MEADOWS, VICTORIA GAIL Ph.D. Investigating methylcitric acid cycle enzymes in *Bacillus subtilis* and polyketide synthase enzymes in *Bacillus amyloliquefaciens*. (2022)
Directed by Dr. Jason Reddick.123 pp.

Bacterial secondary metabolites from *Bacillus* species have an important role in drug discovery and development. This dissertation focuses on research to characterize two enzymes from the *Bacillus* species: *Bacillus subtilis* and *Bacillus amyloliquefaciens* and their biosynthetic pathways. The first chapter of this dissertation will focus on broad topics of understanding secondary metabolites and enzymology of *Bacillus* species. The second chapter will focus on the enzyme 2-methylcitrate dehydratase, MmgE, from the methylcitric acid cycle of *B. subtilis*. This enzyme is linked to energy production when the bacteria go through the process of sporulation due to stress. The enzyme was tested for substrate tolerance with the prediction that it could be used in bioorganic chemistry to add chiral alcohol groups to enoyl acids, and the dehydration of the beta-hydroxyacids to form olefin products with specific double bond geometries. The enzyme, MmgE, was shown to have limited substrate tolerance outside of the natural substrates. The maximum velocity rate for MmgE with malic acid is $V_{max} = 0.00149 \pm 2.009\ \frac{mM}{min}$. The $K_m = 2.78 \pm 0.58\ mM$. Comparability, MmgE with citric acid is $V_{max} = 0.051 \pm 0.007\ \frac{mM}{min}$. The $K_m = 0.186 \pm 0.075\ mM$. The third and last chapter will focus on characterization of the cytochrome P450 enzyme, PksS, from the biosynthetic pathway of the polyketide bacillaene. The *trans*-acyltransferase polyketide synthase product, bacillaene, is used as an antibiotic and antifungal in the biofilm of the bacteria against predators in the surrounding area. We made attempts to knockout *baeS*/*pksS* in their respective *Bacillus* species to determine the product of the reaction between the knockout mutations and the cytochrome P450 enzyme, PksS.

INVESTIGATING METHYLCITRIC ACID CYCLE ENZYMES IN *BACILLUS SUBTILIS* AND POLYKETIDE SYNTHASE ENZYMES IN *BACILLUS AMYLOLIQUEFACIENS*

by

Victoria Gail Meadows

A Dissertation
Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro

2022

Approved by

Dr. Jason Reddick
Committee Chair

DEDICATION

To my mom, thank you for always being there and for your continued support.

APPROVAL PAGE

This dissertation written by Victoria Gail Meadows has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Co-Chair

Dr. Jason Reddick

Committee Members

Dr. Nicholas Oberlies

Dr. Norman Chiu

Dr. Bruce Banks

October 7, 2022

Date of Acceptance by Committee

October 7, 2022

Date of Final Oral Examination

ACKNOWLEDGEMENTS

This dissertation was completed and made possible by the people who have stuck by me and supported me through the years. I am extremely fortunate to have had the privilege of working with Dr. Jason Reddick for almost eight years. I was given the opportunity to always be heard and be creative with my work during my studies. I cannot imagine working for a better mentor during this time. I would also like to thank my committee members: Dr. Nicholas Oberlies, Dr. Norman Chiu, and Dr. Bruce Banks for their support.

I have had several mentors over my time at UNCG. It has been such a pleasure working with all of you. The skills and techniques that I have learned from my mentors over the years are immeasurable. I would like to thank Tyler Graf for the constant support, advice, and friendship given throughout the years. In addition, to always being there to help fix our HPLC when it decided to break down and teach me how to troubleshoot. I would also like to thank Dr. Daniel Todd for all the time we spent with the mass spec and always being willing to help me with my many projects. I would like to thank Dr. Huzefa Raja for always being there to answer questions and showing me new techniques in the lab. I would like to thank Dr. Reynaldo Diaz for the constant help in the lab whenever something broke and helping teach me how to troubleshoot. For the millions of times, we would have to repair our HPLC because it broke down again.

I want to thank my fellow lab mates for all the support they have given me during this time. I want to thank Dr. Brittany Kiel for the support and patience that was shown to me when I first joined the lab and was learning techniques. I especially want to thank Brittney Ballentine for all the constant encouragement, jokes, and friendship I have received over the last several years that have made being in the lab enjoyable. I want to thank Ignacio “The Cardinal” Gallardo for

the never ending support I have received from our friendship, coffee runs to Tate Street, and making working long hours in the lab more bearable than they could have been. I am so grateful to Isaac Fernandez for all the work that was put into attempting to make knockout plasmids and always being willing to try new things in the lab as we learned those skills together. I want to give a very special thanks to Marve Basurto Campos, the Batman of the group, for all the work she has put into our project, especially on Transformation Tuesdays, coffee runs, and friendship. Working in the Reddick lab has always been a fun and friendly environment due to the wonderful people I have had the chance to work with over the years.

I also want to thank my department family. Over the years I have had constant support from the people in this department whenever there were either successes or failures in the lab. I especially want to thank Dr. Jennifer Simpson as there is no one else I can imagine or would have wanted to go through this experience with, and I could not have done this without you. I would like to thank Herma Pierre for always being there to show support in everything I do and for the advice I have received. I would like to give a special thank you to Tangela “The Cardinal” Johnson, Frank Morales Shnaider, David “The Cardinal” Tanas, Logan Brown, and Abraham “The Cardinal” Ustoyev for being the best group of people and support system I could have asked for during my time at UNCG. I would also like to thank Josh Frost and Abbie Horchar for their constant friendship.

Lastly, I want to thank my family for their continued support. To my grandparents who always pushed me to get a higher education and think of getting a PhD. Their continued love and support throughout my life. I want to especially thank my mom for always being there to support me over these years. I want to thank my mom and stepdad for always being part of my support

system, supplying me with food, and listening to me talk about my work even if they didn't understand it. I couldn't be where I am today without that love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

BACKGROUND

Bacillus species

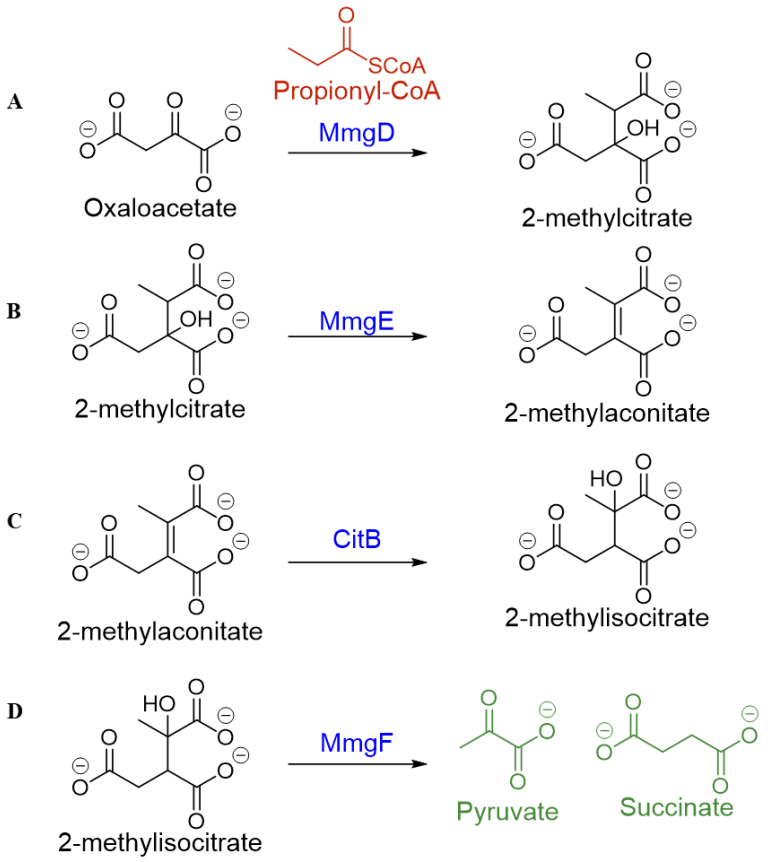
Bacillus subtilis and *Bacillus amyloliquefaciens* are Gram-positive, rod-shaped bacterial species that are widely used throughout scientific research.[1,2] *Bacillus subtilis* produces large amounts of proteins, making it invaluable for studying enzymes for industrial purposes and DNA modifications. *Bacillus subtilis* is a versatile bacterium that can be found in a variety of locations, including soil and aquatic habitats.[3,4] *B. subtilis* initially, called *Vibrio subtilis*, was discovered by Christian Gottfried Ehrenberg in 1835, and then was renamed by Ferdinand Cohn in 1872. *B. amyloliquefaciens* was first isolated from soil in 1943 by Juichiro Fukumoto. *B. amyloliquefaciens* is also associated with plant growth.[5,6] During the 20th century, *B. subtilis* would be used in medicine to treat gastrointestinal diseases.[7]

When *Bacillus* species experience difficult environmental or metabolic conditions, it undergoes the process of sporulation, creating an endospore.[8] This spore will be durable and resistant to the stressors in the surrounding area. Sporulation is the last effort the cell will make for survival.[9,10] The endospore preserves DNA and remains dormant to survive harsh conditions until it can return it the vegetative state. These hostile conditions can include extreme temperature, lack of nutrients, and predation. Asymmetric division will start at one end of the cell, resulting in the smaller forespore and larger mother cell.[9] When released from the mother cell, the spore will have a thick peptidoglycan spore cortex.[11] Despite being dormant, this allows the spore to survive in extreme conditions for an extended period.[8]

Methylcitric Acid Cycle in *Bacillus subtilis*

During the early process of sporulation and absence of glucose, the mother cell-specific sigma factor, σ^E, is activated to express multiple genes. One set of these genes includes the mother cell metabolic gene (*mmg*) operon which encodes for and is responsible for the expression of the methylcitric acid cycle and beta oxidation of fatty acids.[12,13] Once the mother cell-specific sigma factor is activated, important cycles like fatty acid degradation, the citric acid cycle, and the methylcitric acid cycle are all active. Studies have shown that methylcitric cycle is a source of additional energy during sporulation.[14] The *mmg* operon consists of six genes that encode for six enzymes: *mmgABCDE* and *yqiQ*, which has been reclassified as *mmgF* (Figure 1).[15] The first half of the operon consisting of the genes, *mmgABC,* that encode for the enzymes that will produce propionyl-CoA through branched fatty acid degradation. *B. subtilis* is not able to use propionate as a carbon source like *E. coli*, so propionyl-CoA is made through fatty acid degradation.[16]

Figure 1. Products of the methylcitric acid cycle.



A) Oxaloacetate and propionyl-CoA with MmgD will produce 2-methylcitrate. B) 2-methylcitrate with MmgE will produce 2-methylaconitate. C) 2-methylaconitate with CitB will produce 2-methylisocitrate. D) 2-methylisocitrate with MmgF will produce pyruvate and succinate.

The enzyme MmgA was discovered to be an acetoaceyl-CoA thiolase.[17] MmgB was found to be a 3-hydroxybutry-CoA dehydrogenase.[18] The enzyme MmgC was discovered to be an acyl-CoA dehydrogenase.[19] The genes, *mmgDEF*, which encode for the enzymes that will make up the methylcitric acid cycle, take propionyl-CoA from the fatty acid degradation and convert it to pyruvate.[20] The enzyme MmgD is a bifunctional citrate/methylcitrate synthase which will produce 2-methylcitate from propionyl-CoA/acetyl-CoA and oxaloacetate.[21] MmgE,

which is a 2-methylcitrate dehydratase, will form 2-methylaconitate from 2-methylcitrate.[22] 2-methylaconitate with CitB will produce 2-methylisocitrate, which can then be broken down by MmgF to succinate and pyruvate.[23] Evidence has shown MmgF to be a 2-methylisocitrate lyase. Succinate can then enter the citric acid cycle, or Krebs cycle, to regenerate oxaloacetate.[15,24]

The enzyme, MmgE, has been shown to have some substrate tolerance. MmgE reacts with the substrate citrate, which lacks the methyl group of 2-methylcitrate. The reaction produced *cis*-aconitate from citrate and two products from racemic 2-methylcitrate. The enzyme was shown to be able to react with one isomer from each pair of enantiomers of 2-methylctirate. Michaelis-Menten kinetics showed that the enzyme had significant activity with the substate citrate. The maximum velocity rate for MmgE with citrate is $V_{max} = 0.051 \pm 0.007\ \frac{mM}{min}$. The $K_m = 0.186 \pm 0.075\ mM$. The $k_{\text{cat}} = 170.3\ \frac{1}{min}$. In addition, HPLC analysis of enzyme reactions showed that racemic mixtures of the natural substrate of 2-methylcitrate would produce two products. Reactions preceding with propionyl-CoA, oxaloacetate, MmgD, and MmgE would only produce one enantiomer of 2-methylcitrate. MmgE converted that single substrate to just one of the products observed from the racemic 2-methylcitrate mixture. This physiological substrate was expected to be (*Z*)-2-methylaconitate, as described in reports for the methylcitric acid cycle from other species like *E. coli*.[15]

Polyketide Synthase Enzymes in *Bacillus amyloliquefaciens*

Natural products are often studied and used as inspiration in drug discovery.[4,25,26] Polyketides are a class of natural products that exist in plants, fungi, and bacteria are grouped together based on their biosynthetic pathway.[27] Polyketides can be used as antibiotics, antifungal, anticancer, and as immunosuppressants. Polyketide synthases (PKS) are multifunctional megaenzymes that are responsible for the biosynthesis of the polyketide. Polyketide synthases

enzymes work together as an assembly line through different modules to produce a specific carbon chain. PKS can be broken down into three different subgroups: types I-III PKS.[28] Type I PKS deals with multifunctional proteins that have individual domains and lead to macrolide biosynthesis, predominantly found in bacteria and fungi.[27] Type II PKS have individual monofunctional proteins and will lead to aromatic polyketides. Type II PKS can be found in bacteria. Type III PKS biosynthesis flavonoids and stilbenes.[28] Both type I and II start with acetyl-coenzyme A (CoA) or propionyl-CoA as the starter unit, while type III will start with cinnamoyl-CoA and use coenzyme A esters instead of an acyl carrier protein (ACPs) for functionality.

These reactions involve the synthesis and elongation of basic carbon chains from various carboxylic acid units via Claisen condensations. The Claisen condensation involves an enolate formed by decarboxylation. This negative charge will then attack a secondary carbonyl carbon, which will cause the phosphopantetheine thiol group to leave and will be left with the product of the newly formed ketone group. This will continue with each extender unit until the polyketide chain is complete.[29]

Each module contains a different set and order of this suite of enzymes depending on what polyketide carbon chain is being made. For instance, with type I PKS, it will pass to the next ACP and begin the next module until the process is complete. The condensation reaction is performed using a ketosynthase (KS) to elongate the chain. This all occurs at the acyl carrier protein (ACP) domain. There are two classes of modular polyketide synthesis acyl transferase systems, the first being *cis*-AT systems, and the second being *trans*-AT systems. With *cis*-acyltransferase polyketide synthase systems, the acyltransferase (AT) domain is present in each module of PKS. The acyltransferase will transfer the new building block to the ACP domain for

elongation. *Cis*-AT systems are easier to correlate with the make-up of the enzymatic modules and the final polyketide product. The *cis*-AT systems can show what the extender unit is for the polyketide by the acyltransferase domain, which can help in determining the overall structure of the polyketide. In addition, substrate preference can be shown in these systems. These systems have pathways where the modules are duplicated, with slight changes on which reduction enzymes are being used in each module.[30] A prominent example of a natural product derived from a *cis*-AT system is erythromycin.[31–33] Contrariwise, *trans*-AT polyketide synthase systems do not have the acyltransferase as part of the module, but as a separate enzyme. These systems have a lack of correlation between the modules and the final polyketide product, along with lack of the AT unit in the modules. *Trans*-AT PKSs rely on independent genes encoding acyltransferases, resulting in an unbound protein from the mega enzyme complex, for the building blocks of the polyketide.[34,35] This condensation with the ketosynthase results in the formation of a ketone. The resulting ketone can be modified further such as reduction to an alcohol via ketoreductase (KR), dehydration of the alcohol to an alkene via a dehydratase (DH) or reduction to an alkane via an enoyl reductase (ER). These different domains will modify the chain only if they are present in the module. Once the chain has been completed, the polyketide will be released from the acyl carrier protein via a thioesterase (TE).

Once the polyketide is released from the thioesterase, it can continue to go through post polyketide synthase modifications. This will increase the chemical diversity and biological activity of polyketides. These modifications can involve simple or complex pathways. Methyl transferases to add methyl groups can be added by S-adenosyl-L-methionine (SAM) mechanisms. Glycosylation modifications occur when saccharides are added to the polyketide by specific glycosidase enzymes.[36] Cytochrome P450 enzymes can perform simple and complex

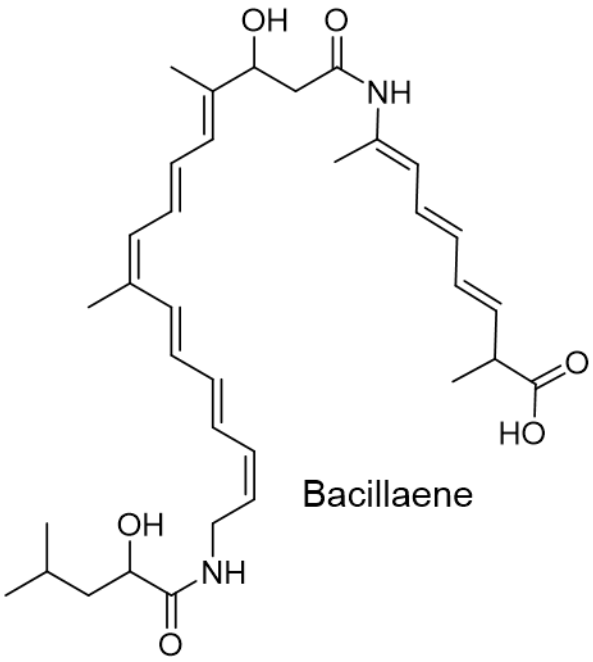
modifications such as hydroxylation and rearrangement reactions.[37,38] There is a large diversity within the class of cytochrome P450 enzymes accounting for the multiple types of reactions by the enzymes. The name of the enzyme for the cytochrome P450's came from the change in color from 420 nm to 450 nm when Fe is reduced and binds with carbon monoxide. This was discovered in the early 1960's by Sato and Omura.[39] These enzymes can be found in a wide variety of places in nature, from humans to bacteria. P450 enzymes will go through a catalytic cycle with iron and an electron source, like NAD(P)H.

Today, polyketides are important to biochemical research due to their antibiotic, antifungal, and anticancer activity. *B. amyloliquefaciens*, when in the presence of predators like *Myxococcus xanthus*, and *Streptomyces coelicolor* will produce antibiotics and antifungals as a defense.[40–42] Several polyketides have been identified from this species to fight against the fire blight disease causing bacteria *Erwinia amylovora*.[43] Unlike wild-type *B. subtilis*, laboratory strains like strain 168 are no longer capable of fighting off predators.[44,45] Wild-type *Bacillus* species will produce the antibiotics and antifungals in the biofilms and create time for the creation of spores. These antibiotics are so important to *Bacillus* species that roughly 5 % of the *B. subtilis* genome is devoted to antibiotic production.[46]

Bacillaene, produced by a *trans*-acyltransferase polyketide synthases system, is isolated from *B. subtilis* and *B. amyloliquefaciens*. The *pksX* gene cluster from *B. subtilis* 168 was sequenced in 1995 by Mayerl and coworkers and was determined to have similarities to proteins known to synthesize polyketides or used in fatty acid metabolism.[47] During this time they isolated bacillaene but were not able to solve the structure. They noted that the isolated bacillaene had antibiotic properties as it inhibited prokaryotic protein synthesis.[48] In 1997, the complete genome of *B. subtilis* 168 was sequenced.[46] The structure of bacillaene was solved in

2006 by Clardy and coworkers.[49] Only the structure of bacillaene and dihydrobacillaene were solved, but other analogs were noted, such as double bond isomers, dehydrated derivatives, and hydroxylated derivatives.[49] Over a decade later, the stability of the bacillaene family of compounds remains an issue for studying its biosynthetic pathway and biological activity.

Figure 2. *Bacillus* species polyketide Bacillaene.



Bacillaene is the product of the *pksX* gene cluster from *B. subtilis*. The *pksX* gene cluster was initially sequenced in 1995 and the structure of bacillaene was elucidated in 2006. The full biosynthesis has yet to be fully understood.

Bacillaene is the product of the *pksX* gene cluster which encodes a *trans*-acyltransferase (*trans*-AT), type I polyketide synthases/nonribosomal peptide synthases (Figure 2). These *trans*-AT type systems make up almost half of the known PKS but due to their complexity are less understood then the *cis*-AT type. The biosynthesis of bacillaene includes several unusual features like β-branching, a *trans*-enoylreductase, and a cytochrome P450 enzyme. The β-branch of bacillaene is installed by the HMG-CoA gene cassette that includes the genes *pksF-pksI* and

acpK.[49] The *trans*-enoylreductase, PksE, was identified to reduce the enoyl substrate between the 14'-15' carbons of the PksJ domain during the biosynthesis of dihydrobacillaene. Reddick and coworkers discovered that the cytochrome P450 enzyme, PksS, from *B. subtilis* would react with dihydrobacillaene and not bacillaene to give an unknown product.[50] It was initially thought that bacillaene was the polyketide synthase product. Piel then went on to confirm that dihydrobacillaene, which lacks a double bond found in bacillaene, is the immediate product of the polyketide synthase.[51] Walsh demonstrated that the *trans*-enoylreductase, PksE, reduces the double bond to produce the biosynthetic intermediate that ultimately produces dihydrobacillaene over bacillaene.[52]

This dissertation characterizes the findings of my research, which studies the secondary metabolism of the *Bacillus* species, specifically *B. subtilis* and *B. amyloliquefaciens*. Specific emphasis is shown on understanding specifically the cytochrome P450 PksS/BaeS post polyketide modifying enzymes of the polyketide bacillaene, full characterization of the enzyme MmgE from the methylcitric acid cycle, and the development of new techniques to aim in these goals.

CHAPTER II: INVESTIGATING MEYTHLCITRIC ACID CYCLE ENZYMES FROM *BACILLUS SUBTILIS*

ABSTRACT

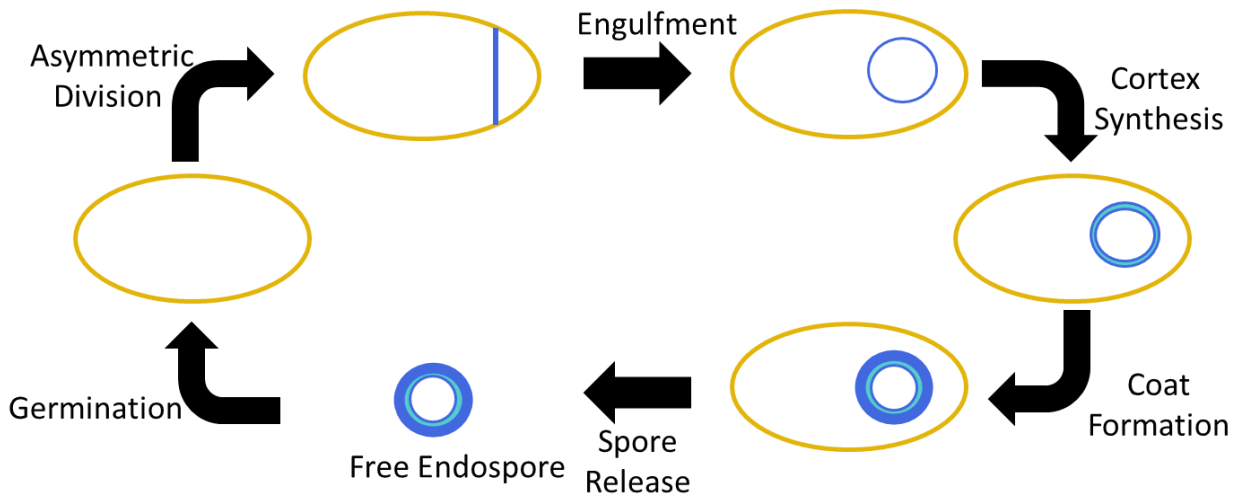
The enzyme, 2-methylcitrate dehydratase, or MmgE, from *Bacillus subtilis* was analyzed for its potential broad substrate tolerance in the reversible dehydration of hydroxy acid structures. Substrate tolerance would suggest potential uses in bioorganic synthesis to stereoselectively add alcohol groups to new compounds, in addition to dehydration reactions to form alkene groups with specific *E* or *Z* configurations. *In vitro*, MmgE has been shown to react with the commercial mixture of 2-methylcitrate and hypothesized to produce the (*Z*) and (*E*)-2-methylaconitate. Reactions involving MmgD and MmgE produce only one product which is expected to be the (*Z*)-2-methylaconitate. In previous studies, it was shown that reactions between MmgE and citrate, which lacks the methyl group of the natural substrate, would produce *cis*-aconitate. To explore the substrate tolerance, MmgE was reacted with a range of substrates with varying conditions, testing incubation time, incubation temperature and buffer conditions. In addition, the reversibly of the enzyme was investigated between *cis*-aconitate and citrate. We have gone on to show that MmgE has limited substrate tolerance and reversibility. The maximum velocity rate for MmgE with malic acid is $V_{max} = 0.00149 \pm 2.009\ \frac{mM}{min}$. The $K_m = 2.78 \pm 0.58\ mM$. Comparability, MmgE with citric acid is $V_{max} = 0.051 \pm 0.007\ \frac{mM}{min}$. The $K_m = 0.186 \pm 0.075\ mM$.

INTRODUCTION

Bacillus is a well characterized model Gram-positive bacteria species used in a wide variety of academic and biotechnological research.[1,53] *Bacillus subtilis*, *Bacillus veleznesis*, and *Bacillus*

amyloliquefaciens are used widely due to their production of enzymes. *B. subtilis* can be found in an array of areas, including soil and aquatic habitats.[46] When *Bacillus* species are in hostile environments, they will undergo the process of sporulation.[8] These hostile environments can include extreme temperatures, lack of nutrients, oxidation, radiation, dehydration, and predators.[8] Sporulation is the last effort the cell will make to survive the tough conditions. During this time, the cell will create an endospore that will be released from the mother cell.[9] The endospore can survive the harsh conditions for an extended period. The endospore will contain all the DNA of the mother cell and remain dormant to survive until it can return to the vegetative state (Figure 3).

Figure 3. Sporulation cycle of bacteria.

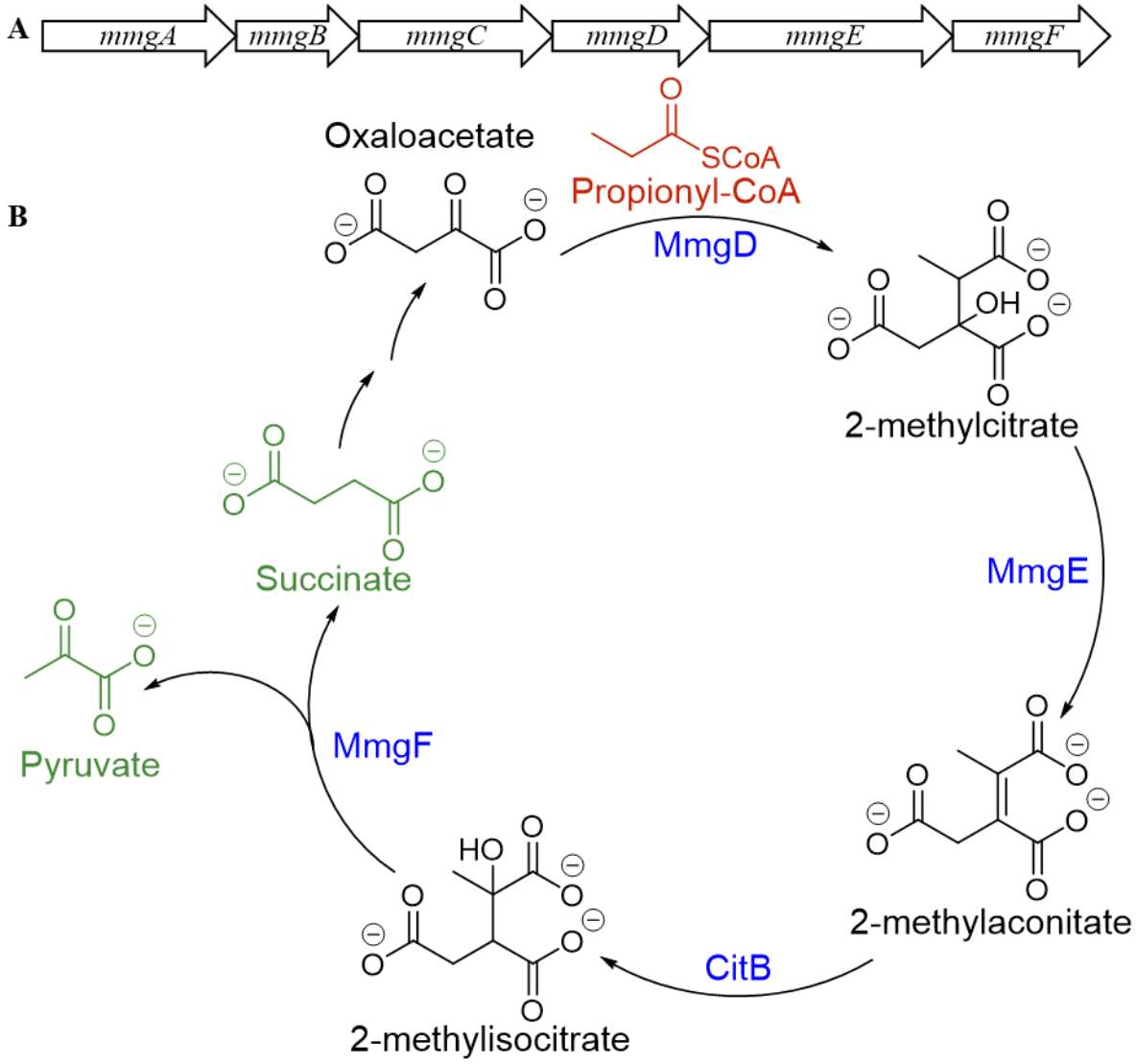


The bacteria will live in the vegetative state until a stress response triggers sporulation. These stress triggers can include lack of nutrients, extreme temperatures, and predators in the surrounding areas. Once sporulation begins, the mother cell will replicate the DNA and asymmetric division will begin at one end of the cell. The spore will be engulfed by the mother cell and a thick cortex coat will be formed around the spore. Once the coat is formed, the endospore will be released from the mother cell and can survive the harsh conditions for an extended period of time.

The cell will exist at stage 0, the vegetative cell, with normal cell growth. At stage I, the DNA of the mother cell will replicate two sets of chromosomes. Stage II consists of asymmetric division that will start at one end of the cell after DNA replication is finished. Stage III will start the engulfment of the forespore. Engulfment will occur when a membrane is formed around the forespore. The formation of the cortex made of peptidoglycan will define stage IV. Stage IV consists in the forespore becoming heat resistant. Stage VIII wraps up with the spore being released from the mother cell into the surrounding environment.[5,11] The fully formed endospore can go back into the vegetative state when nutrients are in the ecosystem, during the process called spore germination. Once sporulation begins, the mother cell-specific sigma factor, σ^E, is activated and will express multiple genes including the mother cell operon, or *mmg* operon.[54,55]

The mother cell metabolic gene (*mmg*) operon from *B. subtilis* encodes for the methylcitric acid cycle (*mmgDEF*) and part of fatty acid beta-oxidation (*mmgABC*) when the cell enters the early sporulation state and when glucose is absent (Figure 4).[55,56]

Figure 4. The mother cell metabolic gene pathway.



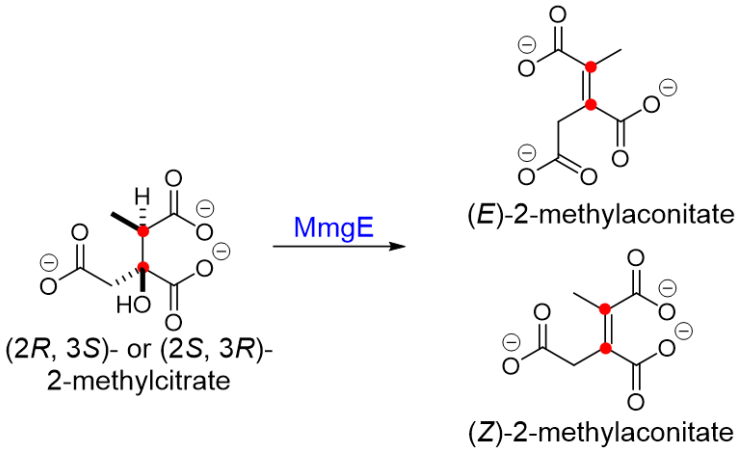
A) The mother cell metabolic gene (*mmg*) operon. The gene cluster that encodes for fatty acid beta-oxidation and the methylcitric acid cycle. B) The methylcitric acid cycle. The methylcitric acid cycle starts with 2-methylcitrate synthases, MmgD, which will catalyze the condensation of propionyl-CoA and oxaloacetate to yield 2-methylcitrate and coenzyme-A. From there, 2-methylcitrate is dehydrated by 2-methylcitrate dehydratase, MmgE, to produce 2-methylaconitate. 2-methylaconitate is rehydrated by aconitase, CitB, to produce 2-

methylisocitrate. The product is cleaved by 2-methylisocitrate lyase, MmgF, to produce pyruvate and succinate, which can then enter the citric acid cycle.

