps that can convert PMM to 2-HPP have yet to be identified.

1.4.3 Dehydrophos, a compound with an elusive structure and unusual biosynthetic pathway

1.4.3.1 Determination of the structure of dehydrophos

Dehydrophos is a broad spectrum antimicrobial tripeptide isolated from the fermentation extracts of *Streptomyces luridus* NRRL 15101 originally discovered and by Eli Lily and Company in 1984 under the name A53868 (3). The original proposed structure, glycyl-leucyl-2-amino-2-propenyl-phosphonate, was reassigned in 1988 to A53868A, again by Eli Lily and Company (51) (Fig 1.9). In 2007, Whitteck and colleagues synthesized A53868A to confirm the structure of this compound and were surprised to find that it is not the compound produced by *S. luridus* NRRL 15101. The

new structure of A53868A was reassigned for a third time to dehydrophos (Fig 1.9), revealing the structure as an interesting O-methylated vinylphosphonic acid coupled to leucine and glycine (140).

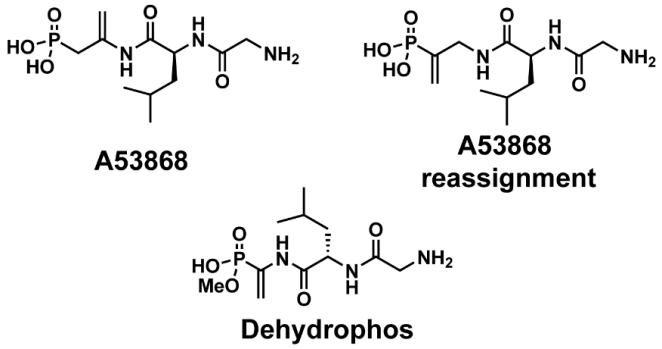


Figure 1.9. Original assignments of A53868 by Eli Lily and Company and final structure of dehydrophos.

1.4.3.2 Antimicrobial activity of dehydrophos

The broad-spectrum antimicrobial activity was originally reported by Johnson and colleagues at Eli Lily and Company and initial *in vivo* studies were done to show clearance of *Salmonella typhimurium* infected chickens (3). Later studies identified the importance of both the vinyl group and methylester on the bioactivity of dehydrophos (74), while *Salmonella enterica* knockout studies further investigated the mode-of-action of dehydrophos (21). Like many other Trojan horse antibiotics, the leucine-glycine dipeptide of dehydrophos is cleaved after import into the cell via the Opp peptide transporter. The peptidases, PepA and PepB, appear to be the major contributors to dehydrophos cleavage; however, PepD and PepN can supplement PepA and PepB mutants to maintain dehydrophos susceptibility. Upon *in vitro* cleavage of dehydrophos by PepA, methylacetylphosphonate was identified as a product of the reaction (21). This

molecule is a pyruvate mimic that has potent inhibitory effects on pyruvate dehydrogenase (103) and 1-deoxy-D-xylulose 5-phosphate (DXP) synthase (128).

1.4.3.3 Biosynthesis of dehydrophos involves several unusual enzymatic steps

The biosynthesis of dehydrophos is an interesting tale of unusual enzymology. The minimal 16 gene biosynthetic cluster for dehydrophos production was identified in 2011 by Circello and colleagues using deletion mutants of a dehydrophos producing fosmid clone (21). The cluster includes genes homologous to those used to make 2HEPn (Fig 1.10): PepM (DhpE), phosphonopyruvate decarboxylase (DhpF), and a type-III iron-dependent alcohol dehydrogenase (DhpG) (13, 21).

An Fe(II)-dependent and α-ketoglutarate-dependent dioxygenase (DhpA) was found to hydroxylate the carbon alpha to the phosphorus atom of 2HEPn, yielding dihydroxyethylphosphonic acid (DHEP). It was proposed that the next three enzymatic steps of dehydrophos biosynthesis are related to serine biosynthesis in *Escherichia coli* (Fig 1.10) (21). DHEP is a molecular mimic of glycerate, and DhpB was a predicted glycerate kinase homolog that was shown to phosphorylate DHEP on the distal hydroxyl group, similar to how glycerate kinase phosphorylates the hydroxyl group on the second carbon. DhpC, a phosphoglycerate dehydrogenase homolog was shown to oxidize 1-hydroxy-2-phosphorylethylphosphonate (HP-EP) to 1-oxo-2-phosphorylethylphosphonate (OP-EP) (20).

DhpH is a two-domain protein, containing an aminotransferase domain (DhpH-N) and an oxidase domain (DhpH-C). It was initially proposed that DhpH-N would form amino-vinyl-phosphonic acid (ΔAla(P)), and this unstable intermediate would be

channeled to the active site of DhpH-C to catalyze the leuciyl-tRNA-dependent peptidyltransferase activity to attach leucine. However, when full length DhpH or DhpH-N was incubated with pSer(P) and different keto-acids, OP-EP was not observed in the reaction, rather, acetylphosphonate was the sole product observed (13). Contrastingly, in 2010, Circello and colleagues hypothesized that the PLP-dependent aspartate aminotransferase homolog, DhpD, would catalyze the aminotransferase reaction needed to form 1-amino-2-phosphorylethyl phosphonate (pSer(P)); however, in 2014, Bougioukou and colleagues showed that DhpD does not catalyze this transamination, rather DhpD catalyzes the transamination of acetylphosphonate to L-alanine-phosphonate (L-alaP) (21).

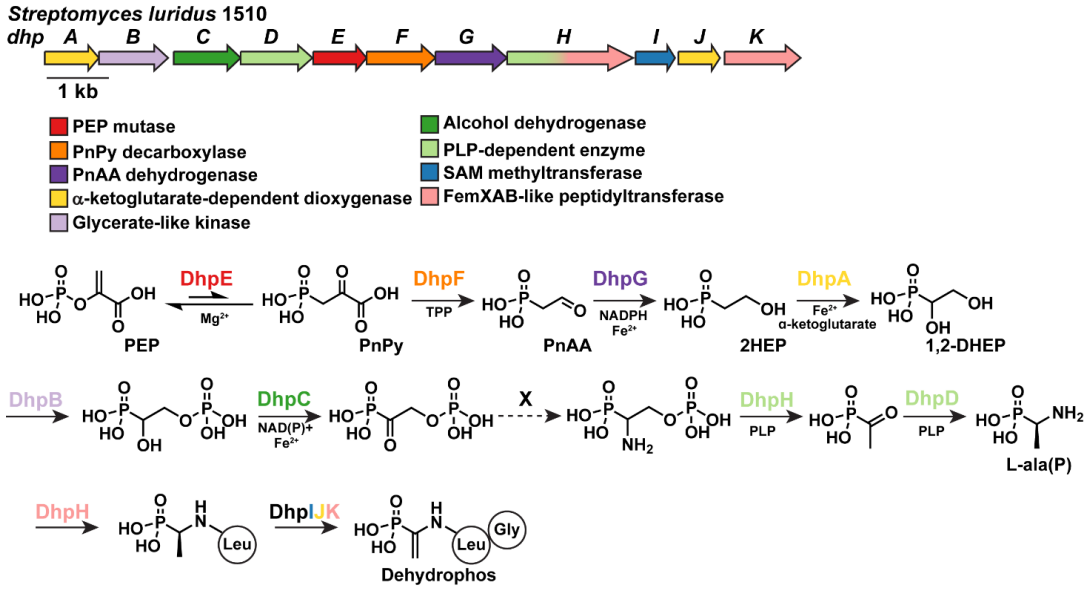


Figure 1.10. Dehydrophos biosynthetic gene cluster from *Streptomyces luridus* 1510. Current proposed biosynthetic route to dehydrophos production.

1.4.4 Phosphinothricin is a commercially available herbicide

Phosphinothricin, also known as glufosinate, is an herbicide trademarked under the names Finale®, Basta®, and Liberty®. This compound mimics the tetrahedral intermediate of glutamine synthase (32) and is a potent herbicide due to inhibition of this critical enzyme required for nitrogen assimilation. Many crops have been engineered to contain the phosphinothricin acetyltransferase (*pat*) gene, which confers self-resistance of the producing organism to phosphinothricin. This means that farmers can grow crops that are resistant to phosphinothricin, however, the invading weeds, which are not resistant to this potent herbicide are killed upon treatment with phosphinothricin.

1.4.4.1 Phosphinothricin: discovery of the different compounds with the same chemical warhead

The biosynthetic logic (93, 117) and characterization of the unusual enzymology (111) of phosphinothricin biosynthesis have been covered in three reviews. Here, I will cover advances in the biosynthesis of phosphinothricin that were not covered in the 2009 or 2013 reviews, including the sequencing and bioinformatic characterization of the bialaphos and phosalacine gene clusters (10), as well as the discovery and characterization of the desmethylphosphinothricin gene cluster (59).

Phosphinothricin is the only described phosphinic acid natural product that contains a carbon-phosphorus-carbon bond, distinguishing it from the more common phosphonic acids, which have only one carbon-phosphorus bond. Phosphinothricin is commonly attached to amino acids to form a tripeptide that has been isolated from bacteria. A phosphinothricin tripeptide, which has phosphinothricin coupled to two L-

alanine amino acids, was originally isolated from spent culture broth of *Streptomyces viridochromogenes* Tü494 (6) and *Streptomyces hygroscopicus* as SF-1293, later to be renamed bialaphos (104). In 1984, phosalacine, in which phosphinothricin is coupled to alanine and leucine, was isolated from culture extracts from *Kitasatospora phosalacinea* KA-338 (108). Seven years later, the compound named trialaphos was isolated from *S. hygroscopicus* KSB-1285, was isolated as a tetrapeptide of phosphinothricin attached to three molecules of L-alanine (63). More recently, a compound has been isolated from *Nonomureae* sp. NRRL B-24552 that looks like phosphinothricin, but lacks amino acids coupled to the molecule. This natural product also has a hydrogen-phosphorus-carbon bond instead of the second carbon-phosphorus bond found in phosphinothricin. For these reasons, this newly isolated compound was called Desmethylphosphinothricin (59).

1.4.4.2 Phosphinothricin biosynthesis involves nearly identical biosynthetic pathways for all producing organisms

DNA sequencing of assorted plasmids helped determine the genes needed for phosphinothricin biosynthesis (11); however, complete sequencing of the *S. viridochromogenes* Tü494 bialaphos biosynthetic gene cluster was not completed until 2005 (10, 118). Bioinformatic comparisons of the sequenced bialaphos (*S. viridochromogenes* Tü494 and *S. hygroscopicus* as SF-1293), phosalacine (*K. phosalacinea* KA-338), and desmethylphosphinothricin (*Nonomureae* sp. NRRL B-24552) biosynthetic gene clusters revealed a common gene set that is involved in desmethylphosphinothricin biosynthesis, an intermediate involved in the biosynthesis of phosphinothricin (10, 59). Unlike the biosynthetic gene clusters in *S. viridochromogenes* Tü494 and *S. hygroscopicus* as SF-1293, the *Nonomureae* sp. NRRL B-24552 gene

cluster does not contain non-ribosomal peptide synthase genes that would be responsible for coupling phosphinothricin to other amino acids, nor does this gene cluster contain a homolog of PhpK, the radical-SAM-dependent methyltransferase responsible for methylation of desmethylphosphinothricin to form the second C-P bond (47, 59, 138). Therefore, it is likely that desmethylphosphinothricin may be the final product of this gene cluster (59).

1.5 PHOSPHONOACETATE

1.5.1 Fosfazinomycin: convergence of two small molecules to form one interesting phosphonic acid

Fosfazinomycin is an unusual phosphonate natural product containing a phosphorus-hydrazide linkage. Compounds A and B, which differ in structure due to the presence (fosfazinomycin A) or absence (fosfazinomycin B) of valine, were originally isolated from the spent culture broth of *Streptomyces lavendofoliae* 630 (105). Antifungal activity, including hyphal swelling and branching abnormalities, was reported for fosfazinomycin (36).

1.5.1.1 Discovery of the fosfazinomycin gene cluster

The biosynthetic gene cluster for fosfazinomycin was identified using enzymatic modification of metabolites produced by *Streptomyces* sp. WM6372 using DhpI (31), the substrate-tolerable radical S- adenosylmethionine (SAM)-dependent O-methyltransferase that is responsible for methylester formation in dehydrophos (85). To do this, partially-purified extracts were treated with DhpI, SAM, or CD_3-labeled SAM. The chemical shifts of the two metabolites produced by *Streptomyces* sp. WM6372 changed after DhpI

treatment, suggesting modification of these molecules. The structure of these molecules were subsequently determined to be methyl phosphonoacetate (Me-PnA) and methyl 1-hydroxyphosphonoacetate (Me-HPnA) (31). Upon recognition that Me-HPnA is a component of fosfazinomycin, extracts of *Streptomyces* sp. WM6372 were monitored using high resolution mass spectrometry for the presence of fosfazinomycin. Confirmation of the presence of fosfazinomycin in the extracts of *Streptomyces* sp. WM6372 allowed for the de-orphaning of the gene cluster responsible for fosfazinomycin production in not only *Streptomyces* sp. WM6372, but also in *Streptomyces* sp. XY332 (31) (Fig 1.11).

1.5.1.2 Formation of fosfazinomycin requires convergence of two biosynthetic pathways

Work by Huang and colleagues have determined that the biosynthesis of fosfazinomycin is different from most other phosphonic acids and involves the convergence of two different biosynthetic pathways to form one bioactive molecule (50). The first involves the biosynthesis of Me-HPnA, reminiscent of most phosphonate biosynthetic pathways (50). The second pathway involves the formation of a methylated hydrazide-containing arginine molecule (49, 50). The mechanism behind the linkage of these two molecules remains unsolved. I will discuss the formation of these two molecules in the following sections.

1.5.1.3 Biosynthesis of methyl 1-hydroxyphosphonoacetate

Elucidation of the biosynthetic pathway for Me-HPnA formation, starting from phosphonoacetaldehyde (PnAA), was completed through *in vitro* enzyme analysis. The

fosfazinomycin gene cluster contains nine genes in an operon-like organization (FzmA-I) that appears to be used to generate Me-HPnA (146). However, it appears as though only four of those genes are required for the formation of Me-HPnA. This cluster contains PepM and phosphonopyruvate decarboxylase homologs, which can generate phosphonoacetaldehyde. It is hypothesized that FzmG, a non-heme Fe(II)-dependent and 2-oxoglutarate-dependent dioxygenase, may have dual-functions during the biosynthesis of fosfazinomycin (Fig 1.11). It was shown that FzmG will oxidize PnAA to PnA when incubated with ferrous iron and alpha-ketoglutarate. The same enzyme will also catalyze the hydroxylation of Me-PnA, forming Me-HPnA at ~1000-fold more efficiency than using PnA as a substrate, suggesting that an enzyme likely methylates PnA between the action of FzmG. This enzyme was identified to be FzmB, a radical SAM-dependent O-methyltransferase. FzmB can use both PnA and HPnA with almost equal efficiency *in vitro*. Thus, the higher catalytic efficiency of FzmG towards Me-PnA suggests that FzmB acts after formation of PnA, but before hydroxylation of PnA (50).

1.5.1.4 Biosynthesis of hydrazino-arginine

Formation of the methylated hydrazide-containing arginine appears to be more complex. A second, ten-gene operon-like cassette (*fzmJ-S*) encoded directly downstream of *fzmA-I*, appears to be involved in N-N bond formation (49, 146). Labeling studies suggest that aspartic acid is used for N-N bond formation during the biosynthesis of fosfazinomycin, although aspartate is not part of the final molecule. A flavin-dependent N-oxidase, FzmM, was identified in the gene cluster which shares homology to the *N*-oxidase involved in cremeomycin biosynthesis (131). FzmM activity was tested *in vitro*, and it was found that aspartic acid is readily oxidized to form nitrosuccinic acid, which

then spontaneously decarboxylates to form nitropropionic acid. Prior to this decarboxylation, the 3-carboxymuconate cycloisomerase homolog, FzmL, is proposed to generate nitrous acid and fumarate through the oxidation of nitrosuccinic acid. When aspartic acid is incubated in the presence of FzmM and FzmL, fumarate and nitrous acid are observed (49). It is still unknown as to how nitrous acid is added to aspartic acid to yield hydrazinosuccinate, but this is a suspected intermediate in fosfazinomycin biosynthesis due to hydrazinosuccinate being a substrate of the acetyltransferase, FzmQ. FzmR is an adenylosuccinate lyase homolog, that hydrolyzes *N*-acetylhydrazinosuccinate into acetylhydrazine and fumarate (49). It is unknown how the hydrazine moiety is added to arginine, but the radical SAM-dependent enzyme, FzmH, was shown to methylate hydrazino-arginine (50). FzmI, an aminoacyl-tRNA peptidyl transferase, can use arginine, hydrazino-arginine, methylhydrazino-arginine, and fosfazinomycin B as substrates (49).

1.5.1.5 Joining the two small molecules into fosfazinomycin

In total, most of the biosynthetic pathway of fosfazinomycin has been elucidated. However, arguably the most interesting biosynthetic steps await characterization: the formation of the nitrogen-nitrogen bond and the formation of the nitrogen-phosphorus bond. It is possible that comparative genomics may provide insights into the biosynthesis of these bonds. Another phosphorus-containing natural product has been identified that contains a phosphorohydrazate linkage. FR-900137 was purified from the culture extracts of *Streptomyces unzenensis* and shows antimicrobial activity against both Gram-positive and Gram-negative organisms (77). This molecule has structural similarities to fosfazinomycin, in that the molecule also contains an unusual phosphorohydrazate

linkage, however, there is no carbon-phosphorus linkage to classify this molecule as a phosphonic acid (79). One could speculate that the phosphorus-nitrogen bond in FR-900137 and fosfazinomycin may be installed in a similar biosynthetic fashion and comparative genomics may provide insight into the genes needed for this unsolved transformation.

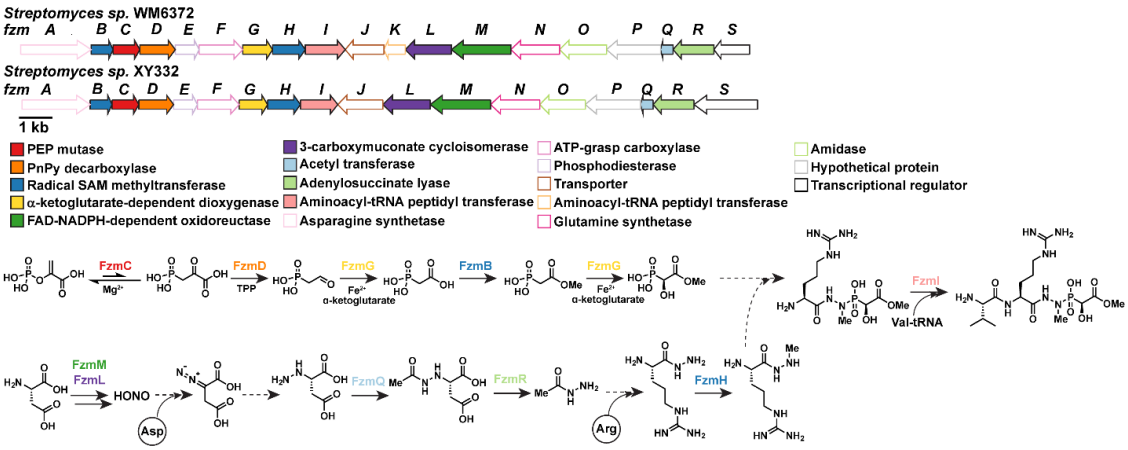


Figure 1.11. Fosfazinomycin biosynthetic gene clusters from *Streptomyces* sp. WM6372 and *Streptomyces* sp. XY332. Current proposed fosfazinomycin biosynthetic pathway based on currently published data.

1.6 DIHYDROXYPROPYLPHOSPHONIC ACID – A COMMON INTERMEDIATE IN THE BIOSYNTHESIS OF VALINOPHOS AND PHOSPHONOTHRIXIN

1.6.1 Discovery and bioactivity of the herbicide phosphonothrixin

Phosphonothrixin was originally isolated from the fermentation broth of *Saccharothrix* sp. ST-888 in 1995 (69, 132) during a search for new compounds showing herbicidal activity. This new phosphonic acid showed promise as an herbicide, as the compound readily bleached the leaves of different gramineous and broadleaf weeds, killing every one of the 12 different weeds tested at a rate >50% if applied directly to the

leaf (132). To this day, we still do not know the mode-of-action of this herbicidal compound. Lange and colleagues speculated that phosphonothrixin may target 1-deoxyxylulose 5-phosphate reductoisomerase (DXR) because of the structural similarity to deoxyxylulose phosphate (DXP), however phosphonothrixin was unable to inhibit isoprenoid biosynthesis in peppermint secretory cells, indicating that DXR is not the physiological target (84).

1.6.2 Valinophos, a novel phosphonic acid discovered using genome mining

Valinophos is a recently described phosphonate discovered through genome mining and was isolated from the fermentation broth of *Streptomyces durhamensis* NRRL B-3309 (59). This compound was not isolated in sufficient quantities for biological activity testing, however, a proposed pathway intermediate, (R)-2,3-dihydroxypropylphosphonic acid (DHPPA), isolated in abundance from *S. durhamensis* extracts, showed inhibitory activity against *Mycobacterium smegmatis* (59).

1.6.2.1 The valinophos and phosphonothrixin gene clusters share homologous genes

The gene clusters for valinophos and phosphonothrixin were both described in 2015, by Ju and colleagues and Lin and colleagues respectively (59, 89). Through examination of the *pepM* genomic neighborhood, the gene clusters for these two molecules were putatively assigned (Fig 1.12). Analysis of these gene clusters suggests that there would be a novel driving force for the thermodynamically unfavorable PepM reaction, as phosphonopyruvate decarboxylase and phosphonomethylmalate synthase homologs are absent from both the biosynthetic gene clusters. Lin and colleagues hypothesize that the two-component transketolase homolog, predicted to condense

pyruvate with phosphonopyruvate in a thiamine pyrophosphate-dependent reaction, would form an intermediate compound that can be subsequently reduced to form phosphonothrixin (89). Alternatively, it is possible that phosphonothrixin biosynthesis proceeds through a similar biosynthetic route to valinophos, as five of the biosynthetic genes are conserved between the two gene clusters (Fig 1.12). As the proposed valinophos gene cluster is missing TPP-dependent enzyme homologs, it is unlikely that valinophos proceeds through a similar biosynthetic scheme proposed by Lin and colleagues.

1.6.2.2 Proposed biosynthesis of valinophos and phosphonothrixin

Ju and colleagues hypothesized that the PepM reaction would be driven forward through the reduction of phosphonopyruvate by VlpB to form 2-hydroxy-3-phosphonopropanoic acid (phosphonolactate) using a short-chain dehydrogenase homolog (59). Subsequently, phosphorylation of the carboxylate via VlpD would form 2-hydroxy-3-oxo-3-phosphorylethylphosphonic acid. This compound can undergo two sequential reductions using VlpC, an *N*-acetyl γ-glutamyl phosphate reductase homolog, and VlpE, an alcohol dehydrogenase homolog, to form DHPPA (Fig 1.12). After this common intermediate, the valinophos and phosphonothrixin biosynthetic pathways are likely to diverge. It is proposed that the two remaining ATP-grasp like proteins will install the ester-linked *N*-acetyl-L-valine to form the final natural product found in *S. durhamensis* extracts (59). During phosphonothrixin biosynthesis, one could imagine that a TPP-dependent enzyme may catalyze the condensation of pyruvate with DHPPA to afford the final natural product, phosphonothrixin (Fig 1.12).

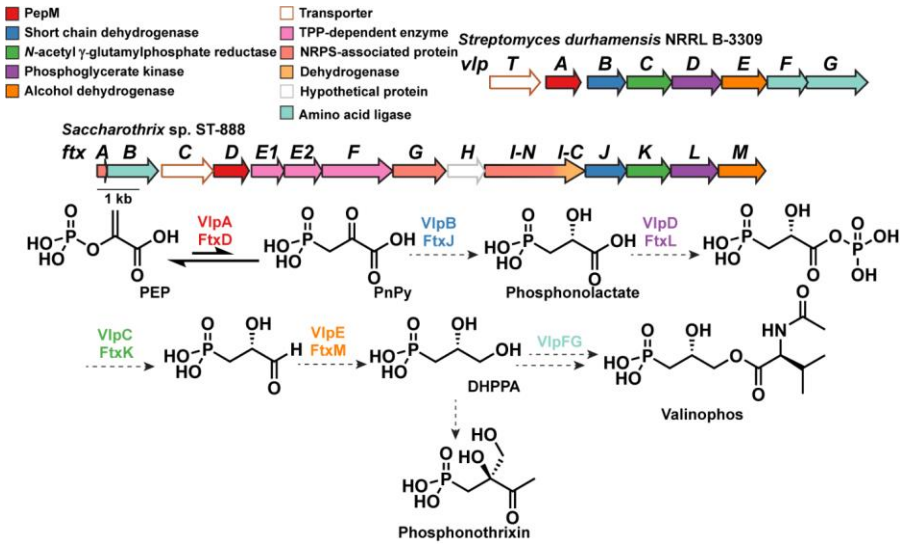


Figure 1.12. The gene clusters from *Streptomyces durhamensis* NRRL B-3309 (valinophos) and *Saccharothrix* sp. ST-888 (phosphonothrixin) are shown. Homologous enzymes are highlighted a similar color. No step of phosphonothrixin or valinophos biosynthesis have been characterized *in vitro* or *in vivo*, thus the biosynthetic pathways described above are hypothetical.

1.7 SUMMARY AND OUTLOOK

Within a ten-year study, the number of phosphonic acid natural products reported within the literature nearly doubled due to the use of genome mining to prioritize strains for phosphonate discovery (59). Even though an immense amount of characterization remains to be done to characterize the bio-transformations within each gene cluster, the knowledge that we have gained in this process has been invaluable. Through the accumulation of knowledge about the interesting bio-transformations involved within phosphonate biosynthesis, we can now quickly identify strains containing gene clusters that may be of interest. This large genome-mining study targeted only actinobacteria to group the number of potential phosphonic acid natural product gene clusters, thus it would be interesting to examine the gene cluster present in all of nature to see if different organisms may also contain the biosynthetic machinery to produce different

phosphonates than those in the actinobacteria. My thesis will describe the work done to characterize different phosphonic acid gene clusters that have been identified during this large screening campaign for novel phosphonic acids (59).

Chapter 2 of my thesis will detail the *in vitro* characterization of DhpA, FzmG, and HpxV, the non-heme iron-dependent and α-ketoglutarate-dependent dioxygenases that all function on small phosphonic acid substrates. **Chapter 3** of my thesis will detail experiments done to characterize the oxime-forming *N*-oxidases found in phosphonocystoximic acid and hydroxyphosphonocystoximic acid biosynthesis. Until recently, the running hypothesis in the field dictated that 2AEP was solely used by organisms as modifications to their membranes as phosphonolipids or phosphonoglycans. As I will show in **Chapter 3** of my thesis, 2AEP is believed to be the substrate of the oxime-forming enzymes, PcxL and HpxL, involved in the biosynthesis of phosphonocystoximic acid (34). Our current understanding is that the phosphonocystoximic acids, nor hydroxynitrilaphos are incorporated as lipid modifications in the bacteria that produce these molecules. Finally, **Chapter 4** of this thesis will cover the work I have done to investigate the biosynthesis of phosphonothrixin. Little is known about the biosynthesis of this potent herbicide, yet this compound has been known for over 20 years (69, 132).

1.8 REFERENCES

1. Allen KD, Wang SC. 2014. Initial characterization of Fom3 from Streptomyces wedmorensis: The methyltransferase in fosfomycin biosynthesis. *Arch. Biochem. Biophys.* 543:67–73.

2. Altincicek B, Kollas A-K, Sanderbrand S, Wiesner J, Hintz M, et al. 2001. GcpE Is Involved in the 2-C-Methyl-d-Erythritol 4-Phosphate Pathway of Isoprenoid Biosynthesis in Escherichia coli. *J. Bacteriol.* 183(8):2411–16.

3. Antibiotic A53868 and process for production thereof. 1984.

4. Barry RJ, Bowman E, McQueney M, Dunaway-Mariano D. 1988. Elucidation of the 2-aminoethylphosphonate biosynthetic pathway in Tetrahymena pyriformis. *Biochem. Biophys. Res. Commun.* 153(1):177–82.

5. Baumann H, Tzianabos AO, Brisson JR, Kasper DL, Jennings HJ. 1992. Structural elucidation of two capsular polysaccharides from one strain of Bacteroides fragilis using high-resolution NMR spectroscopy. *Biochemistry (Mosc.)*. 31(16):4081–89.

6. Bayer E, Gugel KH, Hägele K, Hagenmaier H, Jessipow S, et al. 1972. Stoffwechselprodukte von Mikroorganismen. 98. Mitteilung. Phosphinothricin und Phosphinothricyl-Alanyl-Alanin. *Helv. Chim. Acta*. 55(1):224–239.

7. Bender MA, Kumarasamy N, Mayer KH, Wang B, Walensky RP, et al. 2010. Cost-Effectiveness of Tenofovir as First-Line Antiretroviral Therapy in India. *Clin. Infect. Dis.* 50(3):416–25.

8. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, et al. 2018. GenBank. *Nucleic Acids Res.* 46(D1):D41–47.

9. Bentley SD, Chater KF, Cerdeño-Tárraga A-M, Challis GL, Thomson NR, et al. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature*. 417(6885):141–47.

10. Blodgett JA, Zhang JK, Yu X, Metcalf WW. 2016. Conserved biosynthetic pathways for phosalacine, bialaphos and newly discovered phosphonic acid natural products. *J. Antibiot. (Tokyo)*. 69(1):15–25.

11. Blodgett JAV, Zhang JK, Metcalf WW. 2005. Molecular Cloning, Sequence Analysis, and Heterologous Expression of the Phosphinothricin Tripeptide Biosynthetic Gene Cluster from Streptomyces viridochromogenes DSM 40736. *Antimicrob. Agents Chemother.* 49(1):230–40.

12. Bondad ND, Pantastico EB. 1972. Ethrel-induced ripening of immature and mature-green tomato fruits. *Econ. Bot.* 26(3):238–44.

13. Bougioukou D, Mukherjee S, van der Donk W. 2013. Revisiting the biosynthesis of dehydrophos reveals a tRNA-dependent pathway. *PNAS.* 110(27):10952–57.

14. Bowman E, McQueney M, Barry RJ, Dunaway-Mariano D. 1988. Catalysis and thermodynamics of the phosphoenolpyruvate/phosphonopyruvate rearrangement. Entry into the phosphonate class of naturally occurring organophosphorus compounds. *J. Am. Chem. Soc.* 110(16):5575–76.

15. Challis GL. 2014. Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* 41(2):219–32.

16. Chen CCH, Han Y, Niu W, Kulakova AN, Howard A, et al. 2006. Structure and Kinetics of Phosphonopyruvate Hydrolase from Voriovorax sp. Pal2: New Insight into the Divergence of Catalysis within the PEP Mutase/Isocitrate Lyase Superfamily,. *Biochemistry (Mosc.).* 45(38):11491–504.

17. Chen CCH, Zhang H, Kim AD, Howard A, Sheldrick GM, et al. 2002. Degradation Pathway of the Phosphonate Ciliatine: Crystal Structure of 2-Aminoethylphosphonate Transaminase,. *Biochemistry (Mosc.).* 41(44):13162–69.

18. Cho S-H, Kim S-Y, Tomita T, Shiraishi T, Park J-S, et al. 2017. Fosfomycin Biosynthesis via Transient Cytidylylation of 2-Hydroxyethylphosphonate by the Bifunctional Fom1 Enzyme. *ACS Chem. Biol.* 12(8):2209–15.

19. Cioni JP, Doroghazi JR, Ju K-S, Yu X, Evans BS, et al. 2014. Cyanohydrin Phosphonate Natural Product from Streptomyces regensis. *J. Nat. Prod.* 77(2):243–49.

20. Circello BT, Eliot AC, Lee J-H, van der Donk WA, Metcalf WW. 2010. Molecular Cloning and Heterologous Expression of the Dehydrophos Biosynthetic Gene Cluster. *Chem. Biol.* 17(4):402–11.

21. Circello BT, Miller CG, Lee J-H, Donk WA van der, Metcalf WW. 2011. The Antibiotic Dehydrophos Is Converted to a Toxic Pyruvate Analog by Peptide Bond Cleavage in Salmonella enterica. *Antimicrob. Agents Chemother.* 55(7):3357–62.

22. De Clercq E. 2003. Clinical Potential of the Acyclic Nucleoside Phosphonates Cidofovir, Adefovir, and Tenofovir in Treatment of DNA Virus and Retrovirus Infections. *Clin. Microbiol. Rev.* 16(4):569–96.

23. De Clercq E, Férir G, Kaptein S, Neyts J. 2010. Antiviral Treatment of Chronic Hepatitis B Virus (HBV) Infections. *Viruses*. 2(6):1279–1305.

24. Dev S. 1999. Ancient-modern concordance in Ayurvedic plants: some examples. *Environ. Health Perspect.* 107(10):783–89.

25. Dumora C, Lacoste A-M, Cassaigne A. 1983. Purification and Properties of 2-Aminoethylphosphonate: Pyruvate Aminotransferase from Pseudomonas aeruginusa. *Eur. J. Biochem.* 133(1):119–25.

26. Dyhrman ST, Benitez-Nelson CR, Orchard ED, Haley ST, Pellechia PJ. 2009. A microbial source of phosphonates in oligotrophic marine systems. *Nat. Geosci.* 2(10):696–99.

27. Eliot AC, Griffin BM, Thomas PM, Johannes TW, Kelleher NL, et al. 2008. Biosynthesis of the Potent Antimalarial Compound FR900098. *Chem. Biol.* 15(8):765–70.

28. Falagas ME, Vouloumanou EK, Samonis G, Vardakas KZ. 2016. Fosfomycin. *Clin. Microbiol. Rev.* 29(2):321–47.

29. Falagas ME, Vouloumanou EK, Togias AG, Karadima M, Kapaskelis AM, et al. 2010. Fosfomycin versus other antibiotics for the treatment of cystitis: a meta-analysis of randomized controlled trials. *J. Antimicrob. Chemother.* 65(9):1862–77.

30. Gallant JE, DeJesus E, Arribas JR, Pozniak AL, Gazzard B, et al. 2009. Tenofovir DF, Emtricitabine, and Efavirenz vs. Zidovudine, Lamivudine, and Efavirenz for HIV. *N. Engl. J.Med.* 354(3):251-260.

31. Gao J, Ju K-S, Yu X, Velásquez JE, Mukherjee S, et al. 2014. Use of a Phosphonate Methyltransferase in the Identification of the Fosfazinomycin Biosynthetic Gene Cluster. *Angew. Chem. Int. Ed.* 53(5):1334–37.

32. Gill HS, Eisenberg D. 2001. The Crystal Structure of Phosphinothricin in the Active Site of Glutamine Synthetase Illuminates the Mechanism of Enzymatic Inhibition. *Biochemistry (Mosc.).* 40(7):1903–12.

33. Glawischnig E, Mikkelsen MD, Halkier BA. 2003. Glucosinolates: Biosynthesis and Metabolism. In *Sulphur in Plants*, ed YP Abrol, A Ahmad, pp. 145–62. Springer Netherlands.

34. Goettge MN, Cioni JP, Ju K-S, Pallitsch K, Metcalf WW. 2018. PcxL and HpxL are flavin-dependent, oxime-forming N-oxidases in phosphonocystoximic acid biosynthesis in *Streptomyces*. *J. Biol. Chem.* 293(18):6859–68.

35. Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 17(6):333–51.

36. Gunji S, Arima K, Beppu T. 1983. Screening of Antifungal Antibiotics According to Activities Inducing Morphological Abnormalities. *Agric. Biol. Chem.* 47(9):2061–69.

37. Halkier BA, Gershenzon J. 2006. Biology and Biochemistry of Glucosinolates. *Annu. Rev. Plant Biol.* 57(1):303–33.

38. Halkier BA, Hansen CH, Mikkelsen MD, Naur P, Wittstock U. 2002. Chapter Thirteen The role of cytochromes P450 in biosynthesis and evolution of glucosinolates. In *Recent Advances in Phytochemistry*, ed JT Romeo, RA Dixon. 36:223–48.

39. Hasegawa S, Tamari M, Kametaka M. 1976. Isolation of diacylglyceryl-2-aminoethylphosphonate from bovine liver. *J. Biochem. (Tokyo)*. 80(3):531–35.

40. Hayek T. 1999. The angiotensin-converting enzyme inhibitor, fosinopril, and the angiotensin II receptor antagonist, losartan, inhibit LDL oxidation and attenuate atherosclerosis independent of lowering blood pressure in apolipoprotein E deficient mice. *Cardiovasc. Res.* 44(3):579–87.

41. Hendlin D, Stapley EO, Jackson M, Wallick H, Miller AK, et al. 1969. Phosphonomycin, a new antibiotic produced by strains of streptomyces. *Science*. 166(3901):122–23.

42. Hidaka T, Goda M, Kuzuyama T, Takei N, Hidaka M, Seto H. 1995. Cloning and nucleotide sequence of fosfomycin biosynthetic genes of Streptomyces wedmorensis. *Mol. Gen. Genet. MGG*. 249(3):274–80.

43. Hidaka T, Iwakura H, Imai S, Seto H. 1992. Studies on the biosynthesis of fosfomycin. 3. Detection of phosphoenol-pyruvate phosphomutase activity in a fosfomycin high-producing strain of Streptomyces wedmorensis and characterization of its blocked mutant NP-7. *J. Antibiot. (Tokyo)*. 45(6):1008–10.

44. Hidaka T, Mori M, Imai S, Hara O, Nagaoka K, Seto H. 1989. Studies on the biosynthesis of bialaphos (SF-1293). 9. Biochemical mechanism of C-P bond formation in bialaphos: discovery of phosphoenolpyruvate phosphomutase which catalyzes the formation of phosphonopyruvate from phosphoenolpyruvate. *J. Antibiot. (Tokyo)*. 42(3):491–94.

