MINING THE *PSEUDOMONAS VIRULENCE FACTOR* PATHWAY FOR NOVEL SMALL MOLECULES

Ashley Marie Kretsch

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill
in partial fulfilment of the requirements for the degree of Doctor of Philosophy
in the Department of Chemistry.

Chapel Hill
2019

Approved by:

Bo Li

Matthew Redinbo

Marcey Waters

Jeffery Dangl

Robert Nicholas

ABSTRACT

Ashley Marie Kretsch: Mining the *Pseudomonas virulence factor* pathway for novel
small molecules
(Under the direction of Bo Li)

Bioactive small molecules often play an important role in bacterial virulence. Identification and exploration of these molecules will develop an understanding of new mechanisms for infection and potentially identify novel targets to inhibit virulence. Analysis of bacterial genomes identifies a large number of gene clusters that encode for small molecule-synthesizing enzymes (biosynthetic gene clusters), many of which are unexplored. Through genetic manipulation of these clusters, we can discover their small molecule products and determine their biological roles.

We have identified a biosynthetic gene cluster conserved in over 300 strains of *Pseudomonads.* Many of these strains engage in pathogenic or symbiotic relationships with a range of human, animal, and plant hosts. We have shown that deletions within this gene cluster, named the *Pseudomonas virulence factor (pvf),* reduce fly infection by the pathogen *Pseudomonas entomophila.* In particular, the small molecule products of the *pvf* pathway are suggested to play a role in bacterial signaling and activation of virulence.

Through overexpression of the *pvf* cluster in its native *P. entomophila* strain, we have discovered and characterized a number of small molecules that the *pvf* gene cluster is responsible for. These molecules include a class of pyrazine compounds, some of which are new to biology. We use fly infection and promoter-reporter assays to determine the biological roles of these molecules in virulence and cell-to-cell signaling of *P. entomophila.* Through proteomic analysis of the secretome, we have identified over 500 proteins that are differentially

expressed between wildtype and *pvf* deletion strains, many which are known to be important to virulence and competition.

To identify the active molecule responsible for *pvf* signaling activity we developed an optimized bioactivity-guided extraction and purification method. We can compare active purified fraction from spent media extracts of wildtype and *pvf* deletion strains using nuclear magnetic resonance (NMR) and liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS) to identify the active *pvf* molecule. By characterizing the structures and functions of bioactive small molecules from pathogens, our work has the potential to develop an understanding for the production of these compounds and reveal novel and useful antimicrobial targets.

To Chance Crompton. Your confidence in your career inspired me to pursue mine. Wish you were here to celebrate our accomplishments together. Rest in Peace

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Bo Li, for her mentorship and guidance throughout my graduate schooling. Your trust in my work meant I was able to pursue a project I absolutely loved, gain independence and creativity, and push through the days where accomplishments seemed impossible. Most importantly, you taught me the best leaders are those who lead by example; your own work ethic and commitment inspired me to achieve my goals.

I would like to thank my family, particularly my parents. From a young age they saw my potential and gave me every opportunity imaginable to be challenged. They sat through countless hours of sporting events, took me to gifted student summer camps, and travelled the world with me. They soaked up so many tears are the hard days, and celebrated on the good days. As the only female computer scientist when she started working, my mother has been nothing but inspiration to pursue an academic career. I am also honored to be one of three successful sisters, all in different fields, but all with the same tenacity (and stubbornness). My grandparents and aunt have been a huge support system during my time in Chapel Hill, even driving through the middle of the night to let me be surrounded by love and compassion. Spending holidays with them has been a great gift in graduate school. I am so blessed to be raised by a powerhouse of women: my mom, my grandmothers, and my great-grandmothers.

I would like to thank my lab mates and support system in Chapel Hill, particularly my classmates, the First Five: Andy Chan, Erinn O'Neill, Zach Dunn, and Kevin Santa Maria. Starting in a brand-new lab takes a lot of hard work and great minds, and I am thankful for all

TABLE OF CONTENTS

LIST OF TABLES

LIST OF SCHEMES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AHL   Acyl homoserine lactone

ANOVA   Analysis of variance

ATP   Adenosine triphosphate

BGC   Biosynthetic gene cluster

CAS   Cromazol S reagent

COSY   Correlated spectroscopy

DAP   Diaminopimelic acid

DCM   Dichloromethane

dPNO   Dihydropyrazine *N*-oxide

EA   Ethyl acetate

EVC   Empty vector control

HE   Heterologous expression

HMBC   Heteronuclear multiple bond correlation

HPLC   High performance liquid chromatography

HSQC   Heteronuclear single quantum coherence

ImC   Imidazole cyclane

IPTG   Isopropyl β-D-1-thiogalactopyranoside

IR   Infrared

KO   Knockout

L   Liter (mL milliliter µL microliter)

LB   Luria Broth

LC-HRMS   Liquid chromatography coupled with high resolution mass spectroscopy

LFQ   Label-free quantification

mg   Milligram (µg microgram)

mnl	Monalysin (gene)
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
NRPS	Non-ribosomal peptide synthetase
OD	Optical density
OE	Overexpression
ONPG	Ortho-nitrophenyl-β-galactoside
PCA	Principal component analysis
PCR	Polymerase chain reaction
PKS	Polyketide synthetase
PNO	Pyrazine *N*-oxide
pvf	*Pseudomonas virulence factor (genes)*
PVF	Molecule produced by the *pvf* gene cluster responsible for signaling activity
QS	Quorum sensing
Rpm	Rotation per minute
WT	Wildtype
T6SS	Type VI secretion system
TFA	Trifluoroacetic acid
TOSCY	Total correlated spectroscopy
UV	Ultraviolet
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1 Introduction.

1. 1 The new Gold Rush for antibiotics and signaling molecules by genome mining in microbes

The rise and fall of traditional antibiotics and strategies to identify novel therapeutics

According to the World Health Organization (WHO) **"Antibiotic resistance is one of the biggest threats to global health, food security, and development today".[1]** Since their discovery, the promising ability of antibiotics to combat infectious bacteria has prompted their widespread and inappropriate use. With their deployment in medicine and agriculture, resistance has spread into community and environmental transmission. As a result, there has been a dangerous rise in the presence of "superbugs", pathogens that have developed resistance against every deployed antibiotic, while the development of new therapeutics has severely lagged behind.[2]

The discovery of penicillin ushered in the Golden Age of antibiotic discovery from the 1940s through the 1960s, when many of the antibiotics currently available were identified from bacteria and fungi. These compounds, known as natural products or secondary metabolites, are not required for survival, but produced by organisms for a selective advantage in its environment. After this, synthetic improvements of these scaffolds brought the subsequent Golden Age of medicinal chemistry into the new millennium. As pathogens continue to develop resistance to current antibiotic scaffolds, new antibiotics with novel chemistry or modes of action are desperately sought after. Unfortunately, traditional screening methods are overmined, and rediscovery is the underwhelming result. High rediscovery rates has led to only six new antibiotic classes introduced to the clinic since the 1960s.[3] Now, it is crucial to identify new

methods for natural product discovery and novel antimicrobial therapies. This section addresses

strategies to achieve these goals: enhancing our access to novel natural products with new

discovery methods and shifting the paradigm from bioactivity-based discovery to prioritizing

organisms and natural production from novel environments, niches, and biological roles.

Recent expansion of discovery methods has enlarged the library of natural product-

producing bacteria and fungi, as well as identify undiscovered natural products in known strains.

The biosynthesis of natural products are encoded in bacterial and fungal genomes, typically in

clusters of genes known as biosynthetic gene clusters (BGCs). We can mine the ever-growing

database of genome sequences to identify novel natural products. Even in well-characterized

natural product producers, such as *Streptomyces*, less than half of the putative biosynthetic

clusters have been correlated with their products. Many biosynthetic pathways for natural

products include genes with highly conserved domains, such as non-ribosomal peptide

synthetases (NRPSs) and polyketide synthetases (PKSs). Prediction algorithms such as

antiSMASH use these conserved domains to identify putative biosynthetic clusters.[4] Although a

few hundred biosynthetic gene clusters are already known across bacterial and fungal

genomes, predictions suggestion that over ten thousand uncharacterized biosynthetic gene

clusters are present in currently sequenced species, with approximately 2.5 biosynthetic gene

clusters per megabyte of genomic DNA.[5] Genome mining is the leading strategy for natural

product discovery in this decade.[6, 7] Biosynthetic gene clusters can be identified without native

production of the corresponding natural product, or even without culturing of the producing

bacteria.

One important paradigm shift in the field of natural products is the increase in

understanding of the broad biological roles of natural products. For example, there are

numerous small molecule contributors to virulence besides antibiotics. These include

siderophores, which chelate and scavenge iron in iron-limiting environments, and signaling

molecules, which activate virulence pathways. Most conventional antibiotics target cell viability,

exerting a high selective pressure for resistant populations. Rather than target cell survival, scientists have turned to components of pathogenesis as new targets in producing the next generation of treatments. This approach has a number of advantages, including preservation of microbiome environments, expansion of antimicrobial targets, and the potential for decreased chance of resistance.

Target based discovery

The increase in available genomes has led to a rise in orphan gene clusters without a known biosynthetic product. In fact, the number of orphan gene clusters far surpasses the current resources to investigate them. Therefore, it is important to rapidly connect genes to their respectively molecules or to prioritize BGCs with novel chemistries or activities. Currently, prediction models are iterative and based on previously identified gene clusters. This can hinder efforts to discover novel chemistries from enzymes with unknown functions. In order to correlate potential activity of uncharacterized small molecules, BGCs can be connected to nearby transporter or resistance genes. These genes can help identify novel BGCs not easily identified with bioinformatic prediction tools, but that potentially contain enzymes with novel functions. In addition, some resistance genes are duplicated house-keeping genes with modifications at the antibiotic binding site. If the activity of the resistance gene is known, it could provide insights into the biological activity of the natural product before the structure is determined.

Target-directed genome mining has been an effective method to correlating antibiotics and their modes of action in the case of novobiocin (gyrase C), platensin (FabB/F), and griselimycin (DnaN).[6] In addition, the Moore group mined 86 *Salinospora* strains for duplicated house-keeping genes colocalized with putative natural product BGCs to prioritize a PKS/NRPS cluster with a putative fatty acid synthase resistance gene. They were able to identify a duplicated 20S proteasome B-subunit gene associated with the *sal* BGC which confers

resistance to salinosporamide A and discovered that the thiotetronic acid antibiotic biosynthetic pathway contains two copies of the resistance gene.[8]

Unlocking natural products from "cryptic" gene clusters

Secondary metabolites, as compared to primary metabolites, are not essential for growth and replication, but provide a selective advantage for the survival of the producing organism. Therefore, these products are often not constitutively expressed, and instead are induced under specific conditions. In some cases, small molecules are only produced in extremely low concentrations and can be overshadowed by other metabolites in chromatography experiments or not produced in sufficient quantities for purification and structure elucidation. Therefore, genetic or environmental manipulation is required to unlock these natural products.

There are several methods developed to increase the production of these natural products so that sufficient quantities can be isolated and purified for characterization. The most widely used is expression of the BGC in a heterologous host that does not natively harbor the BGC of interest. Notable and well-studied heterologous hosts are *Escherichia coli*, *Saccharomyces cerevisiae, Streptomyces coelicolor* [9] and *Pseudomonas putida.*[10] These strains can be genetically modified to remove production of their own secondary metabolites, increasing the resources for the expressed cluster of interest and decreasing background in chromatography experiments. When promoters and regulators of a cluster are unknown, they can be exchanged with an inducible promoter using homologous recombination in the native or heterologous host. Interestingly, similar to a modified heterologous host, activators of known pathways can be modified. Recently, the repair of the GacA transcriptional regulators in *Pseudomonas fluorescens* Pf01 lead to the identification of a novel lipopeptide gacamide A.[11] Another tool involves using chemical elicitors and signaling molecules to activate silent gene clusters. For example, sub lethal concentrations of the antibiotic trimethoprim have been used to activate several previously unknown pathways in *Burkholderia thailandensis.*[12]

Although the power of bioinformatic predictions and genome mining are immense, purification of the small molecule products is essential to confirm structures. A dual pronged approach, combining *in vitro* biosynthetic studies with *in vivo* metabolomics experiments, can accelerate identification of small molecule products and verify bioinformatic predictions. Structure elucidation can be crucial for synthetic production or analysis of analogs and the lack thereof is a surprisingly common consequence of bioactivity guided discovery. For example, the potent phytotoxin albicidin from *Xanthomonas albilineans* is a DNA gyrase inhibitor at nanomolar concentrations, but due to its low production rate, had never been identified. A group has recently used an optimized heterologous expression platform in combination with *in vitro* studies fueled by bioinformatic predictions to determine the structure of albicidin and a putative biosynthetic pathway.[13]

Exploring the intersection: discovery from microbial and microbial-host interactions

Many discovery efforts to isolate natural products have relied on analysis of monocultures. Recent studies identified a wealth of novel small molecules through the interactions of microbes with each other, their host, and the environment.[14] In nature, microorganisms rarely live in isolation in rich media. More biologically relevant environments, such as co-cultures or specific medias, could activate biosynthetic pathways that are otherwise suppressed in a laboratory setting. A class of new macrolide antibiotics were isolated from a co-cultivation of *Penicillium fuscum* and *Penicillium clavigerum*.[15] The compounds, named the berkleylactones, were active against gram positive bacteria, including four methicillin-resistance *Staphylococcus aureus* strains. The berkleylactones exhibit a novel mode of action compared to other known macrolides.

Further, recent discoveries have found that the rhizosphere and human microbiota offer a wealth of bacteria that are inaccessible by traditional natural product discovery methods. By

studying BGCs in human-associated bacteria from the human microbiome projects, the Fishbach group discovered that BGCs for clinically relevant thiopeptides are widely distributed. They also identified and characterized a thiopeptide antibiotic lactocillin from a prominent bacteria in the vaginal microbiota.[16] Lactocillin and other thiopeptides were expressed *in vivo* and shown to be potent against a range of gram-positive vaginal pathogens. This is a strong example of the production of drug-like molecules by the human microbiota. Additionally, a novel thiazolidine-containing cyclic peptide antibiotic luodunin, produced by a human nasal commensal bacterium *Staphylococcus lugdunensis,* was shown to prohibit colonization by *S. aureus.*[17]

Advances in microbiome research has allowed for exploration of pathogenic bacteria that infect the gut, including recent studies of the pyrrolo [4,2] benzodiazepine (PBD) derivative tilivalline. This molecule is produced by *Klebsiella oxytoca* and was previously identified as the causative factor in antibiotic-associated hemorrhagic colitis.[18] Heterologous expression has been utilized to explore tilivalline biosynthesis. In these studies, salicyclic acid was identified as an inhibitor of tilivalline biosynthesis towards a potential application in therapy. Another disease-associated natural product is colibactin, produced by *E. coli* among others. Many studies have identified pieces of the colibactin puzzle: insights into the structure of this elusive compound, mode of action and biosynthesis.[19-21] Together, these colibactin studies aid future efforts to understand the impact of *E. coli* on human health and disease.

Recent investigations have looked at biological roles of secondary metabolites specifically to explore effects on the bacterial host. For instance, it has been shown that the phytotoxin coronatine produced by plant pathogen *Pseudomonas syringae* acts a signaling molecule in plants.[22] Coronatine mimics the plant hormone jasmonic acid isoleucine, which promotes the opening of the stomata and allows *P. syringae* to enter the leaves, enhancing bacterial growth. Combining genetic studies with bioinformatics can relate compounds with interesting chemistry to biosynthetic clusters and important biological roles. One group studied

lipopetides from the pathogens *P. syringae* B728a and DC3000 and biocontrol agents *P. fluorescens* Pf01 and SBW25. NMR of an isolated compound showed a structure identical to the genome predicted structure and random mutagenesis linked the predicted biosynthetic cluster for this lipopeptide to surfactant activity.[23]

New technologies will also contribute to the advancement of the metabolomics field. The Dorrestein lab have developed technologies to explore novel chemistries in metabolite samples as well as the ability to analyze the chemical environment of live bacterial interactions in real time. Desorption electron spray ionization mass spectrometry, or NanoDESI-MS, allows rapid collection of metabolite data with high sensitivity, and investigation of the spatiotemporal dynamics of metabolite production. Secreted molecules have a major impact on the phenotypic development of microbial populations, but are challenging to analyze on solid media. Only the more common compounds are accessible, and standard extraction protocols require a large investment of time and money in order to monitor temporal interactions between species. NanoDESI-MS is useful for the study of spatial or temporal changes in metabolite production in co-culture conditions or with the introduction of an elicitor. NanoDESI-MS was used to analyze the interaction between *Streptomyces coelicolor* and *Bacilus subtilis* py79.[24] They found that *B. subtilis* elicits pigment production and aerial hyphae formation in *S. ceolicolor,* which had previously been observed. Interestingly, previously unknown factors were regulated by the interactions between the organisms were identified, including a calcium-dependent antibiotic from *S. coelicolor*. NanoDESI-MS provides real time profiling of co-culture interactions revealing novel metabolites only present under these conditions.

In addition to pathogens, bacteria with biocontrol and plant disease suppression properties provide a rich source of natural products. Natural products produced by these bacteria are often specific for host or microbial interactions, therefore eliciting these factors are difficult in laboratory environments. The natural environments, such as the rhizosphere, are complex mixtures of microbes secreting a complex mixture of metabolites. Novel methods in

species identification and mass spectrometry facilitates natural product discoveries in complex microbial environments.

Within a microbial community, studies comparing soil that is conductive and suppressive to fungal infections have determined that certain species as more prevalent in suppressive soils.[25] Analysis of these prioritized strains led to the discovery of thanamycin, a potent lipopeptide. One method to identify disease suppressive bacteria in soil samples is a high density 16S ribosomal DNA (rDNA) oligonucleotide microarray known as PhyloChip. Surprisingly, in the study conducted by Mendes and coworkers, unclassified phyla represented 16% of bacterial families associated with disease suppression, demonstrating an untapped source of potential contributors to disease suppression.[25]

Over 99% of bacteria present in soil are unculturable. Some can be grown by cultivation *in situ* or with specific growth factors. iChiP grows single cells in media that is plated directly onto their soil environment, to maintain growth factors and nutrients. This can increase growth recovery from 1% to almost 50%. The antibiotic teixobactin was recently discovered through a screen of extracts from new bacteria isolated and cultured with this technique.[26] Teixobactin uniquely inhibits cell wall synthesis by binding lipid II and III, resulting in cell lysis. Exposure to this antibiotic showed no resistance in the many pathogens tested including MRSA and *Mycobacterium.tuberculosis*.

Discovering novel signaling molecules

Virulence factors play a major role in pathogenesis by invading the host, causing disease, or evading host defenses. Examples include toxins that disrupt host membranes, siderophores that scavenges for host iron resources, and biofilms that protect the pathogen from its host. One way these virulence pathways are regulated is through small signaling molecules, which are key aspects of pathogenesis and are currently not well-studied. Although much of natural product discovery focuses on virulence factors or small molecules with antibiotic

or therapeutic roles, it is crucial to study pathogenesis in the context of cell communication and downstream effects of signaling.

Bacteria use diffusible small molecules for communication that directly interact with receptors. This form of cell density dependent communication is known as quorum sensing (QS). Although the acylated homoserine lactones (AHLs) are the most well-studied QS molecules, other small molecule signals have been discovered in recent years. Interestingly, low concentrations of other secondary metabolites, such as antibiotics, can also modulate gene expression and act as elicitors to induce silent gene clusters.[14]

The AHL prototype for quorum sensing consists of an AHL synthetase (LuxI) and a cognate receptor LuxR. Many proteobacteria harbor LuxR receptors without a LuxI-type synthetase, suggesting that other molecules might acts as QS molecules and bind these 'orphan' LuxRs. Together, the Bode and Heermann groups identified a class of dialkylresorcinols and cyclohexanediones that were sensed by a LuxR homolog in *Photohabdus asymbiotica.*[27] The biosynthetic pathway for these molecules and the QS system is important for virulence in *P. asymbiotica*, a known human pathogen. They also identified a family of pyrones that act as signaling molecules in the insect pathogen *Photorhabdus luminescens.*[28] These α-pyrones are detected by an orphan LuxR-type receptor at low nanomolar concentrations.

Streptomyces bacteria produce over 70% of commercially available antibiotics, therefore understanding signaling and regulation in *Streptomyces* is important for expanding the repertoire of small molecules. The most common autoregulators in *Streptomyces* are γ-butyrolactone–type molecules including A factor, but only about 60% of strains are known to use these molecules. From heterologous expression of the SCP1 plasmid from *Streptomyces coelicolor* A3(2), a family of 5 new 2-alkyl-4-hydroxymethylfuran-3-carboxylic acids (AHFCAs), collectively termed Mm furans (MMFs), were identified as inducers of the production of the antibiotic methylenomycin.[29] Other *Streptomyces* may produce this furan family as a signaling molecule for the production of other novel metabolites. In addition, (4S,10R)-10-hydroxy-10-

methyl-9-oxo-dodec-2-en-1,4-olide, a novel autoregulator known as avenolide identified in

Streptomyces avermitilis, controls production of avermectin at an minimum effective

concentration of 4 nM.[30]

1.2 **Genome mining in *Burkholderia* and *Pseudomonas* to identify the *Pseudomonas virulence factor***

Pseudomonas bacteria thrive in diverse environments and interact extensively with eukaryotic hosts and other microbes. These bacteria produce a variety of small molecules that possess unique structures and functions, including siderophores, phytotoxins, antibiotics, and quorum-sensing molecules.[31-34] The genomes of these bacteria contain many uncharacterized biosynthetic gene clusters, including the *Pseudomonas virulence factor* (*pvf*) cluster (Figure 1.1). The *pvf* cluster was initially identified through genetic screens and is implicated in the virulence of the animal pathogen *Pseudomonas entomophila* L48 and the plant pathogen *Pseudomonas syringae* pv. *syringae* UMAF0158.[35-38] Disruption of *pvfB-D* in *P. entomophila* and *P. syringae* significantly decreases the virulence of these strains against adult flies and tomato plants, respectively.[37, 38] The *pvf* cluster was also found to be important for the biocontrol activity of a *Pseudomonas fluorescens* strain.[39] Additional studies found that the *pvf* cluster does not encode enzymes that synthesize toxins to directly harm the host, but that synthesize signaling molecule(s) that regulate virulence factors, including the pore-forming toxin monalysin in *P. entomophila* and the phytotoxin mangotoxin in *P. syringae.*[35, 38, 40]
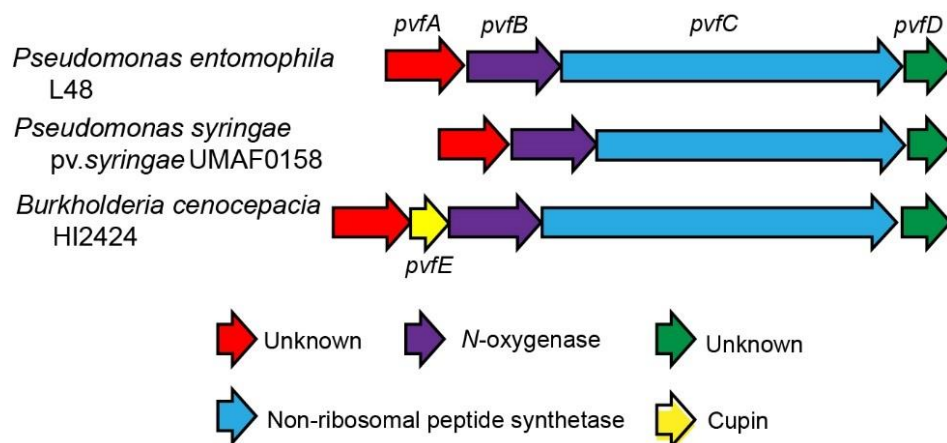
Figure 1.1. Genetic organization of the *Pseudomonas virulence factor (pvf)* in three biologically relevant bacterial strains. L48 is an insect pathogen, UMAF0158 is a plant pathogen, and HI2424 is a human pathogen. Alignment provided by Gina Morgan.

We first identified the *pvf* cluster through genome mining of the human pathogen *Burkholderia cenocepacia*, which is associated with poor clinical outcomes in patients with cystic fibrosis.[41] *B. cenocepacia* is capable of thriving in diverse and nutrient poor environments and is involved in both beneficial and pathogenic relationships with organisms in its environment.[42, 43] To determine the prevalence of the *pvf* cluster, we mined the available genomes using BLAST against the *pvf* cluster from *P. entomophila*. We found that the *pvf* cluster is present in over 300 sequenced strains (Figure 1.2). This cluster is widely conserved in *Pseudomonas,* including a significant number of pathogens and strains with biocontrol activity. Interestingly, the cluster is harbored by every sequenced strain of *P. syringae,* one of the most widespread plant pathogens, which are noted for their diverse and host-specific interactions with different plant species.[44]

The *pvf* operon is typically comprised of four genes: a putative nonribosomal peptide synthetase (NRPS, *pvfC*), a putative diiron *N*-oxygenase (*pvfB*), two genes of unknown function (*pvfA* and *pvfD*) and a fifth gene of unknown function *(pvfE)* present in some strains. NRPSs are a large family of multidomain enzymes that are responsible for the production of many bioactive natural products, such as the antibiotics vancomycin and daptomycin and the virulence factor pyoverdine.[45-47] PvfB shares homology with AurF and CmlI, which convert aryl amines into aryl nitro compounds in the biosynthesis of the antibiotics aureothin and chloramphenicol, respectively.[48-51]
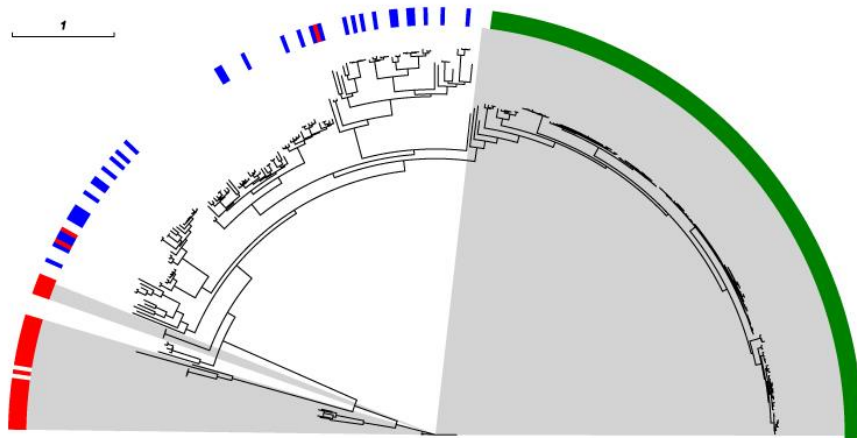
Figure 1.2. Phylogenetic tree of the NPRS PvfC. The *pvf* cluster is located in >300 strains of bacteria, including animal pathogens (red), plant pathogens (green), and strains with biocontrol activity (blue). Figure provide by Kevin Santa Maria

During our studies, homologues of *pvfA−D* were identified in *Burkholderia cenocepacia* H111 as part of the ham cluster (*hamA−G*, Figure 1.3), which was recently shown to be responsible for the synthesis of fragin, an antifungal molecule.[52] Deletion of the fragin biosynthetic gene *hamF* that is absent in the *pvf* cluster led to the discovery of valdiazen, a valinol diazeniumdiolate. Valdiazen exhibits moderate signaling activity toward the *hamA* promoter and is proposed to be the signaling molecule that autoregulates the expression of *ham* genes.[52] However, this study also found that valdiazen could not be identified in *pvf*-containing *Pseudomonas* strains that do not contain *hamF* or *hamG* homologues. Therefore, we propose *pvf*-encoded enzymes from *Pseudomonas* produce different bioactive molecule(s) than valdiazen and set out to identify these small molecules using *P. entomophila* as a model.
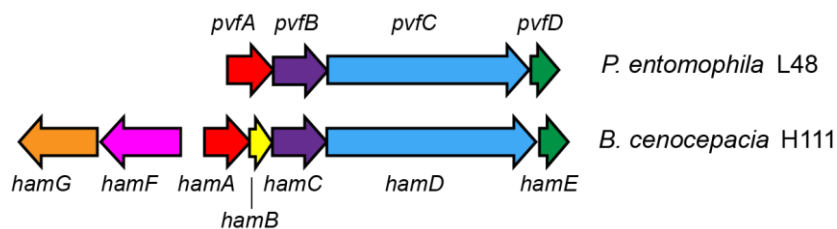
Figure 1.3. Comparison of *pvf* to ham gene cluster. Homologous genes in *pvf* and *ham* are shown in the same color. The ham cluster contains additional genes that are absent in *pvf* (hamG—aminotransferase, hamF—starter condensation domain, hamB—cupin domain). Alignment provided by Gina Morgan.

1.3 *P. entomophila* as a model to study the *pvf* cluster

    P. entomophila was first identified from a *Drosophila melanogaster* female, and is able to kill insects from at least three different orders.[53] *P. entomophila* has also been isolated from the rhizosphere and found to have potential biocontrol properties. [54] As a species, it is highly adaptive and a promising model to study host-pathogen interactions. For our studies, *P. entomophila* is an excellent model to study the *pvf* pathway because it has a well-studied infection model and a fully sequenced genome. [55]

    The genome of *P. entomophila* L48 has about 1000 unique genes when compared to five other *Pseudomonas* species.[55] These include numerous genes related to virulence and pathogenesis, hydrolytic activity (chitinases, lipases, proteases, and hydrolases), adaptation to the environment, as well as over 500 transporters and regulators. The antiSMASH webtool, which predicts biosynthetic gene clusters based on conserved domains in biosynthetic enzymes, found 12 putative gene clusters in the *P. entomophila* L48 genome (Figure 1.4). Among these, two are completely unknown, and eight have been shown to be important for virulence in previous studies.[38, 53, 56] These include the lipopeptide entolysin, the siderophores pyroverdine and pseudomonine, the *pvf* pathway and an aryl polyene. In addition to small molecule virulence factors, genetically encoded proteins can be virulence factors, such as lipases that break down host membranes. Several of the most abundant secreted proteins have also been shown to be virulence factors in *P. entomophila*: the metalloprotease AprA, the toxin monalysin, and the type VI secretion system (T6SS).[40]

Identified secondary metabolite regions

The following regions are from record NC_008027.1 (Pseudomonas entomophila L48):

Region	Type	From	To	Most similar known cluster		Similarity	MIBiG BGC-ID
Region 1	Nrpsfragment	102142	132047	Mangotoxin	NRPS	71%	BGC0000387
Region 2	Arylpolyene	340631	384238	APE Vf	other	40%	BGC0000837
Region 3	Nrps	1872154	1924904	Pyoverdine	NRPS	4%	BGC0000413
Region 4	Nrps	2225800	2284461				
Region 5	Nrps	2408664	2451337				
Region 6	Nrps	2597735	2648440	Pseudomonine	NRPS	100%	BGC0000410
Region 7	Nrps	2823588	2885033	Bicornutin	NRPS	100%	BGC0001135
Region 8	Betalactone	3128018	3160045	Pyoverdine	NRPS	4%	BGC0000413
Region 9	Nrps	3199204	3279073	Entolysin	NRPS	95%	BGC0000344
Region 10	Nrps	3438009	3516451	Pyoverdine	NRPS	15%	BGC0000413
Region 11	Bacteriocin - Nrpsfragment - Nrps	3531164	3618247	WLIP	NRPS	50%	BGC0000462
Region 12	T1pks	5845294	5888056	Micromonolactam	polyketide	100%	BGC0000095

Figure 1.4. antiSMASH analysis of the *Pseudomonas entomophila* L48 genome (NC_008027.1).

Lipopetides and surfactant activity

P. entomophila is predicted to produce three lipopeptides: entolysin,[38, 56] and two uncharacterized lipopeptides. Entolysin has been studied in *P. entomophila* as the lipopeptide surfactant responsible for hemolytic activity in *P. entomophila*. Genetic studies suggest that the biosynthetic gene cluster contains a LuxR family regulator and export system, but the regulatory proteins lack any homoserine lactone binding domain.[56] The regulation of these lipopeptides has yet to be identified, but could be controlled by the *pvf* signaling molecule.

Aryl polyene

antiSMASH analysis predicts that *P. entomophila* contains a biosynthetic gene cluster for aryl polyenes, a widely distributed gene cluster in Gram-negative bacteria.[5] Aryl polyenes are pigments that are structurally similar to carotenoids. Analysis of biosynthetic genes for aryl polyenes and carotenoids shows complementary distribution.[57] Similar to carotenoids, aryl polyenes were shown to protect *Variovorax paradoxus* from reactive oxygen species.[57]

Siderophores and iron scavenging

Like many *Pseudomonas* species, *P. entomophila* has adapted to withstand unfavorable environments, including iron-scarce environments such as an insect host.[53] The genome of *P. entomophila* harbors numerous ton-B receptor genes that are involved in iron scavenging and virulence. *P. entomophila* produces at least two siderophores, a pyroverdine and pseudomonine.[58] *P. entomophila* has also been shown to uptake a number of exogenous pyroverdine molecules, utilizing iron trapped by siderophores from other bacteria.[58] Interestingly, *in vitro* studies with the proteins involved in pseudomonine biosynthesis found that the promiscuity of the enzymes led to production of a family of siderophores, including anguibactin and acinetobactin.[59]

Proteases and toxins

AprA is the most abundant protein in *P. entomophila's* secretome.[60] AprA is an alkaline metalloprotease that is 51 kDA and secreted via a type one secretion system (T1SS). AprA is known to cleave pro-monalysin to its active form outside of the bacteria.[40] Previous studies suggest that AprA also degrades antimicrobial peptides produced by the fly as part of the fly immune response, which could also have a more direct effect on cytotoxicity.[60] AprA is solely responsible for *P. entomophila's* protease activity on skim milk plates.[60]

Monalysin is a pore forming toxin that interferes with host cell membranes.[40] Monalysin production is regulated by *GacA* and *pvf* signaling systems; knockouts of the *GacA* and *pvfC* genes reduce the production of monalysin.[40] The monalysin promoter will be used as a positive control for our global proteomic and transcriptomic studies of *pvf.*

Type VI Secretion System

P. entomophila is the only *Pseudomonas* strain that is pathogenic in multicellular organisms and lacks a type three secretion system (T3SS).[61] Components of the Type VI secretion system

16

(T6SS), including Vgr, Rhs and Hcp proteins, are among the most abundant secreted proteins,[40] and in other species are known virulence factors. Type VI secretion is a widespread mechanism for protein transport in Gram-negative bacteria using a syringe-like protein apparatus. The *P. entomophila* genome harbors a genetic island containing T6SS components, as well as a number of solo Vgr/Hcp pairs.[61] Deletion of *ttsJ*, a component of the T6SS needle, did not affect fly infection in *P. entomophila*.[40] Other studies in *Pseudomonas* and *Burkholderia* suggest that although T6SS might not play a direct role in infection, it is likely involved in competition with other bacteria.[62-64]

Chapter 2 Discovery of novel pyrazine *N*-oxides by overexpressing the *Pseudomonas virulence factor* pathway

2.1 Introduction

Pathogens will express an array of virulence factors, including small molecules, during infection.[44, 53, 65, 66] However, the same strains grown in the lab will only produce a fraction of these metabolites. Increasing availability of bacterial genomes and bioinformatic analysis has revealed the discrepancy between secondary metabolite production in standard laboratory conditions and the number of biosynthetic pathways encoded in the genome[5] The remaining secondary metabolites are produced by biosynthetic gene clusters with unknown promoters, also considered to be silent, or "cryptic", gene clusters. In addition, many bacterial small molecules, especially signaling molecules, are natively produced at low levels.[67] Natural product discovery from bacterial cultures in rich media can lead to rediscovery of known metabolites.[68] Overcoming this challenge involves controlling the expression of cryptic biosynthetic gene clusters through inducible promoters, elicitors, or biologically-relevant culture conditions.
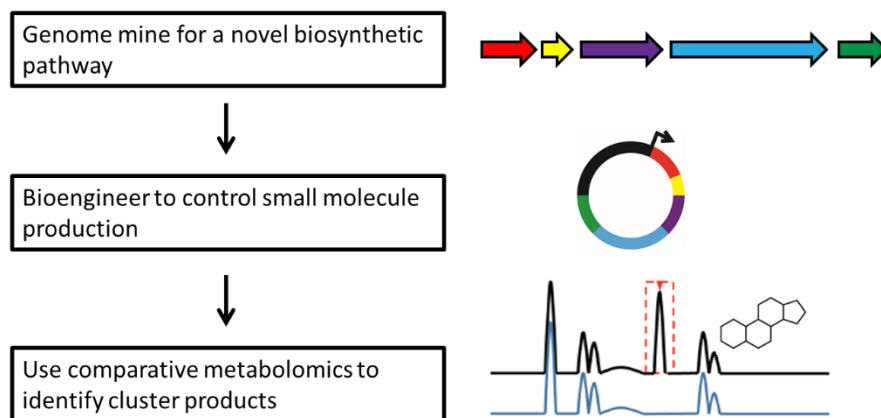
Figure 2.1. Genes-to-molecules strategy to identify small molecule products from cryptic gene clusters.

We use a genes-to-molecules strategy to express cryptic gene clusters and elucidate the small molecules intermediates and products generated by their encoded proteins (

Figure *2.1*). Similar genome-driven discovery methods have been used previously to discover the cryptic polyketide burkholderic acid in *B. thailandensis*.[69] The work in this chapter highlights my work in developing this strategy for the lab.

This system has three components: (1) inducible expression of the gene cluster, (2) genetic knockout of the gene cluster, and (3) complementation of the genetic knockouts with inducible expression. For inducible expression, we clone the gene cluster from the genome and incorporate it into an expression plasmid that (i) contains an inducible promoter for control of small molecule production and (ii) can replicate in the native or host bacteria. For genetic knockout, we use the homologous recombination to remove the operon of interest. Finally, complementation of the expression plasmid into the knockout strain will recover production of the small molecules biosynthesized by encoded proteins.

I constructed both native and heterologous expression strains to characterize the products of the *pvf* pathway. The benefit for using the native system includes (1) presence of substrates and enzyme cofactors for production of the natural product, (2) potential to determine downstream biological and signaling roles. The challenge of using a native strain is that overexpression of *pvf,* which is implicated in signaling, could activate many downstream pathways and result in the production of other metabolites, making it difficult to distinguish the molecules that are directing biosynthesized by *pvf*-encoded proteins. To overcome this potential challenge, I used heterologous expression in *P. aeruginosa* PAO1, a strain which does not natively harbor the *pvf* pathway, to confirm that observed molecules are specifically produced by *pvf*-encoded enzymes. I expect that expression of *pvf* in the heterologous strain will provide a cleaner background than the native expression strain for determining *pvf*-specific products.

We combined *in vitro* reconstitution and in cell metabolomics to identify small molecule products of the *pvf* cluster, a two-pronged approach that our lab has successfully employed in a

separate study.[70] *In vitro* studies can help identify enzyme substrates and pathway intermediates and provide rationale and verification for the biosynthetic products and intermediates. If substrates are identified by *in vitro* methods, isotopically-labelled substrate will be added to cultures to validate that molecules of interest are produced of the particular biosynthetic proteins. Together these techniques can paint a picture of the small molecule products from our genome-mined *pvf* clusters and their biosynthesis.

Through overexpression of the *pvf* operon, we identified and characterized a new family of natural products, (dihydro)-pyrazine *N*-oxides ((d)PNOs). We found that PvfC and PvfB are responsible for the biosynthesis of this unique family of compounds. With this evidence, we propose a biosynthetic pathway for (d)PNOs and suggest a shared biosynthetic intermediate for (d)PNOs and valdiazen.

2.2 Materials and Methods

General methods

Source of chemicals and biological reagents. Chemicals were purchased from Sigma or Fisher Scientific unless otherwise stated. Polymerase chain reactions (PCR) were conducted using Q5 DNA polymerase (New England Biolabs, NEB) and PCR products were isolated using either the PCR cleanup kit (QIAGEN) or DNA gel extraction kit (Zymo Research) after excision from an agarose gel containing ethidium bromide. Ligations were performed using T4 DNA ligase (NEB) overnight at 16 °C. *E. coli* electrocompotent cells were transformed by electroporation. Plasmids were isolated from overnight cultures using the GeneJet plasmid miniprep kit (Thermo Scientific) and verified by DNA sequencing (Eton Biosciences).

Bacterial strains and culture conditions. *Pseudomonas entomophila* L48 and *Pseudomonas aeruginosa* PAO1 were grown in Lennox Luria Broth (Low salt LB) or M9 minimal media (47.75 mM $Na_2HPO_4 \cdot 7H_2O$, 22.04 mM KH_2PO_4, 8.56 mM NaCl, 18.7 mM NH_4Cl, 2 mM $MgSO_4$, 100 µM CaCl, 22.2 mM glucose) at 30 °C for all experiments. *E. coli* TOP 10 or DH5α was grown in Miller Luria Broth at 37 °C. Antibiotics were used at the following concentrations in growth media: gentamicin, 50 mg ml^{-1} for liquid cultures and 150 mg ml^{-1} for solid media, ampicillin 100 mg ml^{-1}. Antibiotics and IPTG were purchased from GoldBio.

Inducible *pvf* expression in native and heterologous *Pseudomonas* strains

Inducible expression of *pvf* in *P. entomophila*. The pPSV35 *Pseudomonas* expression vector carrying *pvfABCD* under the IPTG-inducible lacUV5 promoter (pPSV35-*pvfABCD*) was constructed previously.[38] *P. entomophila ΔpvfC* was transformed with pPSV35-*pvfABCD* and the empty vector control pPSV35 using the electroporation method described by Choi *et al.* [71] Briefly, a sample of 6 mL overnight cultures of *Pseudomonas* were divided into four 1.5 mL microcentrifuge tubes and the cells were harvested at 6000 x g for 5 minutes. After

removal of the supernatant, the cells were washed twice with 1 mL of 300 mM sucrose per tube. The cell pellets were combined and resuspended in 100 µL of 300 mM sucrose to generate electrocompetent cells. An aliquot of 50 µL cells in sucrose was gently mixed with 1 µL of 200 ng µL^{-1} plasmid and transformed by electroporation (1 mm gap, 2.0 keV). These transformed cells were recovered in 1 mL LB for 2 hours while shaking at a speed of 225 rpm. A sample of 150 µL of the cell suspension was plated on LB agar containing gentamycin and incubated at 28 °C for 1 day until colonies appeared.

Heterologous expression of *pvf* in *P. aeruginosa*. *P. aeruginosa* PAO1 cells were transformed with pPSV35 plasmids containing various *pvf* constructs using biparental mating. *E. coli* RHO3 cells, diaminopimelic acid (DAP) auxotrophs of the *Δasd E. coli* mobilizer strain,[72] were transformed with each plasmid. RHO3 cells harboring the plasmid were grown overnight in LB supplemented with 400 µg ml^{-1} DAP and gentamicin at 37 °C. The *P. aeruginosa* PAO1 recipient was grown in parallel overnight in LB at 37 °C. A sample of 5 mL of each overnight culture was harvested by centrifugation at 3500 x g for 5 minutes. Culture supernatants were removed and the cells were resuspended in 500 µL sterile filtered 10 mM MgCl$_2$. Samples of 25 µL of the RHO3 donor strain and 25 µL of the *P. aeruginosa* recipient strain were added to the center of an LB agar plate containing DAP. Plates containing the bacterial mixture were dried at room temperature for 1 hour. Plates were then incubated facing up for 24 hours at 37 °C. Cells at the center of the plate were scooped and resuspended in 60 µL of 10 mM MgCl$_2$, plated on LB agar containing gentamycin for selection. After the plates were incubated at 37 °C overnight, colonies were selected and verified for containing the desired plasmids by colony PCR.

Comparative metabolomics

Growth conditions. Overnight cultures of *P. entomophila* or *P. aeruginosa* that harbor pPSV plasmids containing various *pvf* constructs were grown in 3 mL at 28 °C to saturation. A

sample of 200 μL of overnight culture was used to inoculate 50 mL of Lennox LB containing

gentamycin and incubated at 30 °C with 225 rpm shaking. Expression of *pvf* genes was induced

at OD_{600} = 0.3 by adding a final concentration of 1 mM IPTG and grown at 30 °C and shaking at

225 rpm for 24 hours. Four biological replicates were grown for each strain.

Extraction of culture supernatants using dichloromethane. Cultures (50 mL) were

spun at 5000 x g to remove cells. The supernatant was filtered through a 0.45-micron filter and

the pH was adjusted to 5.0 with 6 N HCl solution. Metabolites were extracted three times with

one-third volume of dichloromethane (DCM) each time. The organic layers were pooled and

dried with sodium sulfate. The extracts were evaporated to dryness in a round bottom flask, and

the dried extracts were resuspended in 5 mL of methanol and transferred to a scintillation vial.

This resuspension was again evaporated to dryness and stored at −20 °C.

Analysis of metabolite extracts by LC-HRMS. The extracted metabolites from each 50

mL of culture were resuspended in 250 μL of 50:50 water:acetonitrile mixture, centrifuged at

maximum speed (>10000 x g) for 10 minutes to remove any particulates, and diluted four-fold

before analysis. A 10 μL sample was analyzed by liquid chromatography-coupled high-

resolution mass spectrometry (LC-HRMS) using the 6520 Accurate Mass Quadrupole Time-of-

Flight Mass Spectrometer (Agilent Technologies). For the liquid chromatography, mobile phase

A contains water and 0.1% formic acid and mobile phase B contains acetonitrile and 0.1%

formic acid. Metabolites were separated on a Thermo Scientific Hypercarb 100 x 2 mm column

at a flow rate of 0.5 mL/min using a gradient of 2% B for 2 min, 2−98% B over 16 min, 98% B for

two min. All absorbances in the range of 190−450 nm were recorded. Mass spectrometry

analysis by Electrospray Ionization (ESI) was carried out under positive ion mode using the

following parameters: gas temperature 325 °C, drying gas 10 L/min, nebulizer 45 psi,

fragmentor 175 V, skimmer 65 V, capillary cap 3500 V, octopole RF 750 V.

Comparative analysis of LC-HRMS data. The metabolic profiles of positive and control

samples were assessed visually by comparing total ion chromatograms followed by analysis

using the web-based metabolomic platform XCMS.[73] Lists were constructed from the comparison of 1) *P. entomophila* L48 Δ*pvfC* and *P. entomophila* Δ*pvfC* + pPSV35-*pvfABCD* and 2) *P. aeruginosa* PAO1 + pPSV35 and *P. aeruginosa* PAO1 + pPSV35-*pvfABCD*. These lists were cross-referenced to identify mass peaks present in both positive samples (*P. entomophila* L48 Δ*pvfC* + p*PSV35-pvfABCD* and *P. aeruginosa* PAO1 + pPSV35-*pvfABCD*) and absent in both negative controls (*P. entomophila* L48 Δ*pvfC* and *P. aeruginosa* PAO1 + pPSV35). The extracted ion chromatograms of each mass peak of interest were manually analyzed for peak shape and ion intensity.

Metabolite purification and structural characterization

Growth conditions. Eight 1 L LB-gentamycin cultures were inoculated with 2 mL overnight cultures of *P. entomophila* L48 Δ*pvfC* + pPSV35-*pvfABCD*. The cultures were incubated at 30 °C and induced with 1 mM IPTG when OD_{600} reached 0.3. After 40 hours of growth at 30 °C, the culture supernatant was separated from bacterial cells by centrifugation and filtration. The supernatant was adjusted to a pH of 5.0 and extracted three times with a one third volume of DCM. The organic layers were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The metabolites extracted were resuspended in 5 mL of methanol, transferred to a scintillation vial, dried down, and stored at −20 °C.

Purification by preparatory HPLC. Metabolites **1**, **2**, and **3** were purified from the large-scale extraction using preparatory high-performance liquid chromatography (HPLC, Varian Prostar with 330 PDA variable wavelength detector). Dried DCM extract was resuspended in 80 % ACN/water (1 mL per 1 L culture extracted) and injected on a Phenomenex Luna C18 column in 1 mL injections. Compounds of interest were separated using mobile phase A (water, 0.1% TFA) and mobile phase B (acetonitrile, 0.1% TFA) over a gradient of 5 minutes at 5% B, 4 min 5−15% B, 33 min 15−25% B, 7 min 25−95% B, and 6 minutes 95%, at a flow rate of 18 mL/min.

All absorbance in the range of 190−450 nm were recorded. Fractions containing a compound of interest based on UV and subsequent MS analysis were combined separately and dried under high vacuum. The retention times were 17 min (**1**), 20 min (**2**), and 40 min (**3**).

Direct injection mass spectrometry. The collected HPLC fractions were analyzed by HRMS under positive ion mode. A sample of 35 µL of each fraction was directly injected onto the mass spectrometer using a mobile phase containing 50% acetonitrile and 0.1% formic acid at a flow rate of 0.4 mL min^{-1}. MS analysis was conducted using the following parameters: gas temperature 325 °C, drying gas 10 L min^{-1}, nebulizer 45 psi, fragmentor 175 V, skimmer 65 V, capillary cap 3500 V, octopole RF 750 V.

Analysis of metabolites by tandem mass spectrometry. A 10 µL sample prepared as previously described was analyzed by liquid chromatography-coupled high-resolution mass spectrometry (LC-HRMS) using the 6520 Accurate Mass Quadrupole Time-of-Flight Mass Spectrometer (Agilent Technologies). For the liquid chromatography, mobile phase A contains water and 0.1% formic acid and mobile phase B contains acetonitrile and 0.1% formic acid. Metabolites were separated on a Thermo Scientific Hypercarb 100 x 2 mm column at a flow rate of 0.4 mL/min using a gradient of 2% B for 2 min, 2−95% B over 26 min, 98% B for two min. All absorbances in the range of 190−450 nm were recorded. Mass spectrometry analysis by Electrospray Ionization (ESI) was carried out under positive ion mode using the following parameters: gas temperature 300 °C, drying gas 10 L/min, nebulizer 45 psi, fragmentor 175 V, skimmer 65 V, capillary cap 3500 V, octopole RF 750 V. Eight precursor ions were selected for fragmentation; the same ion was excluded after three spectra with an exclusion release after one minute. Collision energy was dependent on ion size with 3V/100Da and an offset of 10V. Isolation was set to a medium width of approximately 4 *m/z*.

(^{15}N)-M9 cultures and extractions. To produce ^{15}N enriched metabolites for (^{1}H, ^{15}N) HMBC NMR analysis, 1 liter of *P. entomophila ΔpvfC* + pPSV35-*pvfABCD* cultures were grown under the same conditions as the non-labeled cultures. Low salt LB media is replaced with M9

minimal media supplemented with ^{15}N ammonium chloride ($^{15}NH_4Cl$). Metabolites were extracted and purified as described in the section "Purification by preparatory HPLC". LC-HRMS analysis on the extract indicated the mass to charge ratios of metabolites **1**, **2**, and **3** increased by 2 (Figure 2.8), indicating incorporation of two ^{15}N into each metabolite that is consistent with the predicted chemical formula.

Structural characterization. Purity of isolated **1**, **2,** and **3** was analyzed by LC-HRMS as described in the section "Analysis of metabolite extracts by LC-HRMS". Purified samples of **1**, **2,** and **3** were characterized by NMR and **1** was also characterized by IR. NMR experiments of **1** and **2** were conducted in deuterated chloroform, and experiments of **3** were conducted in deuterated methanol. Each compound was analyzed by ^{1}H, (^{1}H, ^{1}H) COSY, (^{1}H, ^{13}C) HSQC, (^{1}H, ^{13}C) HMBC, and ^{13}C NMR experiments, using Bruker 600 MHz spectrometer equipped with a cryoprobe unless specified. ^{15}N-enriched compounds were analyzed by (^{1}H, ^{15}N) HMBC experiments. Finally, solid state infrared spectroscopy experiments for **1** were conducted with a Thermo Scientific Nicolet iS5 FT-IR.

Chemical synthesis of PNO A (2) and PNO B (3)
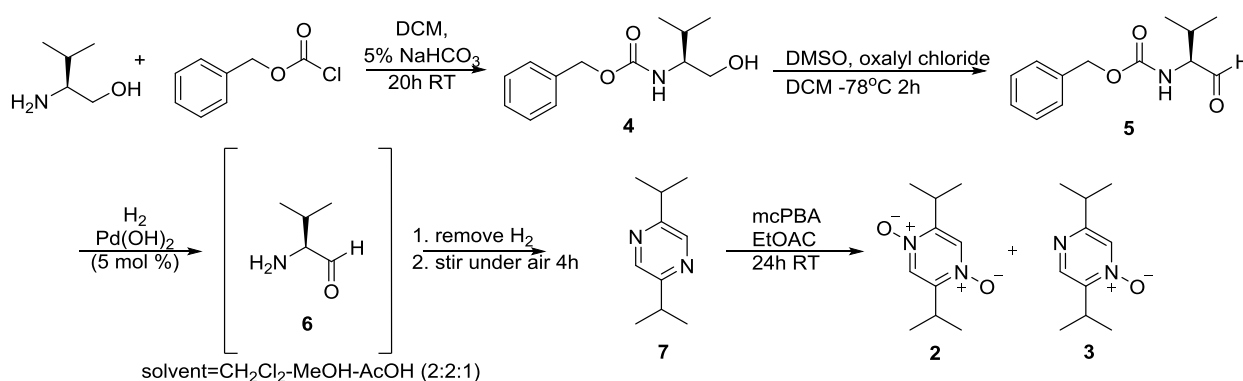
Scheme 2.1. Synthetic route for 2,5-diisopropylpyrazine-*N,N'*-dioxide (**2**) and 2,5-diisopropylpyrazine-*N*-oxide (**3**).

Synthetic standards of compounds **2** and **3** were produced in four steps starting with L-valinol (Scheme 2.1).

Carbamic acid, *N*-[(1*S*)-1-(hydroxymethyl)-2-methylpropyl]- phenylmethyl ester (4).

Benzyl chloroformate protected valinol (**4**) was produced according to literature procedures.[74] 500 mg (4.84 mmol, 1 equiv.) of valinol was dissolved in 6.5 mL of DCM. A sample of 8 mL of a 5 % aqueous solution of NaHCO$_3$ (0.44 g in 8 mL DI H$_2$O, 5.32 mmol, 1.10 eq) was added. The solution was at pH 10. Upon the addition of 3.74 mL (24.24 mmol, 5 eq.) benzyl chloroformate, the reaction became cloudy. The mixture was stirred at room temperature for 20 h. The aqueous and organic layers were separated and the aqueous layer was extracted with an additional 10 mL of DCM. The organic layers were combined, washed with brine, dried with Na$_2$SO$_4$, filtered, and concentrated to give an opaque, colorless oil. This crude oil was purified using a Biotage silica gel purification system (solvents: hexanes/ethyl acetate) to afford **4** (510 mg, 2.15 mmol, 45%). Clear oil.

Carbamic acid, *N*-[(1*S*)-1-formyl-2-methylpropyl]- phenylmethyl ester (5). Cbz-valinal (**5**) was produced using conditions previously outlined.[75, 76] Under nitrogen atmosphere, a sample of 709 μL (10.04 mmol, 4.2 eq.) of DMSO was slowly added to a stirring solution of 416 μL (4.78 mmol, 2 eq.) of oxalyl chloride in 4.5 mL of DCM at −78 °C. At this temperature the solution was stirred for 30 minutes. A sample of 0.500 g of **4** (2.39 mmol, 1 eq.) dissolved in 4.5 mL of DCM was added slowly over 10 min and stirred for an additional 2 h at −78 °C. A sample of 2.5 mL DIPEA (7.17 mmol, 3 eq.) was added at −78 °C. The temperature was slowly raised to room temperature over 20 min, and the reaction mixture was diluted with 8 mL of DCM. The organic layer was washed with saturated aqueous NH$_4$Cl solution and brine, dried over anhydrous Na$_2$SO$_4$, and concentrated *in vaccuo* to generate a crude oil. The crude oil was purified using a Biotage silica gel purification system to afford **5** (368 mg, 1.56 mmol, 65%). Light yellow oil.

2,5-Diisopropylpyrazine (7). 7 was prepared according to literature procedures.[77] A sample of 368 mg of **5** (1.56 mmol, 1 eq.) was dissolved in 18 mL of 1:2:2 acetic acid:methanol:dichloromethane. 55 mg of Pearlman's catalyst (Pd(OH)$_2$, 5 mol%) was added to

the solution. The reaction mixture was stirred under an atmosphere of hydrogen for 1 h at room temperature. The hydrogen source was removed and the reaction mixture was stirred at room temperature for 4 h while open to the air. The reaction mixture was filtered under vacuum. This mixture was extracted with dichloromethane, washed with water, dried with Na_2SO_4 and concentrated *in vaccuo*. Crude oil was purified using a Biotage silica gel purification system to afford **7** (50 mg, 0.3044 mmol, 19.5%). Light yellow oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.41 (s, 2H) 3.10 (sept, 2H) 1.34 (d, 12 H).

2,5-diisopropylpyrazine-*N,N'*-dioxide (2) and **2,5-diisopropylpyrazine-*N*-oxide (3).** Procedure was modified a from previous report.[78] A sample of 120 mg mCPBA (70% purity, 0.488 mmol, 5.4 eq.) was extracted with 2 mL ethyl acetate and 2 mL brine, dried with Na_2SO_4, and filtered. This filtrate was added to a solution of 15 mg **7** (0.091 mmol, 1 eq.) in 650 µL of ethyl acetate. The solution was stirred at room temperature for 25 hours. Reaction was quenched with 2 mL saturated thiosulfate, washed with $NaHCO_3$ and brine, dried with Na_2SO_4 and concentrated *in vacuuo*. The resulting mixture was resuspended in 4 mL 80% acetonitrile/water and products were separated on a preparative HPLC with mobile phase A (water, 0.1% TFA) and mobile phase B (acetonitrile, 0.1% TFA) using a gradient of 5−95% B over 55 minutes to afford compounds **2** (8.0 mg, 0.041 mmol, 45%) and **3** (1.5 mg, 0.0083 mmol, 9.2%). **2**, ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.03 (s, 2H) 3.54 (sept, 2H) 1.30 (d, 12H) UV: 195, 235, 303, 338 nm; **3**, ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.49 (s, 1H) 8.19 (s, 1H) 3.57 (sept, 1H) 3.03 (sept, 1H) 1.34 (d, 6H) 1.31 (d, 6H) UV: 198, 223, 265, 303 nm.

in vitro reconstitution of PvfC activity

Cloning, expression and purification of *N*-term His-tagged PvfC. *pvfC* was amplified from *P. entomophila* L48 genomic DNA using primers AK139 and AK140 and cloned into the pLICHis vector. The insert was digested and annealed according to the ligation independent

construction method[79]. The resulting plasmid, pLICHis-*pvfC,* was transformed into *E. coli* Top10 and verified by sequencing. Both *E. coli* Bap1 and BL21 DE3 strains were transformed with pLICHis-*pvfC* by electroporation for protein expression and purification.

A sample of 1 L of Luria Broth cultures containing ampicillin were inoculated with 5 mL of overnight cultures and grown at 37 °C until OD_{600} reached 0.5−0.6. Protein expression was induced with 0.5 mM of IPTG and grown overnight at 16 °C. Cells were harvested by centrifugation at 4500 x g. His_6-tagged protein was purified by nickel affinity chromatography and size exclusion chromatography using an Akta FPLC. Cell pellet was resuspended in 15 mL of wash buffer (50 mM HEPES at pH 7.5, 500 mM NaCl, 10% glycerol, 5 mM imidazole) and lysed using a Branson sonicator at 30% amplitude for 1.5 min of total 'on' time, cycling between 0.5 sec on and 1.5 sec off. Cell debris was pelleted at 17,000 x g and supernatant was filtered through a 0.45 μm syringe filter and injected onto a GE 5 mL His Trap HP Column. The His_6-tagged protein was eluted over a gradient of 5−300 mM imidazole. Fractions containing proteins of interest were identified by SDS-PAGE, pooled, and concentrated with an Amicon centrifugal concentrator to 5 mL or less for application to a GE Superdex 200 size exclusion column for further purification into storage buffer (Wash buffer with no imidazole). Protein was flash frozen in beads in liquid nitrogen at −80 °C.

$Na_4[^{32}P]PP_i$ exchange end point assay to determine PvfC adenylation specificity. To test the selectivity of the adenylation (A) domain, PvfC (5 μM) was incubated with 4 mM amino acid substrate, 1 mM $Na_4[^{32}P]PP_i$ (~2×10^6 cpm radioactivity per mL), in 50 mM HEPES (pH=7.5), 1 mM $MgCl_2$, 1 mM ATP, and 4 mM cold sodium pyrophosphate (PP_i). Each proteinogenic amino acid was individually tested as substrates. Approximately 200,000 cpm of 1 mM hot/cold ppi was added to the reaction (100 μL reaction volume). Reaction was initiated with addition of enzyme and incubated for 1 hour at room temperature. Reaction was stopped using 500 μL quenching solution (1.6% activated charcoal, 3.5% $HClO_4$, 0.1 M Na_4PP_i). Mixture was centrifuged to remove unreacted $Na_4[^{32}P]PPi$, washed with 1 mL of water, centrifuged and

resuspended in 500 µL water. A Beckman LS 6500 scintillation counter was used to measure the charcoal-bound radioactivity.

PvfC kinetic analysis. The $Na_4[^{32}P]PP_i$ exchange assay was performed using 0.5 µM PvfC and increasing concentrations of L-valine (0.1 mM, 0.25 mM, 1 mM, 2.5 mM, 5 mM, and 10 mM). The rest of the assay components were kept the same as the end point assay. An aliquot of 100 µL was removed at 0, 5, 10, 15, and 20 minutes from each reaction and immediately quenched with quenching solution for initial velocity and kinetic calculations. Kinetic parameters were calculated using GraphPad Prism.

Precursor feeding with [^{2}H$_{1-8}$]DL-valine

Samples of 3 mL overnight cultures of *P. entomophila* L48 Δ*pvfC* + pPSV35-*pvfABCD* were harvested by centrifugation at 3500 x g for 5 min. The supernatant was removed and pellet was resuspended in an equal volume of sterile water. Samples of 50 mL M9 media cultures containing gentamycin were inoculated with 250 µL of the resuspended bacterial cells. Cultures were grown at 30 ºC induced with 1 mM IPTG when OD_{600} reached 0.4. A sample of 100 µM of all proteinogenic amino acids (excluding valine and using cystine instead of cysteine) and 1 mM L-valine or 1 mM [^{2}H$_{1-8}$]DL-valine (Cambridge Isotopes) were added at inoculation. After a total of 24 hours of growth, the bacterial whole cultures were mixed with 50 mL ethyl acetate, vortexed for thirty seconds, and centrifuged at 4500 x g for 30 minutes to pellet cell debris and separate the aqueous and organic layers. The organic layer was separated from the supernatant, dried with magnesium sulfate then evaporated to dryness in a round bottom flask. The extracted metabolites were resuspended in 5 mL of ethyl acetate and transferred to a scintillation vial. The ethyl acetate was evaporated to dryness and the extraction sample was stored at −20 °C. The L-valine or [^{2}H$_{1-8}$]DL-valine feeding experiments were conducted in

quadruplicate. Samples were analyzed by LC-HRMS as described in the section "Analysis of metabolite extracts by LC-HRMS".

Expression of partial *pvf* cluster and metabolomics

All overexpression constructs missing one or more *pvf* genes were cloned according to the protocol below, with the exception of *pvfACD*. Gene cluster fragments were amplified from *P. entomophila* L48 genomic DNA or pPSV35-*pvfABCD* using primers listed in Table 2. These purified PCR products were digested with restriction enzymes EcoRI and HindIII and then ligated into pPSV35 using corresponding restriction sites. *E. coli* Top10 or *E. coli* DH5α were transformed with the ligation mixtures for cloning. The resulting plasmids, pPSV35-*pvfABC*, pPSV35-*pvfBCD,* pPSV35-*pvfBC,* pPSV35-*pvfCD,* and pPSV35-*pvfC* were isolated using the Thermo GeneJet plasmid miniprep kit and verified by DNA sequencing (Eton Biosciences).

The *pvfACD* construct was cloned using overlap extension PCR using pPSV35-*pvfABCD* as a template. Primers were designed to contain homologous regions overlapping the 3' end of *pvfA* with a 5' portion of *pvfC.* Overlapping PCR products were amplified using primers GM019/GM020 and GM021/GM022, respectively. Reactions were then combined and the full PCR product was amplified using a total of 35 cycles including five two-step cycles before addition of primers GM019/GM022 for the remaining cycles. This purified PCR product was digested with EcoRI and HindIII and then ligated into pPSV35 using corresponding restriction sites. The resulting plasmid, pPSV35-*pvfACD,* was transformed into *E. coli* TOP10 and verified by DNA sequencing.

P. aeruginosa PAO1 was conjugated with *E. coli* RHO3 containing the following vectors to transfer these vectors into *P. aeruginosa* PAO1: pPSV35-*pvfABC,* pPSV35-*pvfBCD*, pPSV35-*pvfBC,* pPSV35-*pvfACD,* pPSV35-*pvfCD,* and pPSV35-*pvfC.* Bacterial growth, dichloromethane extractions, and LC-HRMS analysis was conducted as described above to obtain the metabolic

profile for each culture expressing part of the *pvf* cluster. Three independent cultures were growth for each strain as biological replicates.

P. entomophila and (d)PNOs biological assays

P. entomophila fly infection assay. *Drosophila melanogaster* was challenged with *P. entomophila* strains *via* oral infection. Lennox Luria Broth (4 mL) was inoculated with 10 µL of overnight cultures of *P. entomophila* wildtype or *P. entomophila* Δ*pvfC* + pPSV35. Cultures were incubated at 30 °C for 24 hours. Four hours post inoculation, *P. entomophila* Δ*pvfC* + pPSV35 cultures were supplemented with 50 µM of synthetic compounds or compounds purified from *P. entomophila* Δ*pvfC* + pPSV35-*pvfABCD*. After compound supplementation, cultures were grown for another 24 hours. Bacterial cells were then harvested by centrifugation and resuspended in 150 µL of 1% sucrose to concentrate the cells to an OD_{600} of 100. Concentrated bacterial cells were applied to a filter disk covering the agar surface in a fly culture vial. A total of 30 female adult Oregon R flies (3−5 days old) were starved for 4 hours at 29 °C in empty bottles and then transferred to vials containing bacterial resuspension or 1% sucrose. Flies fed with bacteria or 1% sucrose were incubated at 29 °C and live flies were counted each day over a five-day period. Virulence assays were performed at least three times independently and infection with each bacterial strain was performed in triplicate each time.

Activity against Bacillus subtilis. Samples of 4 mL LB media were inoculated with 50 µL of overnight cultures of *B. subtilis* strain 168 and grown for 24 hours with shaking. These cultures were diluted to an OD_{600} of 0.5 and 100 µL was plated on LB agar plates and allowed to dry. (d)PNOs in methanol or water were applied to sterile filter disks (1 mm) and dried. These filter disks were then placed equidistant from the center of the plate. Plates were incubated at 28 °C until bacteria are grown to confluency. For liquid cultures, 4 mL LB cultures were inoculated with 100 µL of *B. subtilis* strain 168 overnight cultures and incubated for 24 hours.

Stocks of (d)PNOs and water or methanol controls were added as 4 µL of 100 mM stock in methanol or water to reach a final concentration of 100 µM.

      Detection of siderophore production. To test for the presence of iron chelators, a sample of 10µL of Cromazol S (CAS) reagent [80] was incubated with 90µL solution of dPNO at 100 µM, 10 µM, or 0 µM in water at room temperature for 30 minutes. Color change was recorded and imaged.

pH stability of dPNO

Water solutions were prepared at pH 1, 3, 5, 7, 9, 11, and 13 (NaOH and HCl were used to achieve desired pH). A sample of 1 µL of 100 mM dPNO stock was dissolved in 1 mL of each pH solution to reach a final concentration of 100 µM. These samples were incubated for 3.5 hours at room temperature. UV spectra was recorded by a Cary 300 UV-vis spectrophotometer using a 1 cm quartz cuvette.