The *mmg* operon encodes part of a fatty acid β-oxidation pathway (*mmgABC*) and part of the methylcitric acid cycle (*mmgDEF*).[12,13] The methylcitric acid cycle is a pathway that gives bacteria, like *Escherichia coli*, the ability to grow on propionate as the sole carbon source.[16] *B. subtilis* is unable to use propionate, yet it still uses the methylcitric acid cycle. Instead, *B. subtilis* appears to use the methylcitric acid cycle (*mmgDEF*) to metabolize propionyl-CoA released by the β-oxidation of the branched fatty acids encoded by *mmgABC*.[2017,19] The *mmg* operon that forms the methylcitric acid cycle, *mmgDEF*, in *B. subtilis* are homologs of *prpCDB* from *Escherichia coli*.[57] The methylcitric acid cycle starts with 2-methylcitrate synthases, MmgD, which will catalyze the condensation of propionyl-CoA and oxaloacetate to yield 2-methylcitrate and coenzyme-A.[21] From there, 2-methylcitrate is dehydrated by 2-methylcitrate dehydratase, MmgE, to produce (*Z*)-2-methylaconitate.[22,58] 2-methylaconitate is rehydrated by aconitase, CitB, to produce 2-methylisocitrate.[23,59] CitB has been shown to be the sole aconitase enzyme in *Bacillus subtilis* from the citric acid cycle. The product is then cleaved by 2-methylisocitrate lyase, MmgF, to produce pyruvate and succinate, which can then enter the citric acid cycle.[24]

The physiological function of 2-methylcitrate dehydratase, or MmgE, is to produce 2-methylaconitate from either (2*R*, 3*S*) or (2*S*, 3*R*)-2-methylcitrate (Figure 5).[58] The heterochiral set of enantiomers (2*R*, 3*S*)/ (2*S*, 3*R*) are stereoisomers to the homochiral set of enantiomers (2*R*, 3*R*)/ (2*S*, 3*S*).[60] It has been shown that only one of the heterochiral isomers of 2-methylcitrate is made from the precursors with MmgD, propionyl-CoA, and oxaloacetate (Figure 6).[61]

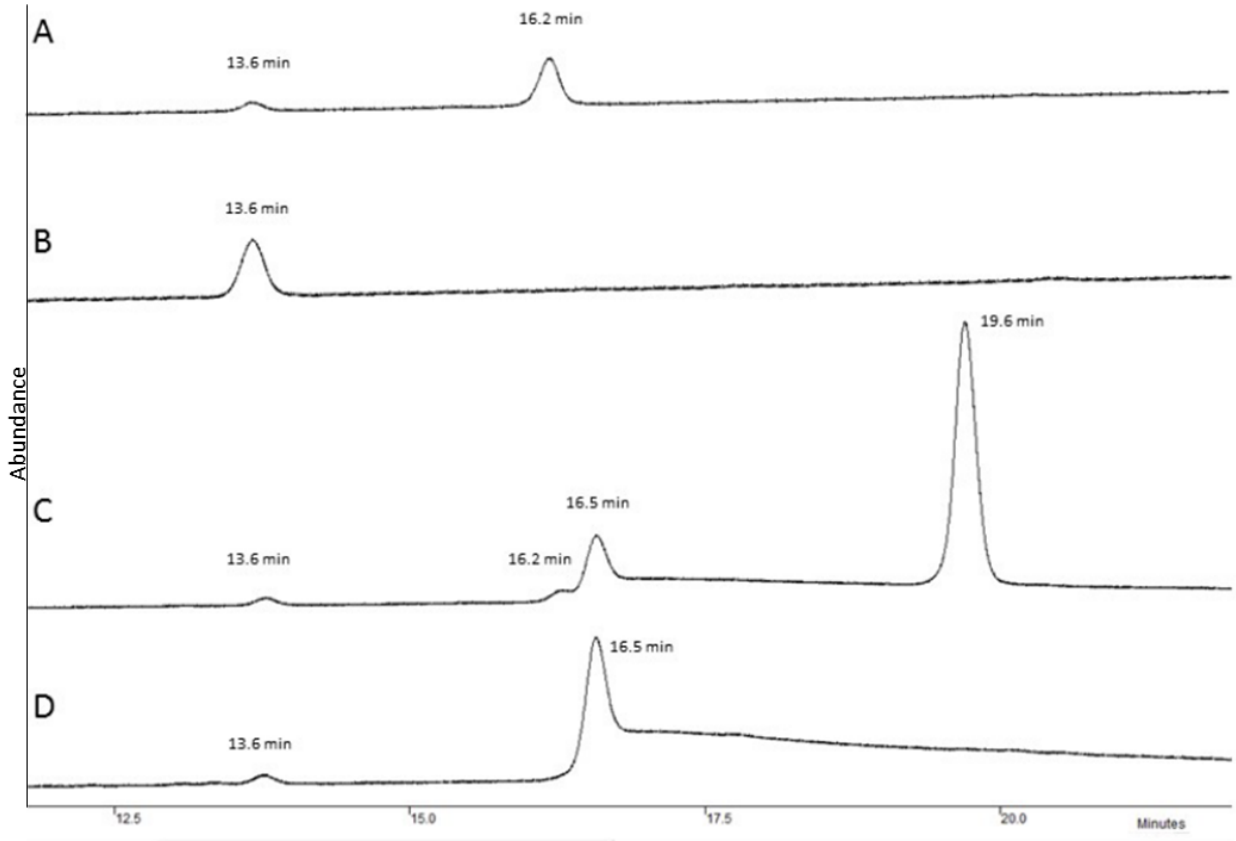
Figure 5. 2-methylcitrate reaction with MmgE.



The relative stereochemistry, instead of absolute chemistry, of 2-methylcitrate is shown for the (2*R*, 3*S*) and (2*S*, 3*R*) stereocenters. The reaction of oxaloacetate and propionyl-CoA with MmgD will produce either the (2*R*, 3*S*) and (2*S*, 3*R*) 2-methyctirate, but it is unknown which one is the product. Only one of those will go on to produce 2-methylaconitate.

At this time, it is unknown whether (2*R*, 3*S*) or (2*S*, 3*R*) 2-methylcitrate is being produced and will go on to react with MmgE. *In vitro* reactions have shown that the commercial mixture of 2-methylcitrate, which has all four isomers, will produce two products, initially hypothesized to be the (*Z*) and (*E*)-2-methylacontiate. The homolog of MmgE, PrpD, in *E. coli* has been thought to do a *syn*-elimination in the production of 2-methylaconitate with one of the homochiral 2-methylcitrate.[57,62] Later on it was shown to go through an *anti*-elimination with one of the heterochiral 2-methylcitrate substrates to produce 2-methylacontiate.

Figure 6. HPLC chromatograms of full reactions with MmgE and MmgD.

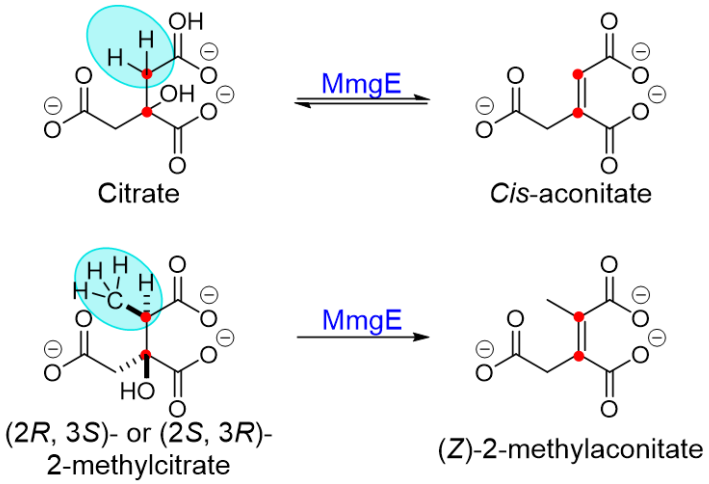


This work was done by Sydney Adams of the Jason Reddick laboratory at UNCG. A) The commerical racemic mixture of 2-methylcitrate. The left peak is the heterochiral 2-methylcitrate and the right peak is the homochiral 2-methylcitrate. The heterochiral 2-methylcitrate includes (2*R*, 3*S*) and (2*S*, 3*R*) 2-methylcitrate. The homochiral 2-methylcitrate includes (2*R*, 3*R*) and (2*S*, 3*S*) 2-methylcitrate. B) MmgD reaction with oxaloacetate and propionyl -CoA. The MmgD reaction will produce one of the heterochiral enantioner of 2-methylcitrate, but it is unknown which one. C) MmgE reaction with the commerical 2-methylcitrate. MmgE can react with one isomer of each the diastereomers, but is unknown which ones. This will lead to two proudcts, the potential (Z) and (*E*)-2-methylaconitate. D) The full

MmgD and MmgE reaction. Oxaloacetate and propionyl-CoA were reacted with both MmgD and MmgE to produce the naturally occurring products of the cycle.

In previous work, it was shown that the enzyme, MmgE, has some substrate tolerance by being able to convert the substrate, citrate, to *cis*-aconitate. Citric acid is missing the methyl group that 2-methylcitrate has on the α-carbon. Citrate was incubated with the MmgE enzyme for one hour at 37 °C and then the sample was analyzed on the HPLC. It was shown that the enzyme can produce *cis*-aconitate from citrate, compared to 2-methylcitrate producing 2-methylaconitate (Figure 7).[63]

Figure 7. Citrate reaction with MmgE.



Citrate lacks the methyl group of 2-methylcitrate, another substrate of MmgE. The hypothesis was that citrate has a less sterically bulky carbon without the methyl group and should fit into the active site with the proton. MmgE can react with citrate to produce *cis*-aconitate. *Cis*-aconitate is very similar in structure to (*Z*)-2-methylaconitate, with the only difference being the methyl group.

In this study we explored the substrate tolerance of 2-methylcitrate dehydratase, MmgE, from the methylcitric acid cycle in *B. subtilis*. Potential uses in bioorganic synthesis would lead

to stereoselectively adding/dehydrating alcohol groups in complex reactions. Another goal of this research was to determine the identity of the reaction products of the methylcitrate stereoisomers with MmgE.

MATERIALS AND METHODS

HPLC Methods

This method was run on a Varian ProStar HPLC with a Synergi 4μm Hydro-RP 80 Å LC column (250x4.6 mm). The solvent system that was used for the HPLC analysis was a mixture of 20 mM sodium phosphate buffer (pH 2.91) and methanol. The "methylcitrate gradient" had a flow rate of 0.7 mL/min and started at 0 % methanol. By 20 minutes the gradient reached 15 % methanol and was increased further to 60 % methanol by 25 minutes. It stayed at 60 % methanol for 10 minutes and then went back to 0 % methanol for another 10 minutes for an overall 50-minute method. This method was used throughout the work shown in this chapter. When analyzing substrate tolerance there was a change of having the gradient set to 0-25 % methanol over 0-30 minutes.

Purification of MmgE

E. coli strain BL21-Star replicating plasmid *mmgE*-pET-200 was streaked on an LB agar plate containing 50 μg/mL kanamycin and left to grow overnight at 37 °C.[22] The following day, a starter culture was made composing of 5 mL of LB, 50 μg/mL kanamycin, and one colony from the above streaked plate of *mmgE*-pET-200. The starter culture was shaken at 220 rpm at 37 °C overnight. The next day, 2 mL of the starter culture was combined with 1 L of LB and 50 μg/mL kanamycin. This was left to shake until the culture had reached an OD between 0.5-0.6 absorbance at 595 nm and was then induced with isopropyl B-D-1-thioglactopyranoside (IPTG) (1 mM final concentration) and was left to grow overnight. The culture was centrifuged at 7480*g*

for 30 minutes, and the pellets were then suspended into the binding buffer, which consisted of 5 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, and sonicated for 3 minutes in 30-second intervals to break the cell membranes. The mixture was centrifuged again at 16000*g* for 30 minutes, and the supernatant was filtered and then loaded onto a nickel nitrilotriacetic acid (Ni-NTA) affinity column (2 mL column bed) that was pre-equilibrated with binding buffer. After the supernatant was loaded onto the column, it was followed by 20 mL binding buffer, then 12 mL wash buffer, which consisted of 60 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, and 6 mL elute buffer, which consisted of 1 M imidazole, 0.5 M NaCl, and 20 mM Tris-HCl. At the elute buffer stage, 1 mL fractions were collected and analyzed with the Bradford assay reagent for protein. The fractions that contained protein were combined so that the buffer was exchanged by either Snakeskin dialysis tubing (22 mm x 7 cm, 7,000 MWCO) and left overnight in 4 L dialysis buffer or using fast protein liquid chromatography (FPLC) HiPrep 26/10 desalting column. Both methods used a dialysis buffer which consisted of 25 mM Tris-HCl (pH 8). The next day, the protein was collected from the dialysis. Glycerol was added to the protein to make a 10 % glycerol stock solution and stored in a -80 °C freezer. The dialysis was replaced by FPLC using a desalting column (HiPrep 26/10 Desalting column) where the protein fractions collected from the column would be collected in a syringe and inserted into the FPLC for dialysis. After the fractions with protein were eluted, the protein was concentrated by being spun down using a Vivaspin (20 5K MWCO PES) in a 14.5 rotor at 4 °C at 8,000*g* for 10 minutes. Glycerol was added to the protein to make a 10 % stock and then stored in the -80 °C freezer.

HPLC analysis of substrate tolerance of MmgE

DL-malic acid, meso-tartaric acid, D-tartaric acid, citraconic acid, mesaconic acid, oxalacetate, maleic acid, citramalic, 3-hydroxy butyric acid, 3-hydroxy-3-methyl pentatonic acid,

fumaric acid, 3-hydroxy glutaric acid, and L-tartaric acid were reacted with MmgE. These substrates were chosen due to the fact they were all 3-hydroxyacids and it was hypothesized that MmgE could dehydrate only at that position. Substrates were tested in a wide range of 0.5-3 mM for substrate tolerance because it was expected that these substrates would not be as efficient as 2-methylcitrate and citrate.

The substrates were incubated with MmgE in 20 mM Tris (pH 7.5) for one hour at 37 °C and were analyzed on HPLC using the method specified above. Initial reactions with meso tartaric acid, D-tartaric acid, L-tartaric acid, and malic acid were analyzed at 37 °C after 1 hour. Samples were taken out of the reaction every 12 minutes for analysis over the course of the hour. After samples were withdrawn, samples were quenched with 20 μL of 1 M sodium phosphate buffer (pH 2.91) and centrifuged down at 13,000*g* for five minutes for removal of the enzyme. The samples were then analyzed on HPLC with standards of substrates as a comparison.

When no reaction occurred, different incubation times, buffer concentrations, and pH's were explored. Both 13 mM HEPES and 100 mM Tris buffer were explored. The reactions were incubated in 100 mM Tris buffer going forward. The pH's tested included 6.5, 7, 8.0, and 8.1 from the original pH of 7.5. The incubation time was elongated from one hour to overnight and then 72 hours. Reactions were also analyzed at both room temperate and 37 °C. After no turnover with the substrates, 10 x and 20 x amount of the enzyme was used in an attempt to push the reactions forward. After analysis it was shown that the only substrate reacting with MmgE was malic acid. After malic acid was incubated with MmgE, no product was formed after a period of one hour. Malic acid was then incubated with MmgE for three nights with a successful turnover to product. A no-enzyme control was analyzed to demonstrate that the slow reaction

was not happening without the enzyme. Malic acid was shown to produce fumaric acid, when compared to a standard of fumaric acid, after incubation for 72 hours at 37 °C.

Michaelis-Menten Kinetics assay of MmgE with Malic Acid

Malic acid with final concentrations of 0, 0.1, 0.25, 0.5, 0.6, 0.8, 1.0, 1.25, 1.75, 2.0, 2.5 mM were mixed with 1.244 mg/mL of pure enzyme in 100 mM Tris buffer (pH 7.5) over the course of 72 hours. Samples were withdrawn (200 μL) at every 5 hours mark and 12 hours mark from the reaction and quenched with 20 μL of 1 M sodium phosphate buffer (pH 2.91) containing 5mM Itaconic acid, which used as an internal integration standard for HPLC. The samples were heated at 85 °C for five minutes and then centrifuged down at 13,000 rpm for five minutes and transferred to new vials. This was to ensure that the reaction was stopped, and the enzyme denatured and removed from the sample. The samples were analyzed on the HPLC with the previous conditions stated above. The peak areas of the malic and fumaric acid were then analyzed and compared to itaconic acid. This work was done with the help of Jazmine Williams, an undergraduate in the Jason Reddick Laboratory.

HPLC analysis of reversibility of MmgE

1 mM *cis*-aconitate was incubated with the enzyme in 100 mM Tris buffer (pH 7.5) for one hour at 37 °C and then analyzed by HPLC. After incubation, the sample was quenched with 10 % final volume 1 M sodium phosphate buffer (pH 2.91), incubated at 85 °C for five minutes and centrifuged at top speed for 5 minutes to remove the enzyme. Standards of 1 mM *cis*-aconitate and citrate were also made to compare to the reactions. *Cis*-aconitate is the thermodynamically less stable form of aconitate compared to *trans*-aconitate, so trace amounts of *trans*-aconitate were found in the *cis*-aconitate standard. This was confirmed with standards of the *trans*-aconitate. *Trans*-aconitate was then incubated with the enzyme in 100 mM Tris buffer

(pH 7.5) for one hour at 37 °C and then analyzed by HPLC. HPLC chromatograms show that *cis*-aconitate with MmgE will produce citrate and will also produce a larger amount of *trans*-aconitate.

NMR analysis of 2-methyl-*trans*-aconitate

The relative limit of detection was determined for the JOEL ECA-500 spectrometer operating at 500 MHz. 3,3-dimethyl acrylic acid was used as the standard for analysis. A range of 0.01-1 mM acrylic acid were analyzed (SI Figures 15-17). The relative limit of detection was hypothesized to be 0.01 mM 3,3-dimethyl acrylic acid (SI Figure 18). At that point, the peaks started to become almost indistinguishable from the baseline.

Cis and *trans*-aconitate were analyzed on NMR as comparisons to 2-methylcitrate as they only lack the methyl group structurally. 10 mg of *cis*-aconitate was dissolved in 750 μL D_2O for analysis on NMR. 10 mg of *trans*-aconitate was dissolved in 750 μL D_2O for analysis on NMR. A mixture of both compounds was made by dissolving 10 mg of each compound in 750 μL D_2O for analysis on NMR. All NMR experiments were collected on a JOEL ECA-500 spectrometer operating at 500 MHz. Experiments analyzed were conventional 1D 1H and 2D-NOSEY.

Reactions of MmgE and 2-methylcitrate in 100 mM Tris buffer (pH 7.5) were incubated for one hour at 37 °C. After incubation, the sample was quenched with 10 % final volume 1 M sodium phosphate buffer (pH 2.91), incubated at 85 °C for five minutes and centrifuged at top speed for 5 minutes to remove the enzyme. The substrates were then analyzed by HPLC for fraction collection of the two product peaks. Fraction collection occurred for peak one between 15.8-16.2 minutes and the second peak was collected between 18.2-18.9 minutes on the method. The contents of each peak were collected into separate tubes and stored in the freezer in between

collections. This work was done with the help of Don Mora and Jazmine Williams, undergraduates in the Jason Reddick Laboratory.

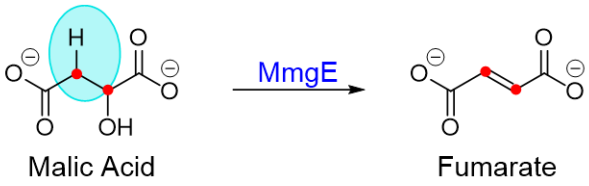
10 mg of (*E*)-2-methylaconitate was ordered from Sigma Aldrich (18802) and dissolved in 750 μL D_2O for analysis on NMR. All NMR experiments were collected on either a JOEL ECS-400 spectrometer operating at 400 MHz, or a JOEL ECA-500 spectrometer operating at 500 MHz. Experiments analyzed on the JOEL ECS-400 spectrometer were HMBC, HSQC, Carbon, 1D-NOE, and proton NMR. Experiments analyzed on the JOEL ECA-500 spectrometer were decoupled carbon, WATERGATE, and 2D-NOESY.

RESULTS AND DISCUSSION

HPLC analysis of substrate tolerance of MmgE

When MmgE reacted with the 2-methylcitrate product of MmgD, only (*Z*)-2-methylaconitate was produced. On the other hand, when MmgE was reacted with a commercially available sample of 2-methylcitrate, the enzyme produced (*Z*)-2-methylaconitate and a second compound, which we expected was the (*E*)-2-methylaconitate isomer. This led to the hypothesis that MmgE could react with one isomer of the heterochiral 2-methylcitrate and one isomer of the homochiral 2-methylcitrate. The physiological product of MmgE would be produced by the heterochiral 2-methylcitrate. MmgE was then reacted with the substrate, citrate, to show reactivity to produce *cis*-aconitate. These reactions showed some substrate tolerance, and it was hypothesized that the enzyme could react with other substrates if their structures included a β-hydroxy acid. DL-malic acid, meso-tartaric acid, D-tartaric acid, and L-tartaric acid were assayed to explore the substrate tolerance of the enzyme.

Figure 8. Malic acid reaction with MmgE.



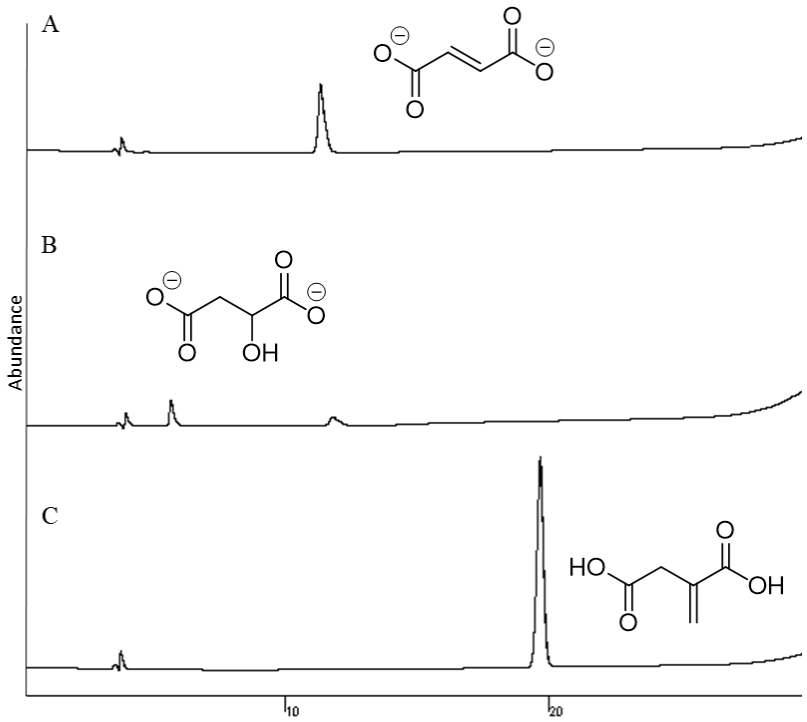
Malic acid was reacted with the enzyme to test the enzymes substrate tolerance. It was hypothesized that MmgE could be used in biotechnological applications in the selective synthesis of olefins or 3-hydroxyacids. It was shown that malic acid with MmgE will produce fumaric acid after incubation for 72 hours at 37 °C.

After exploring different pH of reaction buffers, incubation time, and concentrations, the reactions showed that only malic acid would react with MmgE over the course of three nights (Figure 8). This reaction was shown to be unfavorable compared to citrate, which is closer to the natural substrate. Meso-tartaric acid, D-tartaric acid, L-tartaric acid, and citraconic acid were shown to have no activity with the enzyme (SI Figure 19).

Michaelis-Menten Kinetics of MmgE with Malic Acid

Previous work showed that citrate was a substrate for MmgE. Michaelis-Menten kinetics reactions were run for the enzyme. The Michaelis-Menten kinetics were measured for this alternative reaction and the maximum velocity rate for MmgE with citrate is $V_{max} = 0.051 \pm 0.007\ \frac{mM}{min}$. The $K_m = 0.186 \pm 0.075\ mM$. The $k_{\text{cat}} = 170.3\ \frac{1}{min}$. With citrate, K_m value shows a strong affinity with MmgE. The Michaelis-Menten kinetics were not measured for 2-methylcitrate due to the racemic mixture of the commercial product.

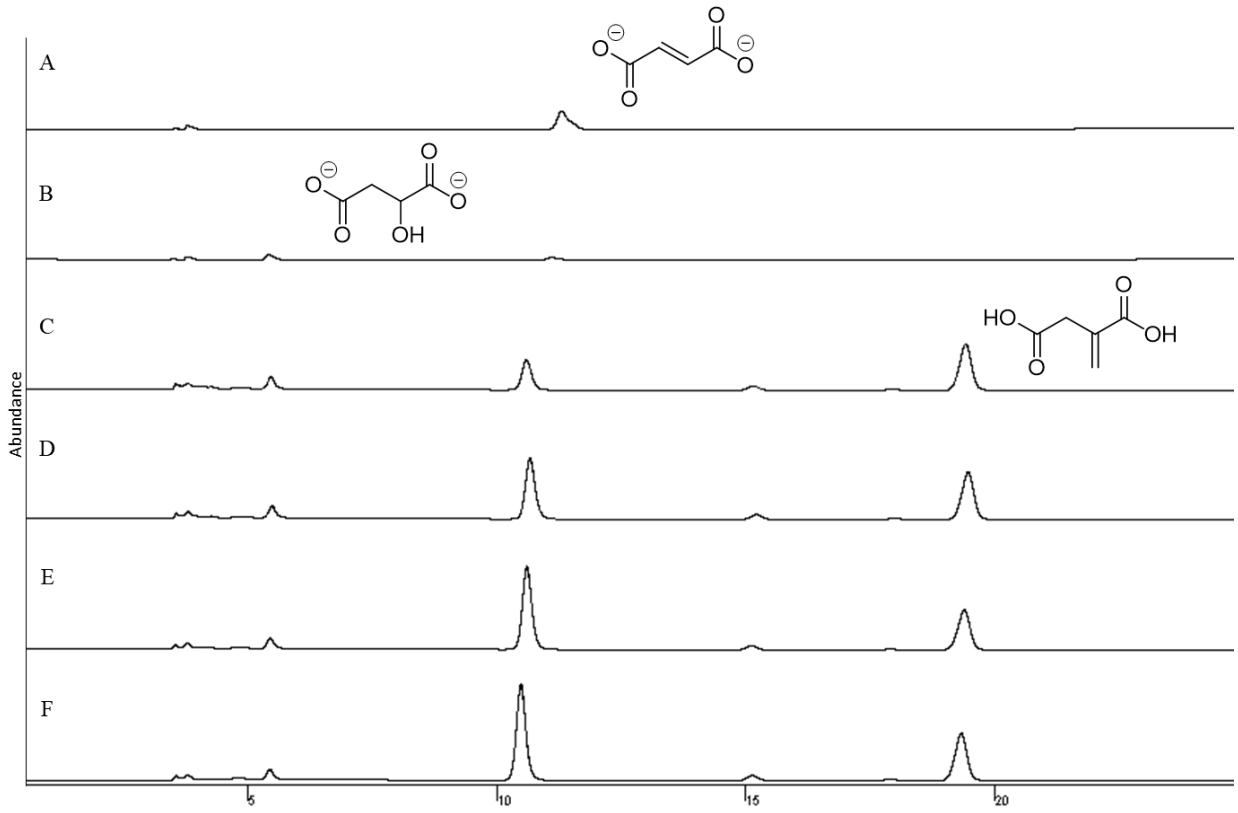
Figure 9. HPLC chromatograms of standards for malic acid kinetics analysis.



A) The standard of 1 mM fumaric acid. B) The standard of 1mM malic acid. C) The standard of 1 mM itaconic acid, which was used as an internal standard. Fumaric acid is the product of the reaction between malic acid and MmgE after a period of 72 hours at 37 °C.

Michaelis-Menten kinetics reactions were measured for Malic acid, at different concentrations stated above, and incubated with MmgE over a period of three nights. The reaction was then analyzed on HPLC using the method stated above. Standards of each substrate were used to compare the reaction for correct identification of the products (Figure 9).

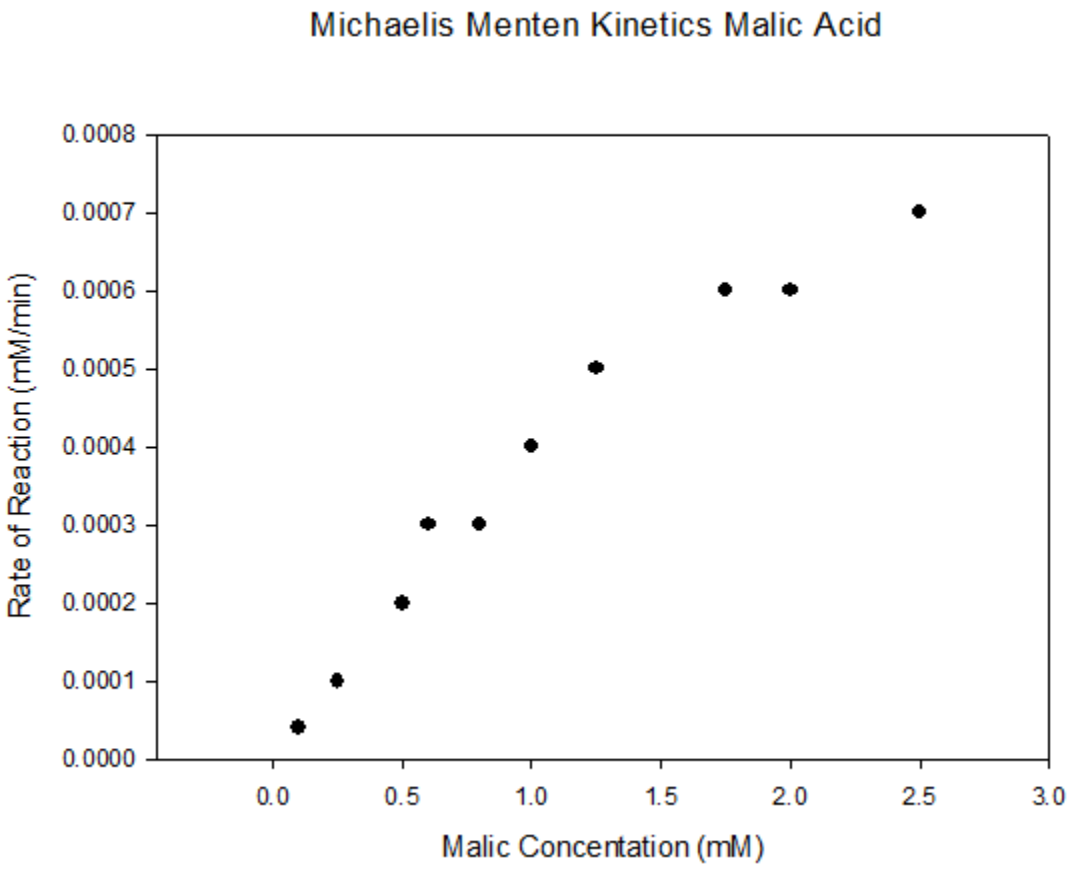
Figure 10. 2.0 mM malic acid reactions over the course of a period of 72 hours at 37 °C.



A) 1 mM fumaric acid standard for comparisons to the reactions. B) 1 mM malic acid standard for comparisons to the reactions. C) 2.0 mM malic acid after 7 hours of incubation with MmgE at 37 °C. D) 2.0 mM malic acid after 19 hours of incubation with MmgE at 37 °C. E) 2.0 mM malic acid after 43 hours of incubation with MmgE at 37 °C. F) 2.0 mM malic acid after 67 hours of incubation with MmgE at 37 °C. Reactions include the internal standard, itaconic acid.

Over the course of the three nights, there was an increase in fumaric acid as the malic acid reacted with MmgE (Figure 10). Itaconic acid was used as an internal standard to compare the peak areas of fumaric acid to malic acid over the course of the three nights.

Figure 11. Michaelis Menten Curve for Malic acid with MmgE.



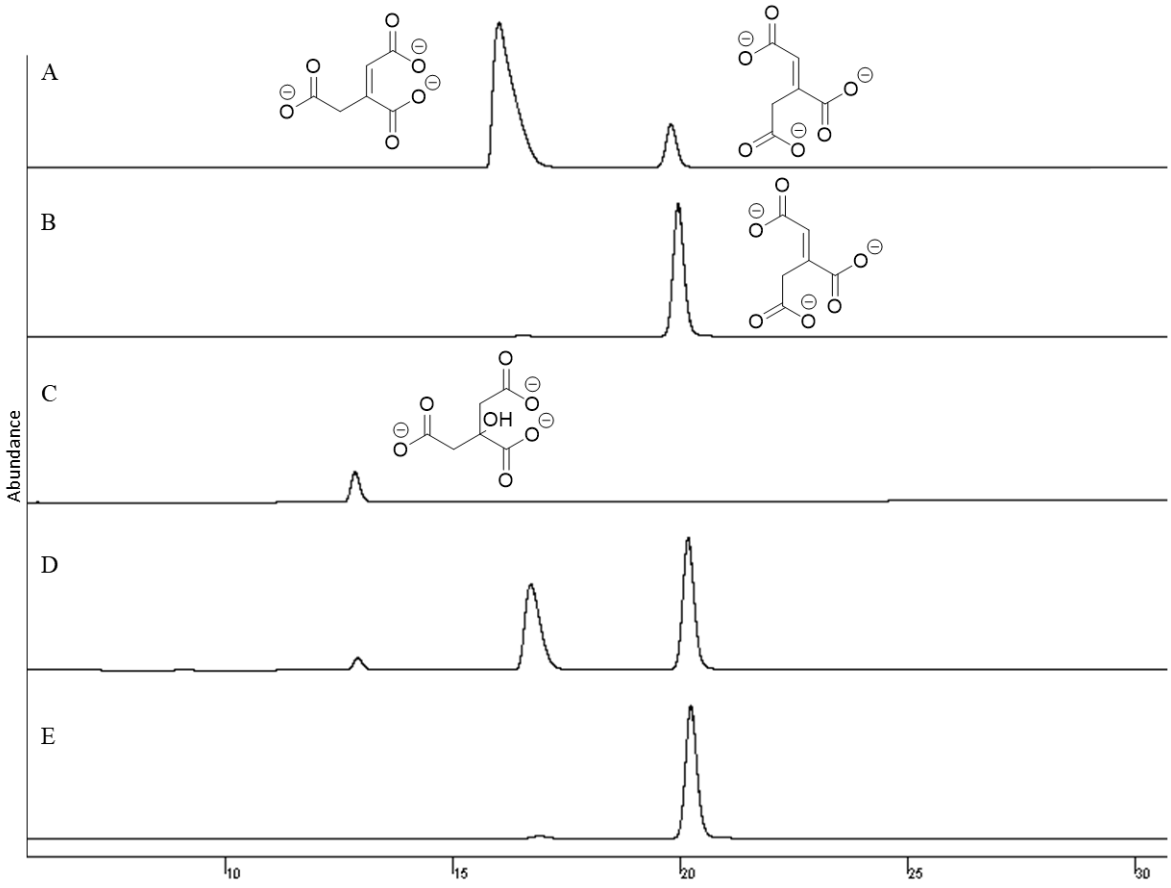
The kinetics shows that the substrate, malic acid, does not have strong affinity for the enzyme, MmgE. This is compared to the substrate citrate.