45. Higgins LJ, Yan F, Liu P, Liu H, Drennan CL. 2005. Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature*. 437(7060):838–44.

46. Horiguchi M, Kandatstu M. 1959. Isolation of 2-Aminoethane Phosphonic Acid from Rumen Protozoa. *Nature*. 184(4690):901–2.

47. Hu K, Werner WJ, Allen KD, Wang SC. 2015. Investigation of enzymatic C–P bond formation using multiple quantum HCP nuclear magnetic resonance spectroscopy. *Magn. Reson. Chem.* 53(4):267–72.

48. Huang K, Li Z, Jia Y, Dunaway-Mariano D, Herzberg O. 1999. Helix swapping between two α/β barrels: crystal structure of phosphoenolpyruvate mutase with bound Mg2+–oxalate. *Structure*. 7(5):539–48.

49. Huang Z, Wang K-KA, Donk WA van der. 2016. New insights into the biosynthesis of fosfazinomycin. *Chem. Sci.* 7(8):5219–23.

50. Huang Z, Wang K-KA, Lee J, Donk WA van der. 2015. Biosynthesis of fosfazinomycin is a convergent process. *Chem. Sci.* 6(2):1282–87.

51. Hunt AH, Elzey TK. 1988. Revised structure of A53868A. *J. Antibiot. (Tokyo)*. 41(6):802.

52. Iguchi E, Okuhara M, Kohsaka M, Aoki H, Imanaka H. 1980. Studies on new phosphonic acid antibiotics. *J. Antibiot. (Tokyo)*. 33(1):18–23.

53. Jayasimhulu K, Hunt SM, Kaneshiro ES, Watanabe Y, Giner J-L. 2007. Detection and identification of Bacteriovorax stolpii UKi2 Sphingophosphonolipid molecular species. *J. Am. Soc. Mass Spectrom.* 18(3):394–403.

54. J. Newman D, M. Cragg G, M. Snader K. 2000. The influence of natural products upon drug discovery. *Nat. Prod. Rep.* 17(3):215–34.

55. Johnen S, Sprenger GA. 2009. Characterization of recombinant thiamine diphosphate-dependent phosphonopyruvate decarboxylase from Streptomyces viridochromogenes Tü494. *J. Mol. Catal. B Enzym.* 61(1):39–46.

56. Johnson T, Newton G, Fahey RC, Rawat M. 2009. Unusual production of glutathione in Actinobacteria. *Arch. Microbiol.* 191(1):89–93.

57. Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, et al. 1999. Inhibitors of the Nonmevalonate Pathway of Isoprenoid Biosynthesis as Antimalarial Drugs. *Science.* 285(5433):1573–76.

58. Ju K-S, Doroghazi JR, Metcalf WW. 2014. Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.*, pp. 1–12.

59. Ju K-S, Gao J, Doroghazi JR, Wang K-KA, Thibodeaux CJ, et al. 2015. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc. Natl. Acad. Sci.* 112(39):12175–80.

60. Kahan FM, Kahan JS, Cassidy PJ, Kropp H. 1974. The Mechanism of Action of Fosfomycin (phosphonomycin). *Ann. N. Y. Acad. Sci.* 235(1):364–86.

61. Kariotoglou DM, Mastronicolis SK. 2001. Sphingophosphonolipids, phospholipids, and fatty acids from aegean jellyfish Aurelia aurita. *Lipids.* 36(11):1255.

62. Karl DM. 2014. Microbially Mediated Transformations of Phosphorus in the Sea: New Views of an Old Cycle. *Annu. Rev. Mar. Sci.* 6(1):279–337.

63. Kato H, Nagayama K, Abe H, Kobayashi R, Ishihara E. 1991. Isolation, Structure and Biological Activity of Trialaphos (Organic Chemistry). *Agric. Biol. Chem.* 55(4):1133–34.

64. Kennedy KE, Thompson GA Jr. 1970. Phosphonolipids: localization in surface membranes of Tetrahymena. *Science.* 168(3934):989–91.

65. Kido Y, Hamakado T, Anno M, Miyagawa E, Motoki Y, et al. 1984. Isolation and characterization of I5B2, a new phosphorus containing inhibitor of angiotensin. *J. Antibiot. (Tokyo).* 37(9):965–69.

66. Kim AD, Baker AS, Dunaway-Mariano D, Metcalf WW, Wanner BL, Martin BM. 2002. The 2-Aminoethylphosphonate-Specific Transaminase of the 2-Aminoethylphosphonate Degradation Pathway. *J. Bacteriol.* 184(15):4134–40.

67. Kim J, Dunaway-Mariano D. 1996. Phosphoenolpyruvate Mutase Catalysis of Phosphoryl Transfer in Phosphoenolpyruvate: Kinetics and Mechanism of Phosphorus−Carbon Bond Formation†. *Biochemistry (Mosc.).* 35(14):4628–35.

68. Kim SY, Ju K-S, Metcalf WW, Evans BS, Kuzuyama T, Donk WA van der. 2012. Different Biosynthetic Pathways to Fosfomycin in Pseudomonas syringae and Streptomyces Species. *Antimicrob. Agents Chemother.* 56(8):4175–83.

69. Kimura T, Nakamura K, Takahashi E. 1995. Phosphonothrixin, a novel herbicidal antibiotic produced by *Saccharothrix* sp. ST-888. II. Structure determination. *J. Antibiot. (Tokyo).* 48(10):1130–33.

70. Kittredge JS, Horiguchi M, Williams PM. 1969. Aminophosphonic acids: Biosynthesis by marine phytoplankton. *Comp. Biochem. Physiol.* 29(2):859–63.

71. Kobayashi S, Kuzuyama T, Seto H. 2000. Characterization of the fomA andfomB Gene Products from Streptomyces wedmorensis, Which Confer Fosfomycin Resistance on Escherichia coli. *Antimicrob. Agents Chemother.* 44(3):647–50.

72. Koguchi T, Yamada K, Yamato M, Okachi R, Nakayama K, Kase H. 1986. K-4, A novel inhibitor of angiotensin. *J. Antibiot. (Tokyo).* 39(3):364–71.

73. Kojo H, Shigi Y, Nishida M. 1980. FR-31564, A new phosphonic acid antibiotic: bacterial resistance and membrane permeability. *J. Antibiot. (Tokyo).* 33(1):44–48.

74. Kuemin M, van der Donk WA. 2010. Structure–activity relationships of the phosphonate antibiotic dehydrophos. *Chem. Commun. Camb. Engl.* 46(41):7694–96.

75. Kuemmerle HP, Murakawa T, De Santis F. 1987. Pharmacokinetic evaluation of fosmidomycin, a new phosphonic acid antibiotic. *Chemioter. Int. J. Mediterr. Soc. Chemother.* 6(2):113–19.

76. Kuntz L, Tritsch D, Grosdemange-Billiard C, Hemmerlin A, Willem A, et al. 2005. Isoprenoid biosynthesis as a target for antibacterial and antiparasitic drugs: phosphonohydroxamic acids as inhibitors of deoxyxylulose phosphate reducto-isomerase. *Biochem. J.* 386(Pt 1):127–35.

77. Kuroda Y, Goto T, Okamoto M, Yamashita M, Iguchi E, et al. 1980. FR-900137, A NEW ANTIBIOTIC. *J. Antibiot. (Tokyo)*. 33(3):272–79.

78. Kuroda Y, Okuhara M, Goto T, Okamoto M, Terano H, et al. 1980. Studies on new phosphonic acid antibiotics. *J. Antibiot. (Tokyo)*. 33(1):29–35.

79. Kuroda Y, Tanaka H, Okamoto M, Goto T, Kohsaka M, et al. 1980. FR-900137, a new antibiotic. II. Structure determination of FR-900137. *J. Antibiot. (Tokyo)*. 33(3):280–83.

80. Kuzuyama T, Hidaka T, Kamigiri K, Imai S, Seto H. 1992. Studies on the biosynthesis of fosfomycin. 4. The biosynthetic origin of the methyl group of fosfomycin. *J. Antibiot. (Tokyo)*. 45(11):1812–14.

81. Kuzuyama T, Seki T, Kobayashi S, Hidaka T, Seto H. 1999. Cloning and expression in Escherichia coli of 2-hydroxypropylphosphonic acid epoxidase from the fosfomycin-producing organism, Pseudomonas syringae PB-5123. *Biosci. Biotechnol. Biochem.* 63(12):2222–24.

82. Kuzuyama T, Shimizu T, Takahashi S, Seto H. 1998. Fosmidomycin, a specific inhibitor of 1-deoxy-d-xylulose 5-phosphate reductoisomerase in the nonmevalonate pathway for terpenoid biosynthesis. *Tetrahedron Lett.* 39(43):7913–16.

83. Lacoste AM, Dumora C, Balas L, Hammerschmidt F, Vercauteren J. 1993. Stereochemistry of the reaction catalysed by 2-aminoethylphosphonate aminotransferase. A 1H-NMR study. *Eur. J. Biochem.* 215(3):841–44.

84. Lange BM, Ketchum REB, Croteau RB. 2001. Isoprenoid Biosynthesis. Metabolite Profiling of Peppermint Oil Gland Secretory Cells and Application to Herbicide Target Analysis. *Plant Physiol.* 127(1):305.

85. Lee J-H, Evans BS, Li G, Kelleher NL, van der Donk WA. 2009. In Vitro Characterization of a Heterologously Expressed Nonribosomal Peptide Synthetase Involved in Phosphinothricin Tripeptide Biosynthesis. *Biochemistry (Mosc.)*. 48(23):5054–56.

86. Lell B, Ruangweerayut R, Wiesner J, Missinou MA, Schindler A, et al. 2003. Fosmidomycin, a Novel Chemotherapeutic Agent for Malaria. *Antimicrob. Agents Chemother.* 47(2):735–38.

87. Liang C-R, Rosenberg H. 1968. On the distribution and biosynthesis of 2-aminoethylphosphonate in two terrestrial molluscs. *Comp. Biochem. Physiol.* 25(2):673–81.

88. Liang CR, Rosenberg H. 1968. The biosynthesis of the carbon-phosphorus bond in Tetrahymena. *Biochim. Biophys. Acta BBA - Gen. Subj.* 156(2):437–39.

89. Lin J, Nishiyama M, Kuzuyama T. 2014. Identification of the biosynthetic gene cluster for the herbicide phosphonothrixin in Saccharothrix sp. ST-888. *J. Antibiot. (Tokyo).*

90. McClung M, Harris ST, Miller PD, Bauer DC, Davison KS, et al. 2013. Bisphosphonate Therapy for Osteoporosis: Benefits, Risks, and Drug Holiday. *Am. J. Med.* 126(1):13–20.

91. McLuskey K, Cameron S, Hammerschmidt F, Hunter WN. 2005. Structure and reactivity of hydroxypropylphosphonic acid epoxidase in fosfomycin biosynthesis by a cation- and flavin-dependent mechanism. *Proc. Natl. Acad. Sci.* 102(40):14221–26.

92. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, et al. 2015. *Minimum Information about a Biosynthetic Gene cluster*. Nature Chemical Biology. https://www.nature.com/articles/nchembio.

93. Metcalf WW, van der Donk WA. 2009. Biosynthesis of Phosphonic and Phosphinic Acid Natural Products. *Annu. Rev. Biochem.* 78:65–94.

94. Miceli MV, Henderson TO, Myers TC. 1980. 2-aminoethylphosphonic Acid metabolism during embryonic development of the planorbid snail helisoma. *Science*. 209(4462):1245–47.

95. Michael CA, Dominey-Howes D, Labbate M. 2014. The Antimicrobial Resistance Crisis: Causes, Consequences, and Management. *Front. Public Health*.

96. Michalopoulos AS, Livaditis IG, Gougoutas V. 2011. The revival of fosfomycin. *Int. J. Infect. Dis. IJID Off. Publ. Int. Soc. Infect. Dis.* 15(11):e732-739.

97. Mine Y, Kamimura T, Nonoyama S, Nishida M, Goto S, Kuwahara S. 1980. IN VITRO AND IN VIVO ANTIBACTERIAL ACTIVITIES OF FR-31564, A NEW PHOSPHONIC ACID ANTIBIOTIC. *J. Antibiot. (Tokyo)*. 33(1):36–43.

98. Munos JW, Moon S-J, Mansoorabadi SO, Chang W, Hong L, et al. 2008. Purification and Characterization of the Epoxidase Catalyzing the Formation of Fosfomycin from Pseudomonas syringae. *Biochemistry (Mosc.)*. 47(33):8726–35.

99. Nakashita H, Kozuka K, Hidaka T, Hara O, Seto H. 2000. Identification and expression of the gene encoding phosphonopyruvate decarboxylase of Streptomyces hygroscopicus *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* 1490(1):159–62.

100. Nakashita H, Shimazu A, Hidaka T, Seto H. 1992. Purification and characterization of phosphoenolpyruvate phosphomutase from Pseudomonas gladioli B-1. *J. Bacteriol.* 174(21):6857–61.

101. Nakashita H, Watanabe K, Hara O, Hidaka T, Seto H. 1997. Studies on the biosynthesis of bialaphos. 17. Biochemical mechanism of C-P bond formation: discovery of phosphonopyruvate decarboxylase which catalyzes the formation of phosphonoacetaldehyde from phosphonopyruvate. *J. Antibiot. (Tokyo)*. 50(3):212–19.

102. Newman DJ, Cragg GM. 2016. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* 79(3):629–61.

103. O'Brien TA, Kluger R, Pike DC, Gennis RB. 1980. Phosphonate analogues of pyruvate. Probes of substrate binding to pyruvate oxidase and other thiamin pyrophosphate-dependent decarboxylases. *Biochim. Biophys. Acta BBA - Enzymol.* 613(1):10–17.

104. Ogawa H, Tsuruoka T, Inouye S, Niida T. 1983. Studies on a new antibiotic SF-1293. *Sci Rep Meiji Seika Kaisha*. 13:42–48.

105. Ogita T, Gunji S, Fukazawa Y, Terahara A, Kinoshita T, et al. 1983. The structures of fosfazinomycins A and B. *Tetrahedron Lett.* 24(22):2283–86.

106. Okuhara M, Kuroda Y, Goto T, Okamoto M, Terano H, et al. 1980. STUDIES ON NEW PHOSPHONIC ACID ANTIBIOTICS. *J. Antibiot. (Tokyo)*. 33(1):24–28.

107. Olivares P, Ulrich EC, Chekan JR, van der Donk WA, Nair SK. 2017. Characterization of Two Late-Stage Enzymes Involved in Fosfomycin Biosynthesis in Pseudomonads. *ACS Chem. Biol.* 12(2):456–63.

108. Omura S, Murata M, Hanaki H, Hinotozawa K, Oiwa R, Tanaka H. 1984. Phosalacine, a new herbicidal antibiotic containing phosphinothricin. Fermentation, isolation, biological activity and mechanism of action. *J. Antibiot. (Tokyo)*. 37(8):829–35.

109. Osbourn A. 2010. Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation. *Trends Genet.* 26(10):449–57.

110. Papatheodoridis GV, Dimou E, Papadimitropoulos V. 2002. Nucleoside analogues for chronic hepatitis B: antiviral efficacy and viral resistance. *Am. J. Gastroenterol.* 97(7):1618–28.

111. Peck SC, van der Donk WA. 2013. Phosphonate biosynthesis and catabolism: a treasure trove of unusual enzymology. *Curr. Opin. Chem. Biol.* 17(4):580–88.

112. Quin LD. 1965. The Prsence of Compounds with a Carbon-Phosphorus Bond in Some Marine Invertebrates*. *Biochemistry (Mosc.)*. 4(2):324–30.

113. Reeves DS. 1994. Fosfomycin trometamol. *J. Antimicrob. Chemother.* 34(6):853–58.

114. Rossolini GM, Arena F, Pecile P, Pollini S. 2014. Update on the antibiotic resistance crisis. *Curr. Opin. Pharmacol.* 18:56–60.

115. Sarkar M, Hamilton CJ, Fairlamb AH. 2003. Properties of Phosphoenolpyruvate Mutase, the First Enzyme in the Aminoethylphosphonate Biosynthetic Pathway in Trypanosoma cruzi. *J. Biol. Chem.* 278(25):22703–8.

116. Sato S, Kudo F, Kim S-Y, Kuzuyama T, Eguchi T. 2017. Methylcobalamin-Dependent Radical SAM C-Methyltransferase Fom3 Recognizes Cytidylyl-2-hydroxyethylphosphonate and Catalyzes the Nonstereoselective C-Methylation in Fosfomycin Biosynthesis. *Biochemistry (Mosc.).* 56(28):3519–22.

117. Schinko E, Schad K, Eys S, Keller U, Wohlleben W. 2009. Phosphinothricin-tripeptide biosynthesis: An original version of bacterial secondary metabolism? *Phytochemistry*. 70(15–16):1787–1800.

118. Schwartz D, Berger S, Heinzelmann E, Muschko K, Welzel K, Wohlleben W. 2004. Biosynthetic Gene Cluster of the Herbicide Phosphinothricin Tripeptide from Streptomyces viridochromogenes Tü494. *Appl. Environ. Microbiol.* 70(12):7093–7102.

119. Schwartz D, Recktenwald J, Pelzer S, Wohlleben W. 1998. Isolation and characterization of the PEP-phosphomutase and the phosphonopyruvate decarboxylase genes from the phosphinothricin tripeptide producer Streptomyces viridochromogenes Tü494. *FEMS Microbiol. Lett.* 163(2):149–157.

120. Seidel HM, Freeman S, Schwalbe CH, Knowles JR. 1990. Phosphonate biosynthesis: the stereochemical course of phosphoenolpyruvate mutase. *J. Am. Chem. Soc.* 112(22):8149–55.

121. Seidel HM, Freeman S, Seto H, Knowles JR. 1988. Phosphonate biosynthesis: isolation of the enzyme responsible for the formation of a carbon–phosphorus bond. *Nature*. 335(6189):457–58.

122. Selmar D. 2010. Biosynthesis of Cyanogenic Glycosides, Glucosinolates and Non-Protein Amino Acids. In *Annual Plant Reviews Volume 40: Biochemistry of Plant Secondary Metabolism*, pp. 92–181.

123. Seto H, Kuzuyama T. 1999. Bioactive natural products with carbon–phosphorus bonds and their biosynthesis. *Nat. Prod. Rep.* 16(5):589–96.

124. Shao Z, Blodgett JAV, Circello BT, Eliot AC, Woodyer R, et al. 2008. Biosynthesis of 2-Hydroxyethylphosphonate, an Unexpected Intermediate Common to Multiple Phosphonate Biosynthetic Pathways. *J. Biol. Chem.* 283(34):23161–68.

125. Shigi Y. 1989. Inhibition of bacterial isoprenoid synthesis by fosmidomycin, a phosphonic acid-containing antibiotic. *J. Antimicrob. Chemother.* 24(2):131–45.

126. Shoji J, Kato T, Hinoo H, Hattori T, Hirooka K, et al. 1986. Production of fosfomycin (phosphonomycin) by Pseudomonas syringae. *J. Antibiot. (Tokyo)*. 39(7):1011–12.

127. Sikorski JA, Gruys KJ. 1997. Understanding Glyphosate's Molecular Mode of Action with EPSP Synthase: Evidence Favoring an Allosteric Inhibitor Model. *Acc. Chem. Res.* 30(1):2–8.

128. Smith JM, Vierling RJ, Meyers CF. 2012. Selective inhibition of E. coli 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *MedChemComm*. 3:65–67.

129. Stapley EO, Hendlin D, Mata JM, Jackson M, Wallick H, et al. 1969. Phosphonomycin. I. Discovery and in vitro biological characterization. *Antimicrob. Agents Chemother.* 9:284–90.

130. Stone MJ, Williams DH. 1992. On the evolution of functional secondary metabolites (natural products). *Mol. Microbiol.* 6(1):29–34.

131. Sugai Y, Katsuyama Y, Ohnishi Y. 2016. A nitrous acid biosynthetic pathway for diazo group formation in bacteria. *Nat. Chem. Biol.* 12(2):73–75.

132. Takahashi E, Kimura T, Nakamura K, Arahira M, Iida M. 1995. Phosphonothrixin, a novel herbicidal antibiotic produced by Saccharothrix sp. ST-888. I. Taxonomy, fermentation, isolation and biological properties. *J. Antibiot. (Tokyo)*. 48(10):1124–29.

133. Umeda T, Tanaka N, Kusakabe Y, Nakanishi M, Kitade Y, Nakamura KT. 2011. Molecular basis of fosmidomycin's action on the human malaria parasite Plasmodium falciparum. *Sci. Rep.* 1:9.

134. Ventola CL. 2015. The Antibiotic Resistance Crisis. *Pharm. Ther*. 40(4):277–83.

135. Wassef MK, Hendrix JW. 1976. Ceramide aminoethylphosphonate in the fungus Pythium prolatum. *Biochim. Biophys. Acta*. 486(1):172–78.

136. Watanabe Y, Nakajima M, Hoshino T, Jayasimhulu K, Brooks EE, Kaneshiro ES. 2001. A novel sphingophosphonolipid head group 1-hydroxy-2-aminoethyl phosphonate in Bdellovibrio stolpii. *Lipids*. 36(5):513–19.

137. Weber MA. 1992. Fosinopril: a new generation of angiotensin-converting enzyme inhibitors. *J. Cardiovasc. Pharmacol.* 20 Suppl 10:S7-12.

138. Werner WJ, Allen KD, Hu K, Helms GL, Chen BS, Wang SC. 2011. In Vitro Phosphinate Methylation by PhpK from Kitasatospora phosalacinea. *Biochemistry (Mosc.)*. 50(42):8986–88.

139. White AK, Metcalf WW. 2007. Microbial Metabolism of Reduced Phosphorus Compounds. *Annu. Rev. Microbiol.* 61(1):379–400.

140. Whitteck JT, Ni W, Griffin BM, Eliot AC, Thomas PM, et al. 2007. Reassignment of the Structure of the Antibiotic A53868 Reveals an Unusual Amino Dehydrophosphonic Acid. *Angew. Chem. Int. Ed.* 46(47):9089–92.

141. Woodyer RD, Shao Z, Thomas PM, Kelleher NL, Blodgett JAV, et al. 2006. Heterologous Production of Fosfomycin and Identification of the Minimal Biosynthetic Gene Cluster. *Chem. Biol.* 13(11):1171–82.

142. Xiong X, Smith JL, Chen MS. 1997. Effect of incorporation of cidofovir into DNA by human cytomegalovirus DNA polymerase on DNA elongation. *Antimicrob. Agents Chemother.* 41(3):594–99.

143. Yamato M, Koguchi T, Okachi R, Yamada K, Nakayama K, et al. 1986. K-26, a novel inhibitor of angiotensin I converting enzyme produced by an actinomycete K-26. *J. Antibiot. (Tokyo).* 39(1):44–52.

144. Yeganeh-Sefidan F, Ghotaslou R, Akhi MT, Sadeghi MR, Mohammadzadeh-Asl Y, Bannazadeh Baghi H. 2016. Fosfomycin, interesting alternative drug for treatment of urinary tract infections created by multiple drug resistant and extended spectrum β-lactamase producing strains. *Iran. J. Microbiol.* 8(2):125–31.

145. Yoshihiro Y, Mitsuru O, Akira H. 1992. Structural elucidation of a novel phosphonoglycosphingolipid in eggs of the sea hare Aplysia juliana. *Biochim. Biophys. Acta BBA - Lipids Lipid Metab.* 1165(2):160–66.

146. Yu X, Doroghazi JR, Janga SC, Zhang JK, Circello B, et al. 2013. Diversity and abundance of phosphonate biosynthetic genes in nature. *Proc. Natl. Acad. Sci.*, p. 201315107.

147. Yu X, Price NPJ, Evans BS, Metcalf WW. 2014. Purification and Characterization of Phosphonoglycans from Glycomyces sp. NRRL B-16210 and Stackebrandtia nassauensis NRRL B-16338. *J. Bacteriol.* 3:6-14.

148. Yun D, Dey M, Higgins LJ, Yan F, Liu H, Drennan CL. 2011. Structural Basis of Regiospecificity of a Mononuclear Iron Enzyme in Antibiotic Fosfomycin Biosynthesis. *J. Am. Chem. Soc.* 133(29):11262–69.

149. Zhanel GG, Walkty AJ, Karlowsky JA. 2016. Fosfomycin: A First-Line Oral Therapy for Acute Uncomplicated Cystitis. *Can. J. Infect. Dis. Med. Microbiol. J. Can. Mal. Infect. Microbiol. Médicale*. 2016.

150. Zhang G, Dai J, Lu Z, Dunaway-Mariano D. 2003. The Phosphonopyruvate Decarboxylase from Bacteroides fragilis. *J. Biol. Chem.* 278(42):41302–8.

Chapter 2: Biochemical characterization of three TauD-family dioxygenases involved in phosphonate biosynthesis

2.1 INTRODUCTION

2.1.1 Non-heme iron and α-ketoglutarate-dependent oxygenases

Non-heme iron and α-ketoglutarate-dependent oxygenases are common enzymes found in all domains of life and are known to catalyze a variety of functions involving C-H bond activation, including, but not limited to: hydroxylation, halogenation, epoxidation, epimerization, dealkylation, desaturation, cyclization, and ring opening (19, 25, 30, 32). With the number of different types of reactions these enzymes catalyze, it is unsurprising that these enzymes can share little sequence identity with one another, with the common feature of these enzymes being a conserved 2-His-1-carboxylate facial triad that coordinates the ferrous iron center (8, 15).

Natural product diversity is often influenced by the different functions carried out by this diverse enzyme superfamily (19, 25, 30, 32) and phosphonic acid natural products are no exception. One of the larger families within the non-heme iron-dependent dehydrogenase superfamily are members of the taurine dioxygenase (TauD) family of enzymes (20, 30). The founding member of the TauD-family is required for use of taurine degradation as a sulfur source (6). TauD catalyzes the hydroxylation of taurine at the alpha position, relative to the sulfonate moiety, which results in fragmentation of the molecule to produce sulfite and acetaldehyde (Fig 2.1) (6). Since the discovery of TauD, numerous family members with roles in the biosynthesis of different natural products

have been discovered, including multiple enzymes required for synthesis of phosphonic acid natural products.

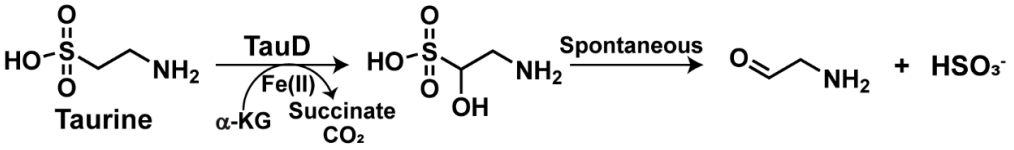


Figure 2.1. TauD reaction and products formed (6).

2.1.2 Non-heme iron and α-ketoglutarate-dependent oxygenases in phosphonate biosynthesis

To date, four different phosphonate biosynthetic pathways have been characterized that encode members of the TauD family of non-heme iron and α-ketoglutarate-dependent oxygenases within their gene clusters that perform a hydroxylation reaction, Fig 2.2. FR-33289 is a hydroxylated derivative of the potent anti-malarial natural product, FR-900098 (23). FrbJ is a non-heme iron and α-ketoglutarate-dependent oxygenase that catalyzes the hydroxylation of FR-900098 to yield FR-33289 (18). Fosfazinomycin is an unusual phosphonic acid natural product that contains a unique hydrazide linkage of a valine-arginine dipeptide to methyl 2-hydroxy-2-phosphonoacetic acid (Me-hPnA). FzmG is a non-heme α-ketoglutarate-dependent oxygenase encoded within the gene cluster that performs two biosynthetic steps within fosfazinomycin biosynthesis (10). First, FzmG catalyzes the oxidation of phosphonoacetaldehyde to phosphonoacetate. Second, after methylation of phosphonoacetate, FzmG catalyzes the hydroxylation of phosphonoacetate-methylester (Me-PnA) (10). Dehydrophos is a broad-spectrum Trojan horse antimicrobial produced by *Streptomyces luridus* (2, 5). A glycine-leucine dipeptide is amide-linked to O-methylated dehydroaminophosphonate (ΔAla(P)) (31), which is a potent inhibitor of

pyruvate dehydrogenase (16). This gene cluster encodes the non-heme α-ketoglutarate-dependent oxygenase, DhpA, which has been shown to catalyze the hydroxylation of 2-hydroxyethylphosphonic acid (2HEP) during the biosynthesis of phosphono-dehydroalanine (2). Hydroxyphosphonocystoximic acid and hydroxynitrilaphos are two unusual natural products produced by *Streptomyces regensis* NRRL WC-3744 (3, 12). Hydroxyphosphonocystoximic contains a thiohydroximate moiety, while hydroxynitrilaphos was the first identified nitrile-containing phosphonic acid. These metabolites contain a hydroxyl group on the carbon α- to the phosphonate moiety, which was expected to be installed via the TauD homolog, HpxV. Here we provide *in vitro* biochemical evidence demonstrating the function of HpxV.

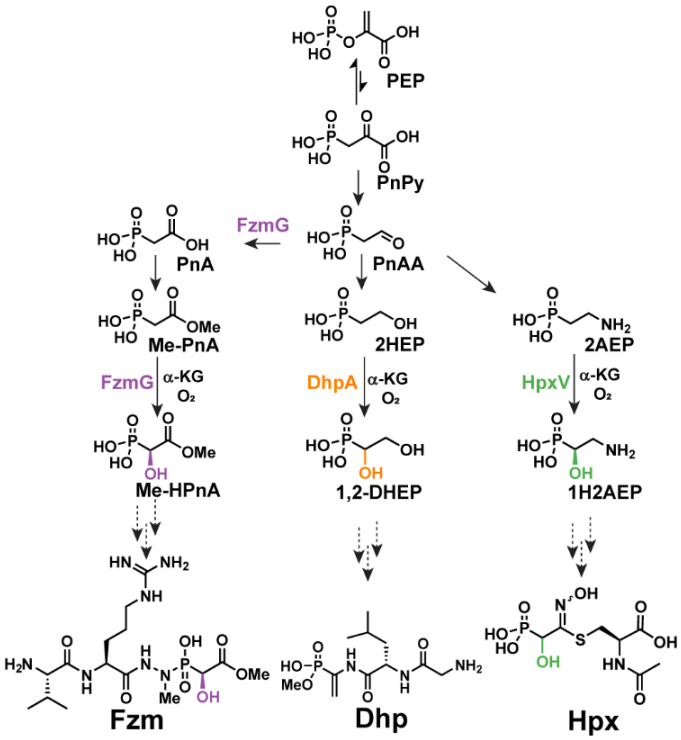


Figure 2.2. Biosynthetic routes to fosfazinomycin A (Fzm), dehydrophos (Dhp), and hydroxyphosphonocystoximic acid (Hpx) highlighting the reactions catalyzed by the TauD-family and non-heme iron, α-ketoglutarate-dependent oxygenases within each biosynthetic pathway that are part of this study.

2.1.3 DhpA, FzmG, and HpxV are evolutionarily distinct from one another

To date, only one of the non-heme iron, α-ketoglutarate-dependent oxygenase that catalyzes the hydroxylation of a phosphonic acid on the carbon α- to the phosphonate group has been structurally characterized. However, this enzyme, PhnY, is a member of the phytanoyl-CoA dioxygenase sub-family (PF05721) of non-heme iron-dependent oxygenases (21) and is only distantly related to the enzymes within this study. No TauD-family oxygenases that catalyze hydroxylation of phosphonic acids on the carbon α- to the phosphorus atom have been structurally characterized. Here, we investigate the differences between DhpA, FzmG, and HpxV that drive the differences in substrate specificity of these enzymes, which catalyze the hydroxylation of small, structurally similar phosphonic acid intermediates early in their respective biosynthetic pathways. Surprisingly, despite their ability to catalyze similar reactions, the enzymes share low amino acid identity for one another and are structurally divergent. This suggests that these enzymes have been recruited into phosphonate biosynthesis from alternative biosynthetic pathways multiple times throughout evolutionary history.

2.2 RESULTS

2.2.1 HpxV uses 2AEP as a substrate and produces (*S*)-1H2AEP

Hydroxynitrilaphos and hydroxyphosphonocystoximate both contain a hydroxylated carbon α- to the phosphorus atom (12). The logical gene to encode a protein to install this hydroxyl group is the non-heme iron and α-ketoglutarate-dependent oxygenase homolog, *hpxV*, which is not present in the organism that produces the related, non-hydroxylated natural product phosphonocystoximate (9, 12). To determine the product formed by HpxV, we incubated 6x-His-tagged HpxV with 2AEP in the presence

of α-ketoglutarate. After incubation, one product was observed (Fig 2.3-2.4), which was determined to be 1H2AEP based on spiking with a 1H2AEP standard (Fig 2.3).

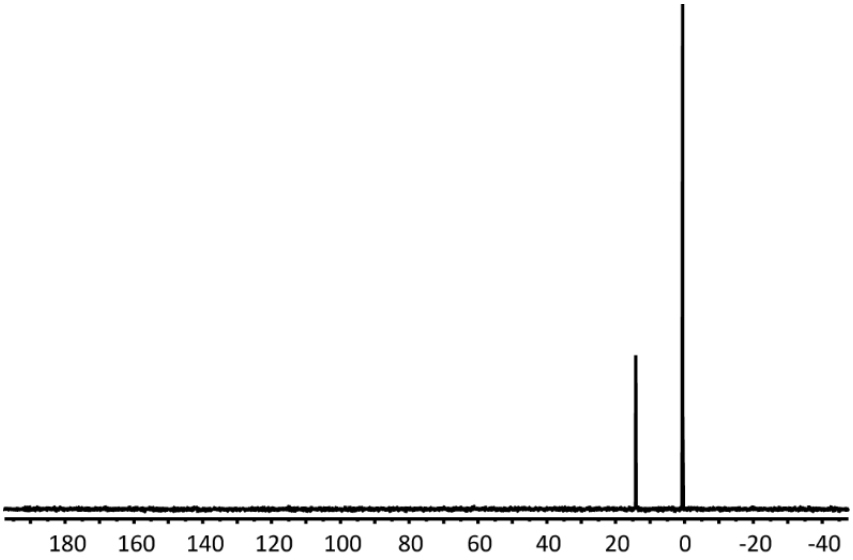


Figure 2.3. ^{31}P NMR (600 MHz) of a 12-hour HpxV enzyme reaction showing production of 1H2AEP from AEP. The peak at 0 ppm corresponds to inorganic phosphate.

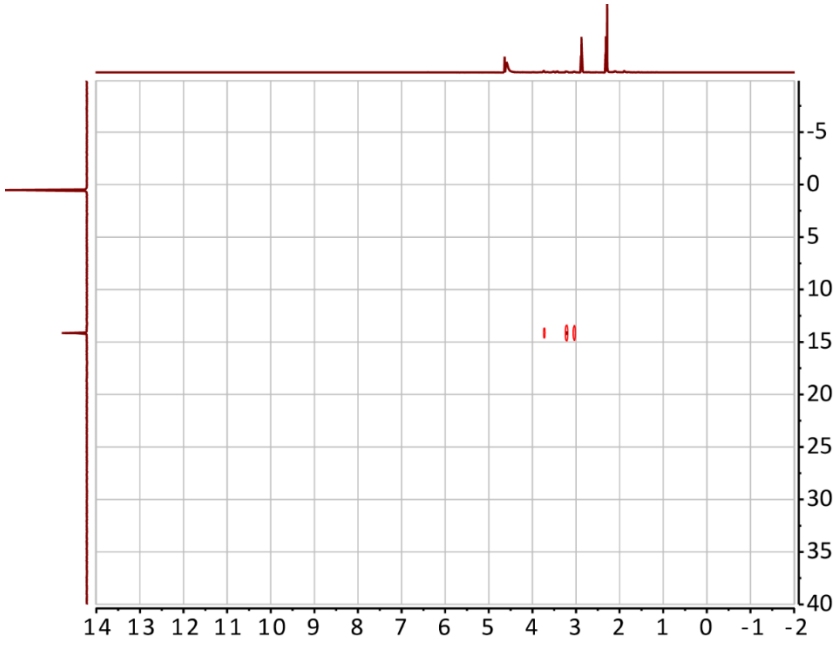


Figure 2.4. ^{31}P-^{1}H two-dimensional HMBC NMR (600 MHz) of a 12-hour HpxV enzyme reaction using 2AEP as the substrate.

To determine the absolute stereochemistry of the product formed by HpxV, we performed polarimetry to determine the specific rotation of enzyme-generated 1H2AEP. Complete turnover of 2AEP to 1H2AEP was observed at 12 hours, determined by ^{31}P NMR (Fig 2.3). The reaction product was purified using anion-exchange chromatography as previously described (7) (Fig 2.5-2.6). The specific rotation of the purified reaction product was determined to be [α] = + 32.9, which is consistent with reported values for (*S*)-1H2AEP [α] = + 31.0 (27).

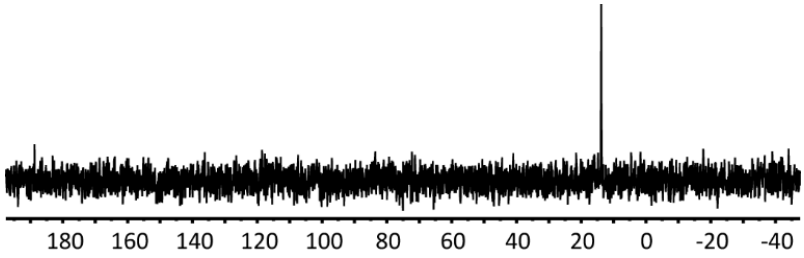


Figure 2.5. ^{31}P NMR (600 MHz) of purified 1H2AEP.