2.3 Results

Identification of several related metabolites from induced expression of *pvf*

To enhance the production of small molecules by *pvf*-encoded enzymes, we engineered

an inducible expression system for the *pvf* cluster and identified small molecules unique to *pvf*

expressing strains. The expression plasmid containing *pvfABCD* under an IPTG-inducible lac$_{UV5}$

promoter was introduced into a *P. entomophila* mutant from which the NRPS gene *pvfC* has

been deleted (*ΔpvfC*), generating the inducible expression strain of *pvf* (*P. entomophila ΔpvfC* +

pPSV-*pvfABCD*). I adapted the transformation method described by Choi *et al* for generating

genetic knockouts in *Pseudomonas aeruginosa*[71] to introduce plasmids into *Burkholderia* and

Pseudomonas. This method is faster and shows a higher efficiency than biparental mating.

Overnight cultures are washed in 300 mM sucrose to generate electrocompetent cells; the

strains can be transformed directly with plasmids containing various constructs without the need

for an additional helper vector.

The culture supernatants of this *pvf* expression strain and the *P. entomophila ΔpvfC*

control were extracted using dichloromethane. We also conducted extractions with ethyl acetate

and adjusted the pH of the supernatant to 5.0 prior to extraction. The extracts were analyzed

using liquid chromatography-coupled high-resolution mass spectrometry (LC-HRMS). By

comparing the metabolic profiles between samples with or without the gene cluster, including

induced and uninduced sample, wildtype and knockout, or full cluster and empty vector control,

we identified compounds potentially biosynthesized by *pvf*-encoded proteins. We used

programs such as XCMS [73] to expedite the process of filtering peaks for further studies (Table

2.1Table *2.2*). XCMS uses principal component analysis, a statistical analysis that determines

the largest statistical differences between two samples.

Table 2.1. Top features from comparative metabolomics between *pvfC* deletion strain (KO) and complementation with pPSV-*pvfABCD* (OE). Statistical analysis by XCMS, Upregulated features are shown, in order of intensity.

fold	pvalue	mzmed	rtmed	maxint	KO_mean	KO_sd	OE_mean	OE_sd
253.39	9.02E-02	**199.1490**	6.211	535726	2.85E+04	4.37E+03	7.21E+06	1.45E+06
2.78	3.82E-02	**197.1332**	6.889	303291	1.48E+06	4.62E+05	4.12E+06	5.56E+05
9.62	3.63E-02	**181.1385**	8.374	264066	2.92E+05	9.72E+04	2.81E+06	3.02E+05
8.77	4.30E-02	181.1011	8.432	264066	3.06E+05	1.01E+05	2.69E+06	3.23E+05
335.63	1.63E-01	331.1973	9.336	208652	6.25E+03	8.99E+02	2.10E+06	7.77E+05
37.50	1.62E-01	315.1997	9.213	202004	3.93E+04	1.38E+04	1.47E+06	5.30E+05
14.03	1.83E-03	315.1508	9.380	158969	5.58E+04	2.98E+04	7.83E+05	2.32E+04
81.11	1.06E-01	213.1548	7.565	150725	1.65E+04	4.19E+03	1.33E+06	3.15E+05
14.53	1.07E-01	213.1646	7.802	150725	1.69E+05	5.81E+04	2.45E+06	5.65E+05
49.45	3.16E-01	196.1212	6.630	106315	5.91E+03	7.93E+03	2.92E+05	2.19E+05
98.66	1.10E-02	199.1224	6.494	104506	4.21E+03	8.88E+02	4.15E+05	1.06E+04
137.47	3.23E-02	181.1372	4.862	99573	6.15E+03	2.04E+03	8.46E+05	6.07E+04
33.78	8.17E-02	199.1487	6.685	94472	3.93E+04	7.69E+03	1.33E+06	2.36E+05
4319.68	4.87E-02	195.1549	10.327	80436	0.00E+00	0.00E+00	8.20E+05	8.90E+04
30.61	1.77E-01	196.1490	6.537	79556	1.38E+04	5.23E+03	4.22E+05	1.65E+05
672.17	7.55E-02	181.1380	6.905	78020	1.12E+03	9.37E+02	7.55E+05	1.27E+05
740.72	1.32E-01	215.1444	8.190	72539	8.27E+02	4.67E+02	6.13E+05	1.81E+05
1216.72	1.81E-01	165.1423	1.026	71339	1.21E+02	1.71E+02	2.63E+05	1.09E+05
36.03	9.62E-02	181.1383	8.836	70181	2.24E+04	1.10E+04	8.06E+05	1.71E+05
3686.63	8.51E-02	200.1518	6.211	60285	1.22E+02	1.73E+02	8.01E+05	1.52E+05
15.02	1.37E-01	199.1422	6.849	58382	2.16E+04	2.15E+03	3.25E+05	9.35E+04
27.83	2.78E-01	199.1499	10.258	55968	1.53E+04	3.50E+02	4.26E+05	2.71E+05
249.61	1.43E-02	212.1160	6.494	49000	7.99E+02	2.96E+02	1.99E+05	6.40E+03
254.59	4.53E-02	195.1835	9.239	47046	1.56E+03	1.94E+02	3.96E+05	3.98E+04
22.99	5.99E-02	199.1492	8.723	46747	1.22E+05	6.92E+03	2.80E+06	3.58E+05
37.75	7.46E-02	199.1387	7.356	46747	7.35E+03	1.59E+03	2.78E+05	4.52E+04
28.81	2.39E-01	212.1442	6.432	43049	1.11E+04	2.45E+03	3.21E+05	1.73E+05
3.93	1.53E-02	211.1495	8.862	42801	8.14E+04	3.01E+04	3.20E+05	2.27E+04
216.52	1.03E-01	181.1374	1.026	41588	6.73E+02	9.51E+02	1.66E+05	3.81E+04
2.12	1.48E-01	100.0787	2.205	38933	2.41E+05	1.51E+04	5.11E+05	9.49E+04
24.78	7.57E-02	332.1496	9.419	37809	1.12E+04	9.87E+02	2.78E+05	4.52E+04
91.36	1.62E-01	332.2002	9.336	37809	4.20E+03	4.72E+02	3.84E+05	1.39E+05
6.88	4.25E-01	245.1714	9.189	36676	9.23E+03	2.32E+03	6.35E+04	6.05E+04
1297.71	1.39E-01	165.1420	4.604	36604	9.79E+01	1.38E+02	2.50E+05	7.86E+04

Table 2.2. Top features from comparative metabolomics between *P. aeruginosa* PAO1 heterologous expression pPSV-*pvfABCD* (HE) and empty vector control pPSv35 (EVC). Statistical analysis by XCMS. Upregulated features are shown, in order of intensity.

fold	pvalue	mzmed	rtmed	maxint	EVC_mean	EVC_sd	HE_mean	HE_sd
6.4	2.18E-01	**181.1375**	8.374	1281609	1.38E+06	5.86E+05	8.89E+06	7.36E+06
7.6	6.52E-02	181.1010	8.417	1281609	1.43E+06	6.20E+05	1.08E+07	4.47E+06
133.4	2.19E-03	**199.1474**	6.292	960096	9.58E+04	5.99E+04	1.28E+07	1.04E+06
122.8	1.66E-02	165.1417	1.026	714518	2.58E+04	4.52E+03	3.16E+06	7.10E+05
292.0	2.77E-02	181.1369	4.925	579061	1.66E+04	1.38E+04	4.86E+06	1.42E+06
286.0	8.29E-02	165.1406	4.618	495854	1.01E+04	9.70E+03	2.90E+06	1.54E+06
54.6	5.66E-02	181.1368	1.026	369708	2.21E+04	1.42E+04	1.21E+06	5.12E+05
64.5	3.43E-02	199.1487	10.426	225120	2.20E+04	2.54E+04	1.42E+06	4.62E+05
85.6	1.68E-01	181.1349	6.797	204031	1.21E+04	2.23E+03	1.03E+06	8.34E+05
8.0	7.43E-02	182.1039	8.416	145784	1.44E+05	6.08E+04	1.16E+06	5.16E+05
9.9	4.13E-02	182.1404	8.364	145784	1.47E+05	5.87E+04	1.45E+06	4.84E+05
51.5	6.81E-03	181.1375	9.003	130608	3.82E+04	1.41E+04	1.97E+06	2.79E+05
227.1	2.01E-03	200.1509	6.297	112666	6.39E+03	5.09E+03	1.45E+06	1.13E+05
44.3	1.68E-04	183.1525	6.274	105751	2.49E+04	2.77E+04	1.10E+06	6.10E+04
77.1	2.10E-02	181.1076	9.053	101186	9.01E+03	3.42E+03	6.95E+05	1.75E+05
1181.4	2.11E-02	164.1340	8.364	99943	6.68E+02	6.15E+02	8.58E+05	2.20E+05
93.7	1.51E-01	149.1101	8.374	93293	7.10E+03	3.75E+03	6.65E+05	5.02E+05
9.0	4.07E-01	196.1196	6.704	89561	1.25E+04	5.69E+03	1.13E+05	1.67E+05
117.7	1.43E-02	166.1446	1.026	77191	2.99E+03	6.14E+02	3.51E+05	7.30E+04
2435.3	4.71E-02	361.2663	4.925	74733	0.00E+00	0.00E+00	4.26E+05	1.66E+05
41.0	6.12E-02	167.1212	7.568	67395	8.65E+03	3.28E+03	3.55E+05	1.56E+05
33.1	9.91E-02	181.2598	8.374	64097	1.29E+04	5.01E+03	4.26E+05	2.44E+05
474.9	2.57E-02	182.1396	4.925	64025	1.10E+03	9.72E+02	5.52E+05	1.56E+05
137.2	2.38E-02	149.1098	1.026	62609	1.68E+03	7.20E+02	2.30E+05	6.22E+04
27.1	7.61E-03	196.1480	6.537	62399	1.07E+04	3.72E+03	2.89E+05	4.33E+04
73.2	8.35E-02	166.1436	4.618	56778	4.50E+03	2.56E+03	3.30E+05	1.74E+05
2361.9	7.12E-02	167.1202	4.578	56230	6.84E+01	1.18E+02	4.37E+05	2.14E+05
775.2	6.81E-02	361.2665	1.010	55882	0.00E+00	0.00E+00	1.36E+05	6.47E+04
11.0	8.14E-02	235.1235	5.033	54787	2.16E+04	4.57E+03	2.37E+05	1.14E+05
14.9	2.25E-01	181.2965	8.542	53749	1.34E+04	6.29E+03	1.99E+05	1.85E+05
474.2	4.53E-02	427.2454	6.737	49168	6.19E+02	5.39E+02	3.21E+05	1.22E+05
135.0	1.85E-02	427.2451	6.114	49086	2.67E+03	2.28E+03	3.61E+05	8.59E+04
154.6	1.01E-01	163.1259	6.944	48536	1.43E+03	9.51E+02	2.21E+05	1.31E+05
883.9	3.19E-02	149.1095	4.914	48348	3.34E+02	5.78E+02	3.98E+05	1.26E+05
2644.5	1.03E-03	199.2749	6.265	43641	7.31E+01	1.27E+02	5.02E+05	2.80E+04
37.8	5.43E-02	182.1396	1.026	43502	3.80E+03	1.39E+03	1.44E+05	5.89E+04
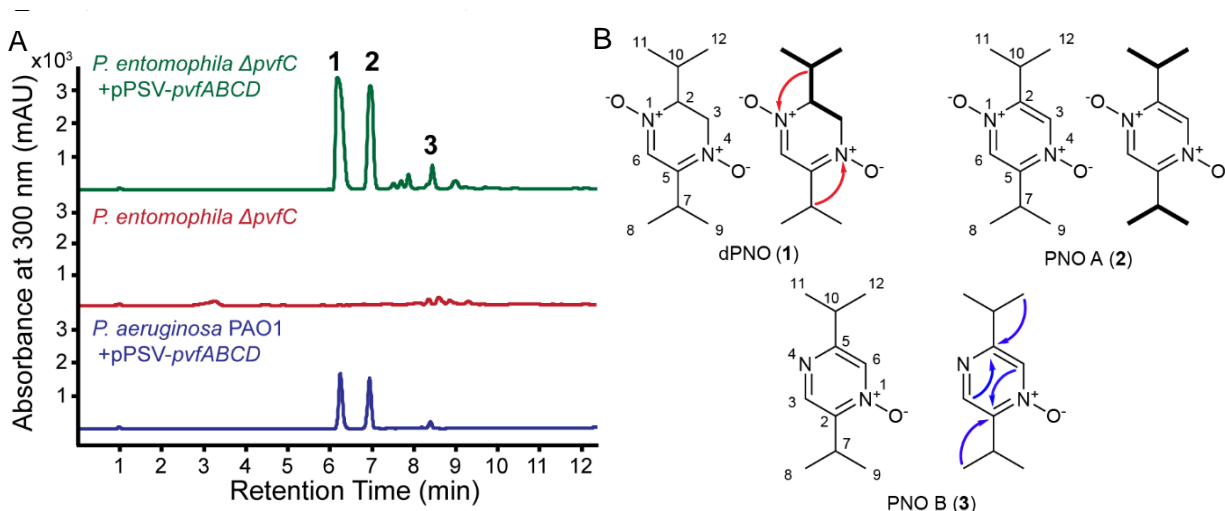
Figure 2.2. (A) Identification of new metabolites produced by *pvf*-encoded enzymes using comparative metabolomics. Overexpression of *pvfABCD* on a pPSV35 vector in *P. entomophila ΔpvfC* (green) led to the production of three new metabolites **1−3**, which are absent from *P. entomophila ΔpvfC* (red). These metabolites were also identified when *pvfABCD* was heterologously expressed in *P. aeruginosa PAO1* (blue). (B) Structural characterization of dPNO (**1**), PNO A (**2**), and PNO B (**3**). COSY correlations are indicated by bonds in bold, and key (^{1}H, ^{13}C)-HMBC and (^{1}H, ^{15}N)-HMBC correlations are indicated by blue and red arrows, respectively.

Three major metabolites are present at high levels in the extract of *ΔpvfC* + pPSV-*pvfABCD*, but are absent in the extract of the *ΔpvfC* control: **1** ([M + H]$^+$, 199.145 *m/z*, $C_{10}H_{18}N_2O_2$), **2** ([M + H]$^+$, 197.126 *m/z*, $C_{10}H_{16}N_2O_2$), and **3** ([M + H]$^+$, 181.135 *m/z*, $C_{10}H_{16}N_2O$) (Figure 2.2Figure *2.4*). The *pvf* plasmid was also introduced into *Pseudomonas aeruginosa* PAO1, which does not natively harbor the *pvf* cluster, yielding a heterologous *pvf* expression strain (PAO1 + pPSV*pvfABCD*). Metabolites 1−3 were also detected in the organic extract of PAO1 + pPSV-*pvfABCD*, but not in the extract of the control strain carrying the empty expression vector (Figure 2.2Figure *2.4*). These results indicate that 1, 2, and 3 are synthesized by the *pvf*-encoded enzymes.

We isolated **1**, **2**, and **3** from 8 L cultures of *Δpvf C* + pPSV*pvfABCD* for structure elucidation. The major metabolites were isolated at a yield of 5 mg/L for **1** and 1−2 mg/L for **2** and **3**. LC-MS analysis of these purifications found that compounds 1−3 were isolated as a pure yield, in suitable quantities for NMR. The structures of these metabolites were analyzed and characterized using a combination of NMR, UV, and IR spectroscopies (Figure 2.2Figure *2.7*,

Table 2.3Table *2.5*, Appendix 1−17). For characterization by ^{15}N NMR, we extracted cultures of

the overexpression strain grown in M9 media supplemented with N[15] ammounium chloride,

N[15]H$_4$Cl. We verified that isolated compounds contained two N[15] atoms by LC-MS (mass shift

+2, Figure 2.8), and conducted further characterization by 2D NMR experiments ([1]H-[15]N

HMBC).
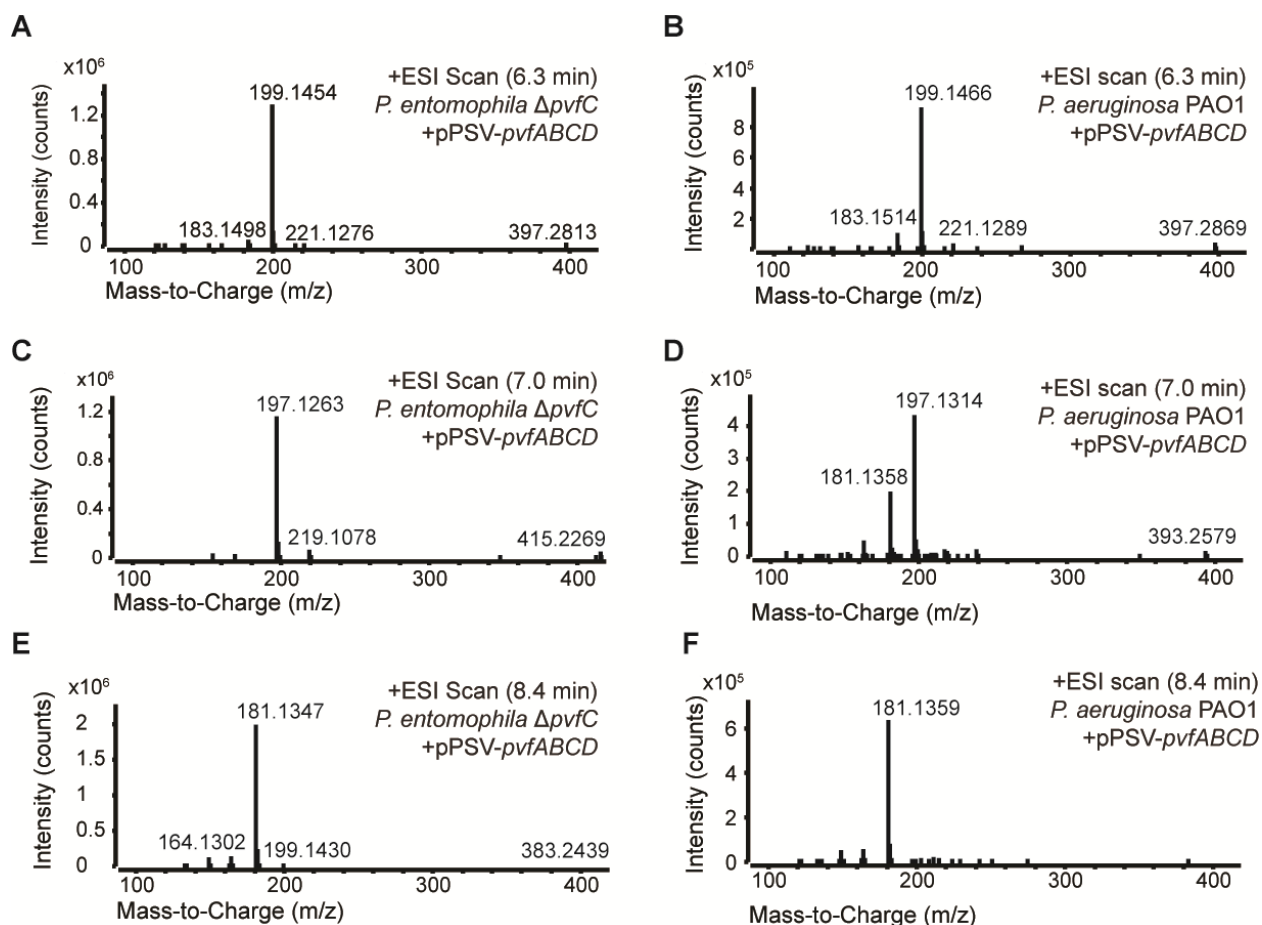
Figure 2.3. High-resolution mass spectra of metabolites **1, 2,** and **3** identified from *P. entomophila ΔpvfC + pPSV-pvfABCD* (Panel A, C, E) and *P. aeruginosa* PAO1 + *pPSV-pvfABCD* (Panels B, D, F). (A), (B) Mass spectrum of dPNO (**1**) at retention time 6.3 min. [M + H]$^+$ calculated for $C_{10}H_{18}N_2O_2$, m/z =199.1441, observed 199.1454. (C), (D) Mass spectrum of PNO A (**2**) at retention time 7.0 min. [M + H]$^+$ calculated for $C_{10}H_{16}N_2O_2$, m/z =197.1245, observed 197.1263. (E), (F) Mass spectrum for PNO B (**3**) at retention time 8.4 min. [M + H]$^+$ calculated for $C_{10}H_{16}N_2O$, m/z =181.1335, observed 181.1347.

Table 2.3. NMR assignments and correlations for dPNO (**1**).

POSITION	Δ$_C$	PROTON	Δ$_H$ (MULTIPLICITY, J[HZ])	COSY	(^{1}H, ^{13}C) HMBC	Δ$_N$	(^{1}H, ^{15}N) HMBC
1-N						−94 (286)	
2	72.0	CH	3.89 (td, [6.5, 4.4])	3a, 3b	6, 10, 11, 12		1, 4
3A, B	57.7	CH$_2$	4.49 (dd, [15.9, 6.8]), 4.16 (dd, [15.9, 4.4])	2, 3b, 3a	2, 5, 6, 10		1, 4
4-N						−121 (259)	
5	145.2						
6	126.8	CH	7.39 (s)		2, 5, 7		1, 4
7	26.4	CH(CH$_3$)$_2$	3.42 (hept, [7.0])	8, 9	5, 6, 8, 9		4
8, 9	18.4, 18.7	CH$_3$	1.15 (d, [7.0])	7	5, 7, 8, 9		
10	28.9	CH(CH$_3$)$_2$	2.48 (m, [13.5, 6.8])	11, 12	2, 3, 11, 12		1
11	19.0	CH$_3$	1.09 (d, [7.1])	10	2, 10, 12		
12	17.7	CH$_3$	1.03 (d, [6.8])	10	2, 11, 10		

Table 2.4. NMR assignments of biologically isolated and synthetic PNO A (**2**).

	SYNTHETIC			BIOLOGICAL		
	δ$_C$	proton	δ$_H$ (multiplicity, J[Hz])	δ$_C$	δ$_H$ (multiplicity, J[Hz])	COSY
1-N, 4-N						
2, 5	152.6			152.8		
3, 6	133.9	C-H	8.03 (s)	134.0	8.09 (s)	
7, 10	26.7	CH(CH$_3$)$_2$	3.54 (hept, [6.8])	26.8	3.55−3.52 (m)	7-8,9, 10-11,12
8, 9, 11, 12	19.6	CH$_3$	1.30 (d, [6.8])	19.6	1.30 (d, [6.9])	8,9-7, 11,12-10

Table 2.5. NMR assignments of biologically isolated and synthetic PNO B (**3**). The two nitrogen shifts were not individually assigned because they are unresolved on the (^{1}H, ^{15}N)-HMBC spectrum. A possible alternative assignment is δ 8.19 for 3-H and δ 8.19 for 6-H.

	SYNTHETIC			BIOLOGICAL					
	$δ_C$	proton	$δ_H$ (multiplicity, J [Hz])	$δ_C$	$δ_H$ (multiplicity, J [Hz])	COSY	(^{1}H, ^{13}C) HMBC	$δ_N$	(^{1}H, ^{15}N) HMBC
1-N								~300	
2	163.4			165.0					
3	144.5	C-H	8.49 (s)	145.9	8.49 (s)		2, 5		1, 4
4-N								~300	
5	149.3			150.7					
6	131.0	C-H	8.19 (s)	132.4	8.19 (s)		2		1, 4
7	33.4	CH(CH$_3$)$_2$	3.03 (hept, [7.0])	34.9	3.03 (hept, [6.9])	8, 9			1, 4
8, 9	20.6	CH$_3$	1.31 (d, [7.0])	22.0	1.30 (d, [6.9])	7	2, 7, 8, 9		
10	26.1	CH(CH$_3$)$_2$	3.57 (hept, [7.0])	27.5	3.56 (hept, [7.2])	11, 12	11, 12		1, 4
11, 12	18.4	CH$_3$	1.34 (d, [6.9])	19.8	1.34 (d, [7.0])	10	5, 10, 11, 12		
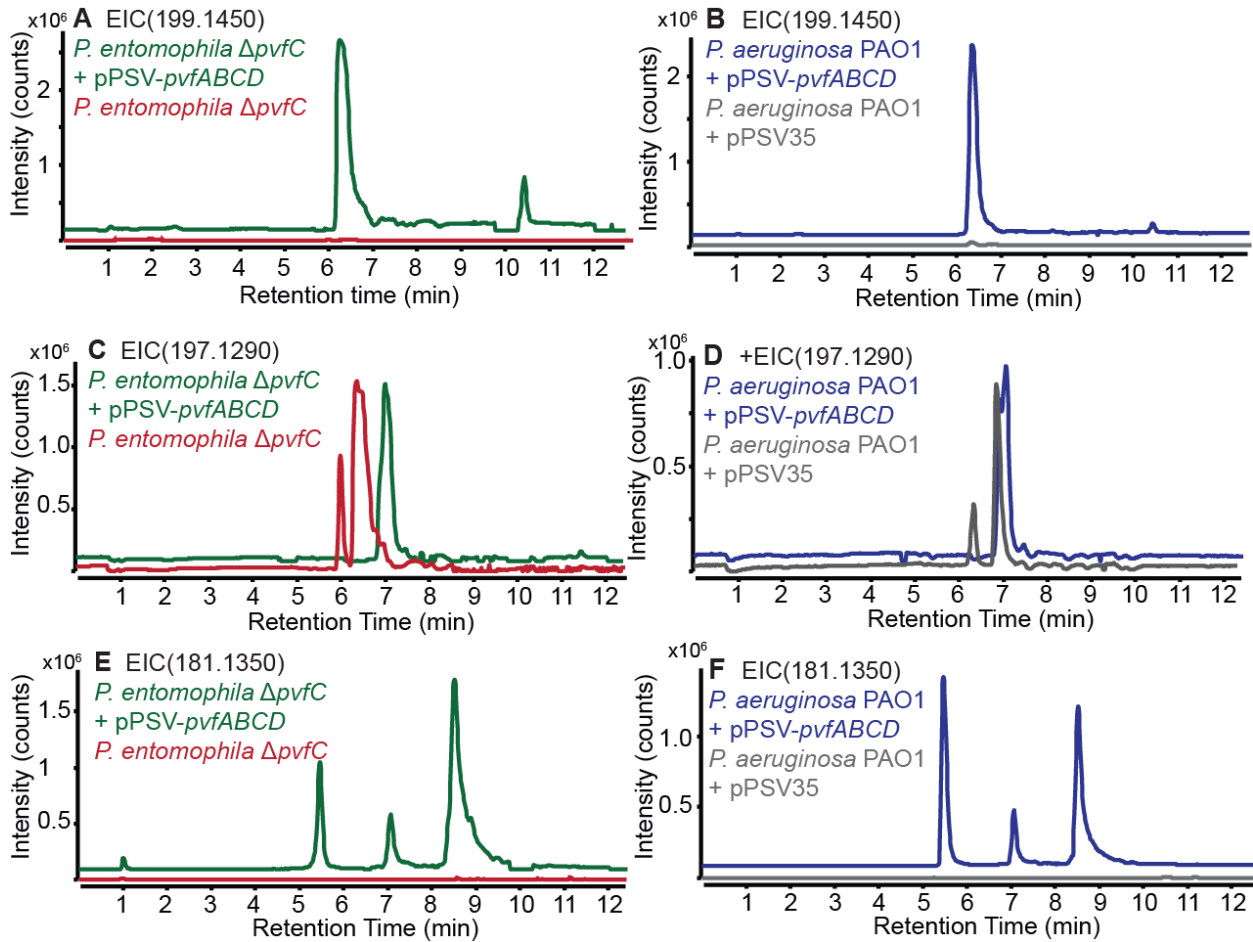
Figure 2.4. Extracted ion chromatograms (EIC) for **1, 2,** and **3** identified from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* (Panel A, C, E, green compared to *P. entomophila ΔpvfC,* red) and *P. aeruginosa* PAO1 + pPSV-*pvfABCD* (Panels B, D, F, blue compared to *P. aeruginosa* PAO1 + pPSV35, grey). (A), (B) EIC of *m/z* = 199.1450, dPNO (**1**) (C), (D) EIC of *m/z* = 197.1290, PNO A (**2**) (E), (F) EIC of *m/z* = 181.1350, PNO B (**3,** peak at 8.7 min). A diketopiperazine with identical *m/z* (197.1290) and similar retention time is present in *pvf* and non-*pvf* expressing cultures (C&D).
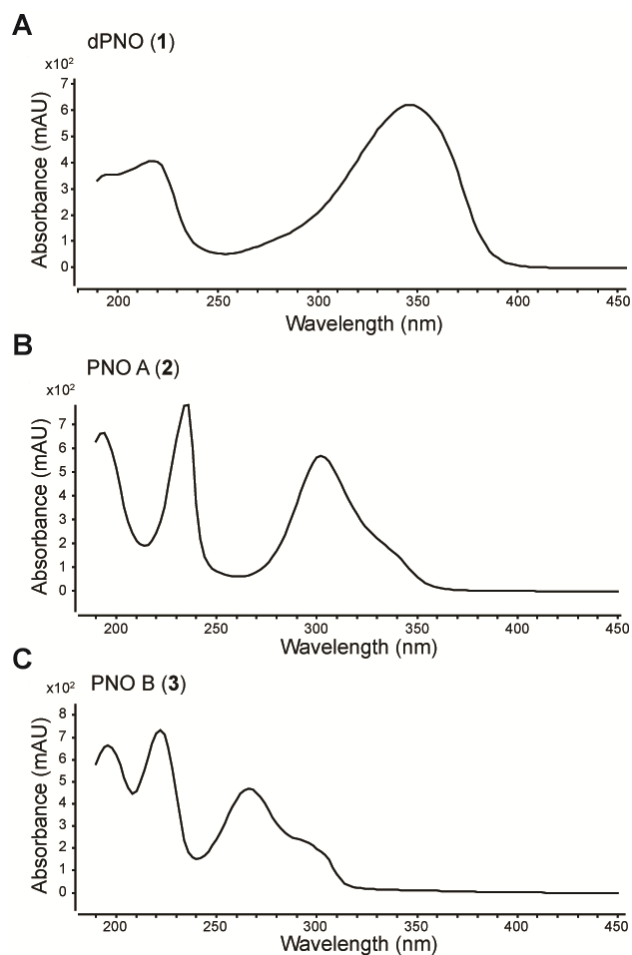
Figure 2.5. UV spectra of metabolites dPNO (**1**), PNO A (**2**), and PNO B (**3**), from extractions of *P. entomophila ΔpvfC* + pPSV-*pvfABCD.* (A) dPNO (**1**), λ_{max} = 219, 348 nm. (B) PNO A (**2**), λ_{max} = 195, 235, 303, 338 nm. (C) PNO B (**3**), λ_{max} = 198, 223, 265, 303 nm.
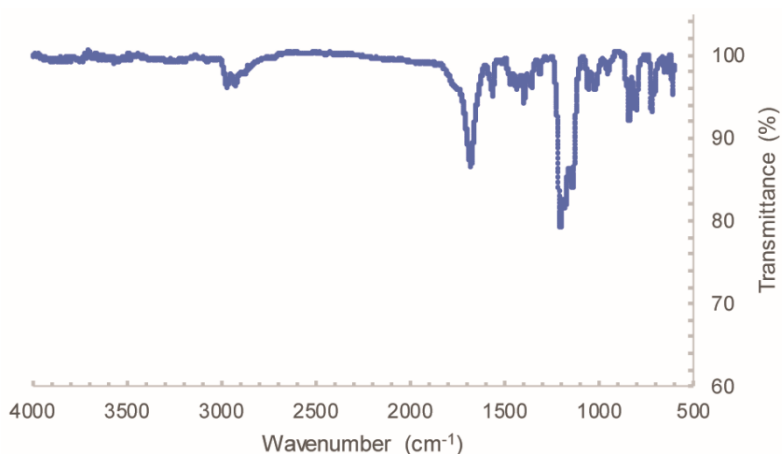
Figure 2.6. IR spectrum of dPNO (**1**). Peaks shown at 2968, 2922, 1679, 1564, 1435, 1398, 1204, 1140, 1056, and 842 cm^{-1}
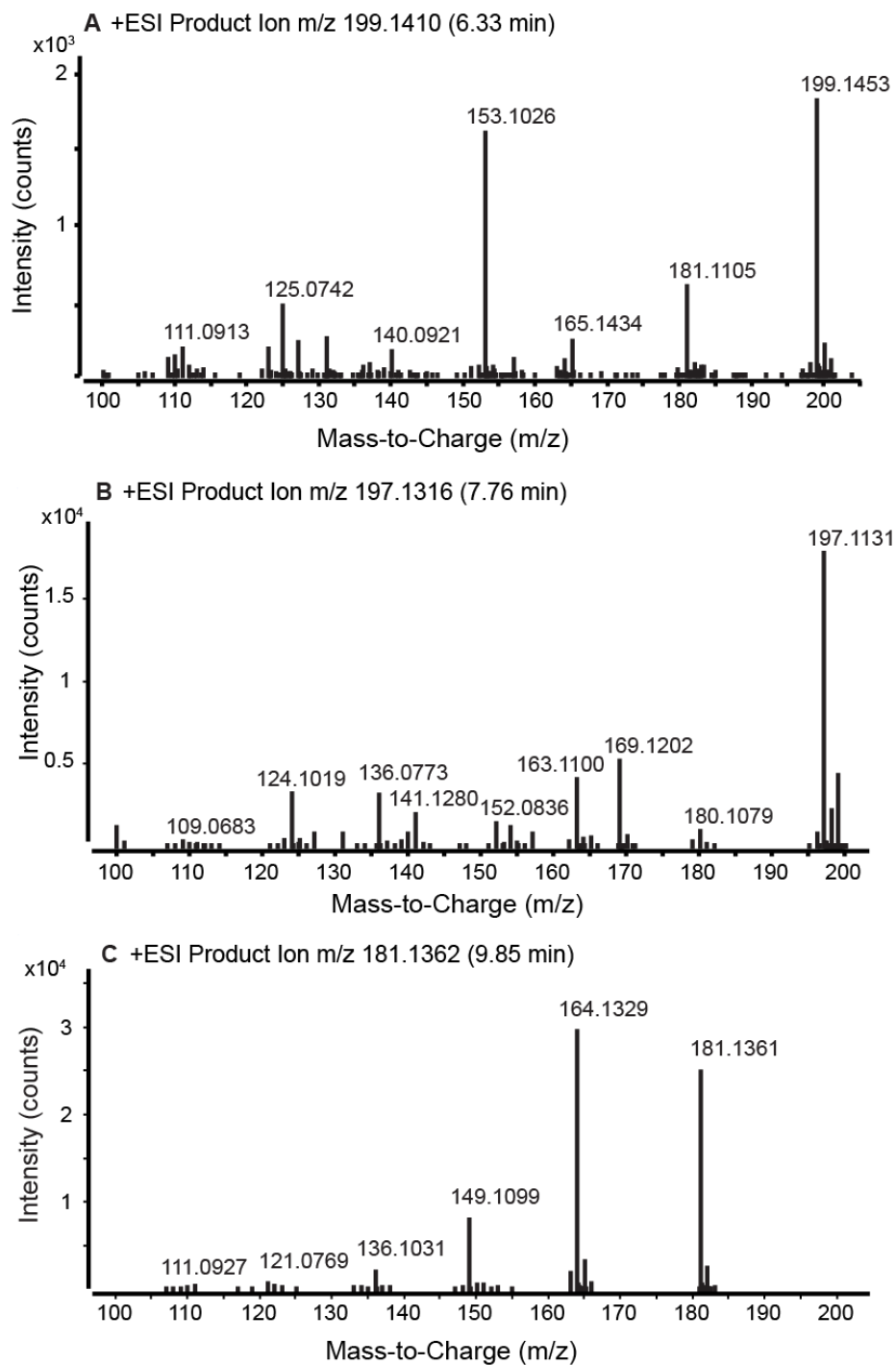
Figure 2.7. Tandem mass spectra metabolites dPNO (**1**), PNO A (**2**), and PNO B (**3**), from cell-free culture extractions of *P. entomophila ΔpvfC* + pPSV-*pvfABCD.*
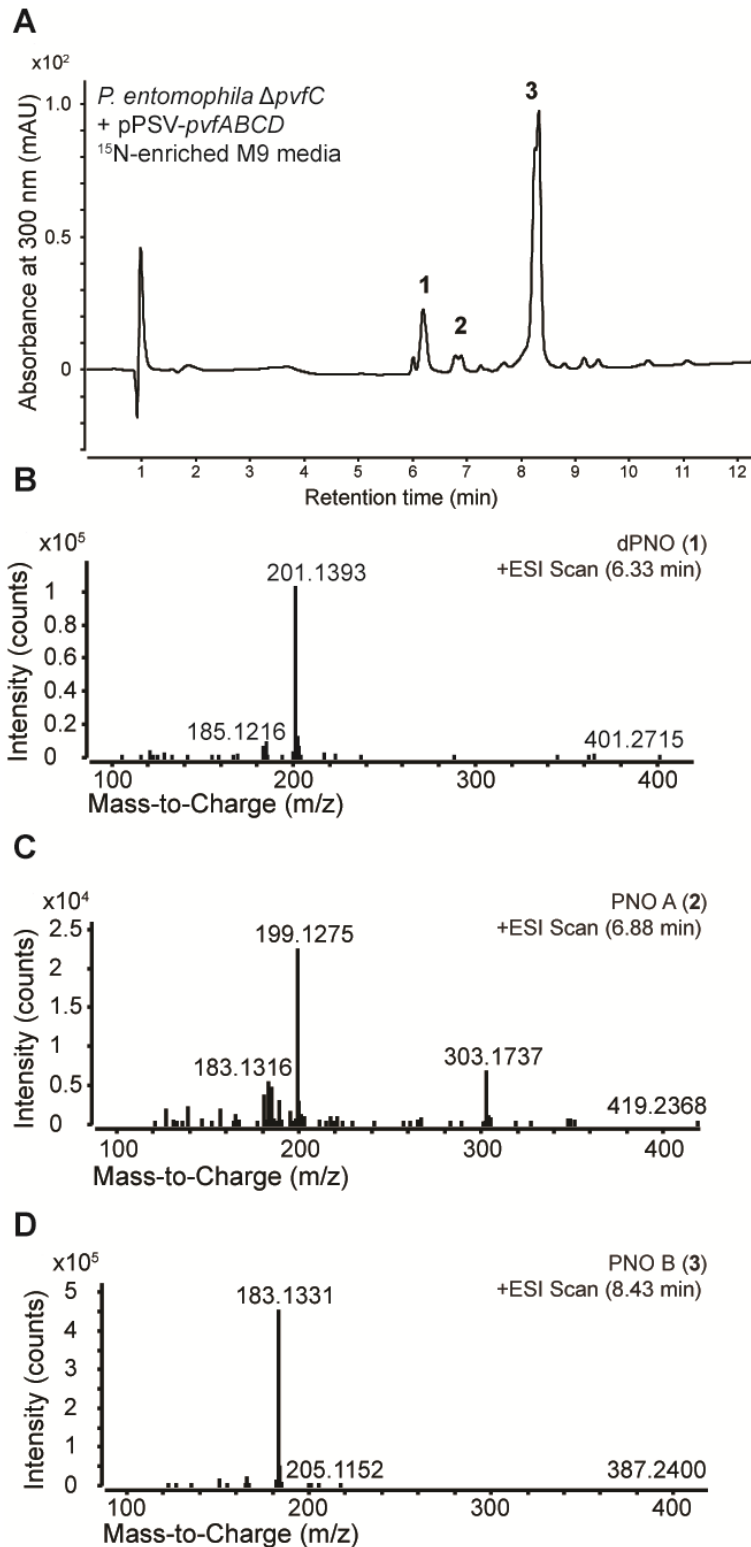
Figure 2.8. Mass spectra of ^{15}N-enriched (d)PNOs. (A) LC trace at 300 nm of ethyl acetate extracts of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* cultures grown in $^{15}NH_4Cl$-enriched M9 media. The metabolic profile of cultures grown in minimal media shows different relative abundances of metabolites **1**, **2** and **3** compared to those grown in Lennox LB. Each (d)PNO metabolite exhibits a mass increase of 2 Da, indicating the incorporation of two nitrogens. (B) dPNO mass spectrum at 6.33 min, *m/z* =201.1393. (C) PNO A mass spectrum at 6.88 min, *m/z* =199.1275. (D) PNO B mass spectrum at 8.43 min, *m/z* =183.1331.

Structural elucidation of (d)PNOs

The ^{1}H spectrum in conjunction with the molecular formula indicates that **2** is symmetric. The COSY correlations between six methyl protons and a single multiplet proton indicate the presence of an isopropyl group (Figure 2.2). The shift at δH 8.09 and UV spectra suggest a heterocyclic pyrazine core (Figure 2.5).[81] The oxygens are likely attached to the pyrazine as *N*-oxides. Together, the structure of 2 was assigned as 2,5-diisopropylpyrazine *N,N'*-dioxide (PNO A).

The chemical formula of **3** contains one less oxygen than **2**, and the NMR spectrum suggests that **3** is asymmetric. The ^{1}H shifts of **3** are largely similar to those of **2**, suggesting **3** possesses a related structure but contains a single *N*-oxide. The COSY, (^{1}H, ^{13}C)-HMBC, and (^{1}H, ^{15}N)-HMBC NMR correlations support the pyrazine core and were used to assign the isopropyl groups to the 2,5 positions of the pyrazine (Figure 2.2, Table 2.5). The structure of **3** was assigned as 2,5-diisopropylpyrazine *N*-oxide (PNO B). The ^{1}H and ^{13}C shifts of PNO A and PNO B are consistent with the shifts reported for these compounds, which have been previously synthesized.[82-85]

The chemical formula indicates that **1** has one less degree of unsaturation than **2** and **3**, which may be due to hydration within the pyrazine ring. The ^{1}H NMR spectrum of **1** contains an additional multiplet proton at δ3.89 and two diastereotopic protons at δ4.16 and δ4.49 (Table 2.3). The connectivity of these atoms was determined based on COSY, (^{1}H, ^{13}C)-HSQC, (^{1}H, ^{13}C)-HMBC, and (^{1}H, ^{1}H)-TOSCY NMR correlations. These correlations and proton splitting patterns support the presence of a 2,3-dihydropyrazine core, and the (^{1}H,[15]N)-HMBC correlations indicate the positions of 2,5-diisopropyl groups (Table 2.3, Figure 2.2). Thus, the structure of **1** was assigned as 2,5-diisopropyl-2,3-dihydropyrazine *N,N'*-dioxide (dPNO).

The structure of **1** contains two conjugated nitrones and is highly unusual. We conducted UV absorbance and IR spectroscopy to further characterize this molecule. We found that **1**

exhibits two UV absorbance maxima at 219 and 348 nm (Figure 2.5) and distinctive IR peaks at 1564, 1435, and 1398 cm^{-1} (Figure 2.6), which is consistent with the reported UV maxima in the 345−370 nm range and strong IR bands at 1530−1570 cm^{-1} and 1440−1470 cm^{-1} for synthetic dinitrone compounds.[86-88] These results further support the proposed structure for 1. The stereochemistry at C2 remains to be determined.

We synthesized **2** and **3** to confirm the proposed structures and obtain materials for biological studies (Scheme 2.1). Products **2** and **3** were synthesized starting with L-valinol to yield a 2,5-diisopropylpyrazine intermediate that was oxidized to yield **2** and **3**. The ^{1}H and ^{13}C NMR and UV spectra of synthetic **2** and **3** are consistent with previous reports,[83-85] and are nearly identical to the spectra for **2** and **3** isolated from *P. entomophila* (Table 2.4Table *2.5*, Appendix 18−21), confirming the structural assignments.
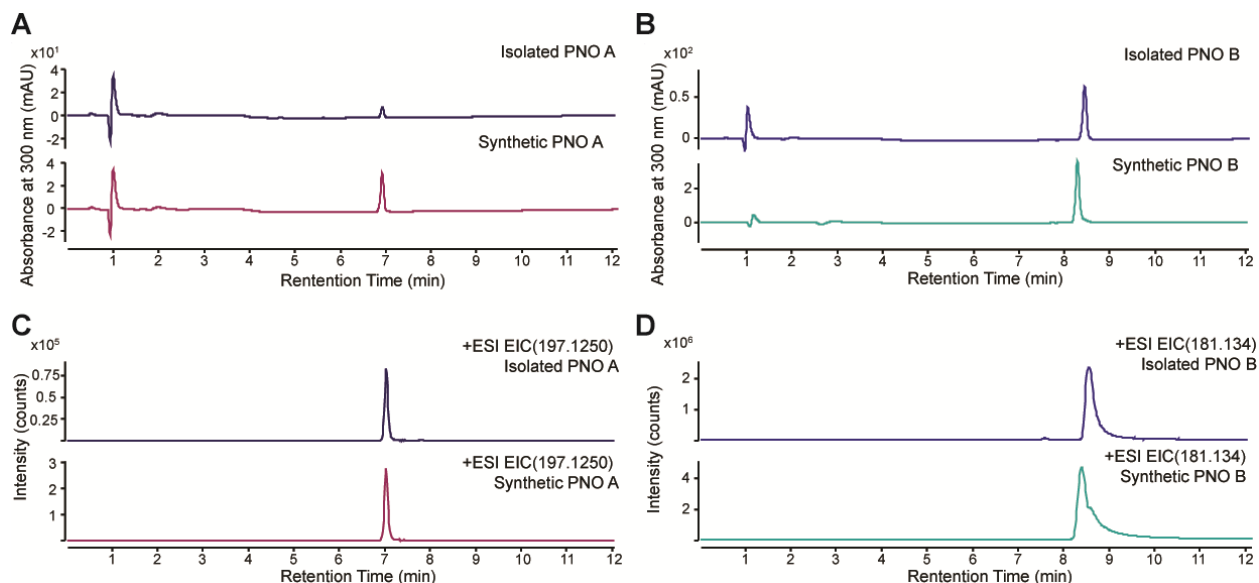
Figure 2.9. Metabolites **2** and **3** isolated from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* (blue) exhibit identical retention times to PNO A and PNO B synthesized from valinol (pink and green). LC-HRMS analysis of biologically isolated and synthetic samples of PNO A (A) UV absorbance at 300 nm, (B) Extracted ion chromatogram at the *m/z* of 197.1250, and PNO B (C) UV absorbance at 300 nm, (D) Extracted ion chromatogram at the *m/z* of 181.134.

Biosynthesis

We conducted biochemical and genetic studies to characterize the biosynthesis of (d)PNOs. Overexpression of different combinations of *pvf* genes revealed that *pvfB* and *pvfC* are necessary and sufficient for the production of (d)PNOs (Figure 2.10). As an undergraduate project for Brigh Turner and Evan Xu, the genes encoding the individual Pvf proteins were amplified from *P. entomophila* L48 genomic DNA and cloned into the pLIC-His vector. All but the heme oxygenase (PvfA) were cloned by this method. Homologous *pvf* genes from *B. cenocepacia* HI2424 and *P. fluorescens* Pf01, were also cloned into the pLIC-His vectors. *E. coli* BL21 cells were transformed with plasmids containing individual *pvf* genes for protein purification.
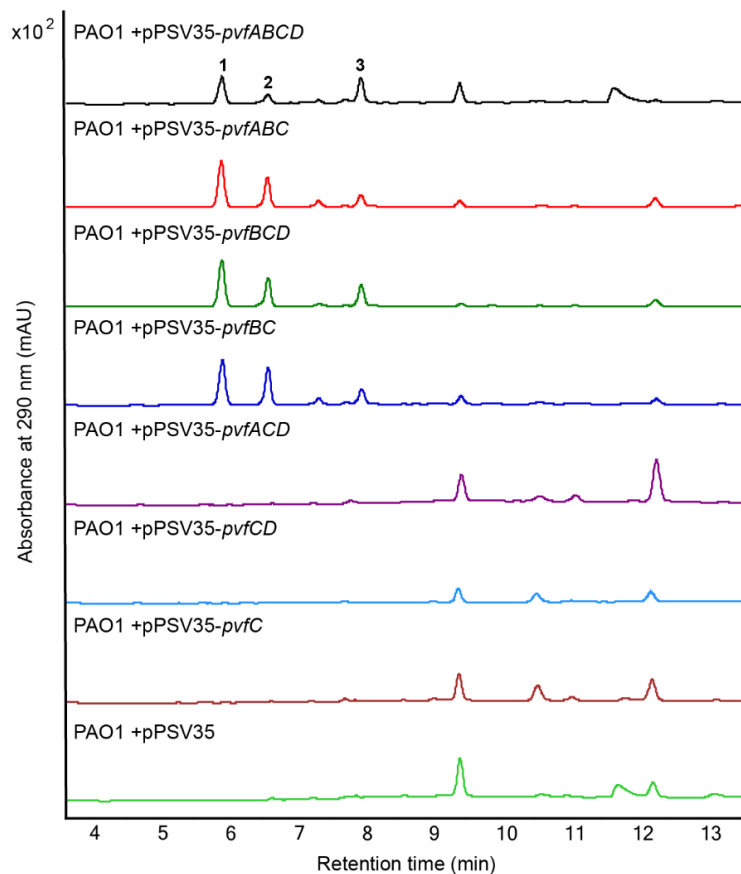
Figure 2.10. The *pvfB* and *pvfC* genes are both necessary and sufficient for the production of (d)PNOs. Liquid chromatography traces at 290 nm are shown for the culture extracts of *P. aeruginosa* PAO1 overexpressing different, partial *pvf* constructs.

The NRPS PvfC contains three domains: adenylation (A), thiolation (T), and reductase (R) (Figure 2.11). The adenylation activity of PvfC was investigated *in vitro* using a ATP-[32P]PPi exchange assay. PvfC selectively activates L-Val among the 20 proteinogenic amino acids (K_M = 3.4 ± 0.7 mM, k_{cat} = 3.4 ± 0.3 s^{-1}, Figure 2.11). The observed K_M is 2−3-fold higher than that of typical NRPS adenylation domains,[89] but is generally consistent with the mM cellular concentration of L-Val in bacteria.[90] Previously, other adenylation domains from the *pvf* pathways in other strains have preferred leucine, specifically from *P. fluorescens* Pf01.
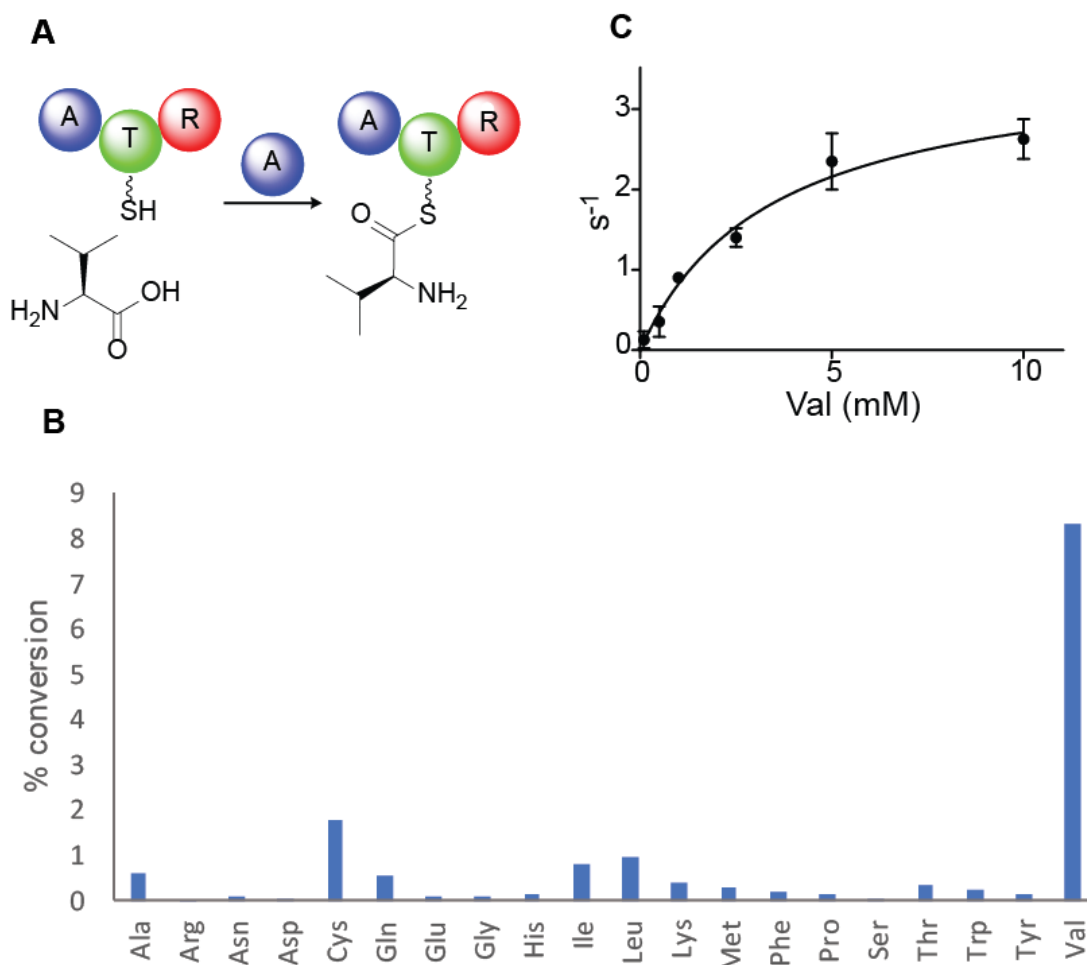
Figure 2.11. PvfC activates and incorporates L-Val into dPNO (**1**), PNO A (**2**), and PNO B (**3**). (A) The adenylation domain of PvfC activates and loads L-Val onto the thiolation domain. (B) Substrate specificity of PvfC towards the 20 proteinogenic L-amino acids. (C) Kinetic measurements of PvfC with L-Val. K_M = 3.4 ± 0.7 mM, k_{cat} = 3.4 ± 0.3 s^{-1}. Domain abbreviations: adenylation (A), thiolation (T), and reductase (R).
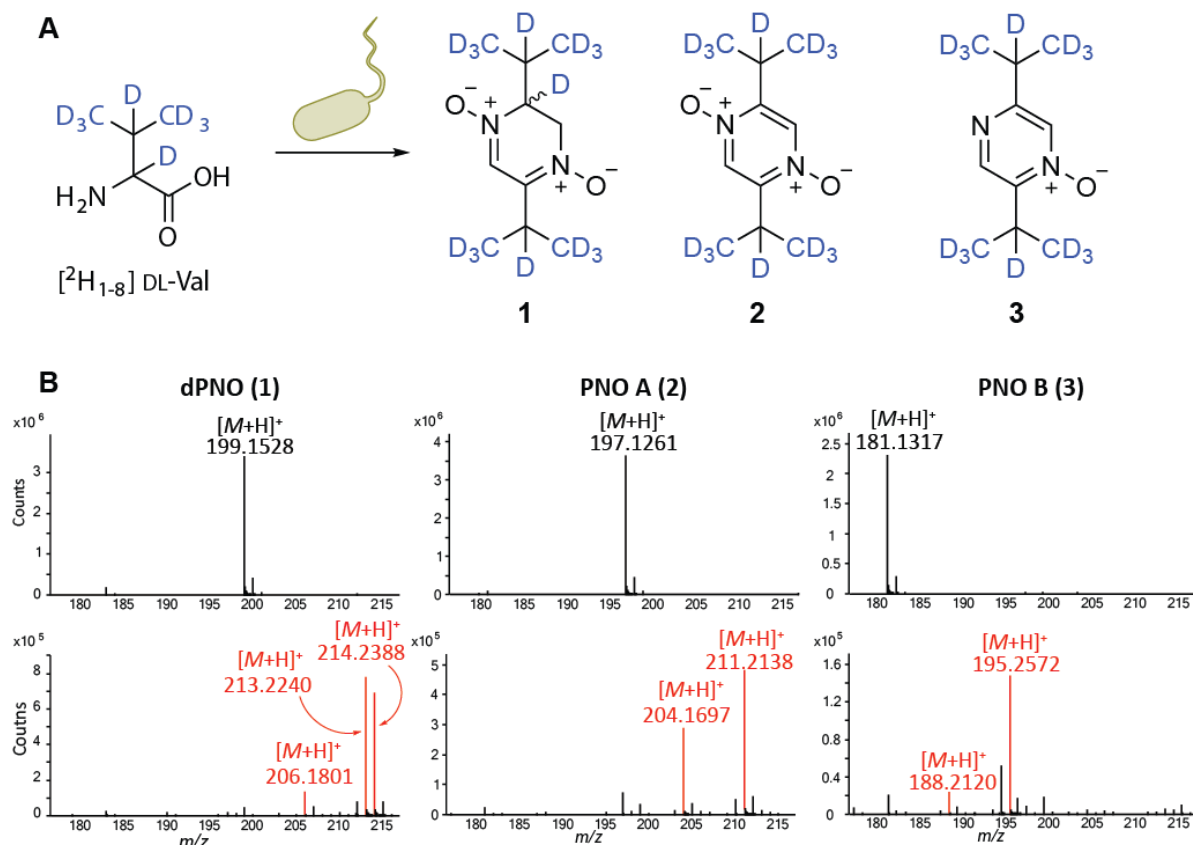
Figure 2.12. Incorporation of $[^2H_{1-8}]_{DL}$-valine into (d)PNOs based on LC/HRMS analysis. Black traces correspond to unlabeled (d)PNOs. Red traces correspond to mass shifts observed for (d)PNOs with $[^2H_{1-8}]_{DL}$-valine supplementation.

The incorporation of Val was also investigated in cells using isotopically labeled Val ($[^2H_{1-8}]_{DL}$Val). A mass increase of +7 Da and +14 Da (and +15 Da for **1**) was observed for **1**, **2**, and **3** (Figure 2.12), when the overexpression cultures were supplemented with $[^2H_{1-8}]_{DL}$Val, indicating the incorporation of one or two $[^2H_{1-8}]_{DL}$Val. This result suggests that L-Val is the physiological substrate for PvfC and is likely the precursor for (d)PNOs.

A PvfB homologue from *P. syringae* pv. *phaseolicola* was previously reported to convert *p*-aminobenzoic acid to *p*-nitrobenzoic acid.[91] However, neither (d)PNOs nor valdiazen contains a *p*-nitrobenzoic acid or aryl nitro products. Thus, *p*-aminobenzoic acid may not be the physiological substrate for PvfB. The structures of (d)PNOs and valdiazen suggest that a *N*-hydroxy-L-Val derivative exists as a precursor to both sets of molecules. PvfB may install the hydroxylamine on PvfC-linked L-Val instead.
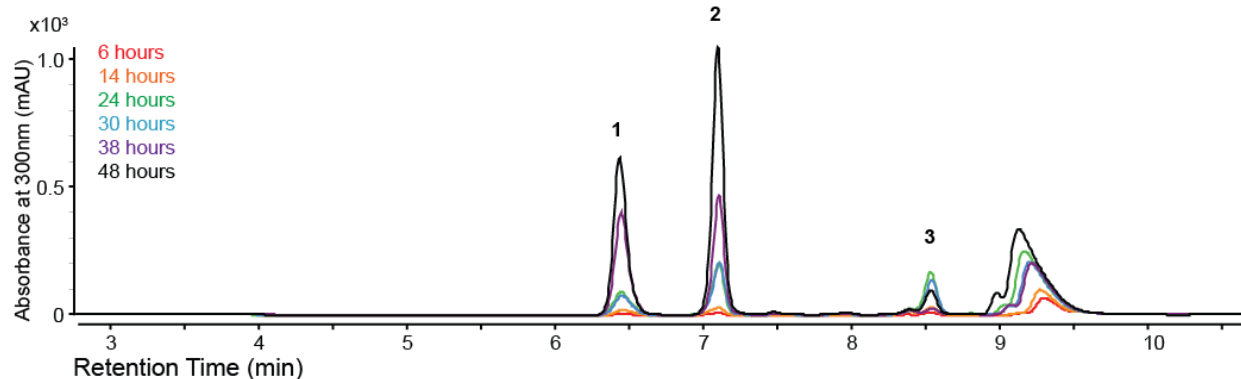
Figure 2.13. Time points suggest PNO B is produced first, followed by PNO A and dPNO. Liquid chromatography traces at 300 nm are shown for cell-free culture extractions of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* at 6 (red), 14 (orange), 24 (green), 30 (blue), 38 (purple) and 48 (black) hours.

We conducted additional metabolomics experiments to understand the timing of the production of these (d)PNOs. We took time point samples for metabolite extraction after IPTG induction in the overexpression strain. We analyzed the production of compounds 1−**3** over the course of 48 hours post inoculation. We found that the optimal production for the dioxides was at 38−48 hours and 24−30 hours for the monoxide, suggesting PNO B might be synthesized first and therefore intermediates of dPNO and/or PNO A (Figure 2.13). However, we also noted an increase in production of PNO B at 48 hours compared to 38 hours. This could be due to conversion of dPNO and/or PNO A to PNO B. We also observe that dPNO is converted to PNO B after several rounds of purification, or if samples were left in acetonitrile at room temperature for extended periods of time (Figure 2.15).

We propose a biosynthetic pathway for (d)PNOs (Figure 2.14). First, PvfC activates and loads L-Val onto the T domain. The amine of the tethered L-Val is oxygenated by PvfB to generate a hydroxylamine. A two-electron reduction of the *N*-hydroxy-L-Val linked to the T domain by PvfC R domain releases the intermediate as an aldehyde. Two molecules of this aldehyde cyclize and dehydrate to create the 2,4-dihydropyrazine *N,N'*-dioxide. Isomerization forms dPNO, while aromatization forms PNO A. Production of the monoxide PNO B may result from transformation from dPNO (Figure 2.15) or cyclization between a *N*-hydroxy-L-Val aldehyde and a L-Val aldehyde. Valdiazen produced by *Burkholderia* may result from a four-electron reduction of the T-domain-linked *N*-hydroxy-L-Val to generate *N*-hydroxy-L-valinol. This molecule can be further modified to the diazoniumdiolate (Figure 2.14).
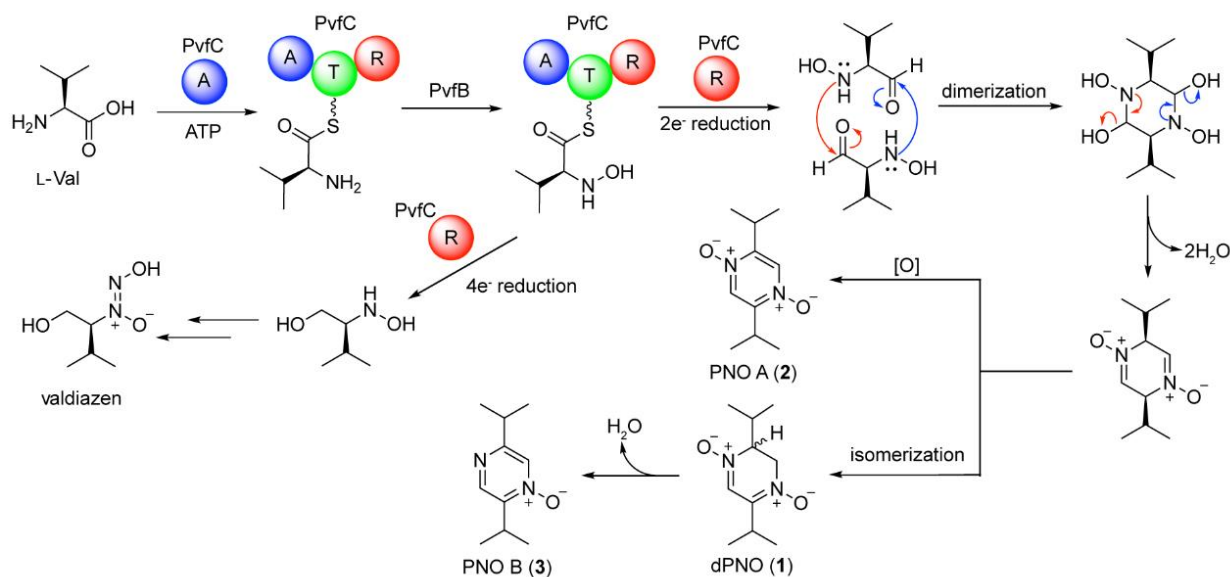
Figure 2.14. Proposed biosynthesis of (dihydro)pyrazine *N*-oxides in *P. entomophila* L48 and valdiazen in *B. cenocepacia* H111.
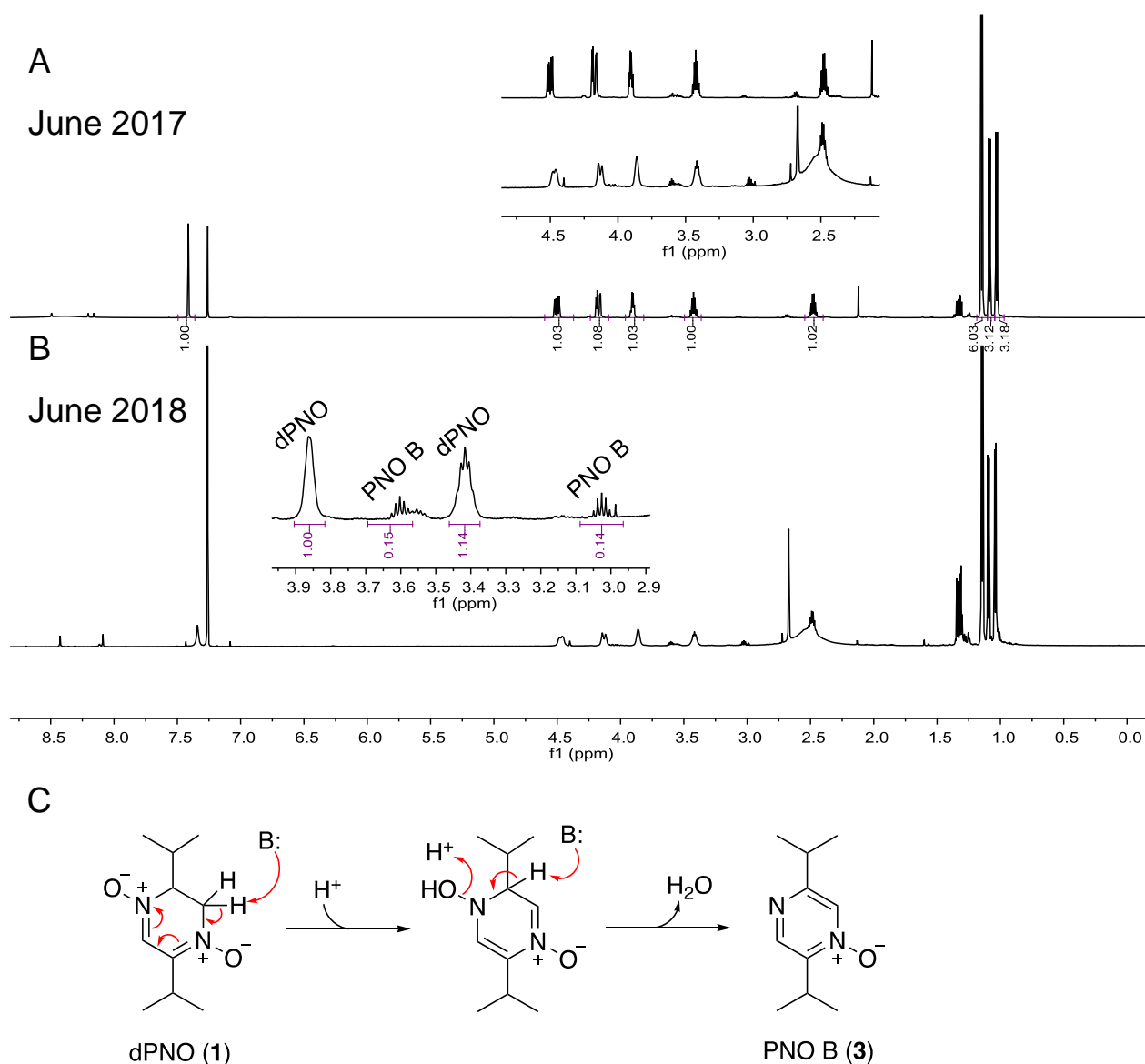
Figure 2.15. Transformation of dPNO (**1**) to PNO B (**3**). ^{1}H NMR spectrum in CDCl₃ (600 MHz) is shown for (A) isolated dPNO from *P. entomophila ΔpvfC* + pPSV-*pvfABCD*, and (B) the same sample after one year of storage. Based on integration of the ^{1}H signal, it is estimated that less than 15% of dPNO has converted to PNO B. (C) A proposed mechanism for the conversion from dPNO to PNO B.