The internal standard, itaconic acid, would remain constant as there was a loss in malic acid as it was converted into fumaric acid. The maximum velocity rate for MmgE with malic acid is $V_{max} = 0.00149 \pm 2.009\ \frac{mM}{min}$. The $K_m = 2.78 \pm 0.58\ mM$. The substrate does not show strong affinity for the enzyme (Figure 11).

HPLC analysis of reversibility of MmgE

Previous work with the enzyme, MmgE, showed that it can convert citrate to *cis*-aconitate, and it was hypothesized that the reverse reaction could take place and that MmgE could turn over *cis*-aconitate to form citrate. *Cis*-aconitate was incubated with MmgE over the course of an hour then analyzed on HPLC.

Figure 12. Reverse reactions of *cis*-aconitate to citrate with MmgE.



A) The standard of *cis*-aconitate, which has trace amounts of *trans*-aconitate. B) The standard of *trans*-aconitate. C) The standard of citrate. D) The full reaction of *cis*-aconitate with MmgE after 1 hour at 37 °C. E) The full reaction of *trans*-aconitate with MmgE after 1 hour at 37 °C.

The chromatogram showed a new peak which corresponded with citrate, but also an increase in the *trans*-aconitate (Figure 12). The reaction would predictably take place through a two-step mechanism for this to occur. Water would attack the double bond in a conjugate addition forming an enolate. This can then lead to the production of citrate where the negatively charged carbon is protonated, or the sigma bond can rotate and dehydrate to produce *trans*-aconitate. Another pathway would be base deprotonation of the β-carbon to create a new double bond (SI Figure 20). *Trans*-aconitate was reacted with MmgE and it was shown that over a longer period it could produce trace amounts of *cis*-aconitate.

NMR analysis of 2-methyl-*trans*-aconitate

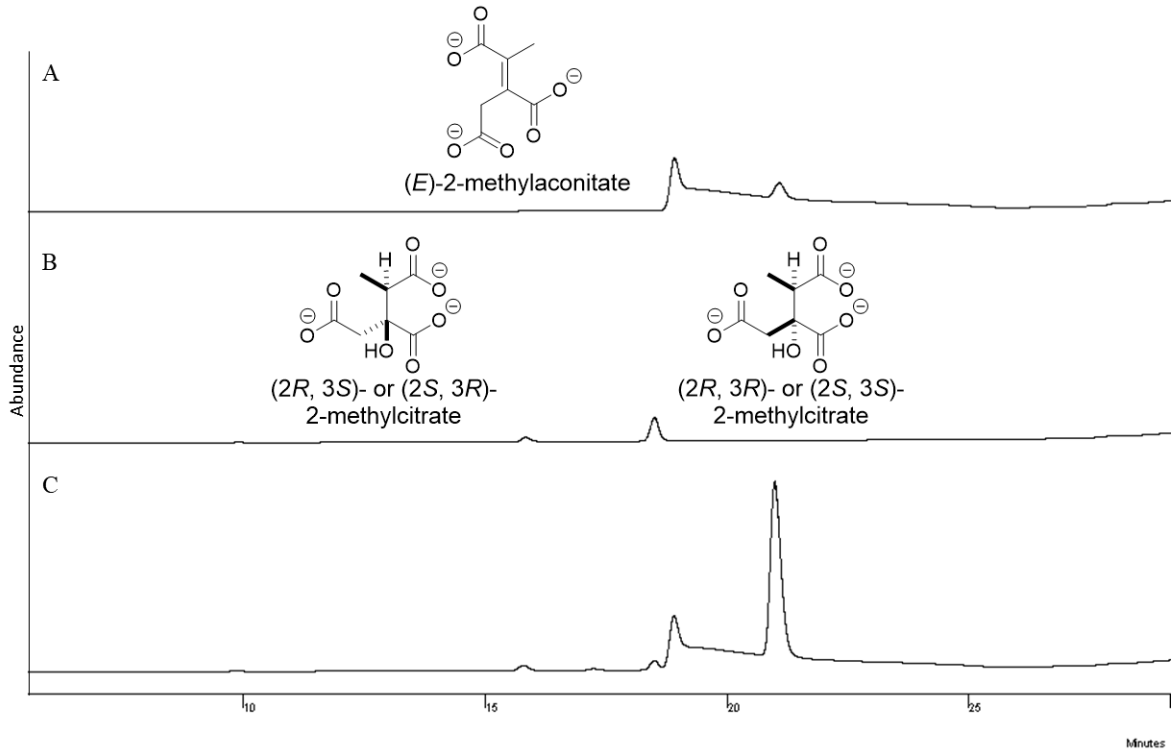
Previous work with MmgE showed that the racemic mixture of 2-methylcitrate would produce two products. Without standards, it was initially difficult to assign the structures of the products. It was hypothesized that the product found only from the reaction preceding with MmgD would be the (*Z*)-2-methylaconitate, with the secondary product being the (*E*)-2-methylcitrate. If the products were double bond isomers, as initially hypothesized, other methods of analysis would not be able to differentiate between the two compounds. NMR analysis of the products would determine the physiological product of the heterochiral 2-methylcitrate, and the product with the homochiral 2-methylctirate.

An analysis of the *cis*-aconitate was compared to the *trans*-aconitate on NMR. These were used to develop the NMR experiment, with the expectation that 2-methylcitate would behave similarly. *Cis* and *trans*-aconitate are similar to the methylaconitate but only lacking the methyl group. These were used to show that the *cis*-aconitate had a NOESY correlation between the methylene and ethylene protons while the *trans*-aconitate would now show a correlation. The ^{1}H NMR analysis of *cis*-aconitate contained four peaks, consistent with both the *cis* and *trans*-

aconitate being present in the mixture (SI Figure 21). The ^{1}H NMR analysis on *trans*-aconitate only contained one set of peaks, consistent with only the *trans*-aconitate being present in the sample (SI Figure 22). A mixture of both *cis* and *trans*-aconitate was analyzed to confirm that their peak shifts would align with the sample of the *trans*-aconitate, as it lacked another compound in the mixture. These peak shifts did correspond with the *cis* and *trans*-aconitate (SI Figure 23) The analysis showed that the *cis*-aconitate had a NOESY correlation between the methylene and ethylene protons (SI Figure 24). The *trans*-aconitate showed no correlation between the methylene and ethylene protons. This led us to believe that there would also be a correlation between methyl and methylene protons on 2-methylcitrate if the alkene was classified as (*Z*) 2-methylcitrate and no correlation if it was (*E*) 2-methylcitrate.

It was previously shown that MmgD makes the (2*R*,3*S*) or (2*S*,3*R*) 2-methylcitrate, but it was not determined which enantiomer was being produced. The product of the reaction between (2*R*,3*S*) or (2*S*,3*R*) 2-methylcitrate and MmgE coeluted with the peak that corresponded with the commercial (*E*) 2-methylaconitate (Figure 13). In a reaction between commercial 2-methylcitrate, containing both pairs of diastereomers, and MmgE, two products were noted and hypothesized to be both the (*Z*) and (*E*) 2-methylaconitate.

Figure 13. Commercial standards of 2-methylcitrate and 2-methylaconitate.



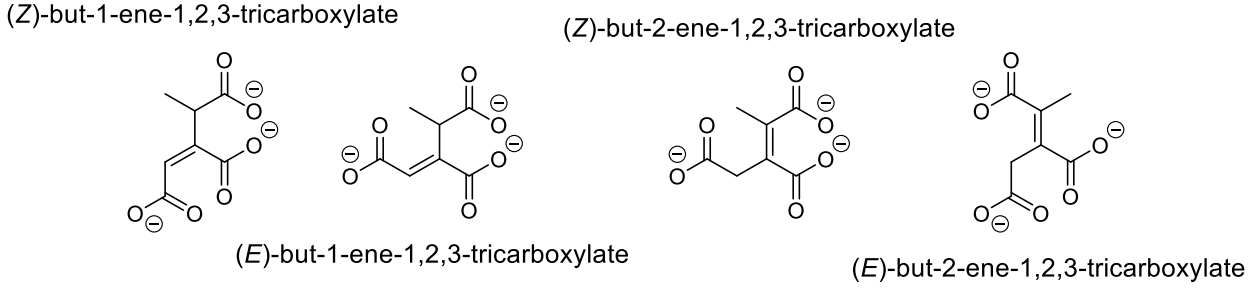
A) The commercial compound of (*E*) 2-methylaconitate from Sigma Aldrich (18802). B) The commercial mixture 2-methylcitrate. This mixture contains both pairs of diastereomers. The wedge and dash bond shows the relative stereochemistry of the compound. C) The reaction between 2-methylcitrate and the enzyme, MmgE. The enzyme will react with one isomer of each diastereomer, but it is unknown which one. The naturally occurring substrate for the enzyme appears to be the heterochiral 2-methylctirate. The heterochiral isomer will go on to produce the left product peak, which before was hypothesized to be the (*Z*) 2-methylaconitate.

Before commercial availability of (*E*) 2-methylaconiate, fraction collection of the two product peaks was done for NMR analysis. Reactions of 2-methylcitrate and MmgE were incubated as stated above and analyzed on HPLC for collection. These collections were analyzed on NMR and the chemical shifts were shown to be in the baseline. Fraction collections continued

in an attempt to collect enough sample for NMR, but not enough sample was collected for a readable proton NMR.

Recently, the (*E*) 2-methylaconitate has become available through the Sigma Aldrich catalog, although there is still no standard of the (*Z*) 2-methylaconitate. This commercial product contained two peaks observed through HPLC. The two peaks of the mixture corresponded with the two product peaks of the commercial racemic mixture of 2-methylcitrate. It was hypothesized that NMR testing could determine which isomer of 2-methylaconitate is the major product of the MmgE reaction. If the (*Z*) 2-methylaconitate is the major product, the hydrogens on the alpha carbon and the methyl will show correlation on a NOESY 2D spectra. If the (*E*)-2-methylaconitate is the major product there will be no correlation of the methyl group on a NOESY 2D spectra.

Figure 14. Potential products of the MmgE reaction with racemic 2-methylcitrate.



(*Z*)-but-2-ene-1,2,3-tricarboxylate, or (*Z*)-2-methylaconitate, has been shown to be one of the products of the reaction between MmgE and one of the heterochiral isomers of 2-methylcitrate. (*Z*)-2-methyaconitate should be the physiological product of the cycle. Either (*Z*) or (*E*)-but-1-ene-1,2,3-tricarboxylate will be the product of the reaction with the homochiral isomer of 2-methylcitrate. (*E*)-but-2-ene-1,2,3-tricarboxylate, or (*E*)-2-methylaconitate should not be a product of the MmgE and heterochiral 2-methylcitrate reaction.

The ^{1}H NMR analysis showed one set of peaks that could only correspond to one compound (SI Figure 25). If both the (*Z*) and (*E*) isomers of 2-methylaconitate were present, two sets of peaks would be shown on NMR analysis. This did not align with the HPLC analysis that showed two peaks in the chromatogram. The ^{13}C NMR analysis showed peaks that aligned with only one of the isomers being present in the commercial mixture (SI Figure 26). NMR analysis was also done the decoupled carbon to confirm that the ^{13}C NMR analysis showed just one of the two isomers (SI Figure 27). The 2D-NOESY showed correlation between the methyl and methylene peaks (SI Figure 28). This would align with the structure of the commercial producing being the (*Z*)-2-methylaconitate. NMR analysis with 2D-ROSEY was also done to confirm the NOESY of the (*Z*)-2-methylaconitate (SI Figure 29). This was then confirmed through 1D-NOE where there was a correlation between the methyl and methylene peaks (SI Figures 30 and 31). From the HSQC and Carbon experiments show correlation that is consistent with the (*Z*) 2-methylaconitate (SI Figure 32). This is consistent with the commercial major product being (*Z*) 2-methylaconitate. This would also suggest that the physiological product of the MmgE reaction with 2-methylcitrate is (*Z*) 2-methylaconitate (Figure 14).

CONCLUSIONS

In this study, we have shown that MmgE has limited substrate tolerance. MmgE can react with one of each diastereomers of 2-methylcitrate to produce both the (*Z*) and (*E*) 2-methylaconitate. The enzyme is also able to perform hydration reactions with *cis*-aconitate to produce citrate. It was also shown that the reaction of MmgE with *cis*-aconitate can produce larger amounts of *trans*-aconitate. MmgE can also react with the unnatural substrate, malic acid, to produce fumaric acid through an extended period. The commercially available 2-methyl-*trans*-aconitate was shown to be an impure mixture, containing two peaks on HPLC. NMR analysis

was shown to correlate the structure with the (*Z*) configuration. Either the (2*R*, 3*S*) or (2*S*, 3*R*) 2-methylcitrate will produce the (*Z*) 2-methylaconiate as the physiological product. Overall, MmgE is not as tolerant as initially thought, but still has a degree of substrate tolerance outside of the natural substrate of methylcitrate.

SUPPLEMENTARY DATA

Figure 15. ^{1}H NMR of 1 mM 3,3-dimethyl acrylic acid [500 MHz, D_2O].

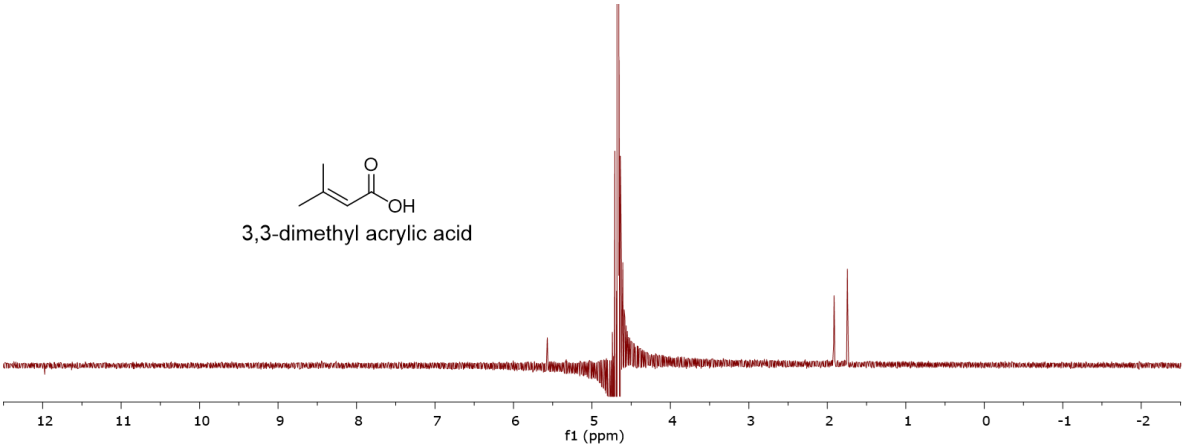


Figure 16. ^{1}H NMR of 0.5 mM 3,3-dimethyl acrylic acid [500 MHz, D_2O].

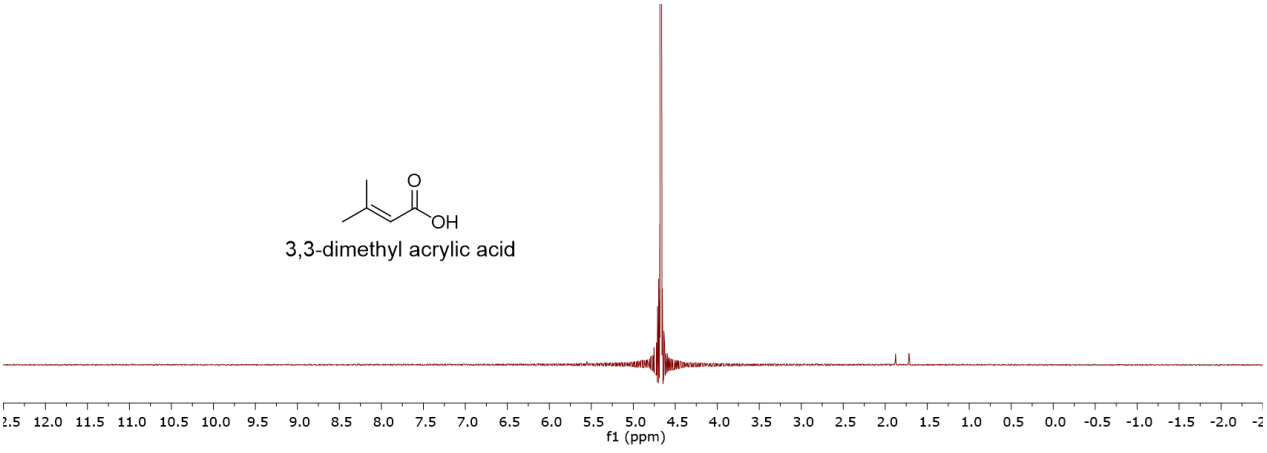


Figure 17. ^{1}H NMR of 0.05 mM 3,3-dimethyl acrylic acid [500 MHz, D_2O].

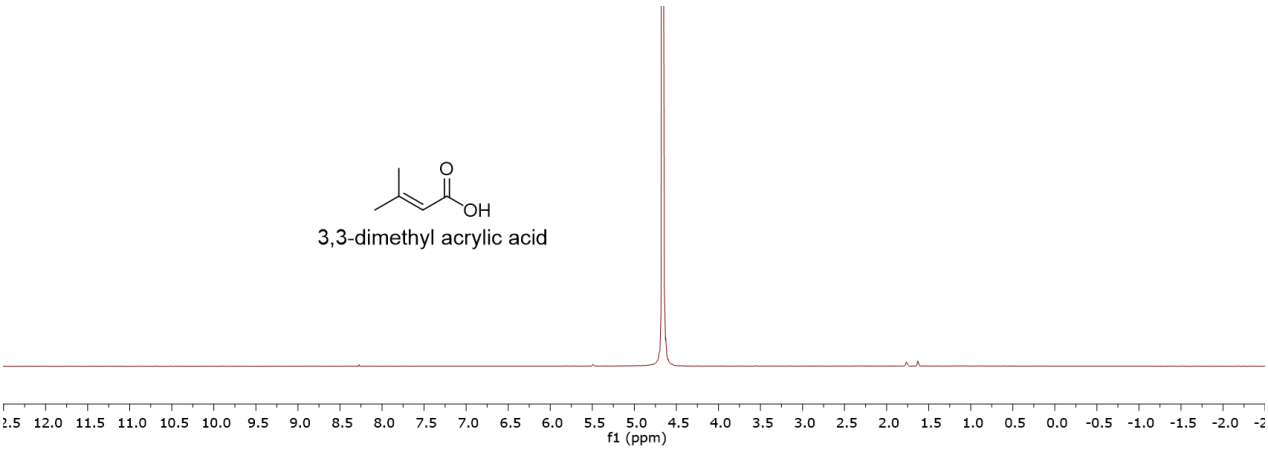


Figure 18. ^{1}H NMR of 0.01 mM 3,3-dimethyl acrylic acid [500 MHz, D_2O].

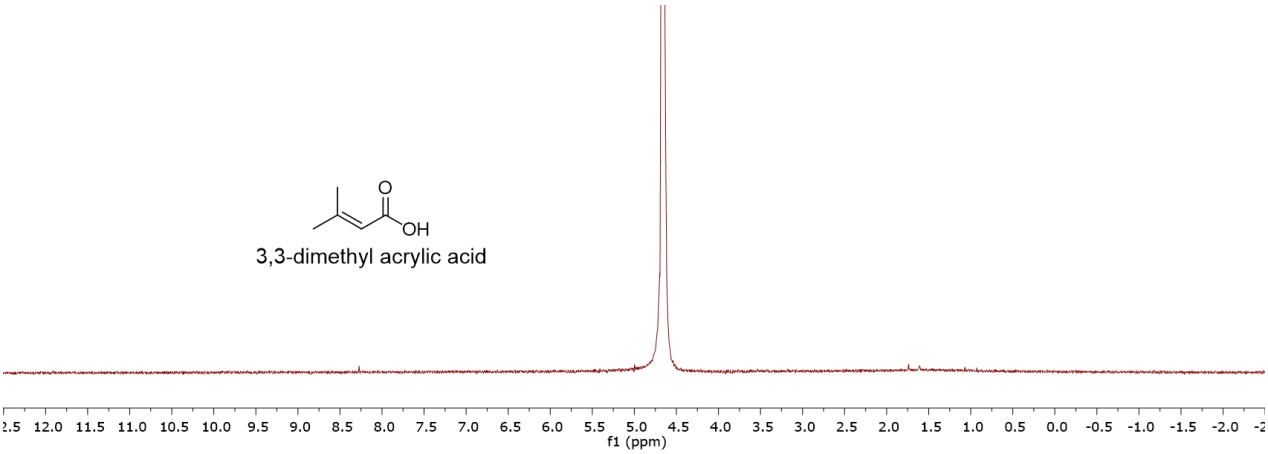
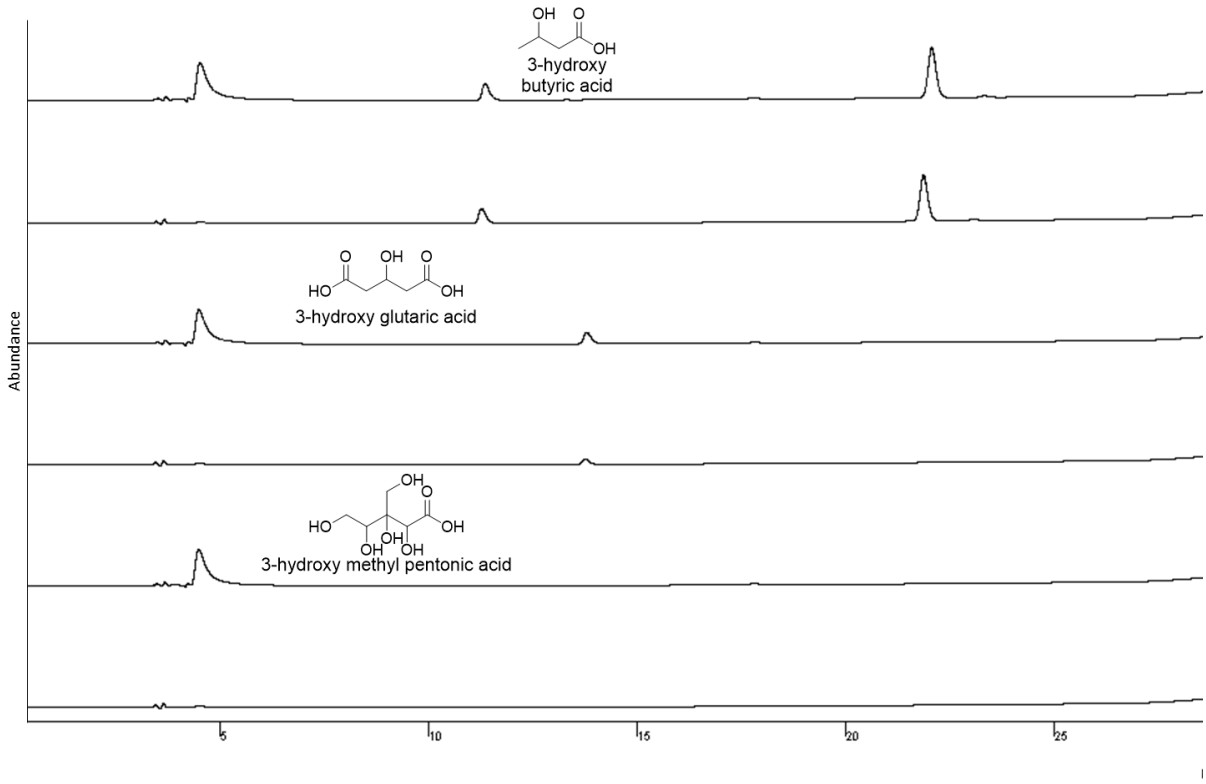
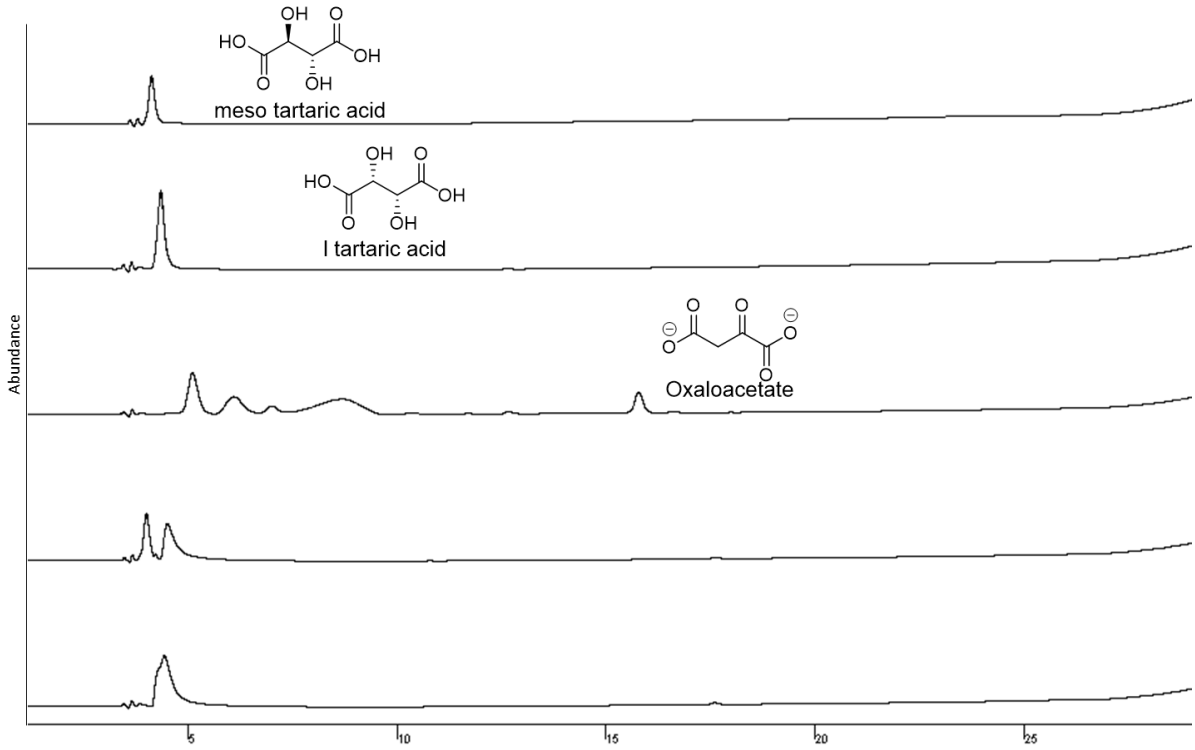


Figure 19. Substrance tolerence with MmgE with hydroxy acids.



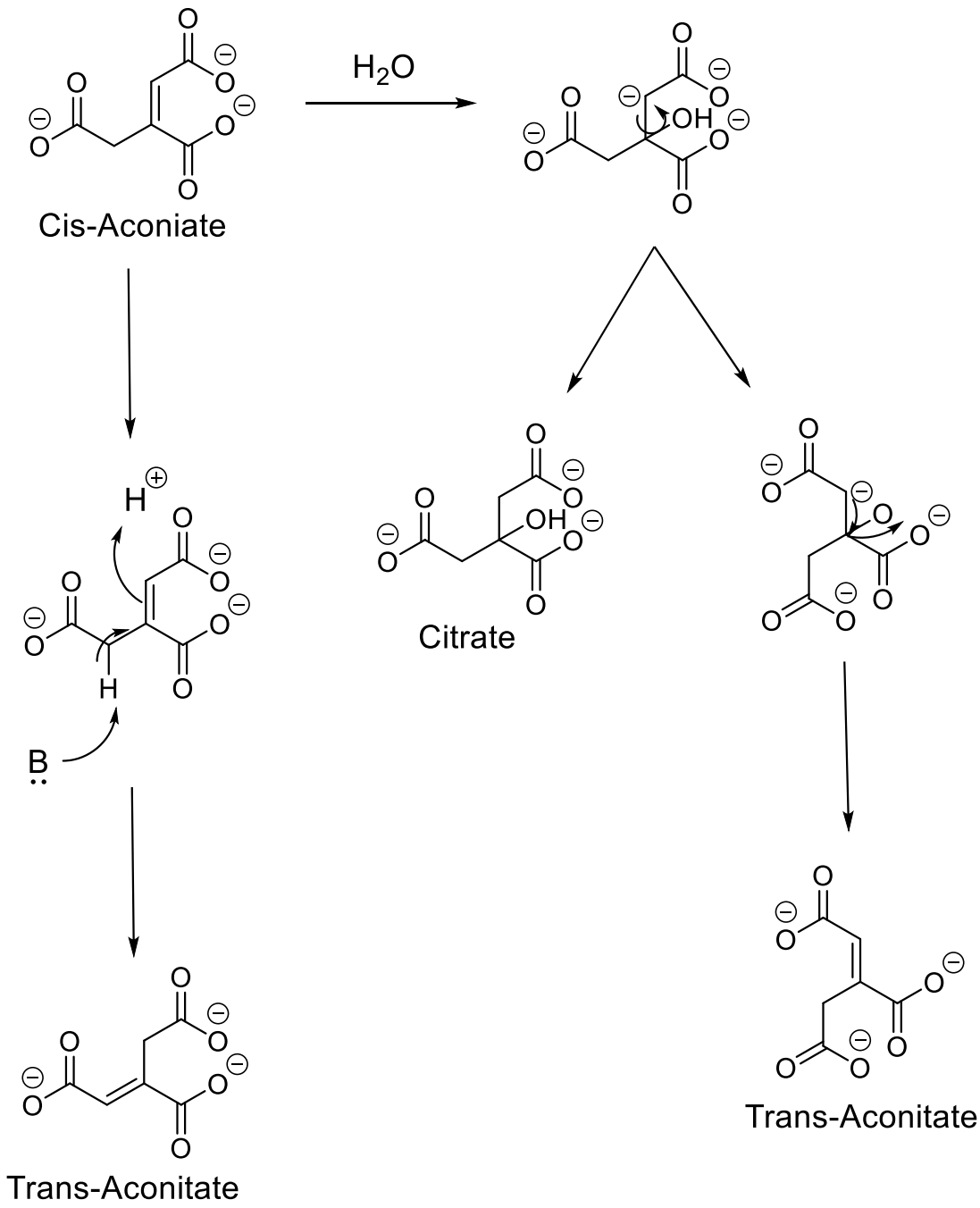
Three different hydroxy acids were analzyed with MmgE after incubation at 37 °C after 1 hour. 3-hydroxy butyric acid is shown as a standard and then after incubation with the enzyme. No additional peaks were shown in the full reaction with MmgE. 3-hydroxy glutaric acid is shown as a standard and then reacted with the enzyme. No additional peaks were shown in the full reaction with MmgE. 3-hydroxy methyl pentonic acid is shown as a standard and then reacted with the enzyme. No additional peaks were observed with the full reaction.

Figure 20. Substrate tolerance with MmgE with tartaric acids.



Meso tartaric acid and l tartaric acid were both tested with the enzyme hypothesizing that it could produce oxaloacetate. Standards of meso tartaric acid, l tartaric acid, and oxaloacetate are shown. The full reactions of meso tartaric acid and l tartaric acid showed no conversion to the product.

Figure 21. Predicted reversible reaction of MmgE with *cis*-aconitate.



Water would attack the double bond in a conjugate addition forming an enolate. This can then lead to the production of citrate where the negatively charged carbon is protonated, or the sigma bond can rotate and dehydrate to produce *trans*-aconitate. Another pathway would be base deprotonation of at the β-carbon to create a new double bond.

Figure 22. ^{1}H NMR of *cis*-aconitate proton [500 MHz, D_2O].

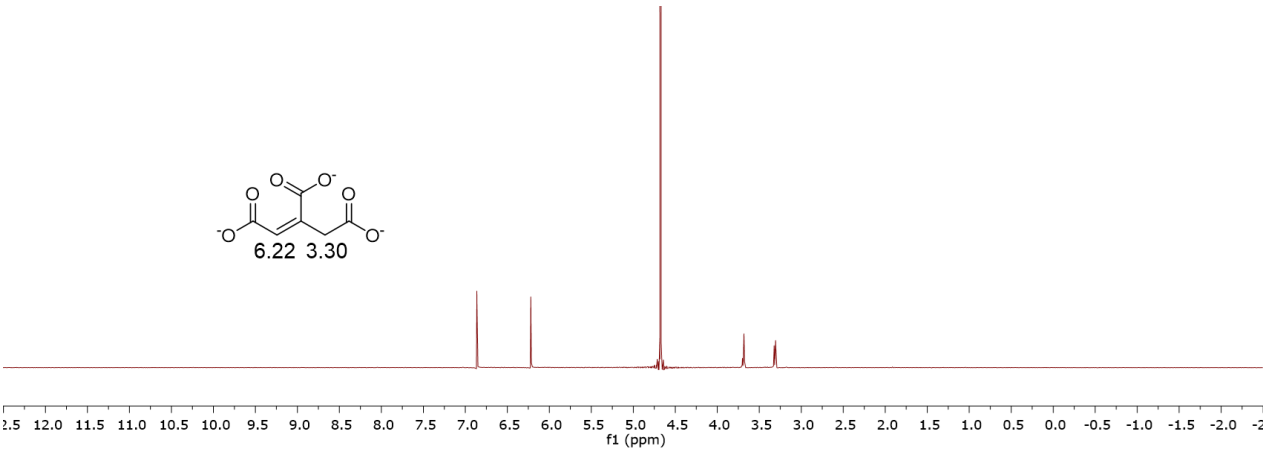


Figure 23. ^{1}H NMR of *trans*-aconiatate proton [500 MHz, D_2O].