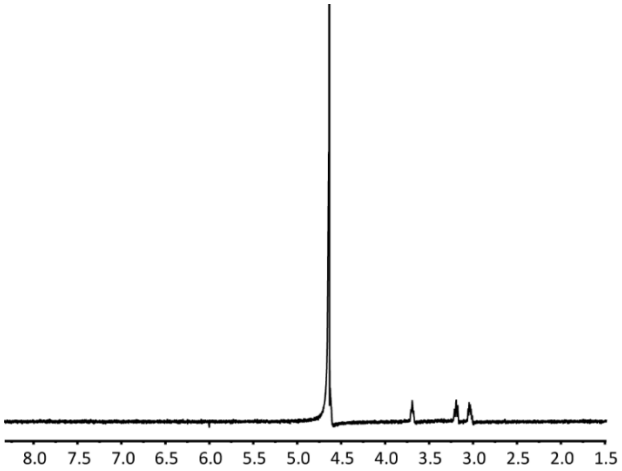


Figure 2.6. ^{1}H NMR (600 MHz) of pure 1H2AEP.

2.2.2 The oxygenases are substrate specific

A previous study has investigated the substrate specificity of FzmG and DhpA. Huang and colleagues showed that FzmG can hydroxylate both Me-PnA in addition to 2HEP, while DhpA only formed product in the presence of 2HEP (10). Knowing this, we wished to know whether DhpA and FzmG could use 2AEP, the substrate of HpxV, and if HpxV could use 2HEP or Me-PnA. We analyzed the substrate specificity of the oxygenases two different ways: (1) by monitoring oxygen consumption of the oxygenases in the presence of different phosphonic acid substrates (Fig 2.7a) and (2) by monitoring product formation of the oxygenases via ^{31}P NMR spectroscopy (Fig 2.7b).

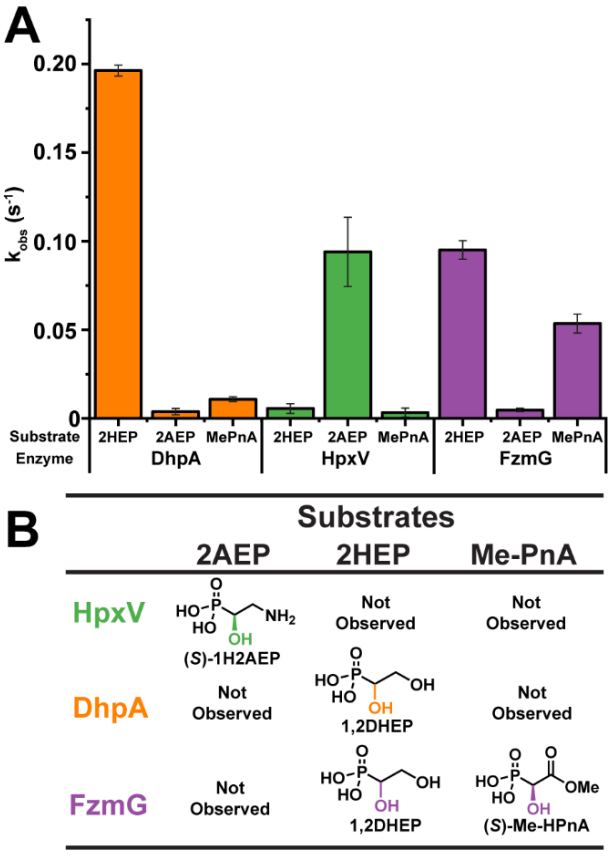


Figure 2.7. Activity assays using DhpA, FzmG, and HpxV showing that these oxygenases are substrate-specific. (A) Oxygen-consumption by the enzymes in the presence of 2HEP, Me-PnA, or 2AEP. The reactions were set up as follows in PBS buffer: 600 μM α-ketoglutarate, 500 μM ascorbic acid, 2 mM phosphonic acid substrate, and 5 μM of enzyme, with the stock reconstituted in four-times the amount of ferrous iron. (B) Product-formation observed by the enzymes in the presence of 2HEP, Me-PnA, or 2AEP. The reactions were set up with 10 mM α-ketoglutarate, 1 mM ascorbic acid, 5 mM phosphonic acid substrate, 5 μM of enzyme, and 200 μM ammonium ferrous iron sulfate in the reactions.

As shown in Fig 2.7, all three oxygenases are highly substrate specific. FzmG was confirmed to utilize 2HEP and Me-PnA as substrates (10) (Fig 2.8-2.9), and even though this enzyme can utilize four different phosphonic acids as substrates (10), FzmG was unable to catalyze product formation or consume oxygen in the presence of 2AEP (Fig 2.7 and Fig 2.10). DhpA was identified to be able to use only 2HEP by monitoring the

substrates tested, while HpxV was only able to form product in the presence of 2AEP (Fig 2.7-2.10).

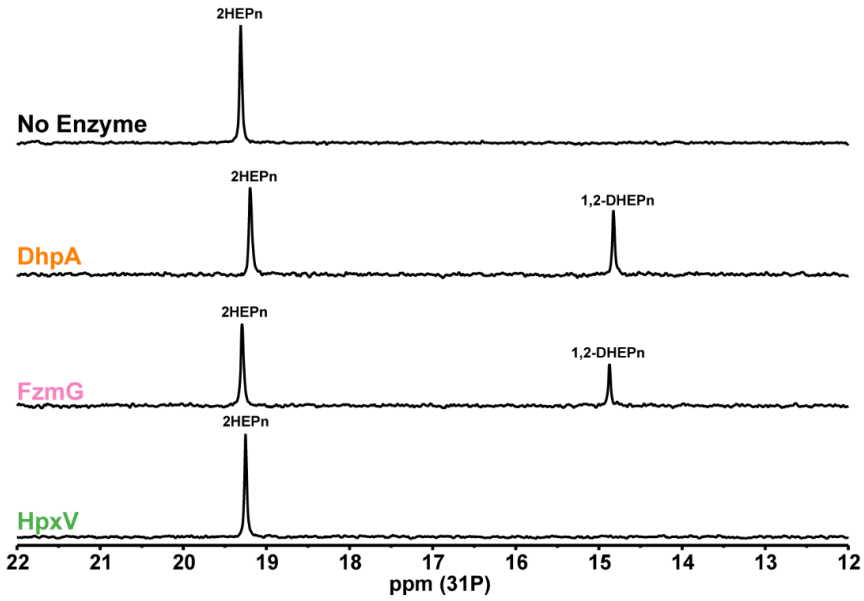


Figure 2.8. ^{31}P NMR analysis of the oxygenases when using 2HEPn as a substrate. Substrates and cofactors were added at the following concentrations in PBS buffer: 10 mM α-ketoglutarate, 1 mM ascorbic acid, 5 mM 2HEPn, 200 μM ammonium ferrous iron sulfate, and 5 μM enzyme.

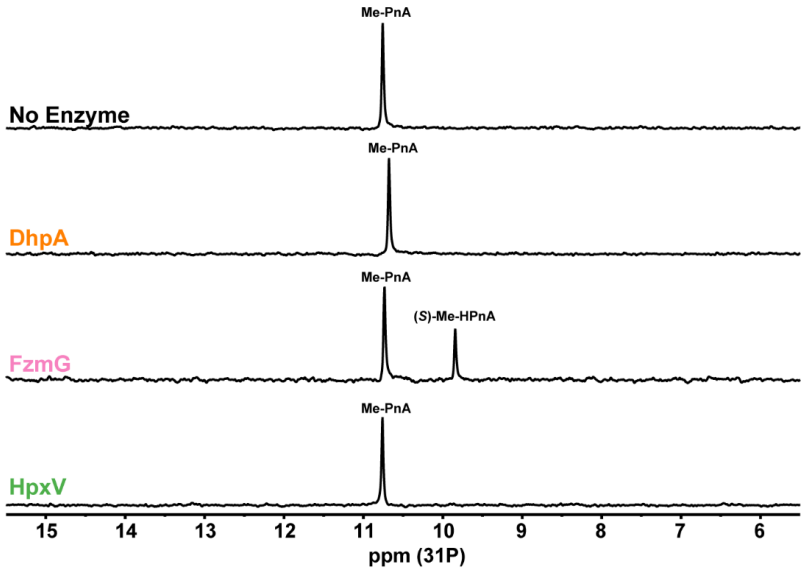


Figure 2.9. ^{31}P NMR analysis of the oxygenases when using Me-PnA as a substrate. Substrates and cofactors were added at the following concentrations in PBS buffer: 10 mM α-ketoglutarate, 1 mM ascorbic acid, 5 mM Me-PnA, 200 μM ammonium ferrous iron sulfate, and 5 μM enzyme.

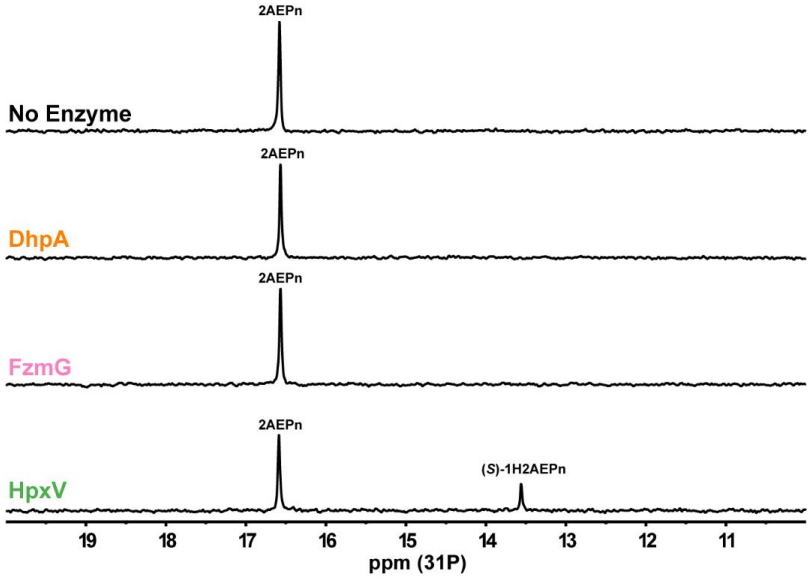


Figure 2.10. ^{31}P NMR analysis of the oxygenases when using 2AEP as a substrate. Substrates and cofactors were added at the following concentrations in PBS buffer: 10 mM α-ketoglutarate, 1 mM ascorbic acid, 5 mM 2AEP, 200 μM ammonium ferrous iron sulfate, and 5 μM enzyme.

2.2.3 HpxV and DhpA catalyze hydroxylation of their phosphonic acid substrates with similar catalytic efficiency to FzmG

Oxygen consumption by HpxV or DhpA was monitored using either 2AEP or 2HEP as substrates, respectively, to determine the kinetic parameters of the oxygenases. Oxygen consumption of DhpA in the presence of 2HEP fits a substrate inhibition model (Fig 2.11). DhpA shares almost identical kinetic parameters to FzmG when utilizing 2HEP as a substrate (Table 2.1); however, in contrast to FzmG, DhpA shows substrate inhibition in the presence of millimolar levels of substrate, which is ten-fold the K_M of the enzyme (Table 2.1) (Fig 2.11).

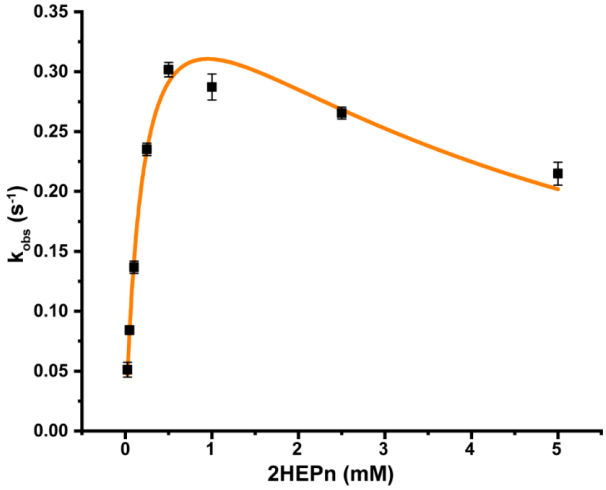


Figure 2.11. Kinetic curve for DhpA using 2HEP as a substrate. The data were fit to a substrate inhibition model. The K_M was determined to be 220 ± 30 μM. while the apparent k_{obs} was observed to be 0.45 ± 0.03 s^{-1}, which gives a catalytic efficiency DhpA for 2HEPn of 2.0 x 10^3 (M^{-1} s^{-1}). The enzyme was inhibited by substrate concentrations approximately ten times the K_M of the enzyme with a K_I of 4.2 ± 0.7 mM.

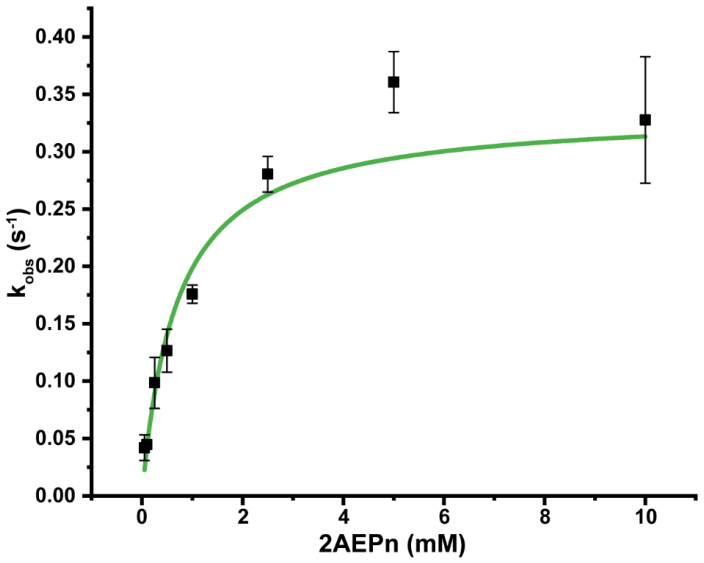


Figure 2.12. Kinetic curve for HpxV using 2-aminoethylphosphonate as a substrate. The data were fit to a Michaelis-Menten model. The K_M was determined to be 700 ± 100 μM, while the apparent k_{obs} was observed to be 0.33 ± 0.03 s^{-1}, which gives a catalytic efficiency HpxV for 2AEPn of 0.47 x 10^3 (M^{-1} s^{-1}).

Oxygen consumption of HpxV in the presence of 2AEP fits a standard Michaelis-Menten model (Fig 2.12). The K_M of HpxV was determined to be approximately three-fold higher than FzmG and DhpA for its substrate (Table 2.1).

Table 2.1. Michaelis-Menten kinetics of the oxygenases in this study determined by monitoring oxygen consumption of the enzymes. The data were fit to either a Michaelis-Menten fit (MM) or substrate inhibition model (SI). N/A indicates that substrate inhibition was not observed.

Enzyme	**Substrate**	**Kinetic Model**	k_{obs}	**Apparent** K_M	k_{cat}/K_M	K_I
			(s^{-1})	*(μM)*	*(M^{-1} s^{-1})*	*(mM)*
HpxV	2AEP	MM	0.33 ± 0.03	700 ± 100	0.4 x 10^3	N/A
DhpA	2HEP	SI	0.45 ± 0.03	220 ± 30	2.0 x 10^3	4.2 ± 0.7
FzmG*	2HEP	MM	0.35 ± 0.02	210 ± 40	1.7 x 10^3	N/A
	Me-PnA	MM	0.27 ± 0.01	80 ± 20	3.4 x 10^3	N/A

The asterisk next to FzmG indicates that these data shown are previously published by Huang, et. al (10).

2.2.4 The oxygenases share common ancestry with non-phosphinic acid oxygenases

All of the oxygenases catalyze the same reaction on similar substrates as part of different phosphonic acid biosynthetic pathways. Comparisons of the amino acid sequences of the oxygenases to one another show that they share very little similarity, with DhpA and FzmG sharing the highest similarity at 42% amino acid identity, while the similarity of all of the other proteins to one another is below 30% amino acid identity (Table 2.2). This suggests that the oxygenases are highly divergent from one another. To test this, we constructed a phylogenetic tree using homologs of the different TauD-family, non-heme iron-dependent enzymes that are known to hydroxylate a phosphonic acid substrate: DhpA, FzmG, HpxV, and FrbJ.

Table 2.2. Percent identity comparisons of the proteins in this study. The amino acid sequences were aligned using MAFFT (13) and the percent identity is shown through analysis through Geneious (14).

	HpxV	DhpA	FzmG
HpxV	100		
DhpA	24	100	
FzmG	25	42	100

This tree shows that DhpA and FzmG share a common ancestor not shared by the other oxygenases (Fig 2.13). Within this clade, we find that there are other uncharacterized oxygenases that have a PEP mutase homolog within their genomic neighborhood, suggesting that they are involved in phosphonate biosynthesis (Fig. 2.14). We can predict the substrates of these enzymes based on the presence of certain genes within the genomic context of the biosynthetic gene cluster. Gene clusters containing an iron-dependent alcohol dehydrogenase homologous to DhpC will likely use 2HEP as a substrate. Those clusters lacking a dehydrogenase homolog, but containing a SAM-

dependent methyltransferfase are hypothesized to use Me-PnA as a substrate. HpxV and other non-phosphonic acid oxygenases share a common ancestor not shared with FzmG and DhpA (Fig 2.13 - 2.14). The most closely related homologs are from *Pseudomonas* (Fig 2.14).

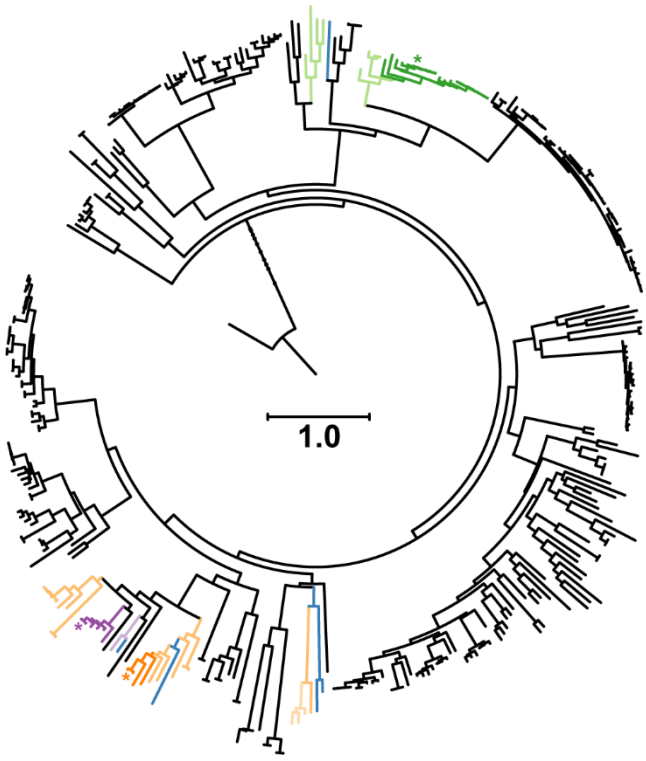


Figure 2.13. Maximum-likelihood phylogeny of homologous enzymes to DhpA, FzmG, and HpxV. The tree shown is a snapshot of a larger phylogenetic tree shown in **Figure 2.14** to show the enzymes in this study. All highlighted branches are oxygenases located within 15 genes upstream or downstream of a PEP mutase homolog, indicating the likely involvement of these enzymes in phosphonate biosynthesis. Asterisks denote the enzymes used in this study: green is HpxV, purple is FzmG, and orange is DhpA. Green branches indicate oxygenases within an Hpx gene cluster context. Light green branches have an oxygenase with a 2AEP aminotransferase homolog within the genomic context. Orange branches indicate oxygenases within a Dhp gene cluster context. Light orange branches have an oxygenase with an iron-dependent dehydrogenase homolog within the genomic context. Purple branches indicate oxygenases within an Fzm gene cluster context. Light purple branches have an oxygenase with a SAM-dependent methyltransferase homolog within the genomic context. Blue branches indicate oxygenases that have no enzyme homologous to the enzymes needed to generate 2AEP, 2HEP, or Me-PnA within their gene clusters.

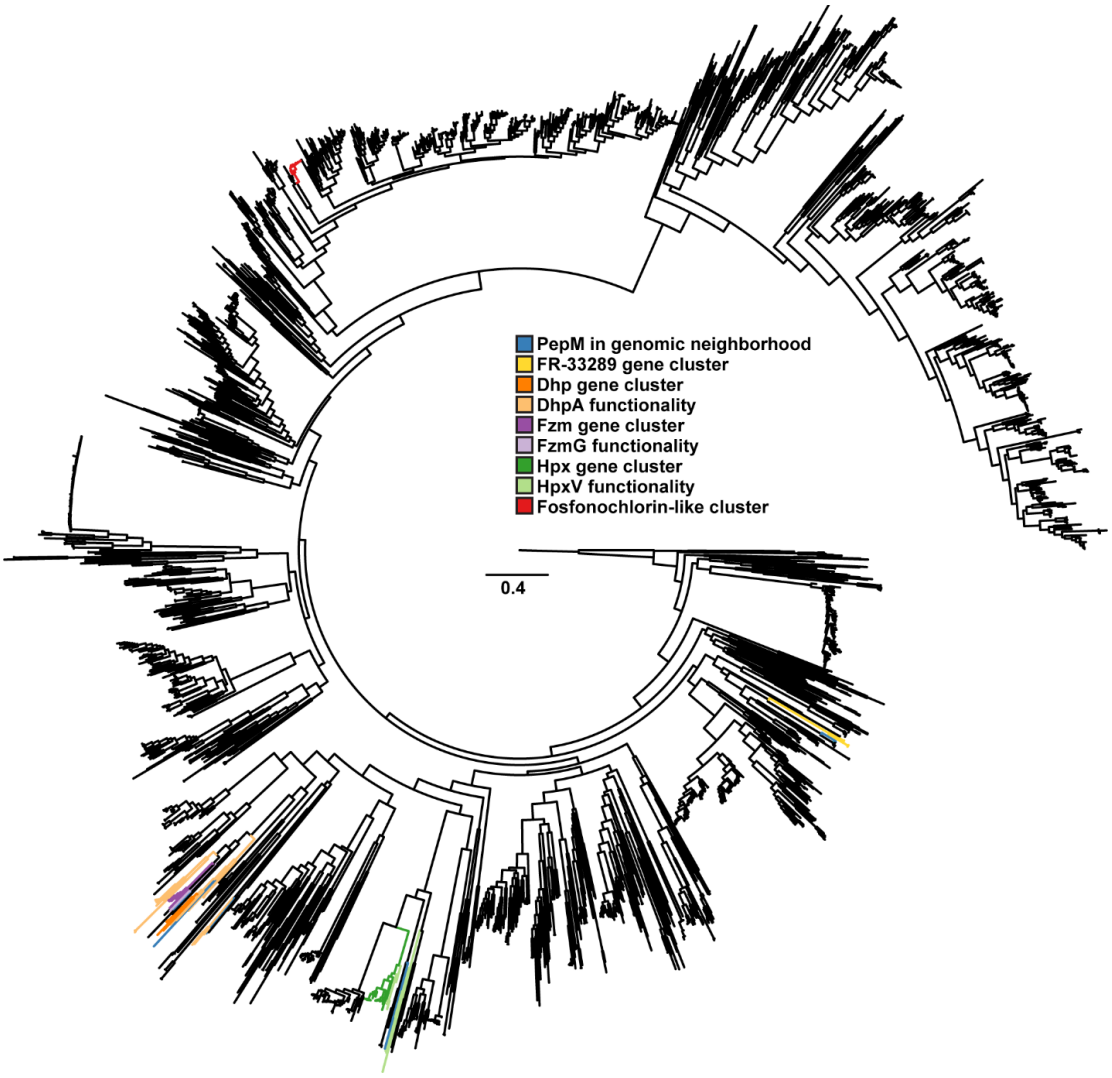


Figure 2.14. Full maximum-likelihood phylogeny of homologous enzymes to DhpA, FzmG, HpxV, and FrbJ that is showing in Fig 2.13. Highlighted branches are TauD-homologs located within 15 genes upstream or downstream of a PEP mutase homolog. The colors of the branches represent different functions of the enzymes shown in the key.

2.3 DISCUSSION

2.3.1 DhpA and HpxV are substrate specific, while FzmG shows substrate promiscuity

DhpA and HpxV only utilize 2HEP and 2AEP, respectively, which are phosphonic acids that are part of dehydrophos and hydroxyphosphonocystoximic acid biosynthesis. Oxygen-consumption and product formation experiments suggest that HpxV utilizes only 2AEP as a substrate, whereas DhpA uses 2HEP. The enzymes cannot use substrates that are not directly involved in the early biosynthetic steps of the pathways of these natural products. This may provide a biosynthetic check-point that helps to ensure the correct natural product is being produced. For example, during dehydrophos biosynthesis, hydroxylation of 2HEP is required for the downstream formation of *L*-aminoethylphosphonic acid, which is the substrate for oxidation to the vinyl functional group (2). Without hydroxylation of 2HEP, the incorrect natural product might be produced. In hydroxyphosphonocystoximic acid biosynthesis, the enzyme following HpxV shows substrate tolerance (9), so it is possible that non-hydroxylated products may be formed *in vivo*.

In contrast, FzmG tolerates different substrates. Previous studies have identified that FzmG can perform two different steps in fosfazinomycin biosynthesis (10), catalyzing the oxidation of an aldehyde to a carboxylic acid and hydroxylation of the carbon alpha to the phosphonate moiety. Therefore, it was unsurprising to us that FzmG catalyzes the hydroxylation of 2HEP as well (10), considering the structural similarity between phosphonoacetaldehyde and 2HEP. FzmG utilizes Me-PnA with the highest

catalytic efficiency out of all the phosphonic acid substrates tested (10), suggesting this is the *in vivo* substrate of the enzyme.

2.3.2 The FzmG and DhpA share a common ancestor not shared with HpxV suggesting that the oxygenases were recruited into phosphonate biosynthesis multiple times

The phylogeny of FzmG is fully consistent with the observation that the enzyme catalyze the oxidation of 2HEP. DhpA and FzmG share a common ancestor not shared with HpxV. This clade appears to have evolved the function of oxidizing 2HEP, as most of the gene clusters have an iron-dependent dehydrogenase homolog that catalyzes the formation of 2HEP from phosphonoacetaldehyde (Fig 2.13) (26). Within this clade, there are a number of proteins that do not appear to be involved in phosphonate biosynthesis because those organisms do not contain a PEP mutase homolog within their genomes. Interestingly, the enzymes that appear to use 2AEP as a substrate all share a common ancestor not shared with FzmG or DhpA (Fig 2.13). Again, this is consistent with the observation that FzmG and DhpA cannot utilize 2AEP as a substrate, whereas HpxV is unable to use Me-PnA or 2HEP, the substrates of FzmG and DhpA, as substrates (Fig 2.7). Combined with the low sequence identity between these proteins (Table 2.2), it is unsurprising to see these enzymes appear to show different ancestry within the oxygenase tree.

HpxV, DhpA, and FzmG all function early within their respective phosphonate biosynthetic pathways, use small, structurally similar phosphonic acid substrates, and catalyze the hydroxylation on the same part of each molecule. Therefore, we suspected that these enzymes would have a similar evolutionary lineage, with all the phosphonate-

specific oxygenases having evolved from one common ancestor. To our surprise, we saw numerous non-phosphonate utilizing homologs within the same clades as the phosphonate-specific enzymes. The homologs of HpxV share common ancestor with non-phosphonic acid producing pseudomonads, as the entire sister clade next to those with a hydroxyphosphonocystoximic acid biosynthetic genomic context are homologous enzymes from *Pseudomonas* species. In contrast, DhpA and FzmG fall within a clade comprised of only actinobacteria.

2.3.3 The phosphonate-specific TauD homologs fall within an uncharacterized sub-group of TauD-family enzymes

The TauD-family of dioxygenases remains largely unexplored enzyme family with a wide variety of potential substrates and functions. The 2601 proteins that make up the clade that include the phosphonate-specific oxygenases have no characterized relatives in the Swiss-Prot database (22). The closest characterized protein is CsiD (A7ZQ90.1), which is a putative gamma-butyrobetaine dioxygenase from *E. coli*, which shares 22%, 29%, and 39% amino acid identity with HpxV, DhpA, and FzmG respectively. This protein is not within the top BLAST retrieval for any of the oxygenases when a BLAST search is performed. This suggests that the phosphonate-specific proteins on this tree represent the first characterized protein sequences from this divergent group of proteins.

2.4 SUMMARY AND OUTLOOK

Hydroxylation is an important part of functionalization of different natural products. Here, we show that phosphonate biosynthesis is no exception. There are many different phosphonate biosynthetic pathways that are predicted to involve the

hydroxylation of a small phosphonic acid substrate. Based on phylogenetic analysis, we can predict the substrates of these enzymes by what common ancestor is shared with the enzymes of interest. RODEO (29) analysis of the different gene clusters containing an HpxV, FzmG, or DhpA homologs suggests there multiple phosphonic acids that will involve the hydroxylation of 2HEP, although the final product is not dehydrophos based on differences within the genomic context. There is clade containing two FzmG homologs that contain a methyltransferase similar to FzmB (10), although the gene cluster does not appear to make fosfazinomycin. These two natural product gene clusters likely produce Me-HPnA. There are some natural product gene clusters that likely will form 1H2AEP as a final product based on enzyme homology to HpxV and the gene clusters contain a 2AEP aminotransferase homolog, although the final natural product is unknown. There are only six natural product gene clusters in which we cannot predict the substrate for the TauD-family oxygenase within the gene cluster. This suggests that based on genomic context, we can predict the likely substrates for most phosphonate-specific oxygenases.

2.5 MATERIALS AND METHODS

2.5.1 Protein overexpression and purification

Plasmids containing *dhpA* (4), *fzmG* (10), or *hpxV* (9) were transformed into *E. coli* Rosetta DE3 cells. Cells were grown in Luria Broth at 37°C until the cultures reached optical density (OD_{600}) 0.6-0.8. The cultures were placed on ice for 30 minutes and protein expression was induced upon addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were shaken at 18°C for an additional 16 hours after induction. Cells were collected via centrifugation at

4500 x g for 15 minutes and the cell pellets were resuspended in ice cold cell lysis buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), 10% glycerol and 20 mM imidazole, pH 8.0) containing 1 mg/mL lysozyme, 100 units DNase I, and 100 μg RNase A, and were incubated for 30 minutes at 4°C. The cells were lysed upon passage through a French Press cell disruptor (Thermo Electron Corporation, Waltham, MA) two times at 1000 psi to lyse cells and the suspension was centrifuged at 17,500 x g for 30 minutes at 4°C. The soluble fraction was mixed with 5 mL of Ni-NTA resin and incubated for 30 minutes at 4°C and the flow through containing unbound protein was discarded. The resin was washed with 20 column volumes of ice cold wash buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), 10% glycerol and 50 mM imidazole, pH 8.0) to elute loosely bound proteins. To elute proteins from the Ni-NTA resin, 15 mL of ice cold elution buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), 10% glycerol and 250 mM imidazole, pH 8.0) was mixed with the resin and was incubated for 15 minutes at 4°C. The flow through containing eluted proteins were concentrated in 10 kDa molecular weight Amicon® Ultra spin filters (EMD Millipore, Burlington, MA). Each protein was exchanged into protein storage buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), and 10% glycerol, pH 8.0) using a PD-10 desalting column (GE Healthcare Life Sciences, Marlborough, MA) following manufacturer instructions. Proteins were flash frozen using liquid nitrogen and stored at -80°C until use.

2.5.2 DhpA and HpxV kinetic assays

Enzyme assays were monitored using a Clark-type oxygen electrode (Hansatech Instruments, Norfolk, England). The reactions were set up with the following

components: air saturated buffer (phosphate buffered saline (PBS): 274 mM sodium chloride, 8 mM sodium phosphate dibasic, 2 mM potassium chloride, and 1.4 mM potassium phosphate monobasic, pH 7.4), 600 μM α-ketoglutarate (buffered to pH 7.5 with sodium hydroxide), 500 μM ascorbic acid, 25 μM – 10 mM substrate (buffered to pH 7.5 using sodium hydroxide), and 1 μM HpxV or 5 μM DhpA. Enzymes were incubated anaerobically with two equivalents of ammonium ferrous iron sulfate for 15 – 20 minutes prior to assays on ice. DhpA was removed from the anaerobic chamber and kept under a nitrogen atmosphere until use in each assay. Activity of DhpA was consistent for upwards of 2 hours upon removal from the anaerobic chamber. HpxV was not as stable, which was noted by a decrease in enzyme activity within 30 minutes after removal of the enzyme from the anaerobic chamber, even with the enzyme being held under a nitrogen atmosphere. Therefore, a new aliquot of HpxV from the -80°C storage stocks were reconstituted anaerobically with fresh ferrous iron sulfate every 30 minutes. The reactions were initiated upon addition of phosphonic acid substrate.

2.5.3 Product formation assays

Product formation was monitored via ^{31}P NMR spectroscopy. The reactions were set up as follows in PBS buffer: 10 mM α-ketoglutarate (buffered to pH 7.5 with sodium hydroxide), 1 mM ascorbic acid,5 mM phosphonic acid substrate (buffered to pH 7.5 using sodium hydroxide), 200 μM ammonium ferrous iron sulfate, and 5 μM enzyme. Reactions were allowed to proceed at room temperature for three hours. Post reaction, EDTA was added to the reactions to a final concentration of 9 mM to chelate the iron. Enzymes were removed from the reactions using a 30 kDa molecular weight Amicon®

Ultra spin filter (EMD Millipore, Burlington, MA) and product formation was determined using ^{31}P NMR spectroscopy.

2.5.4 Oxygen consumption assays

Oxygen consumption of the oxygenases in the presence either 2AEP, Me-PnA, or 2HEP was monitored using a Clark-type oxygen electrode. The reactions were set up as follows in PBS buffer: 10 mM α-ketoglutarate (buffered to pH 7.5 with sodium hydroxide), 1 mM ascorbic acid, 5 mM phosphonic acid substrate (buffered to pH 7.5 using sodium hydroxide), 200 μM ammonium ferrous iron sulfate, and 5 μM enzyme. Reactions were initiated using substrate and the rate of oxygen consumption was measured until completion. The first ten percent of the oxygen-consumption curve was measured in order to obtain steady-state kinetic data on each enzyme. Enzyme reactions that showed a complete cessation of oxygen consumption upon addition of substrate are marked with an asterisk.

2.5.5 Stereochemistry determination of (*S*)-1H2AEP produced by HpxV

(*S*)-1H2AEP was generated using HpxV, the 2AEP-dependent α-ketoglutarate-dependent dioxygenase from *S. regensis* NRRL WC-3744. The reactions contained 50 mM phosphate buffered saline, pH 7.4, 40 mM α-ketoglutarate, 1 mM iron sulfate, 7.5 mM 2AEP, 4 mM L-ascrobic acid, and 200 μM HpxV. The reaction was quenched using 5 reaction equivalents of 100% methanol. Precipitated enzyme was removed by centrifugation and methanol was removed under reduced pressure rotatory evaporation. (*S*)-1H2AEP was purified using weak anion exchange chromatography as previously described (7). Sample concentration of purified material was determined using ^{31}P-NMR, with dimethylphosphinic acid as an internal reference standard. The optical rotation of

the purified product formed via this enzymatic reaction was determined using a Jasco P-2000 digital polarimeter. A sodium lamp was used as the light source (589nm) and was measured through a 5 cm flint glass cell. The value obtained was compared to recorded literature values to determine the correct stereoisomer (27).

2.5.6 Construction of phylogenetic trees

The construct the oxygenase phylogenetic tree, a PSI-BLAST was performed (1) using DhpA (ACZ13452.1), HpxV (WP_030990672.1), FzmG (WP_053787798.1), and FrbJ (ABB90399.1) as the queries were retrieved on March 26, 2018. One iteration was run for each protein and all of the accession numbers between the query protein sequence and the most distantly related phosphonate homolog (FrbJ for DhpA, FzmG, and HpxV and FzmG for FrbJ) were collected. *Escherichia coli* MG 1655 TauD (Accession number) was used as the outgroup in this tree because it was not within the PSI-BLAST retrieval and it has a characterized function. Duplicate accession numbers were removed and the remaining 2601 amino acid sequences were aligned using MAFFT (default settings) (13). A maximum-likelihood phylogenetic tree was generated using FastTree (24). The accession numbers for each amino acid sequence was analyzed using RODEO (29), and if a PEP mutase homolog was found within 15 genes upstream or downstream of the oxygenase, these branches were highlighted on the tree, indicating the likelihood that this protein is involved in phosphonate biosynthesis. Green branches are HpxV homologs that share a phosphonocystoximic acid genomic context. Light green branches indicate that the phosphonate gene cluster associated with the HpxV homolog contains a 2AEP aminotransferase homolog within the gene cluster, but is not a phosphonocystoximic acid biosynthetic gene cluster. Orange highlighting indicates a

DhpA homolog in the genomic context of the dehydrophos gene cluster. Light orange highlights indicate a DhpA homolog that is not in a dehydrophos biosynthetic gene cluster, but contains an iron-dependent dehydrogenase homolog within the gene cluster. Purple highlighting denotes a FzmG homolog in a fosfazinomycin genomic context, whereas purple highlighting indicates a FzmG homolog within a genomic context that does not contain an iron-dependent dehydrogenase, but does contain a homologous enzyme to FzmB, the SAM-dependent methyltransferase responsible for methylation of phosphonoacetate. Yellow highlighting of the branches indicates FrbJ homologs in the genomic context of FR-33289 biosynthetic genes (11). Red highlights predict a TauD-homolog in the genomic context of fosfonochlorin biosynthesis (28). Finally, blue branches denote TauD-family oxygenases in a phosphonic acid biosynthetic genomic context with no known phosphonic acid structure known, nor contain any homologous enzymes to the ones listed above for other highlighting. The Interactive Tree of Life (ITOL) (17) was used to trim the 2601-member tree to show the clade (342 proteins) containing the oxygenases that are the focus of this study (Fig 2.13 – Fig 2.14).