Biological activity of (d)PNOs

We evaluated the activity of (d)PNOs to restore the virulence of the *ΔpvfC* mutant using a fly infection assay.[38] Cultures of the *ΔpvfC* mutant were supplemented with isolated **1**, isolated or synthetic **2**, and/or isolated or synthetic **3**. After 24 h of growth, these cultures were concentrated and fed to adult female flies. Although we initially observed that addition of

isolated dPNOs restored virulence of *ΔpvfC* mutant (Figure 2.16), further purified dPNO and synthetic PNO A and B do not restore the virulence of *ΔpvfC* against adult flies individually or combined (Figure 2.17), while the extracts of the *P. entomophila* wildtype culture restored the virulence of *ΔpvfC* to wildtype levels.[38] Furthermore, the biosynthesis of (d)PNOs does not require PvfD (Figure 2.10). Previously, *pvfD* and a homologue in *P. syringae* were shown to be essential for signaling activity. Together, the evidence suggests that (d)PNOs are potential biosynthetic intermediates or shunt products of the *pvf* pathway. The activity we first observed could have resulted from a minor species in the isolated sample. As some signaling pathways require the synergistic action of multiple metabolites,[92, 93] additional metabolites produced by *pvf*-encoded enzymes may be necessary for the full signaling activity. Identifying the additional metabolites required for the signaling activity will be the focus of future studies.
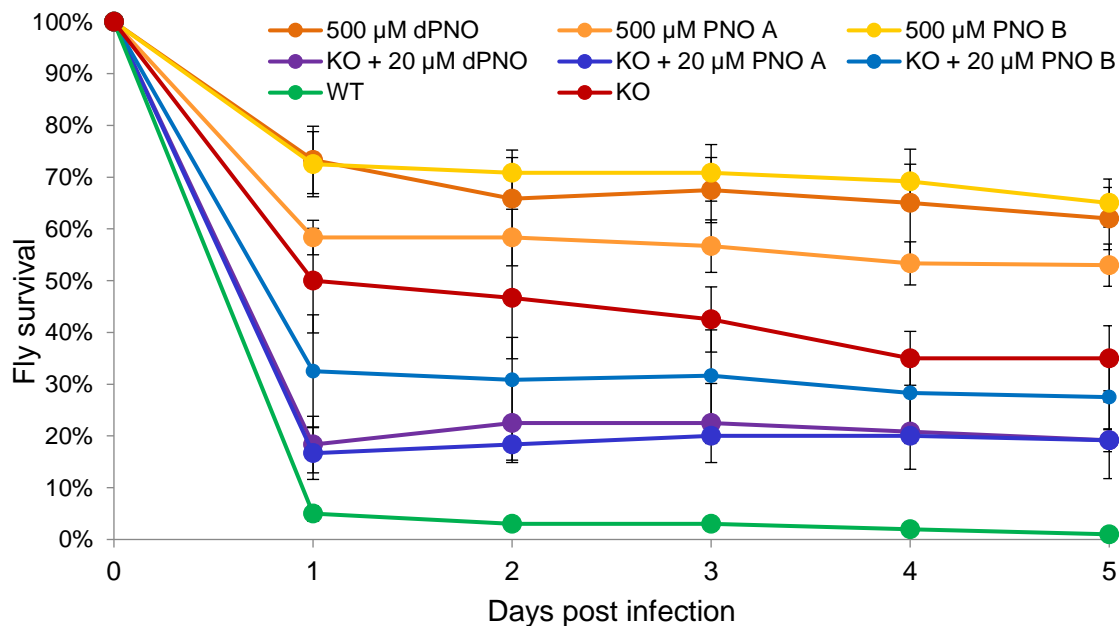
Figure 2.16. Initial isolated material of (d)PNOs restore virulence of *P. entomopila ΔpvfC* against *Drosophila*. Cultures of *P. entomophila ΔpvfC* + pPSV35 was supplemented with 20 µM of isolated dPNO (purple), 20 µM isolated PNO A (blue), or 20 µM isolated PNO B (light blue). Bacterial cultures supplemented with purified compounds were grown for 24 hr before feeding to female flies. In the absence of bacteria, 500 µM of each compound alone in sucrose exhibited little effect on fly growth (burnt orange, orange, and yellow for dPNO, PNO A and PNO B respectively). Representative data are shown for fly survival post bacterial infection over a five-day period.
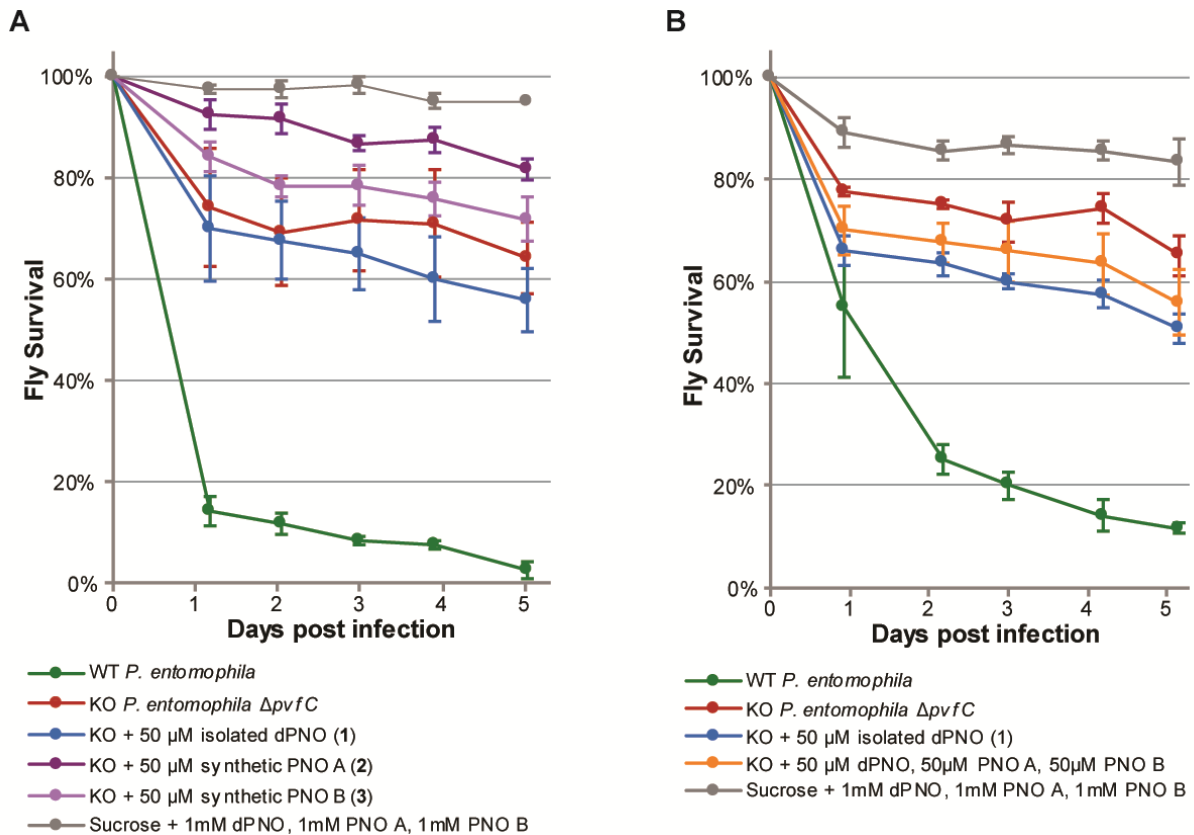
Figure 2.17. (dihydro)Pyrazine *N*-oxides do not complement virulence towards *Drosophila*. (A) Cultures *of P. entomophila ΔpvfC* + pPSV35 was supplemented with 50 µM of isolated dPNO (blue), 50 µM synthetic PNO A (dark purple), or 50 µM synthetic PNO B (lilac). (B) Cultures of *P. entomophila ΔpvfC* + pPSV35 are supplemented with 50 µM of isolated dPNO (blue) or a combination of 50 µM isolated dPNO, PNO A and PNO B (orange). Bacterial cultures supplemented with purified compounds were grown for 24 hr before feeding to female flies. In the absence of bacteria, 1 mM of each compound in sucrose exhibited little effect on fly growth (grey). Representative data are shown for fly survival post bacterial infection over a five-day period, and each graph is a separate experiment.

We conducted further studies to identify a biological role for the (d)PNOs. Due to their high production level in *Pseudomonas* and *E. coli* overexpression *pvf,* these compounds are unlikely to exhibit a toxic effect on Gram-negative bacteria. Thus, we tested their activity against *Bacillus subtilis,* a model Gram-positive bacterium. No significant growth inhibition was observed on solid media up to 100 µM and in liquid cultures up to 1 mM (Figure 2.18). At higher concentrations (100 µM), dPNO exhibited moderate chatechol-like metal-binding activity, as observed as a color shift of CAS dye from blue to purple (

Figure *2.19*). None of the (d)PNOs exhibited surfactant or hemolytic activity, or toxic effects to the flies at concentrations of 1mM.
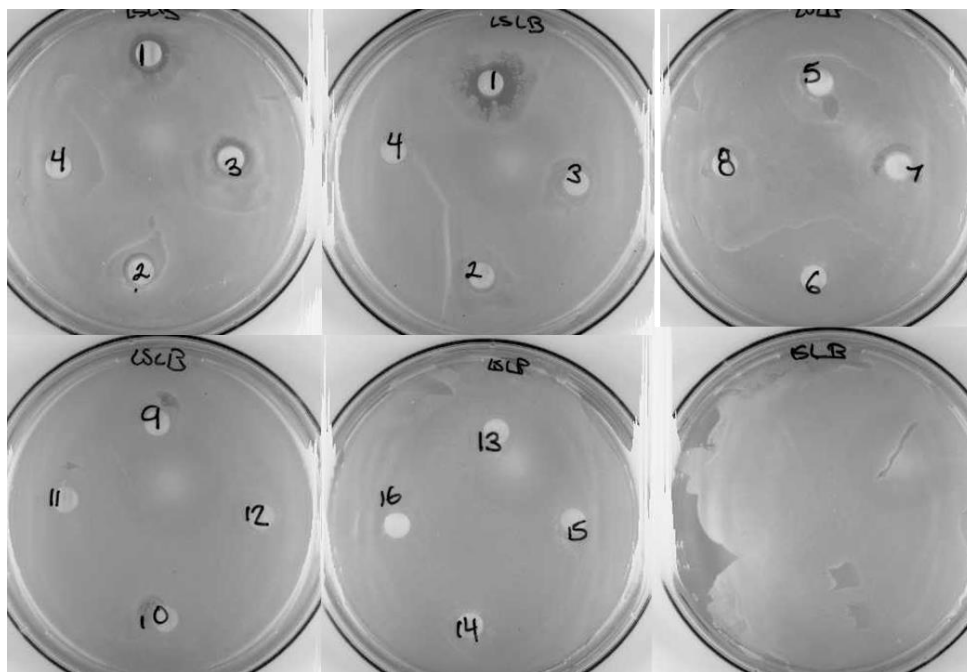
Figure 2.18. (d)PNOs are inactive against Gram-positive bacterium *Bacillus subtilis.*1. 200 µmol dPNO (in MeOH). 2. 100 µmol dPNO + PNOA (in MeOH). 3. 100 µmol PNO A (in MeOH). 4. 20µL Methanol. 5. 500 nmol dPNO. 6. 1 µmol dPNO 7. 2 µmol dPNO. 8. 20µL H_2O. 9. 2 µmol dPNO. 10. 2 µmol dPNO. 11. 2 µmol dPNO. 12. 10µL H2O. 13. 1 µmol synthetic PNO A. 14. 1 µmol biological isolated PNO A. 15. 1 µmol synthetic PNO B. 16. 1 µmol Synthetic di-isopropyl pyrazine. All in H_2O unless stated.

100 µM   10 µM   0 µM   dPNO (**1**)

Figure 2.19. dPNO exhibits moderate catechol-like metal binding activity at 100 µM. Samples of dPNO were incubated with Chromeazurol S (CAS) dye. CAS dye shifts from blue to purple in the presence of a catechol-like chelator.

2.4 Discussion

Although (d)PNOs do not restore the signaling activity of *ΔpvfC*, their unique structures suggest other potential activities toward microbes or multicellular organisms. Volatile pyrazines are well-established pheromones for insects and are distributed in environmental niches ranging from *Pseudomonas* spp. on wine corks to bacteria associated with leaf cutter ants.[94-99] Recently,

a 3,5-dimethylpyrazin-2-ol was identified as a quorum-sensing molecule for *Vibrio cholerae*.[100] In regard to the *N*-oxides, only a few examples have been identified in natural products, such as 4-hydroxy-2-heptylquinoline-*N*-oxide (HQNO) and iodinin (1,6-dihydroxy phenazine *N,N'*-dioxide), which possess signaling, antimicrobial, and redox cycling activities.[101-104] The *N*-oxide derivatives of pyrazines are rare; in particular, the 2,3-dihydropyrazine *N,N'*-dioxide moiety of 1 is unprecedented in natural products, to the best of our knowledge. Unlike iodinin, dPNO does not have connected aromatic rings, yet it is stable at the pH range 1−11 (Figure 2.20) and for several months at room temperature (less than 15% was converted to PNO B after one year in storage, Figure 2.15). Attempts to synthesize dPNO have been unsuccessful; however, overexpression of *pvf* allows the isolation of this unique compound in relatively high yields for further exploration of its chemical properties and biological activities.

We have identified a new biosynthetic pathway for pyrazine derivatives involving a noncanonical NRPS, PvfC, and a diiron *N*-oxygenase, PvfB. The A-T-R domain structure of PvfC resembles two fungal NRPSs, one involved in a general biosynthetic pathway for lysine[105] and the other specific for the biosynthesis of a dityrosyl-pyrazine in *Aspergillus flavus*.[106] Interestingly, in the same work *N,N'*-dioxide species of dityrosyl-pyrazine had also been isolated from *A. flavus*, but no PvfB homologue was found in the gene cluster and the biosynthetic origin of this *N*-oxidation is unknown. Reactions conducted by PvfC and PvfB and their homologues provide essential chemistries for divergent biosynthesis of a variety of structures, including diazeniumdiolates in fragin and valdiazen and (dihydro)pyrazine *N*-oxides.

In summary, through genome-mining in *Pseudomonas* and overexpression of the *pvf* gene cluster, we have identified and characterized a family of pyrazine *N*-oxides, including a novel dihydropyrazine *N,N'*-dioxide metabolite. This work identifies the role of a noncanonical NRPS in the biosynthesis of pyrazine derivatives and suggests a common biosynthetic origin for two

families of molecules. Our work sets the stage for understanding the biosynthesis and bioactivity

of (d)PNOs and identifying new molecules involved in *Pseudomonas* signaling and virulence.
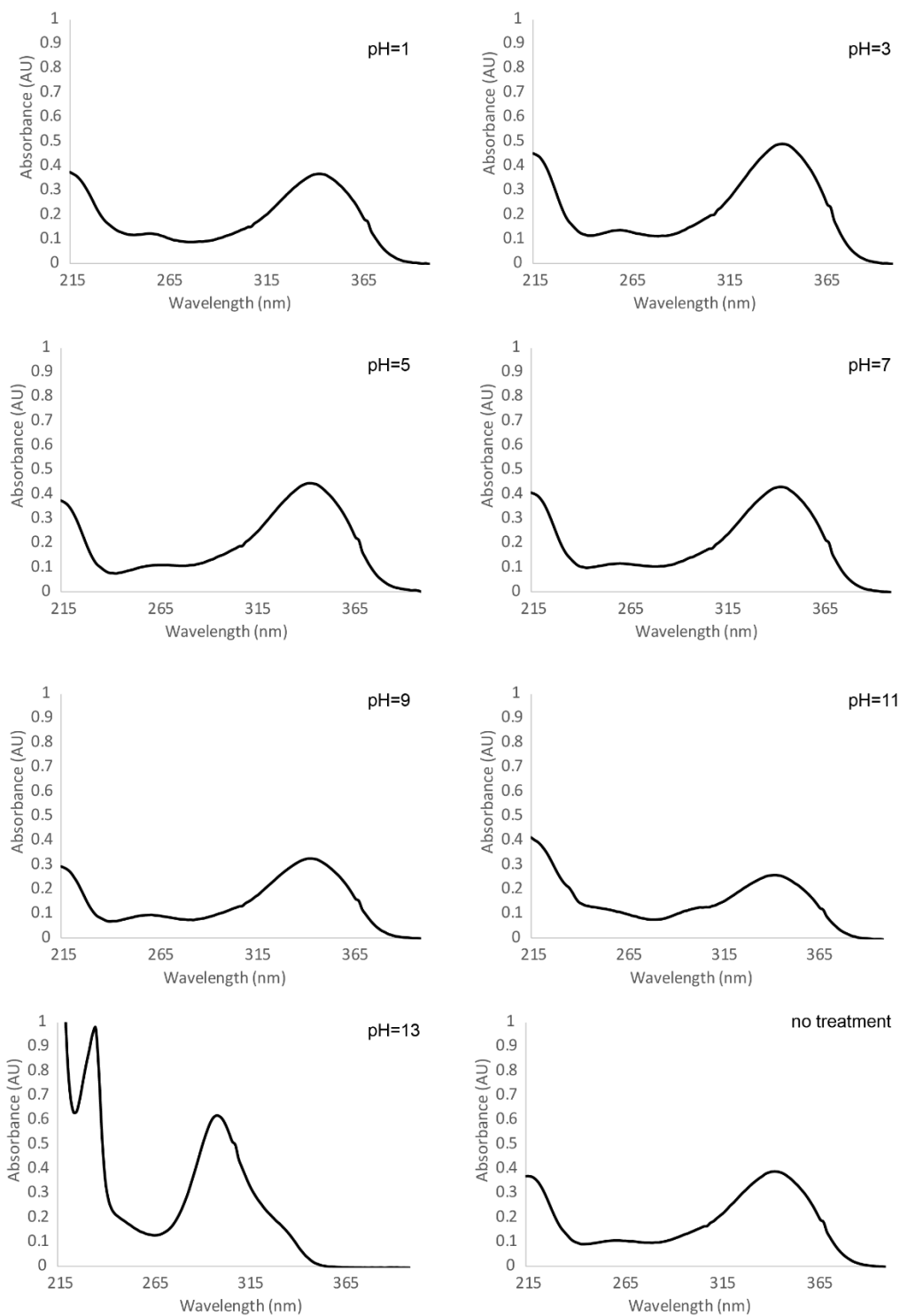
Figure 2.20. pH stability of dPNO. The pH of dPNO solution in water was adjusted to 1, 3, 5, 7, 9, 11, or 13. The UV traces for dPNO remain unchanged at pH 1–11. The UV trace of dPNO at pH 13 resembles that of PNO A, suggesting that dPNO converts to PNO A at pH 13.

58

Chapter 3 The *pvf* signaling pathway regulates the secretome of *Pseudomonas entomophila*

3.1 Introduction

Pseudomonas bacteria use diverse small signaling molecules to communicate changes in their environment to neighboring cells. These cell-to-cell communication molecules, known as quorum sensing (QS) molecules, are produced in response to environmental cures and regulate gene expression to adapt to the environment or interact with eukaryotic hosts. Cell-to-cell communication can take place between bacterial cells of the same species or different species. It regulates a wide range of pathways that contribute to bacterial virulence, including production of virulence factors that enable pathogens to infect the host, biofilm formation, and antibiotic resistance.

Quorum sensing was initially identified in *Vibrio fischeri* as the cell-to-cell communication mechanism that regulated bioluminescence.[107] Since, this communication network has been shown to control behaviors such as conjugation, biofilm formation, natural product synthesis, sporulation, and secretion of toxins and other virulence factors.[108-111] In Gram-negative bacteria, the original QS prototype is the LuxI/LuxR system, where LuxI was a synthetase for an acyl homoserine lactone, or AHL, that was received by LuxR.[112] Later, "LuxR solos" were identified as LuxR transcriptional regulators without a conjugate LuxI.[113] The activity of some of these LuxR solos are also independent of the AHL family of QS molecules,[114] suggesting that these LuxR receptors bind to a different molecular ligand.[115]

In addition to having cell-density dependent activity, the activity of QS molecules is concentration dependent. These small molecules are secreted from the cell and therefore can be received by neighboring bacteria, including cells with genetic deletions of the QS biosynthetic

pathway and other species. Studies have found bacteria that do not harbor biosynthetic pathways for quorum sensing molecules but do contain the receptors and respond in the presence of the QS molecule. Finally, these molecules tend to be autoinducers; at specific concentrations they will also activate their own biosynthesis.

The most well studied quorum sensing molecules are the homoserine lactones, and more commonly acyl homo serine lactones, or AHLs. In *Pseudomonas aeruginosa,* the Las, and Rhl, QS systems control expression of virulence factors.[116] LasR and RhlR are homologues to LuxR in that they sense AHLs (*N*-3-oxo-dodecanoyl-homoserine lactone and *N*-butanoyl-homoserine lactone respectively). In *Pseudomonas,* other examples of QS molecules include 2-heptyl-3-hydroxy-4-quinolone the *Pseudomonas* quinolone signal (PQS), and *cis*-2-dodecnoic acid the diffusible signal factor DSF.[117] The collection of known small molecule signals is diverse, but still limited. For instance, in *Pseudomonas*, not much is known beyond the molecules from *P. aeruginosa*. We also have taken particular interest in novel molecules that control bacterial population, especially in strains where other quorum sensing molecules have not been identified. Interestingly, *Pseudomonas entomophila* does not harbor any AHL synthetase homologs, and therefore do not directly rely on AHL quorum sensing for cell-to-cell communication.

Successful studies have been conducted regarding the use of quorum sensing inhibitors as an anti-virulence strategy.[118, 119] In addition, a strategy to use QS molecule analogs is a viable strategy where early activation of QS pathways could be detrimental to bacterial viability. Quorum sensing inhibitors have also been identified in Nature.[120, 121] Finally, degradation of QS molecules, known at quorum-quenching has been shown to attenuate production of virulence factors and reduce virulence.[122] Like other anti-virulence strategies, QS inhibition has lower chance of resistances and microbiome damage compared to anti-biotic targets.[123] Targeting signaling is also a method to target multiple virulence targets with a single therapeutic.[119, 124]

The collection of proteins and small molecules secreted by *Pseudomonas,* known as the *secretome,* play an important role in infection. They cause virulence by targeting host processes, evading the host immune response, or sequestering nutrients from the environment. Previously, we have identified the *Pseudomonas virulence factor (pvf),* a biosynthetic gene cluster conserved in over 300 strains of *Pseudomonas,* including strains that engage in pathogenic or symbiotic relationships with a range of human, animal, and plant hosts. The *pvf* cluster is implicated in the virulence of the animal pathogen *Pseudomonas entomophila* L48 and the plant pathogen *Pseudomonas syringae* pv*. syringae* UMAF0158.[37, 38] Disruption of the *pvf* cluster *P. entomophila* and *P. syringae* significantly decreases the virulence of these strains against adult flies and tomato plants, respectively. Studies have shown that the *pvf* cluster is involved in the production of signaling molecule(s) that regulate virulence factors, including the pore-forming toxin monalysin in *P. entomophila* and the phytotoxin mangotoxin in *P. syringae*.[36, 40] Therefore, we sought to characterize the signaling molecule produced by the *pvf* cluster, and the effect of the *pvf* cluster on the secretome. *Pseudomonas entomophila* was used as a model because it has a well-studied infection model and a fully sequenced genome. [55]

3.2 Materials and Methods

General methods. See Chapter 2 pg 21.

Inducible expression of *pvf* in *P. entomophila*. See Chapter 2 pg 21.

<u>Determining signaling activity with promoter-reporter strains</u>

Promoter-expression cassettes. All promoter-reporter cassettes were cloned with the following protocol. Promoter regions were amplified from *P. entomophila* L48 genomic DNA using primers listed. The promoter region was considered to be an approx. 500 bp region 5' to the respective protein. The first set of promoter-expression cassettes kept the start site of *gfp/lacZ* the same number of base pairs away from the promoter region as the gene it regulated. As the GFP and LacZ protein have their own ribosomal binding site in this plasmid, we included a few amino acids of the original gene to maintain regulation. This modified promoter region was amplified and cloned into the expression cassette vectors. These purified PCR products were digested with restriction enzymes EcoRI and HindIII and then ligated 5' to GFP or LacZ, on the pUC-GFP and pUC-*lacZ* plasmids respectively, using the corresponding restriction sites.[125, 126]

Wildtype and *pvf* knockout strains of *P. entomophila* L48 were transformed as described in Chapter 2 with a helper vector pTNS3 and the following vectors: pUC-P$_{pvf}$-GFP, pUC-P$_{pvf}$-*lacZ*, pUC-P$_{mnl}$-GFP, pUC-P$_{mnl}$-*lacZ*, pUC-GFP, and pUC-*lacZ*. These transformed cells were recovered in 1 mL LB for 2 hours while shaking at a speed of 225 rpm. A sample of 150 μL of the cell suspension was plated on LB agar containing kanamycin and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and incubated at 28 °C for 1 day until blue colonies appeared. The promoter-Gfp/LacZ cassette is inserted at the single T$_n$7 site, which is located upstream of the *GlmS* gene.[125] Cassettes were verified by PCR of the genomic DNA (Figure 3.1).

Figure 3.1. Example analytical PCR to verify insertion of the promoter-reporter cassette upstream of *GlmS*. Expected band length is 2500bp. Strains are denoted by wildtype (WT) or a deletion mutation (*dpvf*) with a 500bp promoter region amplified prior to the *lacZ* gene. Px denotes no promoter region was inserted as a control.

Beta galactosidase assay. For each experiment, fresh Z-buffer was prepared containing: 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$ and 50 mM beta-mercaptoethanol at pH 7.4. Ortho-nitrophenyl-β-galactoside (ONPG) substrate was prepared by dissolving fresh ONPG in water to a concentration of 4 mg/mL. 1 M K_2CO_3 was prepared as the stop solution.

For time course activity, 50 mL of LB-kan were inoculated with 200 μL of the overnight culture and grown at 30 °C at 225 rpm. At early log phase (OD_{600} 0.2−0.3), wildtype or *pvf* deletion extracts, resuspended in methanol at 10 μL per 1 mL or 1 mg/mL, was added to each culture, with methanol as a control. Immediately, a 1 mL sample was removed, OD_{600} recorded and 1 μL of 34 mg/mL chloramphenicol added on ice. The culture was then returned to the shaker. At appropriate time points the OD_{600} was measured and a sample was taken in the same manner. After all samples were taken, 200 μL of each sample was added to 800 μL of Z buffer for the ONPG assay. Three independent cultures were grown for each strain.

For complementation with extracts, overnight cultures of *P. entomophila* were used to inoculate 2 mL of LB. Cultures were incubated at 28 or 30 °C for 24 hours. Complementation was induced 4 hours post inoculation. Genetic complementation was induced with 1 mM IPTG. Chemical complementation was induced with a solution of compound, extract, or HPLC fraction in methanol at concentrations specified. Controls included 10 μL methanol per mL culture. At 24 hours 200 μL sample was removed, diluted to 1 mL with water, OD_{600} recorded and 1 μL of

34 mg/mL chloramphenicol added on ice. For monalysin promoter strains, two replicates of 20 µL of each diluted sample was added to 980 µL of Z buffer for the ONPG assay. For the *pvf* promoter strains, two replicates of 100 µL of each diluted sample was added to 900 µL of Z buffer for the ONPG assay. For other promoter strains, the full 1 mL of culture was spun down at 3500 x g for 10 min and the cell pellet was resuspended in Z buffer for the ONPG assay. Three independent cultures were grown for each strain.

To each ONPG assay, 100 µL of chloroform and 50 µL 0.1% SDS were added and vortexed. The lysed cells were equilibrated at 28 °C for 5 minutes. Reactions were initiated by the addition of 200 µL of ONPG substrate, vortexed briefly and incubated at 28 °C with shaking. After the appearance of color, 500 µL of stop solution was added; the reactions were kept at 28°C until reactions for all samples were stopped. The time elapsed between ONPG addition and the addition of stop solution was recorded.
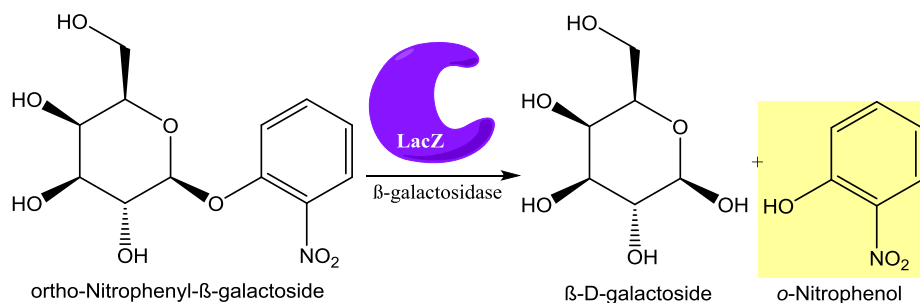
Scheme 3.1. Enzymatic reaction of ONPG cleavage by β-galactosidase.

The absorbance readings at 420 nm and 550 nm were recorded using a BioRad Smart Spec Plus spectrophotometer. The activity of the beta-galactosidase was determined by the following calculation:

$$Miller\ units = \frac{1000x\ A420 - 1.75xA550}{t\ x\ v\ x\ A600}$$

A420 = absorbance at 420 nm (yellow)
A550 = absorbance at 550 nm (background scattering)
A600 = absorbance at 600 nm (cell density)
t = time of reaction
v = volume of cells in reaction

pvf deletion strains and heterologous expression in *E. coli*

pvf **deletion in** *P. entomophila.* Single mutants of *pvfA-D* were generated previously.[38]

For the complete markerless knockout of *pvf*, 400 bp flanks upstream and downstream of the

pvf transcriptional region were amplified from *P. entomophila* L48 genomic DNA and cloned into

the pEXKm5 plasmid (Figure 3.2).[72] *P. entomophila* was transformed with the pExKm-*pvfKO*

knockout vector using the electroporation method described in Chapter 2. These transformed

cells were recovered in 1 mL LB for 2 hours while shaking at a speed of 225 rpm. A sample of

150 µL of the cell suspension was plated on LB agar containing kanamycin and incubated at 28

°C for 1 day until colonies appeared. Knockout isolates were selected and counter selected for

using kanamycin and sucrose respectively and verified by analytical PCR and sequencing

(Figure 3.3).

Figure 3.2. Markerless knockout of the *pvf* cluster using pExKm5 plasmid. DNA fragments flanking the gene cluster are amplified and combined using overlap PCR, then ligated into the multiple cloning site of pExKm5. After transformation into *P. entomophila*, homologous recombination is selected for using kanamycin, and counter selected using sucrose.
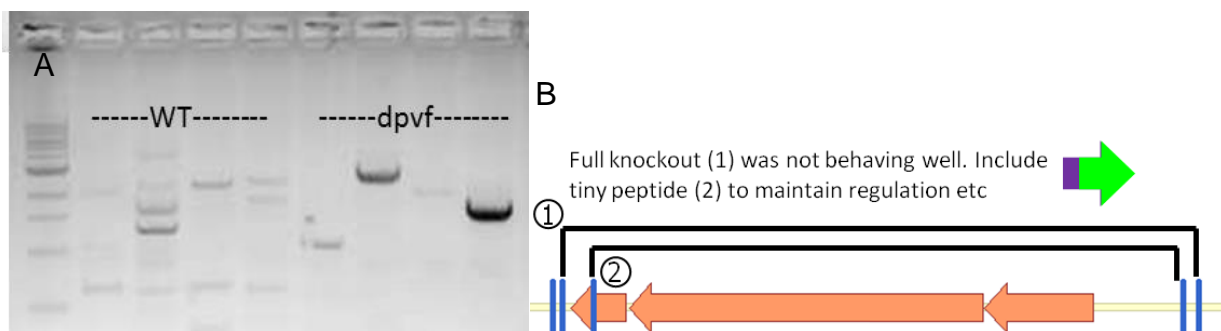
Figure 3.3. (A) Analytical PCR of the genomic DNA verifies deletion of *pvf* cluster using primers both upstream and downstream of the cluster. Expect band sizes are 933, 2448, 1886, and 1495 bp uniquely for the *pvf* deletion strain. The second PCR product was sequenced confirmed. (B) In order to not modify additional genes neighboring the *pvf* cluster, a tiny peptide was retained in the second knockout to maintain regulation but not activity.

Inducible expression of *pvf* in *Burkholderia cenocepacia* HI2424. The *pvf* cluster was amplified from *B. cenocepacia* HI2424 genomic DNA using a touch-down PCR thermocycle, and a high fidelity phusion polymerase. These purified PCR products were digested with restriction enzymes NdeI and HindIII and then ligated downstream of the P_{Rha} promoter on the pSCRhaB2 plasmid using the corresponding restriction sites.[127] *E. coli* RHO3 cells were transformed with sequenced verified plasmid and used to introduce the inducible expression vector *B. cenocepacia* HI2424 via biparental mating as described in Chapter 2. A sample of the mating cell suspension was plated on LB agar containing trimethoprim and incubated at 37 °C until colonies formed.

pvf deletion in *Burkholderia cenocepacia* HI2424. Primers were designed to amplify fours fragments, with overhangs that were complementary to the adjacent fragments for Gibson Assembly (Figure 3.4): a one kilobase fragment upstream of the gene transcriptional region (right flank), trimethoprim resistance gene *dhfrII*, a one kilobase fragment downstream of the transcriptional region (left flank), and a pCR blunt backbone with kanamycin resistance gene. Each were amplified by PCR with Q5 DNA polymerase with *B. cenocepacia* HI2424 genomic DNA as a template for flanking regions, plasmid pSCRhaB2[127] as a template for *dhfrII* and plasmid pCR-Blunt as a template for the backbone. For an overlap extension PCR, the flanking region products, 10 ng, acted as primers to 1 ng of the *dfhrII* cassette. The initial overlap step was a two-step thermocycle with a 60s extension time for ten cycles. The second thermocycle was a 70 − 60 °C touchdown for ten cycles with 60 °C annealing temperature for twenty cycles and two minutes extension time all thirty cycles. In Gibson assembly, about 60 ng (32 pmol) of each of the pCR blunt backbone and overlap extension fragments, 2.8 kb and 3.5 kb respectively, were combined with the 1.33x master mix and incubated at 50 °C for one hour according to the original publication. *B. cenocepcaia* HI2424 was transformed with sequenced plasmids as described in Chapter 2*.* Knockout isolates were selected and counter selected with trimethoprim and kanamycin respectively and verified by analytical PCR and sequencing.
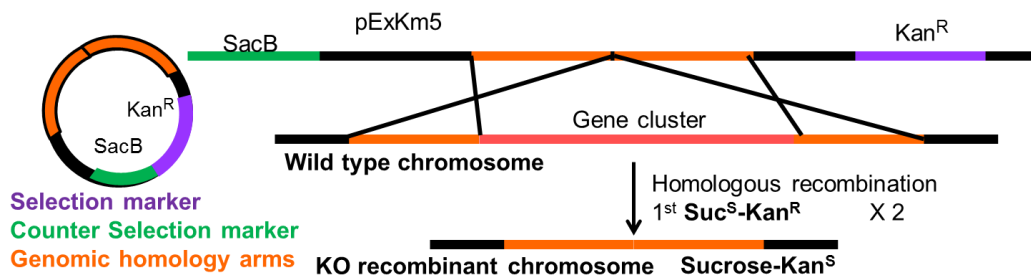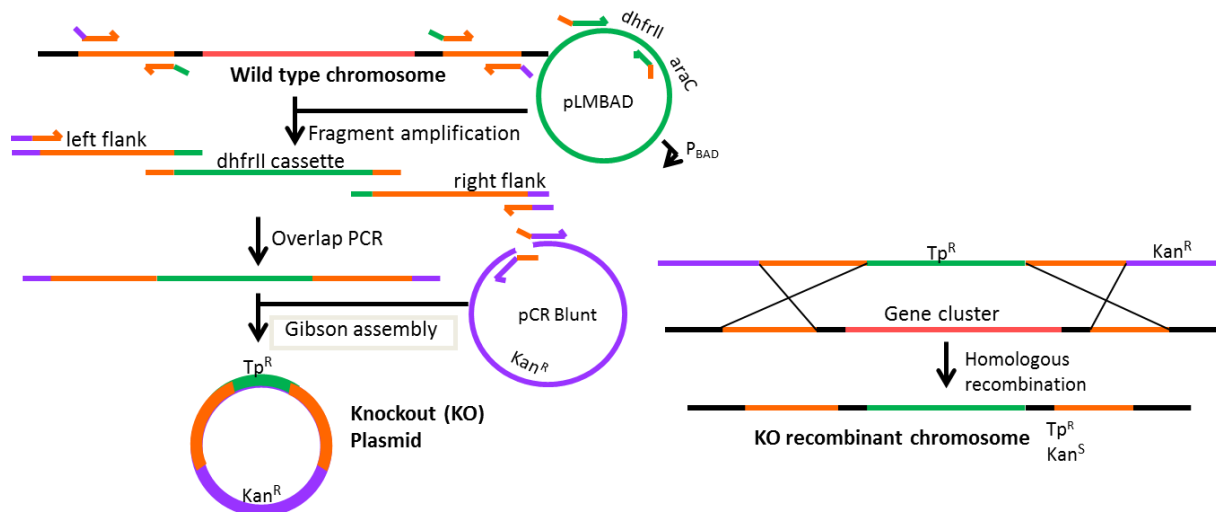
Figure 3.4. Knockout of the *pvf* cluster using Gibson Assembly. DNA fragments flanking the gene cluster are amplified and combined with trimethoprim resistance marker. After transformation into *B. cenocepacia*, homologous recombination is selected for using trimethoprim, and the second event using kanamycin.

Inducible expression of *pvf* in *E. coli*. *E. coli* Bap1 cells, with *sfp* to load the phosphopatetheine post-translational modification onto the NRPS, were transformed with the following vectors or combinations: pPSV35-*pvfABCD$_{PE}$,* pPSV35, pSCRha-*pvfD$_{PE}$,* pSCRha-*pvfAEBCD$_{BCC}$,* pSCRhaB2, pPSV35-*pvfABCD$_{PE}$* + pSCRha-*pvfD$_{PE}$,* pPSV35-*pvfABCD$_{PE}$* + pSCRhaB2, and pPSV35 + pSCRha-*pvfD$_{PE}$*. Growth of *E. coli* expression *pvf,* dichloromethane extracts of cell-free culture supernatant, and LC-HRMS analysis of the extracted were conducted as described in Chapter 2. Three independent cultures were grown for each strain as biological replicates.

Proteomic analysis of secreted proteins

Overnight culture of *P. entomophila* was used to inoculate 50 mL of LB media. Cultures were incubated at 30°C for 24 hours and induced with 1 mM IPTG at four hours post inoculation. After 24 hours cultures were carefully (to reduce cell lysis) spun down to remove cells 3000 x g for 20 minutes and the supernatant was filtered through 0.45-micron filters. Total concentration of protein in the supernatant was determined by Bradford assay (Biorad, Figure 3.5).

	Sample	Abs 595 nm	Conc (µg/mL)
KO	120a	0.4736	36.275
	120b	0.5277	49.792
	120c	0.4656	34.267
WT	122a	0.4706	35.525
	122b	0.4993	42.700
	122c	0.4129	21.092
OE	123a	0.5726	61.033
	123b	0.6320	75.883
	123c	0.5726	61.025
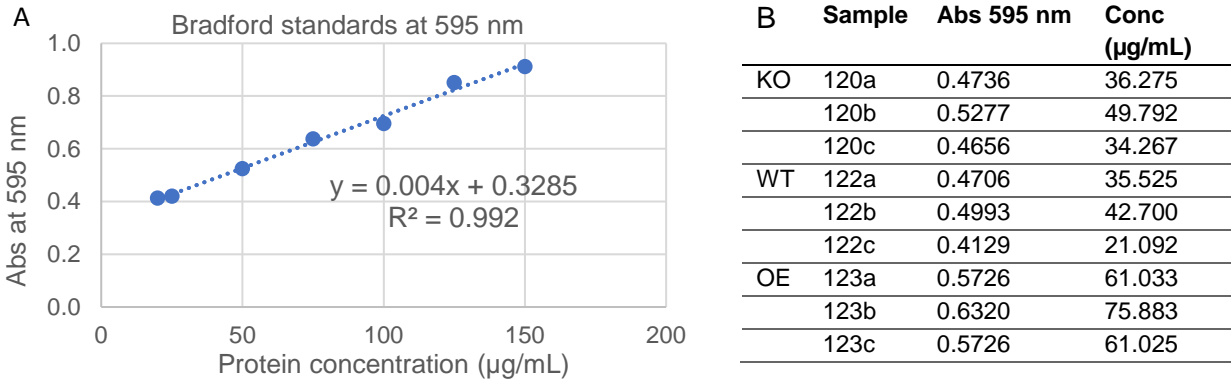
Figure 3.5. (A) Standard curve of BSA (B) concentration of proteomics samples determined by Braford assay.

To precipitate out secreted proteins, 6 mL of supernatant was mixed with 1 mL of 50% trichloroacetic acid (TCA) (1 g/mL). A 3 mL sample was used for proteomics analysis and a 1.5 mL sample was used for visualization on a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel. After incubation for 15 minutes at −20 °C the samples were centrifuged at >20000 x g for 30 minutes. The supernatant was removed and the protein pellet washed with ice cold acetone. For SDS-PAGE analysis, the pellet was resuspended in 100 µL of protein loading dye. For proteomic analysis, the pellet was resuspended in 1 mL 7 M urea and incubated at 16 °C overnight.

Proteomic samples were digested with trypsin and analyzed by liquid chromatography-tandem mass spectrometry on Easy nLC 1000-QExactive HF. Samples were separated in 90 minutes over a gradient from 1 to 32 %B, where mobile phase A contained water and 0.1% formic acid and mobile phase B contained acetonitrile and 0.1% formic acid. The top 15 most intense ions at each retention time were selected for HCD fragmentation. Data were searched against the proteomic database on Uniprot for *P. entomophila* L48 using MaxQuant with 10ppm precurosor ion and 0.02 Da product ion mass tolerance. Carbamidomethylation of Cys was set as a fixed modification. Deamidation of asparagine and glutamine, and oxidation of methionine were set as variable modifications. Peptide false discovery rate of 1% was used to filter all

results. False positives, contaminants and proteins for which one peptide fragment was identified were removed. Label free quantification (LFQ) intensities was used for quantification of each peptide. Statistical significance of each protein feature was determined using Perseus, a statistical program.[128] If 6 or more of the samples did not have a peptide (LFQ intensity was 0), the peptide was removed from analysis. Log_2 LFQ intensities were calculated and missing values were replaced with random numbers drawn from normal distribution.[129] Analysis of variance (ANOVA) multi-sample test determined whether any of the means between wildtype (WT), knockout (KO) and overexpression (OE) were statistically different from each other. For our proteomics data, 1047 proteins were considered statistically significant according to ANOVA tests (p value < 0.05). Proteins were further filtered based on ANOVA p-values and the absolute differences of the log_2LFQ intensities. There were 506 proteins where the ANOVA p value < 0.01. Of these, the set where Log_2int WT-KO > 1 (greater than two-fold difference) and Log_2int WT-OE < 0 contains 94 proteins and the set where Log_2int WT-KO< −1 (greater than two-fold) Log_2int WT-OE < 0 contains 76 proteins.

Biological assays

P. entomophila cultures for biological activity assays. Overnight cultures of *P. entomophila* were used to inoculate 2 mL of LB. Cultures were incubated at 28 or 30 °C for 24 hours. Genetic complementation was induced with 1 mM IPTG at 4 hours post inoculation. For chemical complementation, solutions of compound, extract, or prepHPLC fraction in methanol were added to cultures at the concentration specified four hours post inoculation. Controls included 10 µL methanol per mL culture. Supernatant was recovered by centrifuging cultures at 6000 x g for 5 min.

Motility Swimming assay. Dry soft agar plates (1% tryptone, 0.5% NaCl, 0.5% agar)[130] were point inoculated using a pipette tip dipped in a bacterial culture of *P. entomophila* grown to confluency. These plates were incubated at 28 °C upside down for 24 hours before imaging.

Detection of surfactant production. For visualization, culture supernatant was stained with methylene blue, which does not influence droplet formation.[56] A 50 µL sample of *P. entomophila* culture, diluted to OD_{600} of 0.1 or 0.2, was combined with 50 µl of 0.05% methylene blue. For visualization of small molecule surfactant properties, HPLC fraction were resuspended in methanol to 1 mg/mL. This solution was diluted 1:10 with water and 50 µL was added to 50 µL of 0.05% methylene blue. A sample of 20 microliters of this solution was dropped onto a hydrophilic surface, parafilm, and droplet collapse was visualized by eye and recorded. The diameters of these droplets were also measured and recorded after they had dried at room temperature overnight.

Detection of siderophore production. To test for the presence of iron chelators, 10 µL of Cromazol S (CAS) reagent[80] was incubated with 90 µL of cell-free culture supernatant (OD_{600} of 1 or 2) at room temperature for 30 minutes. For small molecules, HPLC fractions were resuspended in methanol to 1 mg/mL. This solution was diluted 1:10 with water and 90 µL was added to 10 µL CAS reagent. Siderophore activity resulted in a color change from blue to green to yellow or blue to purple. Color change was recorded and absorbance was measured in 96-well plates on a Tecan infinite M1000pro plate reader in the range of wavelengths of 350−700 nm.

Extraction and purification of secreted small molecules

Four 1 L low salt LB cultures were inoculated with 2 mL overnight cultures of *P. entomophila* L48 WT or *Δpvf*. The cultures were grown at 30 °C with shaking at 225 rpm for 24 hours. The culture supernatant was separated from bacterial cells by centrifugation and each liter was extracted twice with half culture volume of ethyl acetate. The organic layer was

separated from the supernatant, dried with magnesium sulfate, then evaporated to dryness in a round bottom flask. The extracted metabolites were resuspended in about 20 mL of ethyl acetate, and transferred to a scintillation vial. The ethyl acetate was evaporated to dryness and the extraction sample was stored at −20°C.

Metabolites were purified from the large-scale extraction using preparatory high-performance liquid chromatography (Agilent PrepStar). Dried extracts were resuspended in 80 % ACN/water (1 mL/100 mg extract) and injected on a Phenomenex Luna C18 column in 1 mL injections. Compounds of interest were separated using mobile phase A (water, 0.1% TFA) and mobile phase B (acetonitrile, 0.1% TFA) over a gradient of 5 minutes at 5% B, 15 min 5−40% B, 15 min 40−95% B, and 6 min 95%, at a flow rate of 15 mL min^{-1}. All absorbance in the range of 190−450 nm were recorded. Fractions were collected every 30 seconds from min 30 to 45, combined separately based on UV, and dried under high vacuum.

3.2 Results

<u>Characterizing the signaling properties of the *pvf* pathway</u>

To study the control of promoters by the *pvf* biosynthetic products, we constructed promoter reporter cassettes. These cassettes, which contained the promoter regions of *pvf*-regulated genes upstream of either *gfp* or *E. coli lacZ*, were inserted into the T_N7 site downstream of *glmS* in the *P. entomophila* genome (

Figure *3.6*). The P_{pvf}-*lacZ* cassette provides a highly sensitive method to study the production of the *pvf* molecules in the wildtype strain. In addition to the *pvf* promoter, promoter regions for the pre-monalysin peptide *mnl* were chosen for this expression study. Previous studies have suggested that monalysin is regulated in part by the *pvf* pathway.[40] To study the effect of *pvf* on the activity of these promoters, the *pvf* and monalysin promoter reporter cassettes were inserted at the Tn7 site in the wildtype, *ΔpvfB*, *ΔpvfC, ΔpvfD,* and *Δpvf P. entomophila*. To verify that changes in activity are not due to disruption of global transcriptional function, the ribosomal S12 promoter region was used as a positive control and the promoterless *gfp* and *lacZ* were inserted as a negative control.

Six *lacZ* hybrid strains containing translational reporters were obtained from Isabelle Vallet-Gely. In this case, the *lacZ* gene is inserted as a translational fusion after the following genes: PSEEN3045 (EtlA, an NRPS involved in the production of entolysin), PSEEN5493 (an acetyl transferase of unknown function), and PSEEN0973 (a putative dihydrodipicolinate synthetase of unknown function). To study the effect of *pvf* on these promoters, an insertional mutation at *pvfC* (IM*pvfC*) was introduced to each reporter strain. We will use our promoter reporters and *lacZ* hybrids to investigate the PVF compounds as signaling molecules.
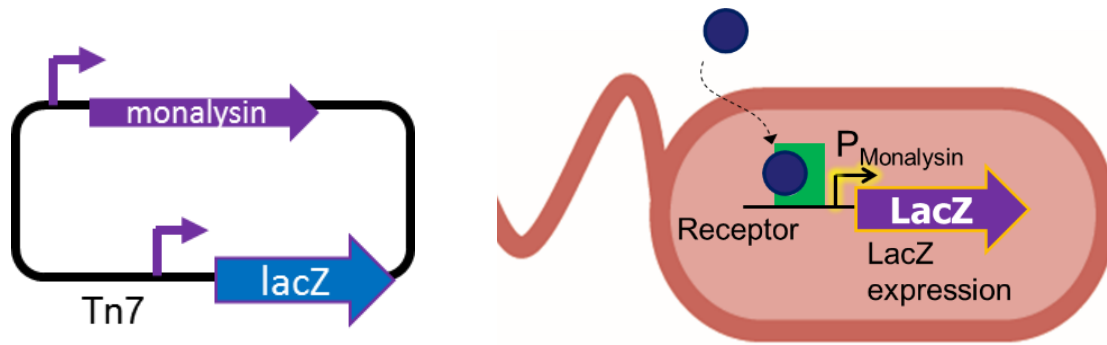
Figure 3.6. Scheme for promoter-reporter cassette to monitor *pvf* signaling activity. (A) A 500bp promoter region including the first few amino acids of the gene (ie monalysin) is amplified and inserted prior to the *lacZ* gene at the Tn7 site of *P. entomophila*. (B) Presence of the pvf molecule will activate the promoter, expressing *lacZ* which can be monitored through an enzymatic cleavage of ONPG to *o*-nitrophenol.

Preliminary experiments conducted by Martina Knechel using the GFP promoter reporters were unsuccessful; the expression of fluorescent protein was not distinguishable against the background fluorescence, and was highly variable in the *Pseudomonas entomophila* cultures. Therefore, we have focused our efforts on the *lacZ* reporters.

The *lacZ* promoter-reporter activity can be assessed both qualitatively and quantitatively. The first relies on the cleavage of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) by β-galactosidase into galactose and 5-bromo-4-chloro-3-hydroxyindole. 5-bromo-4-chloro-3-hydroxyindole will dimerize to form an insoluble indigo dye 5,5'-dibromo-4,4'-dichloro-indigo. This indigo dye can be visualized either on an agar plate on in liquid media. This method can rapidly determine the activation of our promoter. The second is a β-galactosidase activity assay using ONPG (ortho-nitrophenyl-β-galactoside). This is cleaved to give *o*-nitrophenol, a yellow compound that can be quantified using light absorbance at 420 nm. The extent of reaction, in relation to the density of the cells in culture, provides a quantitative activity of the enzyme, in Miller Units.

To verify the role of *pvf* in signaling, we compared the monalysin promoter activity between a wildtype (WT) promoter-reporter strain and a *pvfC* deletion (KO) promoter-reporter strain. We observed a significant difference in activity, particularly after the start of log phase growth (5–10 hours post inoculation, Figure 3.7). Wildtype activity was recovered when the KO

reporter strain was grown with DCM extracts of WT spent media, but not culture extractions of spent media from *pvf* deletion strains. This suggests that the *pvf* cluster is responsible for the monalysin promoter activity. The slight increase with addition of knockout extract, for all promoters, could be a result of other metabolites produced in the stationary phase (24 hr post inoculation) added to lag phase cultures. To verify these characteristics are a result of the *pvf* signaling pathway and not exclusive to the monalysin promoter, we studied other promoters. Similar differences in promoter activity are observed for other promoters (Figure 3.8), specifically the activity of the PSEEN5493-*lacZ* and PSEEN0973-*lacZ* hybrids is decreased in the strains containing an insertional mutation in *pvfC* (IM*pvfC*).
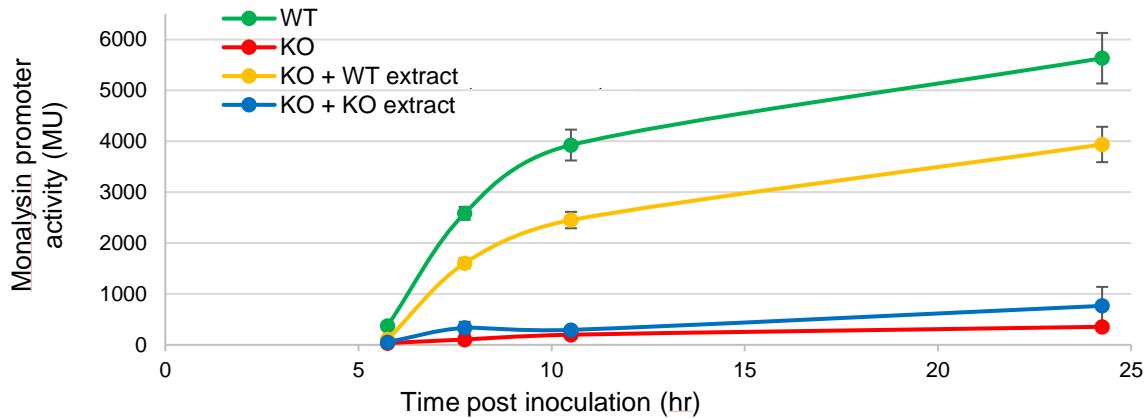
Figure 3.7. Deletion of pvfC affects expression of *mnl* in *P. entomophila*. Quantification of the β-galactosidase activity as a function of bacterial growth is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strain each containing P$_{mnl}$-*lacZ* reporter cassette. Activity of the KO reporter with addition of organic culture extracts from WT (yellow) and *ΔpvfC* strains (blue) is also shown. Extracts were added 4 hours post inoculation as 1 mL culture extracted added to 1 mL reporter culture to be assayed.
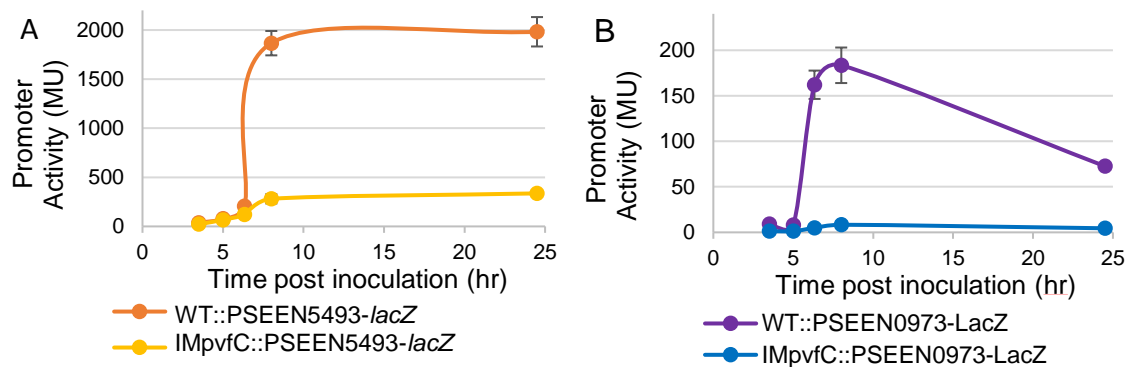
Figure 3.8. A *pvfC* mutation affects expression of other *P. entomophila* genes. Quantification of the β-galactosidase activity as a function of bacterial growth is shown for: (A) the translational fusion PSEEN5493–*lacZ* in WT (orange) and a *pvfC* insertional mutant (yellow), and (B) the translational fusion PSEEN0973–*lacZ* in WT (purple) and a *pvfC* insertional mutant (blue).

74

When the knockout reporter strain was co-cultured with the wildtype *P. entomophila,* we saw recovery of *mnl* promoter activity to wildtype levels (Figure 3.9A). No increase in activity was observed when the knockout reporter strain was co-cultured with genetic knockouts strains *ΔpvfC ΔpvfB* and *ΔpvfD* (Figure 3.9B). As both WT extracts from spent media and co-culturing with the wildtype strain recover promoter activity in *pvf* knockout reporter strains, we suggest the signaling molecule produced by the *pvf* biosynthetic cluster is secreted.

We sought to show that *pvf*-regulated promoters respond to the *pvf* signaling molecule in the concentration-dependent manner, another characteristic of quorum sensing molecules. We observed concentration-dependent increase in promoter activity in the knockout reporter strain with wildtype and overexpression extracts (Figure 3.10), as well as wildtype spent media and fractions from purifications of culture extracts containing the active PVF signal (Figure 3.11).
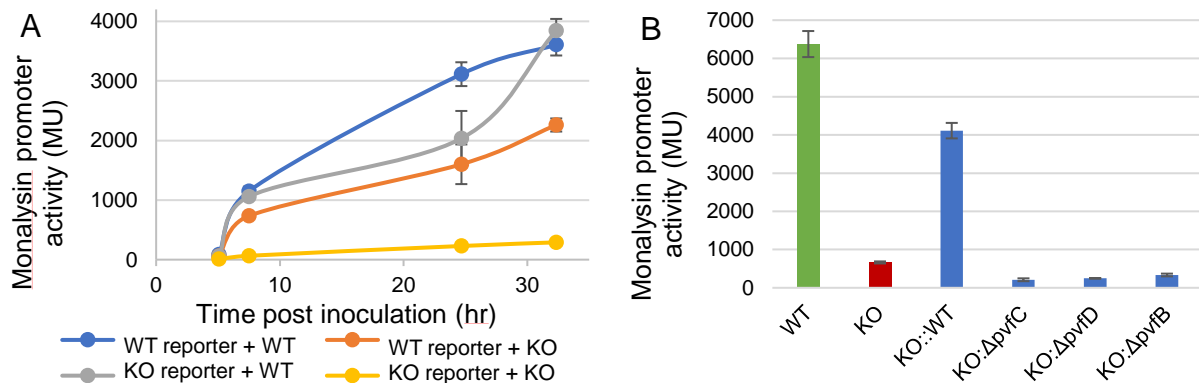
Figure 3.9. Coculture of *ΔpvfC* reporter with WT *P. entomophila* activates monalysin promoter activity. Quantification of the β-galactosidase activity (A) as a function of bacterial growth is shown for *P. entomophila* WT cocultured with *P. entomophila* WT (blue) or *ΔpvfC* (orange)and *pvf* deletion (KO) strain cocultured with *P. entomophila* WT (grey) or *ΔpvfC* (yellow), each containing P*mnl*-*lacZ* reporter cassette. Quantification of the β-galactosidase activity (B) after 24 hours of growth is shown for *P. entomophila* WT reporter strain (green), *ΔpvfC* reporter strain (red), and *ΔpvfC* reporter strain cocultured with WT or *pvf* deletion strains (blue).
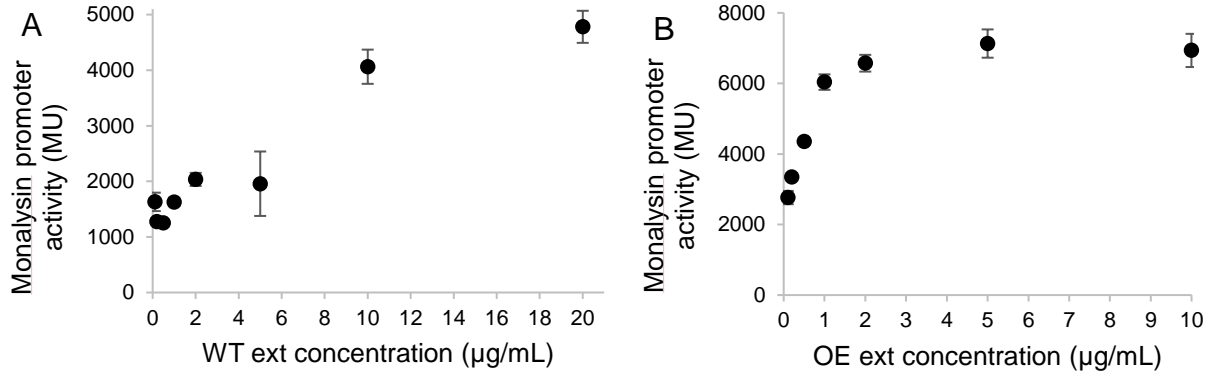
Figure 3.10. The monalysin promoter activity responds to addition of *pvf*-containing extracts in a concentration-dependent manner. Quantification of the β-galactosidase activity after 24 hours post inoculation is shown for *ΔpvfC* deletion strain containing P$_{mnt}$-*lacZ* reporter cassette with addition of increasing concentrations of (A) Wildtype *P. entomophila* extracts, or (B) *P. entomophila ΔpvfC* pPSV-*pvfABCD pvf* overxperssion (OE) extract. For WT extracts, 1 mL of culture extract added to 1 mL reporter culture is an approximate concentration of 77 µg/mL.
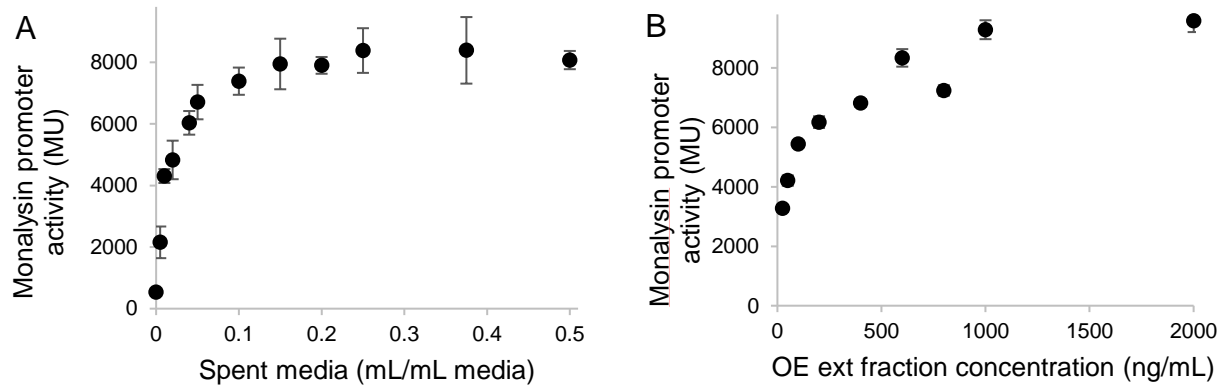
Figure 3.11. The monalysin promoter activity responds to addition of (A) WT spent media and (B) pvf-containing purified extract fractions in a concentration-dependent manner. Quantification of the β-galactosidase activity after 24 hours post inoculation is shown for *ΔpvfC* deletion strain containing P$_{mnt}$-*lacZ* reporter cassette with addition of increasing (A) amount of Wildtype *P. entomophila* spent media from a 24-hour culture in rich media and (B) concentrations of a second-round purification fraction of overexpression (OE) extract (R2F11).

We utilized the *pvf* promoter-reporter to determine if the PVF molecule was an auto-inducer. We first studied the activity of the *pvf* promoter as a function of cell density (Figure 3.12). An increase in activity was observed at the start of the exponential phase, suggesting *pvf* is activated in cell-density dependent manner. Addition of culture extracts from wildtype and overexpression strainss resulted in a five-fold increase in activity of the *pvf* promoter as assessed by the ONPG cleavage assay (Figure 3.13A). Addition of overexpression extracts results in activity levels at only eight hours similar to those seen at stationary phase in wildtype cultures with no induction. Autoinduction was observed with culture extracts from WT *P. entomophila*, *pvf* overexpression strain, and heterologous expression in *E. coli* (Figure 3.13B).
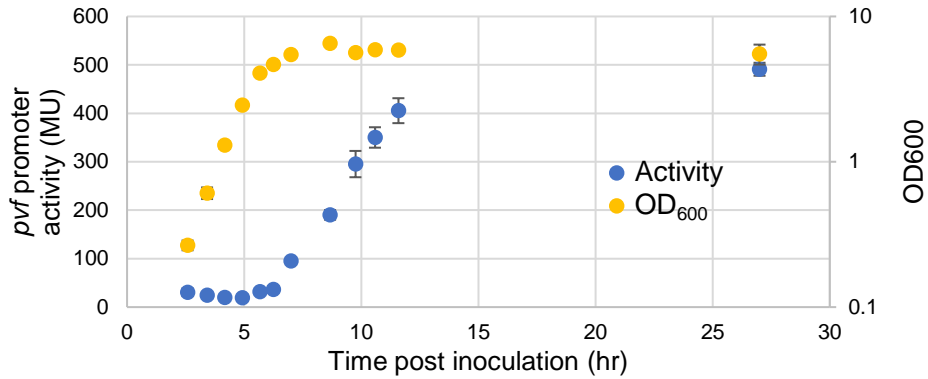
Figure 3.12. The *pvf* promoter activity is cell density-dependent. Quantification of the β-galactosidase activity (blue) as a function of bacterial growth (yellow) of *P. entomophila* WT with a P$_{pvf}$-*lacZ* reporter cassette.
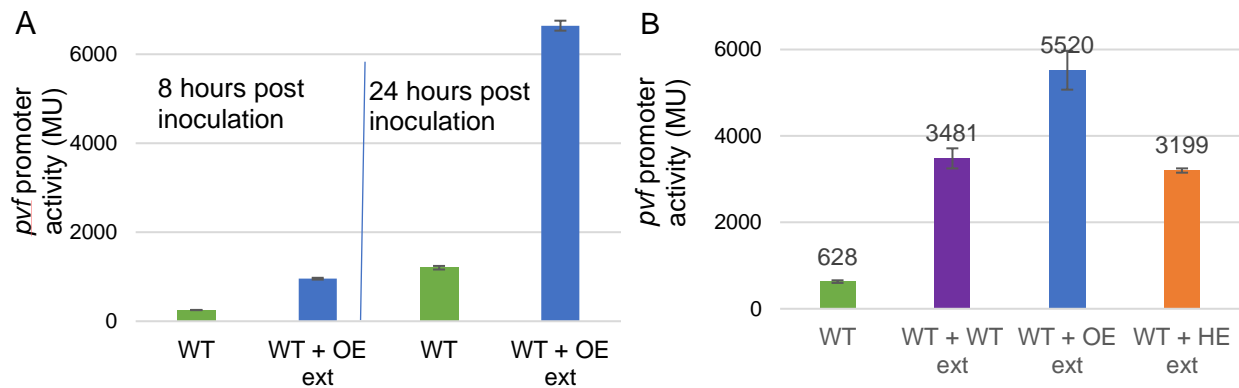
Figure 3.13. PVF small molecule(s) act as an autoinducer. Quantification of the β-galactosidase activity from *P. entomophila* WT strain containing the P$_{pvf}$-*lacZ* reporter cassette is shown(A) 8 and 24 hour post inoculation with (blue) and without (green) addition of 10 µg/mL extract from *P. entomophila* Δ*pvfC* + pPSV-*pvfABCD* overexpression (OE), and (B) Quantification of the β-galactosidase activity 24 hour post inoculation is shown for *P. entomophila* WT strain containing the P$_{pvf}$-*lacZ* reporter cassette (green) and with addition of 5 µg/mL culture extracts from *P. entomophila* WT (purple), *P. entomophila* OE (blue), or *E. coli* Bap1 + pPSV-*pvfABCD* heterologous expression (HE, orange) strains.