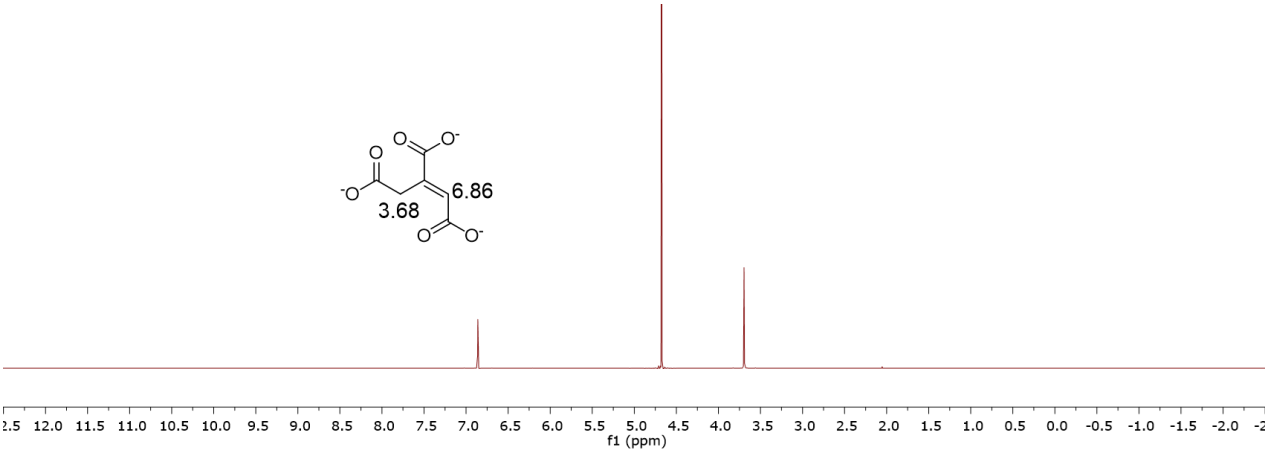


Figure 24. ^{1}H NMR of mixture of *trans* and *cis*-aconitate [500 MHz, D_2O].

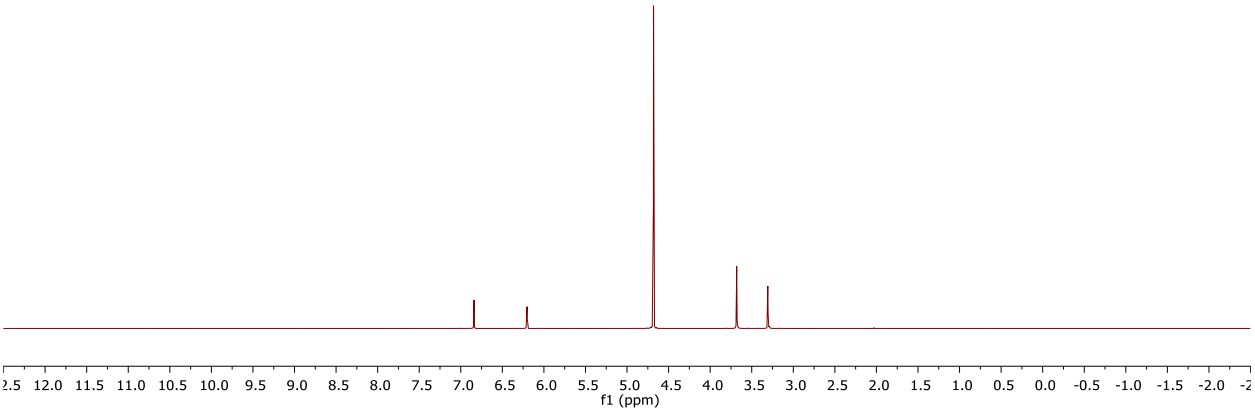


Figure 25. *Cis*-aconitate NOESY NMR [500 MHz, D_2O].

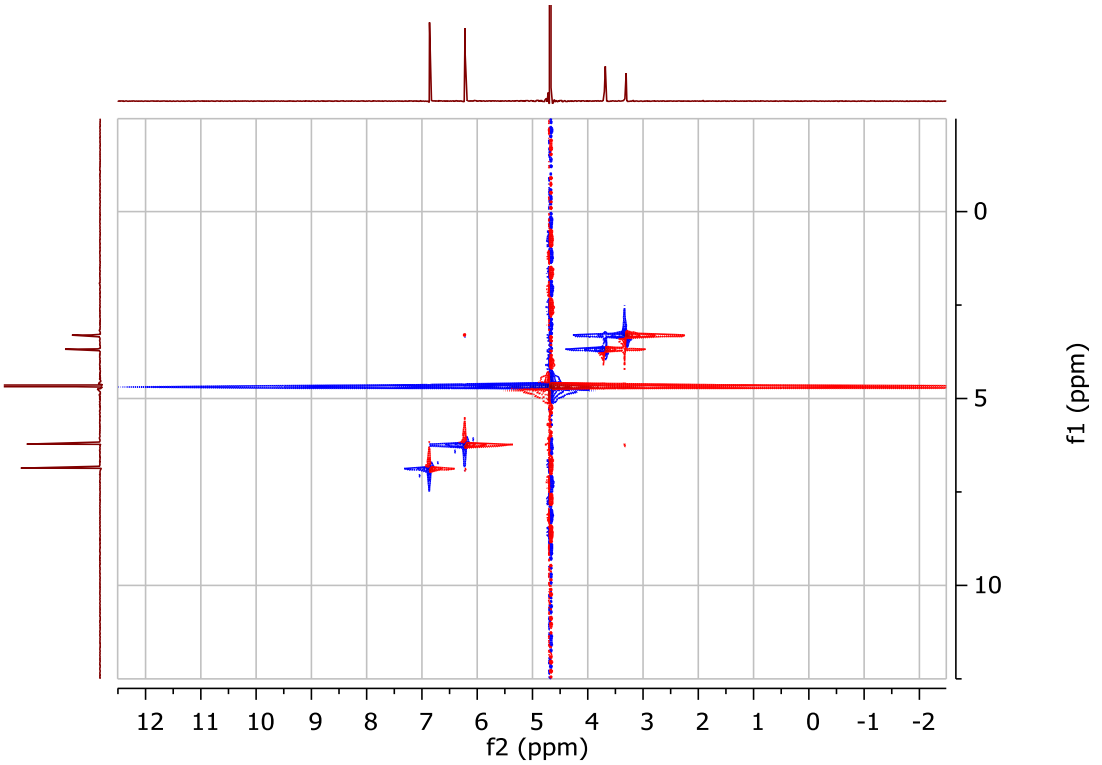


Figure 26. ^1H NMR of (*E*) 2-methylaconitate [500 MHz, D_2O]. Commercial mixture of (*E*) 2-methylaconitate provided by Sigma Aldrich (18802).

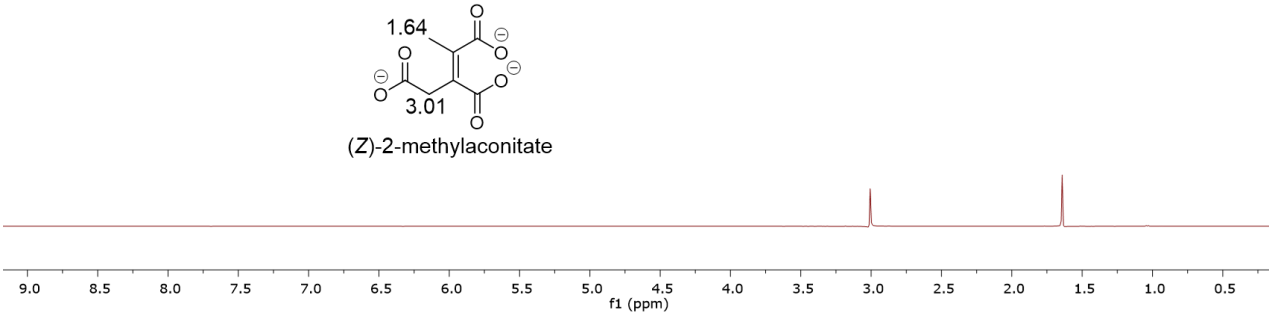


Figure 27. ^{13}C NMR of Methylaconitate [500 MHz, D_2O].

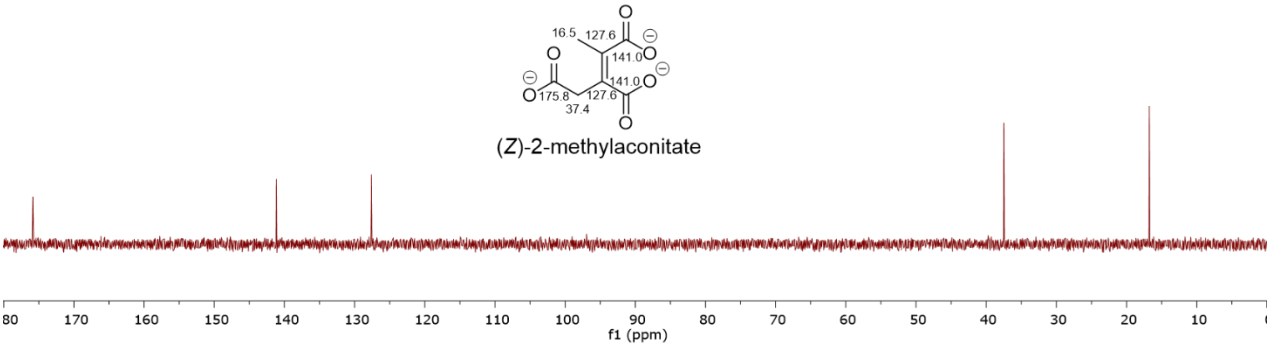


Figure 28. ^{13}C NMR of Methylaconitate [500 MHz, D_2O]. Decoupled carbon NMR.

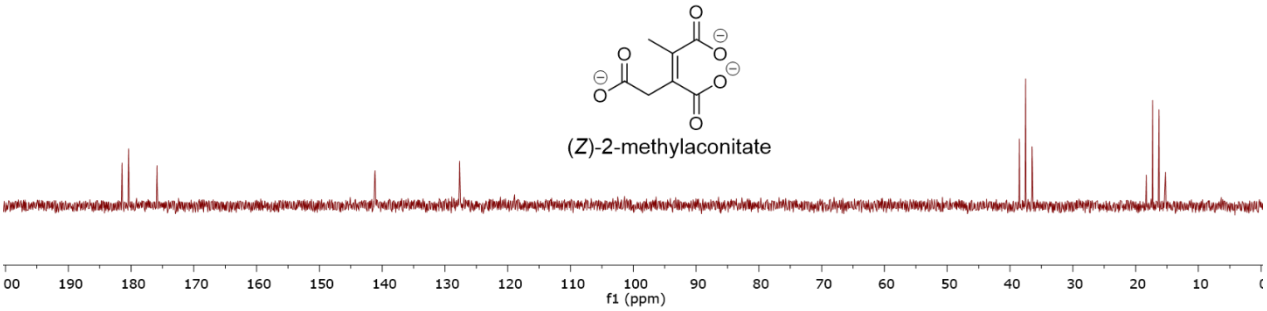
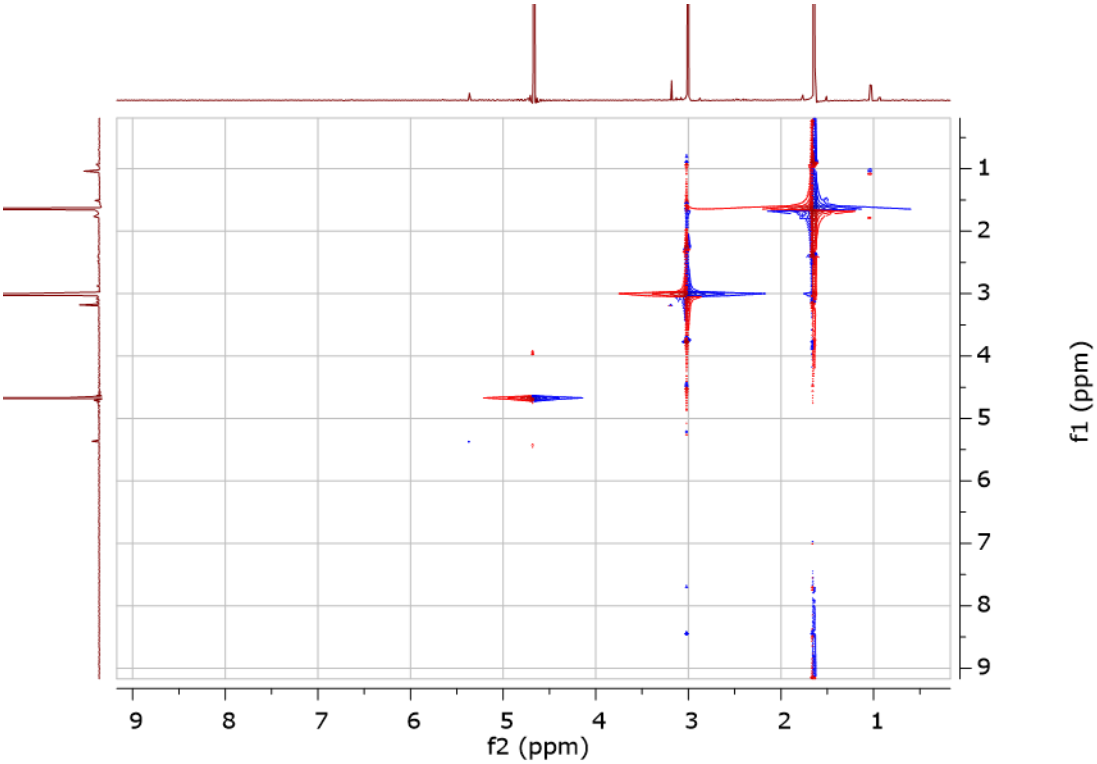


Figure 29. NOESY NMR of 2-methylaconitate [500 MHz, D_2O].



(*Z*)-2-methylaconitate

NOESY was run to determine if there was a correlation between the methyl and methylene of 2-methylaconitate. The commercial mixture of (*E*) 2-methylaconitate was purchased from sigma. If NOESY showed a correlation between the methyl and methylene the bond would be in the *cis* or (*Z*) formation. If there was no correlation between the methyl and methylene the bone would be in the *trans* or (*E*) formation. The NOESY showed a correlation between the methyl and methylene peak, showing the commercial mixture to be the (*Z*) 2-methylaconitate.

Figure 30. ROESY NMR of Methylaconitate [500 MHz, D_2O].

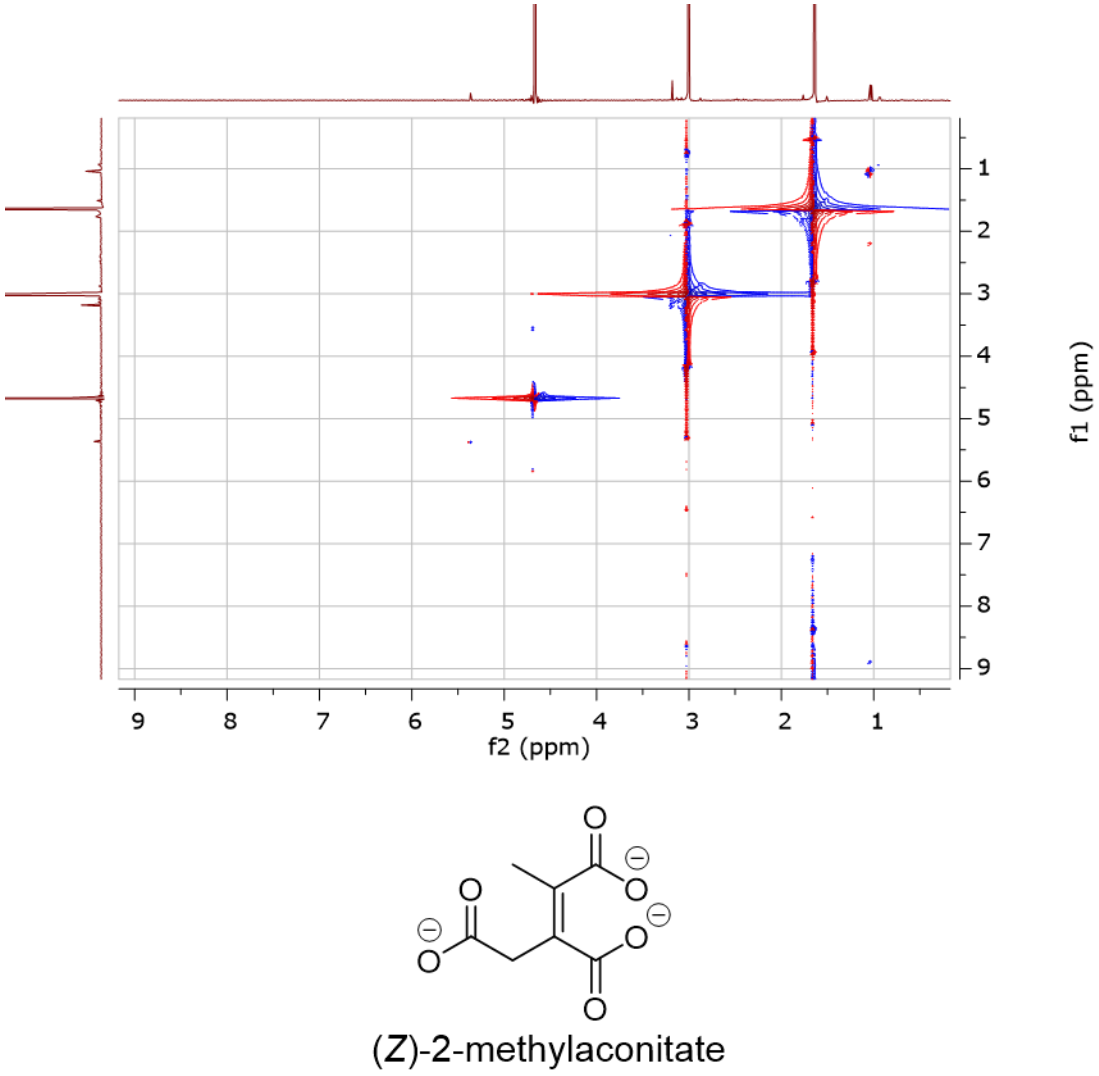


Figure 31. 1D NOE methyl group of Methylaconitate [500 MHz, D_2O].

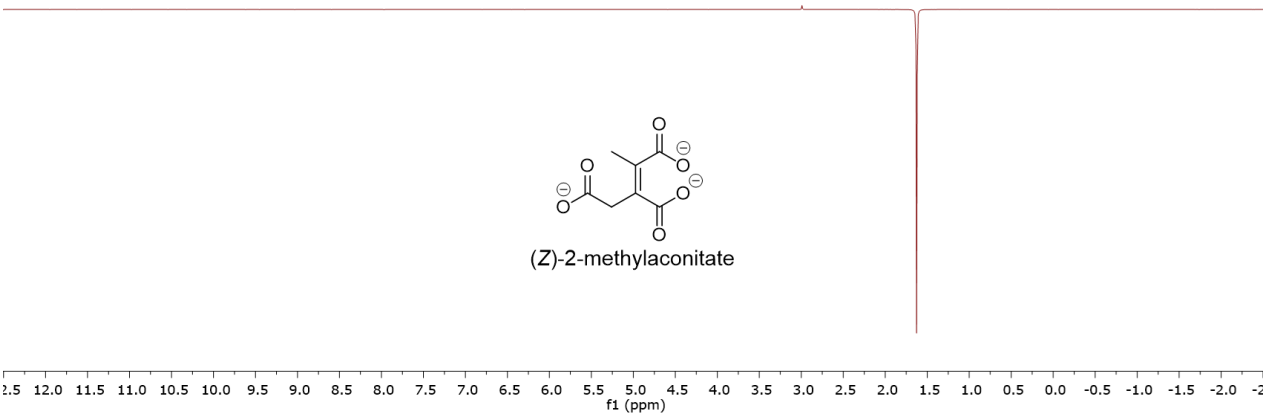


Figure 32. 1D NOE methylene group of Methylaconitate [500 MHz, D_2O].

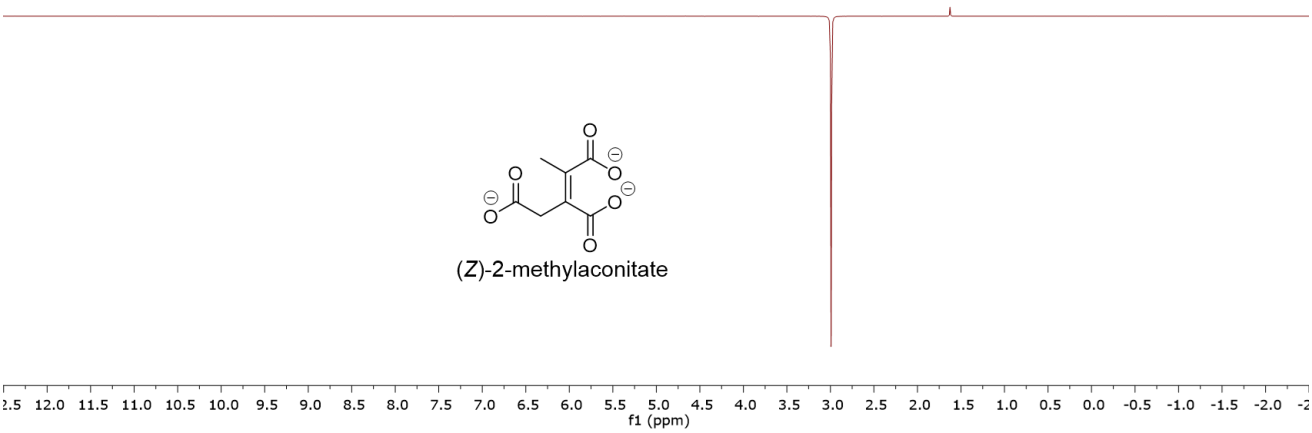
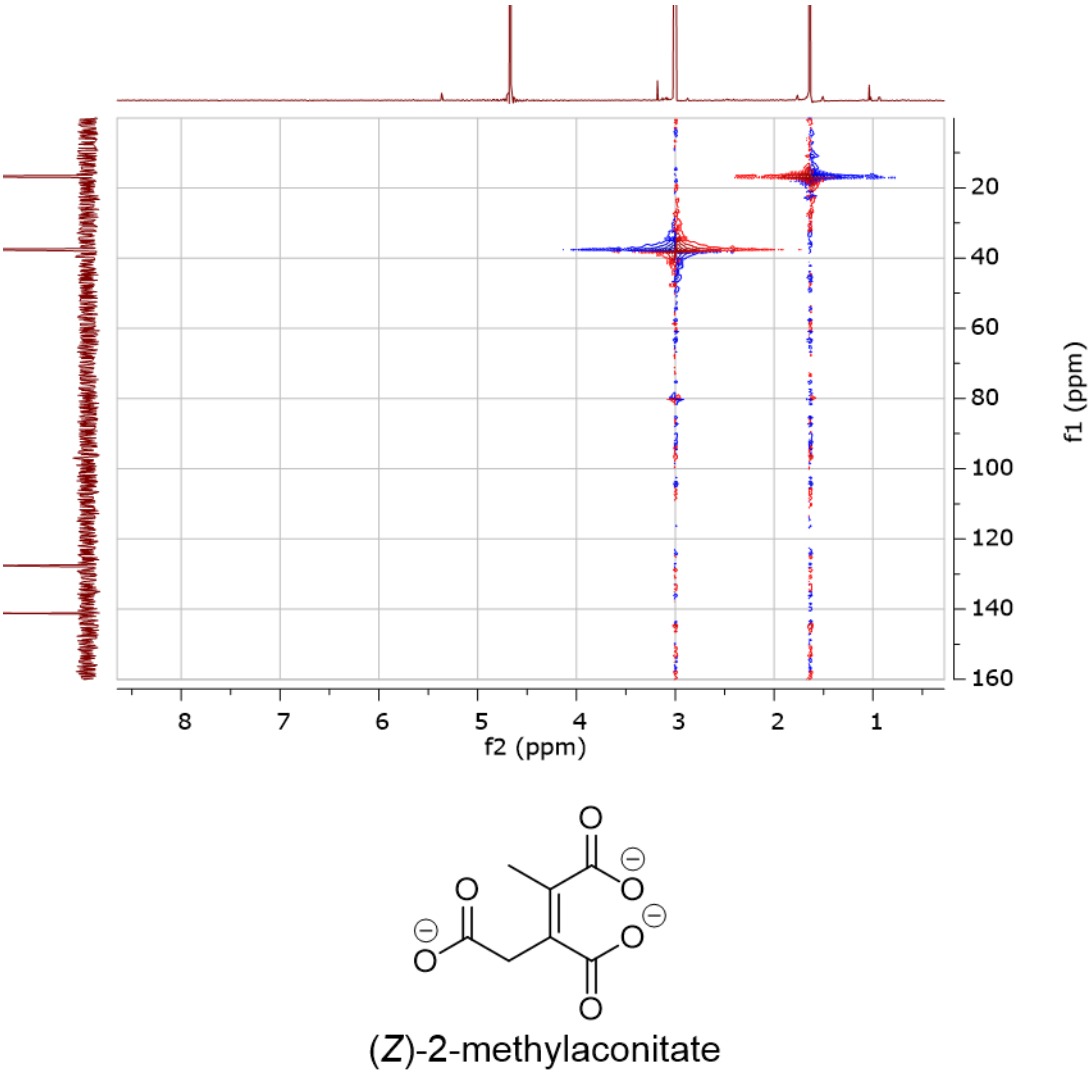


Figure 33. HSQC methylene group of Methylaconitate [500 MHz, D_2O].



CHAPTER III: INVESTIGATING THE CYTOCHROME P450 ENZYME (PKSS) FROM *BACILLUS AMYLOLIQUEFACIENS* IN THE BIOSYNTHESIS OF BACILLAENE

ABSTRACT

Soil bacteria are a wide area of interest due to their ability to produce an array of secondary metabolites, typically antibiotics. The *trans*-acyltransferase polyketide synthase product, bacillaene, is produced by several *Bacillus* species as a defense mechanism against other predatory bacteria and fungi. At this point in time there is an incomplete understanding of bacillaenes biosynthesis due to its complexity and instability. It is known that the polyketide synthase product, dihydrobacillaene will react with the cytochrome P450 enzyme, but the product is unknown. Knockout plasmids were designed to knockout *baeS* in *B. amyloliquefaciens* and *pksS* in *B. subtills* were generated to help determine the product of the reaction between dihydrobacillaene and PksS.

INTRODUCTION

Polyketide Synthases

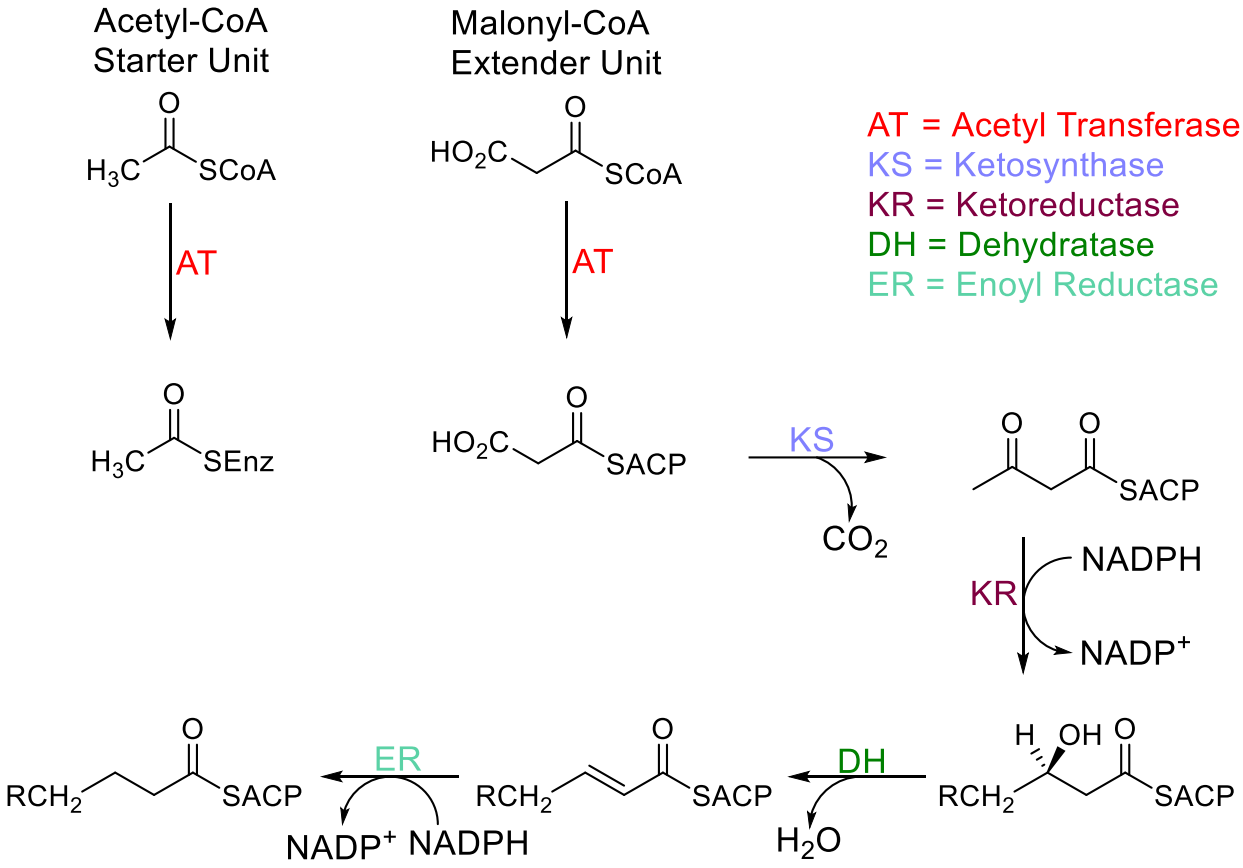
Secondary metabolites are natural products found in bacteria, plants, fungi, and mammals that are often used as a source of drug discovery.[64] These secondary metabolites typically offer their host organism an advantage in survival.[65] Polyketides are a class of natural products found in plants, bacteria, and fungi that are categorized based on their biosynthetic pathway.[27] The biosynthetic pathway of polyketides resembles that of fatty acids but makes up a more complex and diverse group of natural products.[66,67] Today, many drugs have been inspired by polyketides. These drugs include the antibiotic Erythromycin A, an anti-cancer compound neocarzinostain, the immunosuppressant rapamycin, and the cholesterol lowering agent lovastatin.[68] Drugs

derived from polyketides make up roughly 20 % of the medications on the market and around 20 billion dollars worldwide in yearly revenue.[25]

Polyketide synthases (PKSs) are multifunctional enzymes that are responsible for the biosynthesis of the polyketide synthase product. PKSs can be broken down into three subcategories: type I-III.[28] Type I PKSs are multifunctional enzymes with individual domains. These domains of type I PKSs are sorted in modules that will catalyze the elongation of the polyketide carbon chain.[69–71] Type II PKSs has individual monofunctional proteins that will typically lead to aromatic polyketides. Type III PKSs leads to the biosynthesis of flavonoids and stilbenes. The starter unit of the polyketide is typically an acetyl-coenzyme A (CoA) or propionyl-CoA. The polyketide chain will be elongated through a Claisen condensation between an acyl-ACP and malonyl-ACP. This condensation will lead to a decarboxylation, leading to the formation of an enolate. The negative charge on the carbon will then attack a secondary carbonyl carbon which will lead to the formation of a ketone group.[29]

The elongation of the polyketide chain happens as the polyketide synthase enzyme domains work together as an assembly line. Through different combinations of the enzymes, different polyketide chains can be produced (Figure 33). Each polyketide will begin with a starter unit. This process of elongation of the chain occurs at the acyl carrier protein (ACP).

Figure 34. Polyketide biosynthesis.