2.6 REFERENCES

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215(3):403–10.

2. Bougioukou D, Mukherjee S, van der Donk W. 2013. Revisiting the biosynthesis of dehydrophos reveals a tRNA-dependent pathway. *Proc. Natl. Acad. Sci.* 110(27):10952–57.

3. Cioni JP, Doroghazi JR, Ju K-S, Yu X, Evans BS, et al. 2014. Cyanohydrin Phosphonate Natural Product from Streptomyces regensis. *J. Nat. Prod.* 77(2):243–49.

4. Circello BT, Eliot AC, Lee J-H, van der Donk WA, Metcalf WW. 2010. Molecular Cloning and Heterologous Expression of the Dehydrophos Biosynthetic Gene Cluster. *Chem. Biol.* 17(4):402–11.

5. Circello BT, Miller CG, Lee J-H, Donk WA van der, Metcalf WW. 2011. The Antibiotic Dehydrophos Is Converted to a Toxic Pyruvate Analog by Peptide Bond Cleavage in Salmonella enterica. *Antimicrob. Agents Chemother.* 55(7):3357–62.

6. Eichhorn E, Ploeg JR van der, Kertesz MA, Leisinger T. 1997. Characterization of α-Ketoglutarate-dependent Taurine Dioxygenase from Escherichia coli. *J. Biol. Chem.* 272(37):23031–36.

7. Evans BS, Zhao C, Gao J, Evans CM, Ju K-S, et al. 2013. Discovery of the Antibiotic Phosacetamycin via a New Mass Spectrometry-Based Method for Phosphonic Acid Detection. *ACS Chem. Biol.* 8(5):908–13.

8. Friese SJ, Kucera BE, Que Lawrence, Tolman WB. 2006. Self-Assembly of the 2-His-1-carboxylate Facial Triad in Mononuclear Iron(II) and Zinc(II) Models of Metalloenzyme Active Sites. *Inorg. Chem.* 45(20):8003–5.

9. Goettge MN, Cioni JP, Ju K-S, Pallitsch K, Metcalf WW. 2018. PcxL and HpxL are flavin-dependent, oxime-forming N-oxidases in phosphonocystoximic acid biosynthesis in Streptomyces. *J. Biol. Chem.* 293(18):6859–68.

10. Huang Z, Wang K-KA, Lee J, Donk WA van der. 2015. Biosynthesis of fosfazinomycin is a convergent process. *Chem. Sci.* 6(2):1282–87.

11. Johannes TW, DeSieno MA, Griffin BM, Thomas PM, Kelleher NL, et al. 2010. Deciphering the Late Biosynthetic Steps of Antimalarial Compound FR-900098. *Chem. Biol.* 17(1):57–64.

12. Ju K-S, Gao J, Doroghazi JR, Wang K-KA, Thibodeaux CJ, et al. 2015. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc. Natl. Acad. Sci. U. S. A.* 112(39):12175–80.

13. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30(4):772–80.

14. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12):1647–49.

15. Koehntop KD, Emerson JP, Que L. 2005. The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J. Biol. Inorg. Chem. JBIC Publ. Soc. Biol. Inorg. Chem.* 10(2):87–93.

16. Kuemin M, van der Donk WA. 2010. Structure-activity relationships of the phosphonate antibiotic dehydrophos. *Chem. Commun. Camb. Engl.* 46(41):7694–96.

17. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44(W1):W242-245.

18. Li C, Junaid M, Almuqri EA, Hao S, Zhang H. 2016. Structural analysis of a phosphonate hydroxylase with an access tunnel at the back of the active site. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* 72(Pt 5):362–68.

19. Martinez S, Hausinger RP. 2015. Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-dependent Oxygenases. *J. Biol. Chem.* 290(34):20702–11.

20. McCusker KP, Klinman JP. 2009. Modular behavior of tauD provides insight into the origin of specificity in α-ketoglutarate-dependent nonheme iron oxygenases. *Proc. Natl. Acad. Sci. U. S. A.* 106(47):19791–95.

21. McSorley FR, Wyatt PB, Martinez A, DeLong EF, Hove-Jensen B, Zechel DL. 2012. PhnY and PhnZ Comprise a New Oxidative Pathway for Enzymatic Cleavage of a Carbon–Phosphorus Bond. *J. Am. Chem. Soc.* 134(20):8364–67.

22. NCBI. *Basic Local Alignment Search Tool*. http://blast.ncbi.nlm.nih.gov/Blast.cgi.

23. Okuhara M, Kuroda Y, Goto T, Okamoto M, Terano H, et al. 1980. Studies on new phosphonic acid antibiotics. III. Isolation and characterization of FR-31564, FR-32863 and FR-33289. *J. Antibiot. (Tokyo)*. 33(1):24–28.

24. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE*. 5(3):e9490.

25. Purpero V, Moran GR. 2007. The diverse and pervasive chemistries of the alpha-keto acid dependent enzymes. *J. Biol. Inorg. Chem. JBIC Publ. Soc. Biol. Inorg. Chem.* 12(5):587–601.

26. Shao Z, Blodgett JAV, Circello BT, Eliot AC, Woodyer R, et al. 2008. Biosynthesis of 2-Hydroxyethylphosphonate, an Unexpected Intermediate Common to Multiple Phosphonate Biosynthetic Pathways. *J. Biol. Chem.* 283(34):23161–68.

27. Staalduinen LM van, McSorley FR, Schiessl K, Séguin J, Wyatt PB, et al. 2014. Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates. *Proc. Natl. Acad. Sci.* 111(14):5171–76.

28. Takeuchi M, Nakajima M, Ogita T, Inukai M, Kodama K, et al. 1989. Fosfonochlorin, a new antibiotic with spheroplast forming activity. *J. Antibiot. (Tokyo)*. 42(2):198–205.

29. Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, et al. 2017. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* 13(5):470-478.

30. White MD, Flashman E. 2016. Catalytic strategies of the non-heme iron dependent oxygenases and their roles in plant biology. *Curr. Opin. Chem. Biol.* 31:126–35.

31. Whitteck JT, Ni W, Griffin BM, Eliot AC, Thomas PM, et al. 2007. Reassignment of the Structure of the Antibiotic A53868 Reveals an Unusual Amino Dehydrophosphonic Acid. *Angew. Chem. Int. Ed.* 46(47):9089–92.

32. Wu L-F, Meng S, Tang G-L. 2016. Ferrous iron and α-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products. *Biochim. Biophys. Acta BBA - Proteins Proteomics*. 1864(5):453–70.

Chapter 3: Biochemical characterization of a novel class of flavin-dependent, oxime-forming *N*-oxidases involved in the biosynthesis of phosphonocystoximic acids[1]

3.1 INTRODUCTION

3.1.1 Oxime-containing natural products

Oxime-containing small molecules often possess useful biological properties, including antibacterial (25), antifungal (20), insecticidal (5, 33), antitumor (18), immunosuppressive (28), and acetylcholinesterase reactivation activities (15, 46). Some of these bioactive oximes are natural products, with the most commonly identified being the glucosinolates (Fig 1). These amino acid-derived natural products, produced by plants of the order Capparales, contain a sulfonated oxime responsible for the distinctive flavor and odor of many cruciferous vegetables (17). Other natural product oximes include the vibralactoximes, which are biosynthetic derivatives of the lipase inhibitor vibralactone, produced by the fungus *Boreostereum vibrans* (10), the DNA methyltransferase inhibitor, psammaplin A, produced by the sea sponge *Psammaplinaplysilla* (4), and the β-lactam antibiotic, nocardicin A, produced by the actinobacterium *Nocardia uniformis* subsp. *tsuyamanesis* (25, 26) (Fig 3.1).

[1] Most of this chapter was reprinted from Goettge, M. N., Cioni, J. P., Ju, K.-S., Pallitsch, K. & Metcalf, W. W. PcxL and HpxL are flavin-dependent, oxime-forming *N*-oxidases in phosphonocystoximic acid biosynthesis in *Streptomyces*. J. Biol. Chem. 293, 6859–6868 (2018). This work was supported by NIH grant number P01 GM077596 B and the Austrian Science Fund (FWF): P27987-N28. NMR spectroscopy data collected in this study was done so at the Carl R. Woese Institute for Genomic Biology Core Facilities on a 600MHz NMR, funded by NIH grant number S10-RR028833.

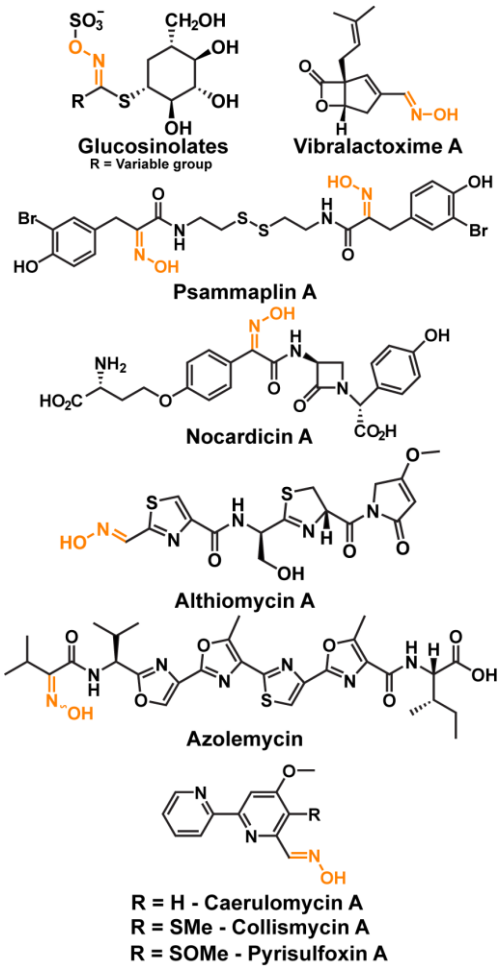


Figure 3.1. Oxime-containing natural products.

Three different enzyme families are known to produce oxime moieties: (1) cytochrome P450-dependent monooxygenases, (2) ferritin-like enzymes, and (3) flavin-dependent *N*-monooxygenases. Glucosinolates are usually made via a pathway involving a cytochrome P450-dependent monooxygenase that catalyzes successive *N*-oxidations to yield an oxime-derivative of an amino acid (13, 17); however, these molecules are probably made via an unknown FAD-NAD(P)H-dependent oxygenase in *Brassica napus* and *Brassica camestris* (6, 7). In bacteria, the oxime in nocardicin A is installed by a cytochrome P450-dependent monooxygenase, known as NocL (25, 26), whereas a metal-dependent *N*-oxygenase, AlmD, is responsible for oxime-formation in althiomycin (12).

Finally, the oxime in the ribosomally-synthesized peptide azolemycin is post-translationally installed by the flavin-dependent monooxygenase, AzmF (31), while the oxime in caerulomycin A, is introduced using a two-component flavin-dependent monooxygenase, CrmH (49).

3.1.2 Phosphonocystoximic acids

Our lab recently identified two new phosphonate natural products with an unusual thiohydroximate moiety (21). These include phosphonocystoximic acid, produced by *Streptomyces* sp. NRRL S-481 and its hydroxylated congener, hydroxyphosphonocystoximic acid, made by a number of *Streptomyces* species, including *Streptomyces regensis* NRRL WC-3744, which also makes hydroxynitrilaphos (11). Significantly, many hydroxyphosphonocystoximate producers make small amounts of phosphonocystoximic acid, as well as a suite of biosynthetic intermediates that differ solely by hydroxylation at the carbon alpha to the phosphorus atom, suggesting similar molecular machinery is used to produce both natural products (21). This idea is strongly supported by analysis of the biosynthetic genes clusters associated with production of these oxime-containing natural products.

Eighteen genes are shared between the clusters responsible for synthesis of phosphonocystoximate (the *pcx* gene cluster) and hydroxyphosphonocystoximate (the *hpx* gene cluster), based on significant amino acid identity between the encoded proteins (Fig 3.2, Table 3.1). The first nine genes (*hpxA*-*hpxI* and *pcxA*-*pcxI*) in both clusters are highly homologous and found in the same order. Nine additional genes are also shared between the two clusters, with their organization being conserved for the most part (*pcxK*-*pcxU* and *hpxK*-*hpxU*). Thus, it seems likely that most of the biosynthetic steps

required for synthesis of the two congeners are similar. However, the gene clusters also show significant differences. Key among these is a gene encoding a putative 2-oxoglutarate-dependent dioxygenase designated HpxV, found solely in *S. regensis*. Based on the presence of this gene and the observed metabolic intermediates produced by the two strains, a scheme for biosynthesis of phosphonocystoximate and hydroxyphosphonocystoximate biosynthesis was proposed (Fig 3.2) (21). The proposed pathways diverge at an early step, with hydroxylation of 2-aminoethylphosphonic acid (2AEPn) by HpxV yielding 1-hydroxy-2AEPn (1H2AEPn) solely in *S. regensis*. Consistent with this model, we recently showed that purified HpxV catalyzes this reaction, generating the *S*-enantiomer of 1H2AEPn (**Chapter 2**). The mechanism of oxime formation in the two pathways, however, remains a mystery because homologs of known oxime-forming enzymes are not found in either biosynthetic gene cluster. Nevertheless, putative NAD(P)H-dependent and flavin-dependent oxygenases produced by the *pcxL* and *hpxL* genes, in *Streptomyces* sp. NRRL S-481 and *S. regensis*, respectively, were suggested as possible candidates for this enzymatic activity (21).

Here we present *in vitro* biochemical evidence that supports this proposal. The data reveal PcxL and HpxL as the founding members of a novel family of oxime-forming enzymes required for production of the hydroxylated and non-hydroxylated forms of thiohydroximate phosphonic acid natural products.

Table 3.1. Comparison of the phosphonocystoximate (Pcx) and hydroxyphosphonocystoximate (Hpx) gene clusters. Genes with the same corresponding fourth letter of the gene are homologous to one another. Percent amino acid identity between the proteins is given based on BLAST. NP = not present within the gene cluster and N/A= not applicable

Pcx Gene	Size (#AA)	Accession Number	Hpx Gene	Size (#AA)	Accession Number	%Identity	Putative Function / Description	Pfam
pcxA	404	WP_030230281.1	*hpxA*	409	WP_030990657.1	57%	Helix-turn-helix domain transcriptional regulator	PF13560
pcxB	60	WP_030230290.1	*hpxB*	201	WP_030646389.1		Hypothetical protein	No Match
pcxC	412	WP_051704820.1	*hpxC*	414	WP_030990660.1	64%	ATP-grasp domain containing protein	PF13535
pcxD	413	WP_051704821.1	*hpxD*	415	WP_048910652.1	70%	ATP-grasp domain containing protein	PF13535
pcxE	484	WP_030230301.1	*hpxE*	456	WP_030990665.1	58%	Condensation domain containing protein	PF00668
pcxF	426	WP_051704822.1	*hpxF*	422	WP_048910653.1	76%	Biotin carboxylase / ATP-grasp domain	PF13535
pcxG	418	WP_030230306.1	*hpxG*	419	WP_030990668.1	70%	Major facilitator superfamily transporter	PF07690
pcxH	434	WP_030230310.1	*hpxH*	447	WP_030646400.1	63%	Peptidase	PF01523
pcxI	451	WP_051704823.1	*hpxI*	447	WP_048910654.1	77%	Peptidase	PF01523
pcxJ	213	WP_030230320.1	NP	N/A	N/A	N/A	Pyridoxamine 5'-phosphate oxidase	PF10590
pcxK	237	WP_030230324.1	*hpxK*	241	WP_030646417.1	58%	Mycothiol-dependent maleylpyruvate isomerase	PF11716
pcxL	742	WP_051704824.1	*hpxL*	750	WP_030990682.1	61%	NAD_binding_9 FAD-NAD(P)-binding	PF13454
pcxM	363	WP_037834389.1	*hpxM*	375	WP_030990679.1	66%	Aminotransferase class-V	PF00266
pcxN	307	WP_030230335.1	*hpxN*	290	WP_030990678.1	55%	Fosfomycin resistance kinase FomA	PF00696
pcxO	215	WP_030230340.1	NP	N/A	N/A	N/A	Thymidylate kinase	PF02223
pcxP	90	WP_030230342.1	NP	N/A	N/A	N/A	Phosphopantetheine-binding protein	PF00550
pcxQ	371	WP_037834390.1	*hpxQ*	395	WP_030990677.1	66%	Uncharacterized nucleotidyltransferase	PF14907
pcxR	549	WP_030230346.1	*hpxR*	563	WP_030990676.1	58%	Non-ribosomal peptide synthetase adenylation domain	PF00501
pcxS	353	WP_030230349.1	*hpxS*	349	WP_030990674.1	72%	2,3-diaminopropionate biosynthesis protein SbnB	PF02423
pcxT	322	WP_037834298.1	*hpxT*	324	WP_030646404.1	67%	2,3-diaminopropionate biosynthesis protein SbnA	PF00291
pcxU	286	WP_030230355.1	*hpxU*	286	WP_030646403.1	72%	Phosphoenolpyruvate phosphomutase	PF13714
NP	N/A	N/A	*hpxV*	322	WP_030990672.1	N/A	TauD taurine catabolism dioxygenase	PF02668
NP	N/A	N/A	*hpxW1*	167	WP_030646419.1	N/A	Phosphonopyruvate decarboxylase alpha subunit	PF02776
NP	N/A	N/A	*hpxW2*	182	WP_030646420.1	N/A	Phosphonopyruvate decarboxylase beta subunit	PF02775
NP	N/A	N/A	*hpxX*	86	WP_030990685.1	N/A	Acyl carrier protein	PF00550
NP	N/A	N/A	*hxpY*	337	WP_030990687.1	N/A	FomB family phosphonate monophosphate kinase	PF13481
NP	N/A	N/A	*hpxZ*	284	WP_030646426.1	N/A	EamA-like transporter family protein	PF00892
orf1	316	WP_037834300.1	NP	N/A	N/A	N/A	3-phosphoglycerate dehydrogenase	PF02826
orf2	367	WP_051704838.1	NP	N/A	N/A	N/A	L-iditol 2-dehydrogenase	PF08240
orf3	365	WP_030230374.1	NP	N/A	N/A	N/A	N-acetyl-gamma-glutamyl-phosphate reductase	PF01118
orf4	418	WP_051704825.1	NP	N/A	N/A	N/A	Phosphoglycerate kinase	PF00162
orf5	373	WP_030230380.1	NP	N/A	N/A	N/A	1,3-propanediol dehydrogenase	PF00465
orf6	202	WP_030230383.1	NP	N/A	N/A	N/A	DUF402 containing protein	PF04167

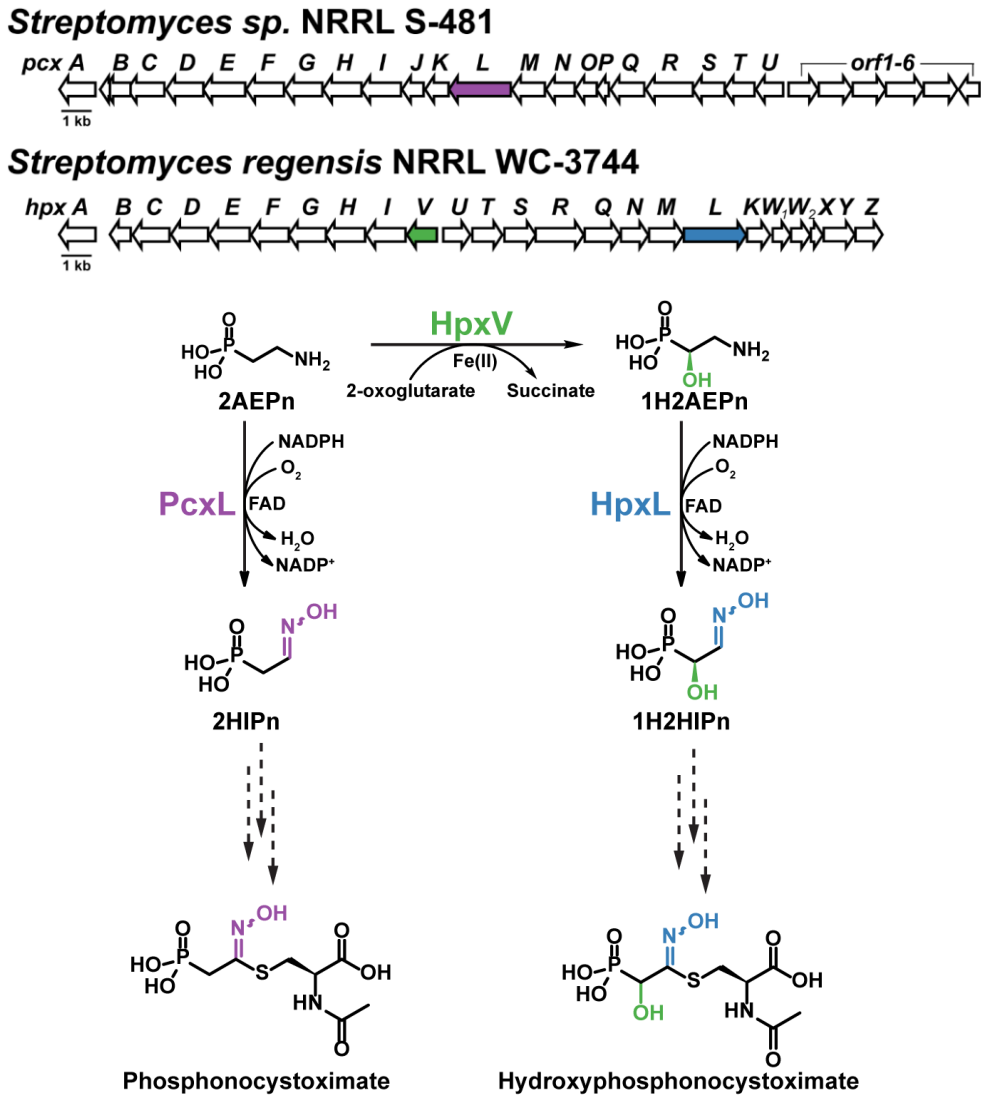


Figure 3.2. Top: The putative biosynthetic gene clusters for phosphonocystoximate and hydroxyphosphonocystoximate, found in *Streptomyces* sp. NRRL S-481 and *Streptomyces regensis* NRRL WC-3744, respectively. Putative gene functions are shown in Table 3.1. Bottom: The proposed biosynthetic route of oxime formation in phosphonocystoximates. HpxV is a 2-oxoglutarate- and ferrous iron-dependent 2AEPn dioxygenase that converts 2AEPn to (*S*)-1H2AEPn. PcxL and HpxL are FAD- and NADPH-dependent amine oxidases that catalyze 2AEPn or 1H2AEPn oxidation to yield 2-hydroxyiminoethylphosphonic acid (**1**) or 1-hydroxy-2-hydroxyiminoethylphosphonic acid (**3**), respectively. Multiple steps are required to convert these oxime-containing intermediates to the final products (hyphenated arrows).

3.2 RESULTS AND DISCUSSION

3.2.1 Bioinformatic analysis suggest HpxL and PcxL are novel oxime-forming amine oxygenases

Several lines of bioinformatics evidence suggest that the PcxL and HpxL proteins are responsible for installation of the oxime moiety in the phosphonocystoximic acids produced by *Streptomyces sp.* S-481 and *S. regensis*. First, conserved FAD- and NAD(P)-binding domains are observed in both proteins, consistent with a role in redox chemistry. Second, both have structural similarity to the *N*-terminal domain of *L*-ornithine *N*-hydroxylase from *Pseudomonas* (PDB entry 3S5W) and *Kutzneria* (PDB entry 4TLX) based on the protein-fold recognition algorithm, Phyre2(24). Taken together, these data suggest that PcxL and HpxL may catalyze FAD- and NAD(P)-dependent amine oxidation. Additional support for the proposed amine oxidase function is provided by their weak similarity (~35% identity) to FAD-NAD(P)H-dependent enzymes involved in the biosynthetic pathways of cremeomycin (42) and fosfazinomycin (19) (Table 1). These enzymes oxidize the primary amine of aspartic acid to form nitrosuccinic acid, which is believed to serve as the substrate for downstream N-N bond formation (42). Finally, the existence of the characterized oxime-forming FAD- and NAD(P)H-dependent amine monooxygenases in the biosynthetic pathways of azolemycin (AzmF) (31) and caerulomycin A (CrmH) (49) is also consistent with the proposed amine oxidase function for PcxL and HpxL. However, it should be stressed that AzmF and CrmH are not related to PcxL and HpxL based on comparison of their primary amino acid sequences (Table 3.2). Moreover, these proteins are significantly smaller than PcxL and HpxL (389/407 amino acids versus 742/750 amino acids). Similarly, the human liver flavin-containing

monooxygenases FMO1 (NCBI accession number NP_001269621.1) and FMO3 (NCBI accession number NP_008825.4), which oxidize phenethylamine to the oxime (30), show no similarity to HpxL and PcxL. Thus, if PcxL and HpxL are indeed FAD-NAD(P)H-dependent enzymes that catalyze oxime formation, they belong to a novel protein family.

Table 3.2. Known, characterized *N*-oxidases and their homology to HpxL, which is encoded by the hydroxyphosphonocystoximic gene cluster. PcxL is found in the homologous phosphonocystoximic biosynthetic gene cluster. FzmM and CreE catalyze the *N*-oxidation of aspartic acid to nitrosuccinic acid during fosfazinomycin (19) and cremeomycin (42) biosynthesis, respectively. AzmF (31) and CrmH (49) are characterized flavin-dependent *N*-oxidases that catalyze oxime formation in azolemycin and caerulomycin biosynthesis. ClmM is the predicted FAD- NAD(P)H-dependent *N*-oxidase of collismycin biosynthesis (16), while AlmD is the metal-dependent *N*-oxidase involved in oxime formation of alithomycin biosynthesis (12). Note that only FzmM and CreE can be considered as bona fide homologs of PcxL and HpxL.

Enzyme	Accession Number	Organism	Natural Product	% Identity to HpxL	E-Value	Length (Amino Acids)
HpxL	WP_030990682.1	*Streptomyces regensis* WC-3744	Hydroxyphosphonocystoximate	100%	0	750
PcxL	WP_051704824.1	*Streptomyces* sp. S-481	Phosphonocystoximate	61%	0	742
AzmF	AMQ23503.1	*Streptomyces* sp. FXJ1.264	Azolemycin	10%	2.9	389
ClmM	CCC55915.1	*Streptomyces* sp. CS40	Collismycin A	9%	1.4	402
CrmH	AFD30961.1	*Actinoalloteichus* sp. WH1-2216-6	Caerulomycin A	8%	1.1	407
AlmD	CCA29200.1	*Myxococcus xanthus* DK897	Althiomycin	8%	3.4	331
FzmM	WP_053787792.1	*Streptomyces* sp. XY332	Fosfazinomycin	30%	6E-14	644
CreE	ALA99202.1	*Streptomyces cremeus*	Cremeomycin	35%	3E-29	666

3.2.2 HpxL and PcxL catalyze NADPH oxidation in the presence of FAD and 2AEPn

To examine the biochemical properties of PcxL and HpxL, 6xHis-affinity-tagged alleles were overexpressed in *E. coli*. Recombinant proteins were then purified and assayed for their ability to oxidize nicotinamide cofactors in the presence of 2AEPn, FAD or FMN (Table 3.3). Both enzymes catalyze oxidation of NADPH and NADH at a

significant rate in the absence of additional substrates. Oxidation of reduced nicotinamides is stimulated by the addition of FAD, but not by addition of FMN, suggesting that the former is the preferred cofactor. It should be noted that the as-purified enzymes contain bound FAD, as indicated by the bright yellow coloration of preparations (Fig 3.3A). Additionally, we conclude that the flavin cofactor bound to the *N*-oxidases is FAD because of the presence of FAD in the protein preparations as determined by LC/ESI mass spectrometry (Fig 3.3B) and the maximum absorbance of the proteins being consistent with the presence of FAD (Fig 3.3C-E). We believe the presence of FAD bound to the proteins upon purification accounts for the basal activity seen in the absence of added FAD. Significantly, oxidation of NADPH, but not NADH, is substantially increased upon addition of 2AEPn, with the highest rates being observed for the combination of NADPH, FAD and 2AEPn. Thus, activity in the absence of 2AEPn is likely due to an uncoupled, oxygen-dependent reaction, which is common for flavin-dependent enzymes (45). Additionally, when this reaction is run in the absence of oxygen, there is no oxidation of NADPH by PcxL, in either the presence or absence of 2AEPn (data not shown).

Table 3.3. Oxidation of nicotinamide cofactors by HpxL and PcxL. The apparent rate in the presence/absence of different flavin cofactors and 2AEPn is shown. **2AEPn*** = rate in the presence of 2AEPn after subtraction of the uncoupled rate (*i.e.* before addition of 2AEPn). 2AEPn was used at a concentration of 5 mM, while FAD and FMN were used at 50 μM, and NADPH or NADH were used at 300 μM. Each experiment was completed in triplicate.

		k_{obs} (s^{-1} x 10^{-3})	
Cofactor(s)	**Substrate**	**HpxL**	**PcxL**
NADPH	none	2.7 ± 0.4	14.4 ± 0.7
	2AEPn	7.2 ± 0.6	232 ± 8
NADPH/FAD	none	12.5 ± 0.1	21 ± 1
	2AEPn	23.2 ± 0.7	310 ± 20
	2AEPn*	**10.7 ± 0.6**	**290 ± 20**
NADPH/FMN	none	5.2 ± 0.5	14.3 ± 0.7
	2AEPn	10.0 ± 0.2	210 ± 20
	2AEPn*	**4.8 ± 0.7**	**200 ± 20**
NADH	none	5.5 ± 0.1	24 ± 1
	2AEPn	1.7 ± 0.2	29 ± 2
NADH/FAD	none	14 ± 2	85 ± 5
	2AEPn	15.0 ± 0.7	83 ± 5
	2AEPn*	**0.1 ± 0.1**	**0 ± 3**
NADH/FMN	none	5.9 ± 0.3	26.6 ± 0.2
	2AEPn	7.0 ± 0.4	27 ± 4
	2AEPn*	**1.0 ± 0.2**	**1 ± 4**

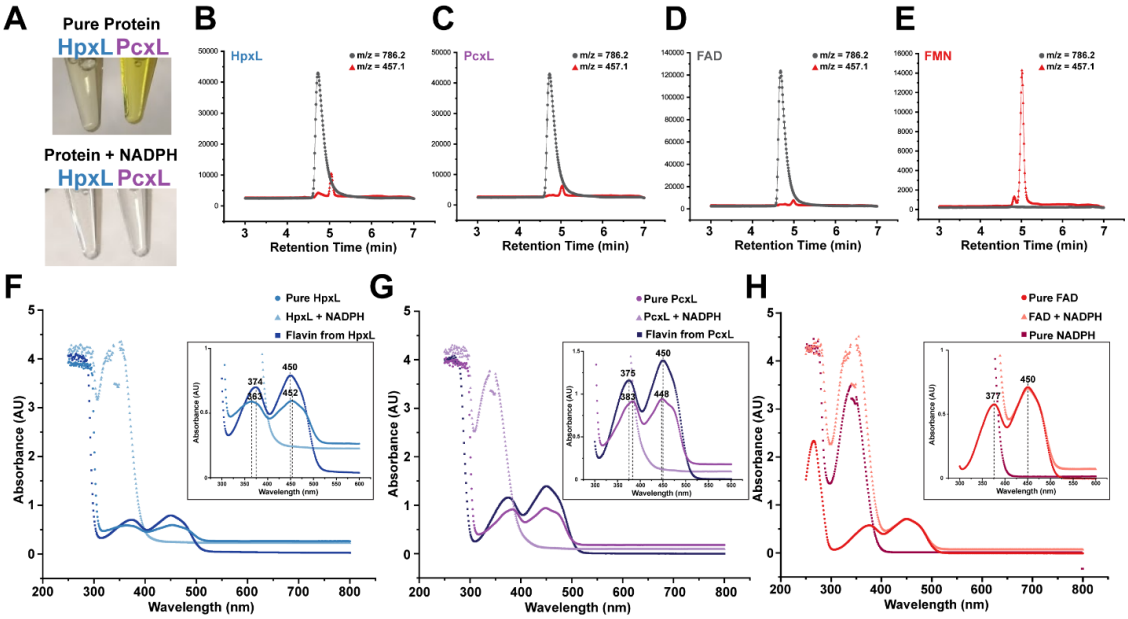


Figure 3.3. Characterization of HpxL and PcxL as FAD-dependent enzymes. (A) The top photograph is of purified 6xHis tagged *N*-oxidases after purification. The bottom photograph shows the enzyme after incubation with NADPH. (B-E) Low resolution mass spectrometry of the flavin isolated from (B) HpxL (37.1 μM) and (C) PcxL (30.9 μM). (D) FAD standards and (E) FMN standards were used at a concentration of 50 μM. (F) HpxL absorption spectra of purified flavin (dark blue squares) or HpxL before (blue circles) and after (light blue triangles) incubation with NADPH. Inset shows the maximum absorption of HpxL at 363 nm and 452 nm. Purified flavin from washed HpxL shows absorption maxima at 374 nm and 450 nm. (G) PcxL absorption spectra of purified flavin (dark purple squares) or the protein before (purple circles) and after (light purple triangles) incubation with NADPH. Inset shows the maximum absorption of PcxL at 383 nm and 448 nm. Purified flavin from washed PcxL shows absorption maxima at 377 nm and 450 nm. (H) Control reactions showing FAD absorption (red circles), NADPH absorption (dark red squares), and a mixture of pure FAD and NADPH (pink triangles). The inset shows the absorption maxima of FAD at 377 nm and 450 nm. The elevated baseline in the spectra shown in (F) and (G) are likely due to light scattering from small amounts of protein precipitation of HpxL and PcxL, as baseline elevation is not observed with the purified flavin that was washed from the proteins. We have observed that HpxL and PcxL are not stable for an extended period of time in solution at room temperature.

3.2.3 HpxL and PcxL oxidize 2AEPn to 2-hydroxyiminoethylphosphonic acid and 2-nitroethylphosphonic acid

A series of NMR experiments were conducted to establish the products of PcxL- and HpxL-catalyzed 2AEPn oxidation. Based on ^{31}P NMR spectroscopy, both enzymes form three phosphorus-containing products after prolonged incubation with NADPH, FAD and 2AEPn (Fig 3.4A and Fig 3.4B). Analysis of these molecules using 2-dimensional proton-phosphorus coupled HMBC NMR spectroscopy allowed assignment of the major peak as a mixture of the (*E*)- and (*Z*)-isomers of 2-hydroxyiminoethylphosphonic acid (**1**), with smaller amounts of 2-nitroethylphosphonic acid (**2**) (Fig 3.5 and Fig 3.6). Accordingly, aldoxime protons display a characteristic proton NMR spectroscopy chemical shift that will be between 6.5 and 8.0 ppm (35). A single proton associated with each of the products, which we assign as (*E*)-1 and (*Z*)-1 at 6.82 ppm and 7.37 ppm, respectively. This assignment is supported by a recent report of the total synthesis of the phosphonocystoximic acids, which involved the synthesis of ethyl protected aldoximes as intermediates. The aldoxime protons for the protected (*E*)-isomer had a chemical shift of 6.82 ppm, while the aldoxime proton associated with the protected (*Z*)-isomer has a chemical shift of 7.41 ppm (34). Our assignment of the third product as 2 is based on the observation of phosphorus correlated protons at 2.03 ppm and 4.48 ppm, which is consistent with literature values of synthetically-produced 2-nitroethylphosphonic acid (3). While both enzymes form a mixture of (*E*)- and (*Z*)-isomers when using 2AEPn as substrate, HpxL preferentially forms (*E*)-1 (Fig 3.5), while PcxL preferentially forms (*Z*)-1 (Fig 3.6).