We also observe activation of the monalysin promoter with addition of culture extracts of *E. coli* and *P. aeruginosa* with heterologous expression of the *pvf* cluster (Figure 3.15−Figure *3.16*), suggesting the *pvf* biosynthetic genes are responsible for the production of the signaling molecule. To determine which *pvf* genes are required for the production of the signaling molecule, we added culture extracts of strains heterologously expressing partial *pvf* gene clusters to the monalysin reporter strain. Only extracts with expression *pvfB, C, and D* together activated the *mnl* promoter, which is consistent with previous studies showing that knockout of *pvfA* did not affect the *pvf* phenotype.
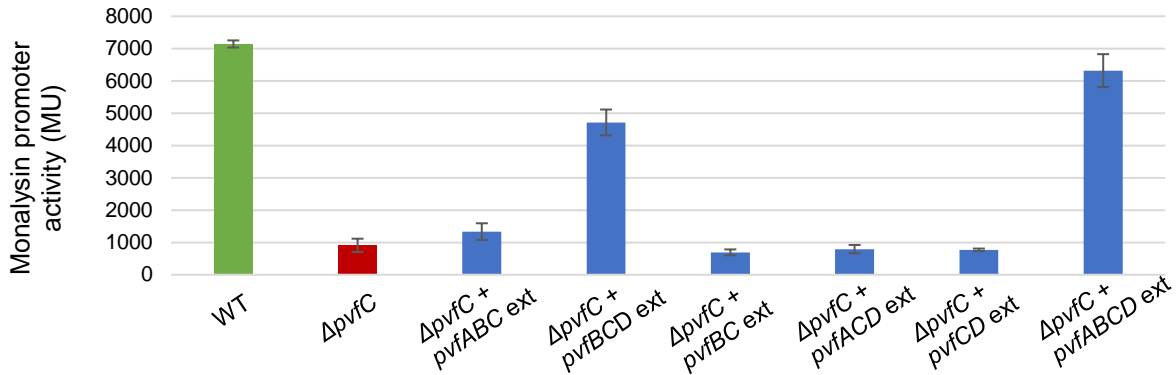
Figure 3.14. The monalysin promoter is activated by addition of *pvfBCD*-containing culture extracts. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strain each containing P*mnl*-*lacZ* reporter cassette. Activity of the KO reporter with addition of extracts heterologously expressing partial *pvf* clusters (blue) is also shown. Extracts were added 4 hours post inoculation. Data provide by Gina Morgan.

We have also studied the role of the *pvf* molecule in inter-species communication. Unpublished work in our group has shown that the NRPS adenylation domains of PvfC from >300 strains that harbor the *pvf* cluster activate either valine or leucine. The NRPS from the *pvf* pathway in *P. entomophila* activates valine, while the homologous NRPS from B*urkholderia cenocepacia* HI2424 and *Pseudomonas syringae* pv. *tomato* DC300 activates valine and leucine, respectively. We found that culture extracts of *B. cenocepacia* HI2424 activated the monalysin promoter in *P. entomophila,* even though HI2424 does not produce monalysin (Figure 3.15). Further, culture extracts from the HI2424 *pvf* deletion strain did not activate the monalysin promoter, suggesting monalysin activation is due to *pvf* expression in HI2424. To validate these results, we found that culture extracts of heterologous expression of the HI2424 *pvf* cluster in *E. coli* also activated the monalysin promoter (Figure 3.16). In contrast, extracts from *P. syringae* did not activate the monalysin promoter, suggesting the *pvf* cluster from *P. syringae* produces a different molecule from *P. entomophila* and each PVF molecule variant is recognized by a specific receptor (Figure 3.17). This result is consistent with observation that the PvfC homolog from *P. syringae* prefers leucine to valine.
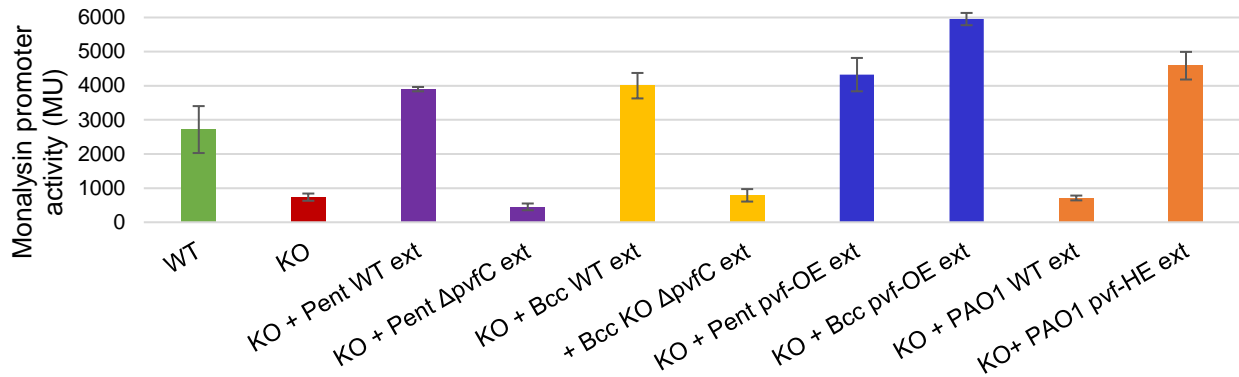
Figure 3.15. Addition of *pvf*-containing extracts activate the monalysin promoter. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strain each containing P*mnl*-*lacZ* reporter cassette. Activity of the KO reporter with addition of extracts from *P. entomophila* WT and *ΔpvfC* (purple)*, B. cenocepacia* WT and *Δpvf* (yellow)*, P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression (OE, blue), *B. cenocepacia* + pSCRha-*pvfABCD* overexpression (OE, blue), or *P. aeruginosa* PAO1 WT (red) and pPSV-*pvfABCD* heterologous expression (HE, red) is also shown. Extracts were added 4 hours post inoculation as 1 mL culture extracted added to 1 mL reporter culture to be assayed.
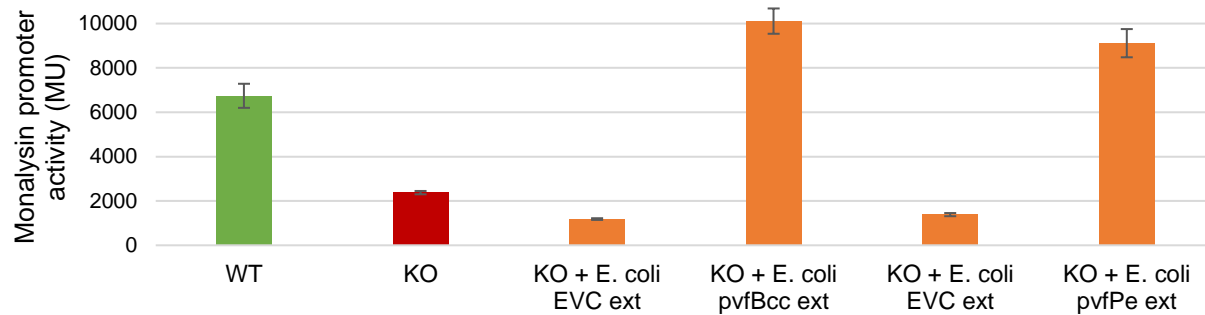
Figure 3.16. Extracts from *E. coli* with heterologously-expressed *pvf* activate the monalysin promoter. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strain each containing P*mnl*-*lacZ* reporter cassette. Activity of the KO reporter with addition of extracts from *E. coli* (orange) with pSCRha empty vector control (EVC), pSCRha-*pvfBCC* heterologous expression, pPSV empty vector control (EVC) and pPSV-*pvfPE* heterologous expression is also shown. Extracts were added 4 hours post inoculation as 1 mL culture extracted added to 1 mL reporter culture to be assayed.
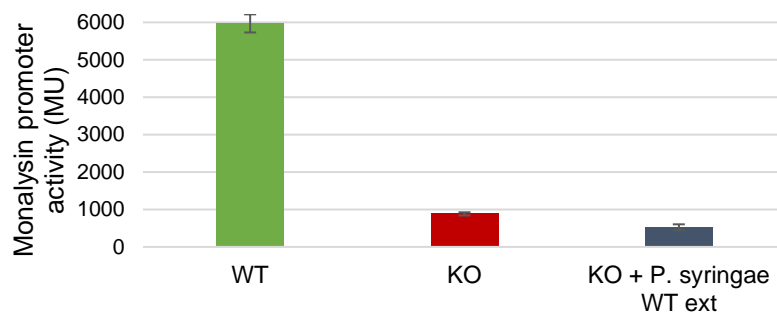
Figure 3.17. Extracts from *P. syringae* do not activate the monalysin promoter. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strain each containing P*mnl*-*lacZ* reporter cassette. Activity of the KO reporter with addition of culture extracts of *P. syringae* (blue) is also shown. Extracts were added 4 hours post inoculation as 1 mL culture extracted added to 1 mL reporter culture to be assayed.

Proteomics

To determine the global regulation of the *pvf* signaling pathway, we analyzed the secreted protein expression profiles of wildtype, *pvfC* deletion, and *pvf* overxpression cultures in rich media by proteomics. Proteins from spent media (Figure 3.18) were analyzed by liquid chromatography-tandem mass spectrometry and identified using the Uniprot database for *P. entomophila* L48. The 1823 putative proteins identified were quantified with label free quantification (LFQ) intensities and Perseus statistical program was used to process the proteomics data. [128]

By analysis of variance (ANOVA), 506 proteins were differentially expressed between the three data sets (p-value < 0.1, Figure 3.19Figure *3.20*). For most of the proteins with differential expression in *ΔpvfC* cultures, expression was restored with inducible expression of *pvf* in the *ΔpvfC* strains (*pvf* overexpression strain). Specifically, 94 proteins showed a two-fold or greater decrease in *ΔpvfC* cultures compared to WT (and was restored in the overexpression strain (LFQ $_{Log2}$int WT-KO >1, WT-OE <0, Appendix 22), and 76 showed a two-fold or greater
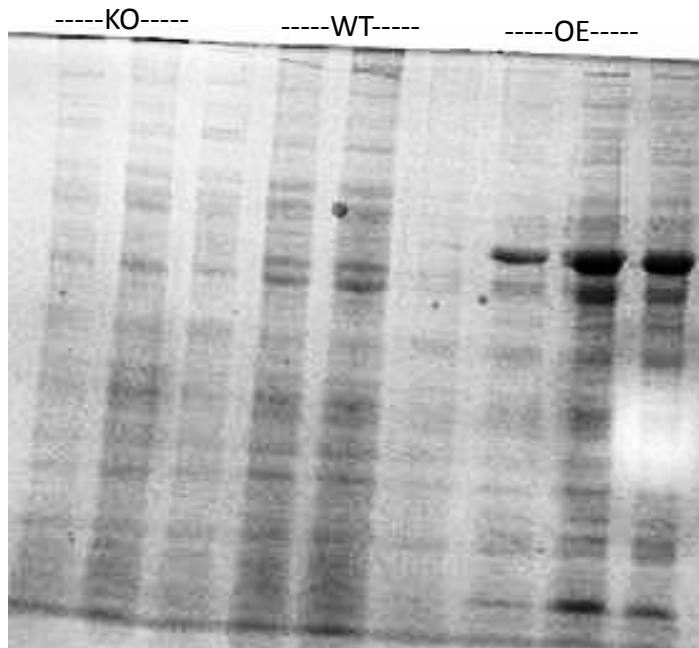
Figure 3.18. Deletion of *pvfC* in *P. entomophila* alters the expression of secreted proteins. SDS PAGE gel of proteins from spent media from 24-hour cultures of *P. entomophila* wildtype (WT), *ΔpvfC* (KO), and *ΔpvfC* + pPSV-*pvfABCD* (OE) in rich media.

increase in *ΔpvfC* cultures compared to WT and was restored in the overexpression strain (LFQ

Log2int WT-KO <−1, WT-OE >0, Appendix 23). We focused on these 180 proteins. One limitation

to our methods is the potential for cells to lyse upon centrifugation and leave cellular proteins in

the supernatant. In our data, we did identify a number of cytosolic proteins, but our stricter p-

value threshold (0.01) eliminated many of these proteins due to higher standard deviations.
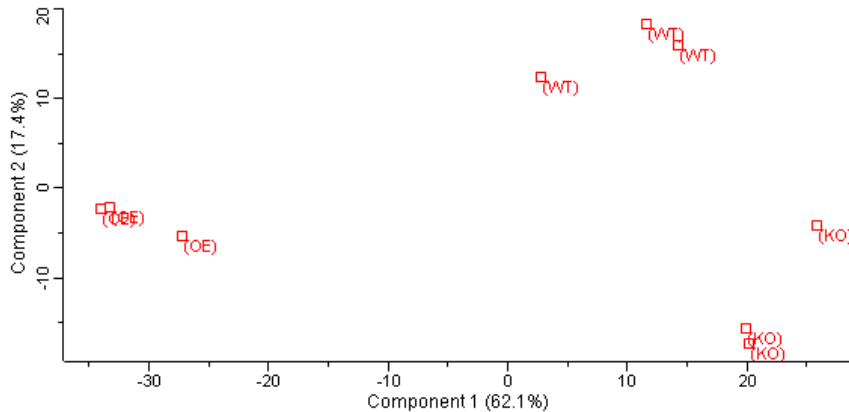
Figure 3.19. Principal component analysis (PCA) clusters biological replicates together, separating the three data sets of *P. entomophila* wildtype (WT), *ΔpvfC* (KO), and *ΔpvfC* + pPSV-*pvfABCD* (OE).
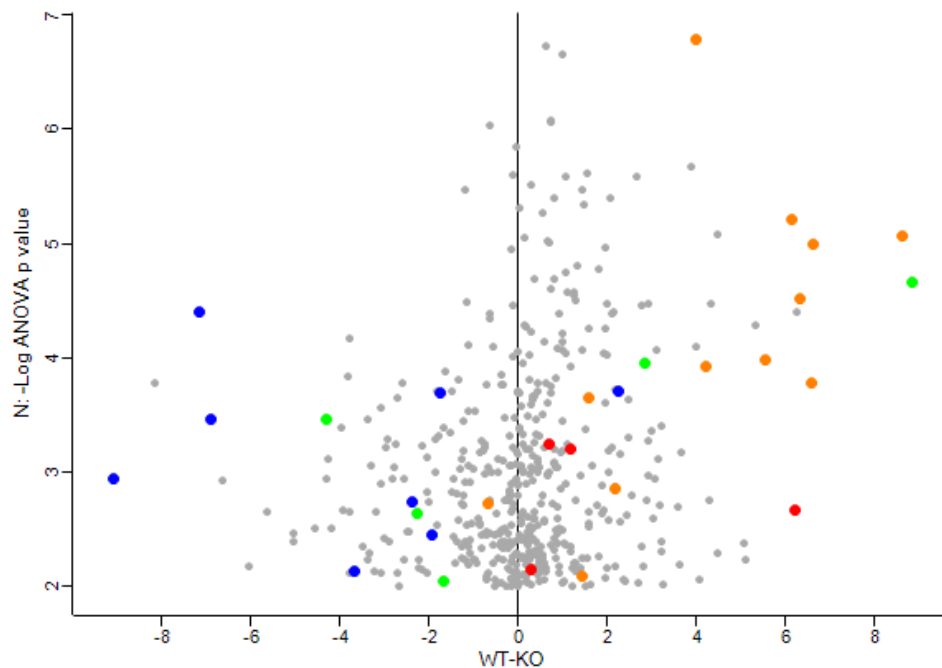
Figure 3.20. 506 proteins are differentially expressed between *P. entomophila* wildtype (WT), *ΔpvfC* deletion strain (KO), and *ΔpvfC* + pPSV-*pvfABCD* overexpression strain (OE). Scatterplot of difference in mean Log2intensity between WT and KO and -LogANOVA p-value, each point is a protein with ANOVA p <0.01. Motility-related proteins are in blue, TonB receptors in green, secretion system proteins in orange, toxins proteases and chitinases in red.

81

Proteins regulated by the *pvf* cluster are predicted to have a range of biological roles based on identification by Uniprot. They include regulatory proteins, such as alginate regulatory protein AlgP and the transcriptional regulator ArgR, which has been established as a virulence regulator in previous genetic studies.[55] In addition, deletion of *pvf* decreases siderophore receptors, toxins, proteases and proteins from the type VI secretion system (T6SS).

One of our top proteomic hits, based on fold change difference between wildtype and *pvf* deletion cultures, was for the pore-forming toxin monalysin, which has previously been shown to be regulated by the *pvf* pathway and is essential for full infection against *Drosophila*.[40] This is consistent with our previous findings that the activity of the *mnl* promoter was reduced in the *pvf* deletion reporter strain compared to the WT reporter strain (Figure 3.7). Similar, the overexpression of *pvf* activates the expression of monalysin, similar to the activation of the *mnl* promoter with *pvf*-expressing extracts. This suggests the small molecule product(S) of *pvf* act as the signaling molecule to activate the monalysin promoter, and the regulation most likely occurs at the transcriptional level.

We also observed a difference in expression of the mostly highly abundant secreted protein, the metalloprotease AprA. AprA is particularly upregulated in the *pvf* overexpression strain, as observed by SDS-PAGE (Figure 3.18). Both of these proteins have been implicated in fly infection; monalysin acts as a pore forming toxin that causes gut damage, and AprA is a protease known to cleave immunity proteins and cleave the premature pre-monalysin after export from the cell. However, using a promoter-reporter assay with the *AprA* promoter, we do not observe a significant difference in AprA promoter activity with deletion of *pvf*.[38] In addition, both the wildtype and *pvf* deletion strains of *P. entomophila* exhibit protease activity. Previous studies have implicated AprA has solely responsible for protease activity on milk agar plates [60]. Therefore, we suggest that AprA is not regulated by *pvf* on the transcriptional level, rather the change in protein expression might be as a result of other regulation mechanisms, possibly related to changes in monalysin production.

Included in our highest upregulated proteins (by fold-change) were proteins involved in type VI secretion (T6SS), a syringe-like protein complex that injects effector proteins into the host cell. This includes the outer membrane-attached tssJ, the 'plunger' tssB/C, rhs, hcp and vgrR genes. Of the vgrR genes that are upregulated, each have the same architecture, with a DUF2345 or COG4253 domain. It has been suggested that VgrR proteins that contain this particular domain have unique function or targets compare to other effector transporters, but at the moment this function has not be elucidated.[131, 132] Type VI secretion in *P. entomophila* is particularly interesting due to the unique lack of a type III secretion system. In addition, the type VI secretion system has been shown not to be important for fly infection.[40] Deletion of ttsJ (PSEEN0535) did not affect the ability of *P. entomophila* to infect adult flies. This suggests that *pvf* control factors not only involved in pathogenesis, but also potential factors involved in competition with the host or neighboring microbes.

In addition to proteins upregulated by *pvf,* our proteomics data indicate that proteins involved in motility are downregulated by *pvf.* Deletion of *pvf* results in an increase in flagellar proteins. To verify the effect of the *pvf* pathway on cell motility we compared the spread of *P. entomophila* WT and *pvf* deletion strains on soft agar plates (Figure 3.21). The *pvf* deletion strain exhibited significantly increased cell motility, and motility was reduced by genetic complementation, when the *pvf* cluster was expressed on the psv-*pvfABCD* plasmid.
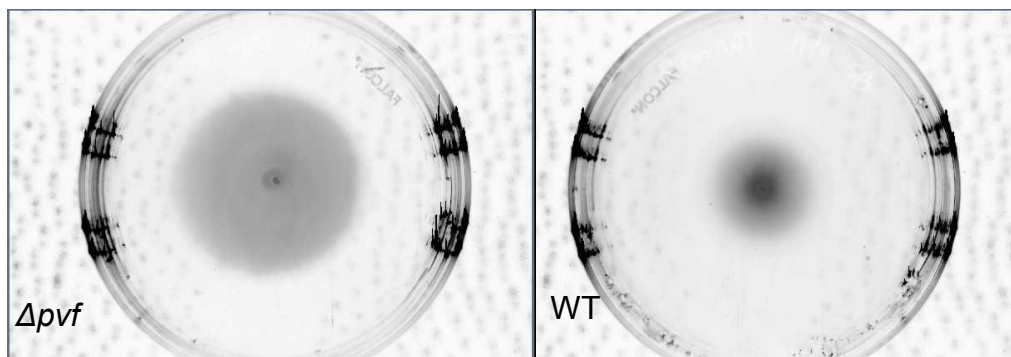
Figure 3.21. Deletion of *pvf* increases the motility of *P. entomophila.* Culture growth on soft agar after 24 hours. Representative data, experiments were completed in triplicate.

Small molecule secretome changes

In addition to upregulation of secreted proteins, we examined the effect of the *pvf* cluster on secondary metabolite production. The proteins that synthesize these metabolites are cytosolic and would not be identified through our proteomics experiment. Through a droplet collapse assay, we found that the *pvf* deletion mutant had reduced surfactant production (Figure 3.22). Previously, the lipopeptide entolysin has been shown to be responsible for hemolytic and biocontrol activity by *P. entomophila*.[56] β-galactosidase activity from the translational fusion *etlB-lacZ* (PSEEN3045-*lacZ*) was significantly reduced with an insertional mutation into *pvfC* compared to WT reporter (Figure 3.23). Another phenotype we observed in the *pvf* deletion strain was reduced siderophore production. In our proteomic analysis, the siderophore transporter BauB showed decreased expression in the *pvf* deletion mutant. We hypothesized that by comparing ethyl acetate extracts of WT and *pvf* deletion cultures, we could identify molecules with surfactant and siderophore characteristics that are regulated by the *pvf* pathway.
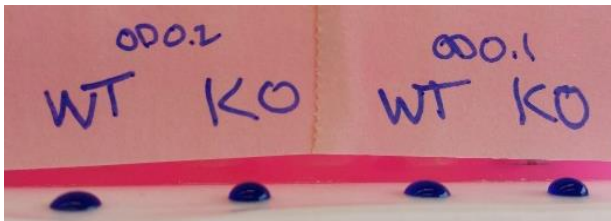
Figure 3.22. A *pvf* deletion mutant has reduced surfactant product by droplet collapse assay. 20 microliter droplets on parafilm of 24-hour culture of *P. entomophila* WT and *pvf* deletion strains at OD_{600} 0.2 and OD_{600} 0.1 are shown, visualized using methylene blue.
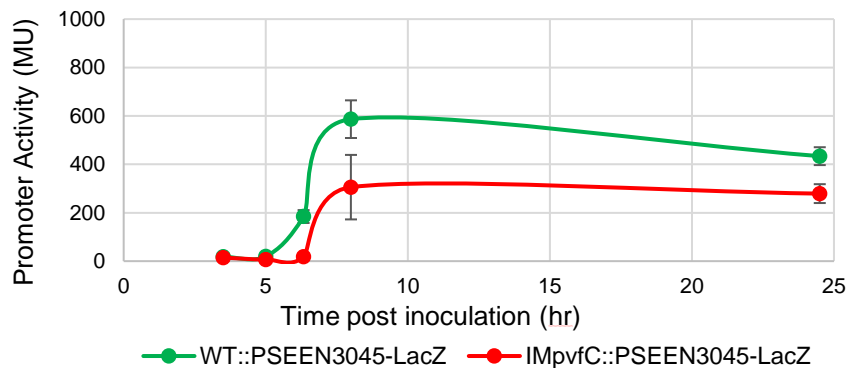
Figure 3.23. A *pvfC* mutation affects expression of entolysin biosynthetic genes in *P. entomophila*. Quantification of the β-galactosidase activity, as a function of bacterial growth, of the translational fusion PSEEN3045–*lacZ* in WT (green) and a *pvfC* insertional mutant (red).

To identify *pvf*-regulated small molecules, the supernatant from a 24-hour culture of *P. entomophila* wildtype (WT) or *Δpvf* (KO) in 4 L LB was extracted with ethyl acetate, and the metabolites from the organic layer were separated by preparatory HPLC. By uv, we observe a significant number of metabolites that were uniquely present in extracts from cultures of the WT strain (Figure 3.24). We took these fractions from purifications of culture extracts of the WT and KO strains and compared their siderophore and surfactant properties at 100 µg/mL. Several of these extract fractions exhibit surfactant and/or siderophore properties (Figure 3.25Figure *3.26*). For siderophore activity, the extract fraction with incubated with CAS dye; the blue dye changes colors to purple for and catechol-like siderophores. Surfactant activity was measured by droplet collapse assay. We are currently purifying these fractions to a single compound for chemical characterization.
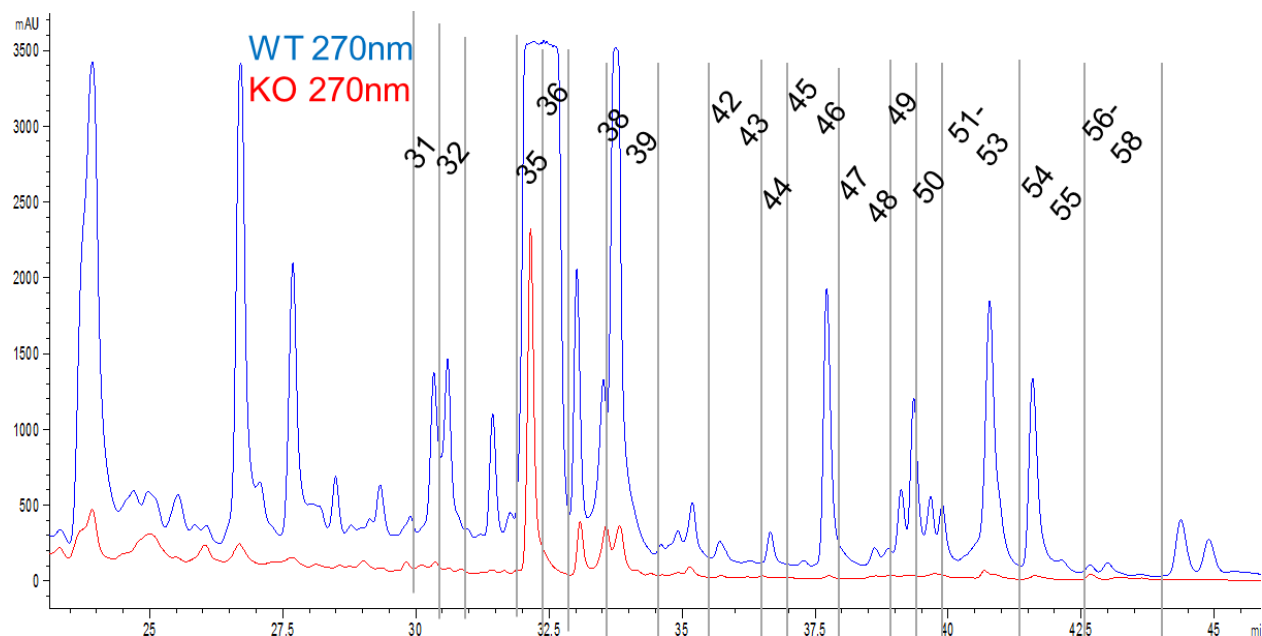
Figure 3.24. HPLC traces at 270nm show a reduction or elimination in the production of several metabolites in the culture extract of the *pvf* KO strain (red) compared to culture extracts from the WT strain (blue). Fractions were collected as indicated for further analysis.
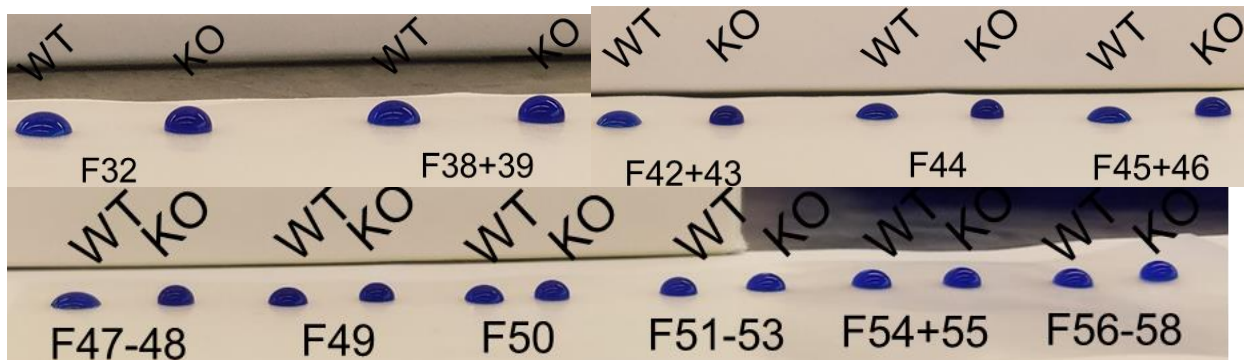
Figure 3.25. Several fractions show production of surfactant metabolites uniquely in WT cultures (F32, F42−47, F49−50). Droplet collapse assays on parafilm tested at 100 µg/mL and stained with methylene blue.

Row 1: F31, <u>32</u>, 33-34, 35, <u>36</u>, <u>37</u>

Row 2: F38-39, 40-41, 42-43, 44, <u>45-46</u>, 47-48

Row 3: F<u>49</u>, 50, <u>51-53</u>, <u>54-55</u>, 56-58, 59-60
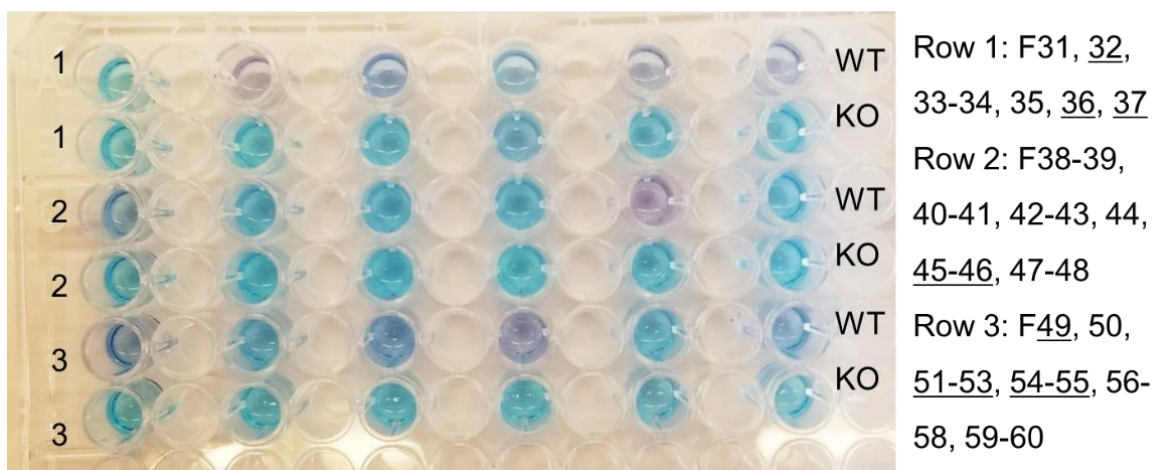
Figure 3.26. Several purified fractions contain iron-binding metabolites uniquely in culture extracts from the WT compared to the *pvf* deletion strain (F32, 36−37, 45−46, 49, 51−53). Purified fraction were tested at a concentration of 100 µg/mL. With CAS dye, changes from blue to purple indicate catechol-like siderophores.

3.4 Discussion

In conclusion, we found that the *pvf* pathway produces a quorum sensing molecule that globally regulates proteins and small molecules. The products of *pvf,* through native expression (WT), overexpression in the native strain (OE), or heterologous expression (HE) activate several promoters in a cell-density and concentration-dependent manner. This is a novel QS molecule, and is particularly interesting in *Pseudomonas entomophila,* which harbors no AHL synthetases.

Similar to other pathogens, such as *Pseudomonas aeruginosa, P. entomophila* might control virulence pathways through a multi-faceted regulation network. For instance, some genes, such as monalysin, are regulated by both *pvf* and the two-component system GacA/S,

while others, such as *pvf,* are regulated by *pvf* but not GacA. Proteomics data also suggests a connection between *pvf* signaling and the transcriptional regulator AlgR.

The *pvf*-regulated proteins of the secretome suggests that the *pvf* signaling pathway plays a number of significant biological roles (Figure 3.27). Proteomics data suggests *pvf* is involved in regulation of the pore forming toxin monalysin, chitinase ChiC, and protease AprA. These proteins play a direct role in causing disease to the bacterial host. In addition, AprA has been shown to cleave immunity proteins from the host's defense. The Type VI secretion system (T6SS) has not been shown to be important for infection in *P. entomophila,* but without a Type III secretion system we hypothesis the T6SS must plays a vital role to inject effector proteins into the host or neighboring bacteria. With the activation of a number of virulence factors, we also observe the deactivation of motility through the *pvf* signaling pathway. This could be for the redistribution of resources as, with only one phage tail, motility is a very resource intensive activity for *P. entomophila.*
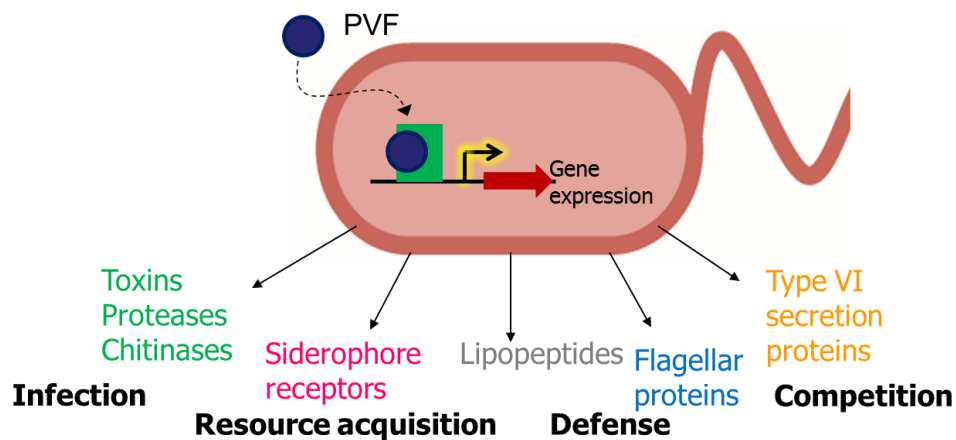
Figure 3.27. The *pvf* signaling pathway regulates proteins with important roles in virulence and competition.

Finally, we observe a change in the production of small molecules, several which has surfactant or metal binding properties. *P. entomophila* is predicted to produce three, or possibly more, surfactants, and only one (entolysin) has been experimentally verified. Similarly, *P. entomophila* is predicted to product the siderophores pseudomonine and pyroverdine, but

studies have suggested the pseudomonine pathway produces several products *in vitro.* Entolysin has exhibited hemolytic activity and is importance for infection. The biological role of siderophores such as pyroverdine in iron-scavenging is well documented.

Together, we have identified a novel quorum sensing system that regulates a significant number of proteins and small molecules in the animal pathogen *Pseudomonas entomophila.* The *pvf* pathway is present in over 300 strains, many which exhibit pathogenetic or biocontrol activities. We hypothesize that *pvf* regulates similar biological processes in these other bacterial strains. Investigations of a novel signaling pathway will provide insights into bacterial signaling, particularly in strains that do not harbor the homoserine lactone QS system. Finally, our work may identify new targets for anti-virulence therapy.

Chapter 4 Efforts toward identifying the active PVF signaling molecule

4.1 Introduction

After the discovery of penicillin by Alexander Fleming in 1928, it was another decade before the compound was isolated and its structure elucidated. Fleming sent his *Penicillium* mold to anyone who requested it in hopes that they might isolate penicillin for clinical use.[133] In 1939, Howard Florey and coworkers scaled up mold cultures and finally purified penicillin from extracts. Since then, activity-guided isolation of natural products is widely used to identify natural products that exhibit biological activities, and a wealth of new techniques have been developed over the past two decades.

Activity-driven discovery involved separating complex mixtures of metabolites by chromatography and testing fractions for biological activity of interest. Active fractions are further purified until a single compound or set of compounds is identified to be responsible for the activity. In recent years, activity-guided purification, combined with genetic manipulations and characterization by high-resolution mass spectrometry, has been an effective strategy for natural product discovery. Methods are now available for isolation, purification, and structural elucidation of natural products from more complex organisms, such as marine invertebrates, or from their symbiotic bacteria.[134] For example, activity driven purification was used in combination with mass spectrometry and metagenomics in the discovery of the anti-HIV compounds, the divamides, and the elucidation of their structures and potential biosynthesis.[135] These compounds were isolated from a symbiotic cyanobacterium *Prochloron didemnid* that live in tunicates, a rich reservoir for bioactive compounds,[135] but detection of these natural products is difficult due to the small size of their producer. High resolution mass spectrometry (HR-MS)

has a much lower limit of detection (pmol) than conventional mass spectrometry and allows for identification of even minute quantities of natural products.

Due to their low abundance in culture, signaling molecules possess significant challenges for isolation and structure elucidation. For example, >1000 liters of *Streptomyces* cultures were necessary for identification of the *Streptomyces* hormone avenolide.[67] Fortunately, several methods to track bioactivity have been developed and are amendable to various chemistries. One possible method could use the ligand binding activity of a quorum sensing (QS) molecule with its receptor, which helps bind autoregulator response elements to DNA sequences. Ligand binding can influence or inhibit binding of the receptor to DNA.[100] This method requires identification of the receptor specific for the QS molecule, which is unknown for the *pvf* signaling pathway. The second method is to use transcriptional or translational reporters. These reporters, when activated, result in the production of a protein, the level of which can be quantified to determine the activity of the promoter of interest. We have constructed reporters using the promoter for the *pvf* biosynthetic pathway, P_{pvf}-*lacZ,* and the promoter for the toxin monalysin known to be regulated by the *pvf* pathway, P_{mnl}-*lacZ.* The level of LacZ proteins can be quantified in a number of ways to estimate the activity of the promoter of interest, as described in Chapter 3. We used these reporters to guide the purification and identification of the active signaling molecule produced by the *pvf* biosynthetic pathway.

4.2 Material and methods

General methods and methods for inducible expression of *pvf* in *P. entomophila* are described in Chapter 2 (pg 21).

Methods for determining activity using promoter-reporter strains are described in Chapter 3 (pg 61).

Activity-guided purification of active signaling molecule from extracts

Growth conditions. Eight 1 L LB-gentamycin cultures were inoculated with 2 mL overnight cultures of *P. entomophila* L48 *ΔpvfC* + pPSV35-*pvfABCD*. The cultures were incubated at 30 °C and induced with 1 mM IPTG when OD_{600} reached 0.3. For wildtype (WT) and *Δpvf* deletion (KO) strains of *P. entomophila,* eight 1 L LSLB cultures were inoculated with 2 mL overnight cultures and incubated at 30 °C. If specified, addition of autoinducer for WT cultures was added four hours post inoculation.

Overexpression extractions. After 24 hours of growth at 30 °C, the culture supernatant was separated from bacterial cells by centrifugation. The supernatant was adjusted to a pH of 5.0 and extracted three times with a one third volume of DCM. The organic layers were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The extracted metabolites were resuspended in 5 mL of methanol, transferred to a scintillation vial, dried, and stored at −20 °C.

WildType and knockout extractions. Supernatant was extracted twice with half volume of ethyl acetate. The aqueous layer was adjusted to pH of 5.0 with HCl and extracted twice with half volume of DCM. The DCM organic layers were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The metabolites extracted were resuspended in 5 mL of methanol, transferred to a scintillation vial, dried down, and stored at −20 °C.

Activity-guided purification. Metabolites were purified from the large-scale extractions using preparatory high-performance liquid chromatography (Agilent PrepStar). Dried extracts were resuspended in 80 % ACN/water (1 mL/100 mg extract) and injected on a Phenomenex Luna C18 column in 1 mL injections. Compounds of interest were separated using mobile phase A (water, 0.1% TFA) and mobile phase B (acetonitrile, 0.1% TFA) over the gradients described below. All absorbance in the range of 190−450 nm were recorded. Fraction were combined separately from multiple rounds and dried under high vacuum. Active Round 1 fractions were injected at 20 mg/mL for Round 2 of purification, Round 2 fractions were injected at 4 mg/mL for Round 3 of purification.

Round 1: a gradient of 5% B for 5 min, 5−40% B for 15 min, 40−95% B for 15 min, and 95% B for 6 min, at a flow rate of 15 mL min^{-1}. Fractions were collected every 30 seconds during min 16−22.

Round 2: a gradient of 5% B for 5 min, 5−20% B for 5 min, 20−35% B for 15 min, 35−95% B for 2 min, and 95% B for 5 min, at a flow rate of 15 mL min^{-1}. Fractions were collected (2b) every 30 seconds during min 14−20 or (2c) every 20 seconds during min 15−19.

Round 3: (a) an isocratic gradient at 18% B for 20 min, collect every 30 sec during min 2−20. (b) a gradient of 5% B for minutes, 5−20% B for 10 min, 20% B for 8 min, 20−95% B for 2 min and 95% B for 6 minutes, at a flow rate of 15 mL min^{-1}. Fractions were collected every 30 seconds during min 15−22.

Methods for structural characterization are described in Chapter 2 (pg 24).

Synthesis of ImC

ImC was synthesized by Gina Morgan in two steps from 2-aminopyridine and 1-methyl-2-bromo-3-butanone (Scheme 4.1). [136] A sample of 94 mg (1 mmol, 1 equiv.) of 2-aminopyridine was dissolved in 4.6 mL MeOH and 100 mg (1.2 mmol, 1.2 eq) NaHCO$_3$ was

added. A sample of 125 µL 1-methyl-2-bromo-3-butanone (165 mg, 1 mmol, 1 eq) was added

dropwise. The solution was heated at reflux for 20 hours. Solvent was removed under vacuum

to afford **1** (160 mg, 1 mmol, 100%). Yellow solid.
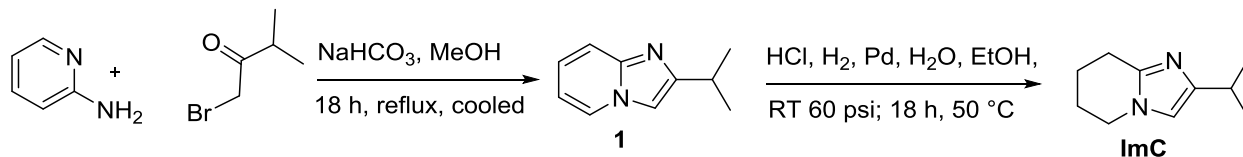
Scheme 4.1. Synthesis of 2-isopropyl-5,6,7,8-tetrahydroimidazo[1,2-a]pyridine, or ImC (imidazole cyclane)

150 mg of **1** was resuspended in 10 mL ethanol/HCl (0.0833 µL). The reaction flask was

placed in a pressure reactor and sealed. Reactor was filled with hydrogen, pressurized, and

flushed with hydrogen three times. Reactor was then pressurized to 60 PSI. The reaction was

heated to 50 °C and stirred for 20 hr. Solvent was removed under vacuum, yielding a

brown/yellow oil. Crude oil was purified using "Round 2" prepHPLC method described above,

collecting fractions from 17.5 min to 18.5 min to afford ImC (60 mg, 0.37 mmol, 44%). Light

yellow oil. ^{1}H NMR (600 MHz, D_2O) δ 6.95 (s, 1H) 4.03 (t, 2H), 2.96 (m, 1H), 2.92 (t, 2H), 2.02

(m, 2H), 1.96 (m, 2H), 1.23 (d, 6H).

4.3 Results

Activity-driven purification of *pvf* overexpression extracts

Construction of an overexpression system in the native *P. entomophila* to identify the biosynthetic product(s) of the *pvf* cluster was described in Chapter 2. Comparative metabolomics comparing a *pvfC* knockout of *P. entomophila* and genetic complementation with the overexpression vector led to the identification of many small molecules produced by *pvf*-encoded enzymes, several of which belonged to a novel family of pyrazine *N*-oxides. Our studies suggest that these PNOs are most likely intermediates or shunt products of the biosynthetic pathway, because neither purified nor synthetic compounds could recover *pvf* signaling activity of the *P. entomophila pvfC* deletion mutant.

To identify the active PVF signaling molecule, we transitioned to an activity-guided discovery approach using the promoter-reporter strains described in Chapter 3. Monalysin promoter activity is significantly decreased in a *pvfC* deletion (KO) promoter-reporter strain in comparison to the wildtype (WT) promoter-reporter strain. Promoter activity was restored to wildtype levels when the KO reporter strain was complemented with dichloromethane (DCM) culture extracts of WT strain, but not with culture extracts from *pvf* deletion strains. I hypothesize that the active molecule(s) can be isolated and identified using reverse phase high-performance liquid chromatography (HPLC) coupled with activity profiling of purified fractions and characterization by LC-HRMS and NMR. First, culture extracts of strains that produce active PVF molecules (either WT or stains containing inducible expression plasmids) are purified and each fraction is added to culture of the KO reporter strain. Fractions that activate the monalysin promoter are further fractionated until the active molecule(s) are isolated as single compounds.

For activity-driven purification, culture extracts (>8 L) of *P. entomophila* overexpressing the *pvf* cluster were injected onto the preparatory HPLC column and separated on a water and acetonitrile gradient containing 0.1% TFA. Active fractions (Figure 4.1A, Round 1 Fractions 3−4

and Figure 4.1B, Round 2 fractions 10−11) were identified and further purified in the next round

with increasingly shallower gradients of water and acetonitrile, ending with an isocratic gradient

at 20% acetonitrile in water (Figure 4.1C).

Figure 4.1. Bioactivity-driven purification of the PVF signaling molecule from overexpression extracts. Activity of fractions from rounds 1−3 purification from extracts of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression (OE) cultures. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strain each containing P$_{mnl}$-*lacZ* reporter cassette. Activity of the KO reporter with addition of purified fractions (blue) is also shown. Fractions were added 4 hours post inoculation at (A) 1 µg/mL (Round 1), (B) 250 ng/mL (Round 2) and (C) 2 µL of 500 µL resuspension per fraction (Round 3).

To verify the isolated compound(s) were product(s) of the *pvf* biosynthetic pathway, we separated extracts of a large-scale culture for *P. aeruginosa* PAO1 + pPSV-*pvfABCD* heterologous expression (HE) strain using the same methods as the overexpression extract. Using the promoter reporter assays, we identified the same active fractions as those from the overexpression strain the first and second rounds of purifications, suggesting the active molecule was the same or similar in the overexpression and heterologous expression strains (Figure 4.3). Fractions from culture extracts of *pvfC* deletion strain did not activate the monalysin promoter, suggesting signaling activity requires the *pvf* biosynthetic pathway. Similar to WT extracts, we observe a dose-dependent promoter activity with addition of an active fraction from the overexpression strain (Round 2, Fraction 11, Figure 4.2). Finally, we verified the activity using another *pvf*-regulated promoter, the PSEEN0973-*lacZ* translational fusion described in Chapter 3. The same fractions purified from overexpression cultures restored *lacZ* expression in the *pvfC* mutant to WT levels, consistent with previous reports and our observations described in Chapter 3 that the signaling molecule can regulate multiple promoters (Figure 4.4).
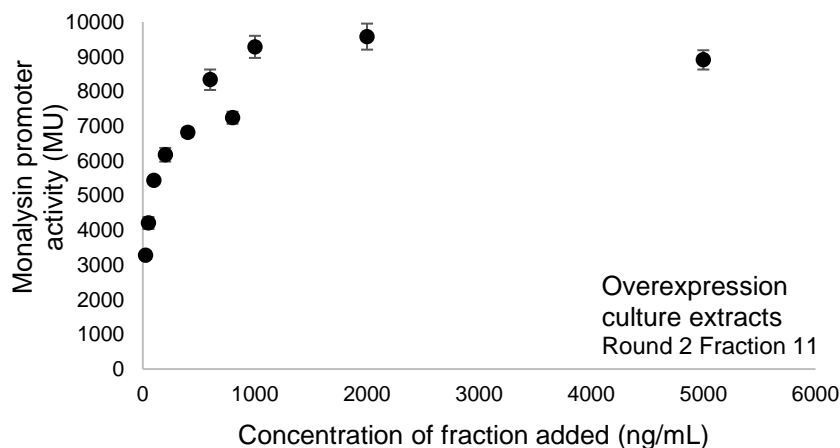
Figure 4.2. Active fractions from overexpression culture extracts activate the monalysin promoter in a dose-dependent manner. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *ΔpvfC* deletion strain containing P$_{mnl}$-*lacZ* reporter cassette, with addition of various amounts of Round 2 Fraction 11 from overexpression culture extracts (OE R2F11).
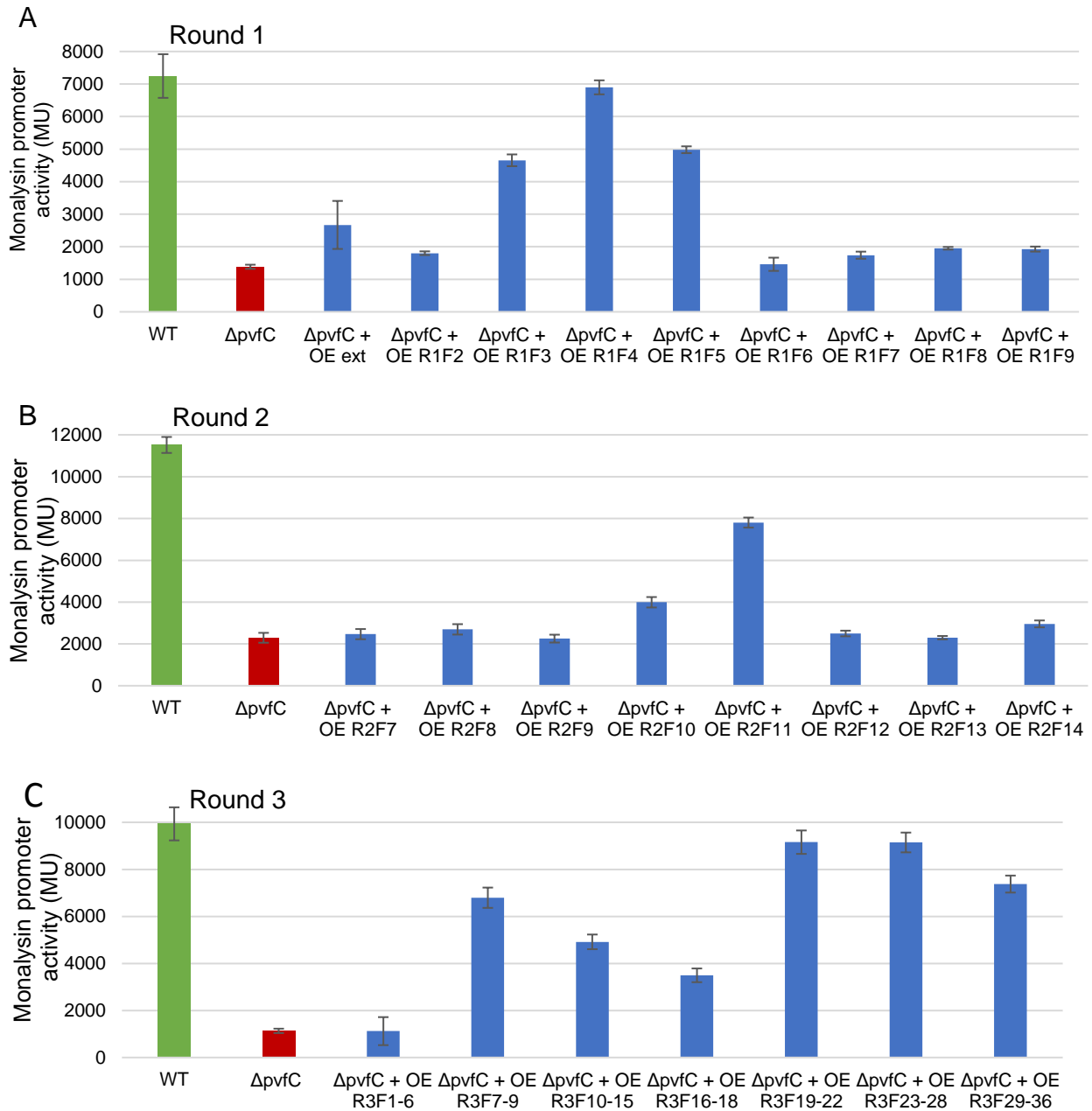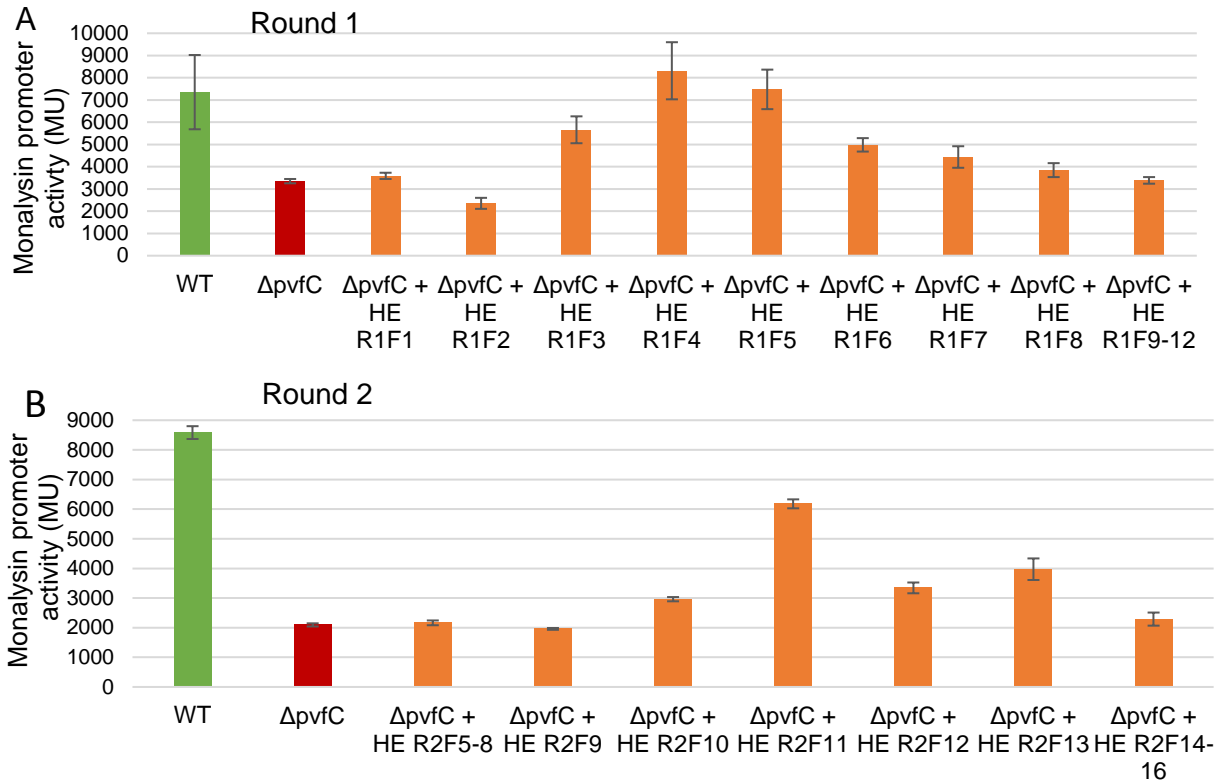
Figure 4.3. Bioactivity-driven purification of the PVF signaling molecule from heterologous expression culture extracts. Activity of fractions from purification rounds 1−2 from culture extracts of *P. aeruginosa* PAO1 + pPSV-*pvfABCD* heterologous expression (HE) strain. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (*ΔpvfC*, red) strain each containing P$_{mnl}$-*lacZ* reporter cassette. Activity of the KO reporter with addition of purified fractions (orange) is also shown. Fractions were added 4 hours post inoculation at 1 μg/mL, (A), Round 1, (B), Round 2.
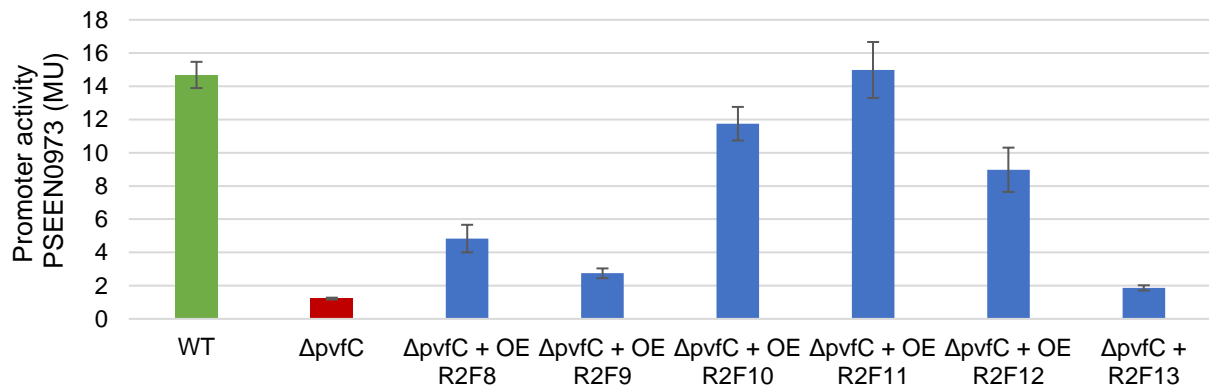
Figure 4.4. The same fractions purified from overexpression cultures that activate the monalysin promoter activate other *pvf*-dependent promoters. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvfC* insertional mutant (red) each containing a PSEEN0973–*lacZ* hybrid. Activity of KO reporter with addition of purified fractions (blue) is also shown. Fractions from purification round 2 of culture extracts of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression (OE) strain were added 4 hours post inoculation at 1 μg/mL.

After three rounds of purification (Figure 4.1), we isolated a single compound (**4**) with a mass of 164 (Figure 4.5Figure *4.6*), with a purity of approximately 90%. We characterized the structure of this compound by NMR experiments, in combination UV spectroscopy and mass spectrometry. The structure was assigned as 2-isopropyl-5,6,7,8-tetrahydroimidazo[1,2-a]pyridine, or imidazole cyclane (ImC,

Table *4.1*−Table *4.2*, Appendix 24−37). We conducted 2-D COSY and HMBC experiments to distinguish ImC from its 3-isopropyl isomer catharsitoxin C, which has been previously isolated from nature.[137] Although similar in structure to catharsitoxin C, ImC is a novel natural product to the best of our knowledge. Isotopic labelling data using $[^2H_{1-8}]_{DL}$-valine suggests incorporation of at least one valine (Figure 4.7), but the biosynthetic origin of the hydrated pyridine is unclear.
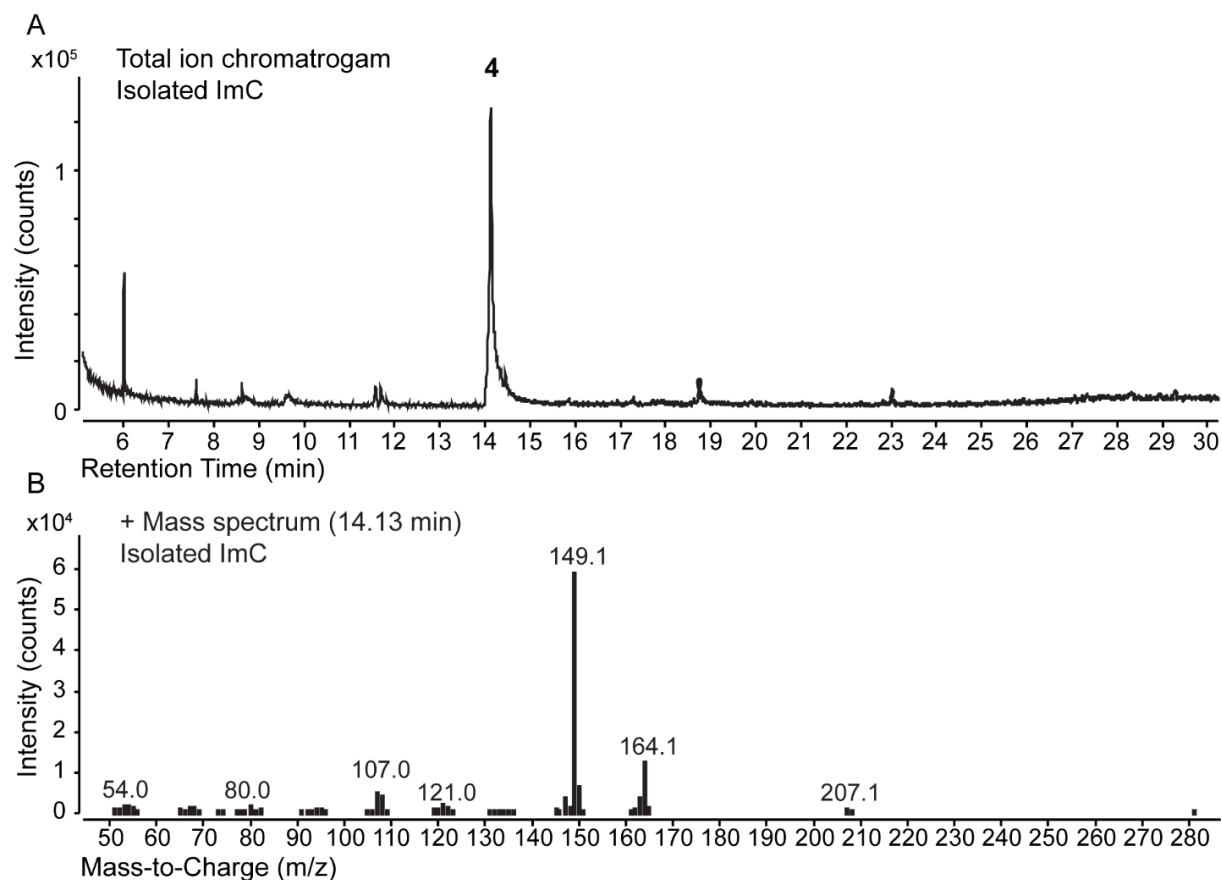
Figure 4.5. Activity-driven isolation and identification of a new metabolite (**4**) produced by *pvf*-encoded enzymes. The metabolite is purified through fractionation of culture extracts of the *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression (OE) strain. Gas chromatography trace suggests a single product is about 90% pure with a mass of 164.1.
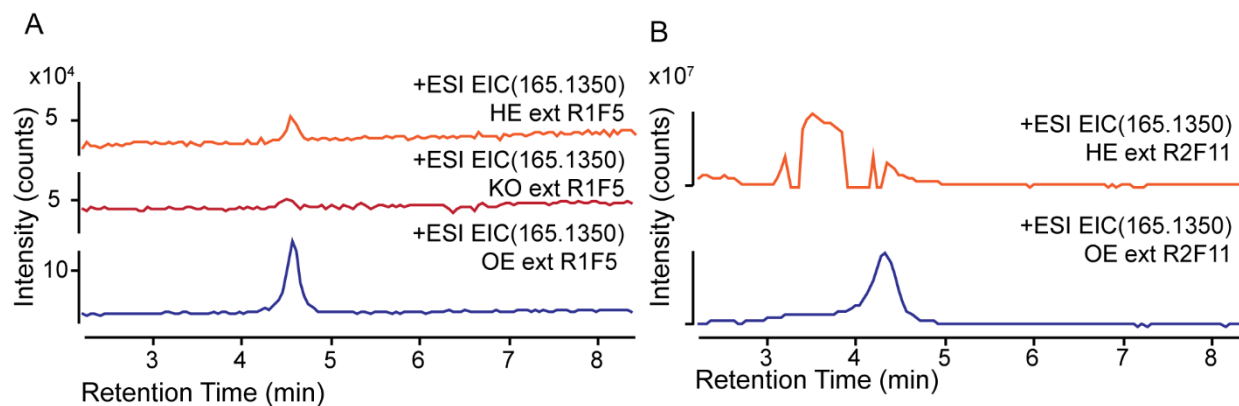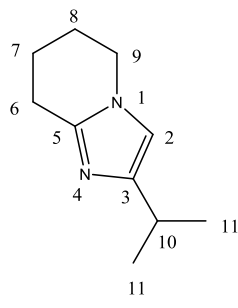
Figure 4.6. Identification of a single compound from bioactivity-driven purification of extracts from cultures of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression strain (OE, blue) *P. aeruginosa* PAO1 + pPSV-*pvfABCD* heterologous expression strain (HE, orange) but not cultures of *P. entomophila ΔpvfC* knockout strain (KO, red). Extracted ion chromatorgrams of *m/z* = 165.135 from (A) purification round 1, fraction 5 (active in OE, HE, not KO) and (B) purification round 2, fraction 11. The difference between the retention times is due to sample overloading.

Table 4.1. NMR assignments and correlations of isolated ImC (**4**) in D$_2$O. 1-N and 4-N cannot be resolved on the (^{1}H, ^{15}N) HMBC (δ_N 206-216 ppm).

POSITION	Δ_C	PROTON	Δ_H (J_{HH}[Hz])	COSY	(^{1}H, ^{13}C) HMBC
1-N					
2	115.1	2C-H	6.92		10
3	143.6				2, 9, 6, 7, 8
4-N					
5	138.8				2, 10, 11
6	20.6	6C-H$_2$	2.89	7	7, 8, 9
7	17.9	7C-H$_2$	1.92	6, 8	9, 8, 6
8	21.1	8C-H$_2$	1.99	7, 9	7, 6, 9
9	45.2	9C-H$_2$	4.00	8	7, 8
10	24.53	CH(CH$_3$)$_2$	2.93	11	11
11	20.6	CH$_3$	1.20	10	10, 11

Table 4.2. NMR assignments and correlations of isolated ImC (**4**) in MeOD.

POSITION	Δ_C	PROTON	Δ_H (J_{HH}[Hz])	COSY	(^{1}H, ^{13}C) HMBC	Δ_N	(^{1}H, ^{15}N) HMBC
1-N						-208 (172)	2, (7), 6, 9
2	117.0	2C-H	7.13		10		
3	145.4				2, 9, 6, 7, 8		
4-N						-214(166)	2
5	140.5				2, 10, 11		
6	22.2	6C-H$_2$	3.03-2.94	7	7, 8, 9		
7	19.5	7C-H$_2$	2.05-1.99	6, 8	9, 8, 6		
8	22.6	8C-H$_2$	2.11-2.05	7, 9	7, 6, 9		
9	46.7	9C-H$_2$	4.09 [5.9]	8	7, 8		
10	26.4	CH(CH$_3$)$_2$	3.03-2.94	11	11		
11	21.8	CH$_3$	1.30 [6.9]	10	11		

Table 4.3 NMR assignments and correlations of synthetic ImC (**4**) in D$_2$O

POSITION	Δ_C	PROTON	Δ_H (J_{HH}[Hz])	COSY	(^{1}H, ^{13}C) HMBC	Δ_N	(^{1}H, ^{15}N) HMBC
1-N						-208 (172)	2, 6, 8, 9
2	115.2	2C-H	6.95		9, 10		
3	143.7				2, 6, 7, 8, 9		
4-N						-214(166)	2, 6, (10)
5	138.8				2, 10, 11		
6	20.6	6C-H$_2$	2.92	7	7, 8, 9		
7	17.9	7C-H$_2$	1.96	6, 8	9, 8, 6		
8	21.1	8C-H$_2$	2.02	7, 9	7, 6, 9		
9	45.2	9C-H$_2$	4.03	8	2, 7, 8		
10	24.6	CH(CH$_3$)$_2$	2.96	11	11, 12		
11	20.7	CH$_3$	1.23	10	11, 12		
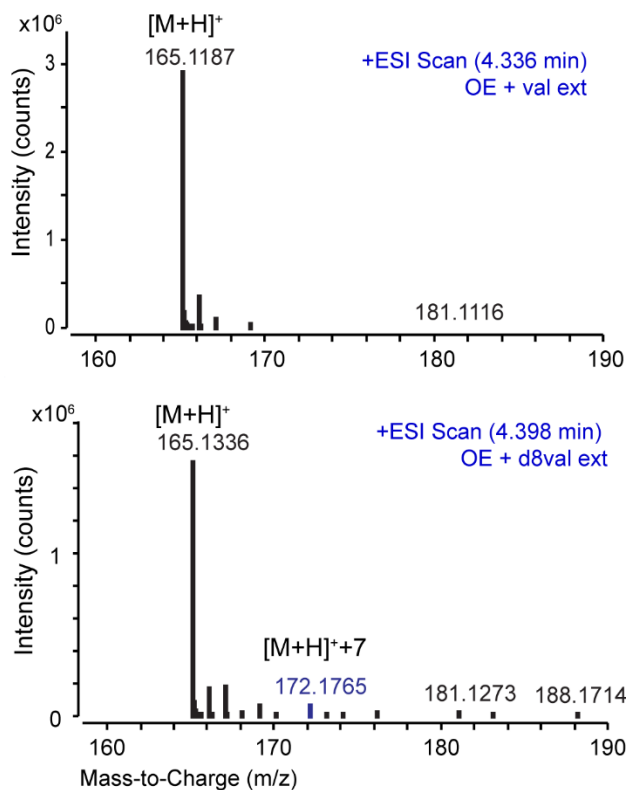
Figure 4.7. Incorporation of $[^2H_{1-8}]_{DL}$-valine into ImC by LC/HRMS analysis. Black traces correspond to unlabeled ImCs. Blue traces correspond to mass shifts observed for ImCs with $[^2H_{1-8}]_{DL}$-valine supplementation. Cultures were grown in minimal M9 media, the rate of $[^2H_{1-8}]_{DL}$-valine incorporation is comparable to PNOs in these conditions.