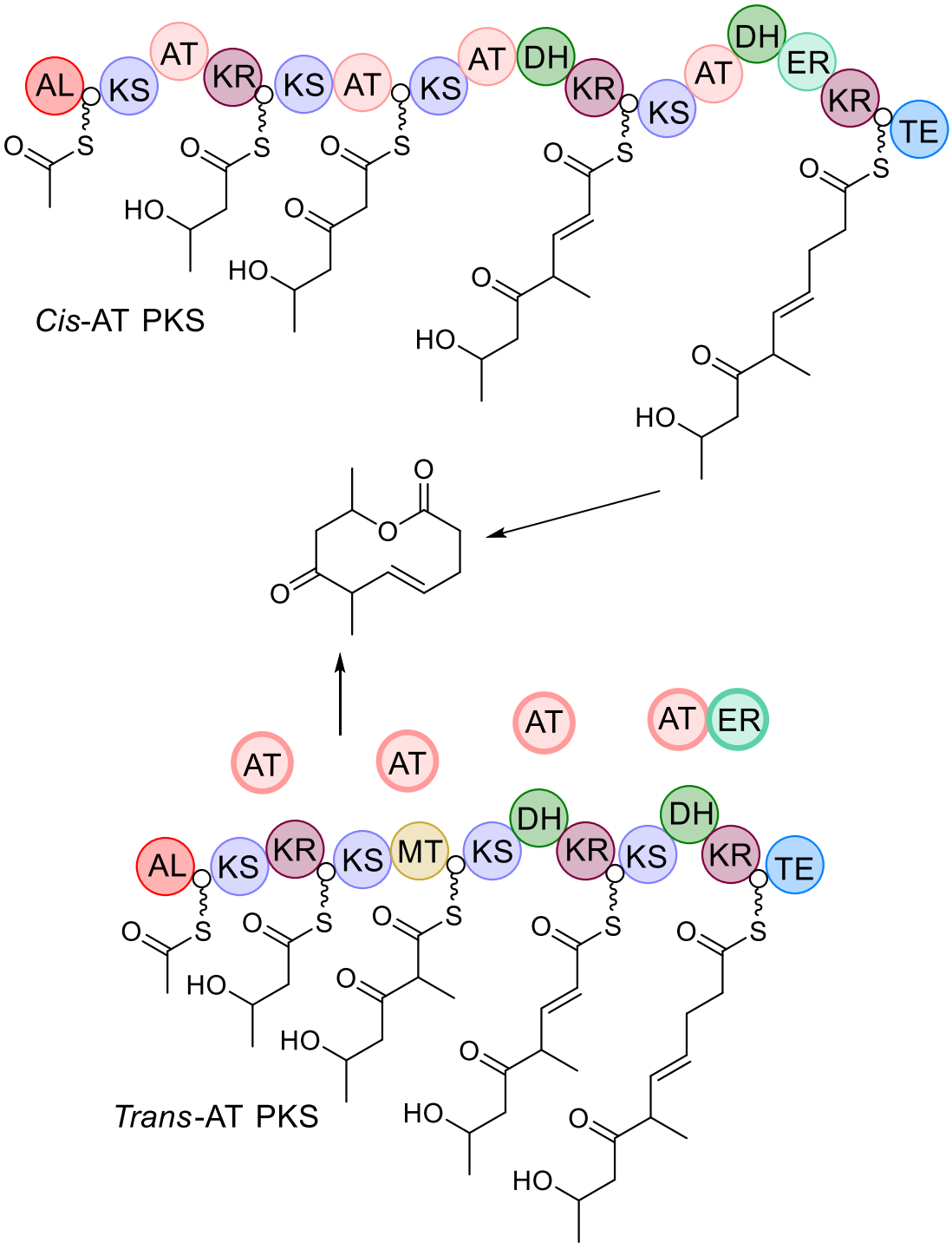


The polyketide will start with a starter unit and an extender unit. For type I polyketides the start unit typically consists of acetyl or propionyl-CoA. The extender unit can consist of different CoA's including methyl malonyl-CoA. These will be transferred to the acyl carrier protein (ACP) by the acyltransferase (AT). Once loaded, the ketosynthase (KS) will catalyze a decarboxylative condensation to form a new carbon-carbon bond and release carbon dioxide. During condensation, a new ketone group will be formed. Through the formation of chain, that ketone group can be modified to increase diversity. The ketone can be reduced to the alcohol by ketoreductase (KR). The alcohol can be reduced to an alkene with dehydratase (DH). The alkene can be reduced further by enoyl reductase (ER). The chain will be extended by cycles of Claisen condensations with any programmed modifications of the beta- position until the polyketide

chain is completed. Once the polyketide chain is complete, it will be released by the thioesterase to produce the polyketide synthase product.

The polyketide intermediate will be attached to the ACP at the phosphopantetheine arm by a thioester linkage.[67] Type I PKSs systems will pass the polyketide synthase intermediate between the ACP units and the next module of enzymes until the chain is complete. The condensation reaction is catalyzed by a ketosynthase (KS) to elongate the chain by adding the extending group and resulting in a ketone. The ketone can be modified further such as reduction to an alcohol via ketoreductase (KR), dehydration of the alcohol to an alkene via a dehydratase (DH) or reducing the alkene to an alkane via an enoyl reductase (ER). Once the polyketide chain has been completed, the polyketide will be released from the ACP via a thioesterase (TE).[72,73]

Figure 35. *Cis* and *trans*-acyltransferase polyketide synthase.



Type I polyketide synthase can be divided into two categories: *cis* and *trans*-acyltransferase polyketide synthase. *Cis*-acyltransferase polyketide synthase can be followed

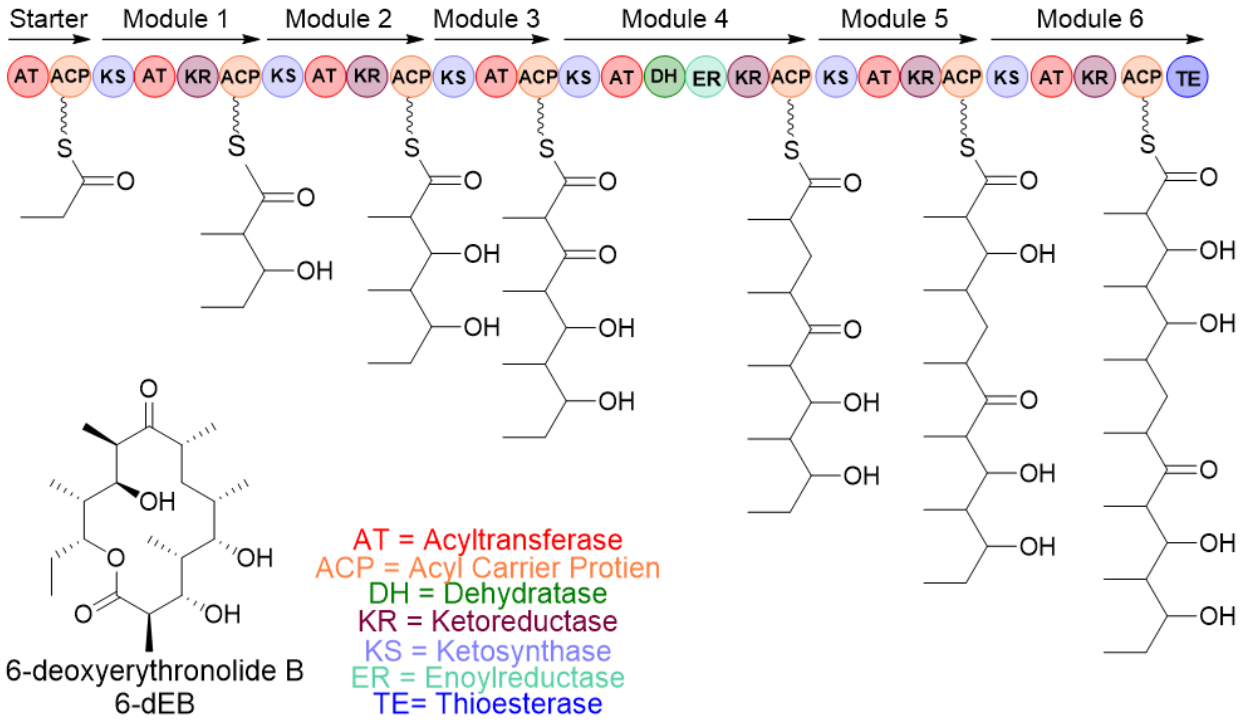
linearly. Looking through the modules, each enzyme corresponds to the elongation and modification of the polyketide chain. The methyl groups are added by using methylmalonyl-CoA as the extender unit. With *trans*-acyltransferase polyketide synthase, the acyltransferase enzymes are not a part of the mega enzyme complex, but an independent enzyme. In addition, other external enzymes, like an enoyl reductase shown above, can act on the polyketide chain that are not a part of the gene cluster, making the overall structure more difficult to predict.[35]

Type I polyketides can then be divided into two subgroups: *cis* and *trans*-acyltransferase (AT) Polyketide synthases (Figure 34). The *cis* and *trans*-acyltransferase seem to have evolved separately from one another.[30] *Cis*-acyltransferase polyketides will have the acyltransferase domain transfer new building blocks to the ACP domain for elongation. *Trans*-acyltransferase polyketides lack the AT domain. *Cis*-AT systems can be easier to correlate with the enzymatic modules for biosynthesis of the carbon chain and the final polyketide product. Each module of the *cis*-AT PKS can be analyzed by the AT sequences. The malonyl monomer that is selected, either malonyl or methylmalonyl, can be determined in these sequences. This determination makes predicting the structure of the final polyketide of *cis*-AT PKS more easily. *Trans*-AT systems have a lack of correlation between the enzymatic modules and the final polyketide product, along with a lack of the AT domain.[74] *Trans*-acyltransferase (*trans*-AT) polyketide synthases make up almost half of the known polyketides but due to their complexity are less understood then the *cis*-AT PKSs.[75]

Type I polyketide synthase systems can co-occur with nonribosomal peptide synthetases (NRPS).[76] Nonribosomal peptide synthetases exists as a mega-enzyme complex that will synthesize amino acid monomers.[77] These monomers can include unusual amino acids, such as D-amino acids, β-amino acids, and α-keto/hydroxy acids.[78] The domains for NRPSs will contain

an adenylation, thiolation and condensation domain.[79] The adenylation domain will create an amino acid thioester. The thiolation domains consists of the peptidyl carrier protein. The condensation domain will attach the amino acid to the peptide chain.[80] Co-occurring PKS/NRPSs systems are made possible by substituting if the acyltransferase or adenylation domain is being used in the module of the biosynthesis.[79]

Figure 36. 6-deoxyerythronolide B (6-dEB) biosynthesis.



6-deoxyerythronolide B is the polyketide synthase product and precursor to erythromycin A. The biosynthesis starts with a propionyl-CoA starter unit. From there methyl-malonyl-CoA units will be used to extend the polyketide chain. The chain will go through six modules to elongate the polyketide chain and create the final product. The chain will go through multiple rounds of extension with the extender unit and modify the ketone group with the KR, DH, and ER. Once released by the thioesterase, the chain will cyclize to form 6-deoxyerythronolide B.

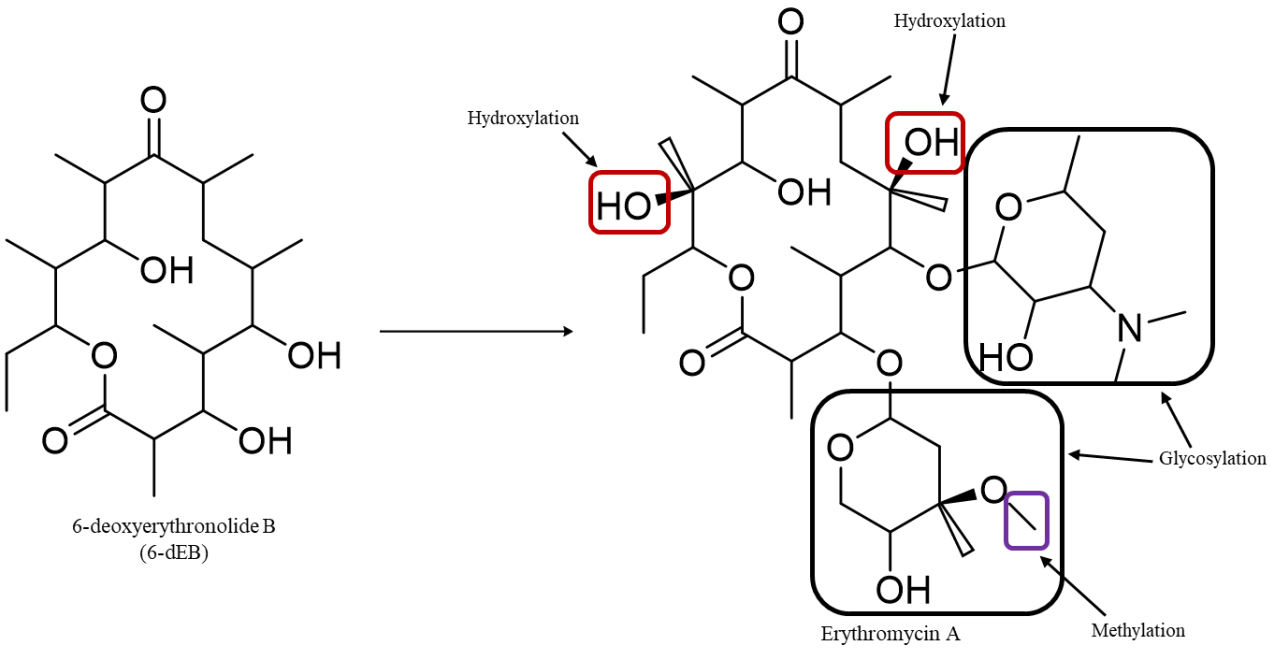
Erythromycin A is considered a model antibiotic for *cis*-AT polyketides produced by the bacteria *Saccharopolyspora erythraea*. The polyketide synthase product of the *eryX* gene cluster is 6-deoxyerythronolide B.[31] The gene cluster encodes for the three 6-deoxyerythronolide B synthases, DEBS, proteins that are responsible for the biosynthesis of the polyketide (Figure 35).[81,82] The *eryAI* gene will encode for DEBS 1, the *eryAII* gene will encode for DEBS 2, and the *eryAIII* gene encodes for DEBS 3.

These three proteins make up the six modules in the biosynthesis of the polyketide. DEBS 1 contains the loading, module 1, and module 2. DEBS 2 contains module 3 and module 4. DEBS 3 contains module 5, module 6, and the release.[83] The initial polyketide synthesis begins with a propionyl-CoA starter unit that will be loaded onto the ACP domain. From there, methyl malonyl-CoA units will be added as the extending unit.[84] The biosynthesis will go through the six modules that will consist of different ketosynthase, ketoreductase, dehydratase, and enoylreductase domains until reaching the thioesterase. At the thioesterase domain, the polyketide synthase product will be released by the TE via cyclization to form 6-deoxyerythronide B. This product is the precursor to erythromycin which will undergo post-polyketide modifications after being released from the synthase.[84,85]

Post Polyketide Synthase Modifications

Once the polyketide synthase product is released from the thioesterase, the chain can continue to go through post polyketide synthase modifications.[86] These modifications can increase the structural diversity and biological activity of polyketide.[87,88] Post-polyketide synthase enzymes will add to the backbone of the polyketide chain.[89] Like with erythromycin A, many post polyketide modifications can include hydroxylations, glycosylations and methylations (Figure 36).[32]

Figure 37. Post polyketide synthase modifications.

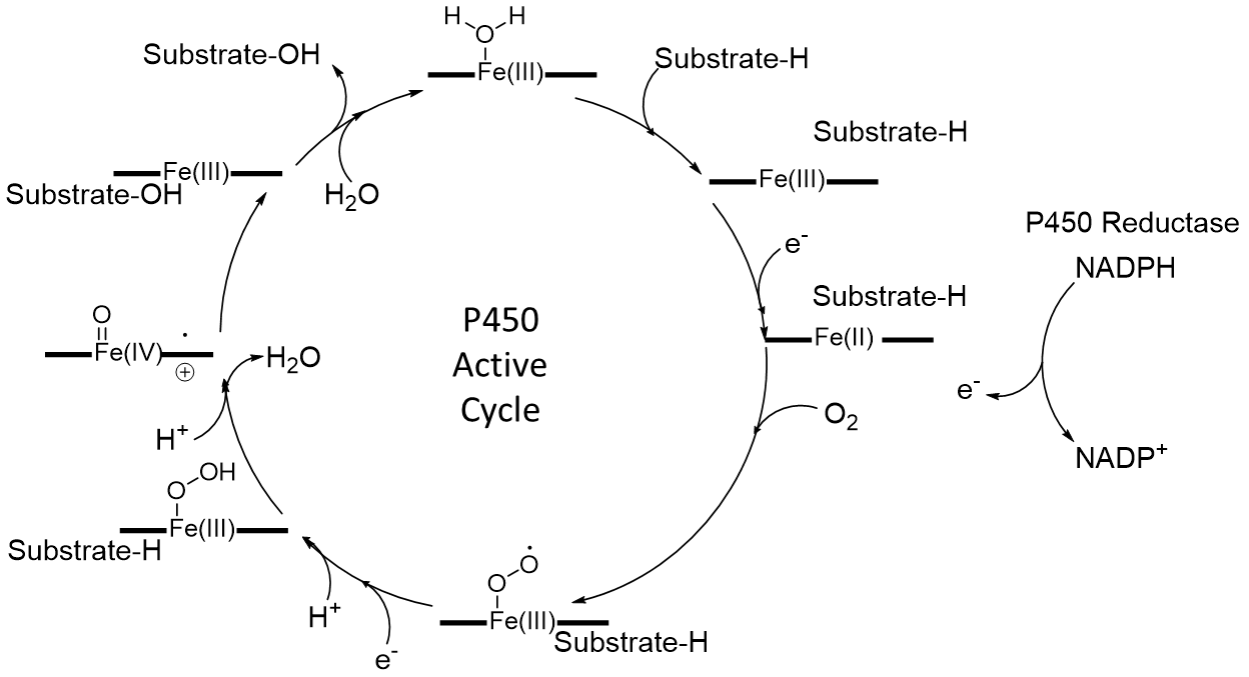


To increase the biodiversity and bioactivity of the polyketides, post polyketide synthase modifications can occur. 6-deoxyerythronolide B, the polyketide synthase product, will go through several rounds of modifications to produce erythromycin A. These modifications include hydroxylations, glycosylations, and methylations.

Cytochrome P450 enzymes make up a superfamily of enzymes that account for substantial amounts of diversity and many diverse types of reactions for oxidation type modifications.[90] Initially discovered in the 1960's, these enzymes account for roughly 75 % of the metabolism of drugs.[91–93] These enzymes prevent toxic drug effects by oxidization of the drug and alternatively can activate other drugs by oxidation.[91] The human cytochrome P450 enzyme CA4 found in the liver and intestine is responsible for over 50 % of drug metabolism.[94] Bacterial cytochrome P450 enzymes have become crucial models in understanding how the superfamily of enzymes function for drug metabolism.

Cytochrome P450 enzymes go through a catalytic mechanism (Figure 37). These enzymes are heme-comprising enzymes that require a redox supporting enzyme from a cytochrome P450 reductase and hydride donor in the form of NAD(P)H.[95]

Figure 38. P450 catalytic cycle.



The substrate will enter the active site and replace water. Electrons will be supplied to the active cycle by the P450 reductase. Cytochrome P450 enzymes need a P450 reductase that supplies the electrons. The reductase is typically paired with NAD(P)H. After several rounds of reduction and protonation, the substrate will be oxidized and released from the active site.

In the first step, the substrate will enter the active site and replaces the water bound to the heme group. Afterwards, electron transfer can occur from a reducing partner, a cytochrome P450 reductase, to reduce the Fe^{III} to the Fe^{II} complex. These reducing partners gain their electrons from a cofactor, like NAD(P)H. O_2 will bind to the complex to make an oxyferrous complex. Another reduction and protonation take place creating Fe^{IV} oxo heme radical species that will react with the substrate. In hydroxylations this Fe oxo species abstracts a hydrogen atom from

the substrate to form a radical substrate and then the new alcohol group on the Fe will hydroxylate the product. Water reenters the active site to protonate the bound oxygen species. The enzyme is restored to the initial state as the product is released and replaced by water.[96]

Cytochrome P450 enzymes can do simple and complex modifications such as hydroxylation, dealkylation, and epoxidation reactions.[38] Due to the chiral nature of the active sites of these enzymes, they can regulate the regioselectivity of the products.[90,97] The enzyme is commonly known for hydroxylation reactions.[98] On top of hydroxylation reactions, the enzymes are capable of more complex reactions. The superfamily of enzymes is also capable of making thioethers, aldehydes, dehalogenation, desaturation, and carbon-carbon coupling.[99] The competition between either hydroxylated or desaturated will depend on the radial species, Compound I (Cpd I), which happens when the iron (IV)-oxo heme reacts with the substrate.[100]

Knockout Mutations

Polyketides and other metabolites are studied in various ways. The biosynthetic pathways can be studied by cloning each gene and characterizing the function of its encoded enzyme *in vitro* with the substrate.[101] The ability to elucidate the function of a gene has become a vital part of biochemical research.[102] The gene can be cloned into a plasmid, and the enzyme purified to be used in further reactions. The substrate and enzyme can react and analyzed by different analytical techniques. Additionally, genes can be knocked out in the model organism to help elucidate its physiological role in the organism. Knocking out a gene can consist of removing the gene entirely from the genome or creating a frameshift so that it cannot be transcribed correctly. The knockout mutation can be studied to identify the function of the gene in the organism when it lacks that function. Another method is complimenting the mutant by reintroducing the gene and restore its function.

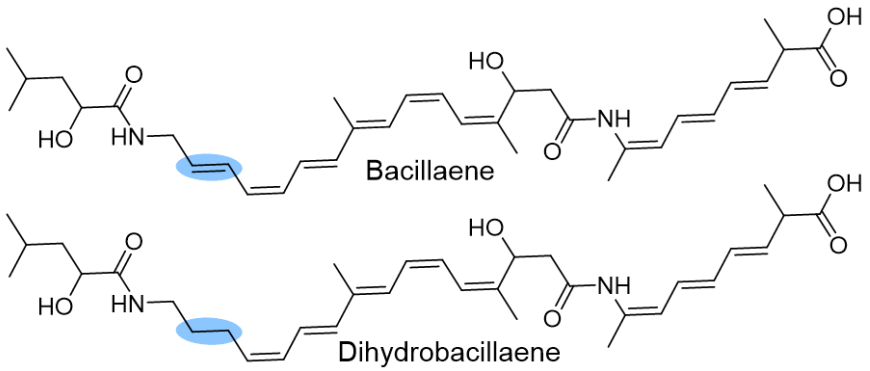
Knockout mutations can be made by DNA recombination events, or genome editing, like CRISPR/Cas9.[103,104] Double crossover knockout mutations consist of removing the gene of interest from the genome. A new gene will be inserted into the genome in place of the gene of interest. For instance, the new gene to be inserted into the genome can be an antibiotic resistance gene. This would cause only bacteria with the knockout mutation to grow on antibiotic plates compared to the wild type strains. This process will eliminate any potential remaining gene function.[105] With single crossover knockout mutations, like Campbell-like insertions, the gene of interest is not removed entirely from the genome.[103] Instead, a fragment of the gene will be integrated into the genome along with part of the plasmid. This causes a disruption in the gene so that it cannot be translated. CRISPR/Cas9 gene editing will lead to a frame shift in the genome.

Clustered regularly interspaced short palindromic repeat (CRISPR)-Cas9 have become widely studied systems from the RNA-guided immune systems of archaea and bacteria for gene editing.[106] CRISPR was initially discovered as an adaptation used by bacteria to fight reinfection from bacteriophages. The system is composed of the target-specific single guide RNA (sgRNA) and the Cas9 protein. The Cas9 protein will target the desired area for cleavage to create a double-stranded break. That break will go through DNA repair to fix the damaged DNA, and the sequence fidelity is lost. This fidelity loss is caused by random insertion and deletion of bases caused by non-homologous end-joining (NHEJ), or DNA being added to the cut site by homology-directed repair (HDR).[107,108] The target specificity of the protein depends on the spacer sequence. The sgRNA will be designed with roughly 20 base pairs that will be complementary to the gene that is desired to knockout. That target sequence will precede the protospacer adjacent motif (PAM).[101]

Bacillaene

The *trans*-acyltransferase polyketide synthase bacillaene, isolated from *Bacillus subtilis*, *Bacillus veleznesis*, and *Bacillus amyloliquefaciens*, was first isolated in 1995 by Mayerl and coworkers, but the structure was not solved until 2006 by Clardy and coworkers.[48,49] Only the structure of bacillaene and dihydrobacillaene were elucidated, but other analogs were noted, such as double bond isomers, dehydrated derivatives, and hydroxylated derivatives (Figure 38).

Figure 39. Structures of bacillaene and dihydrobacillaene.

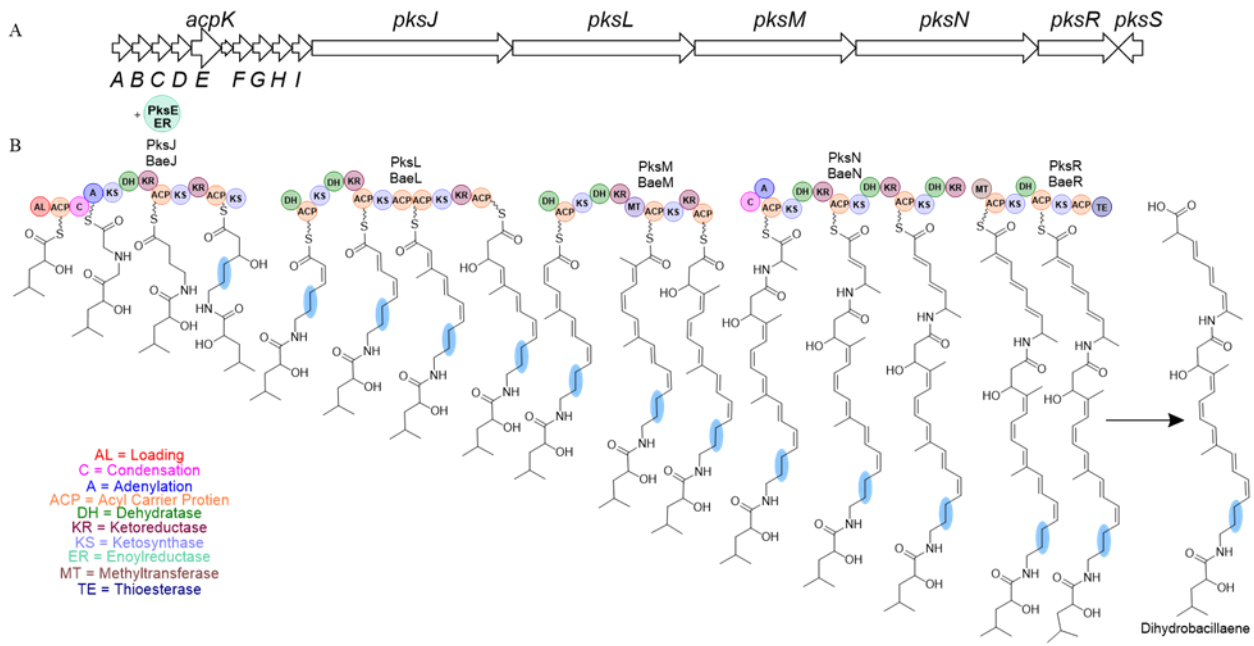


The structures of bacillaene and dihydrobacillaene were elucidated by the Clardy lab in 2006. Other derivatives of the bacillaene family have been noted, like hydroxylated dihydrobacillaene, but those structures were not elucidated.

Today, the absolute configurations of the compounds are still unknown.[109] Studies of these compounds have been incomplete due to their instability at high temperatures and light sensitivity.[109,110] Bacillaene was originally considered the product of the *pksX* gene cluster from *B. subtilis*, which encodes a *trans*-AT, type I polyketide synthases/nonribosomal peptide synthases [49] The biosynthesis includes several unusual features including β-branching, a *trans*-enoylreductase, and a cytochrome P450 enzyme (Figure 39). The β-branch is installed by the HMG-CoA gene cassette that includes *pksF-pksI* and *acpK*.[111,112] The *trans*-enoylreductase, PksE, was identified to reduce the enoyl substrate between the 14'-15' carbons of the

biosynthetic intermediates constructed by PksJ domain during the biosynthesis of dihydrobacillaene. Reddick and coworkers discovered that the cytochrome P450 enzyme, PksS, from *B. subtilis* would react with dihydrobacillaene and not bacillaene to give an unknown product.[113] Piel then went on to confirm that dihydrobacillaene, which lacks a double bond found in bacillaene, is the immediate product of the polyketide synthases.[114] Walsh demonstrated that the *trans*-enoylreductase, PksE, reduces the double bond to produce the biosynthetic intermediate that will go on to produce dihydrobacillaene.[52]

Figure 40. Biosynthesis of bacillaene.

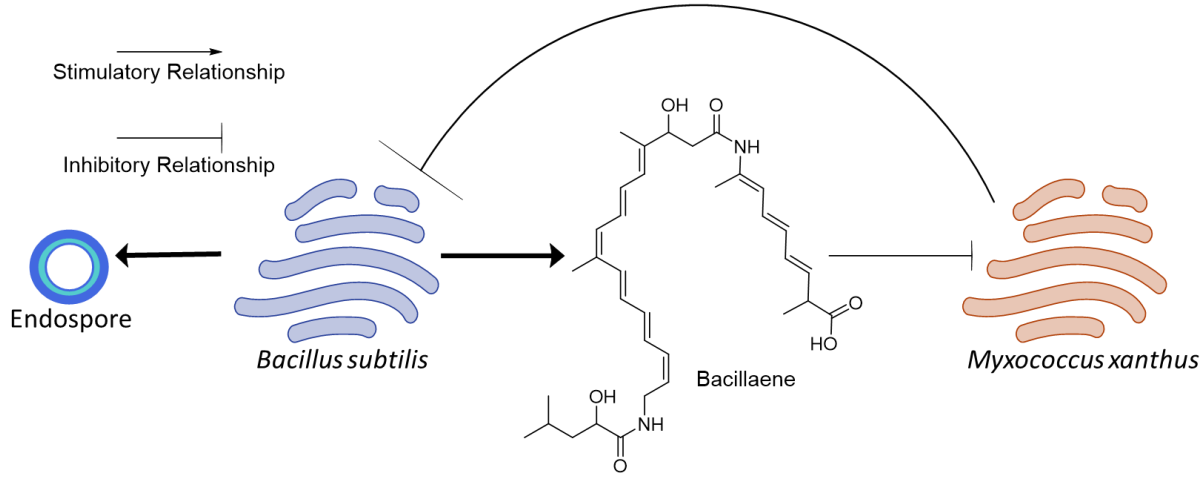


A) The *pksX* gene cluster. The *pksX* gene cluster from *B. subtilis* for biosynthesis of the polyketide bacillaene. The *pksX* genome makes up roughly 2 % of the *B. subtilis* genome. B) The biosynthetic pathway for bacillaene. The *pksX* gene cluster is from *B. subtilis*, the *baeX* gene cluster is from *B. amyloliquefaciens*. The biosynthesis includes several unusual features including β-branching, a *trans*-enoylreductase, and a cytochrome P450 enzyme. The β-branch is installed by the HMG-CoA gene cassette that includes *pksF-pksI* and *acpK*. The *trans*-

enoylreductase, PksE, was identified to reduce the enoyl substrate between the 14'-15' carbons of the PksJ domain during the biosynthesis of dihydrobacillaene. This position is highlighted in blue. The polyketide synthase product, which is released by the thioesterase, is dihydrobacillaene. PksS will react with dihydrobacillaene to give an unknown product.

Non-ribosomal peptide synthases (NRPSs) are multimodular enzymes that deliver amino acids acyl carrier protein and make polypeptides.[115] NRPSs work to assemble amino acids to peptides using a thiotemplated multidomain system similar to the layout of PKSs. Hybrid PKS-NRPSs are multimodular enzymes that will work together in an assembly line to produce natural products having both polyketide and amino acid-derived structures. Bacillaene, a hybrid PKS-NRPS, will incorporate a glycine in the BaeJ/PksJ module and an alanine in the BaeN/PksN module.[116,117]

Figure 41. Bacillaene part in sporulation of *Bacillus* species.



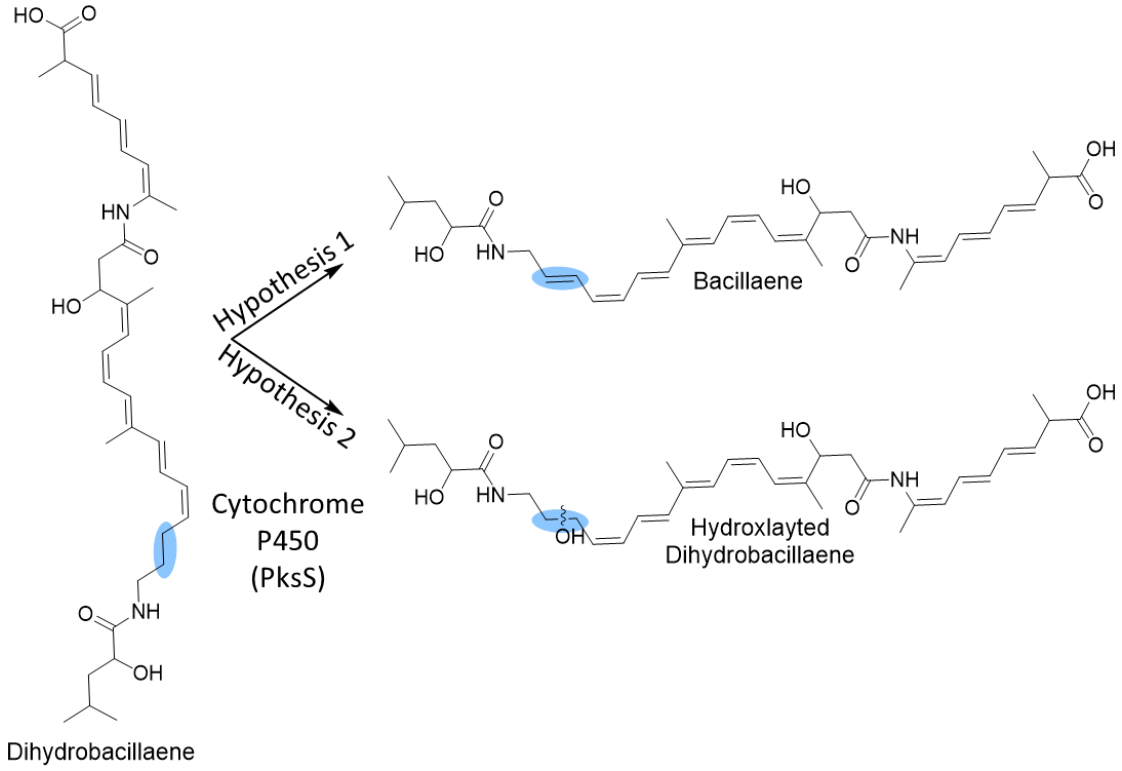
When in the presence of a predatory bacteria and fungi, *Bacillus* will start the process of sporulation. To protect itself from the predators during sporulation, the bacteria will produce bacillaene in the biofilm. Producing bacillaene in the biofilm will protect the bacteria from the predators and give the bacteria time to create endospores.[41]

Bacillaene has remained an interest in studies due to its antibiotic, antifungal, fungistatic and bacteriostatic activity, as well as the ability to inhibit prokaryotic protein synthases.[40–42,48,64,118] Bacillaene was also found to play a role in the sporulation and biofilm formation in *Bacillus* species as part of their defense mechanisms.[110] Bacillaene will inhibit predatory Gram-positive bacteria, Gram-negative bacteria, and fungi so that the *Bacillus* bacteria has enough time to produce spores (Figure 40).[119]

Due to their complexity and instability, the biosynthesis is still not fully understood although glycosylated bacillaenes have been noted to have higher stability.[110,120,121] Research has shown that bacillaenes are a pseudo poison to their host cell. When production of the bacillaenes is increased to a poisonous state, detoxication mechanisms of glycosylation and other degradation will be enforced.[110] Post-polyketide synthase modifications done by enzymes like cytochrome P450's adds to the diversity of this class of secondary metabolites.