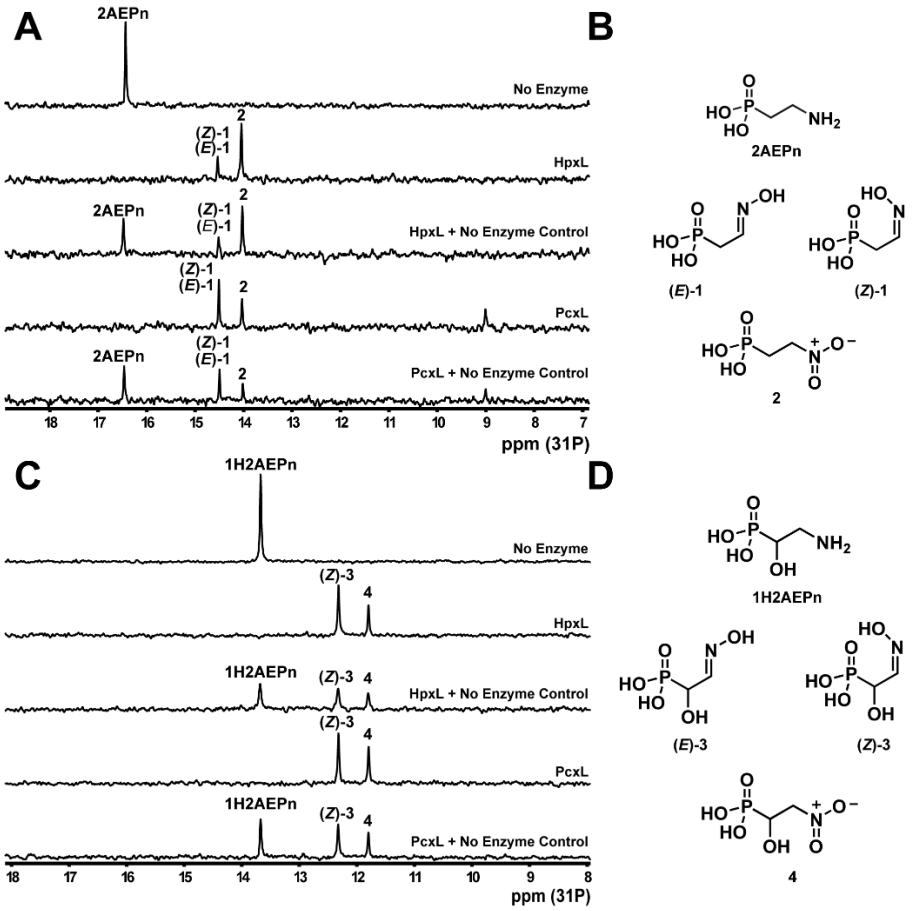


Figure 3.4. ^{31}P NMR-analysis of N-oxidase activity in the presence of 2AEPn or (S)-1H2AEPn. (A) HpxL and PcxL form three major products in vitro when incubated for 16 hours with NADPH, FAD and 2AEPn as the substrate, whereas no product formation is observed in the absence of enzyme. (B) These products have been assigned as the (E)- and (Z)-isomers of 2-iminoethylphosphonic acid (1) and 2-nitroethylphosphonic acid (2) based on the 1H-31P HMBC analysis (Fig. S2 and S3). Note that the (E)- and (Z)-isomers of 1 have identical 31P chemical shifts and thus display only a single peak in the 1D-31P spectrum shown in panel A. (C) HpxL and PcxL both generate two products in vitro using (S)-1H2AEPn as the substrate. (D) These products were assigned as (Z)-1-hydroxy-2-hydroxyiminoethylphosphonic acid (3) and 1-hydroxy-2-nitroethylphosphonic acid (4) based on the 1H-31P HMBC analysis (Fig. S4 and S5). Chemical mixing experiments with the control reaction was performed to confirm that substrate had been completely consumed. Substrates and cofactors were added at the following concentrations: 100 μM FAD, 500 μM NADPH, 3 mM 2AEPn or HpxV-generated (S)-1H2AEPn, 25 μM PcxL or HpxL, 25 mM phosphite, and 10 μM PTDH17x.

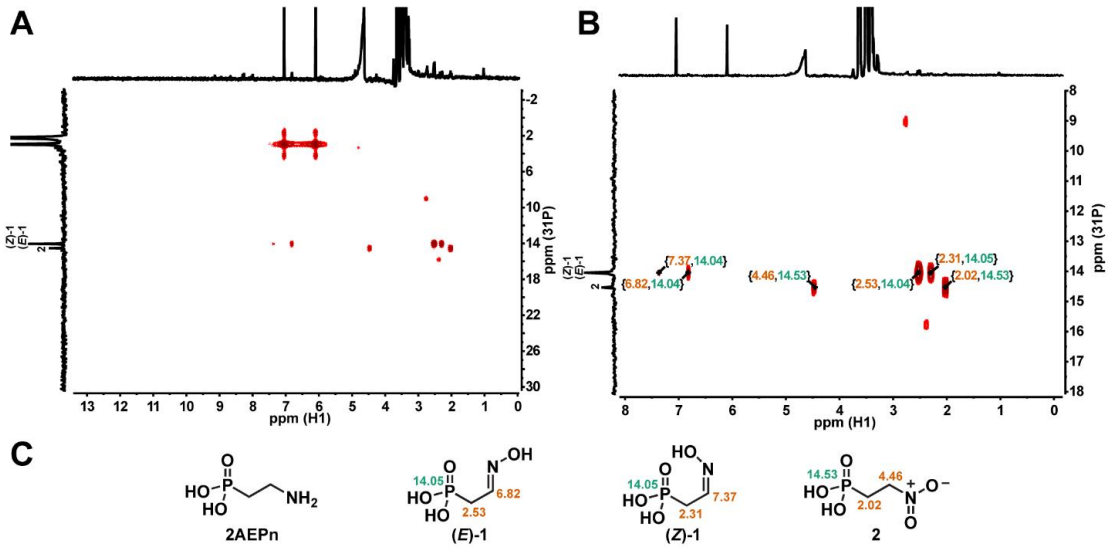


Figure 3.5. HMBC NMR analysis of HpxL incubated in the presence of FAD, NADPH, and 2AEPn for 16 hours. (A) ^{1}H-^{31}P HMBC spectrum of the HpxL reaction with 2AEPn as the substrate showing all phosphorus-containing molecules in the reaction. (B) Magnified region of the ^{1}H-^{31}P spectrum with key proton-phosphorus correlations labeled ({^{1}H chemical shift, ^{31}P chemical shift}). (C) Compounds identified using ^{1}H-^{31}P correlations with the ^{1}H chemical shifts displayed in orange and ^{31}P chemical shift in green. 2AEPn, (*E*)- and (*Z*)-**1**, and **2** are shown. Substrates and cofactors were added at the following concentrations: 100 µM FAD, 500 µM NADPH, 3 mM 2AEPn, and 25 µM HpxL. 25 mM phosphite and 10 µM PTDH17x were also included as a cofactor regenerating system.

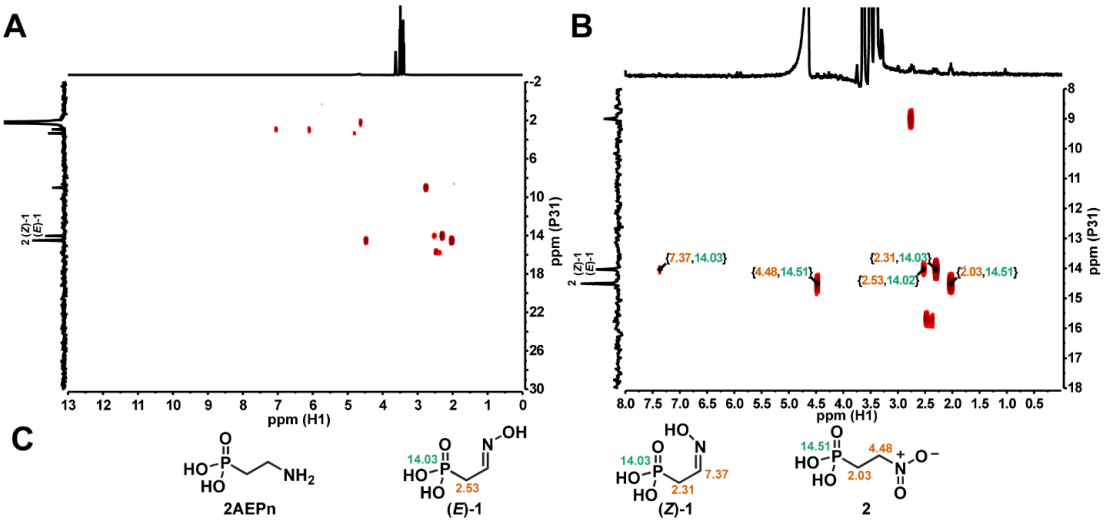


Figure 3.6. HMBC NMR analysis of PcxL incubated in the presence of FAD, NADPH, and 2AEPn for 16 hours. (A) ^{1}H-^{31}P HMBC NMR spectra of the PcxL reaction with 2AEPn as the substrate showing all phosphorus-containing molecules in the reaction. (B) Magnified region of the ^{1}H-^{31}P spectrum with key proton-phosphorus correlations labeled ({^{1}H chemical shift, ^{31}P chemical shift}). (C) Compounds identified using ^{1}H-^{31}P correlations with the ^{1}H chemical shifts displayed in orange and ^{31}P chemical shift in green. 2AEPn, (*E*)- and (*Z*)-**1** and **2** are shown. Substrates and cofactors were added at the following concentrations: 100 µM FAD, 500 µM NADPH, 3 mM 2AEPn, and 25 µM PcxL. 25 mM phosphite and 10 µM PTDH17x were also added as a cofactor regenerating system.

We performed an identical experiment using 1H2AEPn as substrate to determine the products that would be formed by the enzymes. PcxL and HpxL catalyze oxidation of the amine moiety in (*S*)-1H2AEPn as a substrate (Fig 3.4C and Fig 3.4D). In contrast to the products obtained with 2AEPn, the enzymes form only the (*Z*)-isomer of 1-hydroxy-2-hydroxyiminoethylphosphonic acid (**3**) when using (*S*)-1H2AEPn as the substrate. This is evident by the presence of only one set of protons correlating with the phosphorus atom with chemical shifts of 4.15 ppm and 7.44 ppm (Fig 3.7 and Fig 3.8). The enzymes also form smaller amounts of 1-hydroxy-2-nitroethylphosphonic acid (**4**) (Fig 3.7 and Fig 3.8).

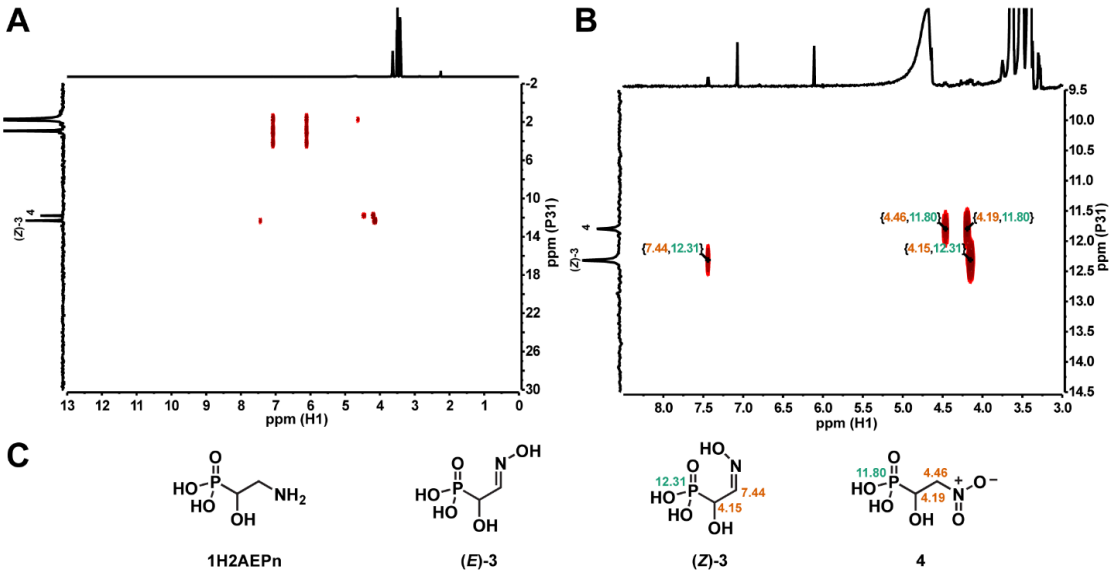


Figure 3.7. HMBC NMR analysis of HpxL incubated in the presence of FAD, NADPH, and (*S*)-1H2AEPn for 16 hours. (A) ^{1}H-^{31}P HMBC NMR spectra of the HpxL reaction with (*S*)-1H2AEPn as the substrate showing all phosphorus-containing molecules in the reaction. (B) Magnified region of the ^{1}H-^{31}P spectrum with key proton-phosphorus correlations labeled ({^{1}H chemical shift, ^{31}P chemical shift}). (C) Compounds identified using ^{1}H-^{31}P correlations with the ^{1}H chemical shifts displayed in orange and ^{31}P chemical shift in green. 1H2AEPn, (*E*)- and (*Z*)-**3** and **4** are shown. Substrates and cofactors were added at the following concentrations: 100 μM FAD, 500 μM NADPH, 3 mM (*S*)-1H2AEPn, and 25 μM HpxL. 25 mM phosphite and 10 μM PTDH17x were also as a cofactor regenerating system.

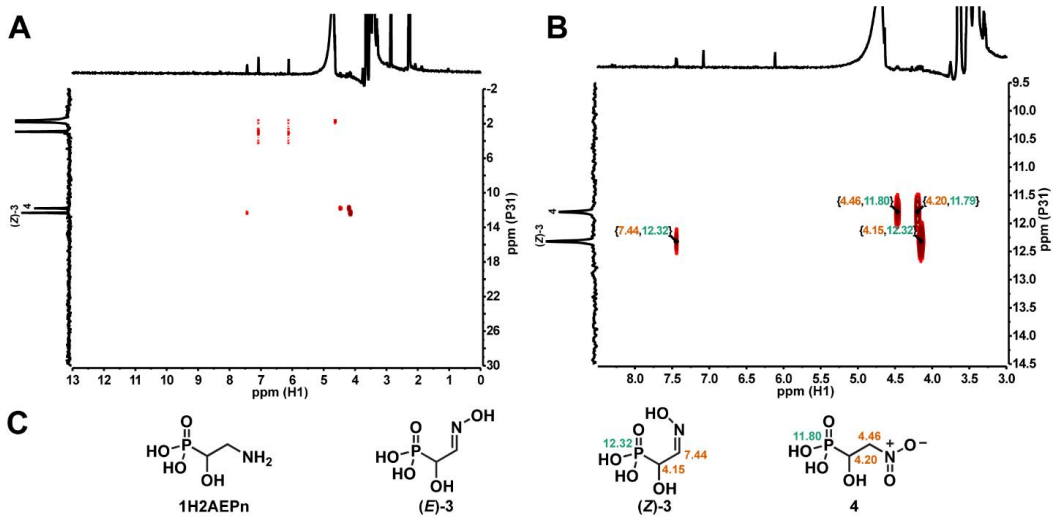


Figure 3.8. PcxL was incubated in the presence of FAD, NADPH, and (*S*)-1H2AEPn for 16 hours. (A) ^{1}H-^{31}P HMBC NMR spectra of the HpxL reaction with (*S*)-1H2AEPn as the substrate showing all phosphorus-containing molecules in the reaction. (B) Magnified region of the ^{1}H-^{31}P spectrum with key proton-phosphorus correlations labeled ({^{1}H chemical shift, ^{31}P chemical shift}). (C) Compounds identified using ^{1}H-^{31}P correlations with the ^{1}H chemical shifts displayed in orange and ^{31}P chemical shift in green. (*S*)-1H2AEPn, (*E*)- and (*Z*)-**3** and **4** are shown. Substrates and cofactors were added at the following concentrations: 100 μM FAD, 500 μM NADPH, 3 mM (*S*)-1H2AEPn, and 25 μM PcxL. 25 mM phosphite and 10 μM PTDH17x were also as a cofactor regenerating system.

To determine the molar ratio of substrates consumed to cofactor used, we determined the stoichiometry of the reaction by monitoring substrate (2AEPn and oxygen) and cofactor (NADPH) consumption during short incubations. We observed that the molar ratio of oxygen to NADPH was approximately 1:1 for both enzymes; however, the molar ratio of NADPH to 2AEPn consumption was much higher: 2.5:1 (NADPH/2AEPn) for PcxL and 4.0:1 for HpxL (Table 3.4). This yields an overall reaction stoichiometry of 2AEPn to NADPH to oxygen consumption of 1/2.5/2.7 for PcxL and 1/4.0/3.8 for HpxL. Significantly, 2-nitroethylphosphonic acid was not observed during the short incubations used in these experiments (data not shown). Thus, the presence of this product in longer incubations is probably due to overoxidation of the substrate.

Table 3.4. Stoichiometry of the PcxL- and HpxL-catalyzed reactions. Oxygen and NADPH consumption were measured in five-minute reactions using a Clark-type electrode and UV/Vis spectrophotometry, respectively, to determine the O_2/NADPH ratio. A separate, three-hour reaction was used to obtain the NADPH/2AEPn ratio with ^{31}P NMR being used to monitor both substrates. Each experiment was performed in triplicate.

Substrate	Enzyme	O_2/NADPH ratio			NADPH/2AEPn ratio			Combined Ratio
		nmol NADPH	*nmol O_2*	*Ratio (NADPH:O_2)*	*nmol NADPH*	*nmol 2AEPn*	*Ratio (NADPH: 2AEPn)*	2AEPn: O_2: NADPH
2AEPn	PcxL	966 ± 0.4	1040 ± 2	1 : 1.1	1970 ± 3	810 ± 30	2.5 : 1	1 : 2.5 : 2.7
	HpxL	163 ± 1	156 ± 2	1.1 : 1	1420 ± 30	350 ± 50	4.0 : 1	1 : 4.0 : 3.8

3.2.4 HpxL and PcxL both preferentially oxidize 1H2AEPn over 2AEPn, but with different stereochemical preferences

To determine the substrate preference of the *N*-oxidases *in vitro*, we performed initial rate kinetics on PcxL and HpxL using 2AEPn and the (*R*)- and (*S*)-enantiomers of 1H2AEPn as substrates (Table 3.5, Fig 3.9). Although both enzymes are capable of using all of these substrates, they manifest significantly different substrate preferences. PcxL shows the highest rate with 2AEPn as substrate, but shows a greater affinity for (*R*)-1H2AEPn. Thus, the catalytic efficiency of PcxL is similar for these two substrates. In contrast, (*S*)-1H2AEPn is a relatively poor substrate for PcxL, while HpxL turns over (*S*)-1H2AEPn and 2AEPn at similar rates, but has a 10-fold higher affinity for the former. Thus, (*S*)-1H2AEPn, which is predicted to be the *in vivo* substrate, is the preferred substrate for this enzyme. (*R*)-1H2AEPn is a relatively poor HpxL substrate and actually inhibits the enzyme at higher concentrations (Fig 3.9F).

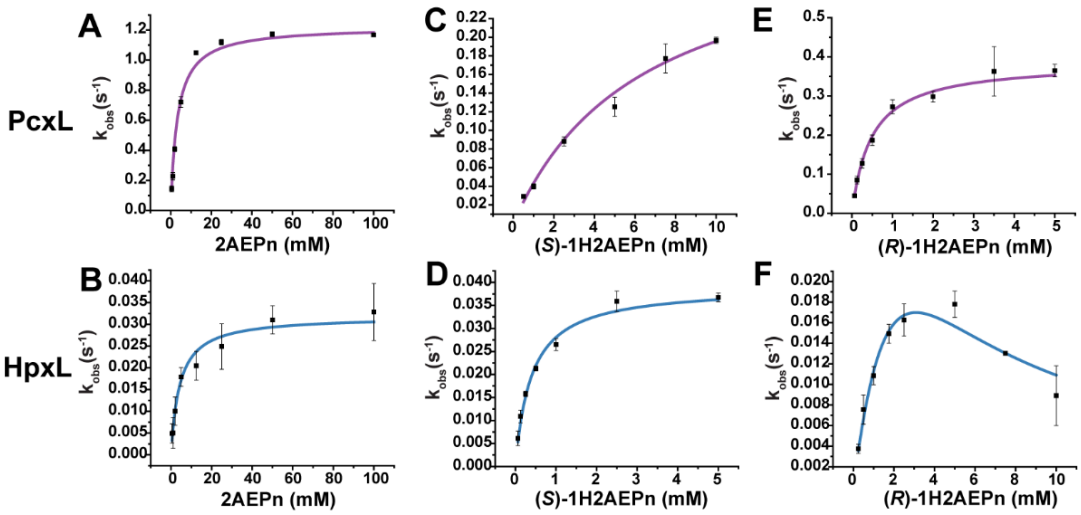


Figure 3.9. Kinetic curves for PcxL and HpxL in the presence of different substrates. The kinetic constant values are shown below in Table 3.5. (A) PcxL and (B) HpxL with 2AEPn as the substrate fit a standard Michaelis-Menten curve. (C). PcxL and (D) HpxL with (*S*)-1H2AEPn fit to the standard Michaelis-Menten equation. (E) PcxL with (*R*)-1H2AEPn fit to the standard Michaelis-Menten equation and (F) HpxL with (*R*)-1H2AEPn fit for substrate inhibition.

Table 3.5. Initial rate kinetics for HpxL and PcxL. Rates were determined by monitoring NADPH consumption via UV-Vis spectrophotometer. V_{max}, K_m, and K_i were determined by fitting the data with either Michaelis-Menten (MM) or substrate inhibition (SI) models. Only best fit is shown. Reaction components included: 50 mM sodium phosphate buffer, pH 7.8, 300 µM NADPH, 50 µM FAD, substrate ranging from 62.5 µM to 150 mM, and enzyme concentrations of 1 µM or 5 µM for PcxL and 5 µM or 10 µM for HpxL. The substrates used were pure 2AEPn, synthetic (*R*)-1H2AEPn, or synthetic (*S*)-1H2AEPn. This experiment was performed in triplicate for each curve.

Enzyme	**Substrate**	**Kinetic Model**	**V_{max} (s^{-1})**	**Apparent K_m (*mM*)**	**k_{cat}/K_m (M^{-1} s^{-1})**	**K_i (*mM*)**
PcxL	2AEPn	MM	1.22 ± 0.02	3.3 ± 0.5	370	n/a
	(*R*)-1H2AEPn	MM	0.39 ± 0.01	0.48 ± 0.02	810	n/a
	(*S*)-1H2AEPn	MM	0.32 ± 0.03	6 ± 1	53	n/a
HpxL	2AEPn	MM	0.032 ± 0.002	4.3 ± 0.8	7.4	n/a
	(*R*)-1H2AEPn	SI	0.06 ± 0.03	4 ± 2	15	2 ± 1
	(*S*)-1H2AEPn	MM	0.039 ± 0.001	0.39 ± 0.03	100	n/a

3.2.5 HpxL and PcxL are members of a large family of putative *N*-oxidases that is common to phosphonate biosynthetic pathways

The low similarity of HpxL and PcxL to other known *N*-oxidases suggests that they are members of larger protein family (Table 3.2). Indeed, BLAST searches of GenBank retrieve thousands of hits using HpxL and PcxL as query sequences (data not shown). Phylogenetic analysis of the top 500 HpxL homologs reveals numerous discrete clades within the family. To provide clues as to the function of the proteins within this family, we examined the genomic context surrounding these homologs using the bioinformatics tool Rodeo, which retrieves adjacent genes and assigns putative functions based on PFAM membership (44). Significantly, almost all genes encoding proteins within the HpxL and PcxL clade fall within gene clusters devoted to the synthesis of phosphonic acids based on co-localization with genes encoding phosphoenolpyruvate mutase (PepM), which encodes the first step in most phosphonate biosynthetic pathways (Fig 3.10A) (21, 32, 48). As described above, the known *N*-oxidases FzmM and CreE are bona fide members of this protein family; however, they are only distantly related to HpxL and PcxL and do not fall within the top 500 homologs shown in Fig 3.10A. Indeed, the BLAST search results using HpxL as the query must be extended to include the top 5000 hits before FzmM and CreE are included.

We performed a similar analysis using previously identified phosphonate biosynthetic gene clusters as the starting point (21). 23% of the phosphonate gene cluster families contain a HpxL homolog, with only three families having characterized products: the phosphonocystoximic acids, hydroxyphosphonocystoximic acids, and fosfazinomycins. Thus, at least 15 novel, uncharacterized phosphonic acid natural

product biosynthetic pathways encode a HpxL homolog. Based on the similarity of these enzymes within the *N*-oxidase tree, we can predict that these novel 15 phosphonic acids will likely involve the oxidation of 2AEPn or 1H2AEPn to form an oxime-containing natural product (Fig 3.10B).

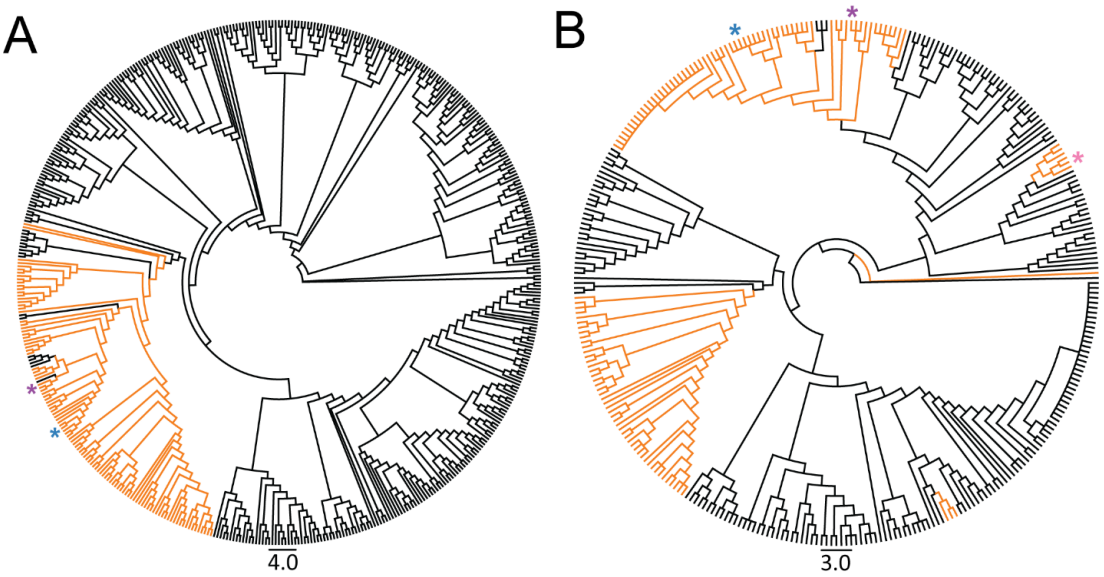


Figure 3.10. Cladograms showing the phylogenetic distribution of putative *N*-oxidases related to HpxL and PcxL. (A) Phylogeny of the 500 closest HpxL homologs. *hpxL*-containing gene clusters that also contain *pepM* are highlighted in orange. (B) Phylogeny of actinobacterial PepM sequences. *pepM*-containing gene clusters that contain an HpxL homolog are highlighted in orange. HpxL is denoted with a blue asterisk, PcxL is denoted with a purple asterisk, and FzmM is denoted with a pink asterisk. Gene clusters were arbitrarily defined to include the 15 genes upstream or downstream of the *N*-oxidase.

3.3 DISCUSSION

3.3.1 PcxL and HpxL are 2AEPn and 1H2AEPn amine-oxygenases

The data presented here show that PcxL and HpxL are novel FAD- and NADPH-dependent oxime-forming amine oxygenases that utilize amine-containing phosphonate substrates. Although these enzymes were shown to oxidize several small amino phosphonates, the substrate preferences of the two enzymes support the proposed

biosynthetic pathways shown in Fig 3.2. Accordingly, we expected that the substrate preference of the *N*-oxidases would be influenced by the presence of the proposed 2AEPn dioxygenase, HpxV, in the gene cluster of *S. regensis* NRRL WC-3744 (Fig 3.2, Table 3.1), with HpxL having a preference for (*S*)-1H2AEPn and PcxL having a preference for 2AEPn (Ju, 2015). The finding that the catalytic efficiency of HpxL is 10-fold higher for (*S*)-1H2AEPn than the other substrates tested is fully consistent with this model. Our conclusions regarding PcxL are less certain, because the enzyme displays a similar catalytic efficiency with both 2AEPn and (*R*)-1H2AEPn. Nevertheless, it is very likely that 2AEPn is the natural substrate given the absence of known genes to produce hydroxylated phosphonate substrates in *Streptomyces* sp. NRRL S-481.

3.3.2 The stoichiometry of the reaction suggests over-oxidation of substrate during formation of the aldoxime

Based on the finding that NADPH and O_2 are always consumed in a 1:1 ratio by PcxL and HpxL (Table 3.4), oxime formation should require two separate hydroxylation steps, each consuming one molar equivalent of NADPH and one molar equivalent of oxygen. Each step would also produce one molar equivalent of H_2O by reduction of O_2 with NADPH. We predict that this reaction would proceed through hydroxyamino- and dihydroxyamino- intermediates, with subsequent loss of water producing a nitroso intermediate, which would tautomerize to afford the final aldoxime product observed in our reactions (Fig 3.11). An infrequent, third NADPH- and O_2-dependent hydroxylation of the dihydroxyamino-, nitroso- intermediates, or the oxime products would then afford small amounts of the nitro-derivatives seen in our data (Fig 3.4 - Fig 3.8). The fact that we see substantially higher levels of O_2 consumption, relative to aldoxime formation, is

consistent with this idea, as is the observation of NADPH oxidation in the absence of phosphonate substrates (*i.e.* the uncoupled reaction). We expect that the latter would produce hydrogen peroxide, although this was not directly tested.

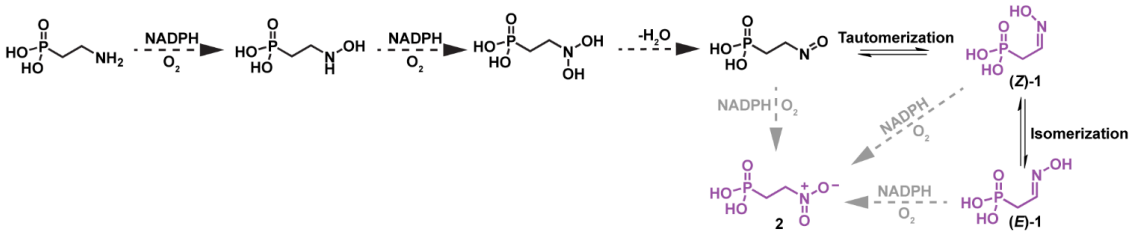


Figure 3.11. Proposed stepwise oxidation of 2AEPn to yield **1** or **2** that may explain the 1:2.5 substrate:NADPH ratio observed for oxime formation *in vitro*. We propose that each *N*-oxidation would require one molar equivalent of NADPH and molecular oxygen. This would require two *N*-oxidations and a dehydration to yield 2-nitrosoethylphosphonic acid, which could tautomerize to form 2-hydroxyiminoethylphosphonic acid. An enzyme-catalyzed overoxidation of either 2-nitrosoethylphosphonic acid or 2-hydroxyiminoethylphosphonic acid could potentially yield 2-nitroethylphosphonic acid (shown with grey arrows). Compounds shown in purple are all products observed in the reactions catalyzed by HpxL and PcxL using 2AEPn as the substrate (Fig 3.4).

3.3.3 PcxL and HpxL are new members of a class of flavin-dependent amine oxidases common in nature

Sequence analysis shows that HpxL and PcxL are members of a very large protein family that is widely distributed in bacteria. Only four members of this family have been biochemically characterized, including HpxL, PcxL and the distantly related enzymes, FzmM and CreE. While the former oxidize amino phosphonates to aldoximes (and small amounts of nitro derivatives), the latter oxidize aspartate to nitrosuccinic acid (19, 42). Given the phylogenetic distance between these distant homologs, it seems reasonable to assume that most members of the family are amine oxidases. Interestingly, members of this presumptive amine oxidase family are common in actinobacterial phosphonate biosynthetic gene clusters, with nearly one quarter of those examined containing a homolog. These enzymes are widely distributed on the PepM maximum-likelihood tree,

suggesting HpxL and PcxL homologs were recruited into phosphonic acid biosynthetic gene clusters multiple times: however, all of these homologs are closely related, falling into the same clade within the amine oxidase tree (Fig 3.10A). Moreover, each of the gene clusters containing these homologs also contain the biosynthetic machinery needed to make 2AEPn or 1H2AEPn. Taken together, these data suggest that numerous phosphonate biosynthetic pathways include 2-hydroxyiminoethylphosphonic acid or 1-hydroxy-2-hydroxyiminoethylphosphonic acid as early intermediates (see Fig 3.4). Thus, although the phosphonocystoximates are the only known oxime-containing phosphonate natural products, it seems certain that numerous additional examples await discovery.

3.4 SUMMARY AND OUTLOOK

Phosphonocystoximic acid and hydroxynitrilaphos are two natural products with interesting functional groups on these molecules. With the discovery of the oxime-forming enzymes, one step of the biosynthesis of these molecules has been elucidated. This is only a small fraction of the genes predicted to be involved in the biosynthesis of the phosphonocystoximic acids (21). In addition, there are many more unanswered questions regarding HpxL and PcxL. Here I will outline my thoughts regarding the future of phosphonocystoximate biosynthesis.

PcxL and HpxL are proteins that catalyze the formation of oxime-containing phosphonic acids and members of a novel group of amine-oxidases. These proteins share little homology to known amine-oxidases outside of phosphonate biosynthesis, with the closest characterized homologs using aspartic acid as the substrate (19, 42). These proteins are substantially longer than other characterized amine-oxidases, yet the conserved NADPH and FAD-binding domains appear to be conserved in the N-terminal

domain of the protein. When performing a BLAST search (2) on the C-terminal domain of these proteins, no proteins with characterized or predicted functions are returned in the search. Therefore, we are unsure whether the C-terminal domain of this protein is required for the oxime-forming function of this protein. Truncation mutants of this protein lacking the C-terminal domain would suggest whether this domain is necessary for function.

PcxL and HpxL homologs appear to be widely distributed in phosphonate biosynthesis (Fig 3.10). Looking within the genomic neighborhood of these phosphonate biosynthetic gene clusters, most of these clusters contain a homologous gene to the 2AEPn aminotransferase (data not shown) responsible for formation of 2AEPn from phosphonoacetaldehyde (8, 14, 27, 29). This suggests the substrate of these enzyme may be 2AEPn, although there is a possibility that these enzymes may utilize additional phosphonic acid substrates. Preliminary testing of alternative substrate utilization (Fig 3.12) by PcxL and HpxL suggest that these enzymes do preferentially catalyze the oxidation of 2AEPn and 1H2AEPn because the catalytic efficiencies of these enzymes with the other phosphonic acid substrates is much lower (Table 3.6). However, these enzymes appear to be able to utilize other phosphonic acid substrates. This suggests there is some substrate tolerance of the enzymes and further investigations may provide a way to produce derivatives of small phosphonic acids with an oxime functional group.

Table 3.6. Initial rate kinetics for HpxL and PcxL using alternative substrates. Rates were determined by monitoring NADPH consumption on a UV-Vis spectrophotometer. V_{max}, K_m, and K_i were determined by fitting the data with either Michaelis-Menten (no substrate inhibition observed) or substrate inhibition models (substrate inhibition model was used when a K_i was determined). Reaction components included: 50 mM sodium phosphate buffer, pH 7.8, 300 µM NADPH, 50 µM FAD, and substrate ranging from 1 mM to 160 mM. Enzyme concentrations of 5 µM for PcxL and 5 µM or 10 µM for HpxL were used. The substrates used were pure phosphonoalanine, l-1-aminoethylphosphonic acid (L-alaP), or aminomethylphosphonic acid (AMPn). This experiment was performed in triplicate for each curve.

Enzyme	Substrate	k_{obs}	K_M	k_{cat}/K_M	K_i
		(s^{-1})	*(mM)*	*(M^{-1} s^{-1})*	*(mM)*
PcxL	Phosphonoalanine	1.61 ± 3.43	161 ± 369	10	11 ± 28
	L-alaP	0.209 ± 0.009	21 ± 4	10	n/a
	AMPn	0.45 ± 0.02	23 ± 3	20	n/a
HpxL	AMPn	0.050 ± 0.002	1.7 ± 0.2	30	45 ± 4

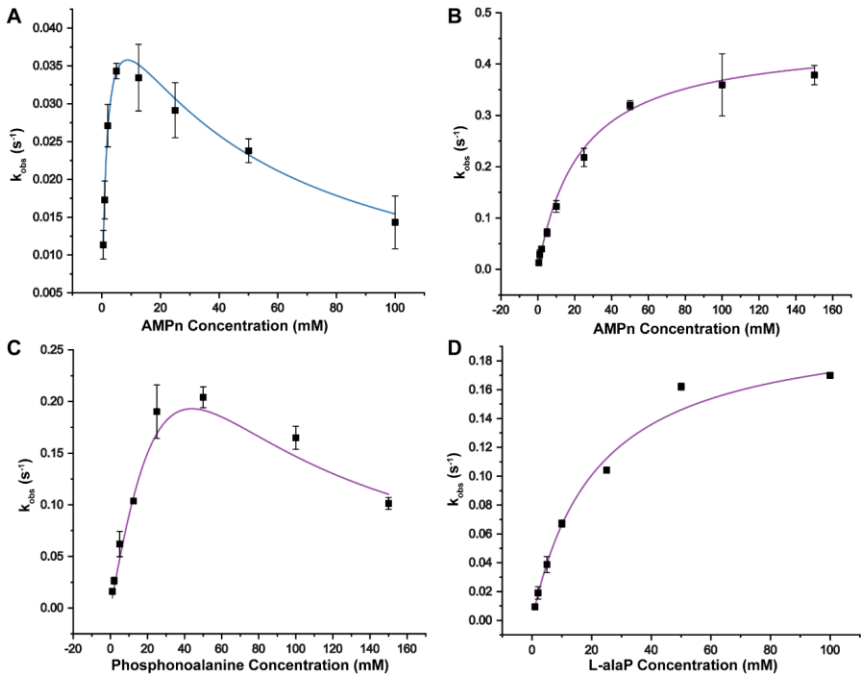


Figure 3.12. Kinetic curves of PcxL or HpxL in the presence of different phosphonic acid substrates. The kinetic constants extracted from each curve are shown in Table 3.6. (A) HpxL using AMPn as a substrate. (B) PcxL using AMPn as a substrate. (C) PcxL using phosphonoalanine as a substrate. (D) PcxL using L-alaP as a substrate.

Streptomyces regensis NRRL WC-3744 has been shown to produce both hydroxynitrilaphos and hydroxyphosphonocystoximic acid (11, 21). We are still unsure of the final natural product of the *hpx* biosynthetic gene cluster and whether this product is bioactive. There is a possibility that hydroxyphosphonocystoximate is an off-pathway intermediate that is produced by the strain and that hydroxynitrilaphos is the final, bioactive molecule produced by these strains. We believe that phosphonocystoximate may be made in a similar biosynthetic route as glucosinolates in plants based on the putative functions of some of the proteins of these biosynthetic gene clusters (11, 21). Based on the biosynthesis of glucosinolates in plants, nitrile-containing compounds are not an uncommon product of these biosynthetic pathways (17). Therefore, we hypothesize that hydroxynitrilaphos may be a by-product that is similar to by-products observed in glucosinolate biosynthesis.