To confirm the structure assignment for ImC, Gina Morgan developed a novel route to synthesized ImC in two steps from 2-aminopyridine and 1-methyl-2-bromo-3-butanone (

Scheme 4.1). MS and NMR spectra of the synthetic compound were identical to that of the isolated compound, confirming the structure assignments (Figure 4.8, Appendix 28−43). Unfortunately, the synthetic compound was could not recover signaling activity in the reporter strain containing the *pvf* knockout (Figure 4.10). One explanation was that the synthetic molecule is different from the isolated molecule. For example, the isolated molecule could contain a *N*-oxide, which could be lost during mass spectrometry detection and would not significantly change the proton spectrum of the molecule. To test this hypothesis, we conducted (^{1}H, ^{15}N)-HMBC experiments and observed that both the purified ImC and synthetic standard

101

had identical ^{15}N shifts (Table 4.2Table *4.3*, Appendix 37 and 43) and therefore were likely

structurally identical. A second possibility is that ImC requires another species in isolated

sample for biological activity, such as coordination to metal ions. To test this possibility, we

analyzed isolated ImC sample by direct injection mass spectrometry, which better detects metal

complexes than chromatographic separations. However, we did not detect any metal complexes

(Figure 4.9). A third possibility is that *pvf* signaling activity requires synergistic action between

ImC and other compounds. To test the possibility, we combined ImC with dPNO that we

previously identified (Chapter 2) and found that combinations of ImC and dPNO did not recover

signaling activity in the monalysin reporter strain containing the *pvf* knockout (Figure 4.11)
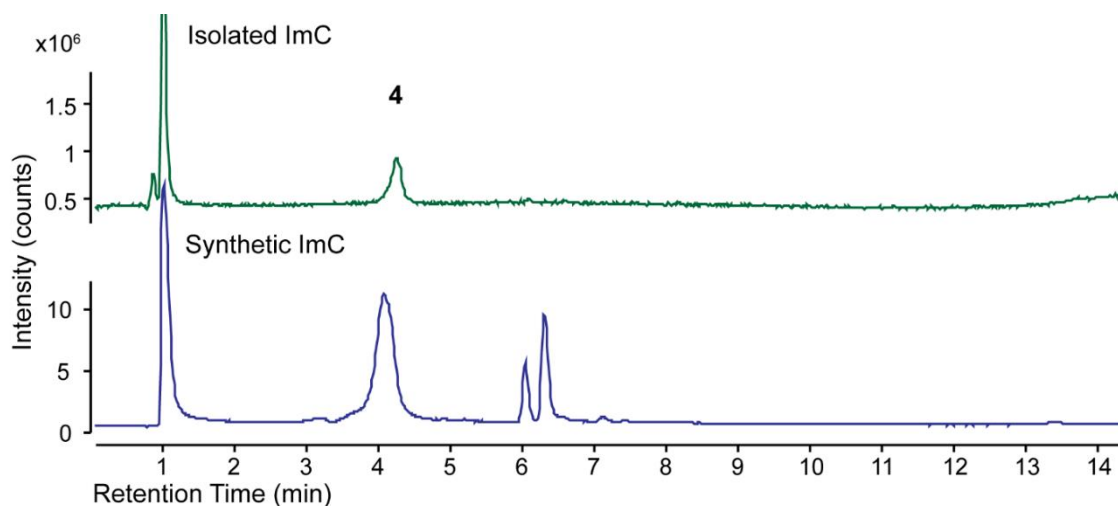
Figure 4.8. Metabolite **4** isolated from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* (green) exhibits an identical retention time to ImC synthesized from 2-aminopyridine and 1-methyl-2-bromo-3-butanone (blue). LC-HRMS analysis of biologically isolated and synthetic samples of ImC, total ion chromatograms. Retention time of ImC is 4.2 min.
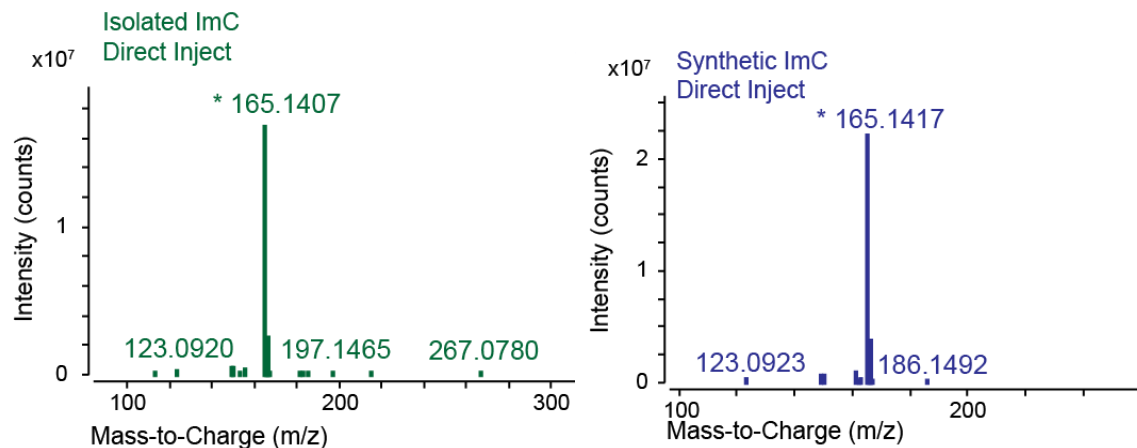
Figure 4.9. ImC isolated from cultures of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* (green) overexpression strain exhibits an identical mass spectrum to synthesized ImC (blue). HR-MS analysis of biologically isolated and synthetic samples of ImC, using a direct inject method to detect metal chelation.
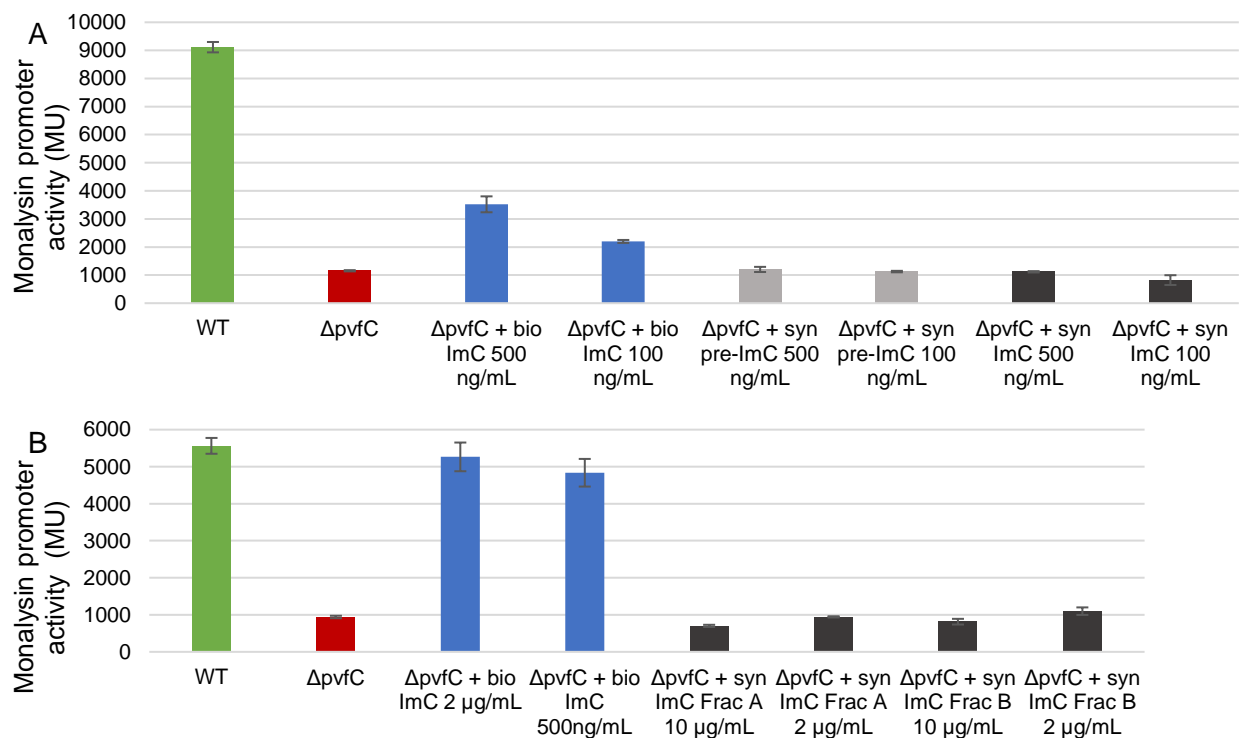
Figure 4.10. Synthetic imidazole cyclane (ImC) does not activate the monalysin promoter. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (*ΔpvfC*, red) strain each containing P$_{mnl}$-*lacZ* reporter cassette. Activity of the KO reporter with addition of (A) biologically isolated (blue), synthetic preImC (**1** in Scheme 4.1, light grey) and ImC (grey) at 500 and 100 ng/mL and (B) biologically isolated (blue) or synthetic ImC (grey) at 10 μg/mL or 2 μg/mL is also shown. Biologically isolated ImC is from purification round 3, fractions 19-22 of culture extracts of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression strain. Two purifications of synthetic ImC (Frac A and B) were tested. Compounds and fractions were added 4 hours post inoculation.
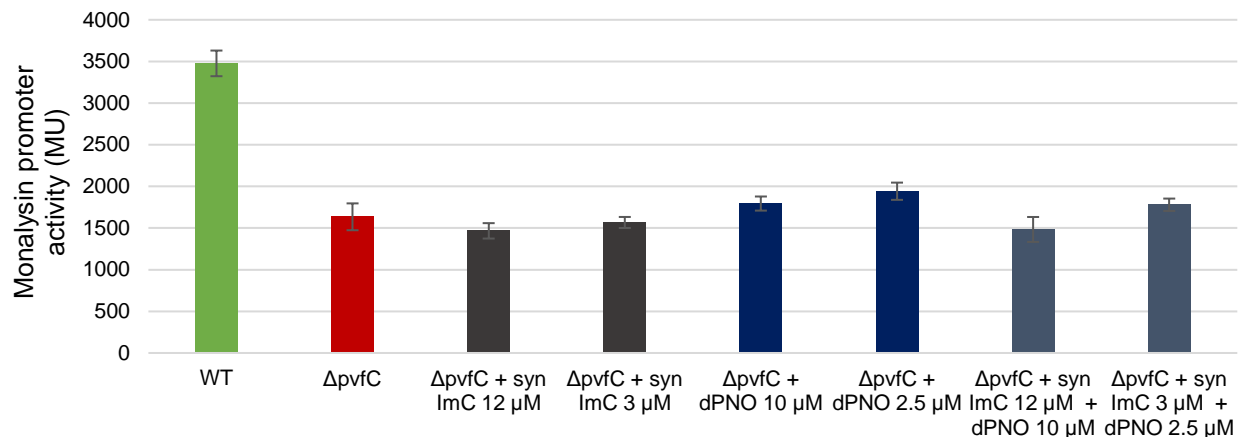
Figure 4.11. Combinations of dPNO and ImC do not activate the monalysin promoter. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (*ΔpvfC*, red) strain each containing P$_{mnl}$-*lacZ* reporter cassette. KO reporter with addition of synthetic ImC (grey), isolated dPNO (dark blue), or combination of ImC and dPNO (grey-blue) at 2.5−12 µg/mL is also shown. Biologically isolated compounds are from culture extracts of *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression strain. Compounds were added 4 hours post inoculation.

To identify genes required for production of ImC and the active compound that remains to be identified, we used overexpression constructs of partial gene clusters described in Chapter 2. Using extracts of these partial gene clusters in the heterologous host PAO1, we determined that *pvfB* and *pvfC* were required for production of ImC by LC-MS (Figure 4.12). We also found that addition of biological purifications of ImC showed lower levels of activity in the *pvfD* knockout reporter strains compared to the *pvfC* and *pvfB* knockout reporter strains (Figure 4.13), suggesting that ImC may be an intermediate prior to modification by PvfD. Therefore, we hypothesized that overproduction of intermediates/shunt products could be due to limited expression of *pvfD,* which is essential for production of the active signaling molecule based on genetic studies.[38]
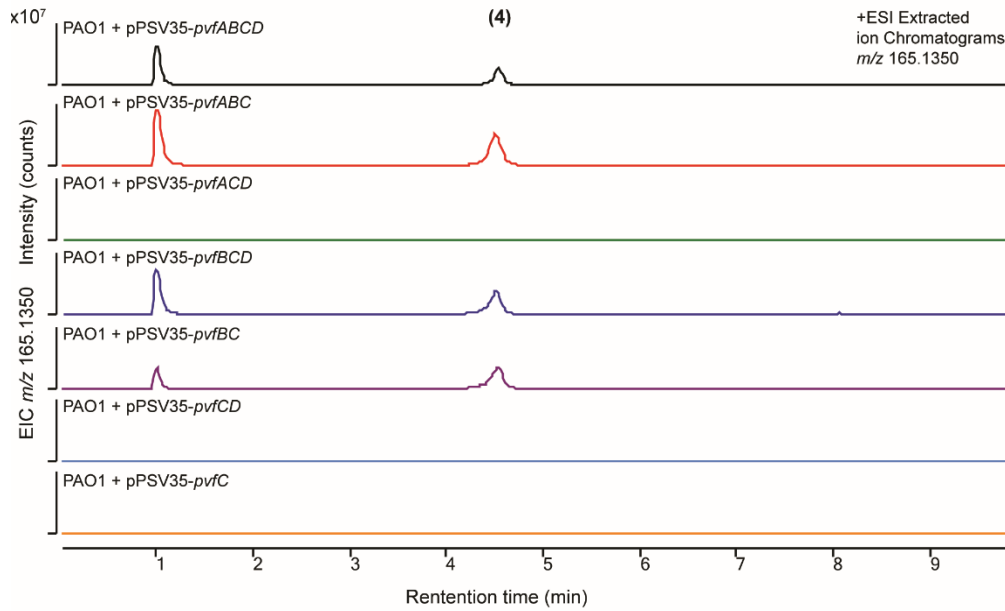
Figure 4.12. The *pvfB* and *pvfC* genes are both necessary and sufficient for the production of ImC (**4**). Liquid chromatography traces with extracted ion chromatograph at *m/z* 165.1350 are shown for the culture extracts of *P. aeruginosa* PAO1 overexpressing different, partial *pvf* constructs.
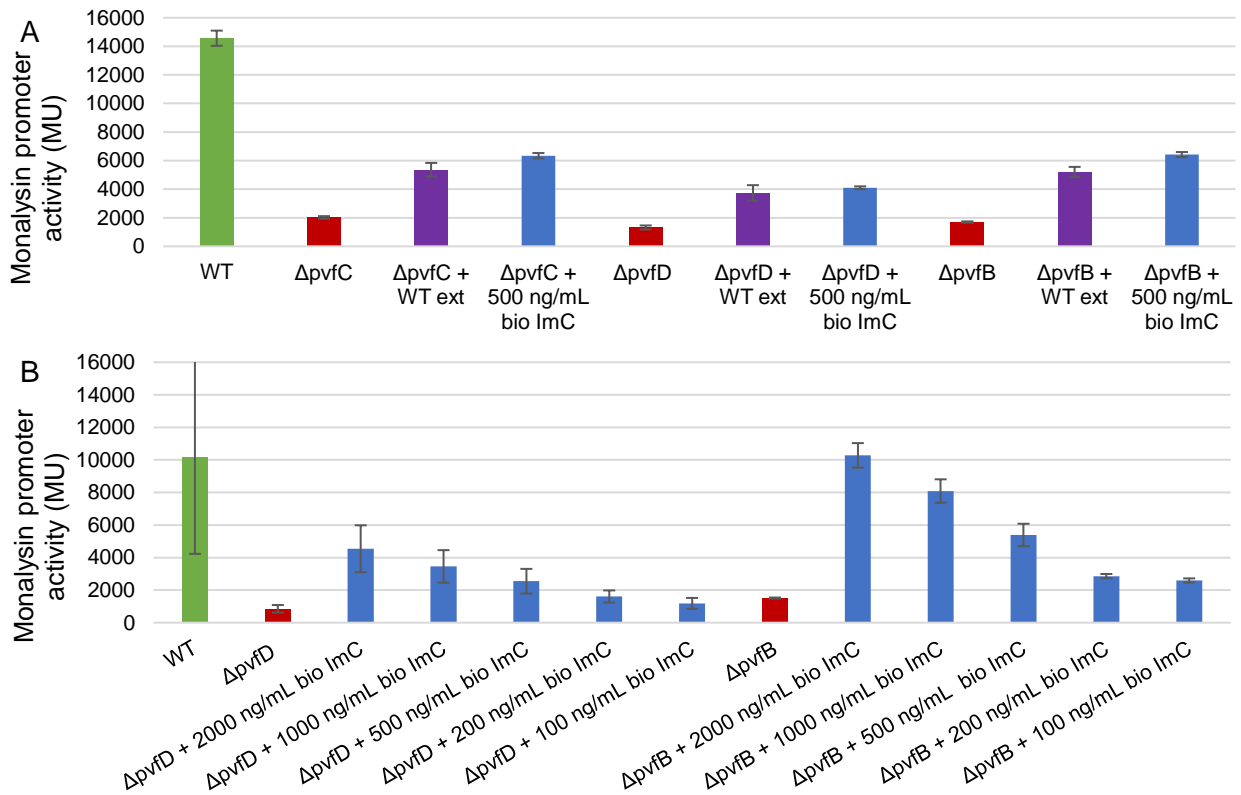
Figure 4.13. Biological purifications of ImC show lower activity in the *pvfD* knockout reporter strain. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing P*mnl*-*lacZ* reporter cassette. (A) The *ΔpvfC, ΔpvfD,* and *ΔpvfB* monalysin reporter strains with addition of *P. entomophila* WT extract (purple) or biologically isolated ImC (OE R3F19−22, blue) at 500 ng/mL and (B) the *ΔpvfD,* and *ΔpvfB* monalysin promoter-reporter strains with addition of biologically isolated ImC (OE R3F19−22, blue) at various concentrations is also shown. Extracts and fractions were added 4 hours post inoculation.

105

Activity driven purification of culture extracts of wildtype and *Δpvf* mutant strains for comparative metabolomics

In an effort to identify the active molecule, we tested several approaches to optimize its production and purification. By modifying the inducible expression system, we could increase the production of active molecule compared to intermediates and shunt products. By using the wildtype (WT) strain, we could eliminate potential intermediates and shunt products produced by the overexpression strain, but the caveat is that the WT produces significantly less of the active molecule. Insights from these experiments combined will assist in isolation and characterization of the active *pvf* signaling molecule.

Using our promoter-reporter assays, we observed that in order for the molecule to be extracted into the organic layer, dichloromethane or ethyl acetate, the pH of the supernatant had to be lowered to 5.0 (Figure 4.14). Organic extracts from supernatant that was unmodified (at pH 8.5) or adjusted to a pH of 12 with sodium hydroxide did not restore wildtype activity to the monalsyin reporter strain in the KO background. The active signaling molecule is surprisingly stable; cell free supernatant and extracts that were exposed to a range of pH from 1−12 and temperatures up to 65 °C for several hours remained active in our reporter system (Figure 4.15). Cultures extracts of the WT also remained active after treatment with EDTA and protease for several hours (Figure 4.15).
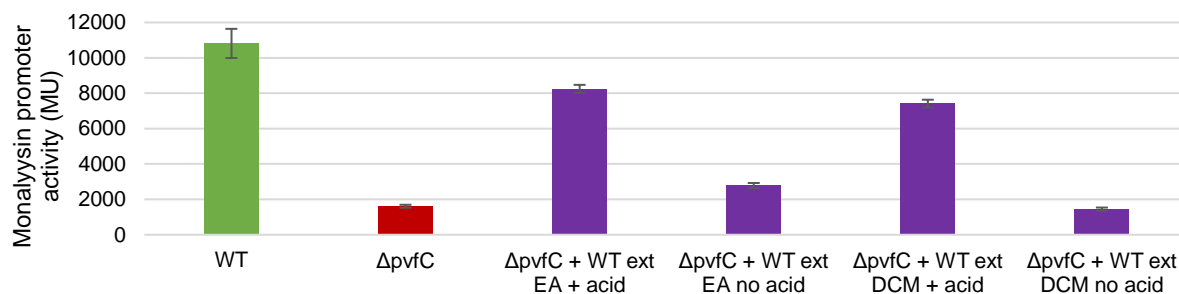
Figure 4.14. The active signaling molecule is extracted into the organic layer only with acidification of spent media. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing P*mnr*-*lacZ* reporter cassette. KO reporter with addition of extracts of WT cultures (purple) is also shown. Spent media was extracted with ethyl acetate (EA) or dichloromethane (DCM), and with or without acidification of spent media from pH 8.5 to 5.0. Extracts were added 4 hours post inoculation at 1 mL culture extracted added to 1mL reporter culture to be assayed.
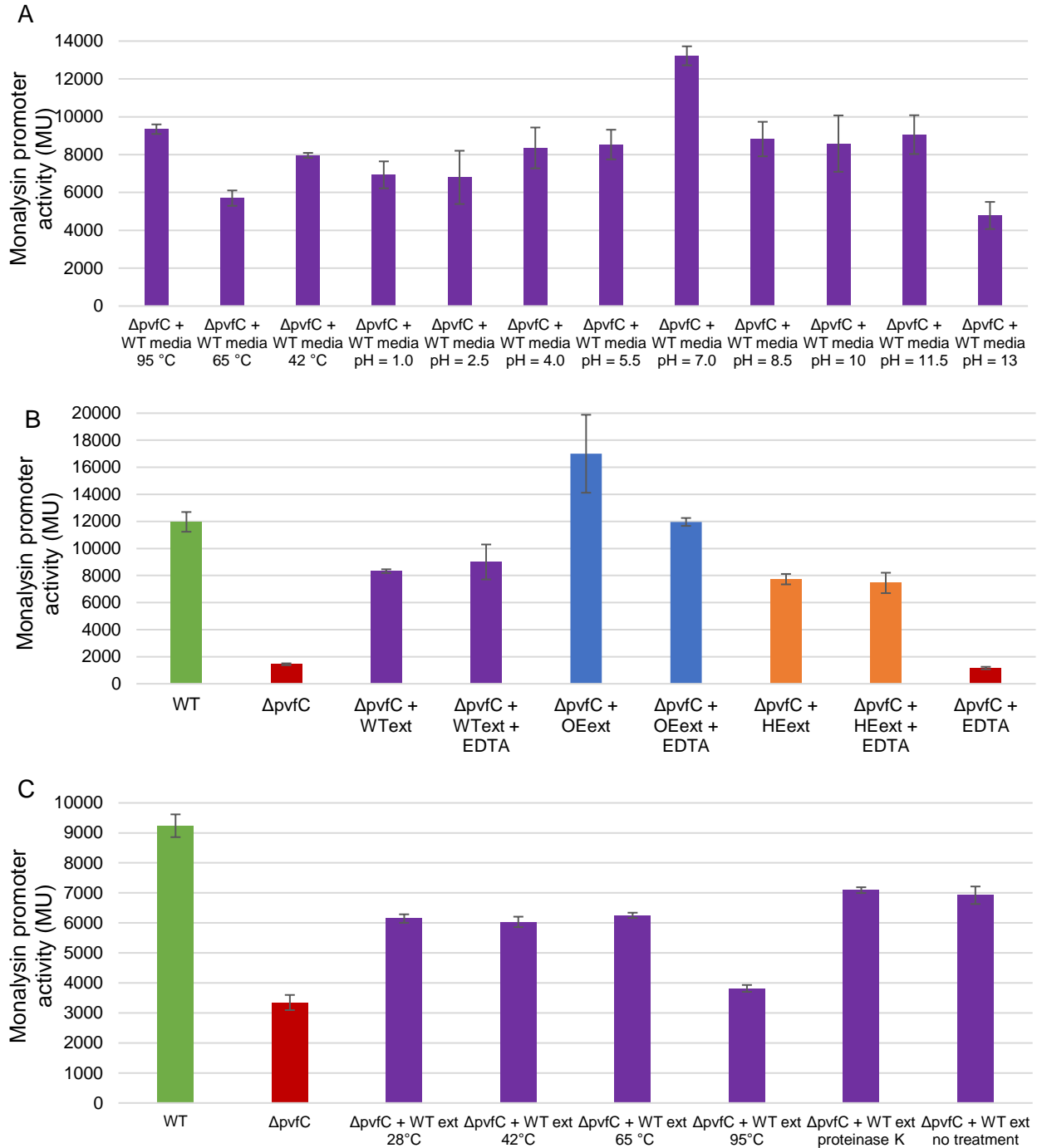
Figure 4.15. The active signaling molecule is stable at temperatures 28−65 °C, a wide pH range and after EDTA or protease treatment. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing the P*mnr*-*lacZ* reporter cassette. KO reporter with addition of (A) WT media or (B,C) culture extracts of WT (purple), *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression (OE, blue) strain, or *P. aeruginosa* PAO1 + pPSV-*pvfABCD* heterologous expression (HE, orange) strain is also shown. (A) The pH of spent media was adjusted, then returned to 7.0 before addition of 0.1 mL of the spent media to 1 mL of the reporter culture. (B,C) Extracts in water were treated with EDTA, heat, or proteinase K for 5.5 hours and then added to reporter strains at 20 µg/mL 4 hours post inoculation.

To increase expression of *pvfD* in the pPSV-*pvfABCD* inducible expression strains, we constructed an additional vector with *pvfD* under the rhamnose-inducible promoter P_{rha} in pSCRhaB2. We transformed *E. coli* Bap1 with this vector and the overexpression vector pPSV-*pvfABCD*. The activity of the culture extracts of Bap1 + PSV-*pvfABCD* + pSCRhaB2 and Bap1 + PSV-pvf*ABCD* + pSCRha-*pvfD* were not significantly different, therefore the additional expression of *pvfD* did not significantly increase production of the final active molecule in *E. coli* (Figure 4.16). This could be due to limited activity of PvfD in *E. coli*. Unfortunately, both the native strain *P. entomophila* L48 and the other heterologous host, *P. aeruginosa* PAO1, harbor trimethoprim resistance, therefore pSCRha-*pvfD* cannot be reliably transformed into these hosts as pSCRhaB2 contains a trimethoprim resistance gene *dhfrii* for selection. Therefore, we designed a new plasmid replacing *dhfrii* with a gene for kanamycin resistance. Construction of inducible expression vectors for *pvf* genes in this new plasmid and subsequent testing for higher activity is currently ongoing.
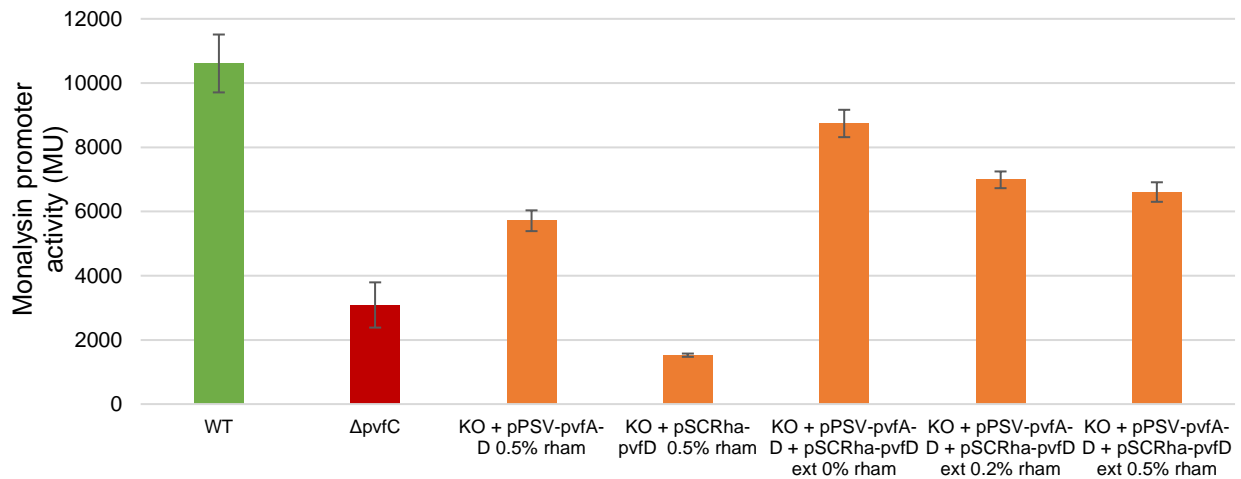
Figure 4.16 Additional expression of PvfD slightly increased production of the active PVF molecule in *E. coli*. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing the P*mnl*-*lacZ* reporter cassette. Supplementation of KO reporter strain with extracts from *E. coli* (orange) with pPSV-*pvfA-D* heterologous expression, pSCRha-*pvfD* tuneable expression, or a combination of the two at various concentrations of the inducer rhamnose for experssion of pSCRha-*pvfD* is also shown.

In addition to ongoing efforts with inducible expression of the *pvf* cluster, we extracted from a 20 L culture of wildtype *P. entomophila* to identify the active molecule while avoiding additional shunt products. However, given the potentially production levels in the native wildtype strain, we sought to optimize our production and extraction methods to yield sufficient quantities for analysis. Importantly, LCMS analysis has not been successful at identifying the active molecule thus far, possibly due to poor ionization, low quantity, and high potency of this molecule. Therefore, we shifted focus to comparative NMR analysis, which require milligram amounts of material.

To verify our isotope labelling experiments contained active molecule, we first tested extract from cultures of the WT strain in minimal media and found they active the *mnl* promoter (Figure 4.17). We also compared yields of the active molecule(s) between extraction methods, as well as organic and aqueous layers from solvent-solvent extractions. For these comparisons, we extracted small molecules from the supernatant of 24-hour cultures. Extracts are resuspended in 10 µL of methanol for every 1 mL of culture. These extracts were used to induce knockout reporter strains, at a ratio of 10 µL/mL (or extract from 1 mL WT culture per 1 mL of the reporter strain). To compare the amount of active molecule(s) present in extracts, the amount of extract added was decreased until the activity was no longer significant compared to the control (knockout reporter induced with methanol).
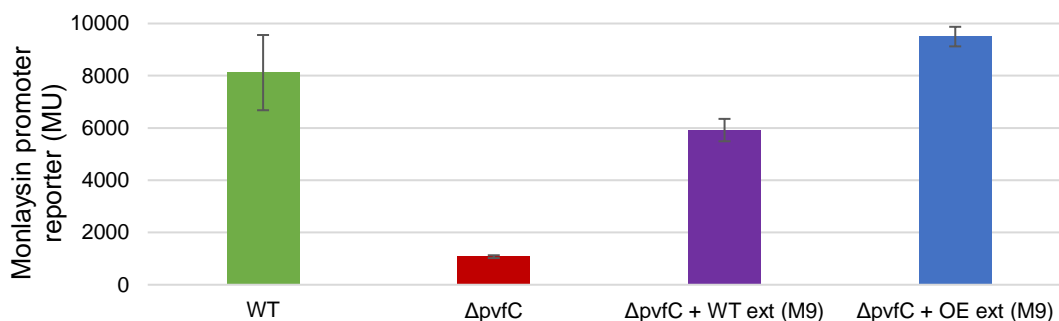
Figure 4.17. The active PVF signaling molecule is produced in minimal media (M9). Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing P$_{mnl}$-*lacZ* reporter cassette. The *ΔpvfC* reporter strain with addition of cultures extracts from WT (purple) and *P. entomophila ΔpvfC* + pPSV-*pvfABCD* overexpression (OE, blue) strains in M9 media is also shown. Extracts were added 4 hours post inoculation at extract from 1 mL culture per 1 mL of the reporter strain.

When we added cell-free supernatant directly to our KO reporter strain, we observed at least twice the activity as the organic extracts, suggesting a large portion of the active molecule remained in the aqueous layer of a solvent-solvent extraction (Figure 4.18). Since both layers contain the active molecule, NMR and LC-MS analysis of the aqueous layer active fractions can be used to validate data from the organic layer active fractions. To obtain metabolites from the aqueous layer, two volumes of acetone were added to crash out salts. This mixture was left overnight and filtered into a round bottom flask. Although by mass the aqueous layer contained more of the active molecule, the mass of the overall extract was orders of magnitude more than the organic extract (10 g/L vs 100 mg/L), requiring significantly more HPLC injections to purify material. Therefore, we chose to primarily use the organic layer extracts for further purification.

We also tried extraction with XAD-4 and XAD-16 resin beads. The XAD resins are generally used for adsorption of organic compounds from aqueous solutions. The beads were rinsed in methanol and swelled with deionized water before incubating with supernatant overnight. The beads were rinsed with water and then the metabolites were eluted three times with methanol. These extracts were messy, solvent intensive, and yielded approximately the same amount of active molecule per volume than our solvent-solvent extractions (Figure 4.19).
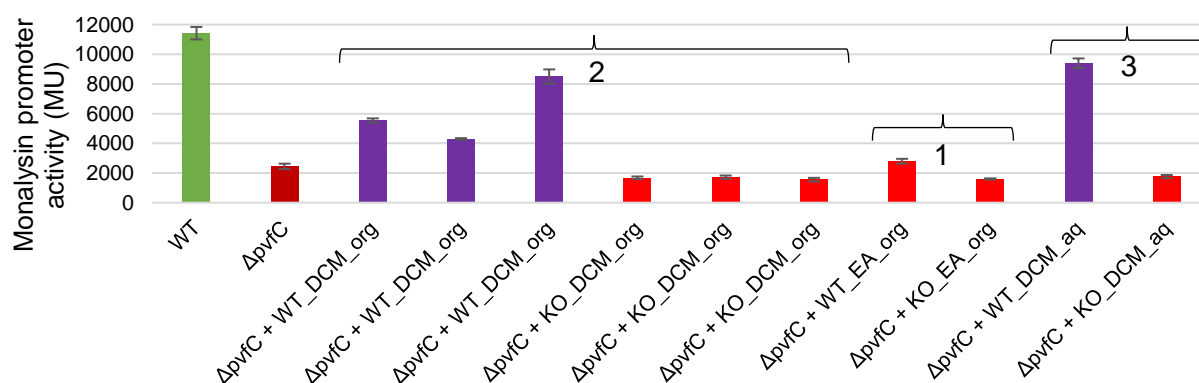
Figure 4.18. A significant fraction of the active PVF molecule remains in the aqueous layer after organic extraction. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing P*$_{mnl}$*-*lacZ* reporter cassette. The *ΔpvfC* reporter strain supplemented with culture extracts from WT (purple) and *ΔpvfC* (bright red) strains is also shown. WT and *Δpvf* cultures were first extracted with ethyl acetate (EA, 1), and the organic layer was separated and dried. The pH of the aqueous layer was adjusted to 5.0 then extracted again with dichloromethane (DCM), and both the organic (2) and aqueous (3) layers were kept and tested. Extracts were added 4 hours post inoculation at extract from 1 mL culture per 1 mL of the reporter strain.
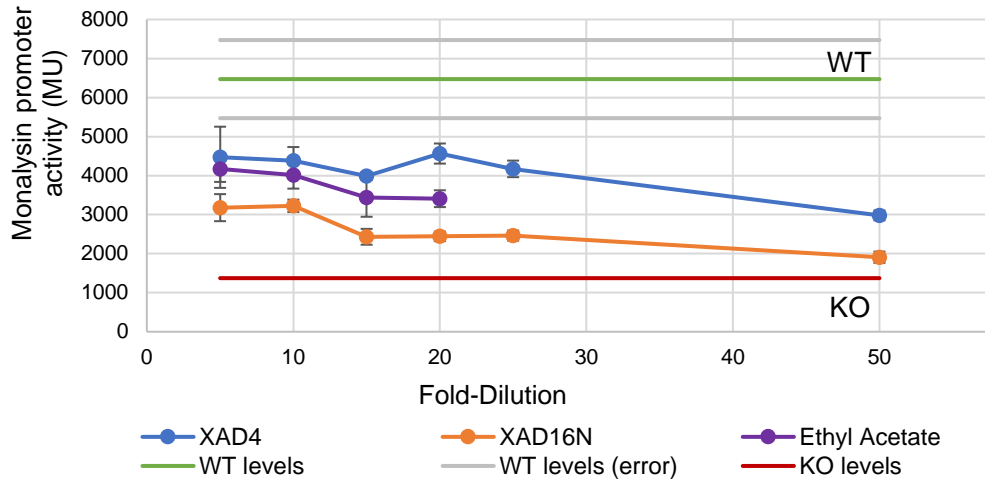
Figure 4.19. Comparison of different extraction methods for effectiveness at extracting the active PVF molecule. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green, error in grey) and *pvf* deletion (red) strains each containing P*mnl*-*lacZ* reporter cassette. KO reporter strain supplemented with of increasing dilutions of culture extracts from WT *P. entomophila* using XAD-4 (blue) resin beads XAD-16N (orange) resin beads or ethyl acetate (purple) is also shown. Extracts were added 4 hours post inoculation, dilution 1 is the equivalent concentration of extract from 1 mL culture per 1 mL of the reporter strain.
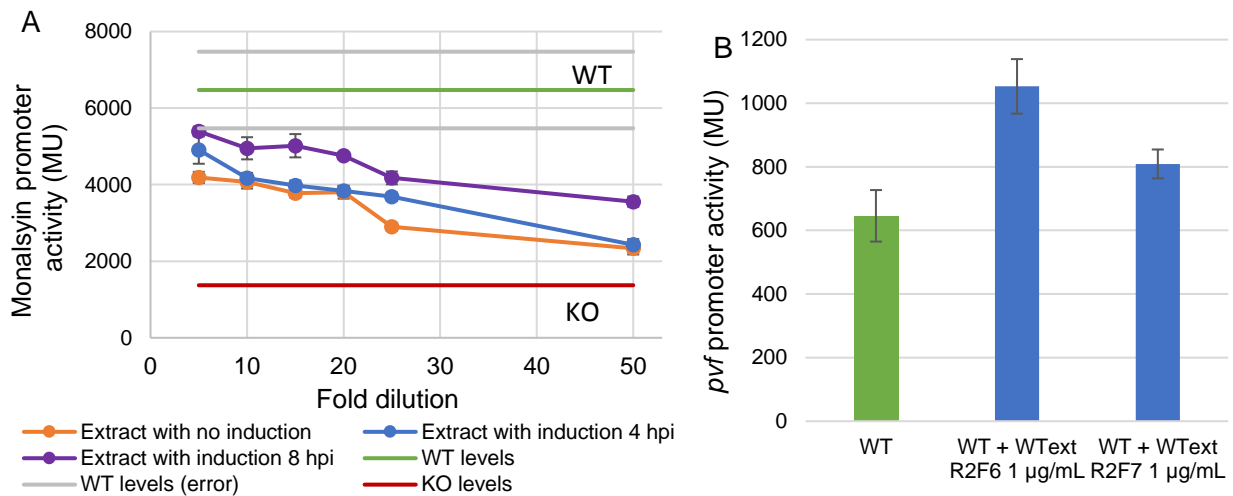
Figure 4.20. Autoinduction with active fractions from culture extracts of wildtype *P. entomophila* increases the overall yield of active PVF molecule(s). (A) Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green, error in grey) and *pvf* deletion (red) strains each containing P*mnl*-*lacZ* reporter cassette. The KO reporter strain supplemented with of increasing dilutions of culture extract of wildtype *P. entomophila* is also shown. Extracted cultures were uninduced (orange) or induced with active fractions from culture extracts of WT at 1 mg/L 4 (blue) or 8 (purple) hours post inoculation. Extracts were added 4 hours post inoculation, dilution 1 is the equivalent concentration of extract from 1 mL culture per 1 mL of the reporter strain. (B) Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) containing P*pvf*-*lacZ* reporter. The WT reporter supplemented with active fractions from culture extracts of WT at 1 µg/mL (blue) is also shown.

111

Our studies of the *pvf* promoter, discussed in Chapter 3, have shown that addition of active PVF molecule will auto-induce the production of PVF in wildtype *P. entomophila* (Figure 3.13). Therefore, we used extracts from a 4 L culture of wildtype to auto-induce PVF production in an additional 12 L culture of wildtype to increase overall yield of active molecule. Comparison between WT cultures with and without supplementation of active fractions suggests that autoinduction of the *pvf* pathway increases the production of active signaling molecule (

Figure *4.20*). For large extractions, we began to use a double extraction method to reduce the metabolite background in the organic extracts. Specifically, since ethyl acetate was ineffective at extracting the active molecule (without acidification), supernatant was first extracted with ethyl acetate to remove metabolites unrelated to the *pvf* pathway. Then the pH of the supernatant was adjusted to 5.0, resulting in a yield of about 50 mg extract per liter culture with the second extraction, compared to about 100 mg/L with only one extraction, reducing the amount of extract for purification of the active molecule by half.

We set out to purify the active PVF molecule from autoinduced wildtype cultures. For the first round of purification, Fractions 1−7 are collected (min 16−19,) and combined for the second round of purification (Figure 4.21A). Fractions 5−8 from the second round of purification (min 16−18, Figure 4.21B) are separately injected for the third round of purification, with the goal of detecting peak(s) common to fractions 5−8 but unique to wildtype cultures. Culture extracts from the *Δpvf* knockout strain were purified in the same manner for comparative metabolomics by LC-MS and NMR. LC-MS did not yield a potential feature of interest (Figure 4.22). In addition, a native molecule, which is likely a valine-proline diketopiperazine based on our previous NMR analysis (m/z = 197.125 with a retention time of 6.8 min), is the most abundant compound in the active fractions but also present the in knockout extracts. Thus, the third purification round was required to separate the active molecule from the abundant diketopiperazine compound (Figure 4.23). To verify that the signaling activity was not unique to the monalysin promoter, these

fractions were tested using the *pvf* promoter-reporter strains and the same fractions exhibited

signaling activity, confirming the pleotropic signaling activity of the PVF signaling molecule

(Figure 4.24). Purification of the aqueous layer from the ethyl acetate extractions (both single

and double extractions) lead to the same active fractions, indicating that it contains the same

active molecule as the organic layer (
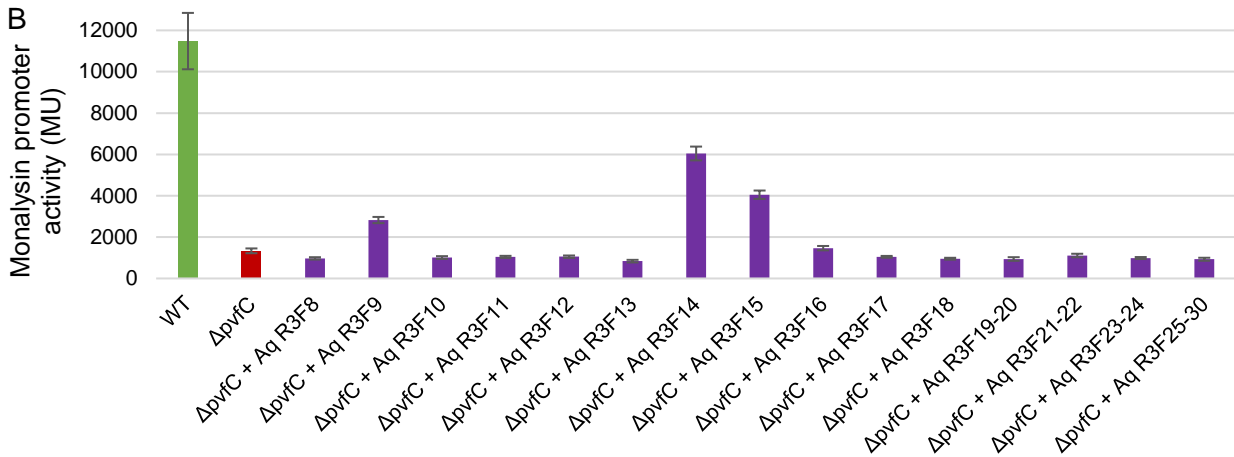
Figure *4.25*). By total ion chromatogram a few potential peaks (Figure 4.22) were unique

to the WT cultures, but no *m/z* feature has been consistently identified in individual purifications

from multiple cultures.
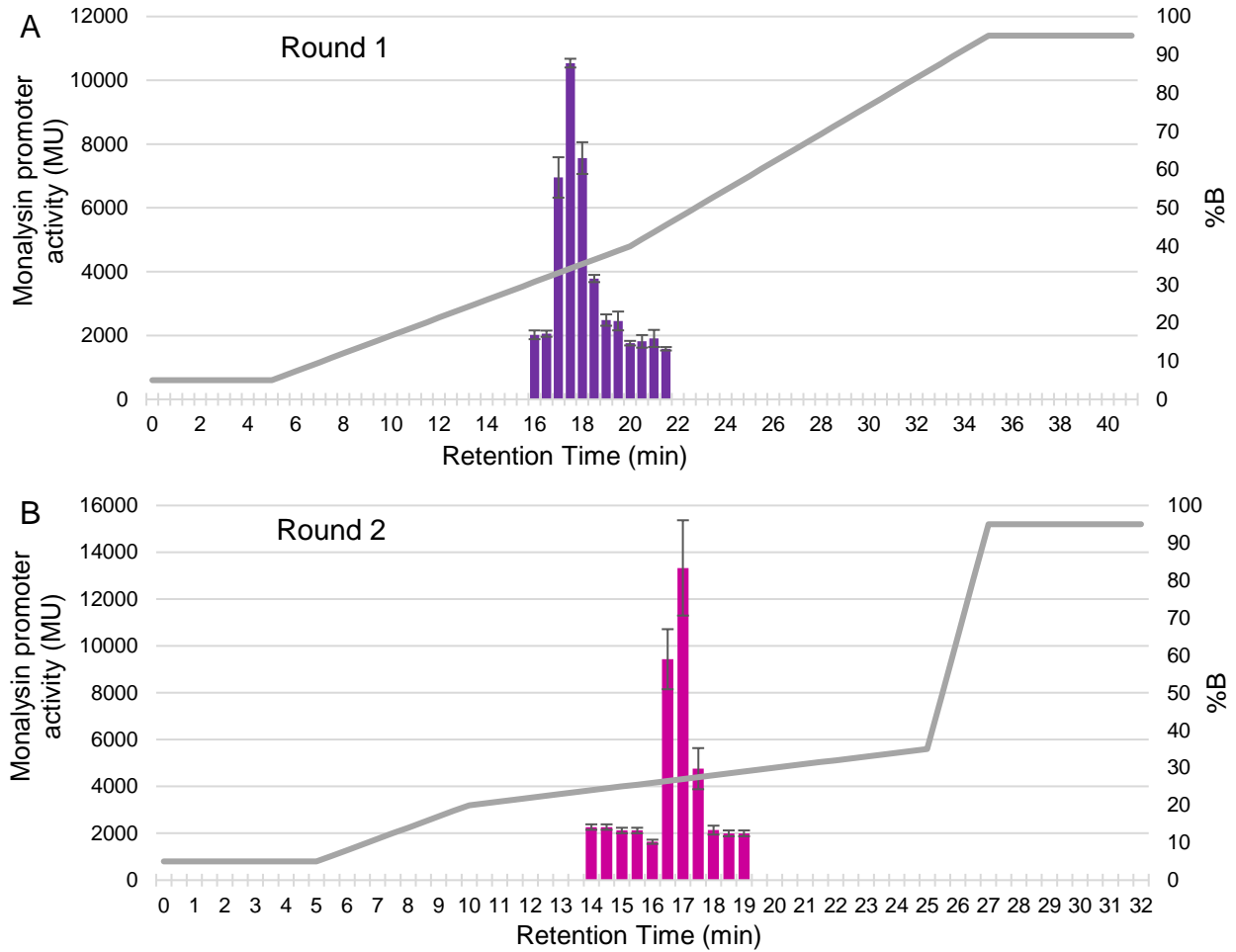
Figure 4.21. Bioactivity-driven purification of the PVF signaling molecule from *P. entomophila* wildtype extracts. Activity and retention time of fractions from Round 1 (A) and Round 2 (B) purification of culture extracts of *P. entomophila* wildtype (WT). Quantification of the β-galactosidase activity 24 hours post inoculation is shown. Cultures of the *pvfC* deletion strain containing the P*mnt*-*lacZ* reporter cassette were supplemented with each dried fractions at 1 µg/mL, 4 hours post inoculation.
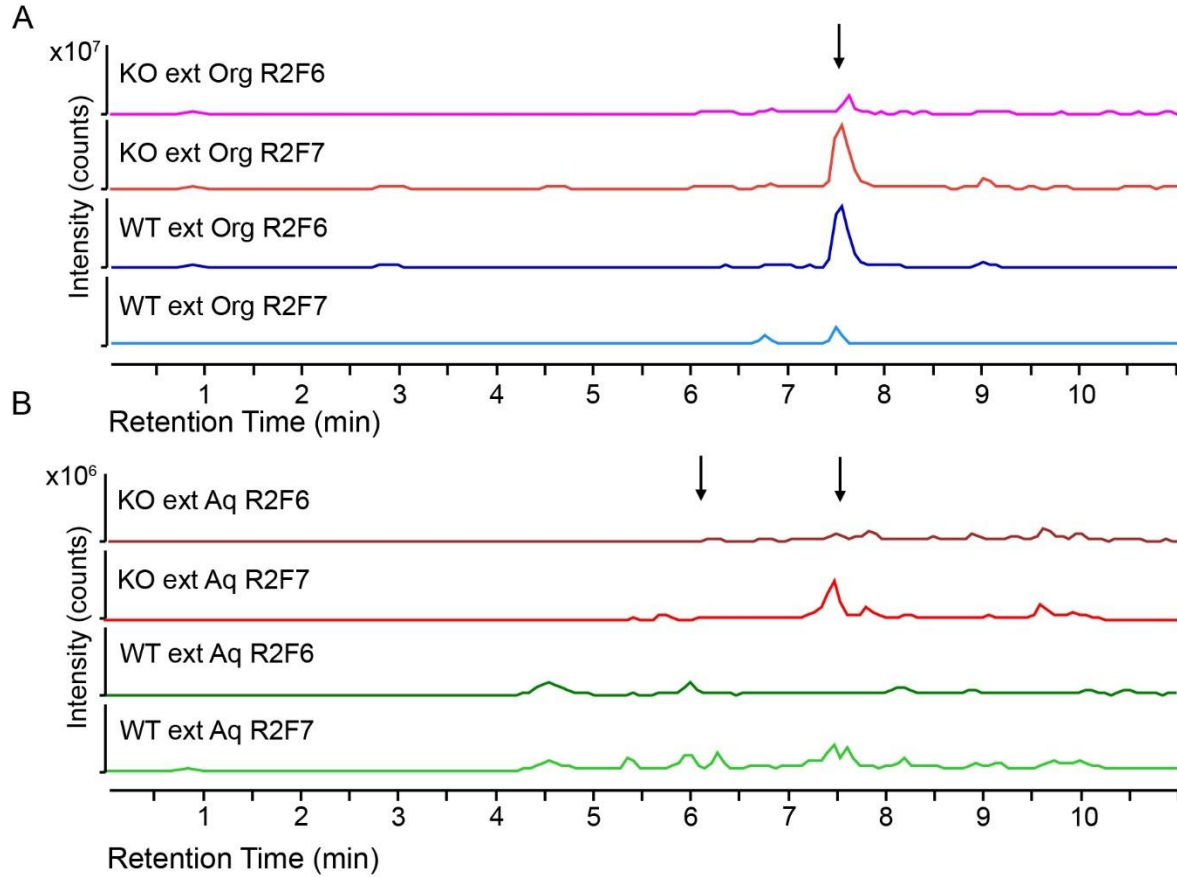
Figure 4.22. Liquid chromatography of active fractions (Round 2, Fractions 6-7) from (A) organic and (B) aqueous culture extracts of *P. entomophila* wildtype (blues and greens) and *Δpvf* deletion (reds) strains. Total ion chromatograms. The molecule present in both WT and KO fractions is indicated at 7.5 min. Peaks of interest present in WT but not KO aqueous active fractions are indicated at 6 min.
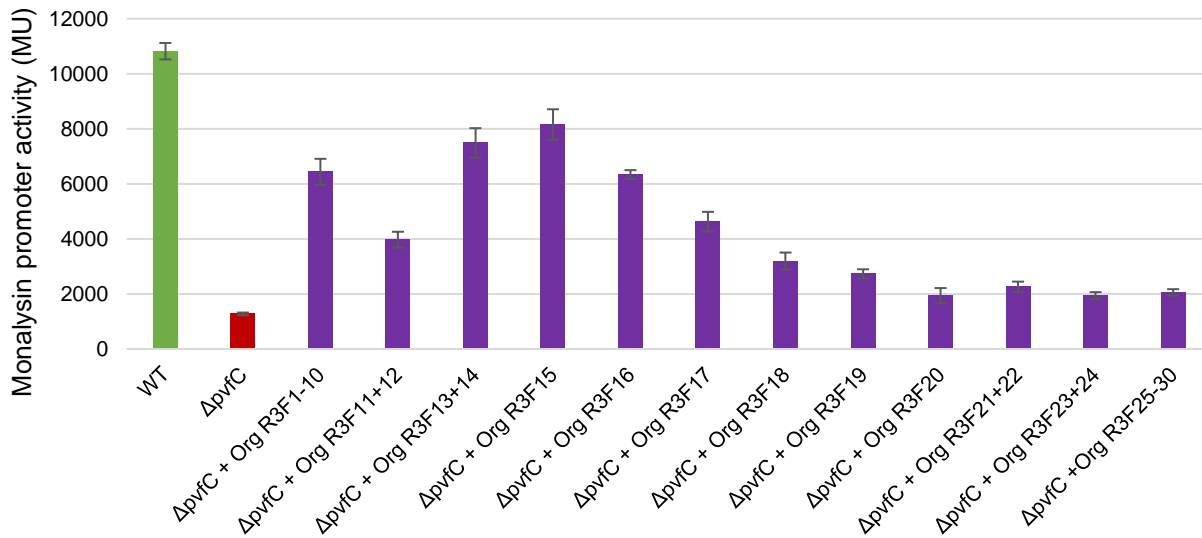
Figure 4.23. Activity of fractions from purification Round 3 of culture extract of *P. entomophila* WT strain. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (red) strains each containing P*mnr*-*lacZ* reporter cassette. The *ΔpvfC* reporter strain with addition of fractions from purification round 3 at 2 µg/mL (purple) is also shown. Fractions were added 4 hours post inoculation.

115
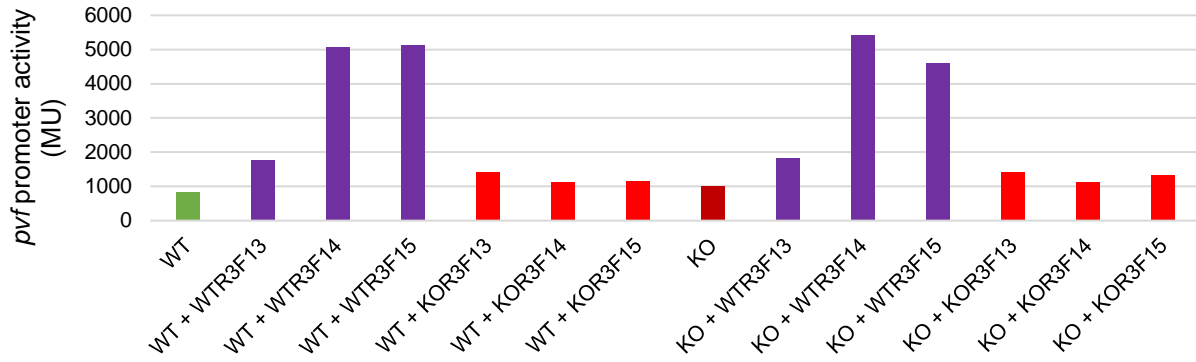
Figure 4.24. The active fractions from culture extracts of wildtype, but not *pvf* deletion strain, activate the *pvf* promoter. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strains each containing the P*pvf*-*lacZ* reporter cassette. The WT and KO reporter strains with addition of purification round 3 fractions of culture extract from WT (purple) and KO (cherry) strains is also shown. Fractions were added at 2 μg/mL four hours post inoculation.
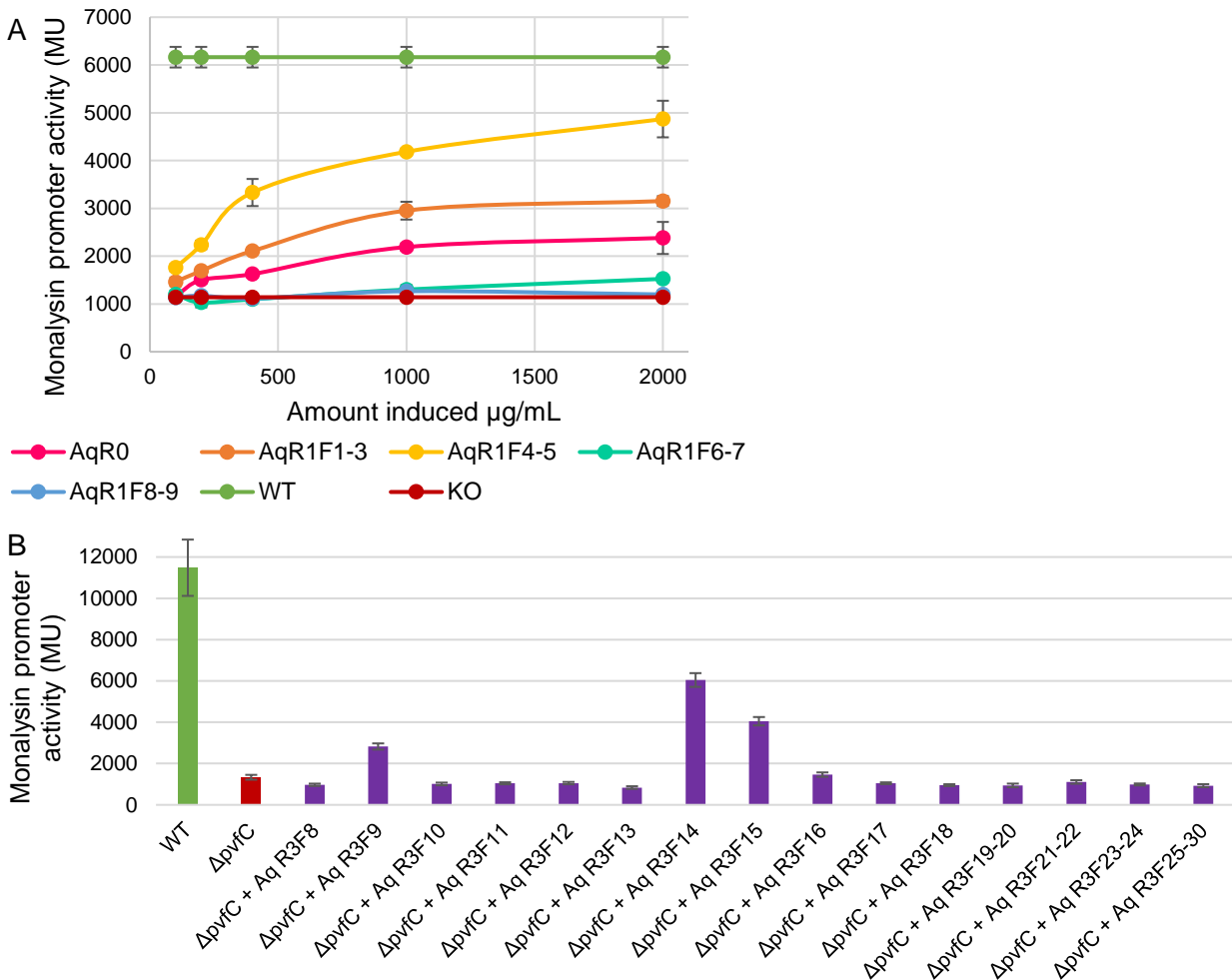
Figure 4.25. Purification of the aqueous layer from the ethyl acetate extractions lead to the same active fractions as the organic layer. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strains each containing the P*mnl*-*lacZ* reporter cassette. The *ΔpvfC* reporter strain with addition of (A) round 1 and (B) round 3 purification fractions of the aqueous layer of extracts of wildtype *P. entomophila* using ethyl acetate is also shown (purple). Fractions at (A) various concentrations or (B) 15 μg/mL were added four hours post inoculation.

116
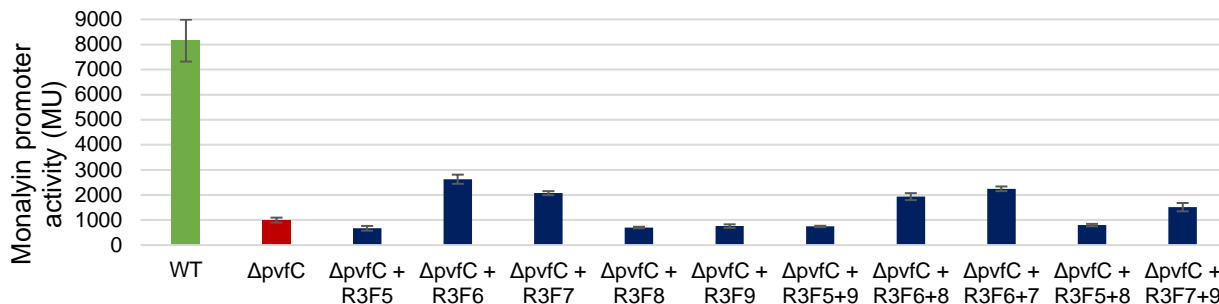
Figure 4.26 Further purification of active biological dPNO samples reveals that the signaling activity is not from dPNO, but an impurity. Combination of fractions does not increase activity. Quantification of the β-galactosidase activity 24 hours post inoculation is shown for *P. entomophila* WT (green) and *pvf* deletion (KO, red) strains each containing the P*mnl*-*lacZ* reporter cassette. The *ΔpvfC* reporter strain with addition of dPNO repurification fractions or combinations of fractions (dark blue) is shown. Fractions from purification round 3 of an active dPNO sample were added at 5 μg/mL at four hours post inoculation. By LC-MS dPNO is present in these fraction R3F9.

Alongside our efforts to purify the active signaling molecule from extracts of the wildtype strain, we also attempted to purify the active molecule from other biological samples that previously exhibited activity. One such sample was isolated dPNO that exhibited activity, discussed in Chapter 2 (Figure 2.16). Since further purified dPNO did not exhibit activity, we hypothesized that an impurity is responsible for the observed activity. To test this hypothesis, we separated the sample of dPNO using the HPLC method for third round purification described above and detected signaling activity from fractions at the same retention time as our WT extracts, but not at the same retention time as dPNO (Fractions 6-7 and Fraction 9 respectively, Figure 4.26). To verify that activity did not require multiple synergistic compounds, we added two fractions together to the KO reporter, but did not observe an increase in activity.

After round 3 of purification, we observe activity in 5 fractions, which corresponds with an entire column volume at 20% acetonitrile (with 0.1% TFA). For a sharper peak, a shallower gradient prior to 20% acetonitrile or an isocratic period at a lower percentage of acetonitrile, might be required. Other efforts to purify the active molecule include alternative solvents for reverse phase chromatography, normal phase chromatography or smaller columns. Initial comparative 2D COSY NMR experiments suggest potential cross peaks of interest observed in WT fractions but not in KO fractions. Scale-up and a mass characterization will be required for structural elucidation.

4.4 Conclusion, Discussion and Future Directions

To identify the small molecule product of the *pvf* biosynthetic pathway responsible for signaling activity, we developed reporter strains using the *pvf* and *mnl* promoters, that are regulated by the *pvf* pathway, as shown in Chapter 3. Through overexpression of the *pvf* cluster and bioactivity-driven purification of large scale extracts, we isolated 2-isopropyl-5,6,7,8-tetrahydroimidazo[1,2-a]pyridine, or imidazole cyclane (ImC), another new compound not observed previously in nature. Discovery of ImC and the (d)PNOs described in Chapter 2 indicate the PvfC and PvfB can produce a diversity of compounds. Isotope labelling experiments suggests that valine is one of the precursors to Imc, but its biosynthesis remains to be elucidated. Isomers of ImC have been isolated from Chinese medicinal plants with potential bioactivity.[137] Although synthetic ImC did not exhibit signaling activities in our reporter assays, it may harbor other biological activity and determination of these activities is ongoing.

Significant progress has been made towards the identification of the elusive PVF signaling molecule, which is likely a novel structure. We showed that this molecule is very stable: it maintains activity after treatments at acidic and basic pHs, or high temperature and with EDTA or protease. We developed an improved large-scale extraction method for this molecule from wildtype cultures and are working towards an optimized third round purification method. Combining comparative metabolomics by NMR and LC-MS will expedite the structure elucidation process: insights obtained using one technique complements analysis in the other. In parallel, we are currently seeking to understand other aspects of the *pvf* pathway to facilitate identification of the PVF molecule, as described below.

Identification of the *pvf* receptor

One ongoing project is to identify the receptor that may play a role in the *pvf* downstream

signaling pathway, which will yield a greater understanding of the *pvf* signaling cascade. In

addition, the receptor could be used to capture and concentrate the active small signaling

molecules produced by the *pvf*-encoded enzymes.

Our first approach to identify receptors involved in the *pvf* signaling cascade was to

conduct T_n5 mutagenesis on the *P. entomophila* PSEEN0973-*lacZ* translational reporter to

identify isolates that exhibited promoter activity levels similar to that of the reporter strain

containing an insertional mutation at *pvfC*. If the T_n5 insertion occurs at a gene related to the *pvf*

signaling pathway, such as a receptor, the promoter activity should decrease to a similar level

as the *pvfC* mutant. We can visualize promoter activity using X-gal, which is cleaved by LacZ

and dimerized to produce a blue dye. We tested several promoter-reporter strains for T_n5

mutagenesis but chose the PSEEN0973-*lacZ* translational hybrid reporter because this reporter

leads to low expression levels of *lacZ* to reduce background, but still high enough for detection.

The insertional mutant strain grew as a lighter blue colony when grown on agar plates with X-

gal, in comparison to the WT PSEEN0973-*lacZ* reporter strain. In comparison, other reporter

strains exhibited more significant differences between the WT and *pvf* mutant reporter strains

with respect to promoter activity (as quantified in Miller Units), but the activities from both strains

were still strong enough to produce a dark blue pigment when plated with X-gal. At the moment,

the biological role of PSEEN0973 is unknown, and we would need to complete a full

mutagenesis screen or a screen on a different *pvf*-regulated promoter reporter strain to confirm

that hits from the initial screen are genes that are part of the *pvf* signaling cascade.

The initial screen was conducted by Martina Knechel while working as an undergraduate

SURF fellow and thesis student. *P. entomophila* containing the PSEEN0973-*lacZ* translational

fusion was subjected to T_n5 mutagenesis. White and lighter blue colonies were subsequently

restreaked to ensure homogeneity. The activity of each mutant, using the ONPG assay, was compared between WT *P. entomophila*-PSEEN0973::*lacZ*, *P. entomophila*-PSEEN0973::*lacZ* with insertional mutation at *pvfC*, and *P. entomophila* without a *lacZ* reporter. Mutants that showed activity levels similar to the *pvf* mutant were identified by arbitrary PCR and sequencing of the region containing the T_n5 insertion.