Our goal was to determine the product of the reaction between PksS and dihydrobacillaene. We hypothesized that the reaction between PksS and dihydrobacillaene would produce either bacillaene or a hydroxylated dihydrobacillaene (Figure 41). The structure of bacillaene is a known compound from the *pksX* gene cluster. Hydroxylated versions of the compounds have been noted but the structures have not been elucidated. Hydroxylated versions might exist outside the proposed hydroxylation at the 14' carbon or 15' carbon.

Figure 42. Predicted products between the reaction of dihydrobacillaene and PksS.



The cytochrome P450 enzyme, PksS, will react with the substrate dihydrobacillaene to produce an unknown product. Cytochrome P450 enzymes are a superfamily of enzymes and are capable of unique chemical reactions. PksS with dihydrobacillaene could potentially produce bacillaene or hydroxylated dihydrobacillaene. Both compounds have been observed by mass, but their structures have not been fully elucidated.

Cytochrome P450 enzymes are commonly known to catalyze hydroxylation's of saturated bonds, but desaturations are also known, and these post-translational modification transformations increase the diversity of polyketides.[100] Since bacillaene and hydroxylated dihydrobacillaene are both known compounds in this complex mixture they are the logical possibilities for the product of the reaction.

MATERIALS AND METHODS

Cloning of the *pksS* Gene

Standard methods were used for restriction endonuclease digestions, ligation, and chemical bacterial transformation. Primers 5'- cgagga**ccatgg**aaatggaaaaattg-3' and 5'-ggggcgttttga**ctcgag**ttttgaaagtgaaacagg-3' were designed to amplify the region of the *pksS* gene in *Bacillus subtilis* 168. Once received, the primers were rehydrated with water to be used to amplify the DNA. To isolate the genomic DNA template, starter cultures of BS168 were grown in LB and left to grow overnight. The starter culture was then used with the Wizard kit to isolate the DNA, and the DNA was left to rehydrate overnight at 4 °C. Polymerase chain reaction (PCR) samples were set up using water, the upstream and downstream primers, Phusion High-Fidelity DNA polymerase, and DNA isolated from *B. subtilis* 168. After PCR, the samples were loaded onto an agarose gel for confirmation of the correct DNA band size. The DNA was electrophoresed at 100 volts, and 90 milliamps for one hour. The gel was then analyzed for the correct DNA band size compared to a 1 kbp DNA standard. The DNA was purified using a PCR purification kit from Qiagen and the concentration evaluated on a Thermo scientific NanoDrop One. The DNA was then transformed into TOP10 *E. coli* and prepared for sequencing.

Purification of PksS

E. coli strain BL21-Star replicating plasmid *pksS*-pET-28a was streaked on an LB agar plate containing 50 μg/mL kanamycin and left to grow overnight at 37 °C. The following day, a starter culture was made composing of 5 mL of terrific broth, 50 μg/mL kanamycin, and one colony from the above streaked plate of *pksS*-pET-28a. The starter culture was shaken at 220 rpm at 37 °C overnight. Terrific Broth consists of 24 g yeast extract, 20 g tryptone, 4 mL glycerol, autoclaved in 900 mL of water and then added potassium phosphate buffer. The

potassium phosphate buffer consisted of 23.135 g KH_2PO_4 g and 125.410 g K_2HPO_4 in 1000 mL of water. The next day, 2 mL of the starter culture was combined with 1 L of terrific broth and 50 μg/mL kanamycin. This was left to shake until the culture had grown for two days at 20 °C and 200 rpm and was then induced with isopropyl B-D-1-thioglactopyranoside (IPTG) (1 mM final concentration). The culture was left to grow overnight at 20 °C and 70 rpm. The culture was centrifuged at 7480 g's for 30 minutes, and the pellets were then suspended into the binding buffer, which consists of 5 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, and sonicated for 3 minutes in 30-second intervals to break the cell membranes. The mixture was centrifuged again at 16000*g* for 30 minutes, and the supernatant was filtered and then loaded onto a nickel nitrilotriacetic acid (Ni-NTA) affinity column (2 mL column bed) that was pre-equilibrated with binding buffer. After the supernatant was loaded onto the column, it was followed by 20 mL binding buffer, then 12 mL wash buffer, which consisted of 60 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, and 6 mL elute buffer, which consisted of 1 M imidazole, 0.5 M NaCl, and 20 mM Tris-HCl. At the elute buffer stage, 1 mL fractions were collected and analyzed with the Bradford assay reagent for protein. The fractions that contained protein were combined so that the buffer was exchanged by either Snakeskin dialysis tubing (22 mm x 7 cm, 7,000 MWCO) and left overnight in 4 L dialysis buffer or using fast protein liquid chromatography (FPLC) desalting column. Both methods used dialysis buffer consisting of 25 mM Tris-HCl (pH 8). The next day, the protein was collected from the dialysis. Glycerol was mixed with the protein to make a 10 % glycerol stock solution and stored in a -80 °C freezer. The dialysis was replaced by FPLC using a desalting column (HiPrep 26/10 Desalting column) where the protein fractions collected from the column would be collected in a syringe and inserted into the FPLC for desalting. After the fractions with protein were eluted, the protein was concentrated by being

centrifuged in a protein concentrator (Vivaspin) in a 14.5 rotor at 4 °C at 8,000*g* for 10 minutes. Glycerol was added to the protein to make a 10 % glycerol stock and then stored in the -80 °C freezer.

LC-MS Methods

The previous HPLC method used by Reddick and coworkers for identification of bacillaene compounds was optimized to be used for a LC-MS method.[50] The previous HPLC used a 5 μM Prevail C-18 column method that began at 95:5 1 % acetic acid: acetonitrile, with acetonitrile increasing to 95 % over the course of 45 minutes at a linear gradient of 0.2 mL/min flow rate. These experiments utilized a Thermo Fisher Scientific Q Exactive Plus that was equipped with electrospray ionization. The method was optimized to the UPLC-MS method beginning at 85:15 water: acetonitrile to 95 % acetonitrile at a linear gradient of 0.2 mL/min flow rate.

The co-culture extracts were analyzed by high-resolution electrospray ionization mass spectrometry with a Thermo Fisher Scientific Q Exactive Plus. This was coupled to an ultra-high performance liquid chromatography system with a BEH C_{18} column (2.1 mm x 50 mm, 1.7 um) with a flow rate of 0.3 ml/min. The gradient started at 15 % acetonitrile with 0.1 % formic acid and increased linearly to 100 % acetonitrile over 8 minutes. The method then re-equilibrated for the injection of the next sample.

Media Optimization to produce the Crude Natural Product Extract

Landy media was used to obtain the natural product extract containing dihydrobacillaene. Landy media (described below) was chosen as it is a minimal media. This minimal media will help stress the cells to the point of sporulation and production of dihydrobacillaene compared to enriched media. The media was optimized based on the previous Landy media used by Reddick

and coworkers. Starter cultures of *B. amyloliquefaciens* (strain ATCC 39374) were grown composing of 5 mL LB and 1 colony of the bacteria. The starter culture was added to the 50 mL culture of Landy media for production of the natural product extract. The cells were grown for a total of 28 hours instead of the previous 24 hours growth period. The optimized Landy media consists of 0.25 g Glutamate and 1 % Landy salts, autoclaved, then filtered 1 % $MgSO_4$ and glucose in 50 mL total volume (pH 6.5). The 10 x Landy salts consisted of 2.5 g KCl, 5 g KH_2PO_4, 0.75 g $FeSO_4$, 0.025 g $MnSO_4$, 0.0008 g $CuSO_4$ in 500 mL water.

Co-Culturing *Bacillus amyloliquefaciens*

In an attempt to increase the production of dihydrobacillaene, co-culturing of *B. amyloliquefaciens* was done with the fungus *Aspergillus fumigatus*. The fungi for co-culturing were provided by the Nicholas Oberlies lab at UNCG. Co-culturing of the bacteria and fungi should create a "war zone" between the two species where they are both fighting for nutrients. This should lead to the bacteria producing dihydrobacillaene in order to protect itself from the fungi. On potato dextrose agar or Landy media agar plates, *A. fumigatus* was spotted on one side, and *B. amyloliquefaciens* was streaked on the other side. The cultures grew at 23 °C, 30 °C, and 37 °C over the span 72 hours, 96 hours, and 120 hours. Three co-culture plates were grown along with monoculture plates of the bacteria and fungi. To extract the dihydrobacillaene, methanol was poured over the plates to kill the bacteria and fungi. The plates were sliced up into smaller pieces and put into a flask. Acetone was added to cover the agar pieces and left to sit for 2-3 hours. The agar acetone mixture was filtered with a Buchner funnel with additional acetone. The filtered extract was concentrated under reduced pressure. The extracts were then analyzed on LC-MS.

Production of the Crude National Product Extract

Production of the dihydrobacillaene natural product crude extract was followed using similar methods as Reddick and coworkers.[50,122] Starter cultures of *B. amyloliquefaciens* ATCC 39374 were grown in 5 mL of LB overnight. Of the overnight starter cultures, 100 µL were then used to inoculate 50 mL cultures of Landy Media and left to grow for 28 hours in the dark at 37 °C. The bacillaene crude extract was extracted with 0.5 mL of extract and 1 mL of MeOH. The mixture was vortexed for one minute and then centrifuged. The methanol was then dried off. The extract was stored in the freezer or used immediately for reactions.

Extraction of Crude Cytochrome P450 Reductase

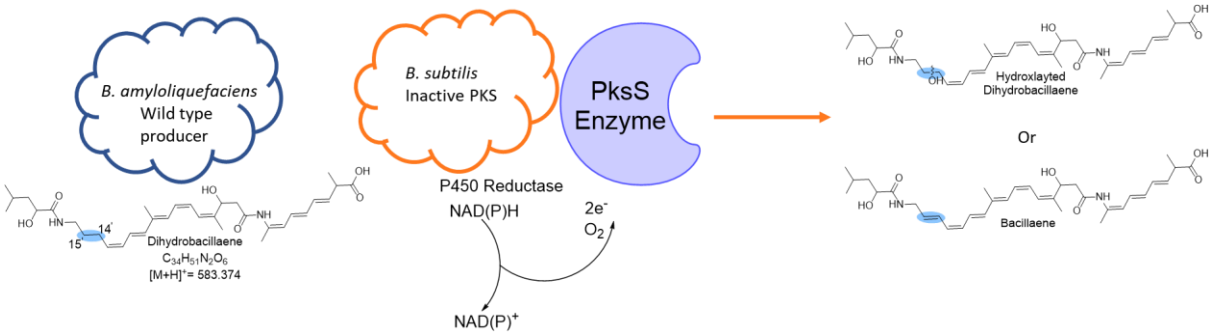
The cytochrome P450 reductase was obtained and used in unpurified form from a crude extract of *B. subtilis* 168. An overnight starter culture of *B. subtilis* 168 was used to inoculate 500 mL of LB and left to shake for two days at 37 °C. The culture was centrifuged (8,500*g* for 30 minutes) and the pellet was resuspended in 10 mL of 50 mM phosphate buffer (pH 8.0) and lysed on ice with lysozyme. The lysate was centrifuged (11,500*g* for 30 minutes) and the supernatant was filtered through a 0.45 µm membrane. Glycerol was added to the filtrate and was stored in -80 °C freezer without further modification for further experiments.

Initial LC-MS Reactions

Initial reaction mixtures were analyzed on reaction mixtures containing the purified PksS enzyme, crude reductase extract, and the crude natural product extract. Varying amounts of PksS, crude natural product extract, crude cytochrome P450 reductase, and NAD(P)H were analyzed via LC-MS. Phosphate buffer (50 mM, pH 8.0, 500 µL) was added to the crude natural product extract before addition to the reactions. Various amounts of each component of the reaction were tested for the optimal reaction conditions. Each reaction was done in pairs, one

with NADH and one with NADPH. Broad ranges of the substrates were used in attempts to optimize the reaction conditions.

Figure 43. Reaction components to determine the product between dihydrobacillaene and PksS.



There are four components to the reaction to determine the product between dihydrobacillaene and the enzyme, PksS. The crude natural product extract from *B. amyloliquefaciens* will produce dihydrobacillaene. The crude cytochrome P450 reductase from *B. subtilis* will give the P450 reductase. The cofactor to the P450 reductase will be either NADH or NADPH. The last component will be the purified cytochrome P450 enzyme, PksS, from *B. subtilis*.

The reactions were incubated at room temperature for one hour and then filtered. An internal standard, Leucine Enkephalin, was used to compare peak areas of dihydrobacillaene, bacillaene, and hydroxylated dihydrobacillaene. A stock concentration of Leucine Enkephalin at 0.4 µg/mL was used to add to reactions to bring it down to a concentration of 0.006 µg/mL.

Bacillaene was determined to be a product in the Heated electro spray ionization (HESI) source of the mass spectrometer having no accurate mass and aligning with the chromatogram peak for the hydroxylated dihydrobacillaene. Different LC methods and mass spectrometry settings were used to try to separate bacillaene from hydroxylated dihydrobacillaene to no avail

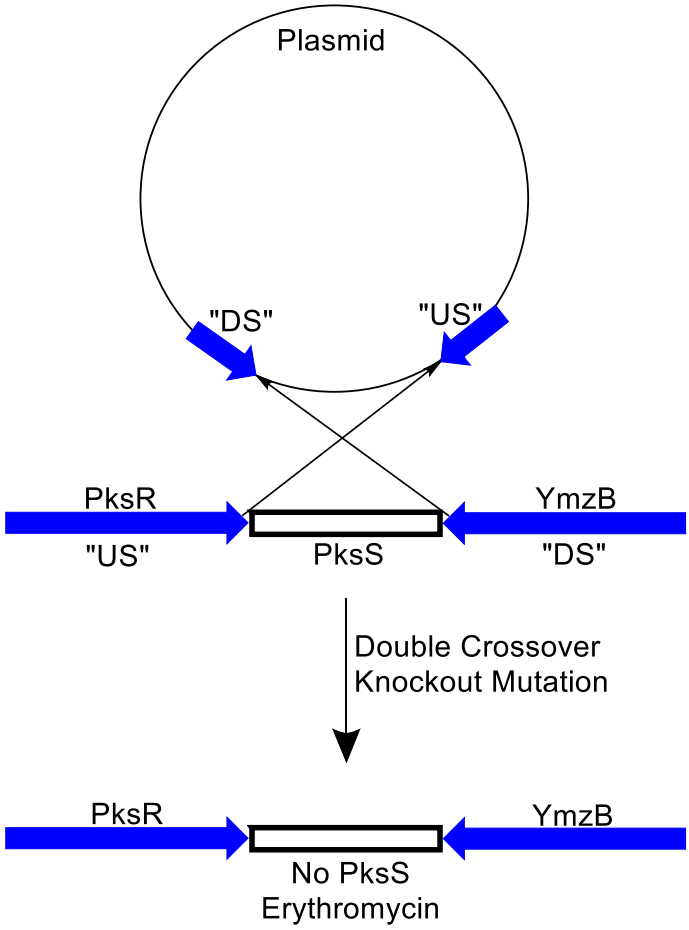
as they kept co-eluting. Hydroxylated dihydrobacillaene and bacillaene appear to be inseparable, perhaps because bacillaene is formed from the hydroxylated dihydrobacillaene in the mass spectrum source. This led to our conclusion that dihydrobacillaene with PksS is producing hydroxylated dihydrobacillaene. The reaction mixtures that included every component: natural product extract, NADH, cytochrome P450 reductase extract, and PksS, resulted in the new peak at around 5.14 minutes with a m/z 599.3990, demonstrating that hydroxylated dihydrobacillaene is the immediate product of the reaction between PksS and dihydrobacillaene. The relative abundance of hydroxylated dihydrobacillaene was found compared to a standard of leucine enkephalin.

Knockout Mutations of *B. amyloliquefaciens* and *B. subtilis*

We designed mutants lacking the *baeS/pksS* genes to eliminate PksS reaction products from the natural product extract, and also to ensure that the reductase extract contained no PksS/BaeS protein. This would ensure that there was no contamination of the P450's product before the purified enzyme was added to the reaction. This work was done in collaboration with Dr. Eric Josephs lab at the Joint School of Nanoscience and Nanoengineering. *ΔbaeS* mutants were made to eliminate the gene *baeS* from the natural product extract from *B. amyloliquefaciens* (ATCC strain 39374). We have substantial evidence that ATCC strain 39374 and 39320 belong to the *B. amyloliquefaciens* species but are currently classified by the ATCC as *B. subtilis*. Knockout plasmids to make mutants of *ΔpksS* were made to eliminate the gene *pksS* from the reductase extract, *B. subtilis* 168. These knockout mutants will be utilized in order to ensure that enzyme reactions contained the P450 reaction product only if the purified P450 enzyme was added. Knockout plasmids to make mutants were made, the PksS crude reductase was produced from *Bacillus subtilis* 168 *ΔpksS* mutant extract and the crude natural product extract was

produced from *Bacillus amyloliquefaciens ΔbaeS* mutant. Although *B. subtilis* strain 168 has the *pksX* gene cluster, this strain lacks *sfp* and it is not capable of producing the polyketides. Even without the *sfp* gene, the other genes responsible for the polyketides are still functional and can cause contamination of the product before addition of the purified enzyme. These knockouts were used to simplify LC-MS chromatograms to produce a clearer picture of the product of the reaction between PksS and dihydrobacillaene.

Figure 44. DNA recombination of the double crossover knockout mutants during transformation into *Bacillus*.



The double crossover knockout mutant plasmids will contain the *pksR* and *ymzB* gene fragments on either side of the erythromycin resistance gene. When transforming into *Bacillus*,

the *pksR* genes and *ymzB* genes will go through DNA recombination so that the erythromycin gene is incorporated into the genome and *pksS* is incorporated into the plasmid.

Double crossover knockout mutations were designed with the pMUTIN4 plasmid, erythromycin resistance gene, and portions of the two genes bordering the gene of interest: *baeR*/*pksR*, and *ymzB* (Figure 42). For the crude natural product extract, the *baeR* gene was used from *B. amyloliquefaciens*. For the crude P450 reductase, the *pksR* gene was used from *B. subtilis*.

Table 1. Double Crossover Knockout Plasmid Primers

Gene	Primer Sequence 5'-3'
pMUTIN Core Forward	CTGCCTCGCGCGTTTCGG
pMUTIN Core Reverse	AATTCTTGAAGACGAAAGGGCCTCG
pMUTIN ErR Forward	ATAAAAAACGCCCGGCGG
pMUTIN ErR Reverse	GATCCTCTAGCACAAAAGAAAAACG
BS168 *pksR* Forward	CCCTTTCGTCTTCAAGAATTACTGTGTCTCTGAAGTGCC
BS168 *pksR* Reverse	TGCCGCCGGGCGTTTTTTATAACTTGCGGTTGCACCTG
BS168 *ymzB* Forward	TTCTTTTGTGCTAGAGGATCTCTCCTCGCCTAATAGGG
BS168 *ymzB* Reverse	CACCGAAACGCGCGAGGCAGGGCTTTCCAAAGATCGGTG
BA *pksR* Forward	CCCTTTCGTCTTCAAGAATTATCGGAAGTGGTGATGCG
BA *pksR* Reverse	TGCCGCCGGGCGTTTTTTATGGGCTTGAAACATTGCCTG

BA *ymzB* Forward	TTCTTTTGTGCTAGAGGATCAACAAACAGCCGGTGCCTG
BA *ymzB* Reverse	CACCGAAACGCGCGAGGCAGGAATTTTCAGGGGCGGGAG

Each gene fragment was amplified by PCR and Gibson-HiFi assembly was used to assemble the four pieces into a new plasmid. Primers are used to construct the four pieces for each plasmid (Table 1). After each gene was cloned, they were then analyzed using gel electrophoresis to make sure that the DNA band size was correct. The DNA was purified using a PCR purification kit from Qiagen and the concentration evaluated on Thermo scientific NanoDrop One.

The Gibson-HiFi assembly reactions were set up according to the guidelines for a 4-6 fragment assembly with a 1:1:1:1 ratio of each gene fragment. The reaction incubated for one hour at 50 °C. The constructed plasmid was then transformed into chemically competent DH5-alpha *E. coli* cells. The cells were plated on 100 µg/mL ampicillin plates and left to grow at 37 °C overnight. The next day, starter cultures were made from one of the colonies, 5 mL of LB and 5 µL of 100 mg/mL ampicillin and left to grow overnight. The cultures were purified with a plasmid Miniprep kit from Qiagen to isolate the plasmid DNA and sent for sequencing to confirm the correct plasmid DNA sequence.

The initial double crossover knockout plasmid sequencing did not prove to be successful. Only two to three fragments were assembled and not all four. The pMUTIN4 plasmid core and erythromycin resistance gene showed in the sequencing along with either the *pksR* or *ymzB* gene, but not both.

Single crossover knockout mutations were designed with the pMUTIN4 plasmid based the original paper.[103] Roughly 200 base pairs from *B. amyloliquefaciens* to knockout *baeS* gene

and *B. subtilis* to knockout *pksS* gene were cloned into the pMUTIN4 plasmid using Gibson-HiFi assembly. The primers for the pMUTIN4 plasmid amplified the plasmid DNA outside of the HindIII and BamHI restriction sites. The primers for the fragment of the *B. amyloliquefaciens baeS* gene fragment amplified roughly 200 bps of the *baeS* gene (Table 2). Each gene fragment was amplified by PCR and then analyzed with gel electrophoresis to make sure the band size was correct.

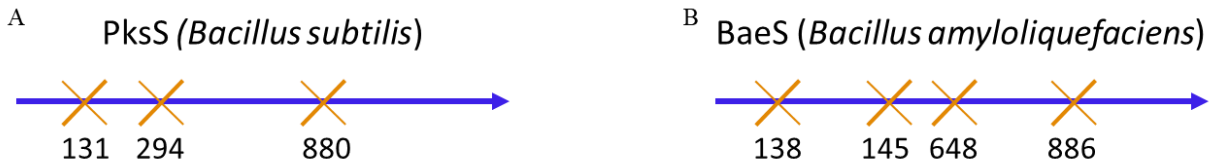
Table 2. Single Crossover Knockout Plasmid Primers

Gene	Primer Sequence 5'-3'
pMUTIN Forward	CATCATGATGGCAAATCGATGATCCCCAGCTTGTTGATAC
pMUTIN Reverse	TTTGGCCGTCCTCTTCAGCGTAATTGTTATCCGCTCACAATT
baeS Forward	TTGTGAGCGGATAACAATTACGCTGAAGAGGACGGCCAAA
baeS Reverse	GTATCAACAAGCTGGGGATCATCGATTTGCCATCATGATGACG

The DNA was purified using PCR purification kit from Qiagen. After purification, the concentration of the DNA was evaluated for Gibson-HiFi assembly. The Gibson-HiFi assembly reactions were set up according to the guidelines in the protocol for a 1:2 vector to insert ratio. The reaction incubated for one hour at 50 °C. The plasmid was then transformed into competent DH5-alpha *E. coli* cells. The cells were plated on 100 µg/mL ampicillin plates and left to grow at 37 °C overnight. The next day, starter cultures were made of the colonies, 5 mL of LB and 5 µL of 100 mg/mL ampicillin and left to grow overnight. The cultures were miniprepped and sent for sequencing to confirm the correct plasmid DNA sequence.

CRISPR/Cas9 knockout mutations were designed using the plasmid pRH030.[123] The plasmid pRH030 was designed for genetic engineering of *Bacillus* species using the CRISPR/Cas9 system. Several knockout plasmids were designed to mutate different areas of the genome of the respected species. CRISPR/Cas9 is not always efficient and targeting different sports in the genome can ensure that one knockout will work. The locations for mutation were chosen using the web tool CHOPCHOP for selection of the target sites.

Figure 45. The CRISPR/Cas9 targets for primer design.



The CRISPR/Cas9 targets were designed at several points along the gene of either *pksS* or *baeS* genes in their respective *Bacillus* species. Each target represents a unique mutation and knockout plasmid, consisting of seven total plasmids. All seven plasmids, each with one of the targets, are being constructed to ensure one successful mutation. A) Three targets were made for *Bacillus subtilis*. B) Four targets were made for *Bacillus amyloliquefaciens*.

For the initial design of the plasmids, oligos for PCR were designed. These oligos were to make up several plasmids and were selected at different parts of the genome. The oligos were amplified with PCR and then purified with a PCR purification kit. The oligos were phosphorylated with T4 PNK. 1 uL of the forward and reverse oligo were added to 1 µL of 10 x T4 ligation buffer, 6.5 µL water, and 0.5 µL T4 PNK. The phosphorylation occurred by allowing the reaction to heat to 37 °C for 30 minutes, then 95 °C for 5 minutes and ramped down to 25 °C at 5 °C/min. The oligo complex was diluted. The ligation of the plasmid included the digested pRH030 plasmid, oligo duplex, T4 ligase, and water. This was allowed to incubate for 10

minutes at room temperature. The constructed plasmid was transformed into *E. coli*. After incubation overnight, the colonies were prepared for sequencing.

Table 3. CRISPR/Cas9 Target Sequences

Species	Target Sequence	Location
B. subtilis	ATCGTATCGGGTAATTAACC	131
B. subtilis	CAAGTGATCTCAGGCGGGTA	294
B. subtilis	CTGTCGTTATGATGGCAAAT	880
B. amyloliquefaciens	GGCAATGGATGCTTCCATCC	138
B. amyloliquefaciens	GATGCTTCCATCCGGGCGAG	145
B. amyloliquefaciens	CGGACAAGATGCGACCAACC	648
B. amyloliquefaciens	TTCGATCTTTGGTGCATCAG	886

Redesigning of the CRISPR/Cas9 plasmids consisted of using Gibson-HiFi assembly to integrate the target sequence into the plasmid. Seven targets were selected for seven knockout plasmids of the *Bacillus* species (Table 3). Four targets were selected for *B. amyloliquefaciens*, and three targets were made for *B. subtilis* (Figure 43). The primers made for each plasmid (Table 4). After PCR, the restriction enzyme DpnI was added to cleave the methylated DNA that remained. The DNA, with DpnI was incubated at 37 °C for 1 hour, then the enzyme was heat deactivated at 80 °C for 20 minutes. The DNA was then cleaned up with a PCR purification kit (Qiagen) and the concentration obtained on the Thermo scientific NanoDrop One. The Gibson-HiFi assembly reactions were set up according to the guidelines in the protocol for a 1:2 vector

to insert ratio. The reaction incubated for one hour at 50 °C. The plasmid was then transformed into competent DH5-alpha *E. coli* cells. The cells were plated on 50 µg/mL kanamycin plates and left to grow at 30 °C overnight. The next day, starter cultures were made of the colonies, 5 mL of LB and 5 µL of 50 mg/mL kanamycin and left to grow overnight.

Table 4. CRISPR/Cas9 Knockout Plasmid Primers

Gene	Primer Sequence 5'-3'
BA138 Forward	GGCAATGGATGCTTCCATCCGTTTTAGAGCTAGAAATAGCAAGT
BA138 Reverse	GGATGGAAGCATCCATTGCCCGTAGGTACATTTTACTCAATTC
BA145 Forward	GATGCTTCCATCCGGGCGAGGTTTTAGAGCTAGAAATAGCAAGT
BA145 Reverse	CTCGCCCGGATGGAAGCATCCGTAGGTACATTTTACTCAATTC
BA648 Forward	CGGACAAGATGCGACCAACCGTTTTAGAGCTAGAAATAGCAAGT
BA648 Reverse	GGTTGGTCGCATCTTGTCCGCGTAGGTACATTTTACTCAATTC
BA886 Forward	TTCGATCTTTGGTGCATCAGGTTTTAGAGCTAGAAATAGCAAGT
BA886 Reverse	CTGATGCACCAAAGATCGAACGTAGGTACATTTTACTCAATTC
BS131 Forward	ATCGTATCGGGTAATTAACCGTTTTAGAGCTAGAAATAGCAAGT
BS131 Reverse	GGTTAATTACCCGATACGATACGTAGGTACATTTTACTCAATTC
BS294 Forward	CAAGTGATCTCAGGCGGGTAGTTTTAGAGCTAGAAATAGCAAGT
BS294 Reverse	TACCCGCCTGAGATCACTTGACGTAGGTACATTTTACTCAATTC

BS880 Forward	CTGTCGTTATGATGGCAAATGTTTTAGAGCTAGAAATAGCAAGT
BS880 Reverse	ATTTGCCATCATAACGACAGACGTAGGTACATTTTACTCAATTC

The cultures were miniprepped to isolate the plasmid DNA and sent for sequencing to confirm the correct sequence. The entire sequence of the plasmids was confirmed by sequencing from the company Plasmidsaurus that uses nanopore technology. Stocks of the knockout plasmids were made by growing starter cultures of 5 mL LB with antibiotic overnight. The next day, 850 µL of culture was added to 150 µL glycerol to be stored in the -80 °C freezer. Plates were grown of *E. coli* that housed the knockout plasmid by streaking from the stock in the -80 °C freezer on LB and left to grow overnight. The DNA of the knockout plasmid was obtained from a miniprep kit. The concentration of the DNA was obtained through Thermo scientific NanoDrop One for transformation into their respective *Bacillus* species.

Transformation of Green Fluorescent Protein into *Bacillus* species

To test the transformation of *Bacillus* species green fluorescent protein (GFP) was transformed into the bacteria. If the transformation was successful, the bacteria should glow green. If the transformation method is not successful, the bacteria will not glow green. The GFP plasmid (Addgene # 112792) was used for the *Bacillus* species.[124]

Transformation of knockout plasmids into *Bacillus* species

Once the knockout plasmids for *baeS* in *B. amyloliquefaciens* and *pksS* in *B. subtilis* were confirmed to have the correct sequence, the plasmids were transformed into their respective *Bacillus* species. After unsuccessful attempts of transforming with the circular plasmid, restriction digest was preformed to linearize the DNA. There has been precedence in the literature for linearizing plasmid DNA for more effective transformations into bacteria.[125] The

cut site of the single crossover knockout *B. amyloliquefaciens* pMUTIN plasmid was chosen to be an unused location in the plasmid in the *lacZ* gene fragment. The restriction digest enzyme ApaI was used to cut the plasmid one time creating a linear piece of DNA for transformation.

To make chemically competent *Bacillus* cells the method was followed from Frank Young laboratory at The University of Rochester, Medical Center. Transformations of both *Bacillus* species were done synchronously. *Bacillus* was plated on LB plates and left to grow overnight at 37 °C. The next day, 3-6 colonies were used to inoculate 20 mL of medium A and then the culture was grown at 37 °C to a 0.1-0.2 OD_{650}. Medium A composed of 81 mL of sterile water, 10 mL 10 x medium A base, and 9 mL 10 x *Bacillus* salts. The 10 x medium A base contains 10 g yeast, 2 g casamino acids, and 900 mL water, with 100 mL 50 % filtered glucose. The 10 x *Bacillus* salts contain 20 g $(NH_4)_2SO_4$, 183 g $K_2HPO_4 \cdot 3H_2O$, 60 g KH_2PO_4, 10 g Na^+ citrate, and 2 g $MgSO_4 \cdot 7H_2O$ in 1000 mL water. The cells were left to shake at 37 °C as the OD_{650} was read every 20 minutes until the growth curve fell into a straight line, or t_0. After t_0 the cells were left to grow for another 60-90 minutes until the OD_{650} read 0.4-0.6. The cells then continued to incubate for another 90 minutes before 50 uL of cell culture was transferred to 450 uL of warmed medium B. Medium B consisted of 10 mL medium A, 0.1 mL 50 mM $CaCl_2 \cdot 2H_2O$, and 0.1 mM 250 nM $MgCl_2 \cdot 6H_2O$. The cultures incubated for 90 minutes at 37 °C before 1 µg of plasmid DNA was added and incubated for an additional 30 minutes. The cell cultures were plated and left to grow overnight at 37 °C.

An additional method for the generation of *Bacillus* competent cells was provided by the Jie Li laboratory at the University of South Carolina. A starter culture of *Bacillus* was grown overnight. The culture was used to inoculate 10 mL of MNGE media to an OD_{600} of 0.1 and left to grow to an OD_{600} of 1.1 at 37 °C. The MNGE media consists of 8.28 mL of sterile water, 920

µL of 10 x MN medium, 1 mL 20 % glucose, 50 µL 40 % glutamate, 50 µL 2.2 mg/ml Fe^{II} ammonium-citrate, 100 µL 5 mg/mL tryptophan, and 30 µL 1 M $MgSO_4$. The 10 x MN media consisted of 136 g $K_2HPO_4 \cdot 3H_2O$, 60 g KH_2PO_4, and 10 g Na-citrate. For the transformation, 400 µL of the cells were added to a glass test tube with 2 µg of linearized plasmid DNA. The cells grew for 1 hour at 37 °C and then 100 µL of expression media was added and left to grow for an additional hour before plating. The expression media consists of 500 µL 5 % yeast extract, 250 µL 10 % casamino acids, 50 µL 5 mg/mL tryptophan, and 250 µL water.