The bioactivity of the phosphonocystoximates and hydroxynitrilaphos is unknown due to the lack of sufficient material isolated from the producing organisms (11, 21). The logical next step would be testing the bioactivity of these compounds to determine whether they show potent bioactivity like other small molecules containing an oxime moiety (37). Glucosinolates, and their degradation products, have been shown to have antifungal (40), antibacterial (1, 38), and insect repellant (43) activities. In particular, it has been noted in numerous studies that the isothiocyanate degradation product is the toxic by-product that displays bioactivity, not the glucosinolate itself. This would be a new mechanism of Trojan horse-like bioactivity for phosphonate natural products if these small molecules have a similar toxic breakdown product. Therefore, it is prudent to

determine whether the phosphonocystoximic acids or hydroxynitrilaphos show bioactivity and if these molecules function in a similar manner to glucosinolates.

3.5 MATERIALS AND METHODS

3.5.1 General experimental procedures

Nuclear Magnetic Resonance (NMR) spectra were collected using an Agilent Technologies 600 MHz NMR spectrometer with D_2O as the lock solvent (minimum 20% v/v). All chemicals used in this study were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted.

3.5.2 *N*-oxidase cloning, overexpression, and purification

The genes encoding the *N*-oxidases were amplified via PCR using primer pairs MG140 (5'-CTGGTGCCGCGCGGCAGCCATATGacgaccgacaattccccagcca-3') and MG141 (5'-AGTGGTGGTGGTGGTGGTGCTCGAGtcatgcgggggtcctttcgatgtcc-3') for *hpxL* and primer pairs MG142 (5'-CTGGTGCCGCGCGGCAGCCATATGcgttccaaccgactcgccatcgtcg-3') and MG143 (5'-AGTGGTGGTGGTGGTGGTGCTCGAGtcacactcccgtggtgatgcgcgcg-3') for *pcxL* from fosmids containing the phosphonocystoximic acid gene clusters (21). The plasmid, pET28B, was amplified using primers pET28-NdeI-R (5'-CATATGGCTGCCGCGCGGCACCAGGCCGCTG-3') and pET28B-XhoI-F (5'-CTCGAGCACCACCACCACCACCACTGAGATCCGG-3'). Uppercase letters represent sequence homologous to the plasmid, pET28B, underlined sequence is an inserted restriction enzyme site, and the lowercase letters represent sequence homology to the gene to be amplified. The primer pair MG140 and MG141 were designed to mutate the GTG start codon of *hpxL* to an ATG start codon for expression in *Escherichia coli*.

The gene fragments were subsequently cloned into pET28B using Gibson Assembly® (New England Biolabs, Ipswich, MA) to yield pMNG016 for *hpxL* and pMNG017 for *pcxL*. The sequences were verified at the UIUC Core Sequencing Facility.

The PcxL and HpxL proteins were affinity-purified after expression in *E. coli*. PcxL and HpxL were overexpressed in transformed *E. coli* Rosetta DE3 cells containing either plasmid pMNG016 or pMNG017. Proteins were purified using one of two methods. Method one involved cells grown in Studier's ZYM5052 autoinduction medium (41) with the following modification: the trace elements solution was replaced with 1 mg/L of each iron(II) sulfate heptahydrate, zinc(II) sulfate heptahydrate, and manganese(II) chloride dihydrate. The cultures were shifted to 18°C when the optical density (OD_{600}) reached 0.6-0.8 and protein was overexpressed for 16 hours. Cells were collected via centrifugation and resuspended in lysis buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), 10% glycerol and 10 mM imidazole, pH 8.0). The cell suspension was mixed with 1 mg/mL lysozyme, 100 units DNase I, and 100 µg RNase A and mixed for 30 minutes at 4°C. The cell suspension was passed two times through a French Press cell disruptor (Thermo Electron Corporation, Waltham, MA) at 1000 psi. The suspension was centrifuged at 14,000 x gravity (g) for a minimum of 30 minutes and the pellet was discarded. The liquid portion was passed over a 5 mL HisTrap FPLC protein column (GE Healthcare Life Sciences, Marlborough, MA) and washed with lysis buffer until the protein was no longer eluting from the column as determined by absorbance at A_{260} nm. Protein was eluted from the column on an ÄTKApurifier FPLC (GE Healthcare Life Sciences, Marlborough, MA) using a gradient from 10 mM – 250 mM imidazole with 60 column volumes over one hour at a flow rate of 5 mL/min.

Fractions containing purified protein, as determined by running the samples on a SDS-PAGE gel, were pooled, and concentrated using a 30 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA). Method two involved cells grown in Luria Broth until the optical density (OD_{600}) reached 0.6-0.8. The cultures were chilled on ice for 30 minutes and protein expression was induced upon the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were shaken at 18°C for 16 hours after induction. Pelleted cells were resuspended in lysis buffer (see above) containing 1 mg/mL lysozyme, 100 units DNase I, and 100 μg RNase A and mixed at 4°C for 30 minutes. The cell suspension was passed two times through a French Press cell disruptor (Thermo Electron Corporation, Waltham, MA) at 1000 psi. The suspension was centrifuged at 14,000 x g for 30 minutes and the pellet was discarded. The soluble fraction containing soluble protein was mixed with 5 mL of loose Ni-NTA resin for 30 minutes at 4°C. The resin was washed with wash buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), 10% glycerol and 20 mM imidazole, pH 8.0) until protein was no longer eluting from the resin as determined by absorbance at A_{260} nm. Protein was eluted from the column by incubating the resin with elution buffer (300 mM sodium chloride, 250 mM sodium phosphate (dibasic), 10% glycerol and 20 mM imidazole, pH 8.0) for 15 minutes and collecting the flow through. This flow through was concentrated using a 30 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA). All samples, regardless of which purification method was used, were exchanged into storage buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), and 10% glycerol) using a PD-10 desalting column (GE Healthcare Life Sciences, Marlborough, MA).

3.5.3 Bioinformatics characterization of HpxL and PcxL

The following amino acid sequences were downloaded from NCBI: HpxL (accession number WP_030990682.1), PcxL (accession number WP_051704824.1), AzmF (accession number AMQ23503.1), ClmM (accession number CCC55915.1), CrmH (accession number AFD30961.1), AlmD (accession number CCA29200.1), FzmM (accession number WP_053787792.1), and CreE (accession number ALA99202.1). These sequences were aligned using the standard settings for a MUSCLE alignment using Geneious (23) (Geneious version 8.1.2), which provided the percent identity between the proteins for Table 3.2.

3.5.4 Cofactor preference of the PcxL and HpxL

Cofactor preference experiments were monitored using an Cary 4000 UV-Vis spectrophotometer (Agilent, Santa Clara, CA). NADH and NADPH consumption was monitored by observing the absorption at 340 nm. Reaction components included: 50 mM sodium phosphate buffer, pH 7.8, 300 μM nicotinamide cofactor (either NADPH or NADH), 50 μM flavin cofactor (FAD or FMN), 10 mM 2AEPn and 10 μM *N*-oxidase (PcxL or HpxL). Reaction components except for substrate and enzyme were mixed in a 200 μL quartz cuvette. Enzyme was added and the background rate of the enzyme was collected for approximately one minute. Reactions were initiated upon addition of 2AEPn and three replicates were collected for each condition.

3.5.5 Flavin identification

HpxL (13.8 mg) and PcxL (5.4 mg) stocks were brought to room temperature and concentrated to 50 μL using a 30 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA) and the flow through was discarded. Protein was washed six

times with 500 μL ddH_2O. At the end of washing, PcxL was colorless, while HpxL retained some bound flavin as noted by the yellow color of the protein. Collected flavin was flash frozen and placed on a lyophilizer until dry. The flavin was resuspended in ddH_2O. We determined the concentration of the flavins based on absorbance at 450 nm, using an extinction coefficient of $\varepsilon = 12{,}000\ M^{-1}cm^{-1}$ for FAD (HpxL extracted flavin = 37.1 μM and PcxL extracted flavin = 30.9 μM). Extracted flavin (10 μL) was injected onto a 150 x 4.6 mm 4-micron Synergi 4u Fusion-RP column connected to an Agilent 1200 series LC/MSD SL mass spectrometer (Agilent, Santa Clara, CA) monitoring for the masses corresponding to FAD (m/z = 786.2) and FMN (m/z = 457.1) in positive mode. The following solvents were used: A = 95% ddH_2O / 5% acetonitrile with 0.1% formic acid and B = 5% ddH_2O / 95% acetonitrile with 0.1% formic acid. The sample was run with a linear gradient from time 0-12 minutes: time 0 min = 5% B, 12 min = 95% B, 15 min = 95% B, 18 min = 5% B, and 25 min = 5% B. Flow rate was 1 mL / minute.

HpxL (3.6 mg) and PcxL (6.1 mg) stocks were used to determine the absorption spectra of the proteins on a Cary 4000 UV-Vis spectrophotometer (Agilent, Santa Clara, CA) scanning from 200 nm – 800 nm. NADPH was added to the protein to a final concentration of 500 μM and the absorption spectrum of the protein was measured after one minute of incubation. Flavin that was washed away from the proteins was also measured.

3.5.6 Stoichiometry determination

Oxygen consumption was monitored on a Clark-type oxygen electrode (Hansatech Instruments, Norfolk, England). NADPH consumption was monitored (A_{340}) on an Cary 4000 UV-Vis spectrophotometer (Agilent, Santa Clara, CA). Reaction components included 50 mM sodium phosphate buffer (air-saturated), pH 7.8, 300 µM NADPH, 50 µM FAD, 50 mM 2AEPn, and 1 µM PcxL or 5 µM HpxL. Reactions were initiated by addition of enzyme and monitored for 5 minutes. Three replicates were performed on each instrument. The number of moles of both oxygen and NADPH consumed by the enzyme in 5 minutes was determined and compared.

NADPH consumption was determined using a Cary 4000 UV-Vis spectrophotometer (Agilent, Santa Clara, CA) and via ^{31}P NMR. 2AEPn consumption was determined using ^{31}P NMR. Reaction components included 50mM sodium phosphate buffer (air-saturated), pH 7.8, 3 mM NADPH, 500 µM FAD, 5 mM 2AEPn, and 25 µM PcxL or 25 µM HpxL. Three replicates were performed. The reactions incubated for 30 minutes at room temperature and were quenched using one reaction volume of 100% methanol. Dimethylphosphinic acid was added to the reaction as an internal quantification standard to yield a final concentration of 10 mM. The number of mols of substrate and NADPH consumed were determined and directly compared.

3.5.7 Generation of (*S*)-1H2AEPn by HpxV

To yield an N-terminal 6xHis-tagged HpxV, *hpxV* was amplified using the primers Orf10_LIC_F (5'-TACTTCCAATCCAATGCAcctactcacccacgcactc-3') and Orf10_LIC_R (5'-TTATCCACTTCCAATGTTATTActacttgctcgaccacatg-3') and was cloned into the ligation independent cloning vector pET His6 TEV LIC cloning vector

(2B-T). N_{6xHis}-HpxV was overexpressed in transformed *Escherichia coli* BL21 DE3 containing the pET His6 TEV LIC cloning vector (2B-T) with *hpxV*. HpxV was overexpressed and purified using method two as described above. (*S*)-1H2AEPn was generated using HpxV, a 2AEPn-dependent α-ketoglutarate-dependent dioxygenase from *S. regensis* NRRL WC-3744. The reactions contained 50 mM phosphate buffered saline, pH 7.4, 40 mM α-ketoglutarate, 1 mM iron (II) ammonium sulfate, 7.5 mM 2AEPn, 4 mM L-ascorbic acid, and 200 μM HpxV. Reactions were incubated for 13 hours at 30°C. The reaction was quenched by precipitating the protein by addition of methanol to yield a 90% methanol precipitation. Precipitated protein was removed via centrifugation and the reaction mixture was concentrated on a rotary evaporator. The concentration of 1H2AEPn generated was determined by ^{31}P NMR spectroscopy, using phosphite as an internal reference standard for quantification at a final concentration of 10 mM. In some experiments synthetic (*S*)-1H2AEPn was used (see below).

3.5.8 Product formation experiments

HpxL or PcxL were incubated with 2AEPn or HpxV-generated (*S*)-1H2AEPn to monitor what products are formed by the *N*-oxidases. The reactions were set up with the following components: 25 mM sodium phosphate buffer, pH 7.8, 500 μM NADPH, 100 μM FAD, 25 μM PcxL or HpxL, 10 μM phosphite dehydrogenase, 25 mM phosphite, and 3 mM of substrate at pH 7.8. Reactions were incubated for 16 hours at room temperature. Product formation was monitored using ^{31}P NMR spectroscopy, ^{1}H NMR spectroscopy, and ^{31}P-^{1}H HMBC spectroscopy. A mutated phosphite dehydrogenase (PTDH17x) that can use either NADH or NADPH as the nicotinamide coenzyme(47) was used to regenerate NADPH in the substrate specificity experiments. This enzyme was

overexpressed in *E. coli* BL21 DE3 cells containing a plasmid with PTDH17x cloned into pET15b. Cells were grown in modified Studier's ZYM5052 medium (as described above) at 37°C until OD_{600} reached 0.6-0.8, the culture was shifted to 18°C, and protein was overexpressed for 16 hours. The cell suspension was passed two-times through a French Press cell disruptor (Thermo Electron Corporation, Waltham, MA) at 1000 psi. The suspension was centrifuged at 14,000 x g for a minimum of 30 minutes and the pellet was discarded. Soluble cell lysate was passed over free Ni-NTA IMAC resin and washed with wash buffer (20 mM Tris-HCl, 500 mM NaCl, and 10% glycerol (% v/v), pH 7.6). Bound protein was washed stepwise with the wash buffer containing increasing levels of imidazole at concentrations of 10 mM, 25 mM, and 50 mM respectively until there was no protein detected by monitoring A_{260} nm. Protein was eluted from the column using wash buffer containing 250 mM imidazole, concentrated using a 10 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA). The protein was exchanged into protein storage buffer (50 mM HEPES, 200 mM KCl, and 10% glycerol (% v/v), pH 7.5) using a PD-10 desalting column (GE Healthcare Life Sciences, Marlborough, MA).

3.5.9 Stereospecific synthesis of (*R*)- and (*S*)-1H2AEPn

Enantiopure (ee> 99%) (*R*)- and (*S*)-1H2AEPn was generated following previously published methods(39).

3.5.10 Kinetic assays

Enzyme assays were monitored using a Cary 4000 UV-Vis spectrophotometer (Agilent, Santa Clara, CA). NADPH consumption was monitored by observation of the absorption at 340 nm. Reaction components included: 50 mM sodium phosphate buffer, pH 7.8, 300 µM NADPH, 50 µM FAD, substrate concentrations ranging from 62.5 µM to 150 mM and enzyme concentrations of 1 µM or 5 µM for PcxL and 5 µM or 10 µM for HpxL. Briefly, reaction components except for substrate and enzyme were mixed in a 200 µL quartz cuvette. Enzyme was added and the background rate of the enzyme was collected for approximately one minute. Reactions were initiated upon addition of substrate. Three replicates were obtained for each concentration of substrate. Due to the time it takes to complete one NADPH consumption curve for HpxL (between 15-30 minutes), all the initial rates are approximate initial rates. The rate over approximately 30 seconds once the enzyme reaches steady-state was read and this was used as the "initial rate".

3.5.11 Construction of phylogenetic trees and assignment of genomic neighborhoods

To construct the *N*-oxidase phylogenetic tree, the top 500 BLAST (2) hits obtained using HpxL (WP_030990672.1) as a query sequence were downloaded on February 18th, 2017. These sequences were aligned using MAFFT with default settings (22) and the maximum likelihood tree was constructed using FastTree using default settings (36). Accession numbers of the 500 amino acid sequences used in the tree were run through RODEO (44) to identify putative *N*-oxidase proteins involved in a phosphonate gene cluster. A positive hit constitutes a phosphoenolpyruvate mutase enzyme being within 15 genes upstream or downstream of the *N*-oxidase.

To construct the PepM phylogenetic tree, five Actinobacteria PepM amino acid sequences were used as a BLAST query (21): *Streptomyces luridus* PepM (ACZ13456), *Streptomyces viridochromogenes* (AAU00071), *Streptomyces fradiae* (ACG70831), *Strepomyces rubellomurinus* (ABB90393), and *Streptomyces regensis* HpxU (WP_030646403.1). The search was limited to Actinobacteria, with Mycobacteria being excluded from the search. All hits for each search were sorted to remove duplicates and the characteristic EDKxxxxxNS motif was manually identified to distinguish actual PepMs from other members of the isocitrate lyase superfamily (9, 32, 48). There were 199 amino acid sequences identified using this approach as putative PepMs. These amino acid sequences were aligned using MAFFT (22) and the maximum likelihood tree was constructed using FastTree (36). RODEO (44) was used to identify phosphonate gene clusters that contain an *N*-oxidase homolog. A positive hit constitutes an *N*-oxidase being within 15 genes upstream or downstream of the PepM.

3.6 REFERENCES

1. Aires A, Mota VR, Saavedra MJ, Monteiro AA, Simões M, et al. 2009. Initial in vitro evaluations of the antibacterial activities of glucosinolate enzymatic hydrolysis products against plant pathogenic bacteria. *J. Appl. Microbiol.* 106(6):2096–2105.

2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215(3):403–10.

3. Anderson VE, Weiss PM, Cleland WW. 1984. Reaction intermediate analogs for enolase. *Biochemistry*. 23(12):2779–86.

4. Arabshahi L, Schmitz FJ. 1987. Brominated tyrosine metabolites from an unidentified sponge. *J. Org. Chem.* 52(16):3584–86.

5. Bacher M, Brader G, Hofer O, Greger H. 1999. Oximes from seeds of Atalantia ceylanica. *Phytochemistry*. 50(6):991–94.

6. Bennett R, Donald A, Dawson G, Hick A, Wallsgrove R. 1993. Aldoxime-Forming Microsomal Enzyme Systems Involved in the Biosynthesis of Glucosinolates in Oilseed Rape (Brassica napus) Leaves. *Plant Physiol.* 102(4):1307–12.

7. Bennett RN, Kiddle G, Hick AJ, Dawson GW, Wallsgrove RM. 1996. Distribution and activity of microsomal NADPH-dependent monooxygenases and amino acid decarboxylases in cruciferous and non-cruciferous plants, and their relationship to foliar glucosinolate content. *Plant Cell Environ.* 19(7):801–12.

8. Chen CCH, Zhang H, Kim AD, Howard A, Sheldrick GM, et al. 2002. Degradation Pathway of the Phosphonate Ciliatine: Crystal Structure of 2-Aminoethylphosphonate Transaminase,. *Biochemistry*. 41(44):13162–69.

9. Chen CM, Ye QZ, Zhu ZM, Wanner BL, Walsh CT. 1990. Molecular biology of carbon-phosphorus bond cleavage. Cloning and sequencing of the phn (psiD) genes involved in alkylphosphonate uptake and C-P lyase activity in Escherichia coli B. *J. Biol. Chem.* 265(8):4461–71.

10. Chen H-P, Zhao Z-Z, Li Z-H, Dong Z-J, Wei K, et al. 2016. Novel Natural Oximes and Oxime Esters with a Vibralactone Backbone from the Basidiomycete Boreostereum vibrans. *ChemistryOpen*. 5(2):142–49.

11. Cioni JP, Doroghazi JR, Ju K-S, Yu X, Evans BS, et al. 2014. Cyanohydrin Phosphonate Natural Product from Streptomyces regensis. *J. Nat. Prod.* 77(2):243–49.

12. Cortina NS, Revermann O, Krug D, Müller R. 2011. Identification and Characterization of the Althiomycin Biosynthetic Gene Cluster in *Myxococcus xanthus* DK897. *ChemBioChem*. 12(9):1411–16.

13. Du L, Lykkesfeldt J, Olsen CE, Halkier BA. 1995. Involvement of cytochrome P450 in oxime production in glucosinolate biosynthesis as demonstrated by an in vitro microsomal enzyme system isolated from jasmonic acid-induced seedlings of Sinapis alba L. *Proc. Natl. Acad. Sci. U. S. A.* 92(26):12505–9.

14. Dumora C, Lacoste A-M, Cassaigne A. 1983. Purification and Properties of 2-Aminoethylphosphonate: Pyruvate Aminotransferase from Pseudomonas aeruginusa. *Eur. J. Biochem.* 133(1):119–25.

15. Eyer P. 2003. The Role of Oximes in the Management of Organophosphorus Pesticide Poisoning. *Toxicol. Rev.* 22(3):165–90.

16. Garcia I, Vior NM, González-Sabín J, Braña AF, Rohr J, et al. 2013. Engineering the Biosynthesis of the Polyketide-Nonribosomal Peptide Collismycin A for Generation of Analogs with Neuroprotective Activity. *Chem. Biol.* 20(8):1022–32.

17. Halkier BA, Gershenzon J. 2006. Biology and Biochemistry of Glucosinolates. *Annu. Rev. Plant Biol.* 57(1):303–33.

18. Hong S, Shin Y, Jung M, Ha MW, Park Y, et al. 2015. Efficient synthesis and biological activity of Psammaplin A and its analogues as antitumor agents. *Eur. J. Med. Chem.* 96:218–30.

19. Huang Z, Wang K-KA, Donk WA van der. 2016. New insights into the biosynthesis of fosfazinomycin. *Chem. Sci.* 7(8):5219–23.

20. Jang J-H, van Soest RWM, Fusetani N, Matsunaga S. 2007. Pseudoceratins A and B, Antifungal Bicyclic Bromotyrosine-Derived Metabolites from the Marine Sponge Pseudoceratina purpurea. *J. Org. Chem.* 72(4):1211–17.

21. Ju K-S, Gao J, Doroghazi JR, Wang K-KA, Thibodeaux CJ, et al. 2015. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc. Natl. Acad. Sci.* 112(39):12175–80.

22. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30(4):772–80.

23. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12):1647–49.

24. Kelley L. *PHYRE2 Protein Fold Recognition Server*. http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index.

25. Kelly WL, Townsend CA. 2002. Role of the Cytochrome P450 NocL in Nocardicin A Biosynthesis. *J. Am. Chem. Soc.* 124(28):8186–87.

26. Kelly WL, Townsend CA. 2005. Mutational Analysis of nocK and nocL in the Nocardicin A Producer Nocardia uniformis. *J. Bacteriol.* 187(2):739–46.

27. Kim AD, Baker AS, Dunaway-Mariano D, Metcalf WW, Wanner BL, Martin BM. 2002. The 2-Aminoethylphosphonate-Specific Transaminase of the 2-Aminoethylphosphonate Degradation Pathway. *J. Bacteriol.* 184(15):4134–40.

28. Kim YS, Jeong H-Y, Kim AR, Kim W-H, Cho H, et al. 2016. Natural product derivative BIO promotes recovery after myocardial infarction via unique modulation of the cardiac microenvironment. *Sci. Rep.* 6:30726.

29. Lacoste AM, Dumora C, Balas L, Hammerschmidt F, Vercauteren J. 1993. Stereochemistry of the reaction catalysed by 2-aminoethylphosphonate aminotransferase. A 1H-NMR study. *Eur. J. Biochem.* 215(3):841–44.

30. Lin J, Cashman JR. 1995. N-Oxygenation of Phenethylamine to the trans-Oxime by Adult Human Liver Flavin-Containing Monooxygenase and Retroreduction of Phenethylamine Hydroxylamine by Human Liver Microsomes. *J. Pharmacol. Exp. Ther*. 282:1269–1279.

31. Liu N, Song L, Liu M, Shang F, Anderson Z, et al. 2015. Unique post-translational oxime formation in the biosynthesis of the azolemycin complex of novel ribosomal peptides from *Streptomyces* sp. FXJ1.264. *Chem. Sci.* 7(1):482–88.

32. Metcalf WW, van der Donk WA. 2009. Biosynthesis of Phosphonic and Phosphinic Acid Natural Products. *Annu. Rev. Biochem.* 78:65–94.

33. Moya P, Castillo M, Primo-Yúfera E, Couillaud F, Martínez-Máñez R, et al. 1997. Brevioxime: A New Juvenile Hormone Biosynthesis Inhibitor Isolated from *Penicillium brevicompactum*. *J. Org. Chem.* 62(24):8544–45.

34. Pallitsch K, Happl B, Stieger C. 2017. Determination of the absolute configuration of (-)-hydroxynitrilaphos and related biosynthetic questions. *Chem. Eur. J.* 23(62): 15655-15665.

35. Pretsch E, Buhlmann P, Badertscher M. 2009. *Structure Determination of Organic Compounds*. 433 pp. 4th ed.

36. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One*. 5(3):e9490.

37. Sánchez-Pujante PJ, Borja-Martínez M, Pedreño MÁ, Almagro L. 2017. Biosynthesis and bioactivity of glucosinolates and their production in plant in vitro cultures. *Planta*. 246(1):19–32.

38. Sotelo T, Lema M, Soengas P, Cartea ME, Velasco P. 2015. In Vitro Activity of Glucosinolates and Their Degradation Products against Brassica-Pathogenic Bacteria and Fungi. *Appl. Environ. Microbiol.* 81(1):432–40.

39. Staalduinen LM van, McSorley FR, Schiessl K, Séguin J, Wyatt PB, et al. 2014. Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates. *Proc. Natl. Acad. Sci.* 111(14):5171–76.

40. Stotz HU, Sawada Y, Shimada Y, Hirai MY, Sasaki E, et al. 2011. Role of camalexin, indole glucosinolates, and side chain modification of glucosinolate-derived isothiocyanates in defense of Arabidopsis against Sclerotinia sclerotiorum. *Plant J. Cell Mol. Biol.* 67(1):81–93.

41. Studier FW. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41(1):207–34.

42. Sugai Y, Katsuyama Y, Ohnishi Y. 2016. A nitrous acid biosynthetic pathway for diazo group formation in bacteria. *Nat. Chem. Biol.* 12(2):73–75.

43. Textor S, Gershenzon J. 2009. Herbivore induction of the glucosinolate–myrosinase defense system: major trends, biochemical bases and ecological significance. *Phytochem. Rev.* 8(1):149–70.

44. Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, et al. 2017. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* 13(5):470-478.

45. Walsh CT, Wencewicz TA. 2012. Flavoenzymes: Versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* 30(1):175–200.

46. Wong L, Radić Z, Brüggemann RJM, Hosea N, Berman HA, Taylor P. 2000. Mechanism of Oxime Reactivation of Acetylcholinesterase Analyzed by Chirality and Mutagenesis. *Biochemistry.* 39(19):5750–57.

47. Woodyer RD, Shao Z, Thomas PM, Kelleher NL, Blodgett JAV, et al. 2006. Heterologous Production of Fosfomycin and Identification of the Minimal Biosynthetic Gene Cluster. *Chem. Biol.* 13(11):1171–82.

48. Yu X, Doroghazi JR, Janga SC, Zhang JK, Circello B, et al. 2013. Diversity and abundance of phosphonate biosynthetic genes in nature. *Proc. Natl. Acad. Sci.*, p. 201315107.

49. Zhu Y, Zhang Q, Li S, Lin Q, Fu P, et al. 2013. Insights into Caerulomycin A Biosynthesis: A Two-Component Monooxygenase CrmH-Catalyzed Oxime Formation. *J. Am. Chem. Soc.* 135(50):18750–53.

Chapter 4: Biosynthesis of the broad-spectrum herbicide, phosphonothrixin

4.1 INTRODUCTION

4.1.1 Phosphonic acids used as herbicides

Phosphonic acids have renowned success as herbicides. Glyphosate, trademarked under the names of RoundUp®, Rodeo®, Accord®, Touchdown®, and Shackle®, is one of the most widely used herbicides available on the market today and was developed by Monsanto in 1970 (8). This synthetic phosphonate is a transition-state analog of phosphoenolpyruvate and targets enolpyruvylshikimate-3-phosphate (EPSP) synthase, which is part of the shikimate pathway of aromatic amino acid biosynthesis (32). There are now a plethora of RoundUp Ready® crops containing a glyphosate-resistant copy of EPSP synthase in order to make weed-treatment easy for farmers (12). Upwards of 90% of cotton, maize, and soybeans planted in 2014 were RoundUp Ready®, glyphosate-resistant crops (3).

Phosphinothricin, also known as glufosinate, is trademarked under the names of Finale®, Basta®, and Liberty®. This natural product was first identified to be produced by *Streptomyces viridochromogenes* Tü494 (2) and *Streptomyces hygroscopicus* (26). Phosphinothricin mimics the tetrahedral intermediate of glutamine synthase (13). Like glyphosate, glufosinate is a potent herbicide (14), with many crops engineered containing the phosphinothricin acetyltransferase gene (*bar* or *pat*) under the trademarks LibertyLink® and WideStrike®. In *Streptomyces*, this gene confers self-resistance to phosphinothricin (23, 35). Because of the wide-spread use of these compounds, resistance to these herbicides is rapidly increasing. Recent studies have found that levels of

glyphosate and glufosinate resistance in weeds is rising, along with almost all other herbicides used on the market today. Resistance has been observed for 22 out of the 25 different classes of herbicides used (20). Currently, estimates on global crop production suggest that weeds cause an annual crop loss of approximately 12% (27) and cost the agriculture industry on average two billion USD each year (20), and this cost will likely rise with an increase in the number of herbicide-resistant weeds. For this reason, we are in need of new herbicidal compounds for use in agriculture.

4.1.2 Phosphonothrixin

Phosphonothrixin was originally discovered in 1995 by Takahashi and colleagues during a search for novel herbicides (18, 34). Like glyphosate and phosphinothricin, this compound shows broad spectrum herbicidal activity against a variety of different weed (34); however, the mode of action of phosphonothrixin remains unknown. Initial reports hypothesized that phosphonothrixin may target 1-deoxyxylulose 5-phosphate reductoisomerase (DXR) because of its structural similarity to the deoxyxylulose phosphate (DXP), but this was shown to not be the target of phosphonothrixin in peppermint secretory cells (19). Due to the lack of understanding of the biosynthesis and mode-of-action of phosphonothrixin, development of this compound into a commercially viable herbicide currently seems unlikely. Identification of the biosynthetic gene cluster in 2014 by Lin and colleagues from *Saccharothrix* sp. ST-888 (21) was a major step for characterization of the biosynthesis of this compound. Upon discovery of the genes required for the biosynthesis of this compound, it might be possible to identify other organisms that contain the potential to produce phosphonothrixin at higher levels than

Saccharothrix sp. ST-888 or to identify a resistance gene for this compound if the target is conserved among diverse organisms.

4.1.3 Biosynthesis of phosphonothrixin

The phosphonate gene cluster in *Saccharothrix* sp. ST-888 appears very different than characterized phosphonate biosynthetic pathways. This gene cluster lacks homologous enzymes to the two known enzymes used to convert phosphonopyruvate into one of two known products, driving the highly unfavorable PEP mutase reaction forward (5): there are no phosphonopyruvate decarboxylase (31) or phosphonomethylmalate synthase (9) homologs within the gene cluster. Lin and colleagues identified the minimal biosynthetic gene cluster through heterologous expression and have proposed a phosphonothrixin biosynthetic scheme that includes the condensation of phosphonopyruvate with pyruvate by a thiamine pyrophosphate-dependent enzyme (21). This chapter will focus on the characterization of the phosphonothrixin biosynthetic gene cluster and will provide evidence for an alternative biosynthetic pathway for phosphonothrixin biosynthesis.

4.2 RESULTS AND DISCUSSION

4.2.1 The phosphonothrixin gene cluster and identification of other organisms with the genetic capacity to produce phosphonothrixin

Sequencing and identification of the phosphonothrixin gene cluster in *Saccharothrix* sp. ST-888 allowed us to identify other organisms that have the potential to produce phosphonothrixin. In total, 16 different organisms contain the genes that share >75% amino acid identity with those found in the *Saccharothrix* sp. ST-888

phosphonothrixin gene cluster (data not shown). The synteny and organization for fourteen of these gene clusters are identical to the *Saccharothrix* sp. ST-888 gene cluster (Fig 4.1). One organism, *Saccharopolyspora spinosa* NRRL 18395, contains many of the genes required for phosphonothrixin biosynthesis, although the amino acid identity of these proteins is much lower than the other 14 organisms with the phosphonothrixin gene cluster.

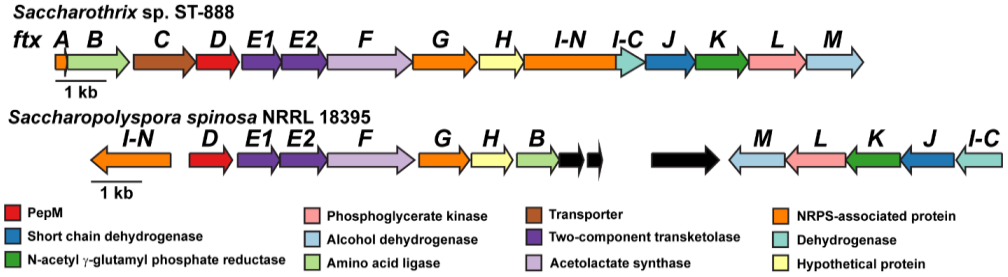


Figure 4.1. The phosphonothrixin gene cluster from *Saccharothrix* sp. ST-888 and putative phosphonothrixin gene cluster from *Saccharopolyspora spinosa* NRRL 18395.

4.2.2 *In vitro* testing of the function of the acetolactate synthase homolog, FtxF[2]

Lin and colleagues suggest that phosphonothrixin biosynthesis proceeds through the condensation of phosphonopyruvate (PnPy) and pyruvate, which can subsequently be reduced to form phosphonothrixin (21). We hypothesized that one of the enzymes that might catalyze this reaction is the thiamine pyrophosphate (TPP)-dependent acetolactate synthase homolog, FtxF. TPP-dependent enzymes catalyze carbon-carbon bond formation through the oxidative condensation of two ketone-containing molecules (6). To test this hypothesis, we incubated 6xHis-tagged FtxF from *Saccharothrix* sp. ST-888 in

[2] *I would like to thank Michael Steven Carter for the* Bacillus subtilis *ALS protein and Shihui Dong for the construction of the FtxF expression plasmid and protein.*

the presence of a previously characterized PEP mutase homolog, RhiH (4), PEP and pyruvate. With this reaction mixture, we do not observe conversion of PEP into any product (Fig 4.2). However, when we incubate PEP with RhiH (4) and the phosphonopyruvate decarboxylase from *Bacteroides fragilis* (Bf-Ppd) (36), we see nearly complete conversion of PEP into phosphonoacetaldehyde (Fig 4.2), which indicates that RhiH is actively converting PEP into phosphonopyruvate. Therefore, it is likely that if the substrate for FtxF were phosphonopyruvate, we should see conversion of PEP into another compound corresponding to the appearance of a new peak in the ^{31}P NMR spectrum. Thus, we do not believe that phosphonopyruvate is the *in vivo* substrate of FtxF.

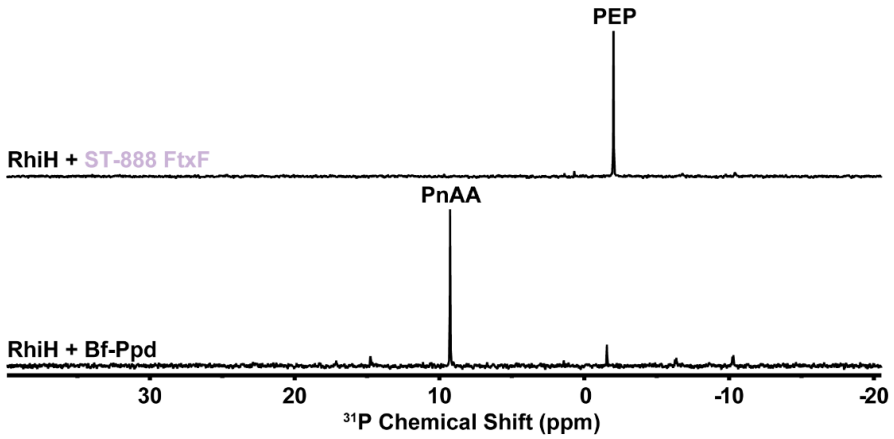


Figure 4.2. Reaction of FtxF from *Saccharothrix* sp. ST-888 with PEP in the presence of RhiH (top). Reaction of RhiH with phosphonopyruvate decarboxylase from *Bacteroides fragilis* (Bf-Ppd) (bottom). The reactions were set up with the following conditions in 50 mM HEPES buffer with 10 mM magnesium chloride, pH 7.5: 10 mM PEP, 50 μg of each enzyme, 1 mM TPP, and 10 mM sodium pyruvate (ALS reaction only).

Dihydroxypropylphosphonic acid (DHPPA) is a compound isolated from the extracts of *Saccharothrix* sp. ST-888 (will be discussed in section 4.2.4). Based on the observation that FtxF does not utilize phosphonopyruvate, I tested the idea that DHPPA

may be a substrate for this enzyme. To do this, I incubated FtxF with DHPPA and sodium pyruvate, but again was unable to detect any product formation (Fig 4.3).

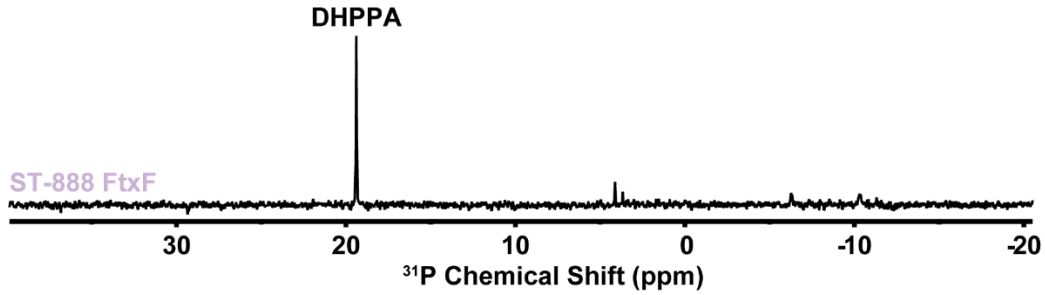


Figure 4.3. Reaction of FtxF from *Saccharothrix* sp. ST-888 with DHPPA and pyruvate. The reactions were set up with the following conditions in 50 mM HEPES buffer with 10 mM magnesium chloride, pH 7.5: 500 μM DHPPA, 50 μg FtxF, 1 mM TPP, and 1 mM sodium pyruvate.