From the initial blue/white screen, we identified multiple clones with Tn5 inserted into PSEEN3202−3203. These genes are part of a four gene cluster PSEEN3202−3205, containing two histidine kinase/response regulators, a CheB-like methyl esterase and a CheR-like methyl transferase. However, this cluster of genes did not appear to be conserved in all *pvf*-containing strains. Only about one third of *pvf*-containing strains harbored all four genes. Among the *pvf*-containing strains, a homolog of PSEEN3205 in *P. alkylphenolica* KL28, identified as part of the *BsmABCD* cluster, was shown to be involved in negative regulation of motility. [138]

We also used a bioinformatics approach by BLAST to identify LuxR-type transcriptional regulators that were present in strains containing *pvf* and not present in strains without. An initial Uniprot search for "LuxR transcriptional regulator" found 25 genes in *P. entomophila,* and 267 "general transcriptional regulators" including LysR, AsnC, AraC, GntR, and TetR regulators in the genome of *P. entomophila* L48. Using pairwise BLAST analysis, we compared the 25 LuxR-type genes to two *pvf*-containing strains, *Burkholderia cenocepacia* J2315, and *P. syringae* pv. *syringae* DC3000 to identify those also conserved in other *pvf*-containing strains. In addition to Uniprot searches, we also conducted homolog search with "pseudoluge" (pseudoluge.pseudomonas.com) in search of regulatory proteins with homologs in *P. syringae* and *P. fluorescens,* which contains *pvf,* but not in *P. aeruginosa* PAO1 that does not harbor *pvf*. Combining these two lists, we were identified the following proteins of interest: Q1IE01 (PSEEN1212), Q1IG11 (PSEEN0437), Q1I7C4 (PSEEN3740), Q1I8E1 (PSEEN3335, EtlR)*.*

Among the four proteins, we identified the LuxR-type regulator, EtlR (PSEEN3335), as a potential protein involved in the *pvf* signaling pathway. EltR is a regulator for entolysin biosynthesis in *P. entomophila.* As shown in Chapter 3, we detected a difference in activity between the *etlA-lacZ* translational hybrid in the wildtype strain and the strain containing an insertional mutation at *pvfC* (Figure 3.23). Our current efforts involved verification of the potential hits PSEEN3202−3203 and PSEEN3335 by genetic knockouts and analysis using the promoter-reporter systems. If the knockout show similar promoter activity to a knockout of *pvf*, these genes likely play a role in the reception and/or transmittance of the *pvf* signal.

The *pvf* signaling molecule as an elicitor

Quorum sensing molecules are known to regulate virulence factors, including the production of small molecules. By using *pvf* as an elicitor, we can identify novel small molecules from *P. entomophila* under *pvf* regulation. In ongoing work discussed briefly in Chapter 3, we have observed several unique molecules present in the wildtype cultures of *P. entomophila* compared to *pvf* deletion mutant cultures. We will isolate these molecules and use LC-HRMS and NMR to determine their structures. Tandem MS-MS fragmentation analysis using Global Natural Product Social Molecular Networking will enable us to determine if these compounds are part of a known family of molecules. Similar to current collaborations in biosynthetic pathways, the Global Natural Product Social Molecular Network allows for rapid comparison of uploaded samples. Molecular network sorts metabolites by their chemistries, where analogs of known metabolites can be identified and novel chemistries prioritized. Through these studies, we can potentially identify novel molecules from known families of compounds or, in combination with activity-guided purification, novel molecules with important bioactivities.

Understanding the *pvf* cluster in *Pseudomonas* and *Burkholderia*

Over 300 sequenced strains contain the *pvf* biosynthetic pathway (by BLAST analysis), including the plant pathogen *P. syringae* and human pathogen *B. cenocepacia.* Our work has demonstrated that the monalysin promoter of *P. entomophila* responds to *pvf*-containing extracts from *B. cenocepacia,* suggesting the produce similar molecules via the *pvf* pathway or are involved in interspecies communication. Our work using *P. entomophila* as a model can help guide studies of the *pvf* pathway in *B. cenocepacia* and *P. syringae.* By characterizing the structures and functions of bioactive small molecules from these pathogens, our work has the potential to develop an understanding for the production of these compounds and reveal novel and useful antimicrobial targets for human therapeutics or agrichemical development.

APPENDIX: NMR SPECTRA AND EXTENDED TABLES
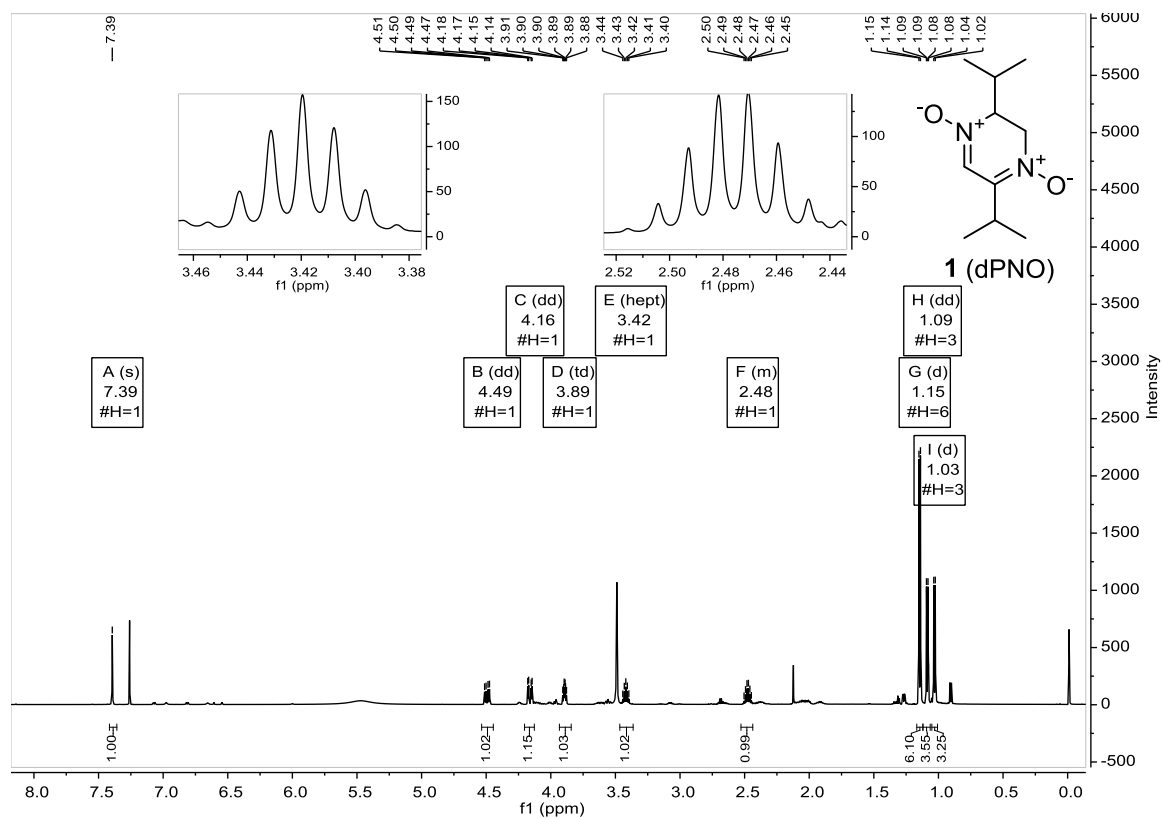
Appendix 1. ^{1}H NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CDCl$_3$ (600 MHz).
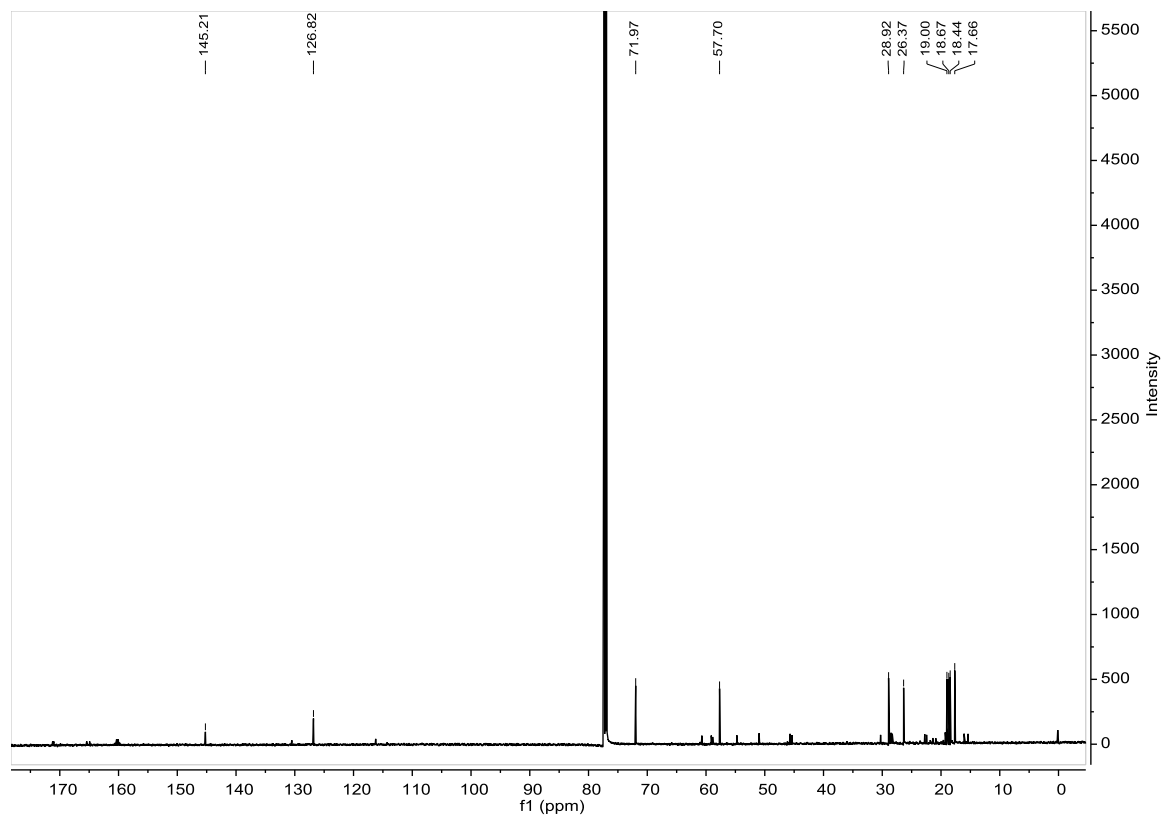
Appendix 2. ^{13}C NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (151 MHz).

1 (dPNO)
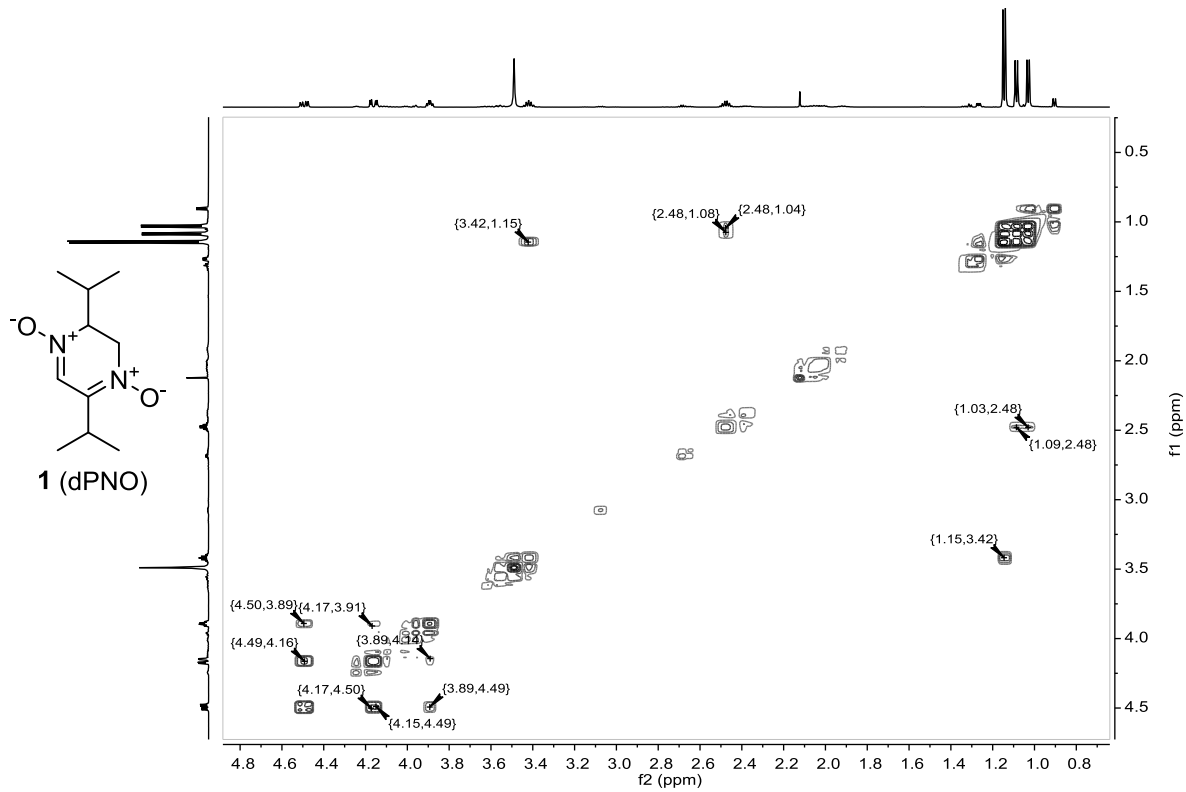
Appendix 3. (^{1}H, ^{1}H)-COSY NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (600 MHz).
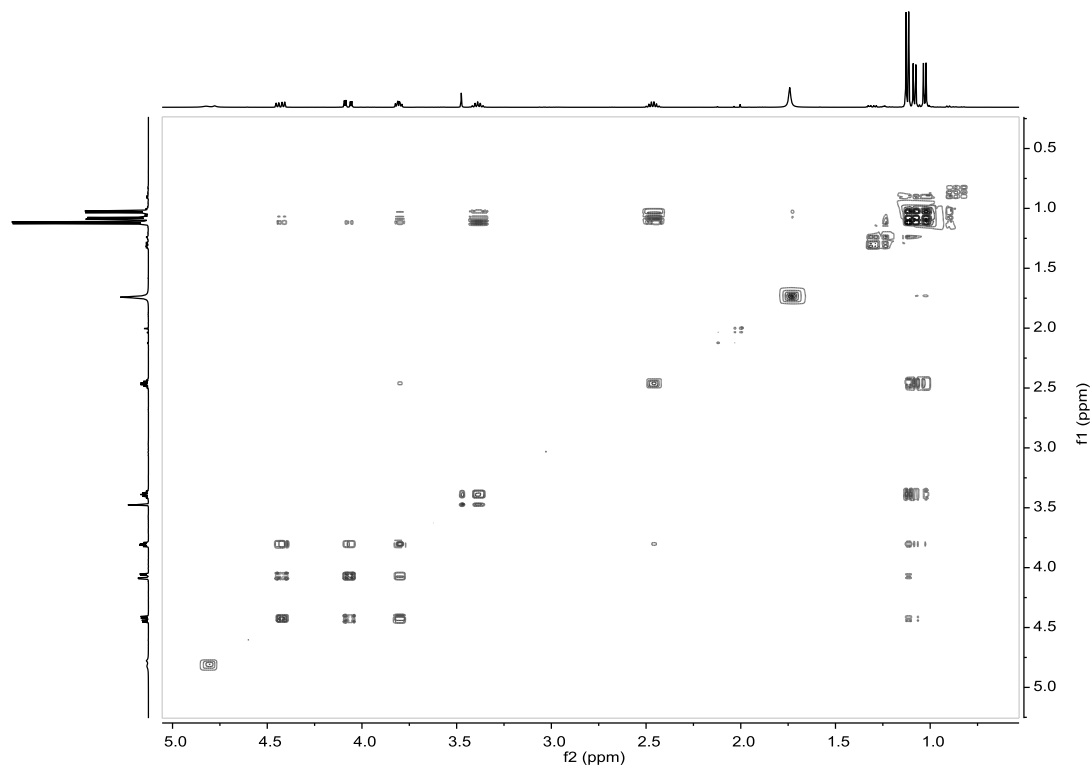
Appendix 4. (^{1}H, ^{1}H)-TOCSY NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (500 MHz).

{1.10,18.72} {1.04,17.90}
{1.15,18.31}
{3.42,26.15} {2.48,28.59}
{4.16,57.65}
{4.49,57.48}
{3.89,71.76}
{7.40,126.61}

1 (dPNO)
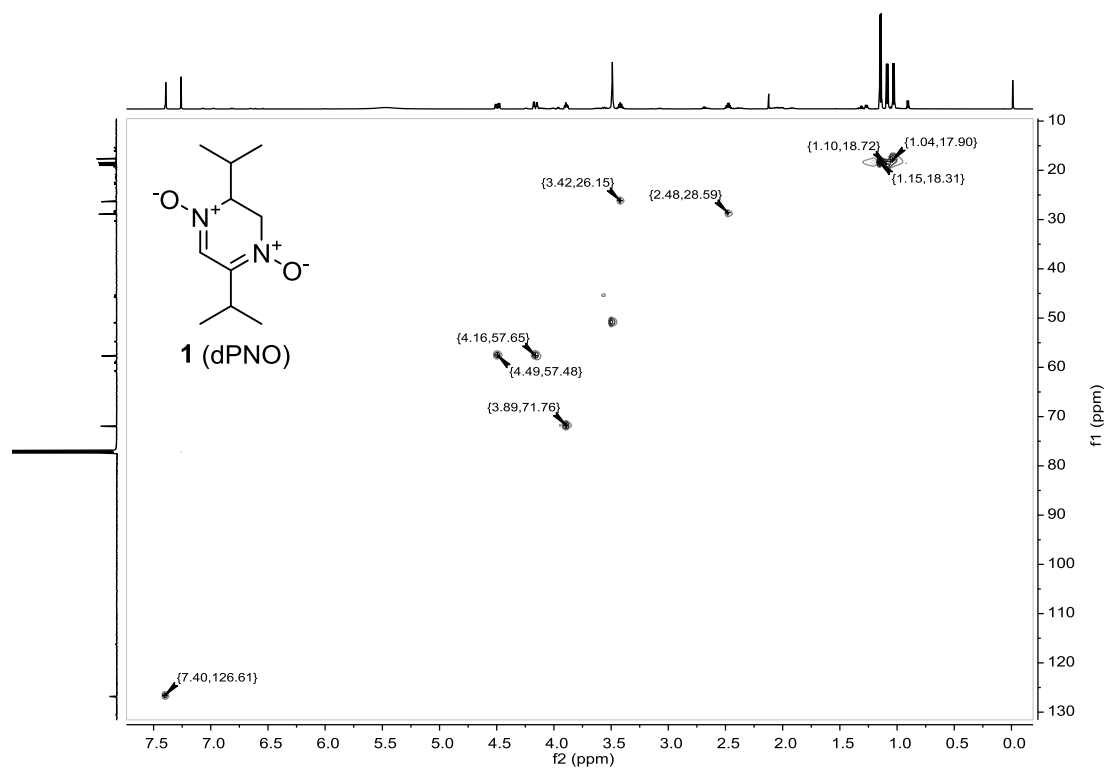
Appendix 5. (^{1}H, ^{13}C)-HSQC NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (^{1}H, 600 MHz, ^{13}C, 151 MHz).
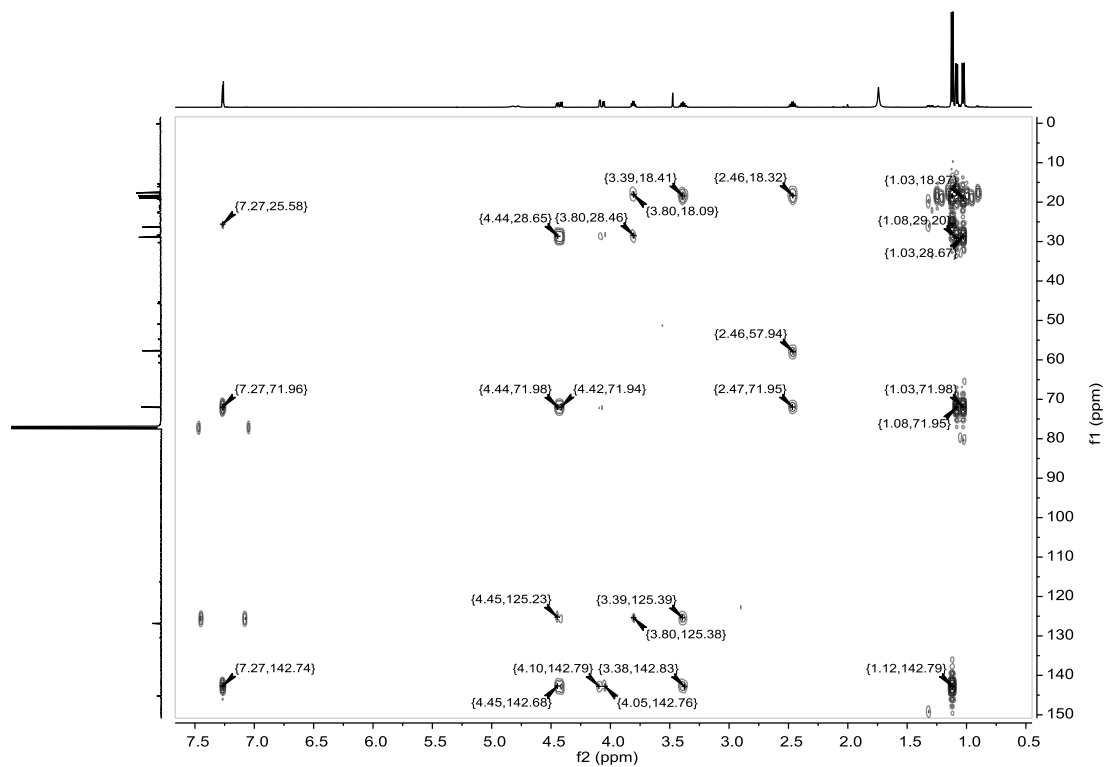
Appendix 6. (^{1}H, ^{13}C)-HMBC NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (^{1}H, 500 MHz, ^{13}C, 126 MHz).
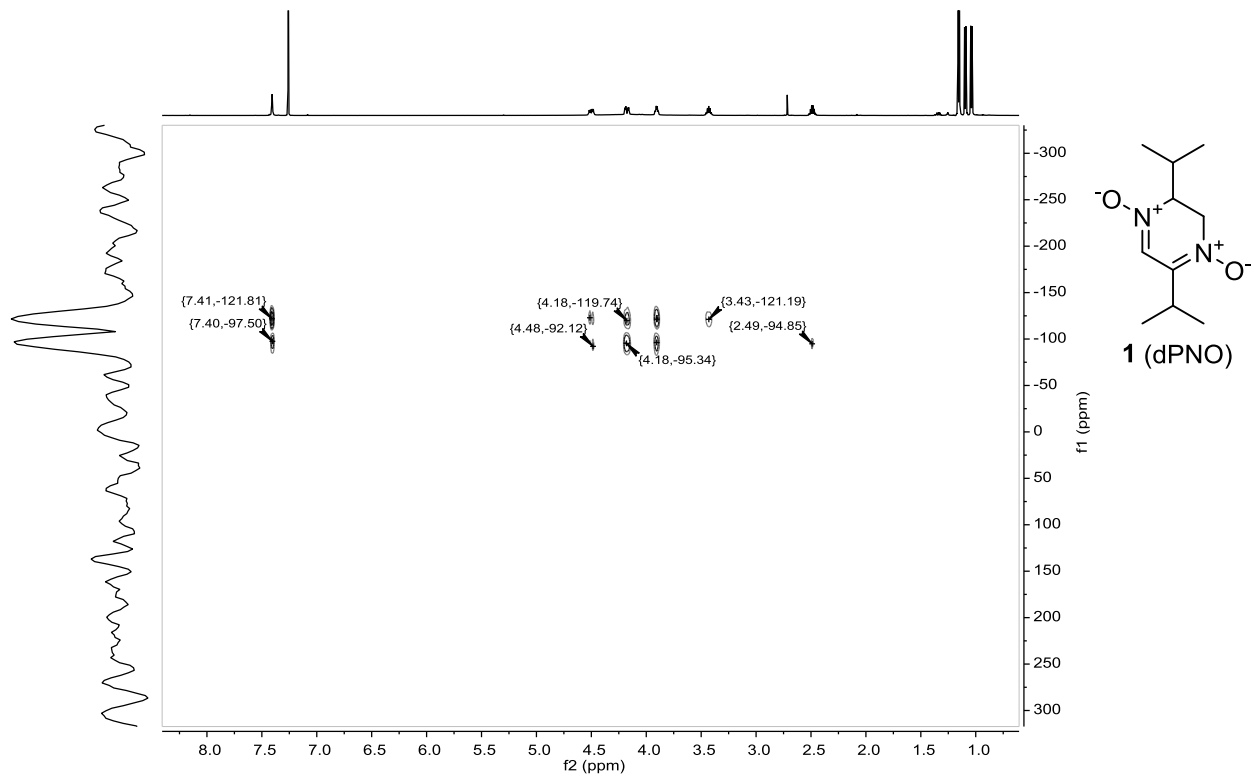
Appendix 7. (^{1}H, ^{15}N)-HMBC NMR spectrum of isolated dPNO (**1**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (^{1}H, 600 MHz, ^{15}N, 61 MHz).
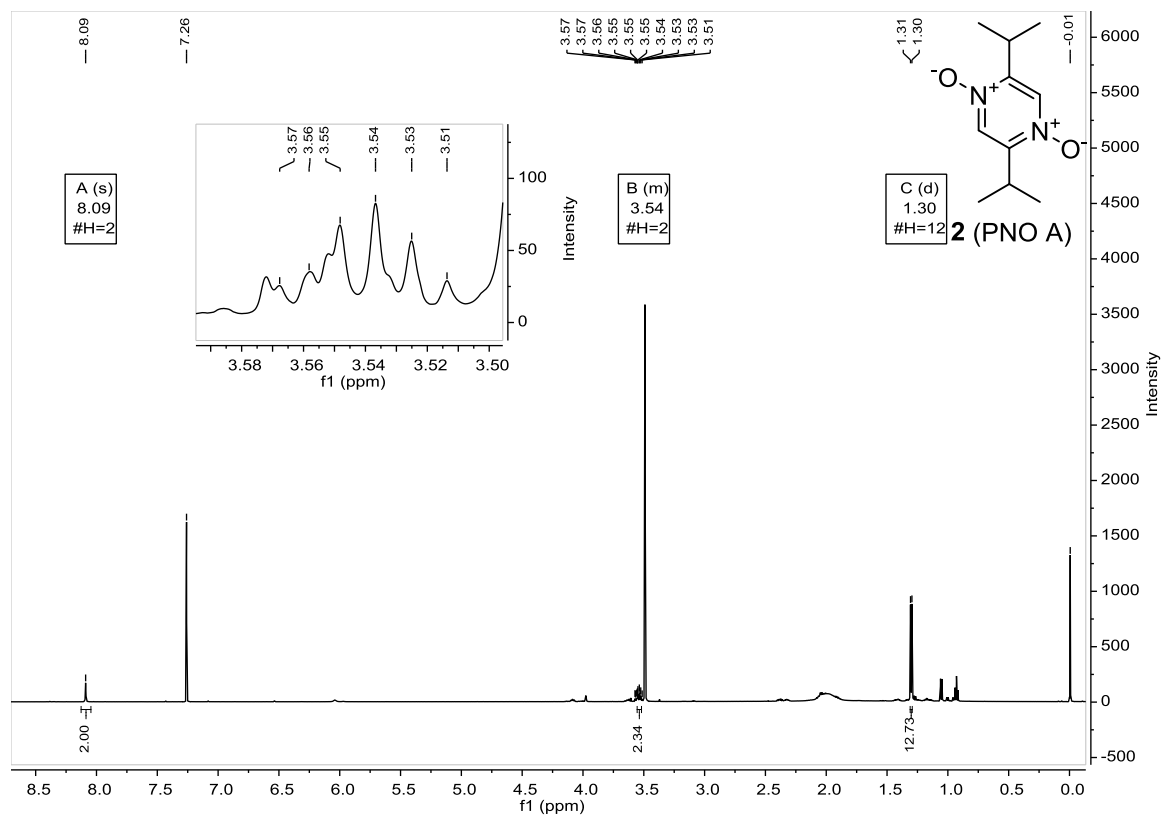
Appendix 8. ^{1}H NMR spectrum of isolated PNO A (**2**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (600 MHz).
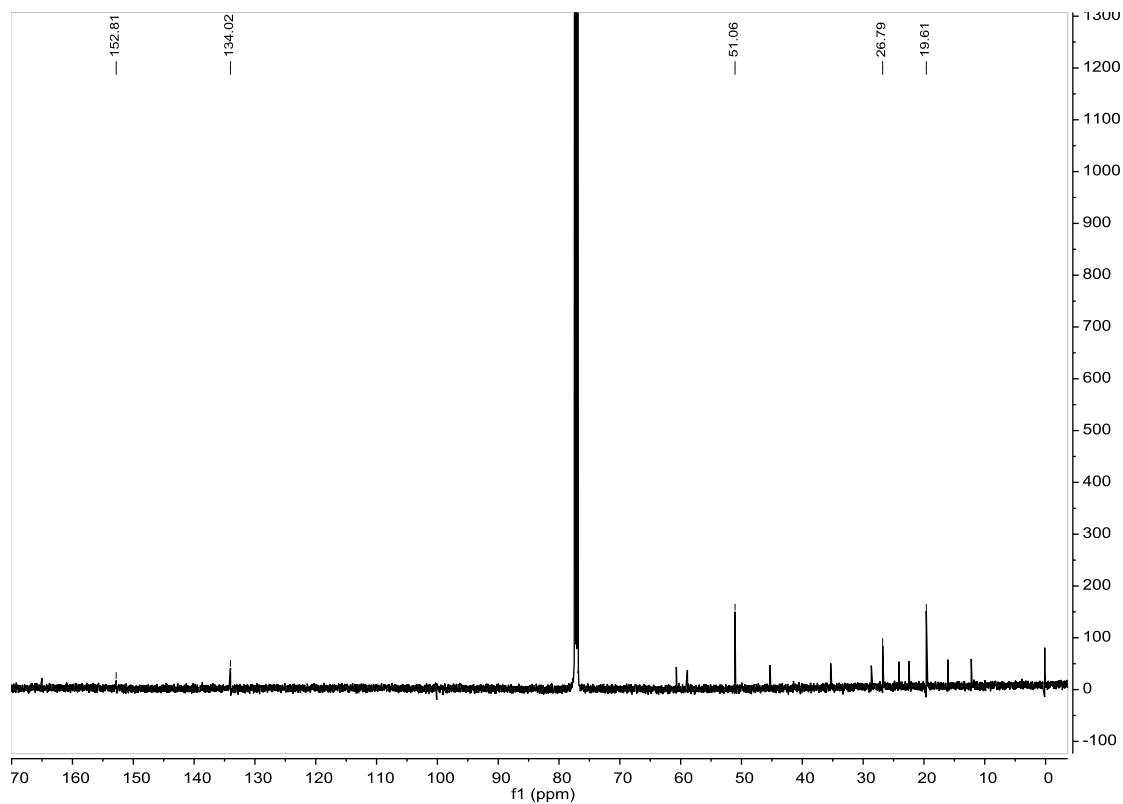
Appendix 9. ^{13}C NMR spectrum of isolated PNO A (**2**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CDCl$_3$ (151 MHz).
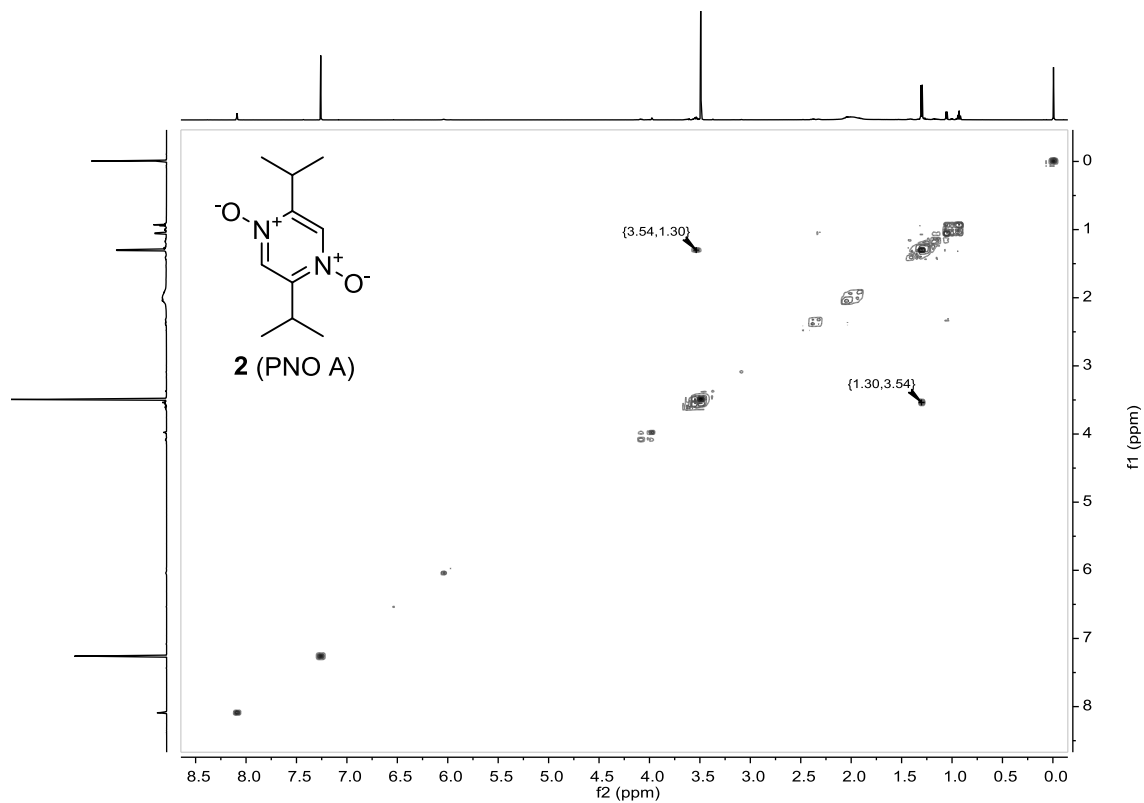
Appendix 10. (^{1}H, ^{1}H)-COSY NMR spectrum of isolated PNO A (**2**) from *P. entomophila ΔpvfC* +pPSV-*pvfABCD* in CDCl$_3$ (600 MHz).
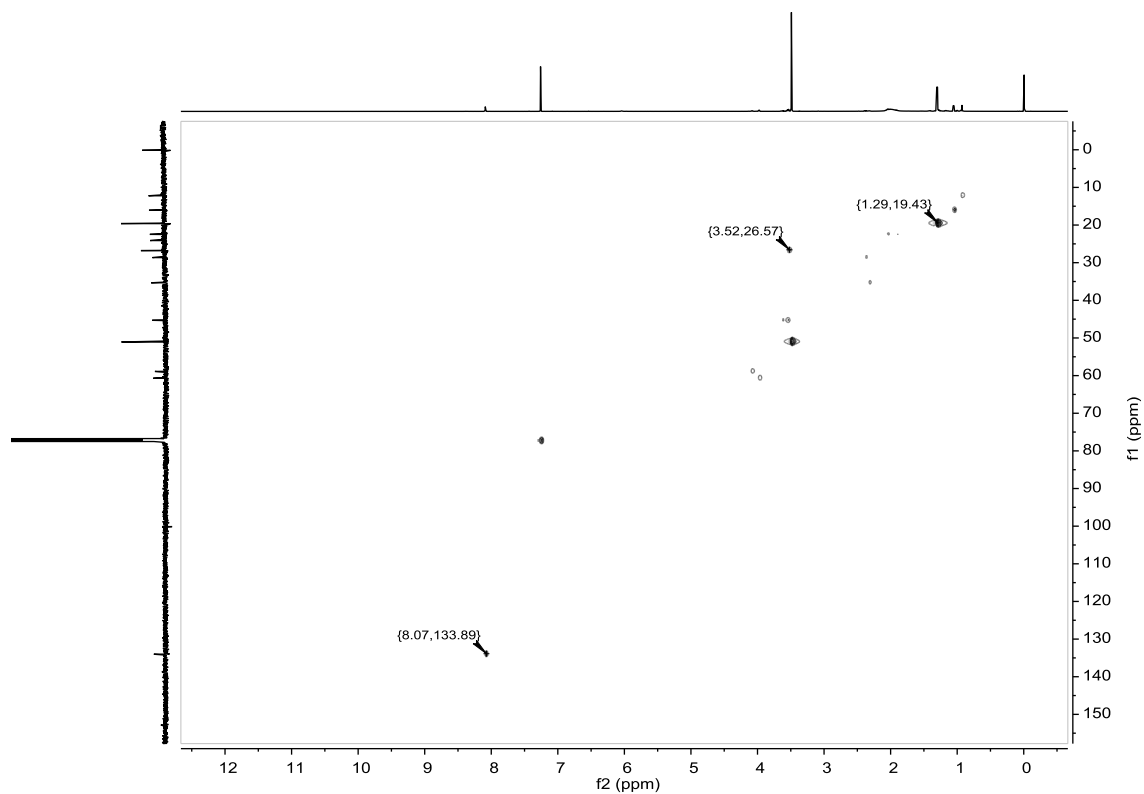
Appendix 11. (^{1}H, ^{13}C)-HSQC NMR spectrum of isolated PNO A (**2**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CDCl$_3$ (^{1}H, 600 MHz, ^{13}C 151 MHz).
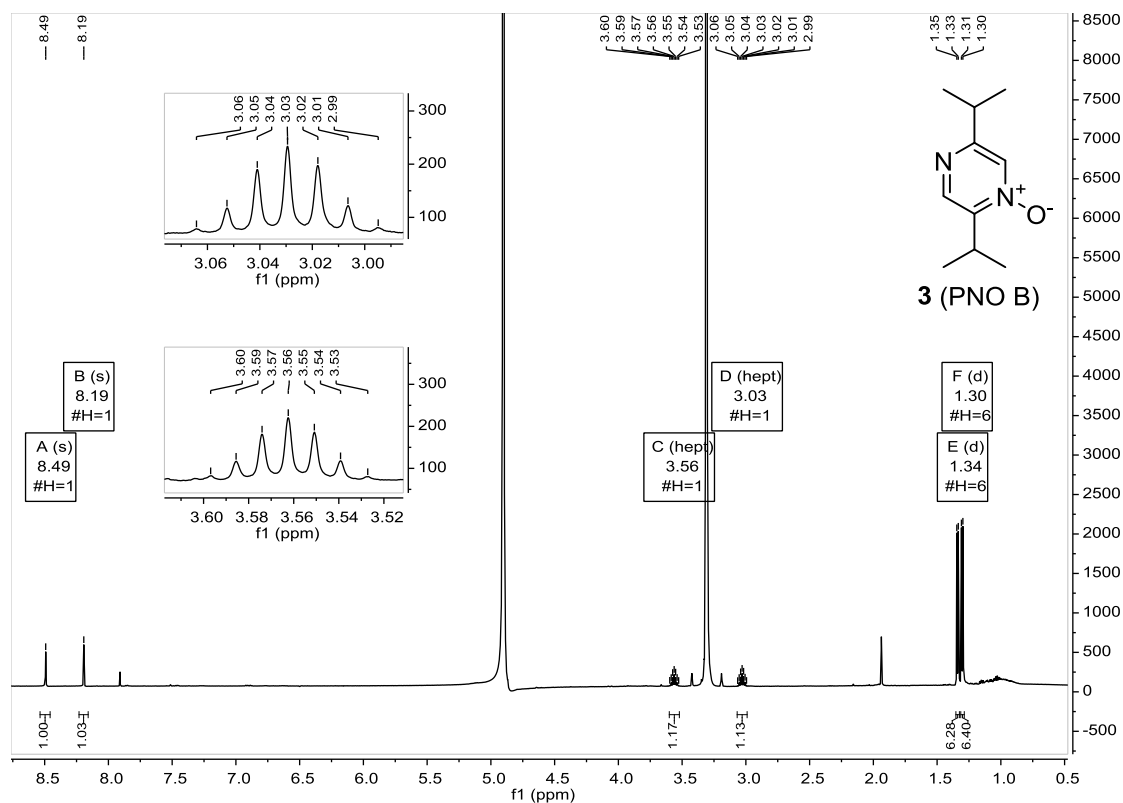
Appendix 12. ¹H NMR spectrum of isolated PNO B (**3**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD₃OD (600 MHz*).
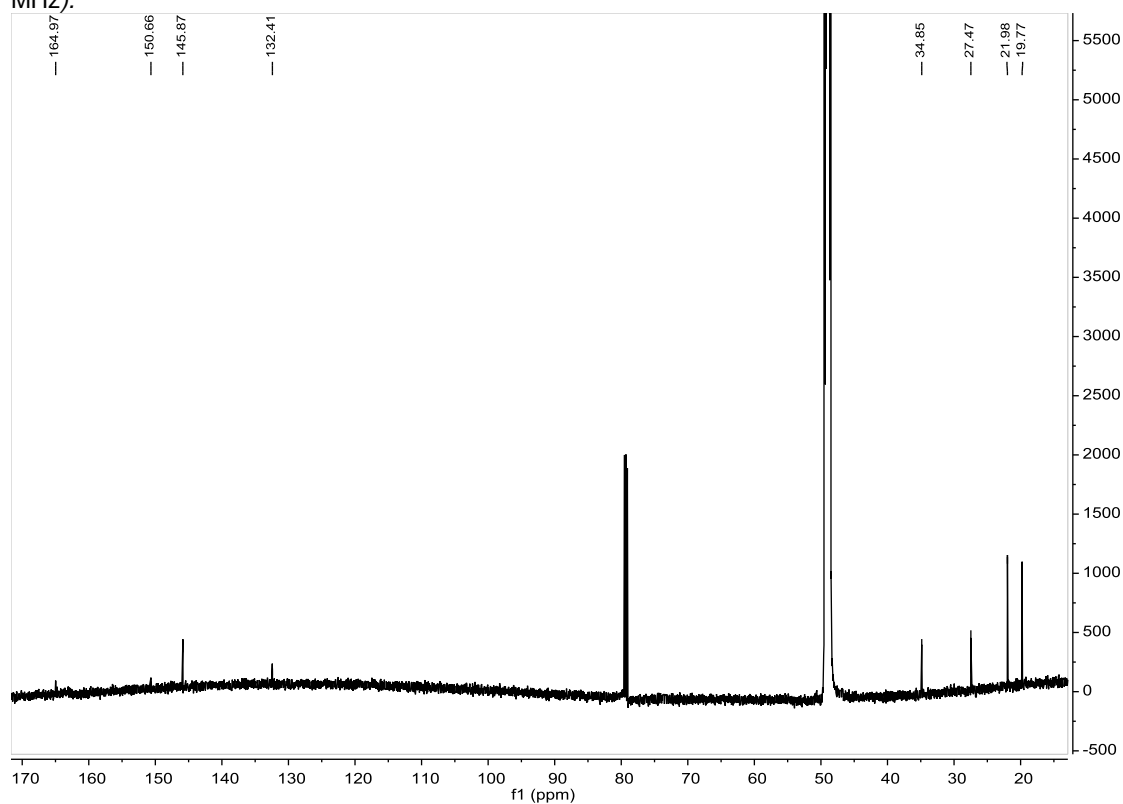
Appendix 13. ¹³C NMR spectrum of isolated PNO B (**3**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD₃OD (151 MHz).
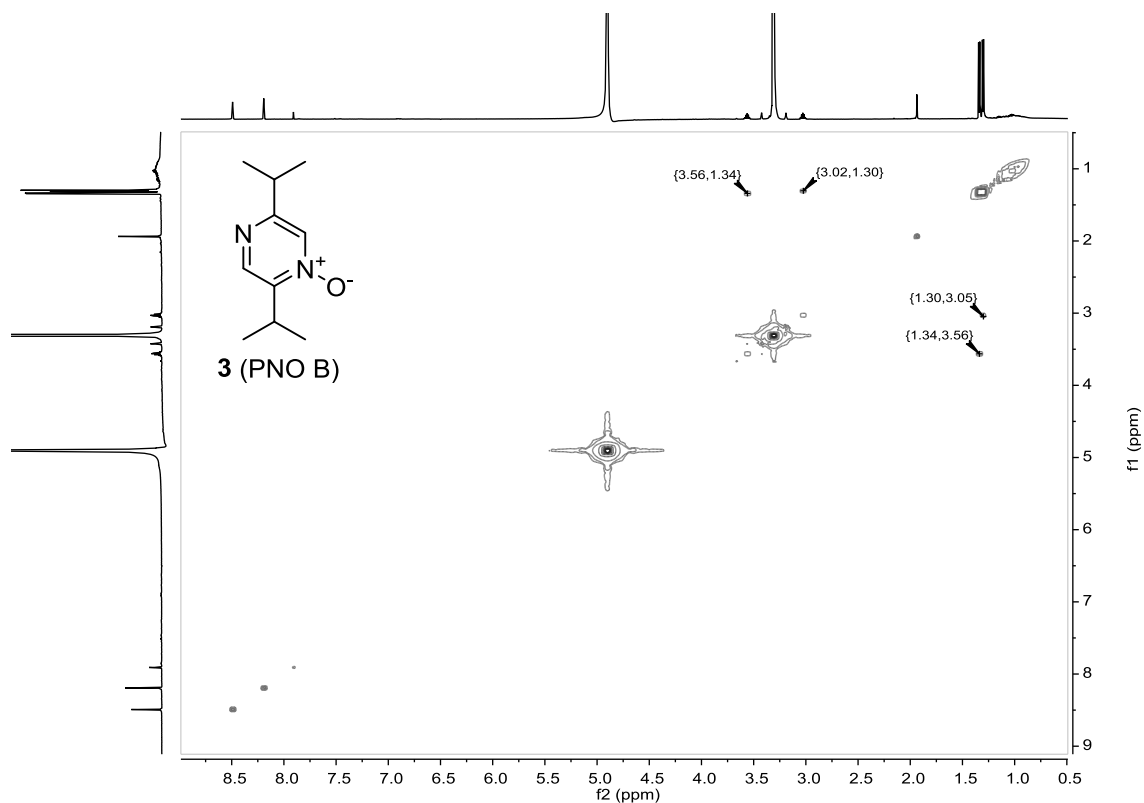
Appendix 14. (¹H, ¹H)-COSY NMR spectrum of isolated PNO B (**3**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD₃OD (600 MHz).
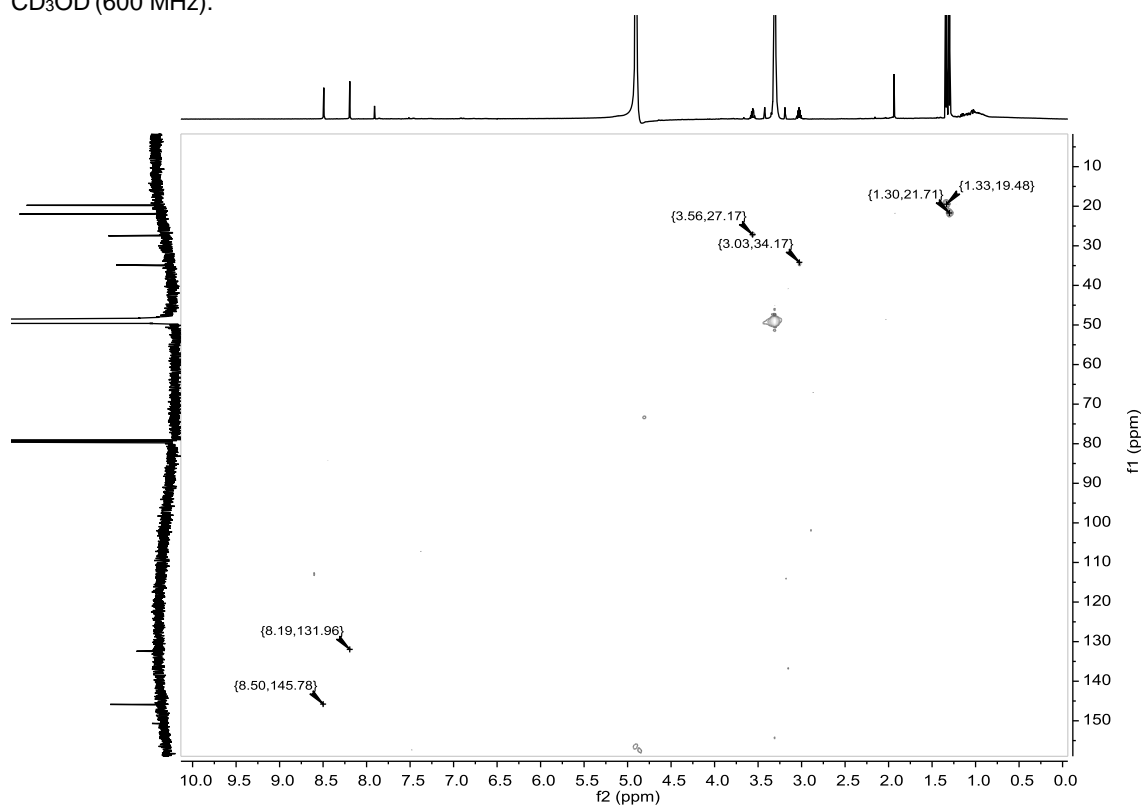
Appendix 15. (¹H, ¹³C)-HSQC NMR spectrum of isolated PNO B (**3**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD₃OD (¹H, 600 MHz, ¹³C 151 MHz).
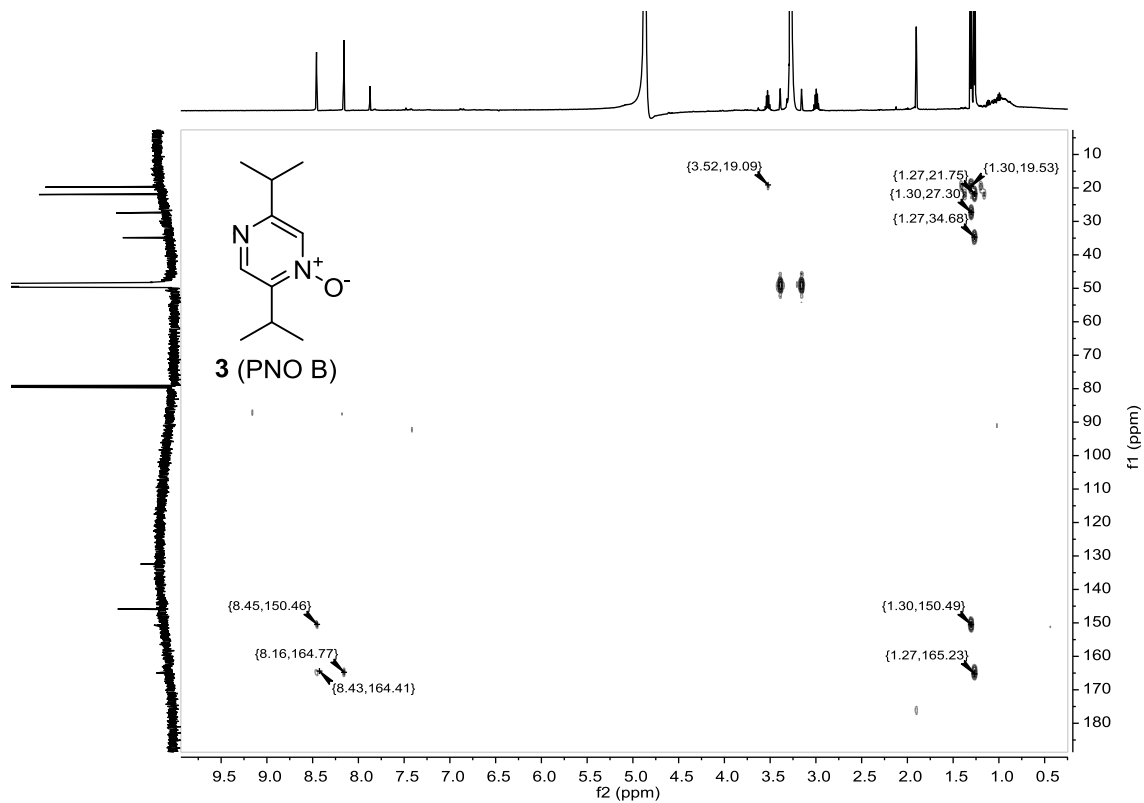
Appendix 16. (^{1}H, ^{13}C)-HMBC NMR spectrum of isolated PNO B (**3**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (^{1}H, 600 MHz, ^{13}C 151 MHz).
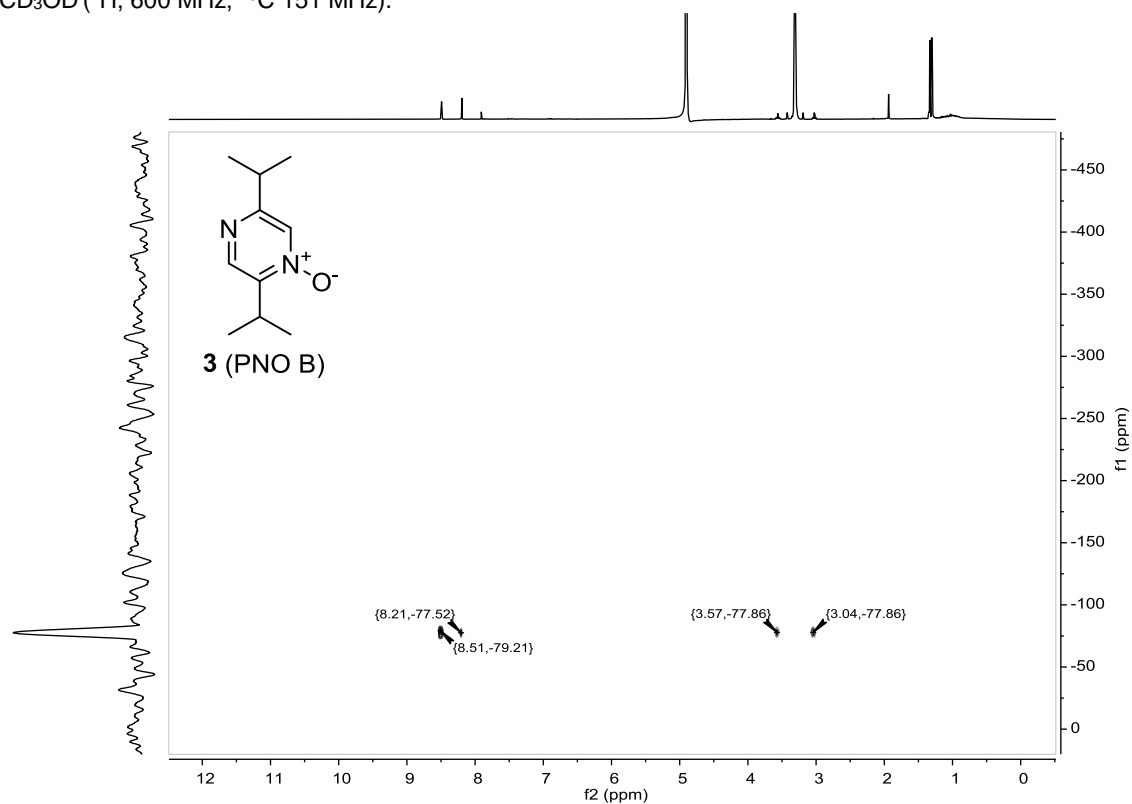
Appendix 17. (^{1}H, ^{15}N)-HMBC NMR spectrum of isolated PNO B (**3**) from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (^{1}H, 600 MHz, ^{15}N, 61 MHz).
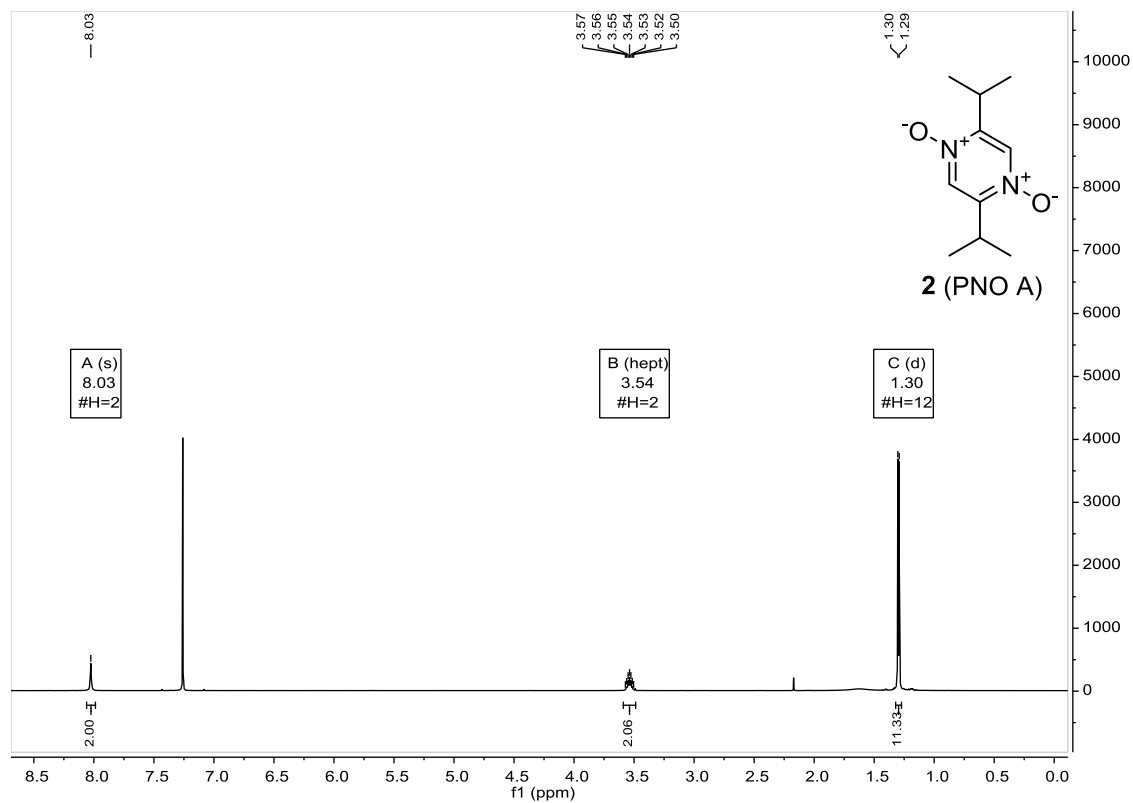
Appendix 18. ^{1}H NMR spectrum of synthetic PNO A (**2**) in CDCl$_3$ (600 MHz).
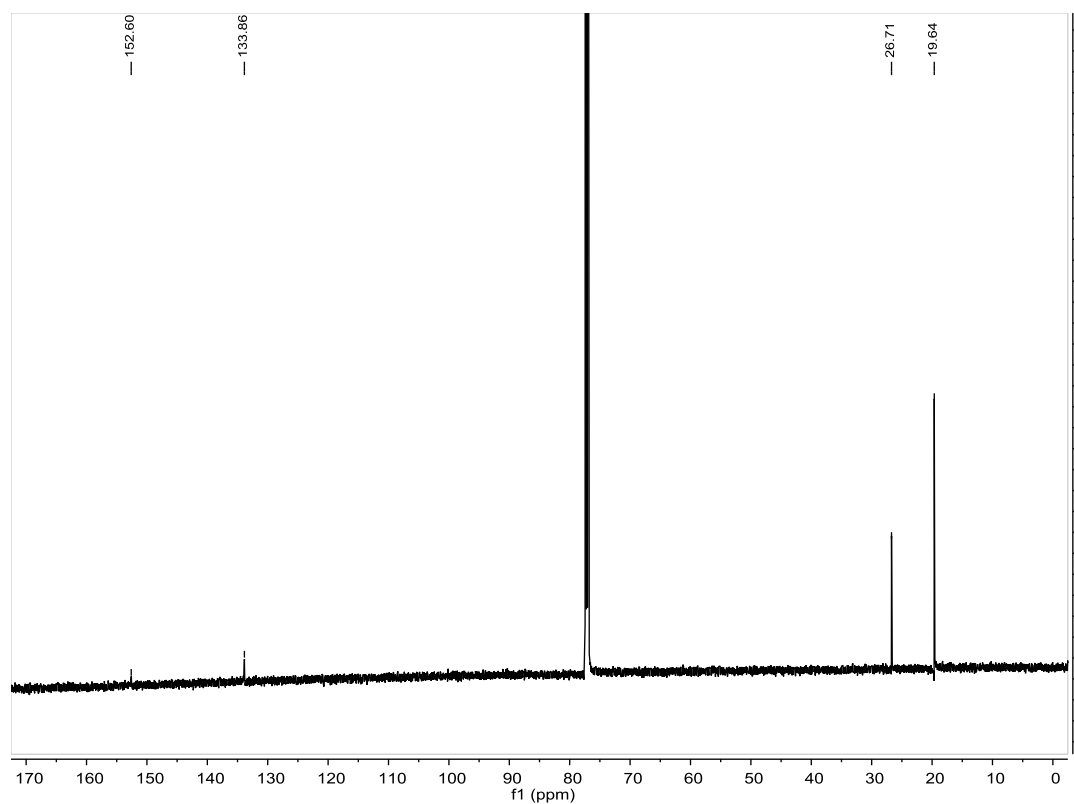
Appendix 19. ^{13}C NMR spectrum of synthetic PNO A (**2**) in CDCl$_3$ (151 MHz).
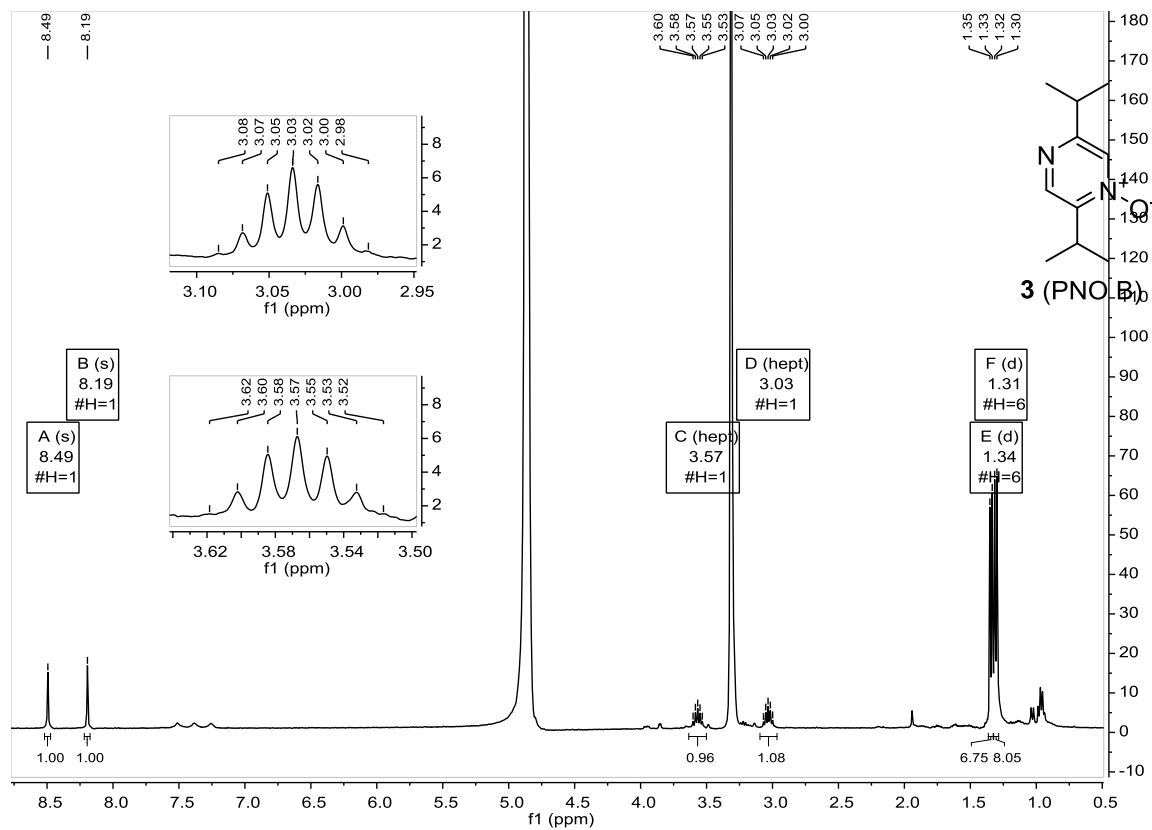
Appendix 20. ¹H NMR spectrum of synthetic PNO B (**3**) in CD₃OD (600 MHz).
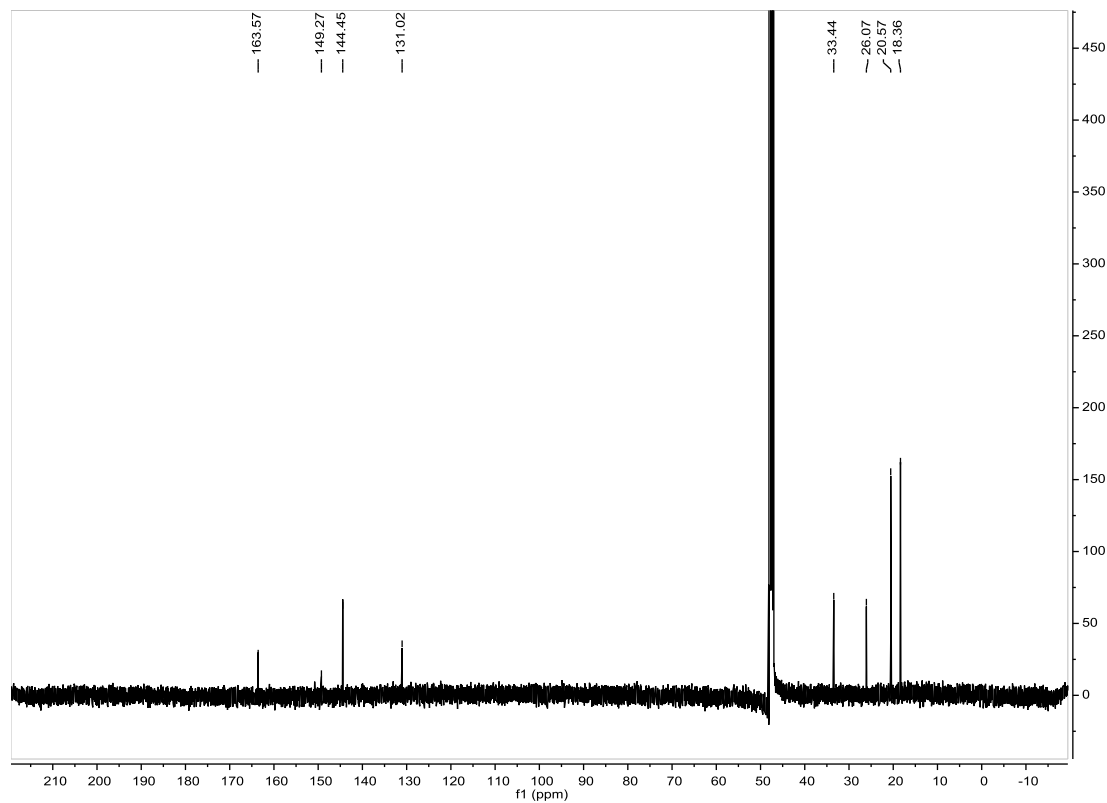
Appendix 21. ¹³C NMR spectrum of synthetic PNO B (**3**) in CD₃OD (151 MHz).