The last attempted method to make *Bacillus* competent cells consists of several changes to the previous methods. A starter culture of the bacteria was grown overnight in LB. The starter culture was added to GCHE media to an OD_{600} of roughly 0.3 and left to shake at 37 °C to around an OD_{600} of 2. At this point, an equal volume of GE media was added to the cells and left to shake for 1 hour at 37 °C. The culture was centrifuged down at 6000 rpm for 5 minutes. The pellet was resuspended in 950 µL of supernatant and 50 µL of 10 % glucose. To 200 µL of the cells, 500 ng of linearized DNA was added and left to shake at 75 rpm and 28 °C for 20 minutes. Afterwards, 500 µL of LB was added to the cells and left to shake for 90 minutes at 37 °C. The cells were then plated on 5 µg/mL erythromycin plates and left to grow overnight. The 500 mL GCHE media consists of 1 g L-glutamate, 4.67 g K_2HPO_4, 3.15 g of KH_2PO_4, 0.4 g trisodium citrate, 0.2 g $MgSO_4$, 0.011 g ferric ammonium citrate autoclaved with 5 g glucose, 0.025g L tryptophan, and 0.5 g casein hydrolysate filtered into the media. The 500 mL GE media consisted of .67 g K_2HPO_4, 3.15 g of KH_2PO_4, 0.4 g trisodium citrate, 0.2 g $MgSO_4$, 0.011 g ferric ammonium citrate autoclaved with 5 g glucose and 0.025 g/L tryptophan filtered into the media. Later on, the 0.1 M potassium phosphate (pH 7) was made into a separate buffer to be added to the media after it was autoclaved.

After unsuccessful attempts to transform the knockout plasmids, additional methods were attempted to transform the plasmids. Glycine (5 %) was added in an attempt to increase the DNA uptake of the competent *Bacillus* cells. Adding glycine has been shown to be successful in making electrocompetent *Bacillus* cells. This addition weakens the cells walls so that the bacteria is more susceptible to the uptake of the plasmid DNA.[126] This work was done with the help of undergraduate Marve Basurto Campos in the Jason Reddick lab.

Reactions with PksE

In the crude natural product extract from wild type ATCC 39374, with the addition of either NADH or NADPH and the crude reductase, there was an increase in dihydrobacillaene. There was an idea that the crude reductase retained the enzyme PksE and other potential enzymes. The increase in dihydrobacillaene was led to be believed to be caused by the additional PksE in the crude reductase extract. The *trans*-enoylreductase, PksE, will reduce the double bond during the initial polyketide synthase in the module *pksJ*/*baeJ* down to the single bond in the biosynthesis of dihydrobacillaene. We thought it possible that PksE was also capable of reducing the double bond to a single bond of bacillaene in vitro to dihydrobacillaene.

To see if PksE would react with dihydrobacillaene from the crude natural product extract, the pBAD plasmid that housed the *pksE* gene was retransformed into TOP10 *E. coli*. PksE was purified by standard protein purification. Reactions were analyzed by LC-MS of PksE and the crude natural product extract. The peak areas of bacillaene, dihydrobacillaene, and hydroxylated dihydrobacillaene were compared to the internal standard. The changes of the crude natural product extract and the reactions with PksE were indistinguishable. It was hypothesized that PksE can only react with the biosynthetic intermediate and not the full structure of bacillaene to

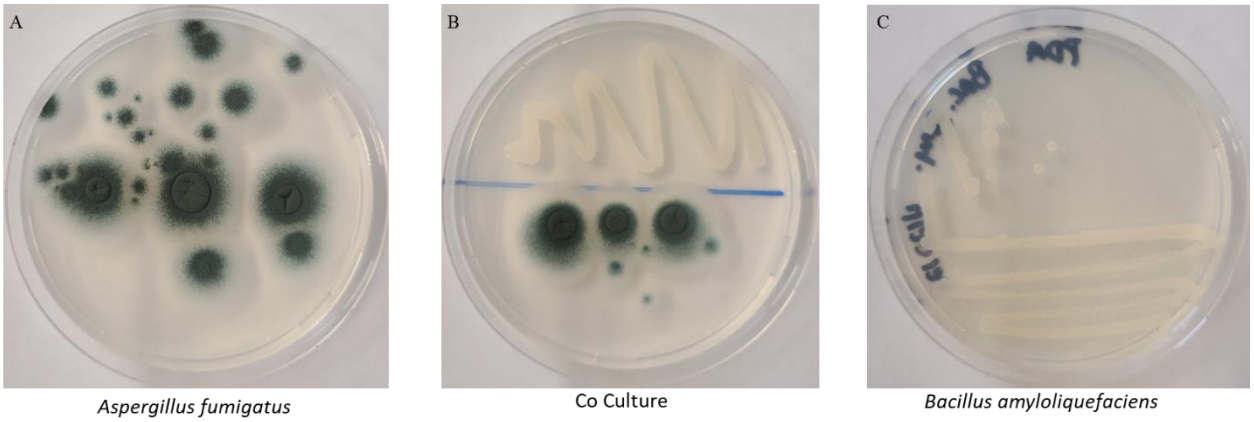
produce dihydrobacillaene. At this time, we do not know what might be causing the apparent reduction.

RESULTS AND DISCUSSION

Co-Culturing *Bacillus amyloliquefaciens*

Dihydrobacillaene is known to be both an antibiotic and antifungal. This hypothesis was that if dihydrobacillaene is used as an advantage for the bacteria in the presence of a predator, co-culturing with a predatory fungus could stress the bacteria more than a lack of nutrients. We expected this stress would cause an increase of dihydrobacillaene to be produced as the bacteria fought for survival. Co-culturing of *B. amyloliquefaciens* and *A. fumigatus* was done in the attempts to achieve a higher yield of dihydrobacillaene compared to liquid Landy media cultures (Figure 44). Landy media is a minimal media that will cause stress to the culture. This minimal media lacks yeast extract, and only has a marginal amount of glucose for the cells to survive.

Figure 46. Co-cultures of *Bacillus amyloliquefaciens* and *Aspergillus fumigatus*.



Co-culturing of the bacteria and fungi was done with the prediction that it could increase the amount of dihydrobacillaene produced by the bacteria. *Bacillus* species produce dihydrobacillaene in their biofilm in the presence of predators as a source of protection. The co-culture was extracted and compared to the monocultures and Landy media for an increase in

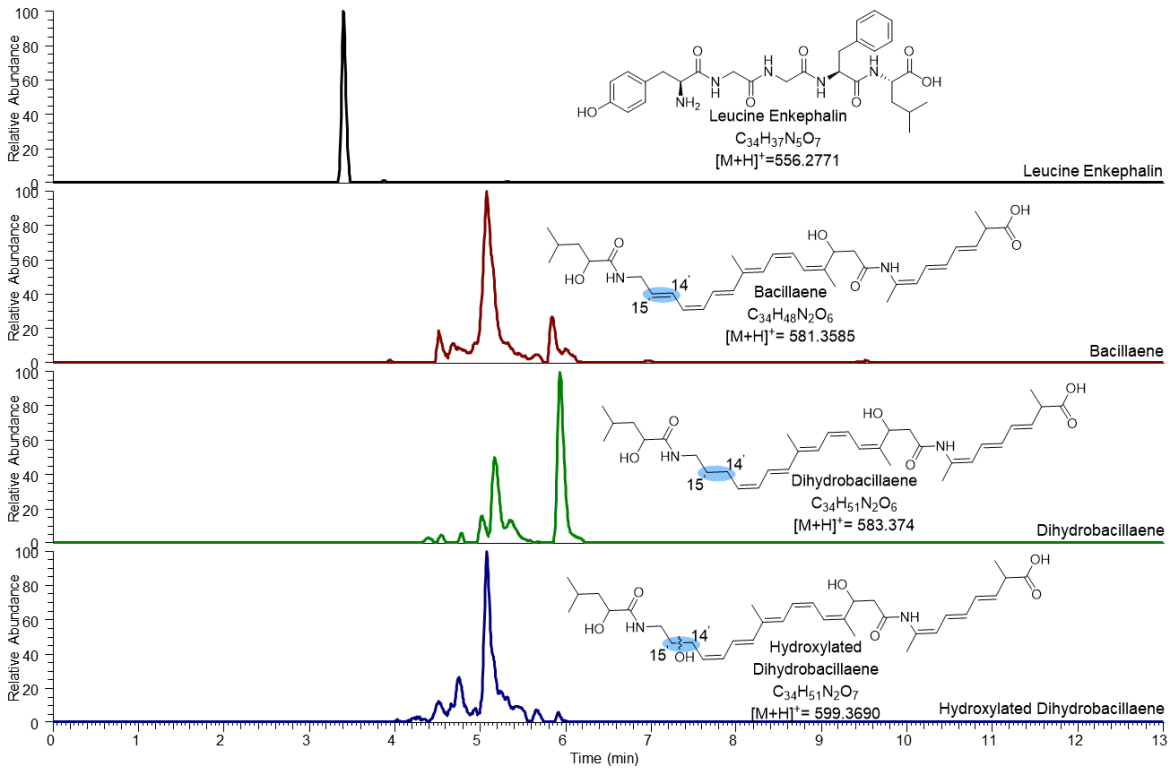
dihydrobacillaene. After LC-MS analysis, it was shown that the co-culturing did not increase the amount of dihydrobacillaene. A) Monoculture of *Aspergillus fumigatus*. B) Co-culture between *A. fumigatus* and *B. amyloliquefaciens*. C) Monoculture of *Bacillus amyloliquefaciens*.

Dihydrobacillaene will be formed in the biofilm of the bacteria when under stress. This stress on the cells was expected to produce the compound of interest, dihydrobacillaene. Wild type *Bacillus amyloliquefaciens* was co-cultured with *Aspergillus fumigatus* at various incubations times, temperatures, and medias. The extract was analyzed on LC-MS for an increased production of dihydrobacillaene compared to the liquid Landy media cultures and monocultures of *B. amyloliquefaciens*. When comparing the co-cultures to the Landy media cultures and monocultures, there was a negligible change in production of dihydrobacillaene. Due to the relatively same amounts of dihydrobacillaene being produced in both the co-cultures and Landy media, the substrate was extracted thereafter from cultures grown in liquid Landy media without co-culturing with fungi.

Initial LC-MS Analysis of Reactions

The initial reactions with the wild type of crude natural product extract, crude cytochrome P450 reductase extract, PksS, and NAD(P)H were analyzed for change in peak area compared to the internal standard. The internal standard of leucine enkephalin was chosen due to some of the structure similarity in addition to the peaks not overlapping on the chromatogram. The peak areas of bacillaene, dihydrobacillaene, and hydroxylated dihydrobacillaene were normalized with the peak areas of constant leucine enkephalin. It was shown before in previous work that the substrate for the enzyme, PksS, is dihydrobacillaene.

Figure 47. Crude natural product extract from *B. amyloliquefaciens* using liquid-chromatography mass-spectrometry.

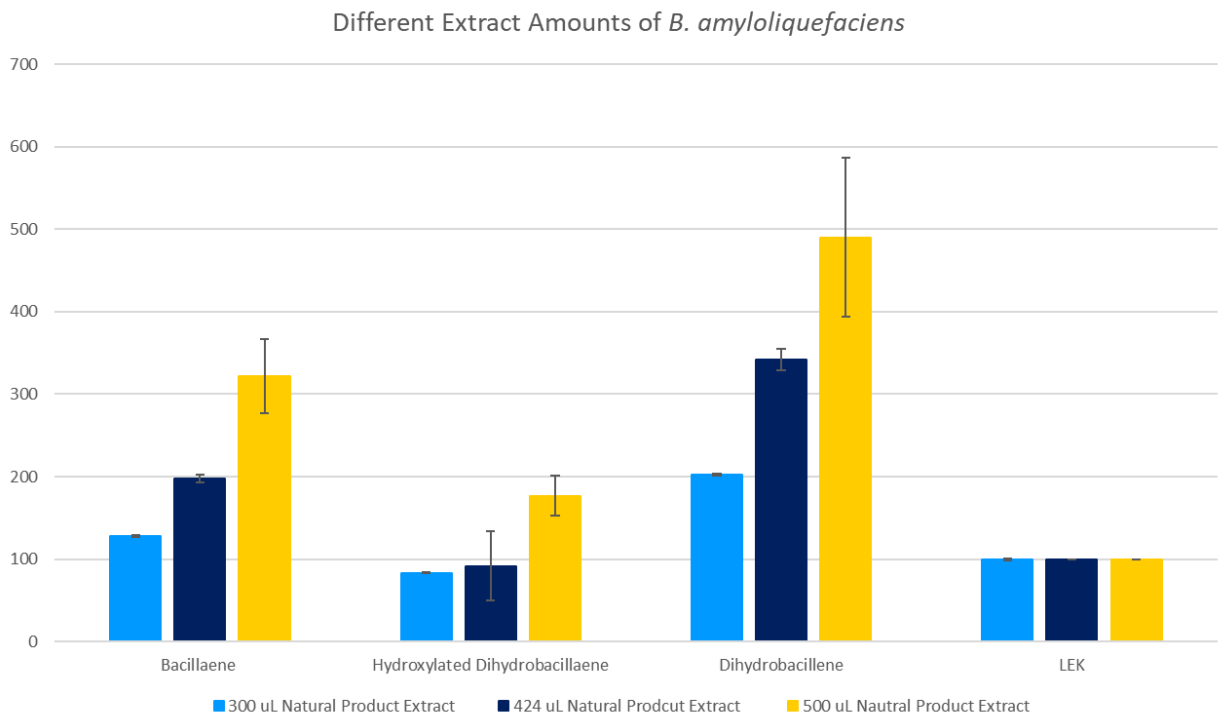


The bacteria were grown in Landy media for 28 hours at 30 °C. Leucine enkephalin was added as an internal standard to determine the relative amount of bacillaene, dihydrobacillaene, and hydroxylated dihydrobacillaene, as there is no standard of these compounds to make a calibration curve and then quantify. Each bacillaene analog was identified by mass in the crude natural product extract.

No quantitative work was done at that time, so while there was a loss in dihydrobacillaene there were no quantitative measurements of any changes in the components still present in the reaction mixture product peak. The initial tests were looking for a reduction in dihydrobacillaene and increase in either bacillaene or hydroxylated dihydrobacillaene when purified PksS, the crude cytochrome P450 reductase, and NAD(P)H were added to the reaction.

Control reactions were done to make sure that the potential increases in the product were from the full reaction. Initial runs of just the natural product extract were analyzed (Figure 45). This was done to get a baseline of the chromatogram without additional components of the reaction.

Figure 48. Standard relative abundance of bacillaene, hydroxylated dihydrobacillaene, and dihydrobacillaene in the natural product extract to leucine enkephalin.

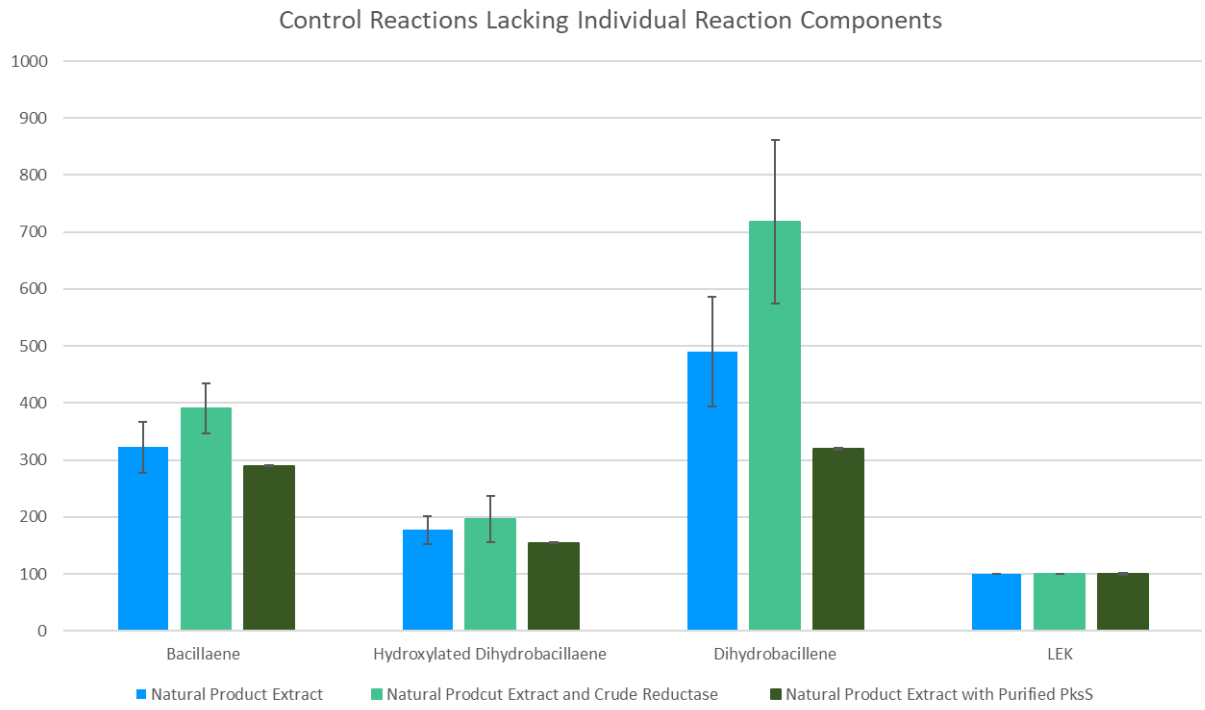


At a total volume of 800 µL, we compared reactions with 300 µL extract (38 % extract) in light blue, 424 µL extract (53 % extract) in dark blue and 500 µL extract (62 % extract) in gold.

The extract of wild type of *B. amyloliquefaciens* showed all three bacillaenes: bacillaene, dihydrobacillaene, and hydroxylated dihydrobacillaene. This showed that the *baeS* gene was still active in the cells before extraction. Comparisons were done between 300 µL, 424 µL, and 500 µL of the crude natural product extract (Figure 46). The comparison between the different

amounts of extract showed that the LC-MS detection and integration was reliably responsive to different quantities of the extracts. When compared to the internal standard, there was an increase in all three bacillaenes as there was an increase in the amount of culture in the sample. Controls were conducted lacking each component of the reaction (Figure 47).

Figure 49. Standard relative abundance of bacillaene, hydroxylated dihydrobacillaene, and dihydrobacillaene to leucine enkephalin.



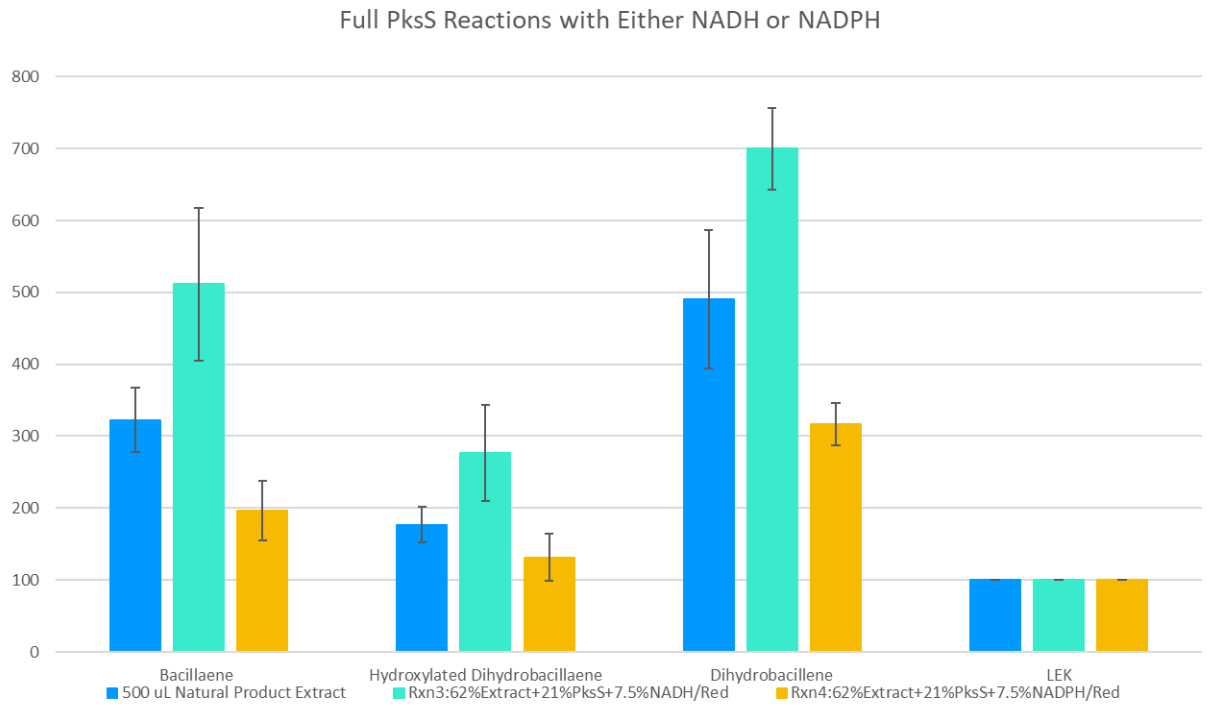
The 500 µL natural product extract is shown in light blue. The control reaction of the natural product extract with the reductase, lacking NAD(P)H and PksS is in light green. The control reaction of the natural product extract with PksS, lacking the reductase, and NAD(P)H is in dark green.

The crude natural product extract was compared to the crude natural product with the crude cytochrome P450 reductase and compared to the crude natural product extract with purified PksS. The enzyme should not be able to function with a lack of the crude cytochrome

P450 reductase and NAD(P)H. The expected outcome of these controls were similar amounts of the bacillaenes to the crude natural product extract. In reactions having only the crude natural product extract with the crude cytochrome P450 reductase, there was a seemingly increase in both bacillaene and dihydrobacillaene. There was also a loss of dihydrobacillaene when comparing the crude natural product extract to the crude natural product extract with the purified PksS.

Full reactions contained the crude natural produce extract, cytochrome P450 extract, purified PksS, and NAD(P)H (Figure 48). The changes in the reaction varied with each reaction.

Figure 50. The full reaction with the natural product extract, reductase, and purified PksS.



The control of the natural product extract is shown in light blue. The full reaction with the natural product extract, crude cytochrome P450 reductase, PksS, and NADH is shown in light green. The full reaction with the natural product extract, crude cytochrome P450 reductase, PksS, and NADPH is shown in yellow. These reactions showed varying amounts of the

bacillaenes in each run. The reactions did not show a loss of dihydrobacillaene, which is the substrate for the reaction. The results showed that knockouts of the *pksS* and *baeS* genes were needed for further studies.

There was no consistent loss of dihydrobacillaene as it should have reacted with PksS. The Reddick group discovered previously that the substrate for the enzyme PksS was dihydrobacillaene, and yet there was no consistent loss of dihydrobacillaene when the enzyme was added to the extract. There was no consistent gain in either the bacillaene or hydroxylated dihydrobacillaene. The predicted product of the reaction would be either bacillaene or the hydroxylated dihydrobacillaene based on typical cytochrome P450 reactions. There was typically still more dihydrobacillaene in the reaction than either bacillaene or hydroxylated dihydrobacillaene. Since dihydrobacillaene is the substrate for PksS, there should have been consistent loss of the compound in the reactions.

It was hypothesized that residual amounts of the population of the bacillaene compounds in the extract shown in the mass spectrometry chromatogram were likely impacted by the presence of *baeS* in the cells before extracting them. With the full pathway still active in the bacteria, there is nothing stopping it from continuing to the product of the reaction, and that product showing on the chromatogram. The three bacillaene variants differ only by a double bond or an alcohol group making them structurally very similar. The data shows the mass equivalent of bacillaene in the hydroxylated dihydrobacillaene peak, which suggests that dehydration is occurring in the mass spectrometer.

Knockout Mutations of *B. amyloliquefaciens* and *B. subtilis*

Several plasmids were designed to knockout *baeS* in *B. amyloliquefaciens* and *pksS* in *B. subtilis*, several knockout plasmids were made to transform back into their respective *Bacillus*

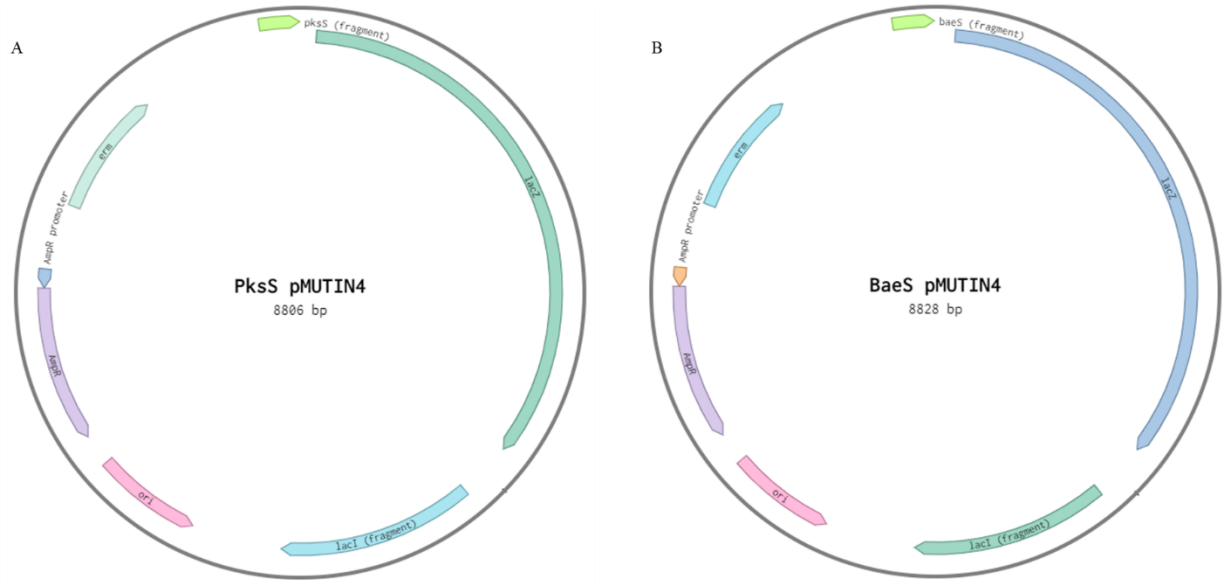
species. Without the *baeS*/*pksS* genes being active, the only compound showing in the chromatograms was expected to be dihydrobacillaene. With purified PksS added to these knockouts, the product of the reaction should be more apparent without trace amounts already shown.

One set of plasmids were designed to generate double crossover knockout mutants lacking *pksS*/*baeS*. These plasmids were made using the pMUTIN4 plasmid, erythromycin resistance gene, and genes on either side of the gene of interest: *pksR*/*baeR* and *ymzB*. Erythromycin resistance was used so that if the bacteria successfully went through DNA recombination with the plasmid, it would have the erythromycin resistance gene. This would cause only bacteria with the resistance gene to grow on erythromycin plates. If the bacteria did not contain the new erythromycin gene from the plasmid, then it would not be able to grow on erythromycin plates. After each gene was cloned, Gibson-HiFi assembly was used to assemble the four DNA pieces. Sequencing showed that not all four DNA pieces would assemble together. The pMUTIN4 plasmid and erythromycin resistance gene were shown to be assembled with one of the other genes, but not both.

When the construction of the double crossover knockout mutation plasmids was shown not to be successful, plasmids designed to create single crossover knockout mutations were made. Primers were designed to amplify the pMUTIN4 DNA excluding the region between the *HindIII* and *BamHI* restriction sites. These sites were selected based on the original pMUTIN paper as the restriction digest sites.[103] Roughly 200 bps of each gene were chosen to clone into the plasmid. The area of the gene was selected to be roughly in the middle of the gene, around 400-600 bp region of *pksS*/*baeS* gene. Roughly 200 bps of each gene were used to clone into the pMUTIN4 plasmid using Gibson-HiFi assembly (Figure 49). Plasmids were purified from *E. coli*

transformants generated after the assembly reaction, they were sent out to Plasmidsaurus for conformation of the sequence of the plasmid.

Figure 51. Single crossover knockout mutation plasmids for *pksS*/*baeS*.

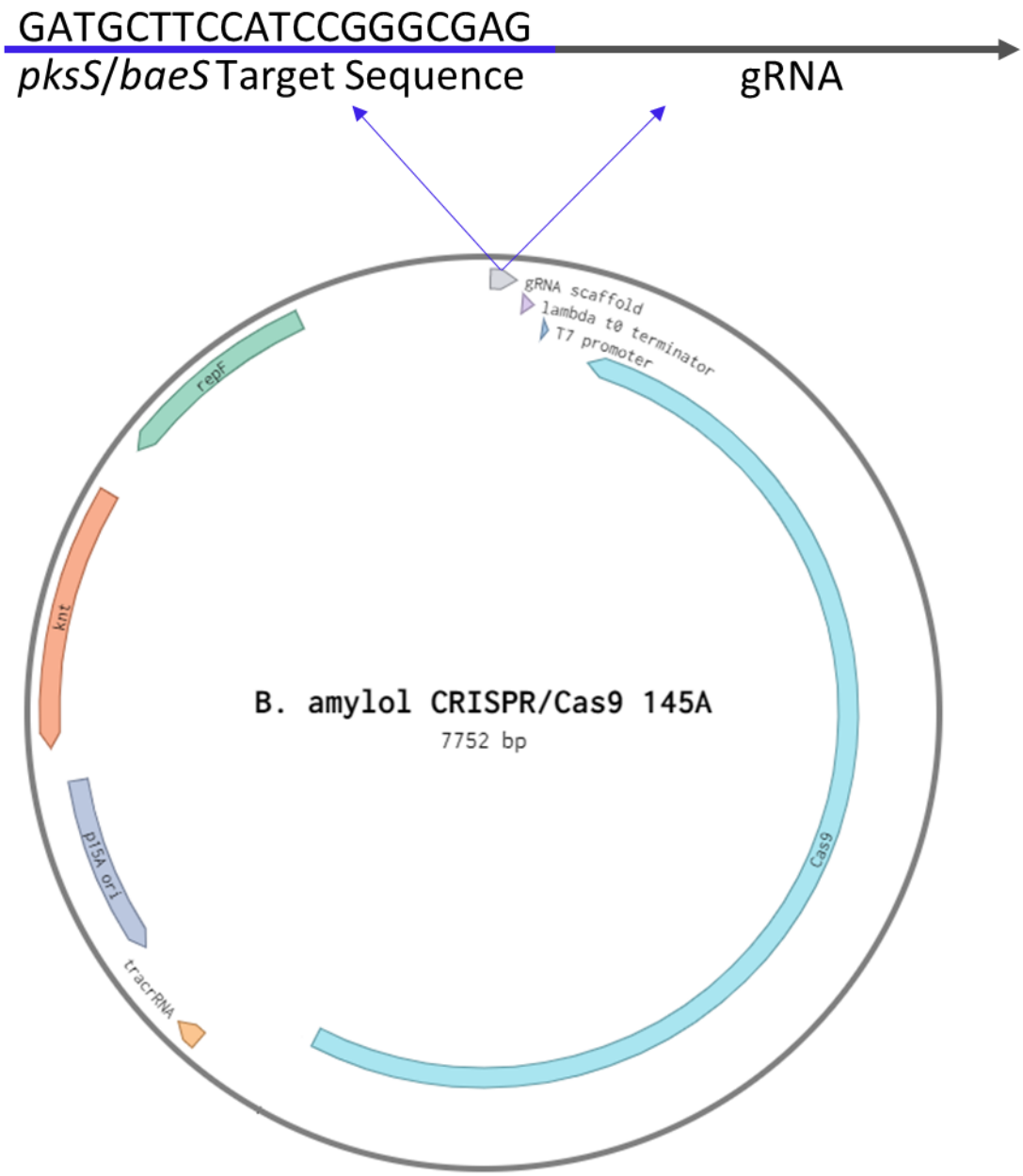


These clones will disrupt the *pksS*/*baeS* gene using a Campbell-type single crossover integration to knockout the genes of interest. A) 213 base pairs from the *pksS* gene were selected to be cloned into the pMUTIN4 plasmid using HiFi-Gibson assembly. B) 219 base pairs from the *baeS* gene were selected to be cloned into the pMUTIN4 plasmid using HiFi-Gibson assembly.

CRISPR/Cas9 knockout mutation plasmids were made for a different mutation strategy complimentary to the recombination-based plasmids described above. Initial plasmids were cloned using a restriction digest enzyme to cut the plasmid and insert the target sequence, roughly 20 bps in length. After several attempts to clone the target sequence into the plasmid, it was shown that the restriction digest enzyme was not cutting the plasmid DNA completely. Less DNA was used, with the same total volume of reaction, in case there was salt inhibition of the enzyme from the plasmid purification kits. After this also did not cause an increase in the amount of DNA being cut by the enzyme, the amount of enzyme being used was increased.

After initial plasmids were attempted with the restriction digest enzymes, the cloning process was switched to using Gibson-HiFi assembly. With Gibson-HiFi assembly, the plasmid could be cloned into two fragments, with one fragment containing the target sequence introduced by way of the PCR primers used to generate the PCR fragments used in the Gibson HiFi assembly. The plasmid was cloned in half, with one half consisting of the target sequence and assembled together using Gibson-HiFi assembly.

Figure 52. CRISPR/Cas9 knockout plasmid.



CRISPR/Cas9 plasmids were cloned using Gibson-HiFi assembly of the pRH030 plasmid (Addgene # 159633) and the CRISPR target sequence. The *B. amyloliquefaciens* CRISPR target 145 is directly before the gRNA scaffold. The additional CRISPR/Cas9 plasmids are identical besides the target site. Each plasmid will have the corresponding target from directly before the gRNA scaffold.

The target sequence would be assembled right before the gRNA scaffold (Figure 50). Seven plasmids were made, four for *B. amyloliquefaciens* and three for *B. subtilis*. The target sequences were chosen to be at several points along the genome of the gene of interest. Once the plasmid was transformed into *E. coli*, it was sent out to Plasmidsaurus for confirmation of the sequence of the plasmid. The sequence data was compared to theoretically made plasmids for conformation.

Knockout Mutation Transformations of *Bacillus* species

Once the knockout plasmids for *baeS* in *B. amyloliquefaciens* and *pksS* in *B. subtilis* were confirmed to have the correct sequence, the plasmids were transformed into their respective *Bacillus* species. Before transforming of the knockout plasmids, a green fluorescent protein (GFP) was used to test the transformation method. Currently, transformation methods are still being optimized.