After determining that FxF does not use DHPPA or phosphonopyruvate as a substrate with pyruvate, I tested the idea that FtxF might be a bonafide acetolactate synthase (ALS), which is the first committed step of branch-chain amino acid biosynthesis. ALS condenses two molecules of pyruvate together to form acetolactate (6). Thus, I incubated FtxF with pyruvate. While we observed formation of acetolactate with FtxF from either *Saccharothrix* sp. ST-888 or *S. spinosa* NRRL 18395, FtxF did not form as much acetolactate as a canonical acetolactate synthase enzyme isolated from *Bacillus subtilis* (Fig 4.4). It should be noted that these reactions are end-point monitoring of product formation and give us little indication of the rate of the reactions, so it is plausible that the *B. subtilis* ALS functions more quickly than FtxF.

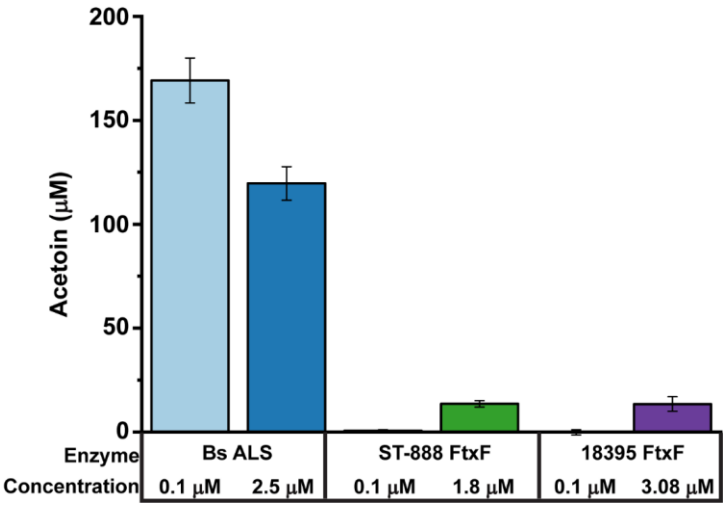


Figure 4.4. Acetolactate production by FtxF from *Saccharothrix* sp. ST-888, *Saccharopolyspora spinosa* NRRL 18395, or an acetolactate synthase from *Bacillus subtilis*. The Westerfeld acetoin detection assay(33) was used to determine if acetolactate was formed by the enzymes by measuring levels of acetoin after chemical oxidative decarboxylation of acetolactate. Reactions were set up with the following components in 100 mM potassium phosphate buffer pH 7.5: 20 mM sodium pyruvate, 237 μM thiamine pyrophosphate, 10 mM magnesium chloride. Reactions were performed in duplicate.

Based on these data, we propose two hypotheses for the function of FtxF. The first is that FtxF may utilize pyruvate as at least one of its substrates and the other may be a phosphonic acid that we have not yet identified. Second, FtxF may only utilize pyruvate as a substrate. Structurally, phosphonothrixin is similar to acetolactate (Fig 4.5). If phosphonothrixin is an inhibitor of acetolactate synthase, which is required for the biosynthesis of branched-chain amino acids in bacteria (22), the host organism would need a way to provide self-resistance to this compound during its biosynthesis. One way to provide this is to encode a copy of the target that is unable to be inhibited by phosphonothrixin. *Saccharothrix* sp. ST-888 encodes several acetolactate synthase homologs within its genome (data not shown), which suggests that FtxF is not the canonical ALS responsible for branched-chain amino acid biosynthesis. Even though activity of the enzyme appears lower than a canonical acetolactate synthase from another

organism (Fig 4.4), there is some function of the enzyme. This is consistent with resistance proteins found in other natural product gene clusters. A producing organism may produce a resistant copy of the target protein that cannot be inhibited by the natural product (15).

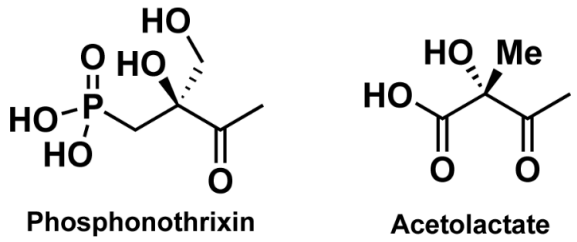


Figure 4.5. Structures of phosphonothrixin and acetolactate.

4.2.3 Phylogeny of FtxF shows it is related to acetolactate synthase proteins

To examine the relatedness of FtxF to proteins with known function, we performed a BLAST (24) search against the Swiss-Prot/UniProt database using *Saccharothrix* sp. ST-888 FtxF (Accession number WP_045300935.1). FtxF shareas a common ancestor with acetolactate synthase proteins (Fig 4.6). However, when performing a BLAST search using FtxF as a query against the non-redundant database, we find that none of the proteins on the Swiss-Prot tree are represented, and only uncharacterized enzymes are retrieved from BLAST when looking for the closest 250 homologs to FtxF (Fig 4.7). These data do not help support nor contradict our two hypotheses proposed for the function of FtxF *in vivo*, because we observe that FtxF shares a common ancestor with canonical ALS proteins; however, these proteins are not the closest homologs of FtxF in the NCBI database. Section 4.2.6 will detail experiments that lead us to believe ALS is the target of phosphonothrixin.

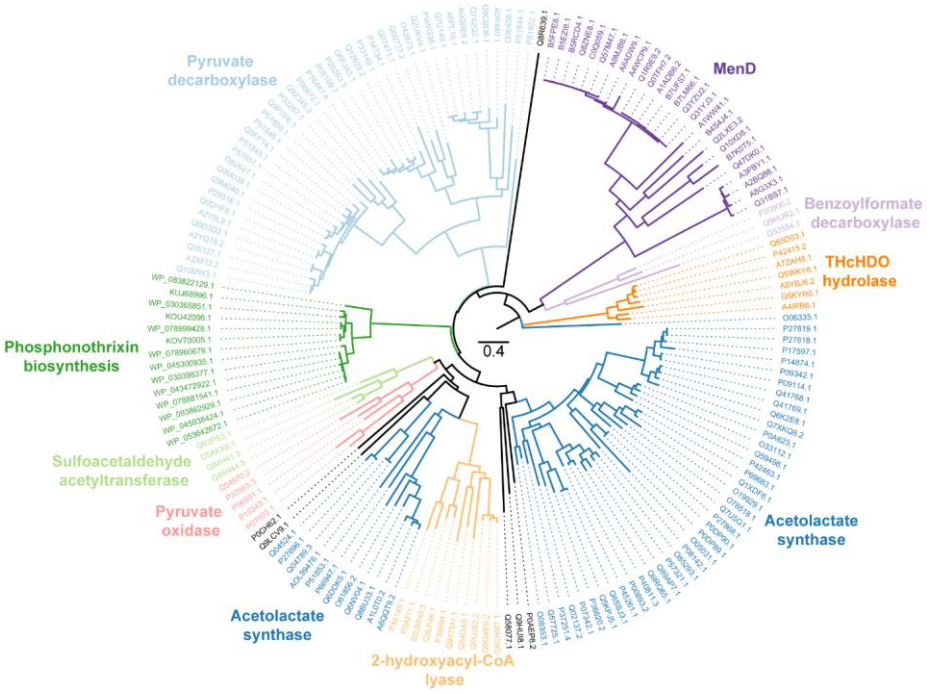


Figure 4.6. Maximum-likelihood phylogenetic tree of FtxF homologs found within the SwissProt/UniProt database. Green represents FtxF proteins. Light green represents sulfoacetylaldehyde acetyltransferases, pink shows pyruvate oxidases, blue represents acetolactate synthase, light blue are pyruvate decarboxylases, orange represents 3D-(3,5/4)-trihydroxycyclohexane-1,2-dione (THcHDO) hydrolases, light orange are 2-hydroxyacyl-CoA lyases, purple are MenD proteins, and light purple are benzoylformate decarboxylases.

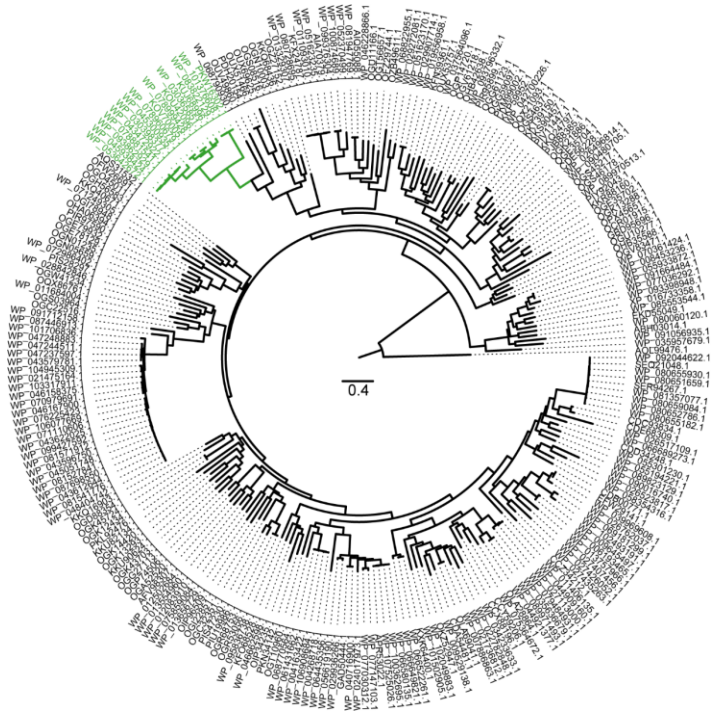


Figure 4.7. Maximum-likelihood phylogenetic tree the top 250 homologs of FtxF in the NCBI non-redundant database(16). Green represents FtxF proteins.

4.2.4 *Saccharothrix* sp. ST-888 produces DHPPA and phosphonothrixin[3]

To further characterize phosphonothrixin biosynthesis, we examined the different phosphonic acids produced by *Saccharothrix* sp. ST-888. When grown in GUBC and R2AS media, *Saccharothrix* sp. ST-888 produces over ten phosphonic acids (data not shown). Two of these, were identified as phosphonothrixin and dihydroxypropylphosphonic acid (DHPPA) (Fig 4.8 and Fig 4.9), based on spectroscopic data (Fig 4.9) (Table 4.1) (16) (Fig 4.9) (Table 4.2) (10).

[3] *I would like to thank Ryuichi Hirota for collaboration on this project. Scale up and initial partial purification of phosphonates was done by him.*

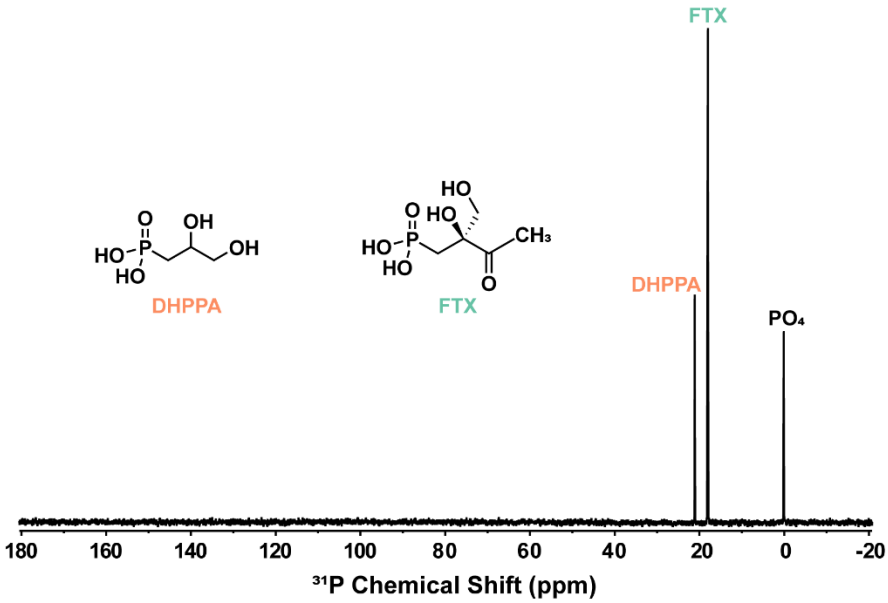


Figure 4.8. ^{31}P NMR spectrum of compounds isolated from *Saccharothrix* sp. ST-888.

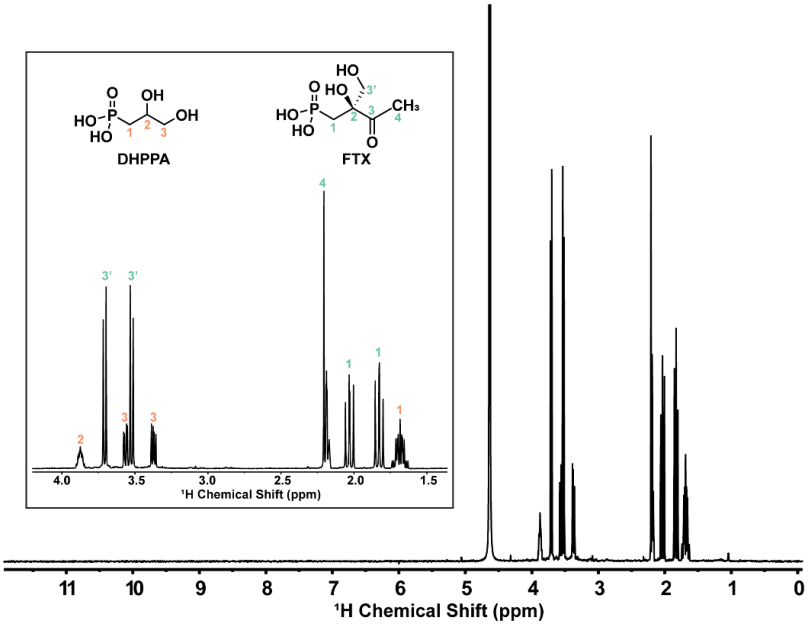
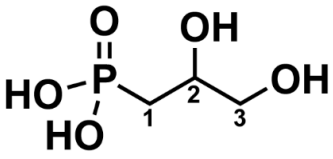
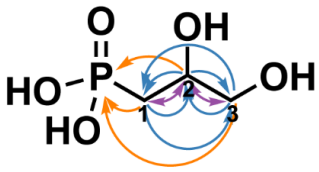


Figure 4.9. ^{1}H NMR spectrum of compounds isolated from *Saccharothrix* sp. ST-888. The inset shows the proton assignments for each molecule.

Table 4.1. ^{1}H (600 MHz), ^{31}P (242 MHz), and ^{13}C (150 MHz) spectroscopic data for DHPPA in D_2O. ^{13}C chemical shifts were inferred from a ^{1}H-^{13}C HSQSAD spectrum. The respective ^{1}H correlations are shown with colored arrows. Orange denotes ^{1}H-^{31}P correlations, blue denotes ^{1}H-^{13}C correlations, while purple denotes ^{1}H-^{1}H correlations. The 2D-NMR experiments are shown in Fig 4.18 – Fig 4.21.



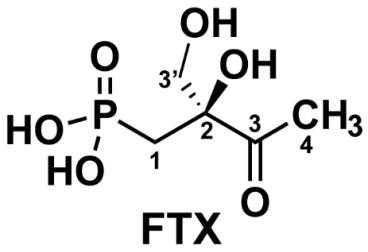
DHPPA



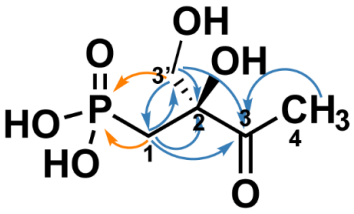
DHPPA

No.	δ_C, mult	δ_H (J in Hz)	COSY	HMBC
1	32.2, CH2	1.68, m (17.4, 10.2, 6.7)	2	2, 3
2	68.3, CH	3.87, m (10.1, 6.7, 3.5)	1, 3	1, 3
3	65.9, CH2	3.37, dd (11.8, 6.8) 3.56, dd (11.8 , 3.6)	2	1, 2
P	21.1, td (J=17.9, 9.3 Hz)	-	-	1,2

Table 4.2. ^{1}H (600 MHz), ^{31}P (242 MHz), and ^{13}C (150 MHz) spectroscopic data for phosphonothrixin in D_2O. ^{13}C chemical shifts were inferred from a ^{1}H-^{13}C HSQSAD and ^{1}H-^{13}C HMBC spectra. Orange denotes ^{1}H-^{31}P correlations while blue denotes ^{1}H-^{13}C correlations. The 2D-NMR experiments are shown in Fig 4.18 – Fig 4.21.



FTX



FTX

No.	δ_C, mult	δ_H (J in Hz)	COSY	HMBC
1	32.8, CH2	1.83, dd (16.8, 15.3) 2.03, dd (17.6, 15.4)	-	2, 3
2	80.8, qC	-	-	-
3	215.4, qC	-	-	-
4	26.0, CH3	2.21, s	-	2, 3
3'	67.7, CH2	3.52, d (11.8) 3.71, d (11.8)	-	2, 4
P	18.03, t (J=17.2 Hz)	-	-	1,3'

Isolation of DHPPA and phosphonothrixin from the same extract suggests that DHPPA is involved in the biosynthesis of phosphonothrixin. DHPPA is also a putative intermediate in the biosynthesis of valinophos (Fig 1.12), isolated from culture extracts of *Streptomyces durhamensis* NRRL B-3309 (16). Comparison of the phosphonate biosynthetic gene clusters from *Saccharothrix* sp. ST-888 and *S. durhamensis* B-3309 revealed a set of five conserved genes (Fig 1.12), which were proposed encode enzymes needed for production of DHPPA during valinophos biosynthesis (16). Based on this idea, it seems reasonable to suggest that DHPPA is also made by *Saccharothrix* sp. ST-888. If so, the published putative phosphonothrixin biosynthetic pathway may be incorrect (21) and phosphonothrixin is biosynthesized in a similar fashion to valinophos (Fig 1.12) (Fig 4.13) (16).

4.2.5 FtxJ is a metal-dependent dehydrogenase that uses DHPPA as a substrate to generate 2OPPA[4]

If DHPPA is an intermediate during the biosynthesis of phosphonothrixin, a plausible route would involve oxidation to 3-hydroxy-2-oxopropylphosphonic acid (2OPPA), which could then be condensed with pyruvate to form phosphonothrixin. Moreover, the FtxJ protein, encoded within the gene cluster is homologous to L-iditol dehydrogenase, which catalyzes the analogous oxidation of sorbitol to fructose in an NAD^+ and Zn^{2+}-dependent reaction (25) (Fig 4.10).

[4] *I would like to thank Matthew Georgiou for construction of the FtxJ expression plasmid.*

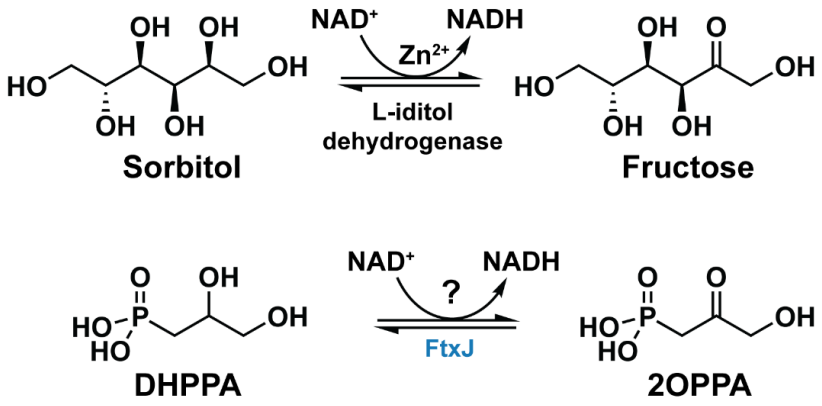


Figure 4.10. (Top) Reaction of L-iditol dehydrogenase, also known as sorbitol dehydrogenase. This is the putative call for FtxJ using BLAST. (Bottom) Reaction catalyzed by FtxJ. The question mark denotes the lack of understanding of the metal-dependence of FtxJ.

With this in mind, I incubated DHPPA with NAD, FtxJ and a mixture of different metals, resulting in conversion of DHPPA to a new compound based on ^{31}P NMR spectroscopy (Fig 4.11). I have putatively assigned this compound 2OPPA based on the chemical shifts of the protons associated with the phosphorus atom (Fig 4.12 and Fig 4.13). Accordingly, only two sets of protons associated with the phosphorus atom and the splitting pattern of these protons is simplified (Fig 4.12) in comparison to DHPPA (Fig 4.9), suggesting a lack of ^{3}J-splitting of the protons with one another. In combination with the lack of a correlation of a proton displaying a chemical shift indicative of an aldehyde (~9-10 ppm) (28), I conclude that the hydroxyl group on the β-carbon is oxidize to the ketone, forming 2OPPA.

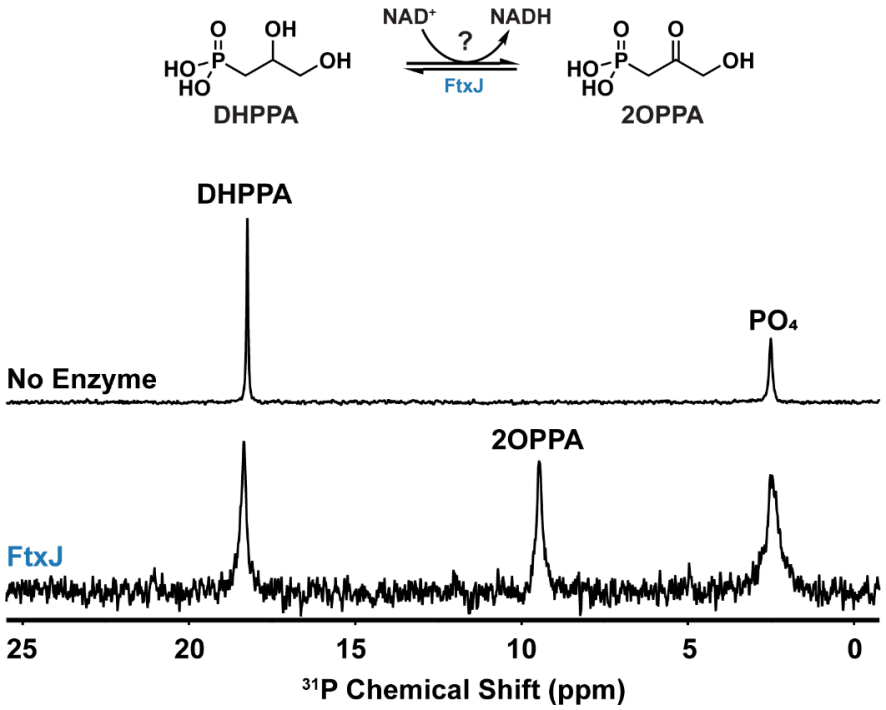


Figure 4.11. (Top) Oxidation of DHPPA to 2OPPA catalyzed by FtxJ. (Bottom) ^{31}P NMR spectroscopy of the FtxJ reactions to monitor DHPPA consumption. The reactions were set up using the following conditions in PBS buffer, pH 7.4: 15 mM DHPPA, 250 μM NAD^+, 10 μM FtxJ, 1x metal mix (50 μM $FeCl_3$, 20 μM $CaCl_2$, 10 μM $MnCl_2$, 10 μM $ZnSO_4$, 2 μM $CoCl_2$, 2 μM $NiCl_2$, 2 μM $CuCl_2$, 2 μM Na_2MoO_4, 2 μM Na_2SeO_3, and 2 μM H_3BO_3). Lactate dehydrogenase was added for use as a cofactor-regeneration system at a final concentration of 4 μg/mL with 10 mM of pyruvate.

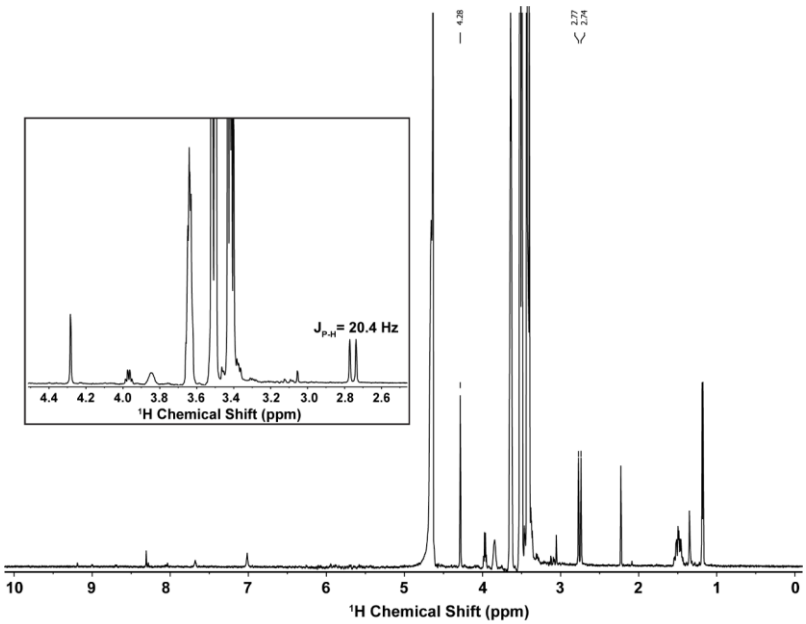


Figure 4.12. ^{1}H spectrum of the FtxJ reaction product from Fig 4.11. Peaks that have been labeled correspond to the product of the reaction. The J-coupling of the doublet is the phosphorus-proton coupling and is consistent with a phosphonate (J_{P-H} = 20.4 Hz).

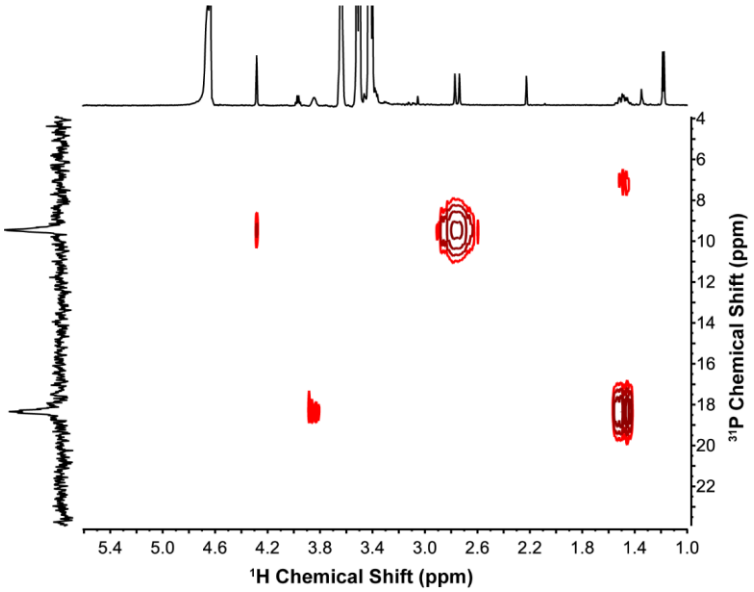


Figure 4.13. ^{1}H-^{31}P HMBC spectrum of the FtxJ reaction zoomed in to show the phosphonates of interest. This figure supports Fig 4.11.

The purified recombinant FtxJ protein has a pink hue, suggesting that the metal used by the protein may not be zinc. I initially tested cobalt, but have found that this is not the metal required for FtxJ function (Fig 4.14), but upon addition of manganese, we see robust conversion of DHPPA to 2OPPA (Fig 4.14).

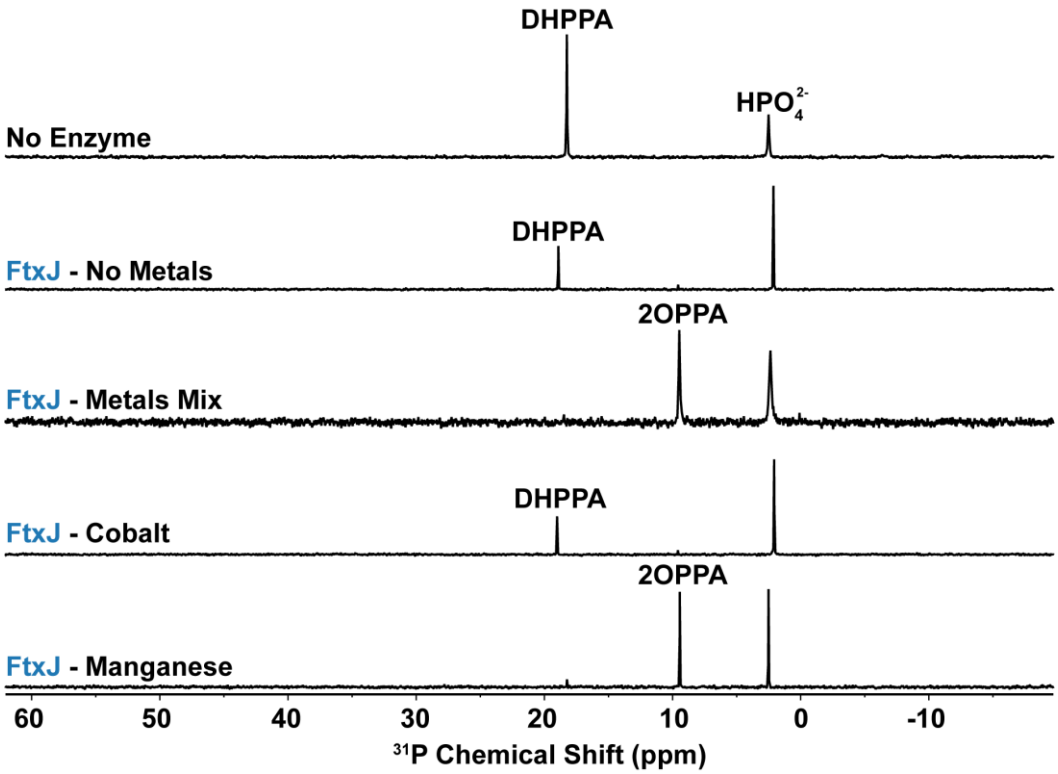


Figure 4.14. ^{31}P NMR spectrum of FtxJ reactions in the absence of metals and with cobalt as the metal. Reaction conditions were set up using the following conditions in PBS buffer, pH 7.4: 5-10 mM DHPPA, 250 μM NAD^+, 10 μM FtxJ, lactate dehydrogenase (4 μg/mL), 10 mM pyruvate, and 10 μM $CoCl_2$ or $MnCl_2$.

4.2.6 FtxE is likely responsible for phosphonothrixin production[5]

TPP-dependent enzymes catalyze a number of different reaction including, but not limited to, oxidative decarboxylation of α-keto acids, carboligation, and C-C bond cleavage (7, 11, 30). If 2OPPA can be condensed with pyruvate utilizing one of the two

[5] *I would like to thank Nektaria Petronikolou for the construction of the FtxE expression plasmid.*

TPP-dependent enzymes encoded within the gene cluster (FtxE or FtxF), phosphonothrixin can be biosynthesized in this one enzymatic step. Therefore, I incubated 6x-HIS-tagged FtxE with the reaction product generated using FtxJ (Fig 4.11) in the presence of TPP and pyruvate. After 24 hours, I observed the formation of a new compound within the reaction mixture (Fig 4.15), which I assigned to be phosphonothrixin based on the chemical shifts of the protons correlated with the phosphorus atom (Fig 4.16). This suggests that the two-component transketolase, FtxE, is involved in the biosynthesis of phosphonothrixin.

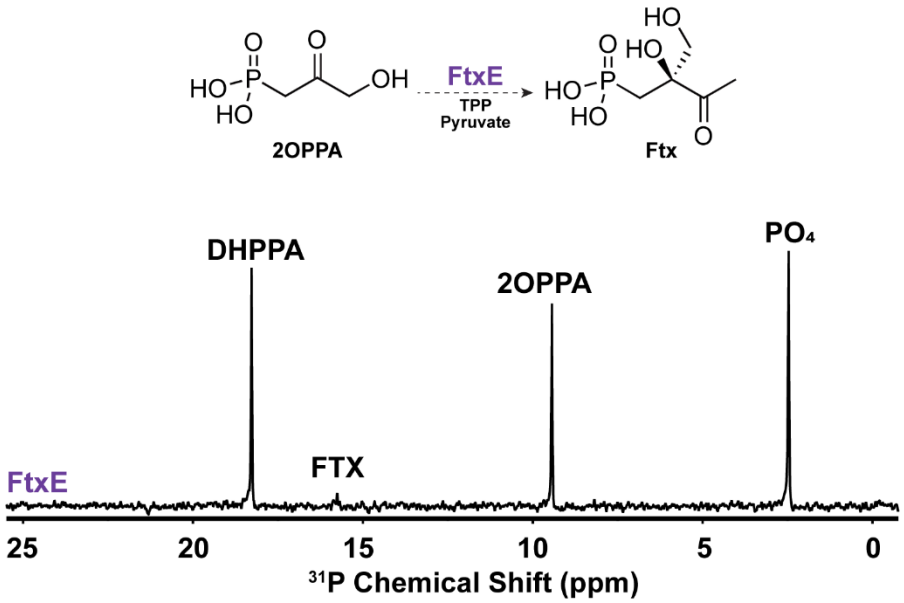


Figure 4.15. (Top) Predicted reaction to form phosphonothrixin (Ftx). (Bottom) ^{31}P NMR spectroscopy of the FtxE reaction to monitor 2OPPA consumption. The reactions were set up using the following conditions in PBS buffer, pH 7.4: reaction from Fig 4.12 pre-treated with Chelex 100 resin to remove excess metals, 10 mM pyruvate, 500 µM TPP, and 10 µM FtxE.

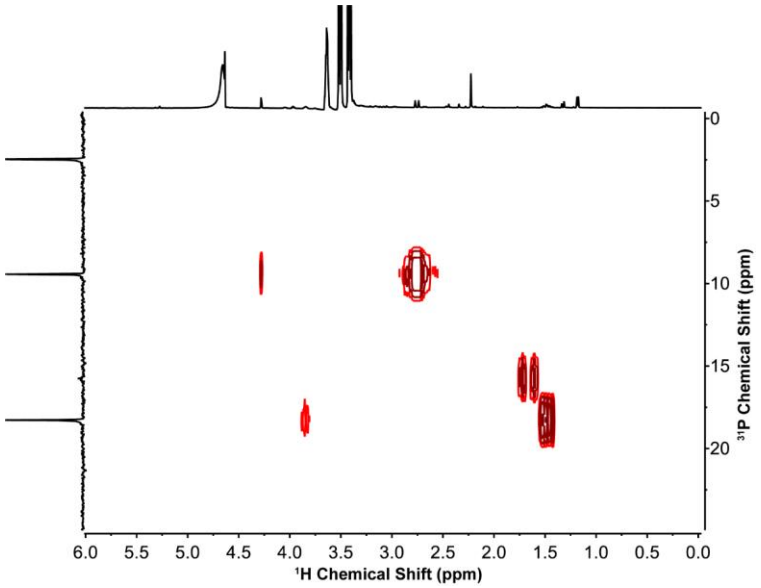


Figure 4.16. ^{1}H-^{31}P HMBC spectrum of the FtxE reaction zoomed in to show the phosphonates of interest. This figure goes in conjunction with Fig 4.15.

4.3 SUMMARY AND OUTLOOK

We still do not understand the full biosynthetic pathway to phosphonothrixin, though the experiments described here suggest that phosphonothrixin is produced using a pathway similar to the valinophos biosynthetic pathway (Fig 4.17) (16). Consistent with this hypothesis, oxidation of DHPPA to 2OPPA is readily catalyzed using FtxJ *in vitro*, suggesting this is the native function of this enzyme *in vivo*. Although our limited NMR data highly suggest the product form by FtxJ is 2OPPA (Fig 4.12 and Fig 4.13), the product of the reaction needs to be isolated and fully characterized structurally and using mass spectrometry.

Initial characterizations of the two-component transketolase homolog, FtxE, suggest that this protein forms phosphonothrixin (Fig 4.15), although the reaction conditions need optimization. There is a possibility that two are missing a cofactor for enzyme function of the transketolases as evident by the weak production of phosphonothrixin (Fig 4.15). Additionally, there may be an inhibitory component left in

the reaction mixture. We used the reaction mixture from Fig 4.11 as the substrate for FtxE. It is possible that any remaining metals or excess lactate formed by lactate dehydrogenase may be inhibitory for FtxE. Therefore, this reaction needs to be performed with pure substrate.

Initial data collected for FtxF suggests that this enzyme does not catalyze the condensation of phosphonopyruvate with pyruvate (Fig 4.2). It is possible that the incorrect keto-acid is being utilized in conjunction with phosphonopyruvate to produce the first intermediate proposed by Lin and colleagues, although the mild activity of FtxF in the presence of pyruvate (Fig 4.4) suggests that this enzyme likely uses pyruvate as a substrate. Thus, the lack of activity with phosphonopyruvate (Fig 4.2) suggests that another phosphonic acid may be the physiological substrate. Alternatively, the enzyme may be responsible for self-resistance.

Many natural product biosynthetic gene clusters encode a resistance gene within the cluster to protect the producing organism from the final product or toxic intermediates (15). An organism may do this in one of two ways. The first involves the modification of toxic intermediates or products so that they are no longer inhibitory. The second involves encoding a copy of the target protein within the gene cluster that is resistant to the final product (15). One hypothesis that cannot be confirmed or refuted with our data is that FtxF may be the resistance protein produced within the phosphonothrixin gene cluster. The close molecular mimicry of phosphonothrixin and acetolactate may allow phosphonothrixin to target acetolactate synthase (Fig 4.5). We do observe canonical acetolactate synthase activity using FtxF (Fig 4.4) and FtxF shares a common ancestor

with bona fide ALS proteins (Fig 4.6); therefore, this enzyme may function as an ALS within the producing organism during the biosynthesis of phosphonothrixin.