Appendix 22. Proteins that are significantly upregulated by the *pvf* signaling pathway

WT-KO	WT-OE	ANOVA -log p value	KO	WT	OE	Description	Gene Name
8.83	-0.26	4.67	21.00	29.83	30.10	Q1IAL9_PSEE4 Ferric siderophore ABC transporter, periplasmic siderophore-binding protein bauB OS	bauB
8.61	-2.22	5.07	21.83	30.44	32.66	Q1IFR7_PSEE4 Uncharacterized protein OS	PSEEN0540
6.62	-5.66	4.99	21.08	27.70	33.36	Q1I3W7_PSEE4 Uncharacterized protein OS	PSEEN5035
6.58	-4.23	3.79	20.46	27.04	31.27	Q1I7Q4_PSEE4 Uncharacterized protein OS	PSEEN3594
6.34	-0.15	4.52	22.13	28.46	28.61	Q1I389_PSEE4 Uncharacterized protein OS	PSEEN5274
6.21	-3.77	2.67	28.69	34.90	38.67	Q1ID47_PSEE4 Alkaline metalloprotease AprA OS	aprA
6.16	-0.96	5.21	22.53	28.68	29.64	Q1IFR6_PSEE4 Uncharacterized protein OS	PSEEN0541
5.55	-0.51	3.98	30.10	35.65	36.17	Q1I387_PSEE4 Hemolysin co-regulated protein 1A OS	PSEEN5276
5.32	-3.37	4.28	22.75	28.06	31.43	Q1IAF1_PSEE4 Uncharacterized protein OS	PSEEN2567
4.45	-4.20	5.08	21.77	26.23	30.43	Q1ICX3_PSEE4 Uncharacterized protein OS	PSEEN1634
4.32	-4.11	4.47	20.70	25.03	29.13	Q1IA91_PSEE4 Putative hydrolase, TatD family OS	PSEEN2638
4.27	-0.89	2.76	20.99	25.26	26.15	Q1IER7_PSEE4 Putative phosphotransferase system FruB, fructose-specific EI/HPr/EIIA components OS	PSEEN0932
4.22	-0.98	3.93	22.18	26.40	27.38	Q1I388_PSEE4 Uncharacterized protein OS	PSEEN5275
4.00	-4.33	6.79	22.46	26.46	30.78	Q1I3W6_PSEE4 Uncharacterized protein OS	PSEEN5036
3.86	-6.64	5.67	20.53	24.39	31.04	Q1IAN1_PSEE4 Uncharacterized protein OS	PSEEN2480
3.23	-0.13	2.31	24.01	27.24	27.37	Q1I3P4_PSEE4 Putative acyl-CoA dehydrogenase OS	PSEEN5113
3.22	-1.91	3.40	23.03	26.26	28.17	Q1I6I2_PSEE4 Uncharacterized protein OS	PSEEN4057
3.22	-1.95	2.39	21.65	24.86	26.82	Q1IH16_PSEE4 Uncharacterized protein OS	PSEEN0042
3.17	-0.51	2.71	26.88	30.05	30.57	Q1I3J9_PSEE4 Sarcosine oxidase (Alpha subunit) oxidoreductase protein OS	soxA
3.13	-0.51	3.12	26.76	29.89	30.41	Q1I3P2_PSEE4 Putative acyl-CoA dehydrogenase OS	PSEEN5115
3.11	-0.25	4.07	20.99	24.10	24.36	Q1I4P5_PSEE4 Response regulator CbrB OS	cbrB
3.07	-0.98	2.94	25.73	28.80	29.78	Q1I3A3_PSEE4 Putative amino acid ABC transporter, periplasmic amino acid-binding protein OS	PSEEN5258
3.04	-2.79	2.60	23.09	26.14	28.93	Q1I714_PSEE4 2-oxoisovalerate dehydrogenase, alpha subunit OS	bkdA1
3.00	-3.30	3.36	22.41	25.41	28.71	NUOCD_PSEE4 NADH-quinone oxidoreductase subunit C/D OS	nuoC
2.92	-0.74	2.98	25.26	28.19	28.93	Q1I3J7_PSEE4 Sarcosine oxidase beta subunit OS	soxB
2.92	-3.66	4.48	22.61	25.53	29.19	Q1I755_PSEE4 Putative anti-sigma F factor antagonist OS	PSEEN3811
2.84	-2.58	3.95	22.88	25.72	28.30	Q1IAM0_PSEE4 TonB-dependent siderophore receptor BauA OS	bauA
2.78	-2.53	4.47	25.56	28.34	30.86	Q1IBM0_PSEE4 Putative monoamine oxidase OS	PSEEN2116
2.66	-0.98	5.59	21.21	23.88	24.86	Q1IA94_PSEE4 Uncharacterized protein OS	PSEEN2635
2.56	-0.86	2.80	25.18	27.74	28.60	Q1I3K0_PSEE4 Sarcosine oxidase, gamma subunit OS	soxG
2.47	-1.85	3.30	24.53	27.00	28.85	DNAJ_PSEE4 Chaperone protein DnaJ OS	dnaJ
2.45	-2.26	2.63	21.32	23.77	26.03	Q1IAH9_PSEE4 Uncharacterized protein OS	PSEEN2538
2.36	-0.46	2.07	21.24	23.60	24.05	Q1I9G9_PSEE4 Putative methyl-accepting chemotaxis transducer OS	PSEEN2935
2.28	-2.15	3.16	26.64	28.92	31.07	Q1IEM6_PSEE4 Uncharacterized protein OS	PSEEN0974
2.27	-0.99	2.58	24.23	26.50	27.49	ACCD_PSEE4 Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta OS	accD

2.25	-0.20	3.70	23.51	25.76	25.97	Q1I437_PSEE4 Putative flagellar motor MotB protein OS	PSEEN4957
2.19	-0.79	2.87	20.60	22.79	23.57	Q1I5H2_PSEE4 Carbamate kinase OS	arcC
2.16	-5.18	3.72	21.18	23.35	28.53	Q1I9J4_PSEE4 Uncharacterized protein OS	PSEEN2906
2.16	-0.53	2.86	21.51	23.67	24.20	Q1IDG7_PSEE4 Uncharacterized protein OS	PSEEN1420
2.12	-2.24	4.41	21.94	24.07	26.30	Q1I9E6_PSEE4 Uncharacterized protein OS	PSEEN2958
2.11	-0.54	3.22	22.10	24.21	24.75	Q1IEW2_PSEE4 Uncharacterized protein OS	PSEEN0885
2.10	-1.42	4.39	26.94	29.04	30.46	Q1I5S5_PSEE4 Uncharacterized protein OS	PSEEN4323
2.05	-4.96	5.40	21.22	23.27	28.23	Q1IAN2_PSEE4 Uncharacterized protein OS	PSEEN2479
2.04	-0.15	3.19	26.04	28.08	28.23	Q1ICS5_PSEE4 Uncharacterized protein OS	PSEEN1688
2.03	-1.09	3.00	28.19	30.22	31.32	Q1IEM7_PSEE4 Putative Dihydrodipicolinate synthase OS	PSEEN0973
2.01	-0.84	2.58	22.30	24.32	25.16	Q1I477_PSEE4 Uncharacterized protein OS	PSEEN4910
2.01	-2.89	2.16	22.62	24.63	27.52	Q1IA95_PSEE4 Uncharacterized protein OS	PSEEN2634
2.00	-0.61	4.03	27.27	29.27	29.87	Q1I4X6_PSEE4 Putative oxidoreductase, short-chain dehydrogenase/reductase family OS	PSEEN4636
1.98	-3.54	4.48	22.32	24.31	27.85	Q1IA49_PSEE4 Putative alcohol dehydrogenase OS	PSEEN2684
1.97	-5.93	4.26	23.20	25.18	31.11	Q1IBG9_PSEE4 Putative isoquinoline 1-oxidoreductase, beta subunit OS	PSEEN2171
1.97	-0.40	3.60	25.37	27.34	27.74	Q1ICD6_PSEE4 Uncharacterized protein OS	PSEEN1838
1.94	-4.77	4.97	22.42	24.37	29.14	Q1I4N1_PSEE4 Uncharacterized protein OS	PSEEN4746
1.94	-1.56	2.08	22.77	24.71	26.27	Q1IBY1_PSEE4 Putative acetyltransferase, GNAT family OS	PSEEN2001
1.93	-1.87	2.67	21.55	23.48	25.35	Q1I4I4_PSEE4 Uncharacterized protein OS	PSEEN4795
1.92	-2.33	4.04	25.96	27.88	30.21	Q1IEM2_PSEE4 Putative dehydrogenase OS	PSEEN0978
1.88	-1.27	2.65	22.53	24.41	25.68	HLDE_PSEE4 Bifunctional protein HldE OS	hldE
1.87	-2.85	2.89	20.45	22.32	25.17	Q1I3J0_PSEE4 Putative oxidoreductase, FAD-binding OS	PSEEN5168
1.82	-0.98	2.27	21.35	23.17	24.15	Q1I6Z1_PSEE4 Transcriptional regulator ArgR, AraC family OS	argR
1.80	-0.29	4.78	24.35	26.16	26.44	Q1ICS0_PSEE4 Uncharacterized protein OS	PSEEN1693
1.60	-1.14	2.02	25.98	27.59	28.72	Q1IB15_PSEE4 Uncharacterized protein OS	PSEEN2341
1.60	-0.03	3.65	25.68	27.28	27.31	Q1IFR5_PSEE4 Uncharacterized protein OS	PSEEN0542
1.59	-2.05	4.26	21.40	22.99	25.04	Q1IAS1_PSEE4 Uncharacterized protein OS	PSEEN2440
1.58	-2.96	2.55	22.22	23.81	26.77	Q1IFI4_PSEE4 Putative acetyltransferase, GNAT family OS	PSEEN0630
1.57	-3.38	2.52	21.65	23.22	26.60	PYRD_PSEE4 Dihydroorotate dehydrogenase (quinone) OS	pyrD
1.54	-8.36	5.61	19.02	20.56	28.92	Q1I9D0_PSEE4 Quinohemoprotein amine dehydrogenase 40 kDa subunit OS	PSEEN2974
1.52	-3.82	3.37	25.51	27.03	30.85	Q1I955_PSEE4 Uncharacterized protein OS	PSEEN3053
1.52	-1.93	2.00	25.26	26.78	28.70	Q1IC44_PSEE4 Uncharacterized protein OS	PSEEN1934
1.46	-2.49	5.35	24.61	26.07	28.56	Q1I8I7_PSEE4 Putative tautomerase OS	PSEEN3286
1.44	-1.97	5.47	25.26	26.70	28.67	Q1IGL9_PSEE4 Thiol-specific antioxidant protein LsfA OS	lsfA
1.42	-3.27	2.09	21.30	22.72	25.99	Q1I511_PSEE4 Uncharacterized protein OS	PSEEN4597
1.42	-1.82	4.07	28.06	29.48	31.31	Q1ID46_PSEE4 Inhibitor of protease aprA OS	aprl
1.41	-1.46	2.98	26.34	27.75	29.21	Q1IDB8_PSEE4 Uncharacterized protein OS	PSEEN1471
1.34	-2.83	4.81	27.38	28.72	31.55	Q1IH17_PSEE4 Uncharacterized protein OS	PSEEN0041
1.30	-0.20	4.50	29.88	31.19	31.38	Q1IF05_PSEE4 Uncharacterized protein OS	PSEEN0836

1.30	-0.43	2.13	26.02	27.33	27.76	Q1IA63_PSEE4 Aromatic amino acid aminotransferase OS	tyrB-2
1.30	-3.99	2.75	21.15	22.45	26.44	Q1IA03_PSEE4 Putative glutathione S-transferase OS	PSEEN2733
1.29	-0.26	2.22	29.68	30.97	31.22	Q1IC20_PSEE4 Putative inosine-uridine preferring nucleoside hydrolase OS	PSEEN1958
1.26	-4.03	4.58	24.47	25.73	29.77	Q1I359_PSEE4 Glycine dehydrogenase (decarboxylating) OS	gcvP-1
1.24	-0.65	2.21	23.16	24.40	25.05	Q1IGY0_PSEE4 Putative cytochrome c-type protein OS	PSEEN0079
1.24	-0.45	4.55	31.12	32.36	32.81	Q1IGQ6_PSEE4 Putative dipeptidase OS	PSEEN0172
1.22	-2.33	2.00	29.47	30.69	33.03	Q1I5H3_PSEE4 Ornithine carbamoyltransferase, catabolic OS	argF
1.21	-1.55	2.17	23.22	24.43	25.97	Q1I3P0_PSEE4 ATP-dependent dethiobiotin synthetase BioD OS	bioD
1.18	-0.13	3.20	28.24	29.42	29.55	Q1I724_PSEE4 Putative exoprotein with autotransporter OS	PSEEN3843
1.18	-3.39	3.20	25.88	27.06	30.45	Q1IDX9_PSEE4 Putative oxidireductase OS	PSEEN1235
1.18	-0.95	2.32	21.17	22.35	23.30	QUEF_PSEE4 NADPH-dependent 7-cyano-7-deazaguanine reductase OS	queF
1.16	-4.84	3.92	22.66	23.82	28.67	Q1I2K9_PSEE4 Cyclopropane-fatty-acyl-phospholipid synthase OS	PSEEN5518
1.11	-2.19	4.58	25.91	27.02	29.21	Q1IAF3_PSEE4 Putative Ribonuclease OS	PSEEN2564
1.10	-4.32	3.24	21.01	22.11	26.43	Q1IF33_PSEE4 Putative cobalamin synthesis protein/P47K family protein OS	PSEEN0805
1.08	-1.11	2.90	27.41	28.49	29.60	Q1IGR7_PSEE4 Alginate regulatory protein AlgP OS	algP
1.08	-1.61	5.59	27.78	28.86	30.47	Q1IDN9_PSEE4 Putative hydrolase probable dienelactone hydrolase OS	PSEEN1342
1.07	-5.77	4.74	21.91	22.98	28.75	Q1ID81_PSEE4 Putative zinc-containing alcohol dehydrogenase OS	PSEEN1511
1.06	-0.70	4.04	25.43	26.49	27.19	Q1ICF1_PSEE4 Uncharacterized protein OS	PSEEN1821
1.05	-0.19	2.37	26.26	27.31	27.50	Q1IBH3_PSEE4 Uncharacterized protein OS	PSEEN2167
1.00	-0.18	4.15	25.04	26.04	26.22	Q1I2X2_PSEE4 Uncharacterized protein OS	PSEEN5399

Appendix 23. Proteins that are significantly downregulated by the *pvf* signaling pathway

WT-KO	WT-OE	ANOVA -log p value	KO	WT	OE	Description	Gene Name
-9.07	3.45	2.95	33.69	24.62	21.16	Q1I720_PSEE4 Negative regulator of flagellin synthesis FlgM OS	flgM
-8.16	2.69	3.79	32.57	24.41	21.72	RL32_PSEE4 50S ribosomal protein L32 OS	rpmF
-7.17	1.16	4.40	29.22	22.06	20.90	Q1I730_PSEE4 Flagellar basal-body rod protein FlgF OS	flgF
-6.90	1.66	3.46	30.76	23.86	22.20	Q1I758_PSEE4 Flagellar hook-length control protein FliK	fliK
-6.63	4.63	2.92	30.99	24.37	19.74	Q1ICA5_PSEE4 DNA-binding protein HU-beta OS	hupB
-6.02	0.93	2.18	29.29	23.27	22.34	Y1604_PSEE4 UPF0434 protein PSEEN1604 OS	PSEEN1604
-5.05	1.86	2.40	29.51	24.46	22.60	Q1I2U6_PSEE4 50S ribosomal protein L33 OS	rpmG
-4.55	0.51	2.51	27.31	22.76	22.24	Q1I3J2_PSEE4 Uncharacterized protein OS	PSEEN5166
-4.30	5.39	3.46	29.79	25.49	20.10	Q1I457_PSEE4 Putative iron ABC transporter, periplasmic iron-binding protein OS	PSEEN4935
-4.28	4.85	2.94	33.05	28.77	23.92	Q1I5F4_PSEE4 DNA-binding protein HU form N OS	hupN
-4.26	2.42	3.12	34.41	30.15	27.73	Q1I3P1_PSEE4 Uncharacterized protein OS	PSEEN5116
-4.17	1.90	2.51	30.26	26.09	24.19	Q1I2Z5_PSEE4 Uncharacterized protein OS	PSEEN5374
-3.96	4.48	3.39	30.03	26.07	21.59	Q1I509_PSEE4 Uncharacterized protein OS	PSEEN4601
-3.83	0.61	3.84	30.20	26.36	25.75	Q1I8W2_PSEE4 Peptidyl-prolyl cis-trans isomerase (PPIase) (Rotamase) OS	ppiC-2
-3.78	0.96	2.66	31.68	27.90	26.94	Q1IA61_PSEE4 Probable thiol peroxidase OS	tpx
-3.77	3.65	4.17	28.84	25.07	21.42	Q1I5R9_PSEE4 Glycerophosphodiester phosphodiesterase OS	PSEEN4329
-3.67	1.44	2.14	27.02	23.35	21.92	Q1I679_PSEE4 Putative phage tail fiber assembly protein OS	PSEEN4167
-3.47	4.15	2.35	30.65	27.18	23.03	RPOZ_PSEE4 DNA-directed RNA polymerase subunit omega OS	rpoZ
-3.39	5.17	3.47	30.53	27.13	21.97	RL6_PSEE4 50S ribosomal protein L6 OS	rplF
-3.39	1.76	2.24	27.45	24.06	22.31	RL30_PSEE4 50S ribosomal protein L30 OS	rpmD
-3.34	1.91	2.29	28.22	24.87	22.96	Q1I4Y8_PSEE4 Uncharacterized protein OS	PSEEN4624
-3.29	0.69	3.06	28.47	25.18	24.49	Q1I803_PSEE4 Putative type II secretion system protein D putative signal peptide OS	PSEEN3480
-3.24	1.22	2.14	28.53	25.29	24.06	Q1I506_PSEE4 Uncharacterized protein OS	PSEEN4605
-3.17	1.97	2.66	31.60	28.42	26.45	Q1I635_PSEE4 Putative outer membrane protein OmpH-like OS	PSEEN4211
-3.09	3.43	3.56	29.11	26.02	22.58	Q1I710_PSEE4 Uncharacterized protein OS	PSEEN3857
-3.09	0.80	2.12	27.98	24.89	24.08	Q1I5K3_PSEE4 Putative peptidase, M23/M37 family OS	PSEEN4398
-3.06	2.47	2.89	27.69	24.63	22.16	RS11_PSEE4 30S ribosomal protein S11 OS	rpsK
-3.01	5.58	2.43	30.09	27.08	21.51	RL13_PSEE4 50S ribosomal protein L13 OS	rplM
-2.92	2.17	3.29	29.42	26.49	24.33	Q1IAZ0_PSEE4 Probable transcriptional regulatory protein PSEEN2368 OS	PSEEN2368
-2.90	3.04	2.40	27.31	24.41	21.37	Q1IG63_PSEE4 Putative metalloendopeptidase, M23/M37 family OS	PSEEN0379
-2.81	1.50	2.95	32.68	29.87	28.37	RL24_PSEE4 50S ribosomal protein L24 OS	rplX
-2.80	0.66	3.05	27.88	25.09	24.43	Q1I3U9_PSEE4 Type IV pili response regulator PilH OS	pilH
-2.76	1.37	3.25	32.39	29.63	28.26	CH10_PSEE4 10 kDa chaperonin OS	groS
-2.72	3.90	3.65	29.03	26.31	22.40	Q1IAX9_PSEE4 Uncharacterized protein OS	PSEEN2379
-2.68	0.79	2.00	28.56	25.88	25.09	Q1IG13_PSEE4 Uncharacterized protein OS	PSEEN0435

-2.58	5.69	3.78	29.18	26.60	20.91	Q1I699_PSEE4 Putative Mu-like phage protein gp38 OS	PSEEN4147
-2.54	0.95	2.94	25.49	22.95	21.99	Q1I2P8_PSEE4 Uncharacterized protein OS	PSEEN5477
-2.49	0.18	2.48	29.17	26.68	26.49	Q1I6Z7_PSEE4 Uncharacterized protein OS	PSEEN3877
-2.39	4.58	2.74	27.84	25.45	20.86	Q1I727_PSEE4 Basal-body rod modification protein FlgD OS	flgD
-2.27	2.23	2.65	27.85	25.58	23.35	Q1IGH5_PSEE4 Putative TonB-dependent siderophore receptor OS	PSEEN0263
-2.24	4.73	2.17	28.57	26.33	21.60	Q1IBE8_PSEE4 Putative acetyl transferase, GNAT family OS	PSEEN2196
-2.15	1.30	3.24	27.88	25.73	24.43	Q1I6H3_PSEE4 Uncharacterized protein OS	PSEEN4067
-2.14	4.95	2.16	29.90	27.76	22.81	RL23_PSEE4 50S ribosomal protein L23 OS	rplW
-2.05	2.78	3.14	26.92	24.86	22.08	Q1I613_PSEE4 Putative 23S rRNA m1G745 methyltransferase RrmA OS	PSEEN4233
-2.03	1.32	2.83	24.56	22.54	21.22	Q1IGK4_PSEE4 Heat shock protein 15 OS	hslR
-1.95	1.95	2.46	32.22	30.28	28.33	Q1I736_PSEE4 Flagellar hook-associated protein FlgL OS	flgL
-1.84	0.44	3.29	26.29	24.45	24.01	Q1IGZ3_PSEE4 Putative electron transport protein, Sco1/SenC family OS	PSEEN0065
-1.83	3.23	3.71	29.41	27.59	24.35	Q1I324_PSEE4 Uncharacterized protein OS	PSEEN5345
-1.82	2.48	2.55	26.92	25.09	22.62	Q1IF50_PSEE4 ATP-dependent zinc metalloprotease FtsH OS	ftsH
-1.73	4.00	3.69	31.95	30.22	26.22	Q1I742_PSEE4 Flagellin FlaG OS	PSEEN3825
-1.66	2.11	2.06	25.11	23.45	21.34	Q1I7R0_PSEE4 Putative TonB-dependent siderophore receptor OS	PSEEN3587
-1.64	3.57	3.89	33.25	31.61	28.05	Q1IC65_PSEE4 Uncharacterized protein OS	PSEEN1912
-1.53	4.52	3.35	32.75	31.22	26.70	RL10_PSEE4 50S ribosomal protein L10 OS	rplJ
-1.49	7.00	3.71	30.12	28.63	21.62	Q1I611_PSEE4 Putative cold shock protein CspE-like OS	PSEEN4235
-1.47	0.36	2.83	25.10	23.62	23.26	APAH_PSEE4 Bis(5-nucleosyl)-tetraphosphatase, symmetrical OS	apaH
-1.46	1.76	2.44	30.93	29.47	27.71	Q1IF54_PSEE4 Transcription elongation factor GreA OS	greA
-1.40	2.36	2.20	27.96	26.57	24.20	RS5_PSEE4 30S ribosomal protein S5 OS	rpsE
-1.34	0.51	2.58	27.43	26.09	25.58	RBFA_PSEE4 Ribosome-binding factor A OS	rbfA
-1.30	4.31	3.02	33.08	31.78	27.47	RL1_PSEE4 50S ribosomal protein L1 OS	rplA
-1.27	3.79	2.35	27.12	25.85	22.06	Q1I2S6_PSEE4 Putative lipoprotein OS	PSEEN5449
-1.26	4.12	3.12	30.93	29.67	25.55	RS8_PSEE4 30S ribosomal protein S8 OS	rpsH
-1.26	1.59	2.73	25.86	24.60	23.01	Q1I700_PSEE4 Carbon storage regulator homolog OS	csrA
-1.26	1.98	2.26	30.71	29.45	27.47	Q1I3X4_PSEE4 Uncharacterized protein OS	PSEEN5028
-1.24	0.83	3.45	29.45	28.21	27.37	Q1IDF7_PSEE4 Putative carboxy-terminal protease for penicillin-binding protein 3 Prc OS	PSEEN1431
-1.23	2.10	2.80	31.34	30.10	28.00	Q1IGA2_PSEE4 Uncharacterized protein OS	PSEEN0340
-1.22	1.58	3.19	25.20	23.98	22.40	Q1I7G2_PSEE4 Putative VacJ lipoprotein OS	PSEEN3699
-1.21	3.08	2.72	31.64	30.42	27.34	Q1I607_PSEE4 Uncharacterized protein OS	PSEEN4239
-1.21	0.59	2.92	27.88	26.67	26.08	Q1I2Y8_PSEE4 Uncharacterized protein OS	PSEEN5382
-1.20	2.65	2.40	29.67	28.47	25.82	Q1IF52_PSEE4 Uncharacterized protein OS	PSEEN0785
-1.17	3.38	2.23	28.35	27.17	23.80	ERPA_PSEE4 Iron-sulfur cluster insertion protein ErpA OS	erpA
-1.14	3.32	4.49	25.02	23.87	20.55	Q1IE29_PSEE4 Putative hydrolase, MutT/nudix family protein OS	PSEEN1183
-1.14	3.13	2.54	25.28	24.14	21.01	Q1IBE0_PSEE4 Cold-shock protein CspD OS	cspD
-1.10	3.69	3.53	27.98	26.88	23.19	YACG_PSEE4 DNA gyrase inhibitor YacG OS	yacG

-1.10	0.68	2.41	29.68	28.59	27.91	FETP_PSEE4 Probable Fe(2+)-trafficking protein OS	PSEEN5200
-1.09	1.06	2.92	27.07	25.98	24.92	Q1IGR8_PSEE4 Peptidyl-prolyl cis-trans isomerase OS	PSEEN0160
-1.08	3.07	2.29	28.92	27.84	24.77	Y1082_PSEE4 UPF0307 protein PSEEN1082 OS	PSEEN1082
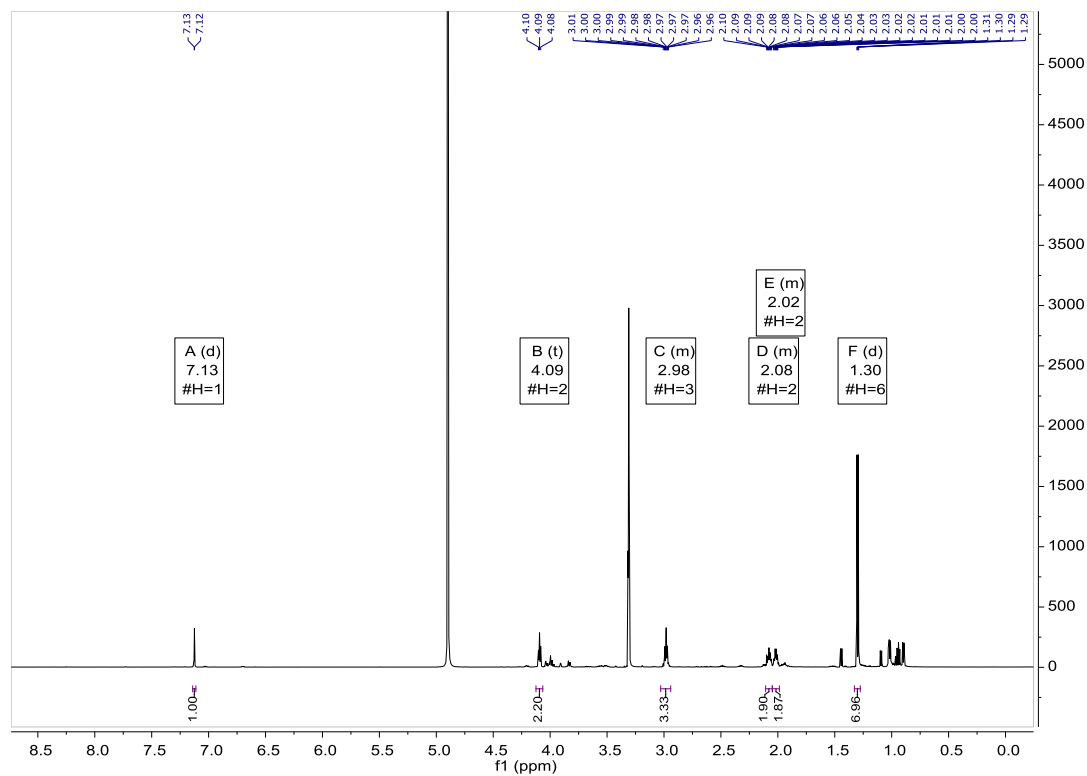
Appendix 24. ^{1}H NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (600 MHz).
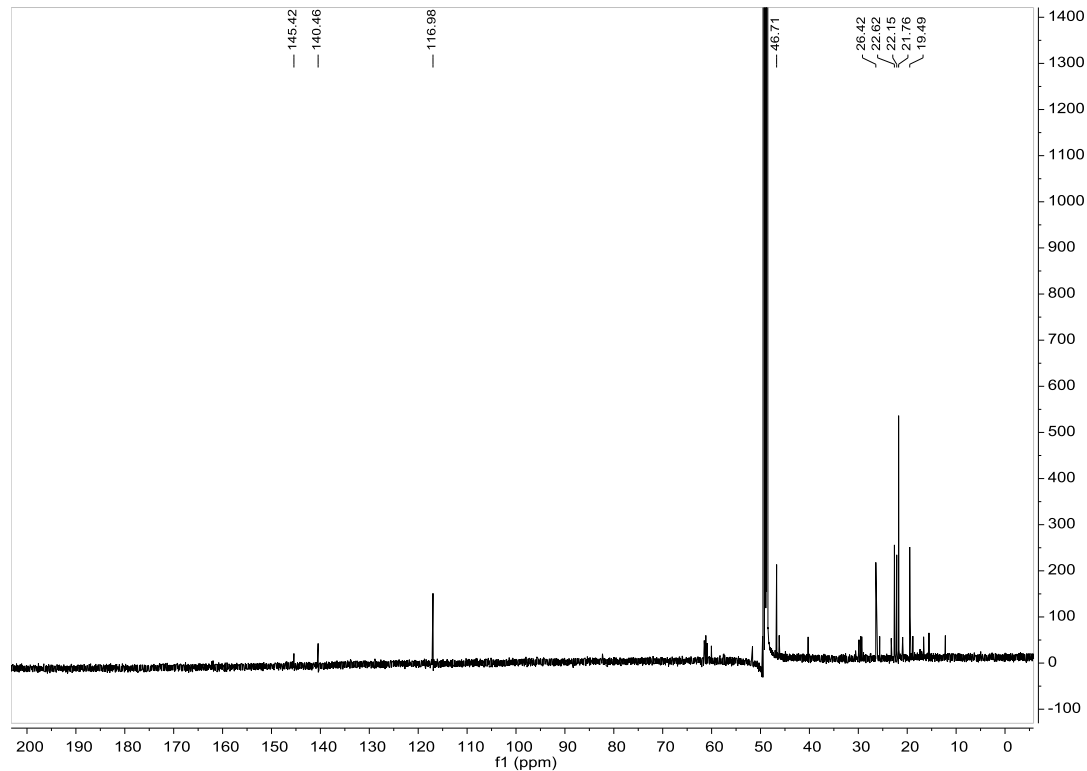
Appendix 25. ^{13}C NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (151 MHz).
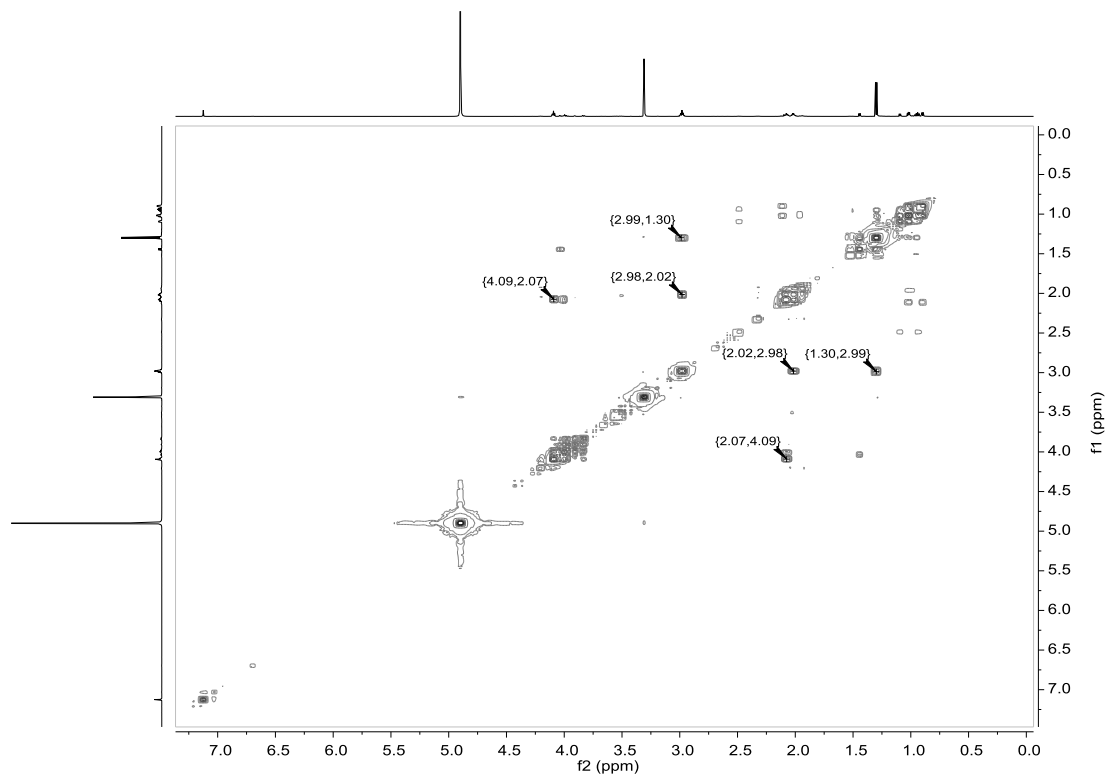
Appendix 26. (^{1}H, ^{1}H)-COSY NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (600 MHz).

Appendix 27. (^{1}H, ^{13}C)-HSQC NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (^{1}H, 600 MHz, ^{13}C 151 MHz).

Appendix 28. (^{1}H, ^{13}C)-HMBC NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila* *ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (^{1}H, 600 MHz, ^{13}C 151 MHz)
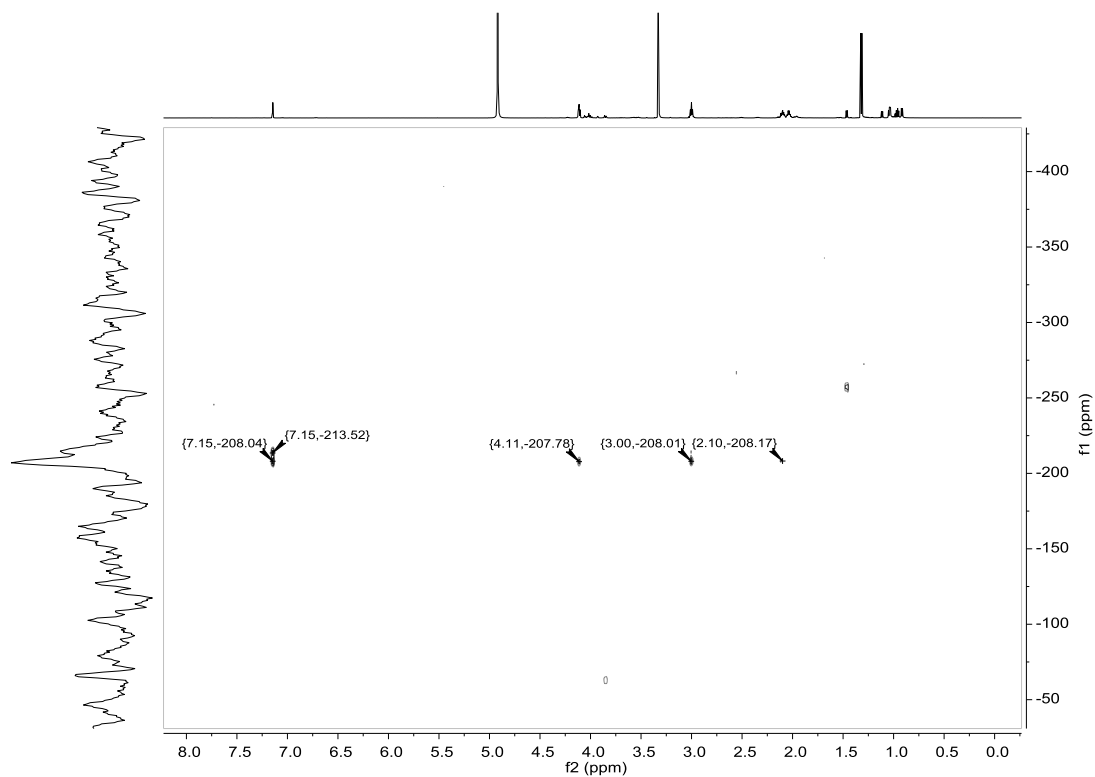
Appendix 29. (^{1}H, ^{15}N)-HMBC NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila* *ΔpvfC* + pPSV-*pvfABCD* in CD$_3$OD (^{1}H, 600 MHz, ^{15}N, 61 MHz).
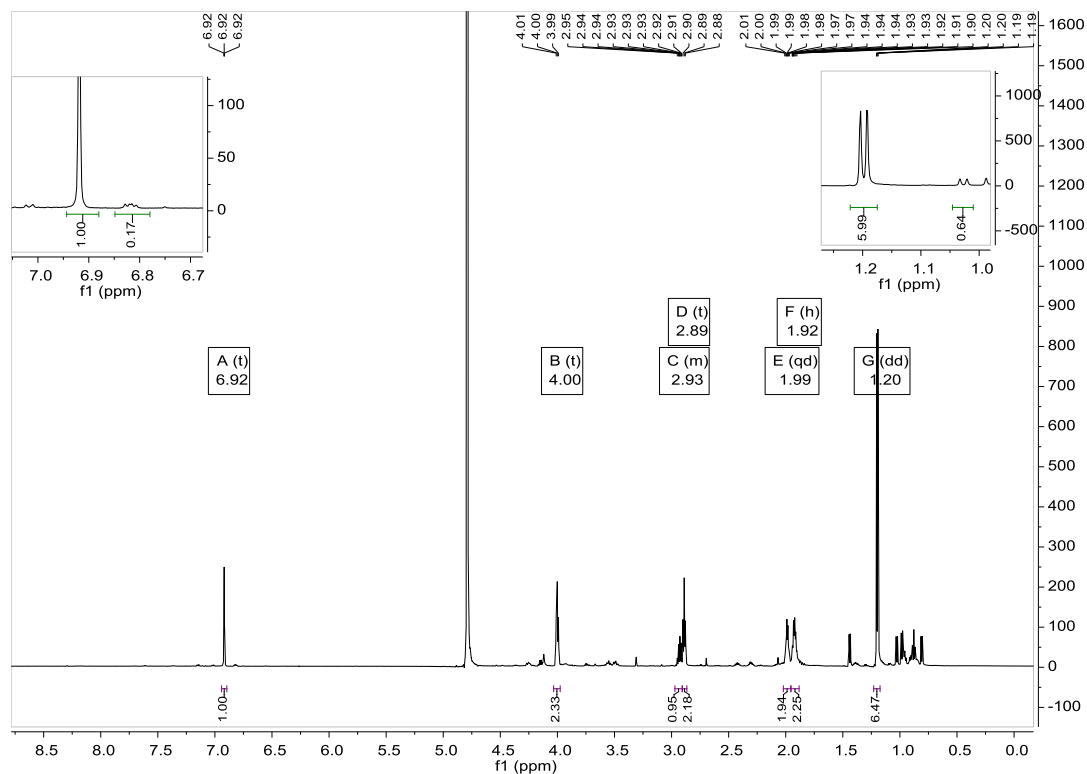
Appendix 30. ^{1}H NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in D$_2$O (600 MHz). By integration, ImC appears to be approx. 85% pure.
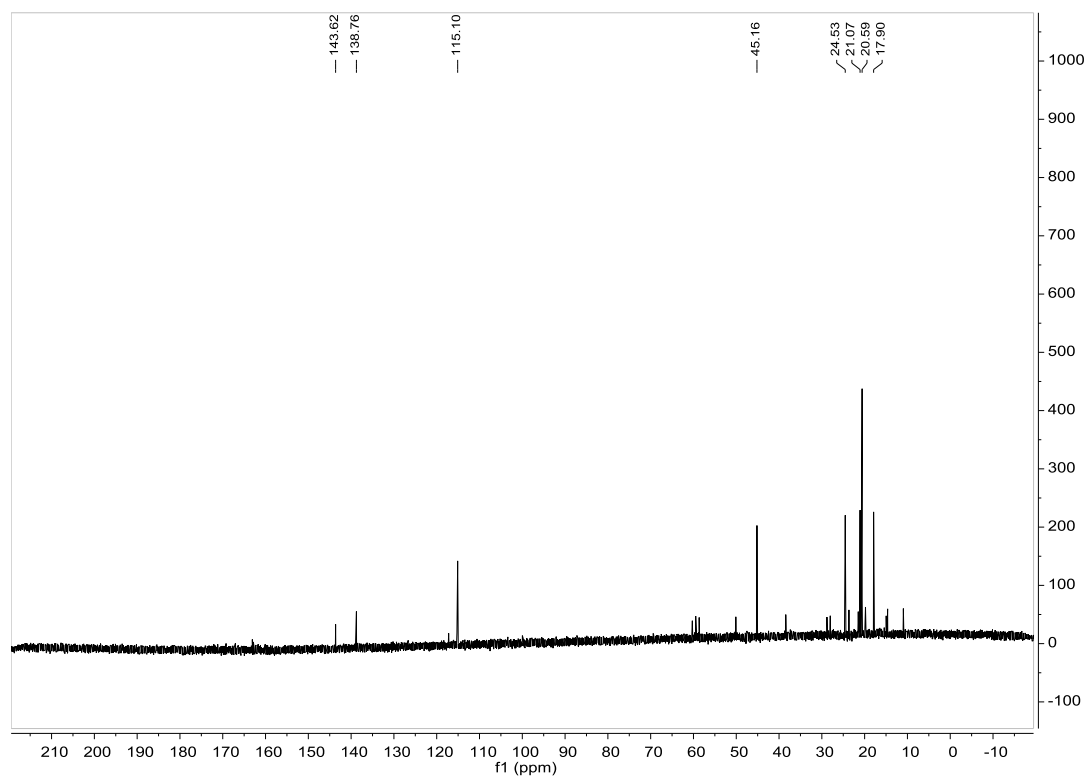
Appendix 31. ^{13}C NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in D$_2$O (151 MHz).
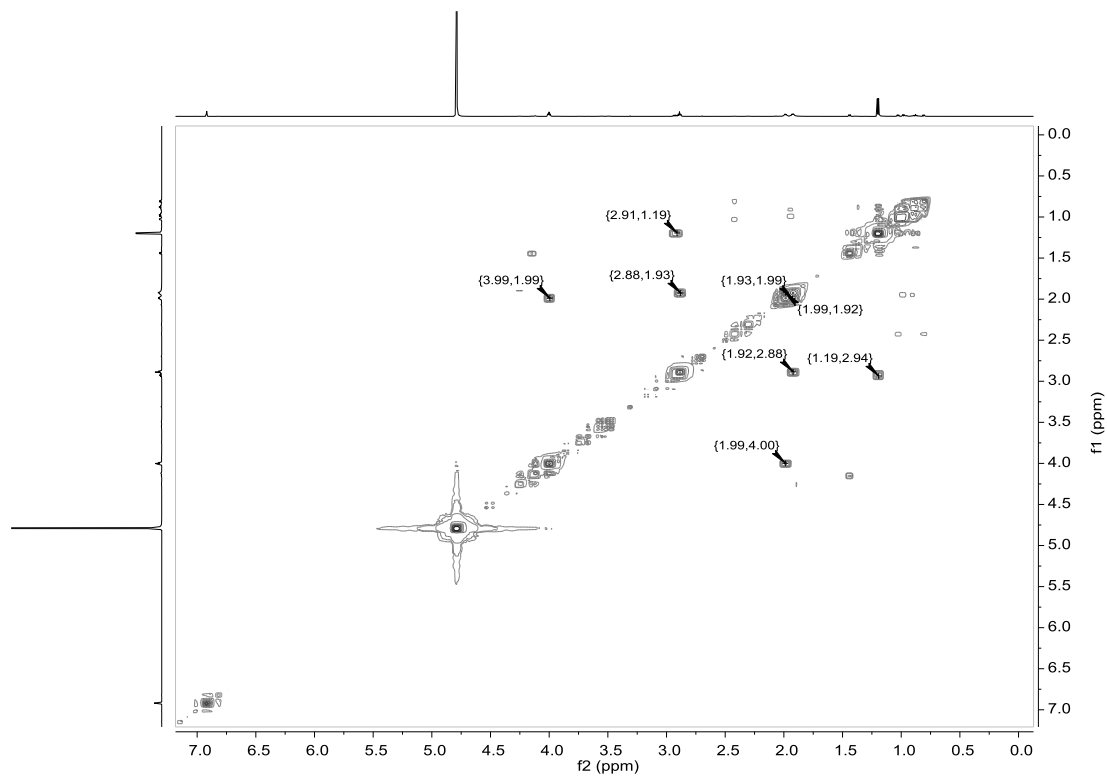
Appendix 32. (^{1}H, ^{1}H)-COSY NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in D$_2$O (600 MHz).
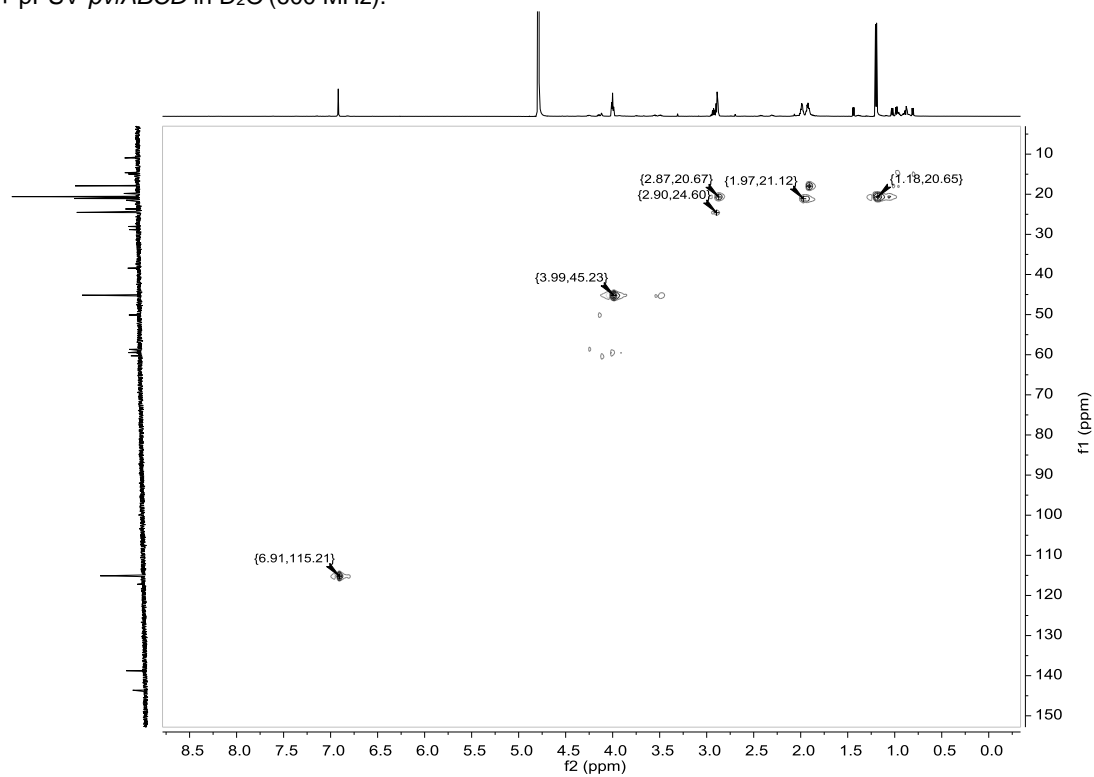
Appendix 33. (^{1}H, ^{13}C)-HSQC NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in D$_2$O (^{1}H, 600 MHz, ^{13}C 151 MHz).
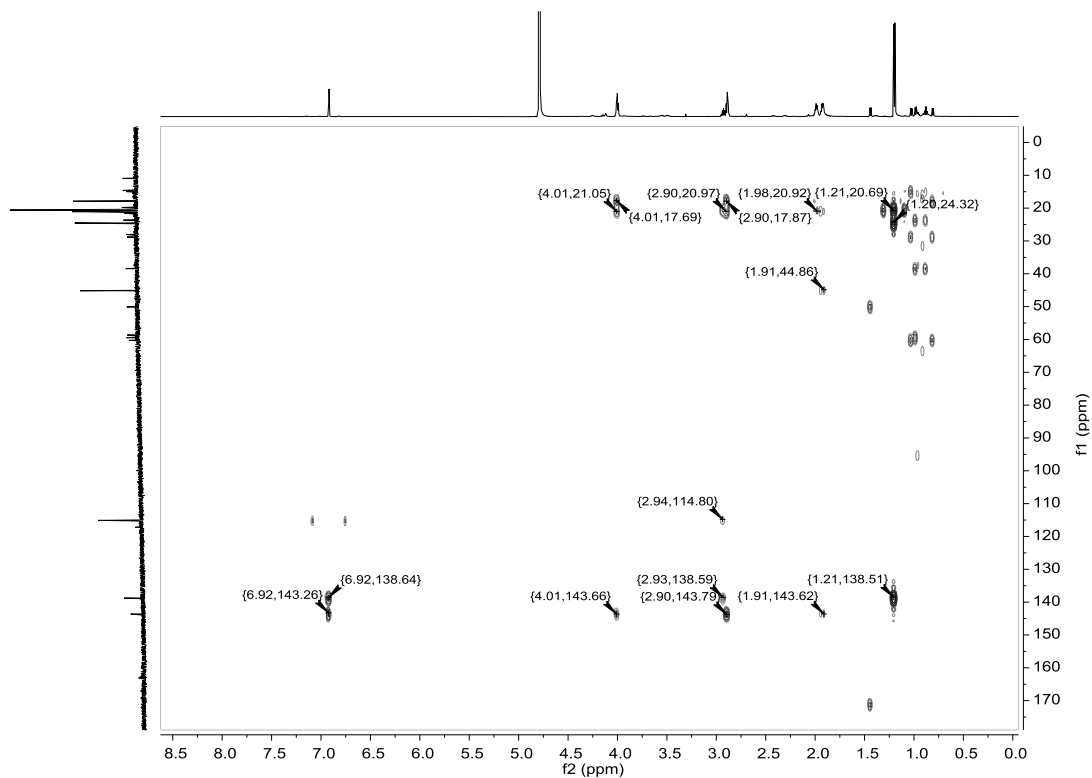
Appendix 34. (^{1}H, ^{13}C)-HMBC NMR spectrum of isolated ImC (**4**), HPLC Round 2 Fraction 11 from *P. entomophila* *ΔpvfC* + pPSV-*pvfABCD* in D$_2$O (^{1}H, 600 MHz, ^{13}C 151 MHz).
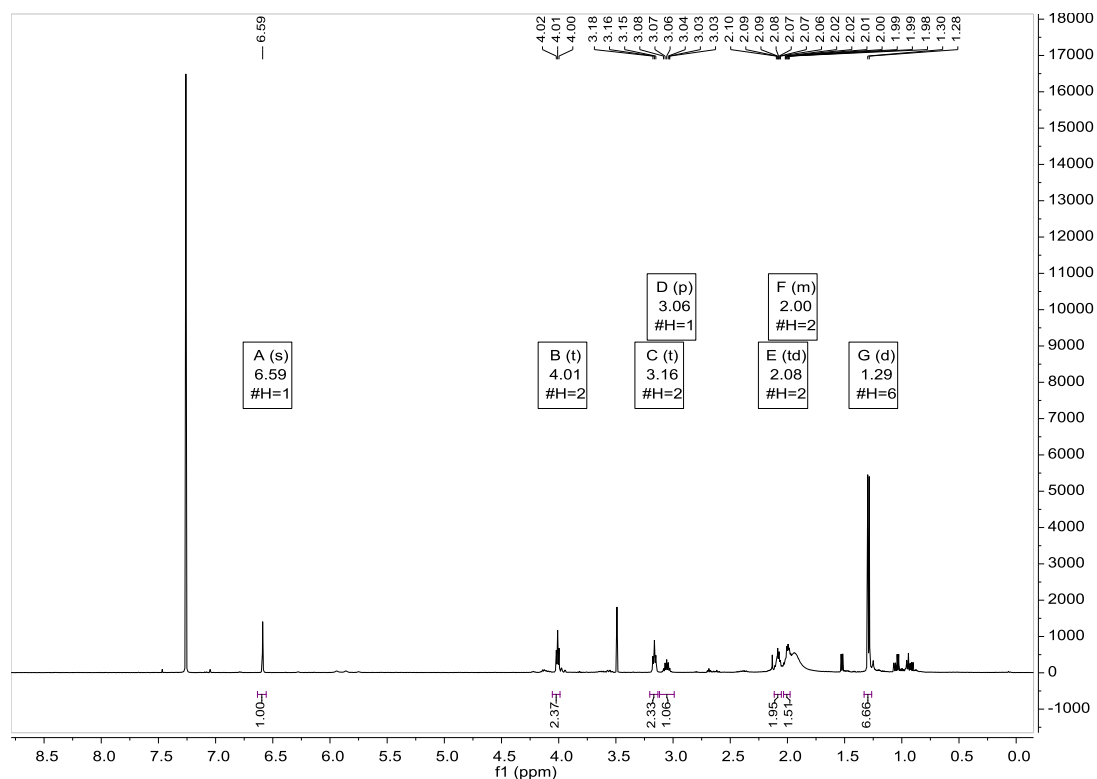
Appendix 35. ^{1}H NMR spectrum of isolated ImC (**4**), HPLC Round 3 Fractions 19−22 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in CDCl$_3$ (600 MHz).
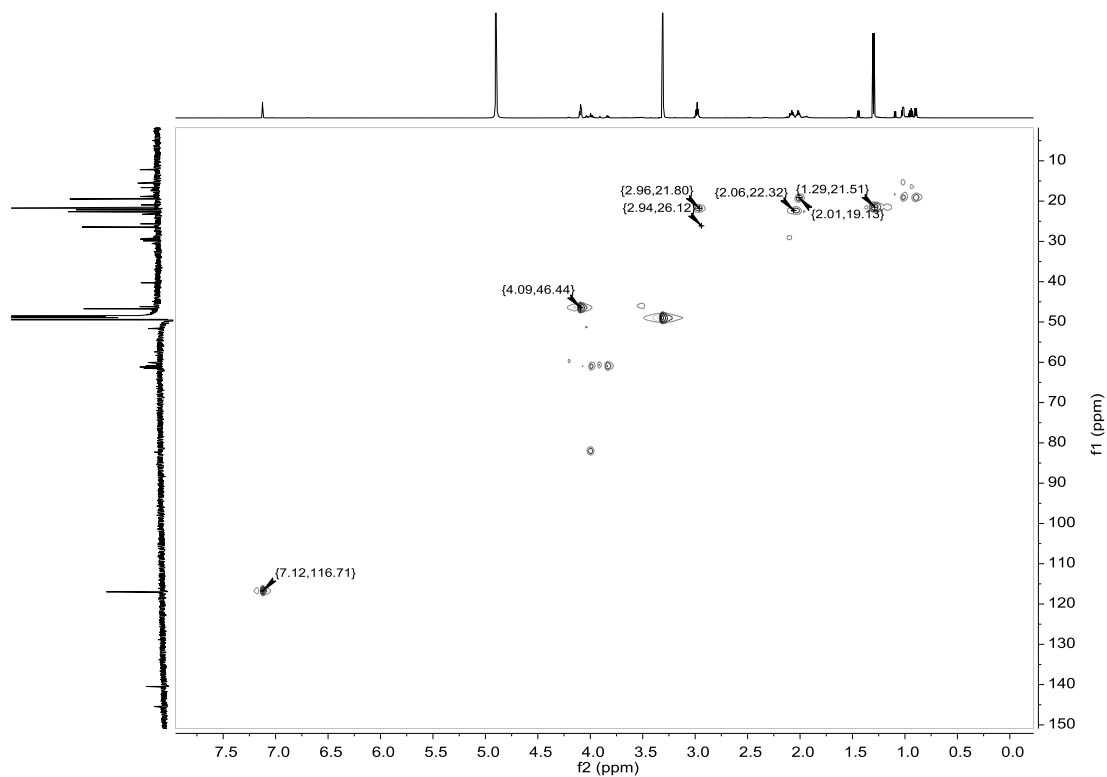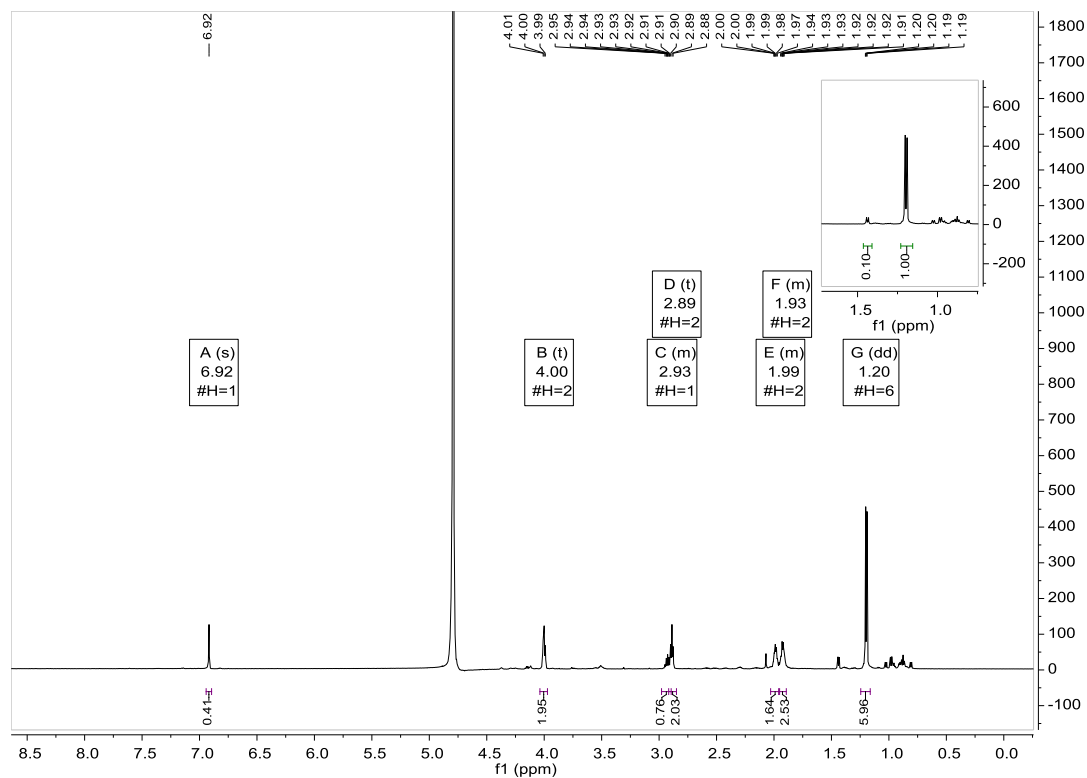
Appendix 36. ¹H NMR spectrum of isolated ImC (**4**), HPLC Round 3 Fractions 19−22 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in D₂O (600 MHz). By integration, ImC appears to be approx. 90% pure.
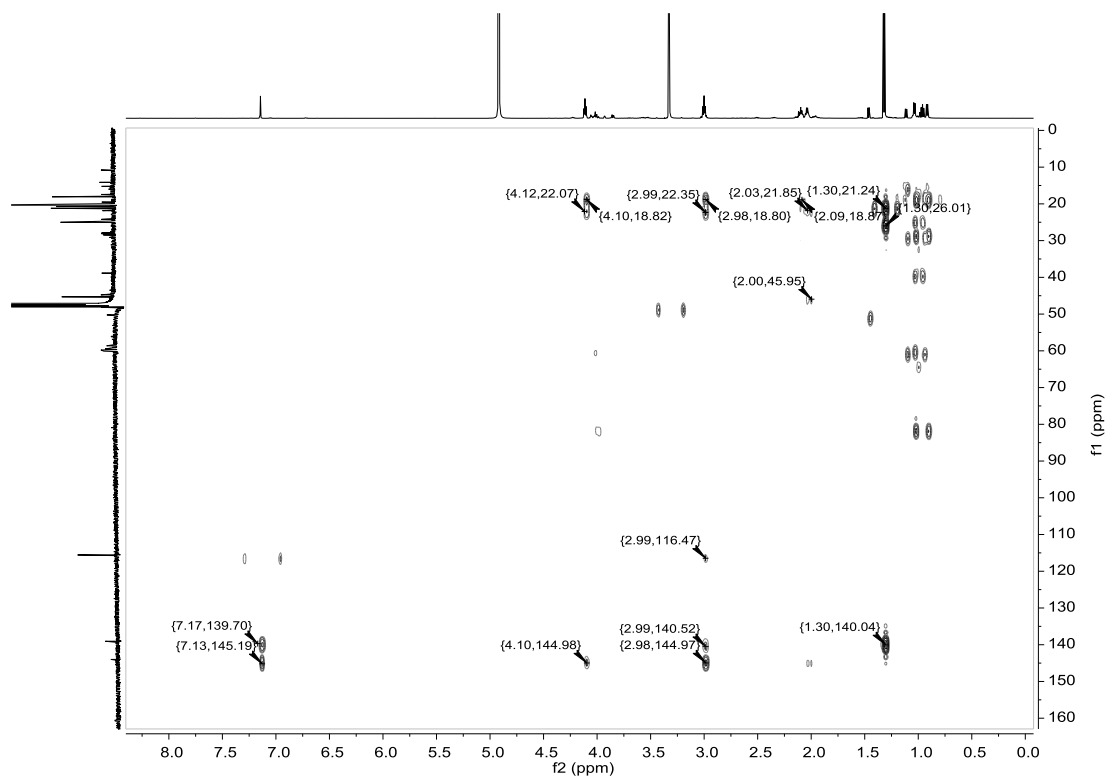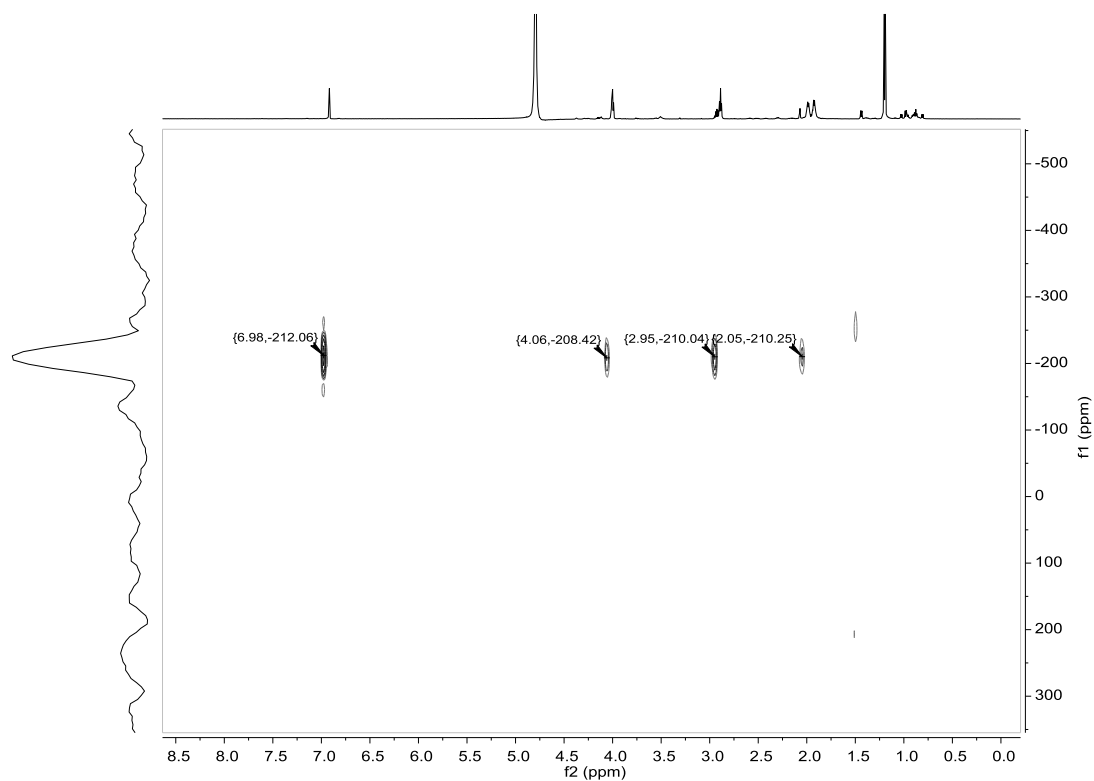

Appendix 37. (¹H, ¹⁵N)-HMBC NMR spectrum of isolated ImC (**4**), HPLC Round 3 Fractions 19−22 from *P. entomophila ΔpvfC* + pPSV-*pvfABCD* in D₂O (¹H, 600 MHz, ¹⁵N, 61 MHz).
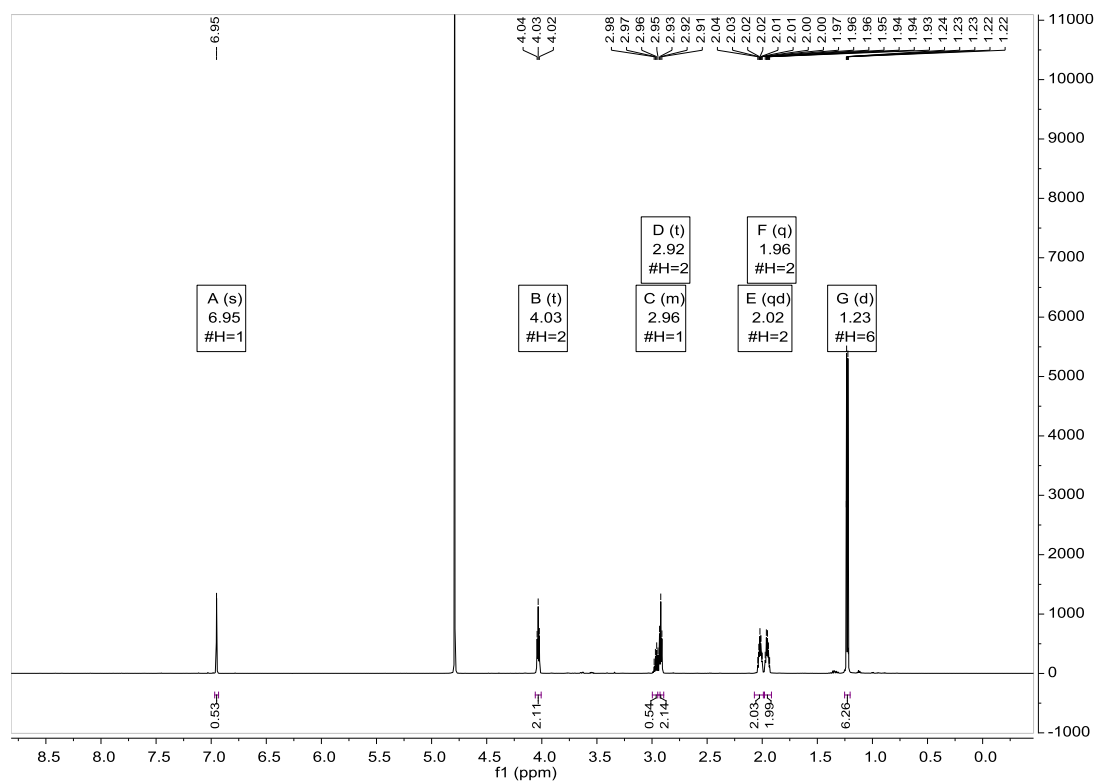
Appendix 38. ^{1}H NMR spectrum of synthetic ImC (**4**) in D$_2$O (600 MHz).
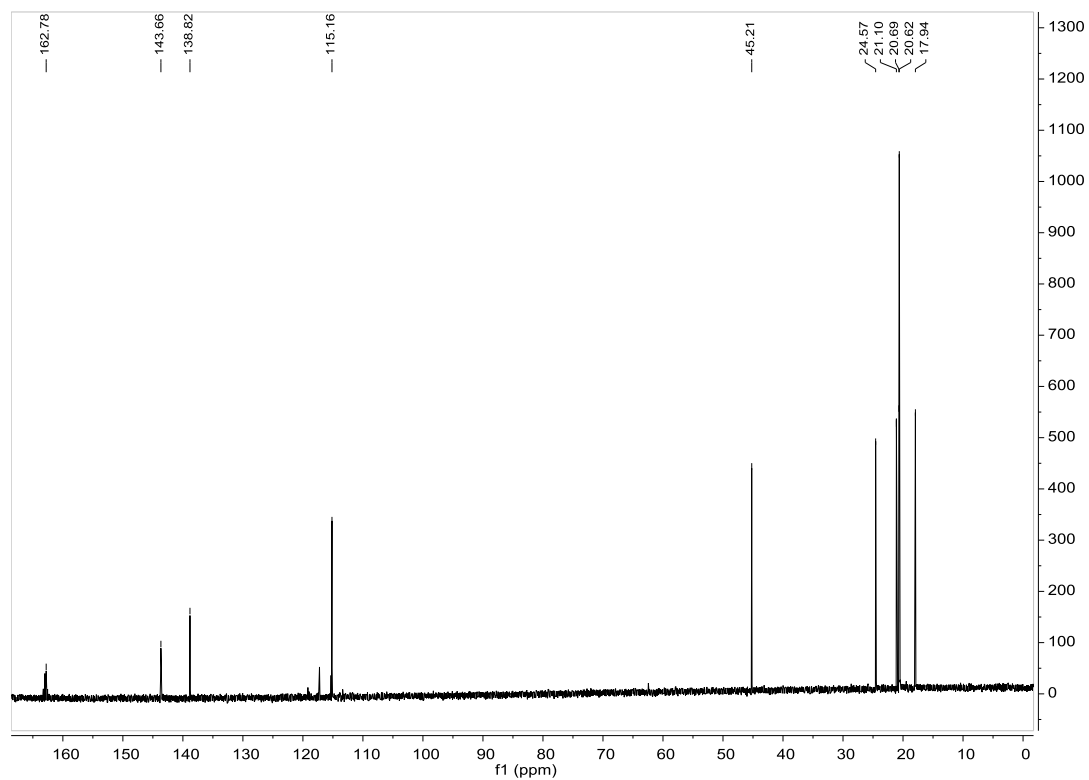
Appendix 39. ^{13}C NMR spectrum of synthetic ImC (**4**) in D$_2$O (151 MHz).
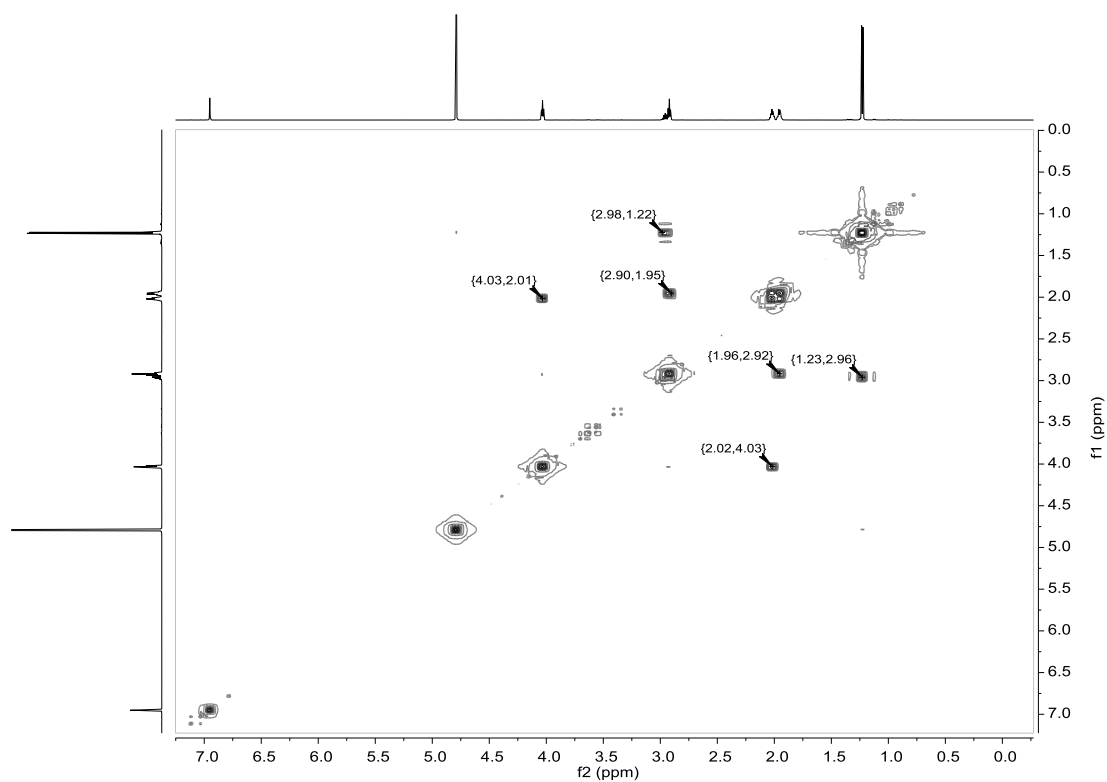
Appendix 40. (^{1}H, ^{1}H)-COSY NMR spectrum of synthetic ImC (**4**) in D$_2$O (600 MHz).
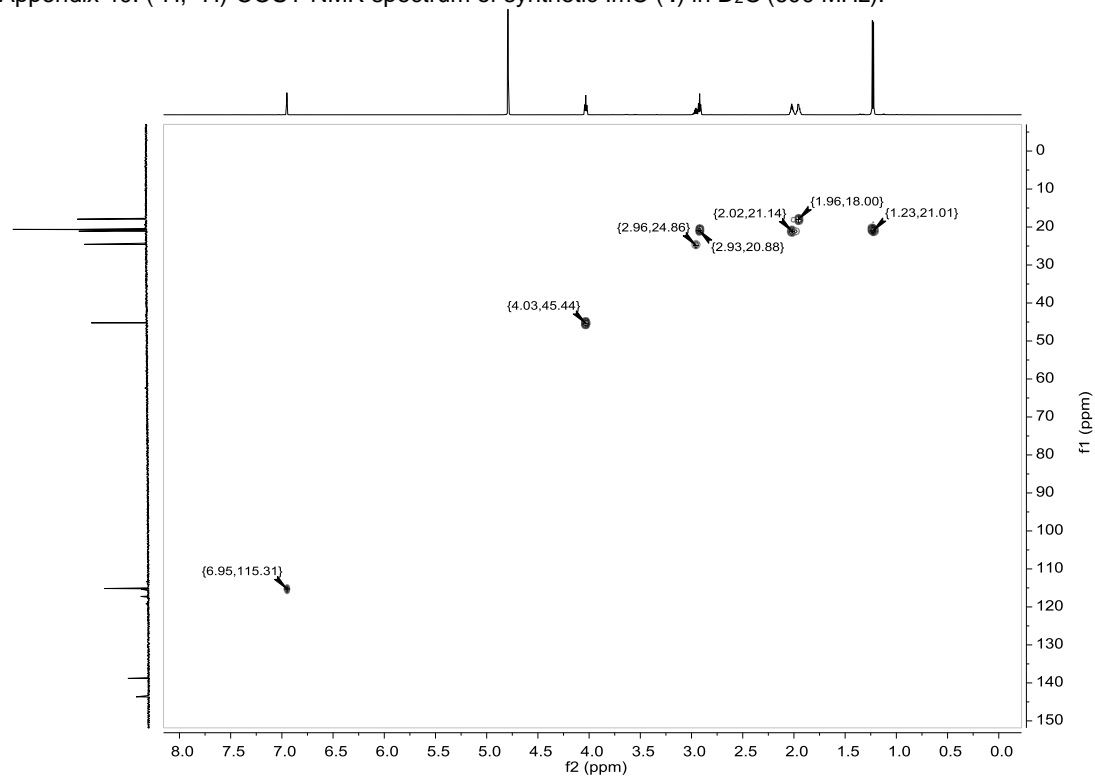
Appendix 41. (^{1}H, ^{13}C)-HSQC NMR spectrum of synthetic ImC (**4**) in D$_2$O (^{1}H, 600 MHz, ^{13}C 151 MHz).
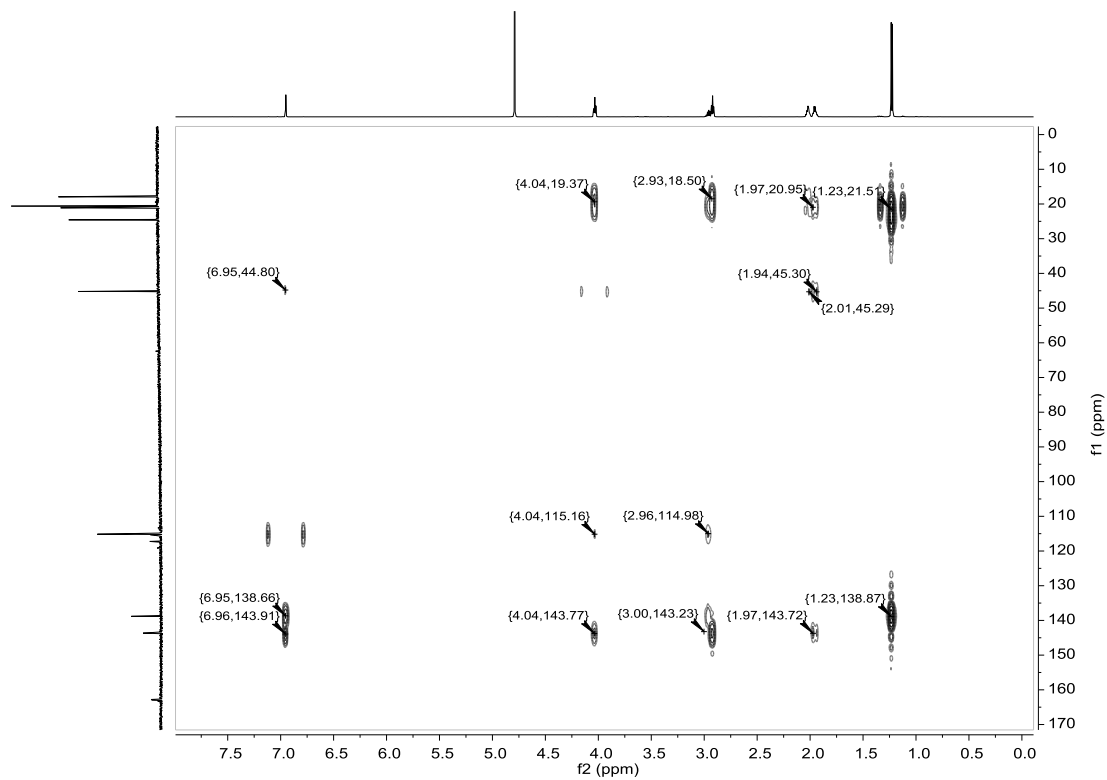
Appendix 42. (^{1}H, ^{13}C)-HMBC NMR spectrum of synthetic ImC (**4**) in D$_2$O (^{1}H, 600 MHz, ^{13}C 151 MHz).
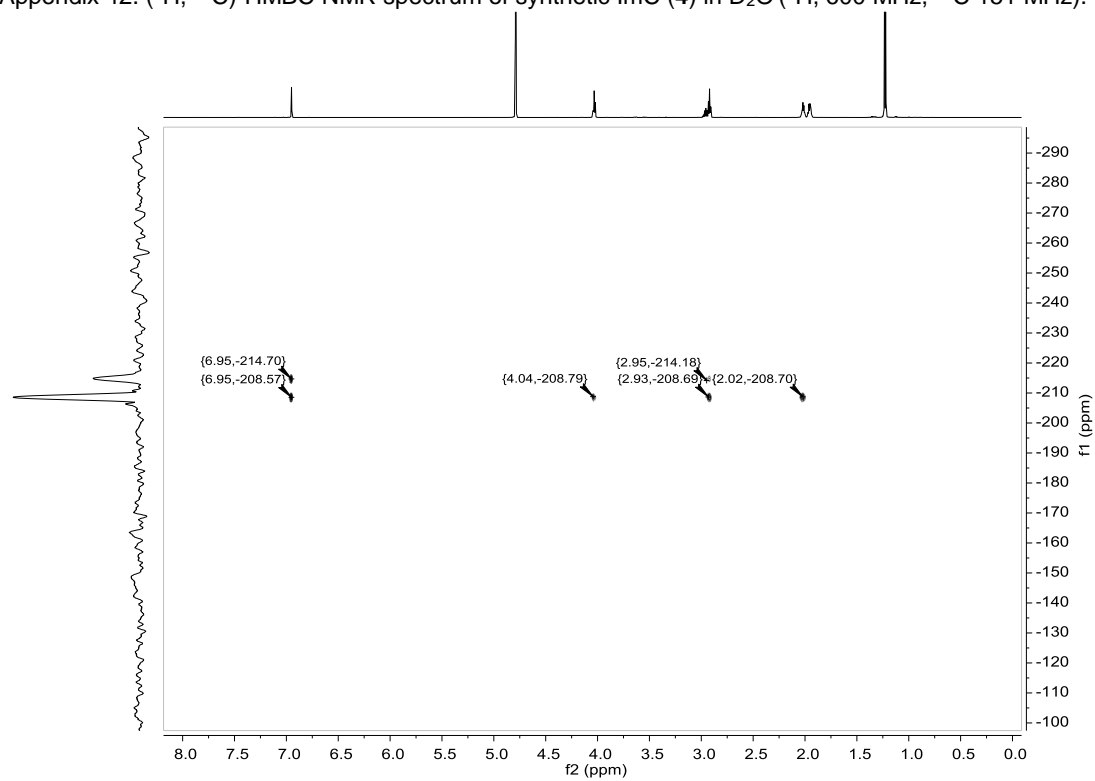
Appendix 43. (^{1}H, ^{15}N)-HMBC NMR spectrum of synthetic ImC (**4**) in D$_2$O (^{1}H, 600 MHz, ^{15}N, 61 MHz).

Appendix 44. List of primers used in these studies

Name	Sequence (5' to 3')	Description
AKp172c	GTCGAATTCCTCGAATATGGC	*pvfABC* fwd (EcoRI) start of *pvfA*
AK180b	CGATAAGCTTTCAGATGAAGCCGATGTCG	*pvfABC* rev (Hind III) end of *pvfC*
GM023	GATCGATCGAATTCATGAACGCCCATGAATACCGCTCCTT	*pvfBCD* fwd (EcoRI) start of *pvfB*
GM024	GATCGATCAAGCTT TCAGCTCAGCAGGTTCTTCACATTCCC	*pvfBCD* rev (Hind III) end of *pvfD*
AK181b	ATCGGAATTCATGAACGCCCATGAATACCG	*pvfBC* fwd (EcoRI) start of *pvfB*
AK180b	CGATAAGCTTTCAGATGAAGCCGATGTCG	*pvfBC* rev (Hind III) end of *pvfC*
GM019	GGATAACAATTTCACACAGGAAACAGCTATGACC	*pvfACD* fragment 1 fwd (EcoRI) start of *pvfA*
GM020	GCGTCTCATTCATGGGGTGACC	*pvfACD* fragment 1 rev (3' *pvfA* into 5' *pvfC*)
GM022	CCCATGAATGAGACGCCTCGACATAC	*pvfACD* fragement 2 fwd (3' of *pvfA* into 5' *pvfC*)
GM023	CAGGGTTTTCCCAGTCACGACG	*pvfACD* fragement 2 rev (HindIII) end of *pvfD*
AKp242	ATCGGAATTCATGAGACGCCTCGACATA	*pvfCD* fwd (EcoRI) start of *pvfC*
AKp175	CGATAAGCTTAAAGCCGCTCC	*pvfCD* rev (HindIII) end of *pvfD*
AKp242	ATCGGAATTCATGAGACGCCTCGACATA	*pvfC* fwd (EcoRI) start of *pvfC*
AK180b	CGATAAGCTTTCAGATGAAGCCGATGTCG	*pvfC* rev (Hind III) end of *pvfC*
AK139	TACTTCCAATCCAATGCGATGAGACGCCTC	*pvfC* fwd (pLICHis)
AK140	TTATCCACTTCCAATGCGCTATCAGATGAAGCC	*pvfC* rev (pLICHis)
AKp182	ATCGGAATTCCCGCGCACGCTTACA	L48 *pvf* promoter region fwd (EcoRI)
AKp183	CGATAAGCTTGATGATGTCCGTTGTGGG	L48 *pvf* promoter region rev (HindIII)
AKp239	CGATAAGCTTATGTGCTCATGATGATGTCCG	L48 *pvf* promoter region with first AAs, rev (HindIII)
AKp184	ATCGGAATTCCCTTCATGGGAAGTGGAAG	L48 *mnl* promoter region fwd (EcoRI)
AKp185	CGATAAGCTTCGTCCTTCCCTCGTTGC	L48 *mnl* promoter region rev (HindIII)
AKp237	CGATAAGCTTCCTTGATCGTCATCGTCCTTC	L48 *mnl* promoter region with first AAs, REV (HindIII)
AKp186	ATCGGAATTCGCGATACTCGCCTTCACC	L48 *AprA* promoter region fwd (EcoRI)
AKp187	CGATAAGCTTGCAGTTGATGTCAGACAGGA	L48 *AprA* promoter region rev (HindIII)
AKp238	CGATAAGCTTCTTTCGACATGCAGATTACTTCC	L48 *AprA* promoter region with first AAs, REV (HindIII)

AKp240	ATCGGAATTCTATCACCAAGGCCTCGCT	L48 *rspL* S12 promoter region fwd (EcoRI)
AKp241	CGATAAGCTTTAGTTGCCATCTACTAGCTCC	L48 *rspL* S12 promoter region with first AAs rev (HindIII)
AKp154	CCTTGAGCTGATCGAACTGCC	L48 *pvf* KO upstream sequencing primer
AKp155	TGATTGGGCACGATCTCTGTT	L48 *pvf* KO downstream sequencing primer
AKp170	ATCGATGCGGCCGCGCACTTTATCCACCAGAAGAAG	L48 *pvf* KO left flank fwd (NotI)
AKp206b	TCATGATGATGTCCGTTGTGG	L48 *pvf* KO left flank rev in frame
AKp207b	ATCATCATGAATGACCTGGAGCCTGATC	L48 *pvf* KO right flank in frame with overlap fwd
AKp171	GTGCATGAATTCGCTAAAGCCGCTCCTACA	L48 *pvf* KO right flank rev (EcoRI)
AKp1.1b	ATCGATCGCATATGGTGAATCAACTGGAAGGGCTGGAGCGAAAAATCGGCGAGC	HI2424 *pvfA-D* fwd (NdeI)
AKp1.3b	CGATCGATAAGCTTTTACTGCGGCTTGAACAGCGCGCGCAGGTTCGTCAT	HI2424 *pvfA-D* rev (HindIII)
AK3.1.2b	AGTGTGCTGGAATTCAGGGCACAGATCGTCGAGC	HI2424 KO Fragment 1 (left flank with pCRblunt overhang) fwd
AK6.2.2b	GCGTCAGACCCCGTAGATGTCTAGGATAAGAATTGCG	HI2424 KO Fragment 1 (left flank with *dhfrII* overhang) rev
AK6.2.1b	CGCAATTCTTATCCTAGACATCTACGGGGTCTGACGC	HI2424 KO Fragment 2 (*dhfrII* with left flank overhang) fwd
AK6.3.2b	GTGACGAACGCGTCGAATAGGGATCCTAAGATATCGCTTAG	HI2424 KO Fragment 2 (*dhfrII* with right flank overhang) rev
AK6.3.1b	CTAAGCGATATCTTAGGATCCCTATTCGACGCGTTCGTCAC	HI2424 KO Fragment 3 (right flank with *dhfrII* overhang) fwd
AK6.4.3b	TGGATATCTGCAGAATTCAGGATGCAGGCGTTCGTCAC	HI2424 KO Fragment 3 (right flank with pCRblunt overhang) rev
AK6.4.2b	GTGACGAACGCCTGCATCCTGAATTCTGCAGATATCCA	HI2424 KO Fragment 4 (pCRblunt with right flank overhang) fwd
AK3.1.3b	GCTCGACGATCTGTGCCCTGAATTCCAGCACACT	HI2424 KO Fragment 4 (pCRblunt with left flank overhang) rev
AKp176	TACTTCCAATCCAATGCGATGACGACCGCCGTCA	HI2424 *pvfB* fwd (pLICHis)
AKp177	TTATCCACTTCCAATGCGCTATCACAGGTAGTCGGACAAGG	HI2424 *pvfB* rev (pLICHis)
AKp178	TACTTCCAATCCAATGCGATGCTGCAACGTACTGCAA	HI2424 *pvfD* fwd (pLICHis)
AKp179	TTATCCACTTCCAATGCGCTATTACTGCGGCTTGAACAGC	HI2424 *pvfD* rev (pLICHis)
AKp190	TACTTCCAATCCAATGCGATGAACGCCGCCGACTA	Pf01 *pvfA* fwd (pLICHis)
AKp191	TTATCCACTTCCAATGCGCTATCAAACCAGGTAGTCGCG	Pf01 *pvfA* rev (pLICHis)
AKp192	TACTTCCAATCCAATGCGATGCCAACCAAAGAGCAACTTA	Pf01 *pvfB* fwd (pLICHis)
AKp193	TTATCCACTTCCAATGCGCTATCATTGGGCCACCTCC	Pf01 *pvfB* rev (pLICHis)
AKp194	TACTTCCAATCCAATGCGATGAGCAATCTGCAACCC	Pf01 *pvfD* fwd (pLICHis)
AKp195	TTATCCACTTCCAATGCGCTATTACACAAACAGCTTCAGCAC	Pf01 *pvfD* rev (pLICHis)

Appendix 45. List of strains used in these studies

Name	Strain	Plasmid
AK075	*P. entomophila* L48 wild type	None
AK076	*P. entomophila* Δ*pvfC* mutant	None
AK079	*P. entomophila* Δ*pvfD* mutant	None
AK080	*P. entomophila* Δ*pvfB* mutant	None
AK081	*P. entomophila* Δ*pvfA* mutant	None
AK123	*P. entomophila* Δ*pvfC* mutant	pPSV35-*pvfABCD*
AK084	*P. aeruginosa* PAO1	pPSV35
AK085	*P. aeruginosa* PAO1	pPSV35-*pvfABCD*
AK082	*E. coli* Top10	pPSV35
AK083	*E. coli* Top10	pPSV35-*pvfABCD*
--	*E. coli* Top10	pLICHis
KT003	*E. coli* Top10	pLICHis-*pvfC*
KT007	*E. coli* Bap1	pLICHis-*pvfC*
KT002	*E. coli* Top10	pLICHis-*pvfB*
GM103	*E. coli* BL21	pLICHis-*pvfB*
AK155	*E. coli* Top10	pPSV35-*pvfABC*
AK177	*E. coli* Top10	pPSV35-*pvfBC*
--	*E. coli* Top10	pPSV35-*pvfBCD*
GM048	*E. coli* Top10	pPSV35-*pvfACD*
GM050	*E. coli* DH5α	pPSV35-*pvfCD*
GM049	*E. coli* DH5α	pPSV35-*pvfC*
--	*E. coli* RHO3	None
GM114	*E. coli* RHO3	pPSV35-*pvfABC*
GM060	*E. coli* RHO3	pPSV35-*pvfBCD*
GM035	*E. coli* RHO3	pPSV35-*pvfACD*
GM037	*E. coli* RHO3	pPSV35-*pvfCD*
GM036	*E. coli* RHO3	pPSV35-*pvfC*
AK176	*P. aeruginosa* PAO1	pPSV35-*pvfABC*
AK179	*P. aeruginosa* PAO1	pPSV35-*pvfBC*
GM061	*P. aeruginosa* PAO1	pPSV35-*pvfBCD*
GM032	*P. aeruginosa* PAO1	pPSV35-*pvfACD*
GM033	*P. aeruginosa* PAO1	pPSV35-*pvfCD*
GM034	*P. aeruginosa* PAO1	pPSV35-*pvfC*
AK114	*E. coli* Top10	pUC-*GFP* miniTn7-kan
AK115	*E. coli* Top10	pUC-*lacZ* miniTn7-kan
AK116	*E. coli* RHO3	pTNS3 helper plasmid
AK137	*E. coli* Top10	pUC-P$_{pvf}$-*GFP*
AK137b	*E. coli* Top10	pUC-P$_{pvf}$-*GFP* (with start aa)
AK138	*E. coli* Top10	pUC-P$_{pvf}$-*LacZ*
AK138b	*E. coli* Top10	pUC-P$_{pvf}$-*LacZ* (with start aa)
AK139	*E. coli* Top10	pUC-P$_{mnr}$-*GFP*
AK139b	*E. coli* Top10	pUC-P$_{mnr}$-*GFP* (with start aa)

AK140	*E. coli* Top10	pUC-P*mnr*-*LacZ*
AK140b	*E. coli* Top10	pUC-P*mnr*-*LacZ* (with start aa)
AK141	*E. coli* Top10	pUC-P*AprA*-*GFP*
AK141b	*E. coli* Top10	pUC-P*AprA*-*GFP* (with start aa)
AK142	*E. coli* Top10	pUC-P*AprA*-*LacZ*
AK142b	*E. coli* Top10	pUC-P*AprA*-*LacZ* (with start aa)
AK143	*P. entomophila* L48::attTn7-P*pvf*-*GFP*	None
AK143b	*P. entomophila* L48::attTn7-P*pvf*-*GFP*	None
AK144	*P. entomophila* L48::attTn7-P*pvf*-*LacZ*	None
AK144b	*P. entomophila* L48::attTn7-P*pvf*-*LacZ*	None
AK145	*P. entomophila* L48::attTn7-P*mnr*-*GFP*	None
AK145b	*P. entomophila* L48::attTn7-P*mnr*-*GFP*	None
AK146	*P. entomophila* L48::attTn7-P*mnr*-*LacZ*	None
AK146b	*P. entomophila* L48::attTn7-P*mnr*-*LacZ*	None
AK147	*P. entomophila* L48::attTn7-P*AprA*-*GFP*	None
AK147b	*P. entomophila* L48::attTn7-P*AprA*-*GFP*	None
AK148	*P. entomophila* L48::attTn7-P*AprA*-*LacZ*	None
AK148b	*P. entomophila* L48::attTn7-P*AprA*-*LacZ*	None
AK149	*P. entomophila* L48 Δ*pvfC*::attTn7-P*pvf*-*GFP*	None
AK149b	*P. entomophila* L48 Δ*pvfC*::attTn7-P*pvf*-*GFP*	None
AK150	*P. entomophila* L48 Δ*pvfC*::attTn7-P*pvf*-*LacZ*	None
AK150b	*P. entomophila* L48 Δ*pvfC*::attTn7-P*pvf*-*LacZ*	None
AK151	*P. entomophila* L48 Δ*pvfC*::attTn7-P*mnr*-*GFP*	None
AK151b	*P. entomophila* L48 Δ*pvfC*::attTn7-P*mnr*-*GFP*	None
AK152	*P. entomophila* L48 Δ*pvfC*::attTn7-P*mnr*-*LacZ*	None
AK152b	*P. entomophila* L48 Δ*pvfC*::attTn7-P*mnr*-*LacZ*	None
AK153	*P. entomophila* L48 Δ*pvfC*::attTn7-P*AprA*-*GFP*	None
AK153b	*P. entomophila* L48 Δ*pvfC*::attTn7-P*AprA*-*GFP*	None
AK154	*P. entomophila* L48 Δ*pvfC*::attTn7-P*AprA*-*LacZ*	None
AK154b	*P. entomophila* L48 Δ*pvfC*::attTn7-P*AprA*-*LacZ*	None
AK157	*E. coli* Top10	pUC-P*rspL*-*GFP*
AK158	*E. coli* Top10	pUC-P*rspL*-*lacZ*
AK166	*P. entomophila* L48 ::attTn7-P*x*-*GFP*	None
AK167	*P. entomophila* L48::attTn7-P*x*-*LacZ*	None
AK168	*P. entomophila* L48::attTn7-P*rspL*-*GFP*	None
AK169	*P. entomophila* L48::attTn7-P*rspL*-*LacZ*	None
AK170	*P. entomophila* L48 Δ*pvfC*::attTn7-P*x*-*GFP*	None
AK171	*P. entomophila* L48 Δ*pvfC*::attTn7-P*x*-*LacZ*	None
AK172	*P. entomophila* L48 Δ*pvfC*::attTn7-P*rspL*-*GFP*	None
AK173	*P. entomophila* L48 Δ*pvfC*::attTn7-P*rspL*-LacZ	None
AK180	*P. entomophila* L48 *entA*::*lacZ*	
AK181	*P. entomophila* L48 *pvfC*::gent[R] *entA*::*lacZ*	
AK182	*P. entomophila* L48 PSEEN5493::*lacZ*	
AK183	*P. entomophila* L48 *pvfC*::gent[R] PSEEN5493::*lacZ*	

AK184	*P. entomophila* L48 PSEEN0973::*lacZ*	
AK185	*P. entomophila* L48 pvfC::gent[R] PSEEN0973::*lacZ*	
AK188	*P. entomophila* L48 Δ*pvfD*::attTn7-P$_{mnr}$-LacZ	
AK189	*P. entomophila* L48 Δ*pvfB*::attTn7-P$_{mnr}$-*LacZ*	
AK159	*E. coli* Top10	pEXKm-*pvf*KO
AK194	*P. entomophila* L48 Δ*pvf*	
AK001	*E. coli* S17.1	pSCRhaB2
AK003	*E. coli* K-12 (NovaBlue)	pSCRhaB3-*pvfA-E* (HI2424)
AK015	*B. cenocepacia* HI2424	
AK023	*E. coli* RH03	pSCRhaB3-*pvfA-E* (HI2424)
AK025	*B. cenocepacia* HI2424	pSCRhaB3-*pvfA-E* (HI2424)
AK037	*E. coli* Top10	pCRBlunt-*pvf*KO-*dhfrII* (HI2424)
AK039	*E. coli* RH03	pCRBlunt-*pvf*KO-*dhfrII* (HI2424)
AK040	*B. cenocepacia* HI2424 Δ*pvf*	
AK159	*E. coli* Top10	pEXKm-*pvf*KO
AK194	*P. entomophila* L48 Δ*pvf*	
AK190	*E. coli* Bap1	pPSV35
AK191	*E. coli* Bap1	pPSV-*pvfA-D* (L48)
AK192	*E. coli* Bap1	pSCRhaB
AK193	*E. coli* Bap1	pSCRha-*pvfA-E* (HI2424)
AK208	*E. coli* Bap1	pPSV35 + pSCRhaB
AK209	*E. coli* Bap1	pPSV-*pvfA-D* + pSCRhaB (L48)
AK210	*E. coli* Top10	pSCRhaB-*pvfD* (L48)
AK211	*E. coli* Bap1	pPSV35 + pSCRhaB-*pvfD* (L48)
AK212	*E. coli* Bap1	pPSV-*pvfA-D* + pSCRhaB-*pvfD* (L48)
AK056	*E. coli* Top10	pSCRhaB-Kan

REFERENCES

1. WHO Antibiotic resistance. (Feb 24),

2. O' Neill, J. In *Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations*, The Review on Antimicrobial Resistance2014.

3. Fair, R. J.; Tor, Y., Antibiotics and bacterial resistance in the 21st century. *Perspectives in Medicinal Chemistry* **2014**, 25-64.

4. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H., antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research* **2015,** *43*, W237-W243.

5. Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; Birren, B. W.; Takano, E.; Sali, A.; Linington, R. G.; Fischbach, M. A., Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **2014,** *158*, 412-421.

6. Ziemert, N.; Alanjary, M.; Weber, T., The evolution of genome mining in microbes – a review. *Natural Product Reports* **2016,** *00*, 1-18.

7. Deane, C. D.; Mitchell, D. A., Lessons learned from the transformation of natural product discovery to a genome-driven endeavor. *Journal of Industrial Microbiology and Biotechnology* **2014,** *41*, 315-331.

8. Tang, X.; Li, J.; Millán-Aguiñaga, N.; Zhang, J. J.; O'Neill, E. C.; Ugalde, J. A.; Jensen, P. R.; Mantovani, S. M.; Moore, B. S., Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. *ACS Chemical Biology* **2015,** *10*, 2841-2849.

9. Pablo Gomez-Escribano, J.; Bibb, M. J., *Streptomyces coelicolor* as an expression host for heterologous gene clusters. *Methods in Enzymology* **2012,** *517*.

10. Lieder, S.; Nikel, P. I.; de Lorenzo, V.; Takors, R., Genome reduction boosts heterologous gene expression in *Pseudomonas putida*. *Microbial Cell Factories* **2015,** *14*, 23.

11. Jahanshah, G.; Yan, Q.; Gerhardt, H.; Zoltán, P.; Pataj, Z. n.; Lä, M.; Pianet, I.; Josten, M.; Sahl, H.-G.; Silby, M. W.; Loper, J. E.; Gross, H., Discovery of the cyclic lipopeptide Gacamide A by genome mining and repair of the defective GacA regulator in *Pseudomonas fluorescens* Pf0-1. *Journal of Natural Products* **2019,** *82*, 301-308.

12. Okada, B. K.; Wu, Y.; Mao, D.; Bushin, L. B.; Seyedsayamdost, M. R., Mapping the trimethoprim-induced secondary metabolome of *Burkholderia thailandensis*. *ACS Chemical Biology* **2016,** *11*, 2124-2130.

13. Cociancich, S.; Pesic, A.; Petras, D.; Uhlmann, S.; Kretz, J.; Schubert, V.; Vieweg, L.; Duplan, S.; Marguerettaz, M.; Noëll, J.; Pieretti, I.; Hügelland, M.; Kemper, S.; Mainz, A.; Rott, P.; Royer, M.; Süssmuth, R. D., The gyrase inhibitor albicidin consists of p-aminobenzoic acids and cyanoalanine. *Nature Chemical Biology* **2015,** *11*, 195-197.

14.   Traxler, M. F.; Kolter, R., Natural products in soil microbe interactions and evolution. *Natural Product Reports* **2015,** *32*, 956-970.

15.   Stierle, A. A.; Stierle, D. B.; Decato, D.; Priestley, N. D.; Alverson, J. B.; Hoody, J.; Mcgrath, K.; Klepacki, D., The Berkeleylactones, antibiotic macrolides from fungal coculture. *Journal of Natural Products* **2017,** *80*.

16.   Donia, M. S.; Cimermancic, P.; Schulze, C. J.; Wieland Brown, L. C.; Martin, J.; Mitreva, M.; Clardy, J.; Linington, R. G.; Fischbach, M. A., A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **2014,** *158*, 1402-1414.

17.   Zipperer, A.; Konnerth, M. C.; Laux, C.; Berscheid, A.; Janek, D.; Weidenmaier, C.; Burian, M.; Schilling, N. A.; Slavetinsky, C.; Marschal, M.; Willmann, M.; Kalbacher, H.; Schittek, B.; Brötz-Oesterhelt, H.; Grond, S.; Peschel, A.; Krismer, B., Human commensals producing a novel antibiotic impair pathogen colonization. *Nature* **2016,** *535*, 511-516.

18.   Von Tesmar, A.; Hoffmann, M.; Abou Fayad, A.; Hüttel, S.; Schmitt, V.; Herrmann, J.; Müller, R., Biosynthesis of the *Klebsiella oxytoca* pathogenicity factor tilivalline: Heterologous expression, *in vitro* biosynthesis, and inhibitor development. *ACS Chemical Biology* **2018,** *13*, 812-819.

19.   Vizcaino, M. I.; Crawford, J. M., The colibactin warhead crosslinks DNA. *Nature Chemistry* **2015,** *7*, 411-417.

20.   Brotherton, C. A.; Wilson, M.; Byrd, G.; Balskus, E. P., Isolation of a metabolite from the *pks* island provides insights into colibactin biosynthesis and activity. *Organic Letters* **2015,** *17*, 1545-1548.

21.   Bian, X.; Plaza, A.; Zhang, Y.; Müller, R., Two more pieces of the colibactin genotoxin puzzle from *Escherichia coli* show incorporation of an unusual 1-aminocyclopropanecarboxylic acid moiety. *Chemical Science* **2015,** *6*, 3154-3160.

22.   Zheng, X. Y.; Spivey, N. W.; Zeng, W.; Liu, P. P.; Fu, Z. Q.; Klessig, D. F.; He, S. Y.; Dong, X., Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. *Cell Host and Microbe* **2012,** *11*, 587-596.

23.   Yang, M.; de Kock, M. J. D.; de Bruijn, I.; van Beek, T. A.; Raaijmakers, J. M.; de Waard, P., Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species. *Molecular Microbiology* **2006,** *63*, 417-428.

24.   Watrous, J.; Yang, J. Y.; Raaijmakers, J. M.; Dorrestein, P. C.; Roach, P.; Heath, B. S.; Bandeira, N.; Moore, B. S.; van der Voort, M.; Alexandrov, T.; Laskin, J.; Gross, H.; Pogliano, K.; Kersten, R. D., Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences* **2012,** *109*, E1743-E1752.

25.   Mendes, R.; Kruijt, M.; Bruijn, I. d.; Dekkers, E.; Voort, M. v. d.; Schneider, J. H. M.; Piceno, Y. M.; DeSantis, T. Z.; Andersen, G. L.; Bakker, P. A. H. M.; Raaijmakers, J. M., Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* **2011,** *332*, 1097-1100.

26.	Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. a.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K., A new antibiotic kills pathogens without detectable resistance. *Nature* **2015,** *517*, 455-459.

27.	Brameyer, S.; Kresovic, D.; Bode, H. B.; Heermann, R., Dialkylresorcinols as bacterial signaling molecules. *Proceedings of the National Academy of Sciences* **2014,** *112*, 572-577.

28.	Heermann, R.; Schubert, K.; Brameyer, S.; Manske, C.; Kopp, Y.; Brachmann, A. O.; Bode, H. B.; Hitkova, I.; Kresovic, D., Pyrones as bacterial signaling molecules. *Nature Chemical Biology* **2013,** *9*, 573-578.

29.	O'Rourke, S.; Corre, C.; Chater, K. F.; Challis, G. L.; Song, L., 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. *Proceedings of the National Academy of Sciences* **2008,** *105*, 17510-17515.

30.	Miyamoto, K. T.; Nishitomi, K.; Iguchi, H.; Nihira, T.; Herawati, E.; Takamatsu, S.; Uchida, M.; Nagamitsu, T.; Omura, S.; Ikeda, H.; Kitani, S., Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. *Proceedings of the National Academy of Sciences* **2011,** *108*, 16410-16415.

31.	Papenfort, K.; Bassler, B. L., Quorum sensing signal-response systems in Gram-negative bacteria. *Nature Reviews Microbiology* **2016,** *14*, 576-588.

32.	Bender, C. L.; Alarcón-Chaidez, F.; Gross, D. C., *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiology and molecular biology reviews* **1999,** *63*, 266-292.

33.	Cornelis, P., Iron uptake and metabolism in pseudomonads. *Applied Microbiology and Biotechnology* **2010,** *86*, 1637-1645.

34.	Gross, H.; Loper, J. E., Genomics of secondary metabolite production by *Pseudomonas* spp. *Natural Product Reports* **2009,** *26*, 1408-1446.

35.	Carrión, V. J.; van der Voort, M.; Arrebola, E.; Gutiérrez-Barranquero, J. A.; de Vicente, A.; Raaijmakers, J. M.; Cazorla, F. M., Mangotoxin production of *Pseudomonas syringae* pv. *syringae* is regulated by MgoA. *BMC microbiology* **2014,** *14*.

36.	Arrebola, E.; Cazorla, F. M.; Romero, D.; Pérez-García, A.; de Vicente, A., A nonribosomal peptide synthetase gene (*mgoA*) of *Pseudomonas syringae* pv. *syringae* is involved in mangotoxin biosynthesis and is required for full virulence. *Molecular plant-microbe interactions* **2007,** *20*, 500-509.

37.	Arrebola, E.; Carrión, V. J.; Cazorla, F. M.; Pérez-García, A.; Murillo, J.; de Vicente, A., Characterisation of the *mgo* operon in *Pseudomonas syringae* pv. *syringae* UMAF0158 that is required for mangotoxin production. *BMC Microbiology* **2012,** *12*.

38.	Vallet-Gely, I.; Opota, O.; Boniface, A.; Novikov, A.; Lemaitre, B., A secondary metabolite acting as a signalling molecule controls *Pseudomonas entomophila* virulence. *Cellular Microbiology* **2010,** *12*, 1666-1679.

39.   Kremmydas, G. F.; Tampakaki, A. P.; Georgakopoulos, D. G., Characterization of the biocontrol activity of *Pseudomonas fluorescens* strain X reveals novel genes regulated by glucose. *PLoS ONE* **2013,** *8*, e61808.

40.   Opota, O.; Vallet-Gély, I.; Vincentelli, R.; Kellenberger, C.; Iacovache, I.; Gonzalez, M. R.; Roussel, A.; van der Goot, F. G.; Lemaitre, B., Monalysin, a novel ß-pore-forming toxin from the *Drosophila* pathogen *Pseudomonas entomophila*, contributes to host intestinal damage and lethality. *PLoS Pathogens* **2011,** *7*, e1002259.

41.   Chiarini, L.; Bevivino, A.; Dalmastri, C.; Tabacchioni, S.; Visca, P., *Burkholderia cepacia* complex species: health hazards and biotechnological potential. *Trends in Microbiology* **2006,** *14*, 277-286.

42.   Mahenthiralingam, E.; Baldwin, A.; Dowson, C. G., *Burkholderia cepacia* complex bacteria: Opportunistic pathogens with important natural biology. *Journal of Applied Microbiology* **2008,** *104*, 1539-1551.

43.   Holden, M. T. G.; Seth-Smith, H. M. B.; Crossman, L. C.; Sebaihia, M.; Bentley, S. D.; Cerdeno-Tarraga, A. M.; Thomson, N. R.; Bason, N.; Quail, M. A.; Sharp, S.; Cherevach, I.; Churcher, C.; Goodhead, I.; Hauser, H.; Holroyd, N.; Mungall, K.; Scott, P.; Walker, D.; White, B.; Rose, H.; Iversen, P.; Mil-Homens, D.; Rocha, E. P. C.; Fialho, A. M.; Baldwin, A.; Dowson, C.; Barrell, B. G.; Govan, J. R.; Vandamme, P.; Hart, C. A.; Mahenthiralingam, E.; Parkhill, J., The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. *Journal of Bacteriology* **2009,** *91*, 261-277.

44.   Xin, X.-F.; Kvitko, B.; He, S. Y., *Pseudomonas syringae*: what it takes to be a pathogen. *Nature Reviews Microbiology* **2018,** *16*, 316-328.

45.   Felnagle, E. A.; Jackson, E. E.; Chan, Y. A.; Podevels, A. M.; Berti, A. D.; Mcmahon, M. D.; Thomas, M. G., Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Molecular Pharmaceutics* **2008,** *5*, 191-211.

46.   Cezard, C.; Farvacques, N.; Sonnet, P., Chemistry and biology of pyoverdines, *Pseudomonas* primary siderophores. *Current Medicinal Chemistry* **2015,** *22*, 165-186.

47.   Grünewald, J.; Sieber, S. A.; Mahlert, C.; Linne, U.; Marahiel, M. A., Synthesis and derivatization of daptomycin: A chemoenzymatic route to acidic lipopeptide antibiotics. *Journal of the American Chemical Society* **2004,** *126*, 17025-17031.

48.   Choi, Y. S.; Zhang, H.; Brunzelle, J. S.; Nair, S. K.; Zhao, H., *In vitro* reconstitution and crystal structure of *p*-aminobenzoate *N*-oxygenase (AurF) involved in aureothin biosynthesis. *Proceedings of the National Academy of Sciences* **2008,** *105*, 6858-6863.

49.   Winkler, R.; Hertweck, C., Biosynthesis of nitro compounds. *ChemBioChem* **2007,** *8*, 973-977.

50.   Li, N.; Korboukh, V. K.; Krebs, C.; Bollinger, J. M., Four-electron oxidation of *p*-hydroxylaminobenzoate to *p*-nitrobenzoate by a peroxodiferric complex in AurF from *Streptomyces thioluteus*. *Proceedings of the National Academy of Sciences* **2010,** *107*, 15722-15727.

51.   Makris, T. M.; Vu, V. V.; Meier, K. K.; Komor, A. J.; Rivard, B. S.; Münck, E.; Que, L.; Lipscomb, J. D., An unusual peroxo intermediate of the arylamine oxygenase of the chloramphenicol biosynthetic pathway. *Journal of the American Chemical Society* **2015,** *137*, 1608-1617.

52.   Jenul, C.; Sieber, S.; Daeppen, C.; Mathew, A.; Lardi, M.; Pessi, G.; Hoepfner, D.; Neuburger, M.; Linden, A.; Gademann, K.; Eberl, L., Biosynthesis of fragin is controlled by a novel quorum sensing signal. *Nature Communications* **2018,** *9*, 1297.

53.   Dieppois, G.; Opota, O.; Lalucat, J.; Lemaitre, B., Chapter 2 *Pseudomonas entomophila:* A versatile bacterium with entomopathogenic properies. *Pseudomonas: Volume 7: New Aspects of Pseudomonas Biology* **2015**, 1-316.

54.   Kamala-Kannan S, L. K., Park SM, Chae JC, Yun BS, Lee YH, Park YJ, Oh BT, Characterization of ACC deaminase gene in *Pseudomonas entomophila* strain PS-PJH isolated from the rhizosphere soil. *Journal Basic Microbiology* **2010,** *50*, 200-205.

55.   Vodovar, N.; Vallenet, D.; Cruveiller, S.; Rouy, Z.; Barbe, V.; Acosta, C.; Cattolico, L.; Jubin, C.; Lajus, A.; Segurens, B.; Benoıˆt Vacherie, B. t.; Wincker, P.; Weissenbach, J.; Lemaitre, B.; Médigue, C.; Boccard, F., Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium *Pseudomonas entomophila*. *Nature Biotechnology* **2006,** *24*, 3.

56.   Vallet-Gely, I.; Novikov, A.; Augusto, L.; Liehl, P.; Bolbach, G.; Péchy-Tarr, M.; Cosson, P.; Keel, C.; Caroff, M.; Lemaitre, B., Association of hemolytic activity of *Pseudomonas entomophila*, a versatile soil bacterium, with cyclic lipopeptide production. *Applied and Environmental Microbiology* **2010,** *76*, 910-921.

57.   Schöner, T. A.; Gassel, S.; Osawa, A.; Tobias, N. J.; Okuno, Y.; Sakakibara, Y.; Shindo, K.; Sandmann, G.; Bode, H. B., Aryl Polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. *ChemBioChem* **2016,** *17*, 247-253.

58.   Matthijs, S.; Laus, G.; Meyer, J. M.; Abbaspour-Tehrani, K.; Schafer, M.; Budzikiewicz, H.; Cornelis, P., Siderophore-mediated iron acquisition in the entomopathogenic bacterium *Pseudomonas entomophila* L48 and its close relative *Pseudomonas putida* KT2440. *Biometals* **2009,** *22*, 951-964.

59.   Wuest, W. M.; Sattely, E. S.; Walsh, C. T., Three siderophores from one bacterial enzymatic assembly line. *Journal of the American Chemical Society* **2009,** *131*, 5056-5057.

60.   Liehl, P.; Blight, M.; Vodovar, N.; Boccard, F.; Lemaitre, B., Prevalence of local immune response against oral infection in a *Drosophila/Pseudomonas* infection model. *PLoS Pathogens* **2006,** *2*, 0551-0561.

61.   Sarris, P. F.; Scoulica, E. V., *Pseudomonas entomophila* and *Pseudomonas mendocina*: Potential models for studying the bacterial type VI secretion system. *Infection, Genetics and Evolution* **2011,** *11*, 1352-1360.

62.   Chen, W.-J.; Kuo, T.-Y.; Hsieh, F.-C.; Chen, P.-Y.; Wang, C.-S.; Shih, Y.-L.; Lai, Y.-M.; Liu, J.-R.; Yang, Y.-L.; Shih, M.-C., Involvement of type VI secretion system in secretion of iron chelator pyoverdine in *Pseudomonas taiwanensis. Scientific reports* **2016,** *6*, 32950.

63.   Jani, A. J.; Cotter, P. A., Type VI secretion: not just for pathogenesis anymore. *Cell Host & Microbe* **2010,** *8*, 2-6.

64.   Basler, M.; Ho, Brian T.; Mekalanos, John J., Tit-for-Tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **2013,** *152*, 884-894.

65.   Lindeberg, M.; Myers, C. R.; Collmer, A.; Schneider, D. J., Roadmap to new virulence determinants in *Pseudomonas syringae*: insights from comparative genomics and genome organization. *Molecular plant-microbe interactions* **2008,** *21*, 685-700.

66.   Liu, X.; Cheng, Y. Q., Genome-guided discovery of diverse natural products from *Burkholderia sp. Journal of Industrial Microbiology and Biotechnology* **2014,** *41*, 275-284.

67.   Kitani, S.; Miyamoto, K. T.; Takamatsu, S.; Herawati, E.; Iguchi, H.; Nishitomi, K.; Uchida, M.; Nagamitsu, T.; Omura, S.; Ikeda, H.; Nihira, T., Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis. Proceedings of the National Academy of Sciences* **2011,** *108*, 16410-16405.

68.   Dickey, S. W.; Cheung, G. Y. C.; Otto, M., Different drugs for bad bugs: antivirulence strategies in the age of antibiotic resistance. *Nature Reviews Drug Discovery* **2017,** *16*, 457-471.

69.   Franke, J.; Ishida, K.; Hertweck, C., Genomics-driven discovery of burkholderic acid, a noncanonical, cryptic polyketide from human pathogenic *Burkholderia* species. *Angewandte Chemie* **2012,** *51*, 11611-11615.

70.   O'Neill, E. M.; Mucyn, T. S.; Patteson, J. B.; Finkel, O. M.; Chung, E.-H.; Baccile, J. A.; Massolo, E.; Schroeder, F. C.; Dangl, J. L.; Li, B., Phevamine A, a small molecule that suppresses plant immune responses. *Proceedings of the National Academy of Sciences* **2018**, 201803779.

71.   Choi, K. H.; Kumar, A.; Schweizer, H. P., A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: Application for DNA fragment transfer between chromosomes and plasmid transformation. *Journal of Microbiological Methods* **2006,** *64*, 391-397.

72.   López, C. M.; Rholl, D. A.; Trunck, L. A.; Schweizer, H. P., Versatile dual-technology system for markerless allele replacement in *Burkholderia pseudomallei. Applied and Environmental Microbiology* **2009,** *75*, 6496-6503.

73.   Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry* **2006,** *78*, 779-787.

74.   Macnevin, C. J.; Moore, R. L.; Liotta, D. C., Stereoselective synthesis of quaternary center bearing azetines and their β-amino acid derivatives. *Journal of Organic Chemistry* **2008,** *73*, 1264-1269.

75.   Sima, M.; Naama, C. R.; Amnon, A., Optically Active *N-* and *C*-terminal building blocks for the synthesis of peptidyl olefin peptidomimetics. *European Journal of Organic Chemistry* **2010,** *2010*, 4671-4686.

76.   Kiyooka, S.; Nakano, M.; Shiota, F.; Fujiyama, R., Dramatic changes in diastereoselectivity with the quantity of titanium tetrachloride used in Lewis acid mediated reactions of allylsilane with α-amino aldehydes. *The Journal of Organic Chemistry* **1989,** *54*, 5409-5411.

77.   Badrinarayanan, S.; Sperry, J., Pyrazine alkaloids via dimerization of amino acid-derived α-amino aldehydes: biomimetic synthesis of 2,5-diisopropylpyrazine, 2,5-bis(3-indolylmethyl)pyrazine and actinopolymorphol C. *Organic & Biomolecular Chemistry* **2012,** *10*, 2126.

78.   Das, S. K.; Frey, J., Regioselective double Boekelheide reaction: first synthesis of 3,6-dialkylpyrazine-2,5-dicarboxaldehydes from DL-alanine. *Tetrahedron Letters* **2012,** *53*, 3869-3872.

79.   Aslanidis, C.; de Jong, P. J., Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Research* **1990,** *18*, 6069-6074.

80.   Schwyn, B.; Neilands, J. B., Universal chemical assay for the detection determination of siderophores. *Analytical Biochemsitry* **1987,** *160*, 47-56.

81.   Pinto, D. C. G. A.; Santos, C. M. M.; Silva, A. M. S., Advanced NMR techniques for structural characterization of heterocyclic structures. *Recent Research Developments in Heterocyclic Chemistry* **2007**, 397-475.

82.   Stadeli, W., ^{15}N-NMR. Studies of aminopyridines, aminopyrimidines and of some diazine *N*-oxides. *Helvetica Chimica Acta* **1980,** *63*, 504-22.

83.   Matsuo, M.; Matsumoto, S.; Kurihara, T.; Akita, Y.; Watanabe, T.; Ohta, A., ^{13}C NMR Spectra of alkyl- and phenylpyrazines and their *N*-Oxides. *Organic Magnetic Resonance* **1980,** *13*, 172-179.

84.   Klein, B.; Berkowitz, J., Pyrazines. I. Pyrazine-*N*-oxides. Preparation and spectral characteristics. *Journal of the American Chemical Society* **1959,** *81*, 5160-5166.

85.   Ohta, A.; Ohta, M., Convenient *N*-oxidation of pyrazines. *Synthesis* **1985,** *2*, 216-217.

86.   Lamchen, M.; Mittag, T. W., Nitrones Part IV. Synthesis and properties of a monocyclic α-dinitrone. *Journal of the Chemical Society C: Organic* **1966**, 2300-2303.

87.   Mazhukin, D. G.; Tikhonov, A. Y.; Volodarsky, L. B.; Konovalova, E. P., Interaction of 1,2-bishydroxylamines with 1,2-dicarbonyl compounds - Preparation and Pproperties of 2,3-dihydropyrazine-1,4-dioxides. *Khimiya Geterosiklichskikh Soedinenii* **1993,** *25*, 514-522.

88.   Dultseva, G. G.; Skubnevskaya, G. I.; Tikhonov, A. Y.; Mazhukin, D. G.; Volodarsky, L. B., Derivatives of dihydropyrazine-1,4-dioxide, 3-imidazoline 3-oxide, and r-phenyl nitrones with functional groups as new spin traps in solution and in the gas phase. *Journal of Physical Chemistry* **1996,** *100*, 17523-15727.

89. Ehmann, D. E.; Shaw-Reid, C. A.; Losey, H. C.; Walsh, C. T., The EntF and EntE adenylation domains of *Escherichia coli* enterobactin synthetase: Sequestration and selectivity in acyl-AMP transfers to thiolation domain cosubstrates. *Proceedings of the National Academy of Sciences* **2000,** *97*, 2509-2514.

90. Bennett, B. D.; Kimball, E. H.; Gao, M.; Osterhout, R.; Van Dien, S. J.; Rabinowitz, J. D., Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology* **2009,** *5*, 593-599.

91. Platter, E.; Lawson, M.; Marsh, C.; Sazinsky, M. H., Characterization of a non-ribosomal peptide synthetase-associated diiron arylamine *N*-oxygenase from *Pseudomonas syringae* pv. phaseolicola. *Archives of Biochemistry and Biophysics* **2011,** *508*, 39-45.

92. Wu, Y.; Seyedsayamdost, M. R., Synergy and target promiscuity drive structural divergence in bacterial alkylquinolone biosynthesis. *Cell Chemical Biology* **2017,** *24*, 1437-1444.e3.

93. Srinivasan, J.; Kaplan, F.; Ajredini, R.; Zachariah, C.; Alborn, H. T.; Teal, P. E. A.; Malik, R. U.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C., A blend of small molecules regulates both mating and development in *Caenorhabditis elegans*. *Nature* **2008,** *454*, 1115-1118.

94. Rajini, K. S.; Aparna, P.; Sasikala, C.; Ramana, C. V., Microbial metabolism of pyrazines. *Critical reviews in microbiology* **2011,** *37*, 99-112.

95. Dickschat, J. S.; Wickel, S.; Bolten, C. J.; Nawrath, T.; Schulz, S.; Wittmann, C., Pyrazine biosynthesis in *Corynebacterium glutamicum*. *European Journal of Organic Chemistry* **2010**, 2687-2695.

96. Dickschat, J. S.; Reichenbach, H.; Wagner-Döbler, I.; Schulz, S., Novel pyrazines from the myxobacterium *Chondromyces crocatus* and marine bacteria. *European Journal of Organic Chemistry* **2005**, 4141-4153.

97. Beck, H. C.; Hansen, A. M.; Lauritsen, F. R., Novel pyrazine metabolites found in polymyxin biosynthesis by *Paenibacillus polymyxa*. *FEMS Microbiology Letters* **2003,** *220*, 67-73.

98. Bañeras, L.; Trias, R.; Godayol, A.; Cerdán, L.; Nawrath, T.; Schulz, S.; Anticó, E., Mass spectrometry identification of alkyl-substituted pyrazines produced by *Pseudomonas* spp. isolates obtained from wine corks. *Food Chemistry* **2013,** *138*, 2382-2389.

99. Silva-Junior, E. A.; Ruzzini, A. C.; Paludo, C. R.; Nascimento, F. S.; Currie, C. R.; Clardy, J.; Pupo, M. T., Pyrazines from bacteria and ants: convergent chemistry within an ecological niche. *Scientific Reports* **2018,** *8*, 2595.

100. Papenfort, K.; Silpe, J. E.; Schramma, K. R.; Cong, J.-P.; Seyedsayamdost, M. R.; Bassler, B. L., A *Vibrio cholerae* autoinducer–receptor pair that controls biofilm formation. *Nature Chemical Biology* **2017,** *13*, 551-557.

101. Szamosvári, D.; Böttcher, T., An unsaturated quinolone *N*-oxide of *Pseudomonas aeruginosa* modulates growth and virulence of *Staphylococcus aureus*. *Angewandte Chemie* **2017,** *56*, 7271-7275.

102.   Hazan, R.; Que, Y. A.; Maura, D.; Strobel, B.; Majcherczyk, P. A.; Hopper, L. R.; Wilbur, D. J.; Hreha, T. N.; Barquera, B.; Rahme, L. G., Auto poisoning of the respiratory chain by a quorum-sensing-regulated molecule favors biofilm formation and antibiotic tolerance. *Current Biology* **2016,** *26*, 195-206.

103.   Filkins, L. M.; Graber, J. A.; Olson, D. G.; Dolben, E. L.; Lynd, L. R.; Bhuju, S.; O'Toole, G. A., Coculture of *Staphylococcus aureus* with *Pseudomonas aeruginosa* drives *S. aureus* towards fermentative metabolism and reduced viability in a cystic fibrosis model. *Journal of Bacteriology* **2015,** *197*, 2252-2264.

104.   Podojil, M.; Gerber, N. N., The biosynthesis of 1,6-phenazinediol 5,10-dioxide (Iodinin) by *Brevibacterium iodinum. Biochemistry* **1967,** *6*, 2701-5.

105.   Ehmann, D. E.; Gehring, A. M.; Walsh, C. T., Lysine biosynthesis in *Saccharomyces cerevisiae*: Mechanism of α-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5. *Biochemistry* **1999,** *38*, 6171-6177.

106.   Forseth, R. R.; Amaike, S.; Schwenk, D.; Affeldt, K. J.; Hoffmeister, D.; Schroeder, F. C.; Keller, N. P., Homologous NRPS-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus. Angewandte Chemie* **2013,** *52*, 1590-1594.

107.   Engebrecht, J.; Nealson, K.; Silverman, M., Bacterial bioluminescence: Isolation and genetic analysis of functions from *Vibrio fischeri. Cell* **1983,** *32*.

108.   Cheng, F.; Ma, A.; Luo, J.; Zhuang, X.; Zhuang, G., *N*-acylhomoserine lactone-regulation of genes mediating motility and pathogenicity in *Pseudomonas syringae* pathovar *tabaci* 11528. *Microbiology Open* **2016,** *6*, e00440.

109.   Nouwens, A. S.; Beatson, S. A.; Whitchurch, C. B.; Walsh, B. J.; Schweizer, H. P.; Mattick, J. S.; Cordwell, S. J.; Cordwell scordwell, S. J., Proteome analysis of extracellular proteins regulated by the *las* and *rhl* quorum sensing systems in *Pseudomonas aeruginosa* PAO1. *Microbiology* **2003,** *149*, 1311-1322.

110.   Arevalo-Ferro, C.; Hentzer, M.; Reil, G.; Görg, A.; Kjelleberg, S.; Givskov, M.; Riedel, K.; Eberl, L., Identification of quorum-sensing regulated proteins in the opportunistic pathogen *Pseudomonas aeruginosa* by proteomics. *Environmental Microbiology* **2003,** *5*, 1350-1369.

111.   Williams, P.; Cámara, M., Quorum sensing and environmental adaptation in *Pseudomonas aeruginosa*: a tale of regulatory networks and multifunctional signal molecules. *Current Opinion in Microbiology* **2009,** *12*, 182-191.

112.   Fuqua, W. C.; Winans, S. C.; Greenberg, E. P., Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of Bacteriology* **1994,** *176*, 269-75.

113.   Malott, R. J.; O'Grady, E. P.; Toller, J.; Inhülsen, S.; Eberl, L.; Sokol, P. A., A *Burkholderia cenocepacia* orphan LuxR homolog is involved in quorum-sensing regulation. *Journal of Bacteriology* **2009,** *191*, 2447-60.

114.   Sokol, P. A.; Sajjan, U.; Visser, M. B.; Gingues, S.; Forstner, J.; Kooi, C., The CepIR quorum-sensing system contributes to the virulence of *Burkholderia cenocepacia* respiratory infections. *Microbiology* **2003,** *149*, 3649-3658.

115.   Truong, T. T.; Seyedsayamdost, M.; Greenberg, P. E.; Chandler, J. R., A *Burkholderia thailandensis* acyl-homoserine lactone-independent orphan LuxR homolog that activates production of the cytotoxin malleilactone. **2015,** *197*, 3456-3462.

116.   Latifi, A.; Winson, M. K.; Foglino, M.; Bycroft, B. W.; Stewart, G. S. A. B.; Lazdunski, A.; Williams, P., Multiple homologues of LuxR and LuxI control expression of virulence determinants and secondary metabolites through quorum sensing in *Pseudomonas aeruginosa* PAO1. *Molecular Microbiology* **1995,** *17*, 333-343.

117.   Zhou, L.; Zhang, L.-H.; Cámara, M.; He, Y.-W., The DSF family of quorum sensing signals: Diversity, biosynthesis, and turnover. *Trends in Microbiology* **2017,** *25*, 293-303.

118.   Suneby, E. G.; Herndon, L. R.; Schneider, T. L., *Pseudomonas aeruginosa* LasR-DNA binding is directly inhibited by quorum sensing antagonists. *ACS Infectious Diseases* **2017,** *3*, 183-189.

119.   Rutherford, S. T.; Bassler, B. L., Bacterial quorum sensing: its role in virulence and possibilities for its control. *Perspectives in medicine* **2012,** *2*.

120.   Kordbacheh, H.; Eftekhar, F.; Ebrahimi, S. N., Anti-quorum sensing activity of *Pistacia atlantica* against *Pseudomonas aeruginosa* PAO1 and identification of its bioactive compounds. *Microbial Pathogenesis* **2017,** *110*, 390-398.

121.   Ismail, A. S.; Valastyan, J. S.; Bassler, B. L., A host-produced autoinducer-2 mimic activates bacterial quorum sensing. *Cell Host and Microbe* **2016,** *19*, 470-480.

122.   Tettmann, B.; Niewerth, C.; Kirschhöfer, F.; Neidig, A.; Dötsch, A.; Brenner-Weiss, G.; Fetzner, S.; Overhage, J., Enzyme-mediated quenching of the *Pseudomonas* quinolone signal (PQS) promotes biofilm formation of *Pseudomonas aeruginosa* by increasing iron availability. *Frontiers in Microbiology* **2015,** *7*, 1978.

123.   Gerdt, J. P.; Blackwell, H. E., Competition studies confirm two major barriers that can preclude the spread of resistance to quorum-sensing inhibitors in bacteria. *ACS Chemical Biology* **2014,** *9*, 2291-2299.

124.   Hentzer, M.; Wu, H.; Andersen, J. B.; Riedel, K.; Rasmussen, T. B.; Bagge, N.; Kumar, N.; Schembri, M. A.; Song, Z.; Kristoffersen, P.; Manefield, M.; Costerton, J. W.; Molin, S.; Eberl, L.; Steinberg, P.; Kjelleberg, S.; Hùiby, N.; Givskov, M., Attenuation of *Pseudomonas aeruginosa* virulence by quorum sensing inhibitors. *EMBO* **2003,** *22*, 3803-3825.

125.   Choi, K.-H.; Schweizer, H. P., Mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nature protocols* **2006,** *1*, 153-161.

126.   Choi, K. H.; Gaynor, J. B.; White, K. G.; Lopez, C.; Bosio, C. M.; Karkhoff-Schweizer, R. R.; Schweizer, H. P., A Tn7-based broad-range bacterial cloning and expression system. *Nature methods* **2005,** *2*, 443-448.

127. Cardona, S. T.; Valvano, M. A., An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* **2005,** *54*, 219-228.

128. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J., The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **2016,** *13*, 731-740.

129. Perseus *Replace missing values from normal distribution*: http://coxdocs.org/doku.php?id=perseus:user:activities:MatrixProcessing:Imputation:ReplaceMissingFromGaussian, 2015.

130. Coutinho, C. P.; de Carvalho, C. C.; Madeira, A.; Pinto-de-Oliveira, A.; Sá-Correia, I., *Burkholderia cenocepacia* phenotypic clonal variation during a 3.5-year colonization in the lungs of a cystic fibrosis patient. *Infection and Immunity* **2011,** *79*, 2950-2960.

131. Flaugnatti, N.; Le, T. T. H.; Canaan, S.; Aschtgen, M.-S.; Nguyen, V. S.; Blangy, S.; Kellenberger, C.; Roussel, A.; Cambillau, C.; Cascales, E.; Journet, L., A phospholipase A1 antibacterial type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Molecular Microbiology* **2016,** *99,* 1099-1118.

132. Boyer, F.; Fichant, G.; Berthod, J.; Vandenbrouck, Y.; Attree, I., Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* **2009,** *10*.

133. Gayne, R., The discovery of penicillin—New insights after more than 75 years of clinical use. . *Emerging Infectious Diseases* **2017,** *23*, 849-853.

134. Ebada, S. S.; Edrada, R. A.; Lin, W.; Proksch, P., Methods for isolation, purification and structural elucidation of bioactive secondary metabolites from marine invertebrates. *Nature Protocols* **2008,** *3*, 1820-1831.

135. Smith, T. E.; Pond, C. D.; Pierce, E.; Harmer, Z. P.; Kwan, J.; Zachariah, M. M.; Harper, M. K.; Wyche, T. P.; Matainaho, T. K.; Bugni, T. S.; Barrows, L. R.; Ireland, C. M.; Schmidt, E. W., Accessing chemical diversity from the uncultivated symbionts of small marine animals. *Nature Chemical Biology* **2018,** *14*, 179-185.

136. Shuji Kitamura, T. D. A., Steve Gonzales, Huerou Yvan, LeScott Alan Pratt, Yoshihisa Nakada Preparation of pyrazole, benzimidazole, indole, or indazole compounds useful as diacylglycerol acyl transferase inhibitors and for the treatment of obesity, hyperlipidemia, diabetes. 2007.

137. Suenaga, K.; Shimogawa, H.; Nakagawa, S.; Uemura, D., Catharsitoxins from the Chinese remedy qiung laug. *Tetrahedron Letters* **2001,** *42*, 7079-7081.

138. Lee, K.; Su Ha, G.; Veeranagouda, Y.; Seo, Y.-S.; Hwang, I., A CHASE3/GAF sensor hybrid histidine kinase BmsA modulates biofilm formation and motility in P*seudomonas alkylphenolica*. **2018,** *25*, 25.