The data presented in this chapter, allow us to propose a biosynthetic pathway for this compound (Fig 4.17). We hypothesize that phosphonothrixin and valinophos share common first biosynthetic steps to form DHPPA (16). DHPPA will be oxidized using the metal-dependent dehydrogenase, FtxJ, to yield 2OPPA (Fig 4.11). FtxE would then catalyze the oxidative decarboxylation of pyruvate and the condensation with 2OPPA to yield phosphonothrixin (Fig 4.15) (Fig 4.17).

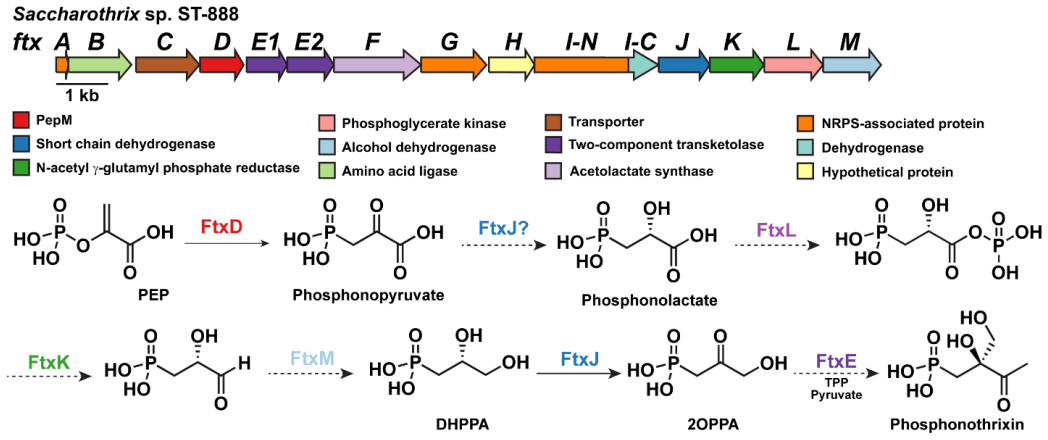


Figure 4.17. (Top) *Saccharothrix* sp. ST-888 phosphonothrixin gene cluster (21). (Bottom) Putative phosphonothrixin biosynthetic pathway. Dashed arrows represent hypothesized reactions based on the similarity to valinophos biosynthesis (16).

Discovery and characterization of the biosynthesis of phosphonothrixin is an important endeavor because of the potential application of phosphonothrixin as an herbicide. Due to the rising need of new herbicidal compounds, if the biosynthesis of phosphonothrixin can elucidated, we will understand the steps in need of optimization for large-scale production of this compound. Additionally, if the target of this compound can be identified and the resistance gene is encoded within the biosynthetic gene cluster,

crops resembling RoundUp Ready® and LibertyLink® crops may be developed, thus helping combat the rise in herbicide resistant weeds that cause massive loss in crops yields yearly.

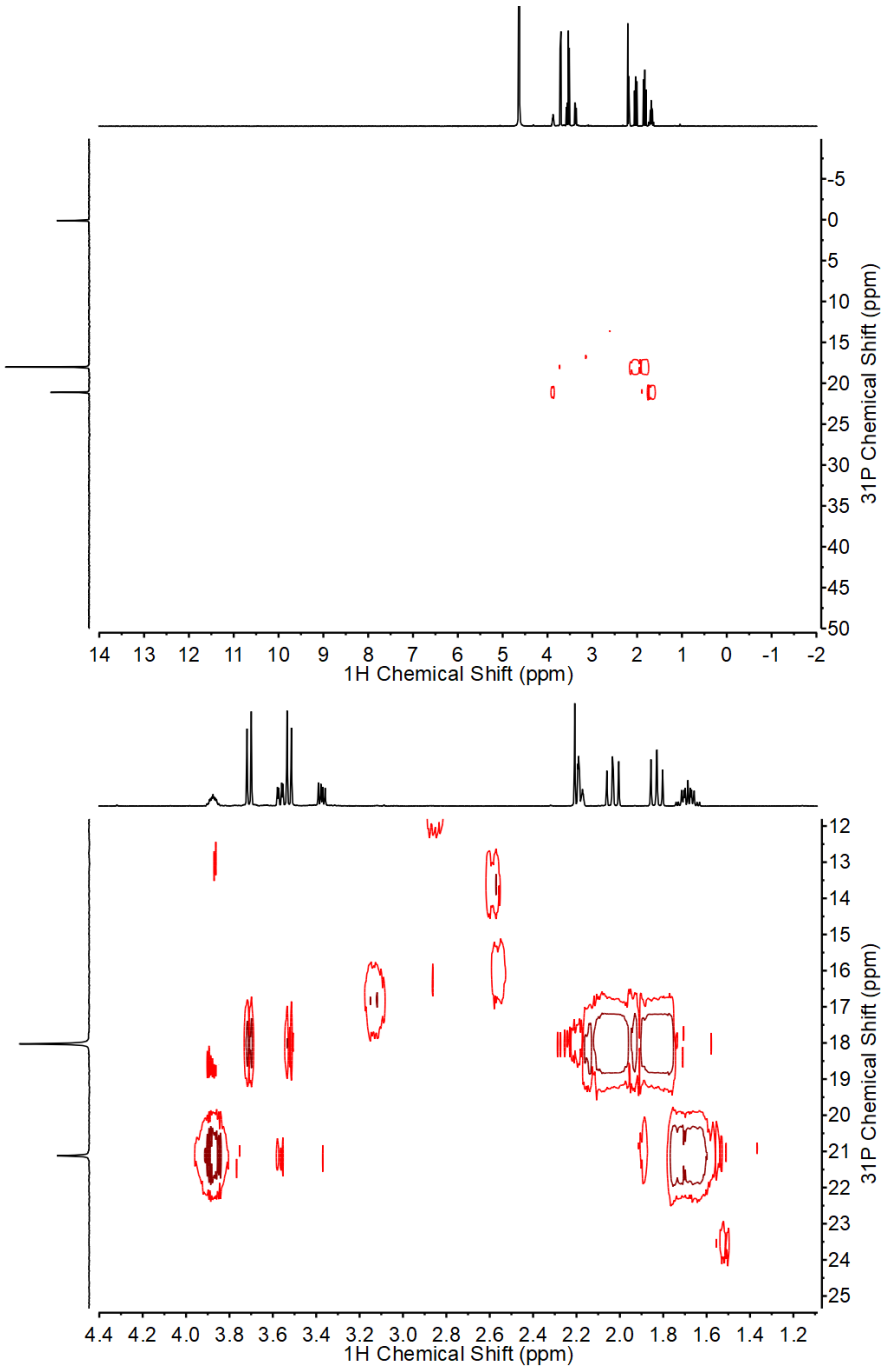


Figure 4.18. (Top) Full ^{1}H-^{31}P HMBC spectrum collected for the co-purified mixture of phosphonothrixin and DHPPA. (Bottom) ^{1}H-^{31}P HMBC spectrum zoomed in to focus on the phosphonates of interest.

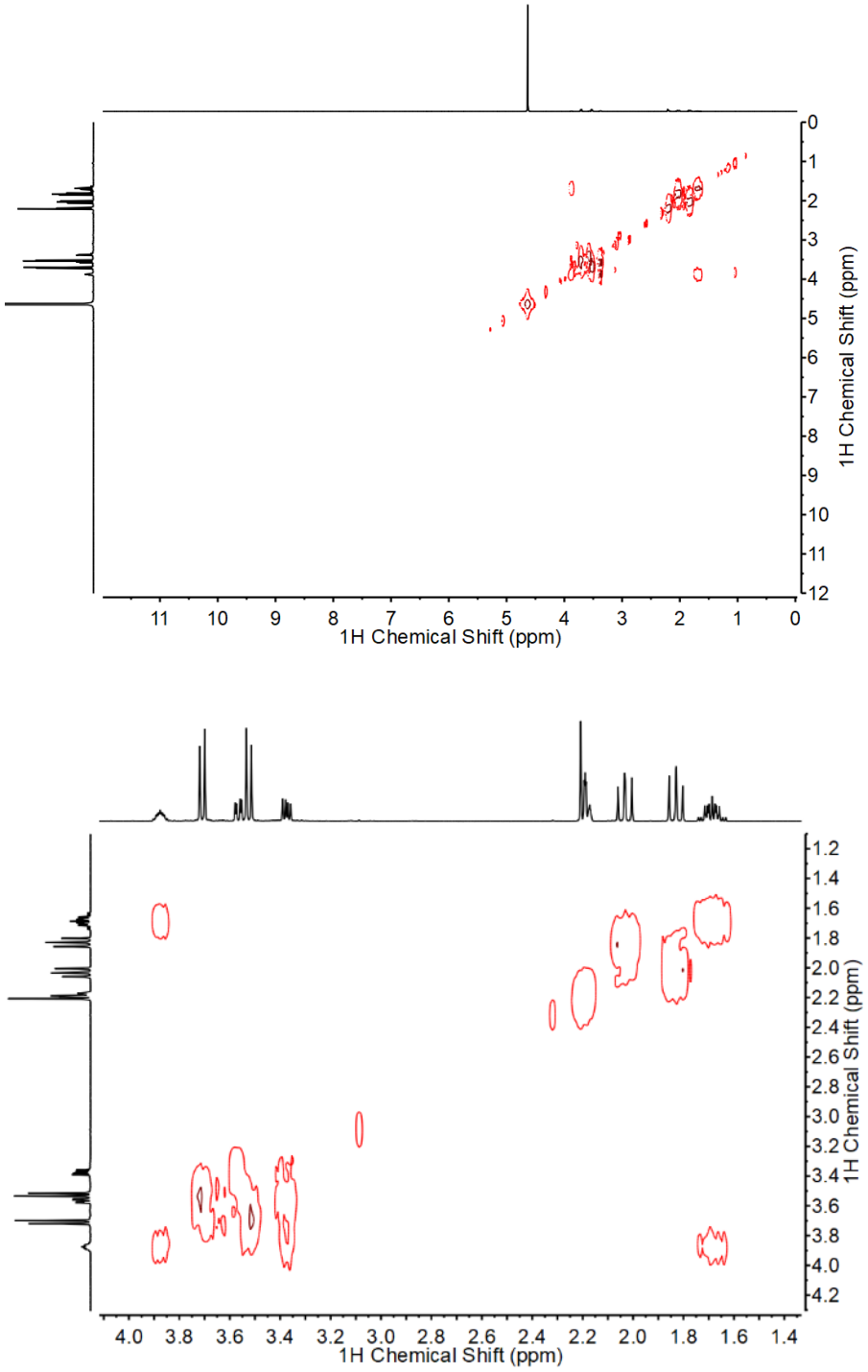


Figure 4.19. (Top) Full ^{1}H-^{1}H COSY spectrum collected for the co-purified mixture of phosphonothrixin and DHPPA. (Bottom) ^{1}H-^{1}H COSY spectrum zoomed in to focus on the phosphonates of interest.

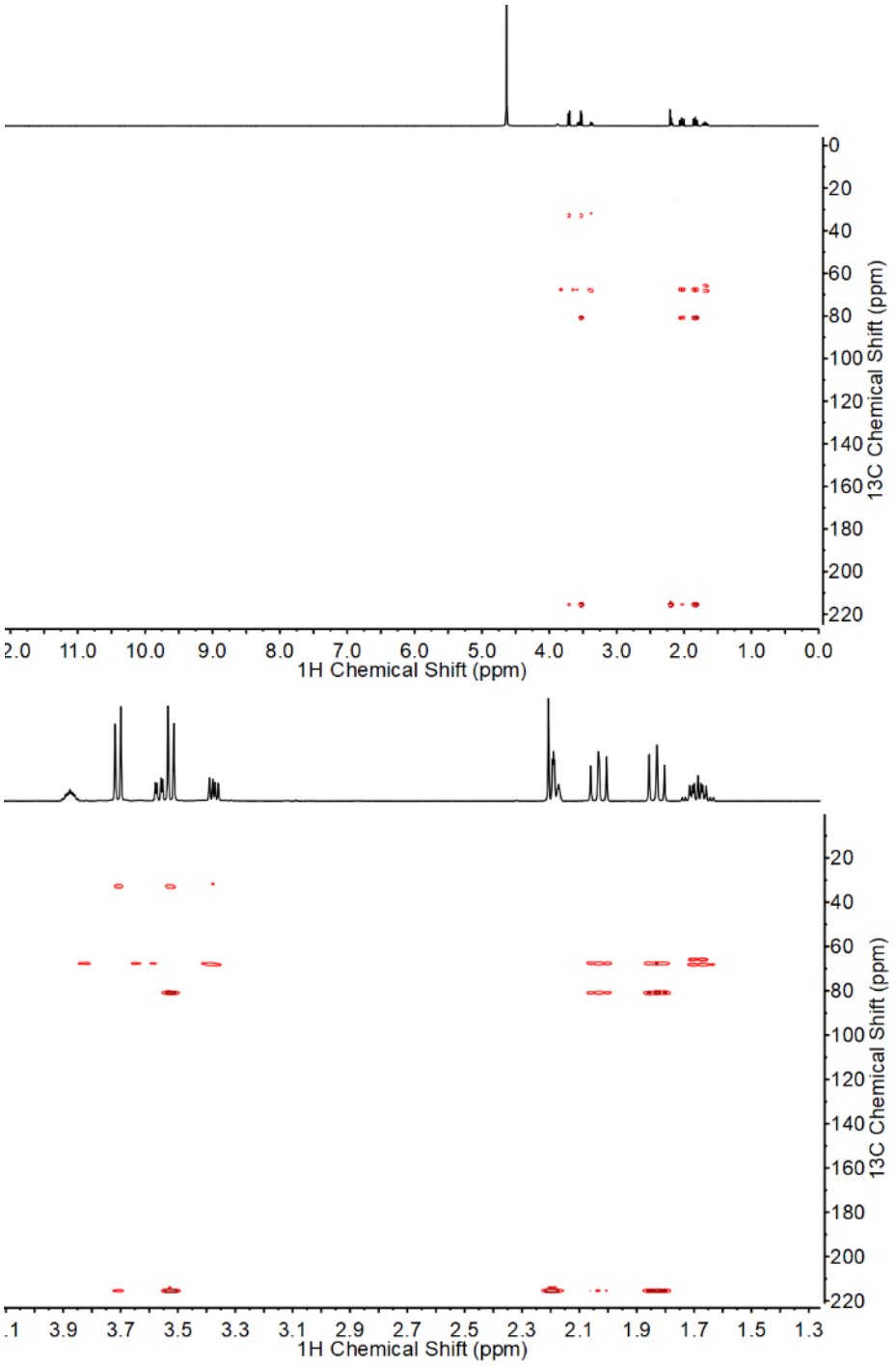


Figure 4.20. (Top) Full ^{1}H-^{13}C HMBC spectrum collected for the co-purified mixture of phosphonothrixin and DHPPA. (Bottom) ^{1}H-^{13}C HMBC spectrum zoomed in to focus on the phosphonates of interest.

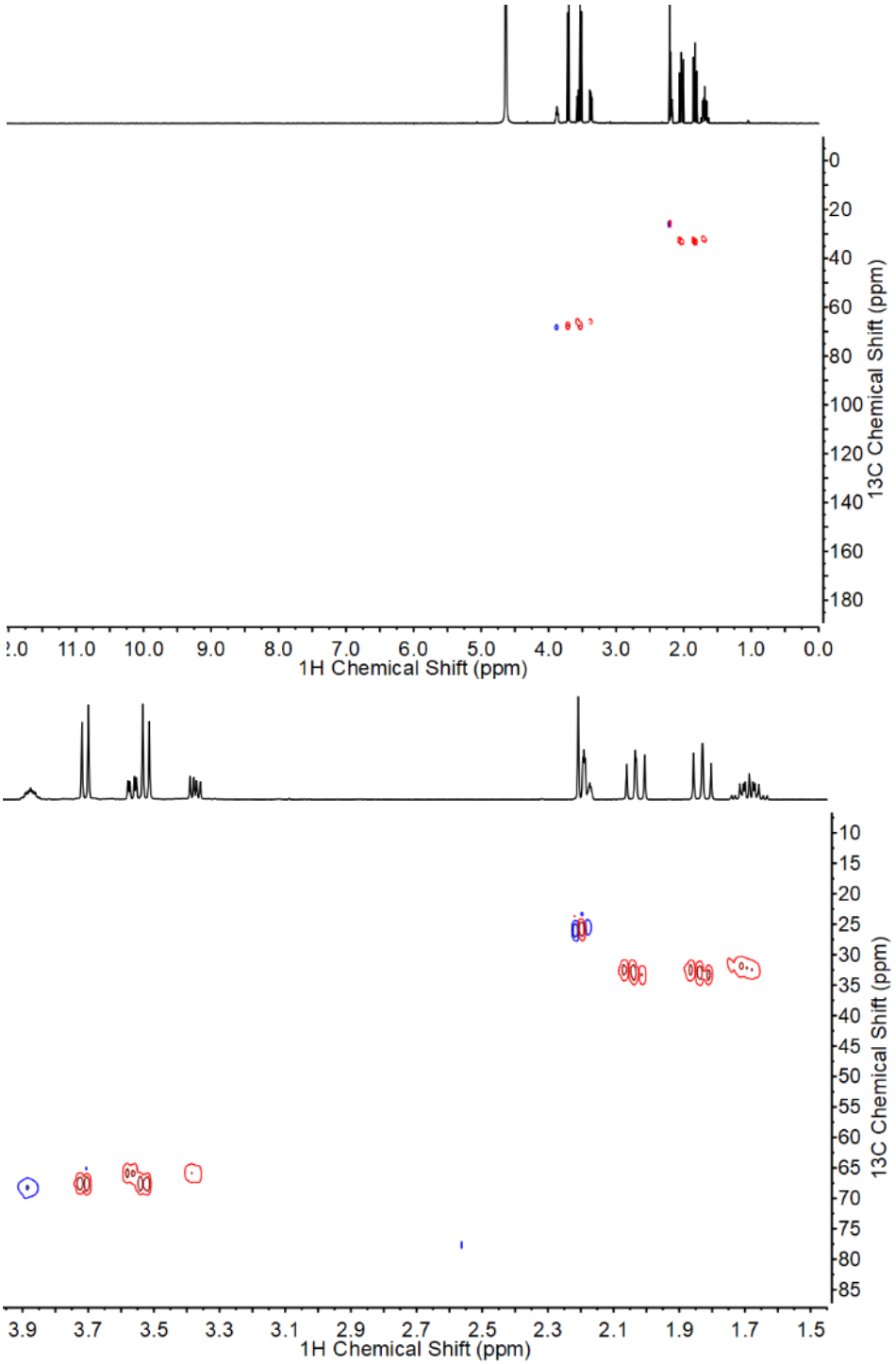


Figure 4.21. (Top) Full ^{1}H-^{13}C HSQC-AD spectrum collected for the co-purified mixture of phosphonothrixin and DHPPA. (Bottom) ^{1}H-^{13}C HSQC-AD spectrum zoomed in to focus on the phosphonates of interest. The phasing is set for red to correspond to carbons containing two protons, while blue represents one or three protons on one carbon atom.

4.4 MATERIALS AND METHODS

4.4.1 General experimental procedures

An Agilent Technologies 600 MHz NMR spectrometer was used to collect all nuclear magnetic resonance spectra (NMR) using D_2O as a lock solvent. Fisher Scientific or Sigma-Aldrich were used for the purchase of all chemicals used in this chapter.

4.4.2 Cloning and protein purification

The FtxF gene from *Saccharopolyspora spinosa* NRRL 18395 was amplified by PCR using primer pair MG036 (5'-CAGCGGCCTGGTGCCGCGCGGCAGCCATATGcgggtagctgactacatcgtcg-3') and MG037 (5'-GATCTCAGTGGTGGTGGTGGTGGTGCTCGAGctaacagcgggtggggatcgtgact-3') from genomic DNA isolated using an UltraClean® Microbial DNA Isolation Kit (MO-BIO Laboratories, San Diego, CA). The plasmid, pET28B, was amplified using primers pET28-NdeI-R (5'-CATATGGCTGCCGCGCGGCACCAGGCCGCTG-3') and pET28B-XhoI-F (5'-CTCGAGCACCACCACCACCACCACTGAGATCCGG-3'). Uppercase letters represent sequence homologous to the plasmid pET28B. The underlined sequence is an inserted restriction enzyme site, and the lowercase letters represent sequence homology to the gene to be amplified. The genes for FtxF were cloned into pET28B using Gibson Assembly® (New England Biolabs, Ipswich, MA) to yield pMNG004. Strains MMG423 was used for expression of RhiH and MMG422 was used for the expression of Bf-Ppd.

All of the proteins were isolated using a similar protocol. Briefly, a 5 mL overnight culture of BL21 (DE3) cells containing the plasmid of interest was used to inoculate a 1 L flask of LB + kanamycin (50 μg/ mL). This culture was grown at 37°C until the optical density (OD_{600}) is between 0.6-0.9. Protein production was induced with addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 μM. Cultures were shaken at 18°C for 16-18 hours. Cells were collected via centrifugation and resuspended in lysis buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), and 10 mM imidazole, pH 8.0). The cell suspensions were mixed with 1 mg/mL lysozyme and 100 units DNase I and incubated on a mutator for 30 minutes at 4°C. To lyse the cells, each cell suspension was passed through a French Press cell disruptor (Thermo Electron Corporation, Waltham, MA) two times at 1000 psi each to lyse the cells. The lysed cell suspension was separated via centrifugation for 45 minutes at 10,000 x g at 4°C. The soluble liquid portion was incubated with 5 mL per liter of culture of loose Ni-NTA resin (Thermo Fisher Scientific) for 30 minutes at 4°C. The resin was washed with approximately 10 column volumes of wash buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), and 20 mM imidazole, pH 8.0) until no protein was observed to be eluting from the resin by monitoring absorbance at A_{260} nm. Proteins were eluted using 15 mL of elution buffer (300 mM sodium chloride, 50 mM sodium phosphate (dibasic), and 250 mM imidazole, pH 8.0). The flow through (containing pure protein) was concentrated using a 10 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA). A PD-10 desalting column (GE Healthcare Life Sciences, Marlborough, MA) was used to exchange the purified proteins into protein storage buffer (200 mM KCl, 50 mM HEPES, and 10% glycerol, pH 7.5).

4.4.3 *In vitro* activity assays using FtxF

To test whether FtxF could use phosphonopyruvate as a substrate, we incubated 6xHis-tagged FtxF from Saccharothrix sp. ST-888 with 10 mM phosphoenolpyruvate, 10 mM sodium pyruvate, 1 mM thiamine pyrophosphate, 50 μg FtxF, and 50 μg RhiH in 50 mM HEPES buffer pH 7.5 containing 10 mM magnesium chloride. Reactions using flavin used 100 μM of either FAD or FMN. The control reaction uses the same components as above, with the exception of phosphonopyruvate decarboxylase from *Bacteroides fragilis* being used at a concentration of 50 μg in place of FtxF. To test whether FtxF could use DHPPA as a substrate, we incubated 6xHis-tagged FtxF from *Saccharothrix* sp. ST-888 with 1 mM DHPPA, 10 mM sodium pyruvate, 1 mM thiamine pyrophosphate, 50 μg FtxF, and 50 μg RhiH in 50 mM HEPES buffer containing 10 mM magnesium chloride. Reactions were incubated at 30°C for 12 hours. Enzymes were removed from the reactions using a 30 kDa molecular weight Amicon spin filter, and D_2O was added at a final concentration of 20% for use as a lock solvent for NMR. Product formation was determined using ^{31}P NMR spectroscopy.

To test whether FtxF can function as a canonical acetolactate synthase, we incubated 6xHis-tagged FtxF from *Saccharothrix* sp. ST-888 or *S. spinosa* NRRL 18395 with 20 mM sodium pyruvate and 237 μM TPP in 100 mM potassium phosphate buffer, pH 7.5 containing 10 mM magnesium chloride. FtxF-ST-888 was used at a concentration of 0.1 μM and 1.8 μM, while FtxF-18395 was used at a concentration of 0.1 μM and 3 μM. Control reactions were performed using 6xHis-tagged AlsS from *Bacillus subtilis* at concentrations of 0.1 μM and 2.5 μM (courtesy of Michael Carter). Reactions were incubated at room temperature for 13 hours. Reactions were quenched using 4 μL of 50%

sulfuric acid for 25 minutes. 20 μL of the product was added to 280 μL of acetoin detection reagent containing 1% α-napthol, 0.1% creatine, and 1 M sodium hydroxide. These reactions were incubated for 25 minutes at room temperature. Absorbance at 530 nm was recorded. A standard curve using 6.25 μM to 200 μM acetoin was made using the same protocol. Reactions were repeated in duplicate.

4.4.4 FtxF phylogeny

To construct the maximum-likelihood phylogeny of FtxF that includes only proteins with characterized functions, a BLAST search was performed against the SwissProt/UniProt database using *Saccharothrix* sp. ST-888 FtxF as the search query (accession number WP_045300935.1) (1). 168 entries were retrieved. To construct the maximum-likelihood phylogeny of the closest homologs to FtxF, a BLAST search was performed against the nonredundant NCBI database to retrieve 250 sequences (1). Sequences were aligned using MAFFT standard settings (17). A maximum-likelihood phylogenetic tree was constructed of each dataset using FastTree (29).

4.4.5 Isolation of phosphonothrixin and DHPPA

This project was done in collaboration with Dr. Ryuichi Hirota from Hiroshima University. *Saccharothrix* sp. ST-888 was grown at 30°C for ~7 days in 20 liters of both GUBC and R2AS broth, yielding 40 liters of spent media (16). Spent media was dried under reduced pressure and resuspended in two liters of water. Methanol was added to a final concentration of 75%. Methanol insoluble material was discarded. The 75% methanol soluble material was dried under reduced pressure and extracted using 100% methanol. Methanol soluble material was dried under reduced pressure rotary evaporation

and suspended in 200 mL deionized water. The water-soluble material was passed over activated charcoal and material not bound to the activated charcoal was precipitated using calcium acetate. Soluble material was separated on a 134.5 cm x 6.4 cm, 4.8 liter resin bed Sephadex® G-25 resin size exclusion column (GE Healthcare Life Sciences). Fractions containing phosphonic acids were further purified using hydrophilic lipophilic balance (HLB) exchange. The fractions containing phosphonic acids were concentrated under reduced pressure and resuspended in deionized water. Weak cation exchange was performed. The fraction containing phosphonic acids was concentrated and suspended in deionized water. This was separated on a 175 cm x 4 cm, 2.4 liter resin bed Sephadex® LH-20 resin size exclusion column (GE Healthcare Life Sciences). Fractions containing phosphonic acids were concentrated and further purified using IMAC-Fe. I received this phosphonate–enriched, partially-purified mixture from Ryuichi Hirota and carried out the co-purification of phosphonothrixin and DHPPA.

From this point forward, I performed the purification of the phosphonic acids. This material was resuspended in 4 mL deionized water and separated on 90 mL Sephadex ® G-10 gel filtration column (GE Healthcare Life Sciences), 160 cm x 2 cm. The column was run at a flow rate of 100 μL per minute at 4°C using deionized water as the solvent. After collection of 40 mL, fractionation of the material was done collecting 0.5 mL fractions. Phosphonic acids eluted between fractions 21-44. Fractions 31-44 were pooled and concentrated under reduced pressure. Sample was resuspended in 5 mM sulfuric acid. One eighth of the sample material (100 μL) was injected onto a 300 mm x 7.8 mm Aminex® HPX-87H ion exclusion chromatography column (Bio-Rad). The column was run under isocratic conditions in 5 mM sulfuric acid in deionized water at a

flow rate of 0.6 mL / min, collecting fractions every minute (0.6 mL fractions). Fractions were neutralized using 1M sodium hydroxide and dried under reduced pressure. Fraction ten contained the purified mixture of DHPPA and phosphonothrixin. Fraction ten was desalted by passing the material over a 160 cm x 2 cm, 90 mL resin bed, Sephadex® G-10 column (GE Healthcare Life Sciences) using the same conditions described earlier. Fractions containing phosphonic acids were pooled and concentrated under reduced pressure.

Structure elucidation of the compounds within fraction ten was performed using ^{1}H NMR (600 MHz), ^{31}P NMR (242 MHz), and ^{13}C NMR (150 MHz). All spectra were collected on a 600 MHz Agilent NMR spectrometer. Experiments performed include one-dimensional ^{1}H and ^{31}P NMR spectroscopy and two-dimensional ^{1}H-^{31}P HMBC, ^{1}H-^{13}C HMBC, ^{1}H-^{13}C HSQCAD, and ^{1}H-^{1}H COSY. Carbon chemical shifts were inferred from the ^{1}H-^{13}C HMBC and ^{1}H-^{13}C HSQCAD experiments. These data are consistent with the chemical shifts reported for DHPPA (16) and phosphonothrixin (10).

4.4.6 FtxJ assays

Pure N-terminally 6xHis-tagged FtxJ (concentration 10 μM) was incubated with 15 mM DHPPA, 250 μM NAD^{+}, and metals (50 μM $FeCl_3$, 20 μM $CaCl_2$, 10 μM $MnCl_2$, 10 μM $ZnSO_4$, 2 μM $CoCl_2$, 2 μM $NiCl_2$, 2 μM $CuCl_2$, 2 μM Na_2MoO_4, 2 μM Na_2SeO_3, and 2 μM H_3BO_3) in PBS buffer, pH 7.4 in addition to lactate dehydrogenase (concentration 4 μg/mL) and 10 mM pyruvate. The reactions testing for product formation in the presence of cobalt used the same conditions as above, with the metal mix replaced with 10 μM $CoCl_2$. The reactions were incubated at room temperature for 18

hours and stopped by removing the enzyme using a 10 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA). A small amount (~100 μL volume) of Chelex® 100 sodium form resin (Sigma-Aldrich) was added to the reaction to remove excess metals. D_2O was added to a final concentration of 20% for use as a lock solvent. Reaction products were identified via ^{1}H NMR (600 MHz) spectroscopy and ^{31}P NMR (242 MHz) spectroscopy.

4.4.7 FtxE assays

Pure 6xHis-tagged FtxE1 and untagged FtxE2 (concentration 10 μM) were incubated with the reaction product of FtxJ with 10 mM pyruvate and 500 μM TPP in PBS buffer, pH 7.4. The reactions were stopped after a 24-hour room temperature incubation using a 10 kDa molecular weight Amicon® Ultra spin filter (EMD Millipore, Burlington, MA). D_2O was added as a lock solvent to a final concentration of 20%. Reaction products were identified via ^{1}H NMR (600 MHz) spectroscopy and ^{31}P NMR (242 MHz) spectroscopy.

4.5 REFERENCES

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215(3):403–10.

2. Bayer E, Gugel KH, Hägele K, Hagenmaier H, Jessipow S, et al. 1972. Stoffwechselprodukte von Mikroorganismen. 98. Mitteilung. Phosphinothricin und Phosphinothricyl-Alanyl-Alanin. *Helv. Chim. Acta*. 55(1):224–239.

3. Benbrook CM. 2016. Trends in glyphosate herbicide use in the United States and globally. *Environ. Sci. Eur.* 28(3).

4. Borisova SA, Circello BT, Zhang JK, van der Donk WA, Metcalf WW. 2010. Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by Bacillus subtilis ATCC6633. *Chem. Biol.* 17(1):28.

5. Bowman E, McQueney M, Barry RJ, Dunaway-Mariano D. 1988. Catalysis and thermodynamics of the phosphoenolpyruvate/phosphonopyruvate rearrangement. Entry into the phosphonate class of naturally occurring organophosphorus compounds. *J. Am. Chem. Soc.* 110(16):5575–76.

6. Chipman DM, Duggleby RG, Tittmann K. 2005. Mechanisms of acetohydroxyacid synthases. *Curr. Opin. Chem. Biol.* 9(5):475–81.

7. Costelloe SJ, Ward JM, Dalby PA. 2008. Evolutionary Analysis of the TPP-Dependent Enzyme Family. *J. Mol. Evol.* 66(1):36–49.

8. Dill GM, Sammons RD, Feng PCC, Kohn F, Kretzmer K, et al. 2010. Glyphosate: Discovery, Development, Applications, and Properties. In *Glyphosate Resistance in Crops and Weeds*, pp. 1–33.

9. Eliot AC, Griffin BM, Thomas PM, Johannes TW, Kelleher NL, et al. 2008. Biosynthesis of the Potent Antimalarial Compound FR900098. *Chem. Biol.* 15(8):765–70.

10. Fields SC. 1998. Total synthesis of (±)-Phosphonothrixin. *Tetrahedron Lett.* 39(37):6621–24.

11. Frank R. W, Leeper FJ, Luisi BF. 2007. Structure, mechanism and catalytic duality of thiamine-dependent enzymes. *Cell. Mol. Life Sci.* 64(7–8):892–905.

12. Funke T, Han H, Healy-Fried ML, Fischer M, Schönbrunn E. 2006. Molecular basis for the herbicide resistance of Roundup Ready crops. *Proc. Natl. Acad. Sci.* 103(35):13010–15.

13. Gill HS, Eisenberg D. 2001. The Crystal Structure of Phosphinothricin in the Active Site of Glutamine Synthetase Illuminates the Mechanism of Enzymatic Inhibition. *Biochemistry*. 40(7):1903–12.

14. Hoerlein G. 1994. Glufosinate (phosphinothricin), a natural amino acid with unexpected herbicidal properties. *Rev. Environ. Contam. Toxicol.* 138:73–145.

15. Hopwood DA. 2007. How do antibiotic-producing bacteria ensure their self-resistance before antibiotic biosynthesis incapacitates them? *Mol. Microbiol.* 63(4):937–40.

16. Ju K-S, Gao J, Doroghazi JR, Wang K-KA, Thibodeaux CJ, et al. 2015. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc. Natl. Acad. Sci.* 112(39):12175–80.

17. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30(4):772–80.

18. Kimura T, Nakamura K, Takahashi E. 1995. Phosphonothrixin, a novel herbicidal antibiotic produced by Saccharothrix sp. ST-888. II. Structure determination. *J. Antibiot. (Tokyo)*. 48(10):1130–33.

19. Lange BM, Ketchum REB, Croteau RB. 2001. Isoprenoid Biosynthesis. Metabolite Profiling of Peppermint Oil Gland Secretory Cells and Application to Herbicide Target Analysis. *Plant Physiol.* 127(1):305.

20. Liebman Matt, Baraibar Bàrbara, Buckley Yvonne, Childs Dylan, Christensen Svend, et al. 2016. Ecologically sustainable weed management: How do we get from proof- of- concept to adoption? *Ecol. Appl.* 26(5):1352–69.

21. Lin J, Nishiyama M, Kuzuyama T. 2014. Identification of the biosynthetic gene cluster for the herbicide phosphonothrixin in *Saccharothrix* sp. ST-888. *J. Antibiot.* 68:357–359.

22. Liu Y, Li Y, Wang X. 2017. Molecular evolution of acetohydroxyacid synthase in bacteria. *MicrobiologyOpen*. 6(6):e00524.

23. Lutz KA, Knapp JE, Maliga P. 2001. Expression of bar in the Plastid Genome Confers Herbicide Resistance. *Plant Physiol.* 125(4):1585–90.

24. NCBI. *Basic Local Alignment Search Tool*. http://blast.ncbi.nlm.nih.gov/Blast.cgi.

25. Ng K, Ye R, Wu XC, Wong SL. 1992. Sorbitol dehydrogenase from Bacillus subtilis. Purification, characterization, and gene cloning. *J. Biol. Chem.* 267(35):24989–94.

26. Ogawa H, Tsuruoka T, Inouye S, Niida T. 1983. Studies on a new antibiotic SF-1293. *Sci. Rep.* 13:42–48.

27. Pimentel D, McNair S, Janecka J, Wightman J, Simmonds C, et al. 2001. Economic and environmental threats of alien plant, animal, and microbe invasions. *Agric. Ecosyst. Environ.* 84(1):1–20.

28. Pretsch E, Buhlmann P, Badertscher M. 2009. *Structure Determination of Organic Compounds*. 433 pp. 4th ed.

29. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One*. 5(3):e9490.

30. Schellenberger A. 1998. Sixty years of thiamin diphosphate biochemistry. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* 1385(2):177–86.

31. Schwartz D, Berger S, Heinzelmann E, Muschko K, Welzel K, Wohlleben W. 2004. Biosynthetic Gene Cluster of the Herbicide Phosphinothricin Tripeptide from Streptomyces viridochromogenes Tü494. *Appl. Environ. Microbiol.* 70(12):7093–7102.

32. Sikorski JA, Gruys KJ. 1997. Understanding Glyphosate's Molecular Mode of Action with EPSP Synthase: Evidence Favoring an Allosteric Inhibitor Model. *Acc. Chem. Res.* 30(1):2–8.

33. Speckman RA, Collins EB. 1982. Specificity of the Westerfeld adaptation of the Voges-Proskauer test. *Appl. Environ. Microbiol.* 44(1):40–43.

34. Takahashi E, Kimura T, Nakamura K, Arahira M, Iida M. 1995. Phosphonothrixin, a novel herbicidal antibiotic produced by Saccharothrix sp. ST-888. I. Taxonomy, fermentation, isolation and biological properties. *J. Antibiot. (Tokyo)*. 48(10):1124–29.

35. Thompson CJ, Movva NR, Tizard R, Crameri R, Davies JE, et al. 1987. Characterization of the herbicide-resistance gene bar from Streptomyces hygroscopicus. *EMBO J.* 6(9):2519–23.

36. Zhang G, Dai J, Lu Z, Dunaway-Mariano D. 2003. The Phosphonopyruvate Decarboxylase from Bacteroides fragilis. *J. Biol. Chem.* 278(42):41302–8.