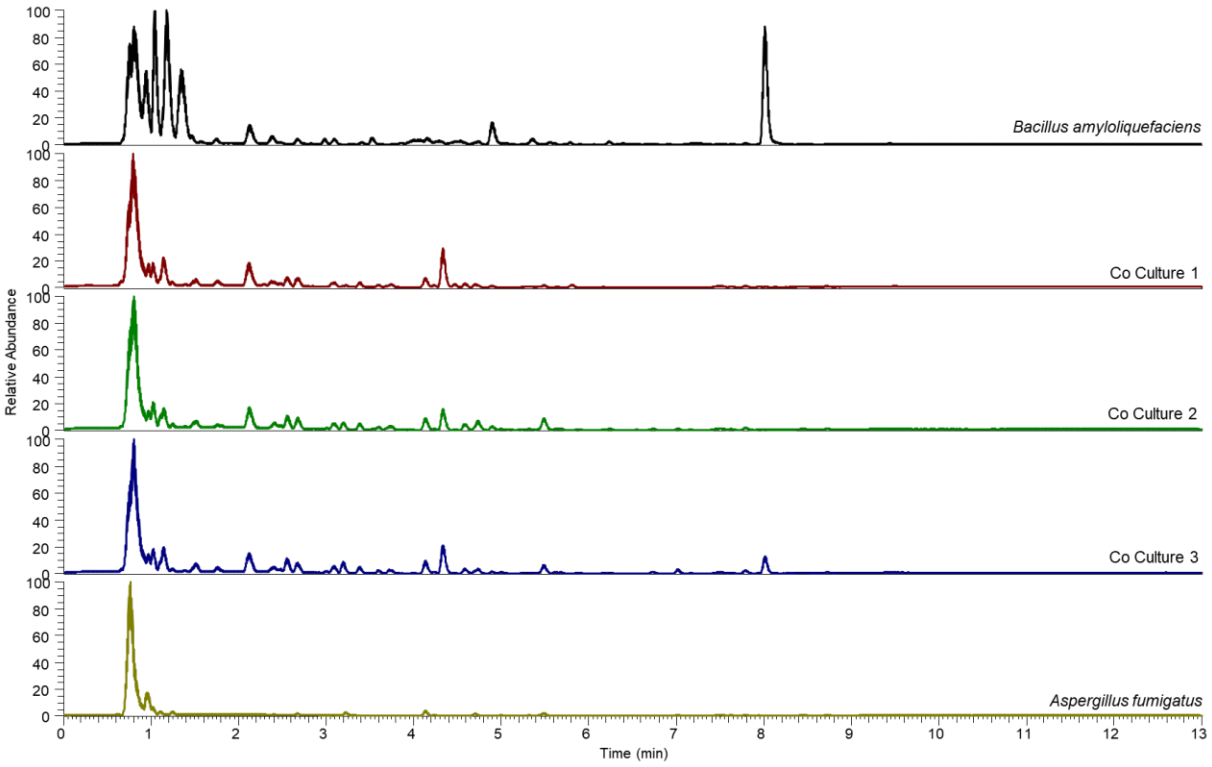
CONCLUSIONS

The biosynthesis of bacillaene is done through *trans*-acyltransferase polyketide synthase/nonribosomal peptide synthases. The polyketide synthase/nonribosomal peptide synthases product is dihydrobacillaene. The biosynthesis starts with an alpha-hydroxyisocaproate (α-HIC) and then glycine is added to create the starting unit. This is then followed by a nonribosomal peptide synthase module that will insert an alanine. The β-branch is formed from the HMG gene cassette that includes *pksF-pksI* and *acpK*. The *trans*-enoylreductase, PksE, will reduce the double bond during the biosynthesis to produce the dihydrobacillaene intermediate. Knockout mutation plasmids were generated to knockout the *baeS* gene in *B. amyloliquefaciens* and *pksS* gene in *B. subtilis*. The double crossover knockout plasmid construction was shown to not be successful through sequencing. Plasmids designed to generate knockout mutations

through CRISPR/Cas9 methodology and single knockout mutations were shown to be successfully generated through sequencing. In initial mass spectrometry analysis, bacillaene and hydroxylated dihydrobacillaene co-eluted. This led to our conclusion that dihydrobacillaene with PksS is potentially producing hydroxylated dihydrobacillaene. This will be determined later on with the knockout mutants.

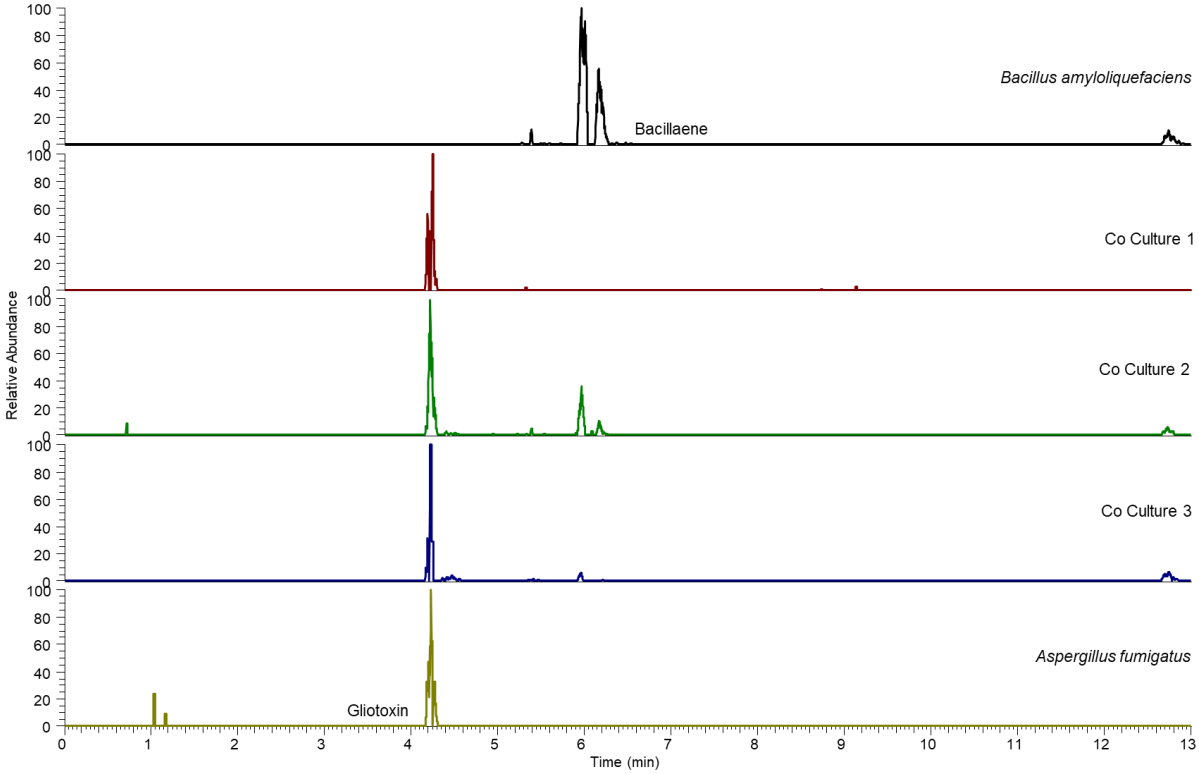
SUPPLEMENTARY DATA

Figure 53. LC-MS of co-culture extract.



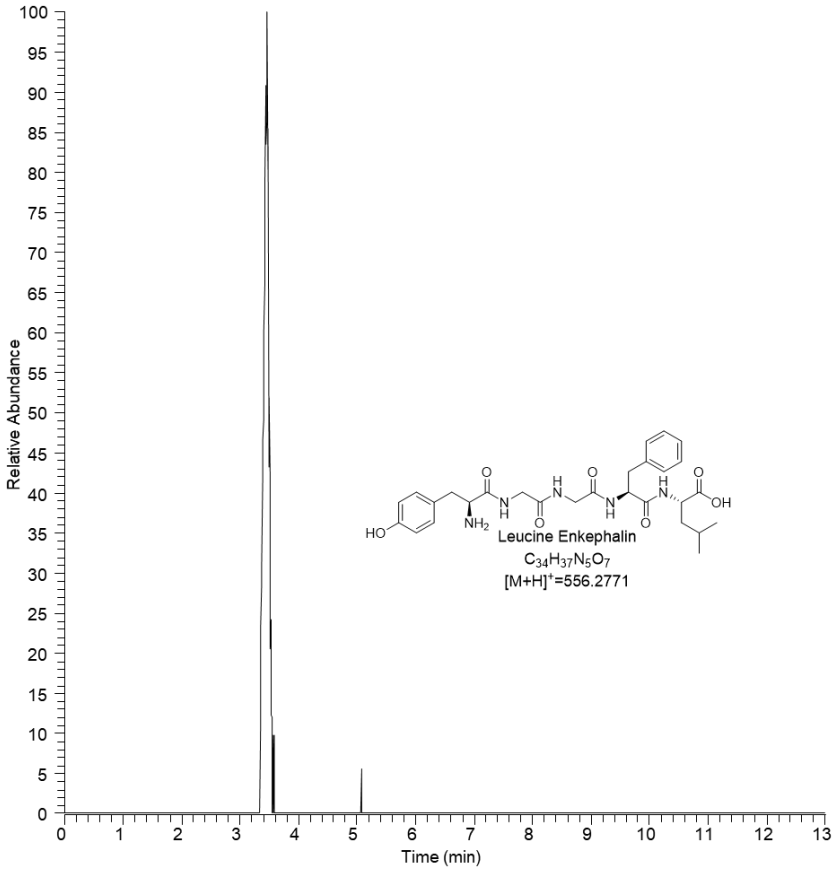
The mass spectrometry chromatographs of the monocultures and co-cultures of the bacteria and fungi are shown above.

Figure 54. Selected mass chromatograms in co-culture extracts.



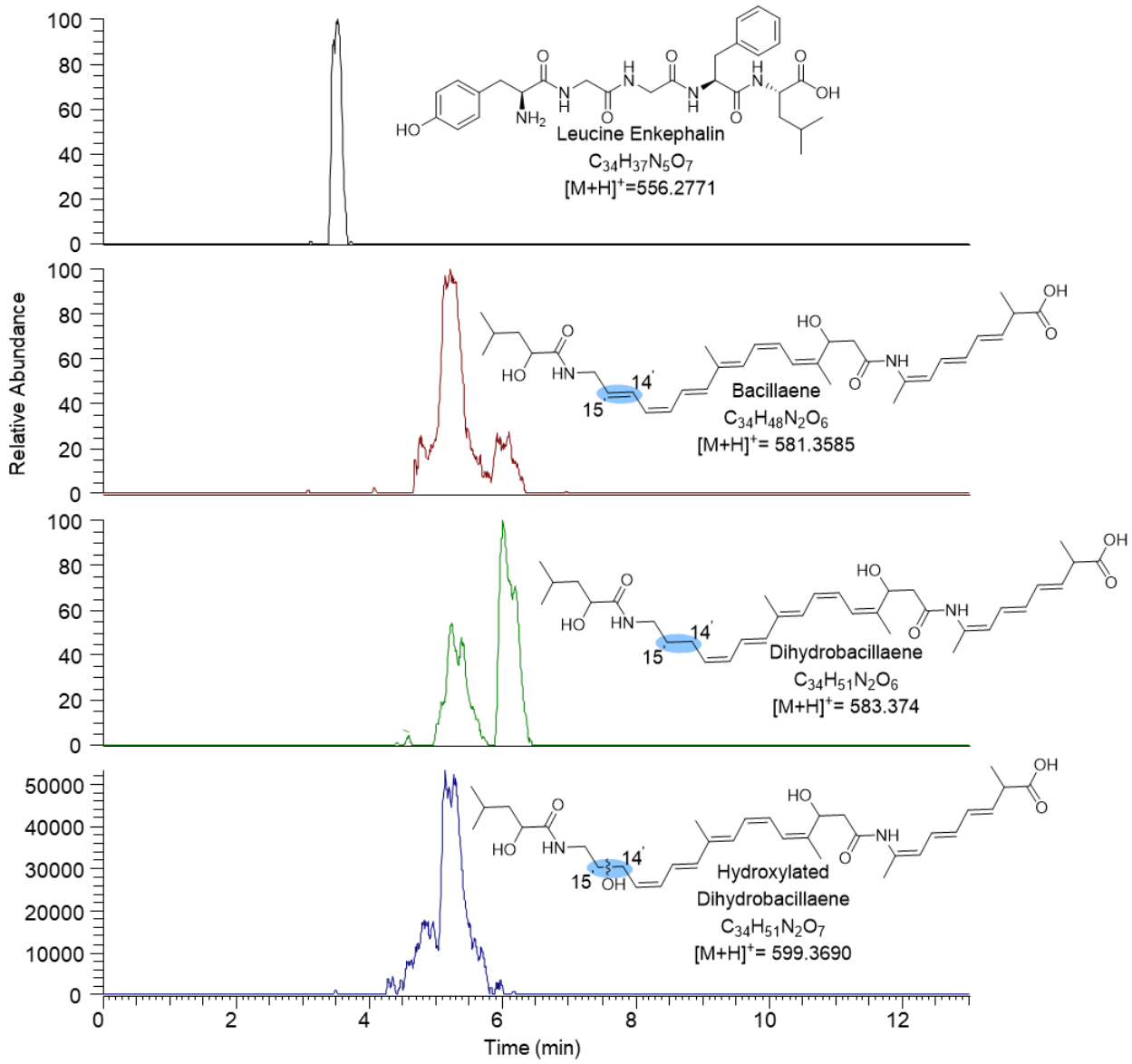
These chromatograms show selection for the masses of bacillaene, dihydrobacillaene, hydroxylated dihydrobacillaene, and gliotoxin. Gliotoxin is produced by *A. fumigatus*. Dihydrobacillaene has literature precedence for being toxic to the cells so a larger amount of dihydrobacillaene was not produced by the co-culturing method compared to the liquid media method.

Figure 55. Leucine enkephalin chromatogram. Leucine enkephalin was chosen as the internal standard for the dihydrobacillaene reactions with PksS.



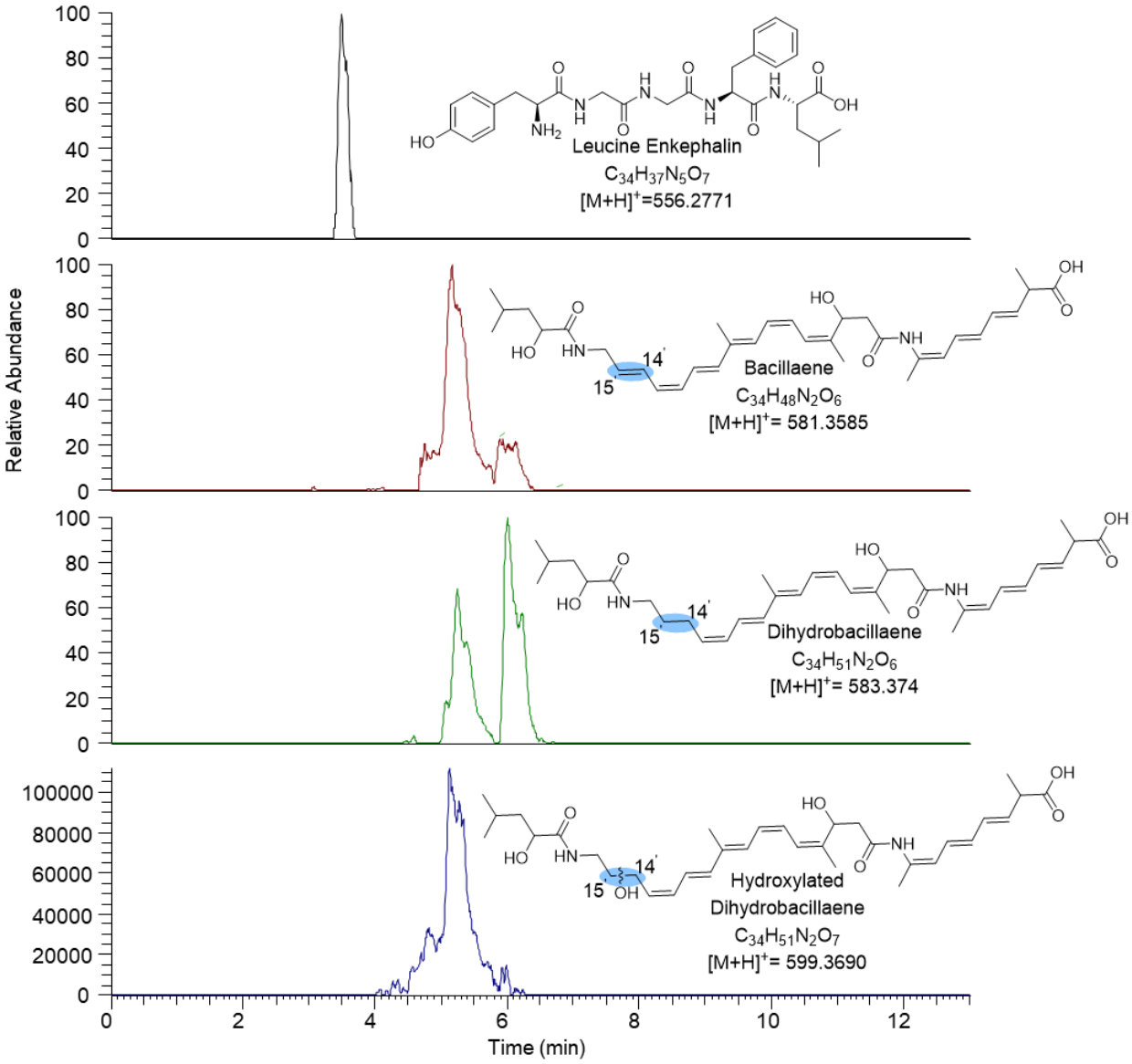
This internal standard would help determine the relative abundance of the bacillaenes. As leucine enkephalin is kept at the same concentration and peak area, it can be compared to the change in peak areas of the three bacillaenes.

Figure 56. Crude 300 µL natural product extract.



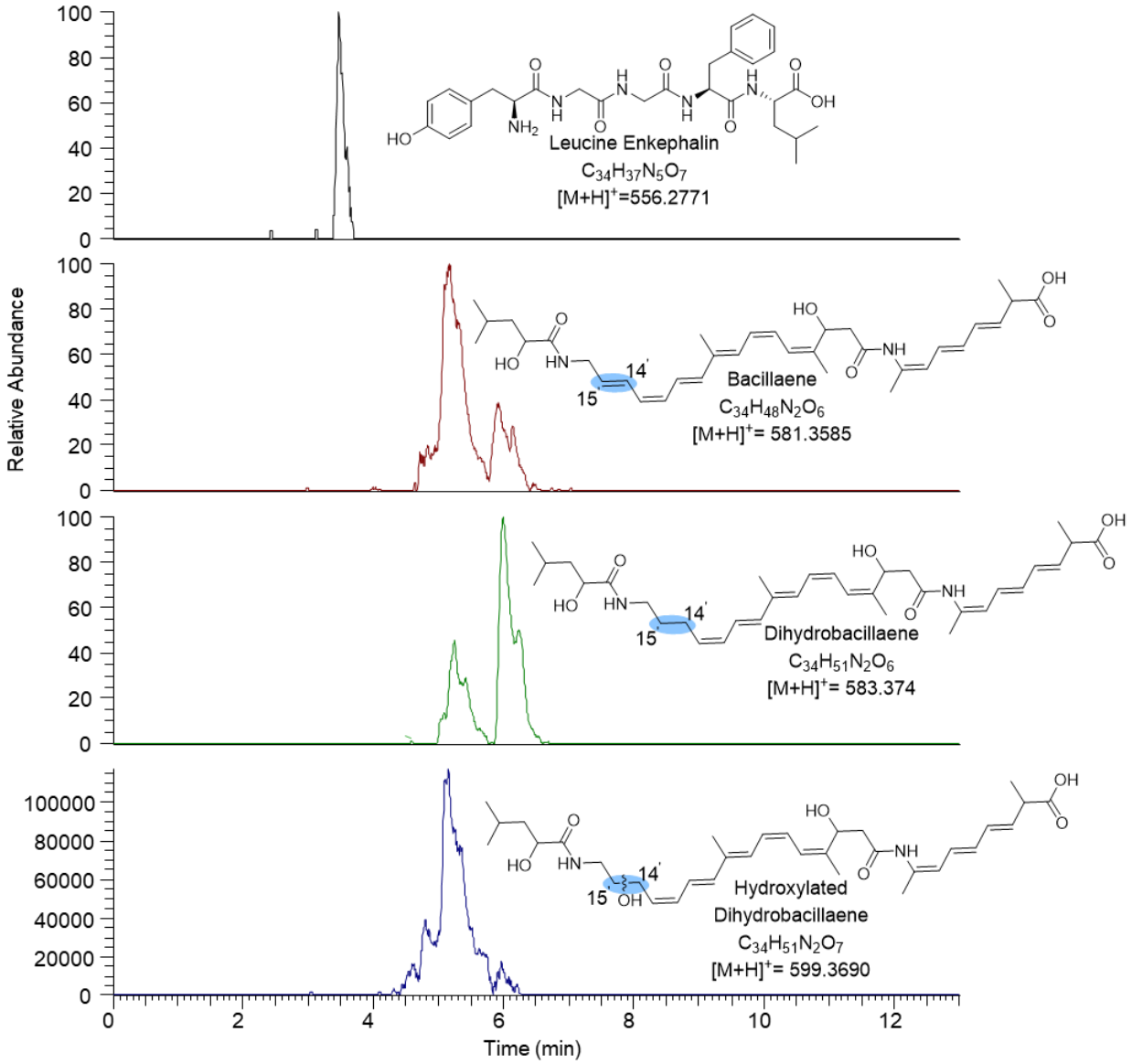
Control samples of the crude natural product extract were done that included various amounts of the crude natural product extract in the buffer. This sample included 300 µL of the extract with 500 µL of the potassium phosphate buffer.

Figure 57. Crude 500 µL natural product extract.



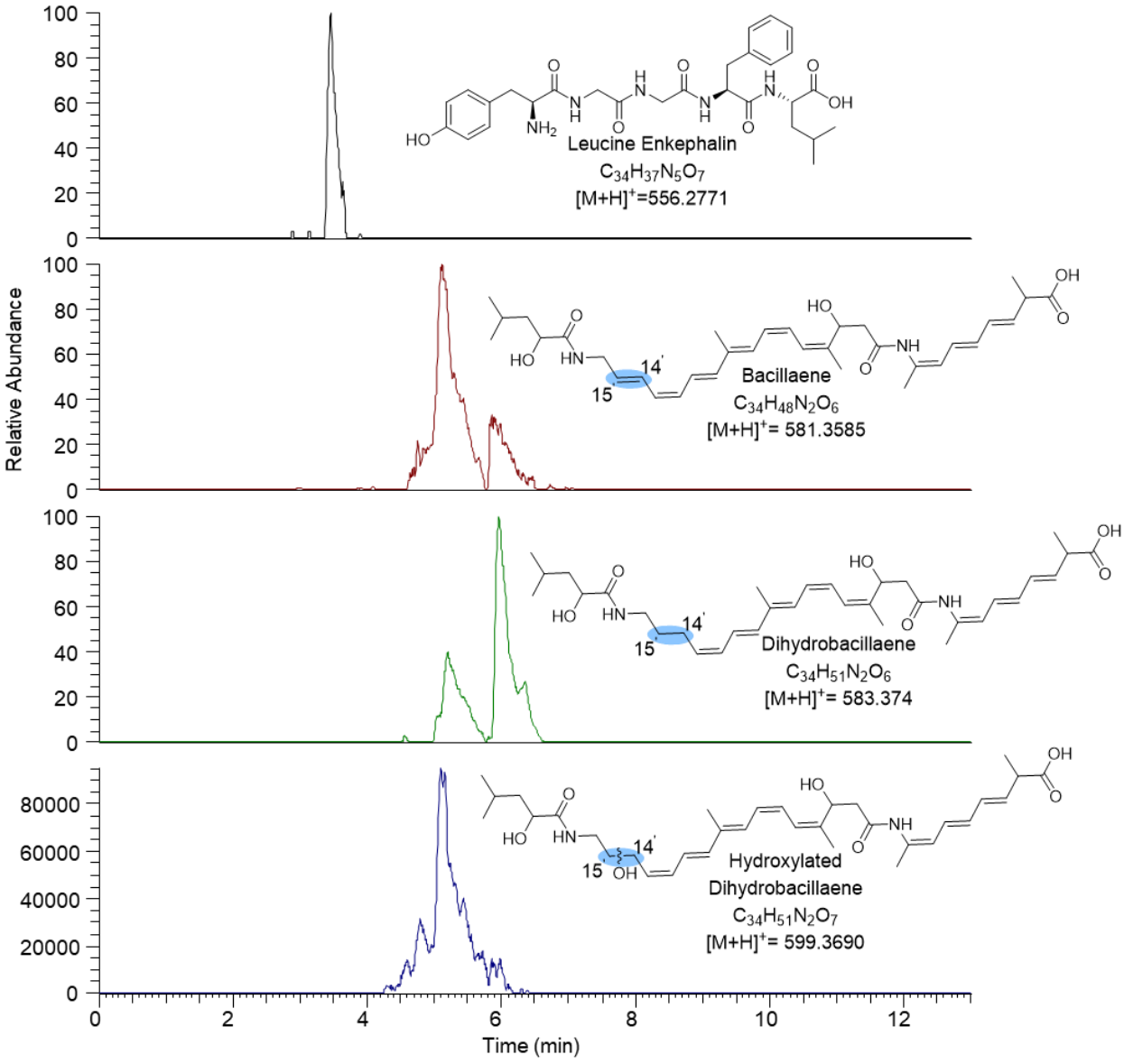
Control samples of the crude natural product extract were done that included various amounts of the crude natural product extract in the buffer. This sample included 500 µL of the extract with 300 µL of the potassium phosphate buffer.

Figure 58. Crude natural product extract with the crude natural product reductase and NADPH.



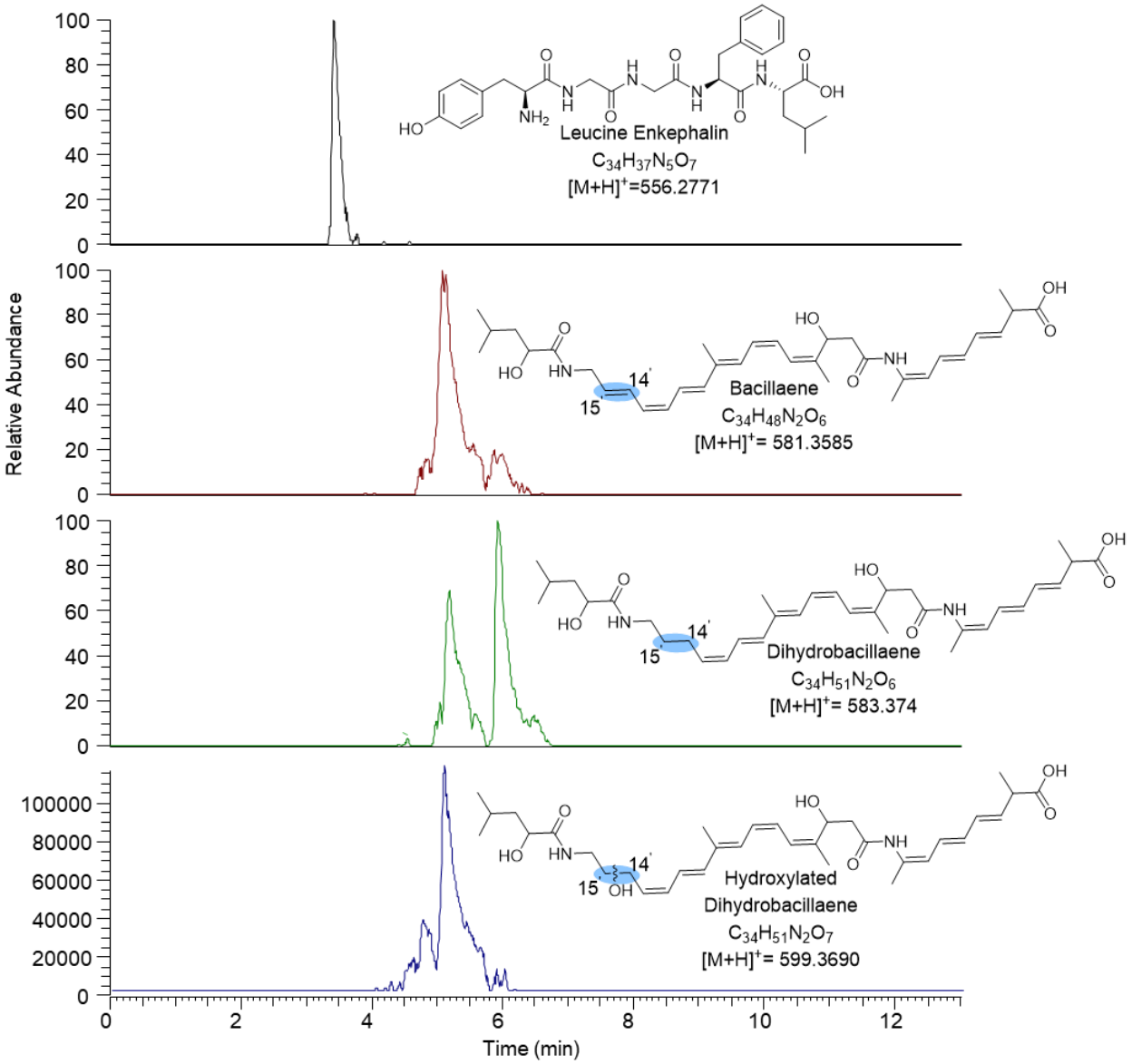
Controls were done that lacked each individual component of the full reaction. This chromatogram shows the control of the natural product extract with the reductase and NADPH. This lacks the purified PksS that will react with the dihydrobacillaene. This was done to ensure that differences seen in the full reaction could be contributed to all parts of the reaction, specifically the enzyme, and not something else.

Figure 59. Crude natural product extract with the crude natural product reductase.



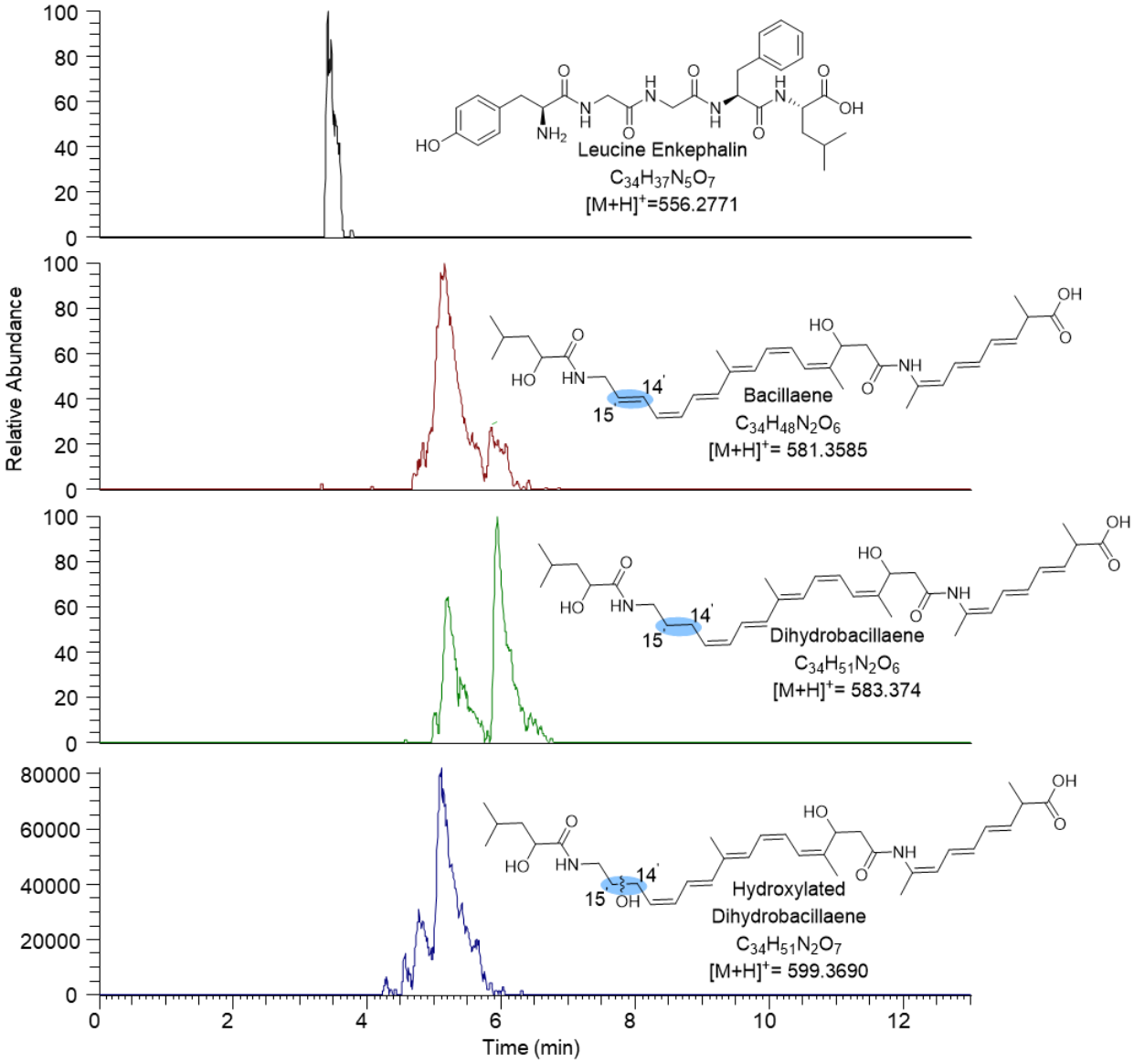
Controls were done that lacked each individual component of the full reaction. This chromatogram shows the control of the natural product extract with the reductase. This lacks the purified PksS that will react with the dihydrobacillaene and NAD(P)H. The reductase will depend on the NAD(P)H for the active cycle of the enzyme. This was done to ensure that differences seen in the full reaction could be contributed to all parts of the reaction, specifically the enzyme, and not something else.

Figure 60. Crude natural product extract with purified PksS.



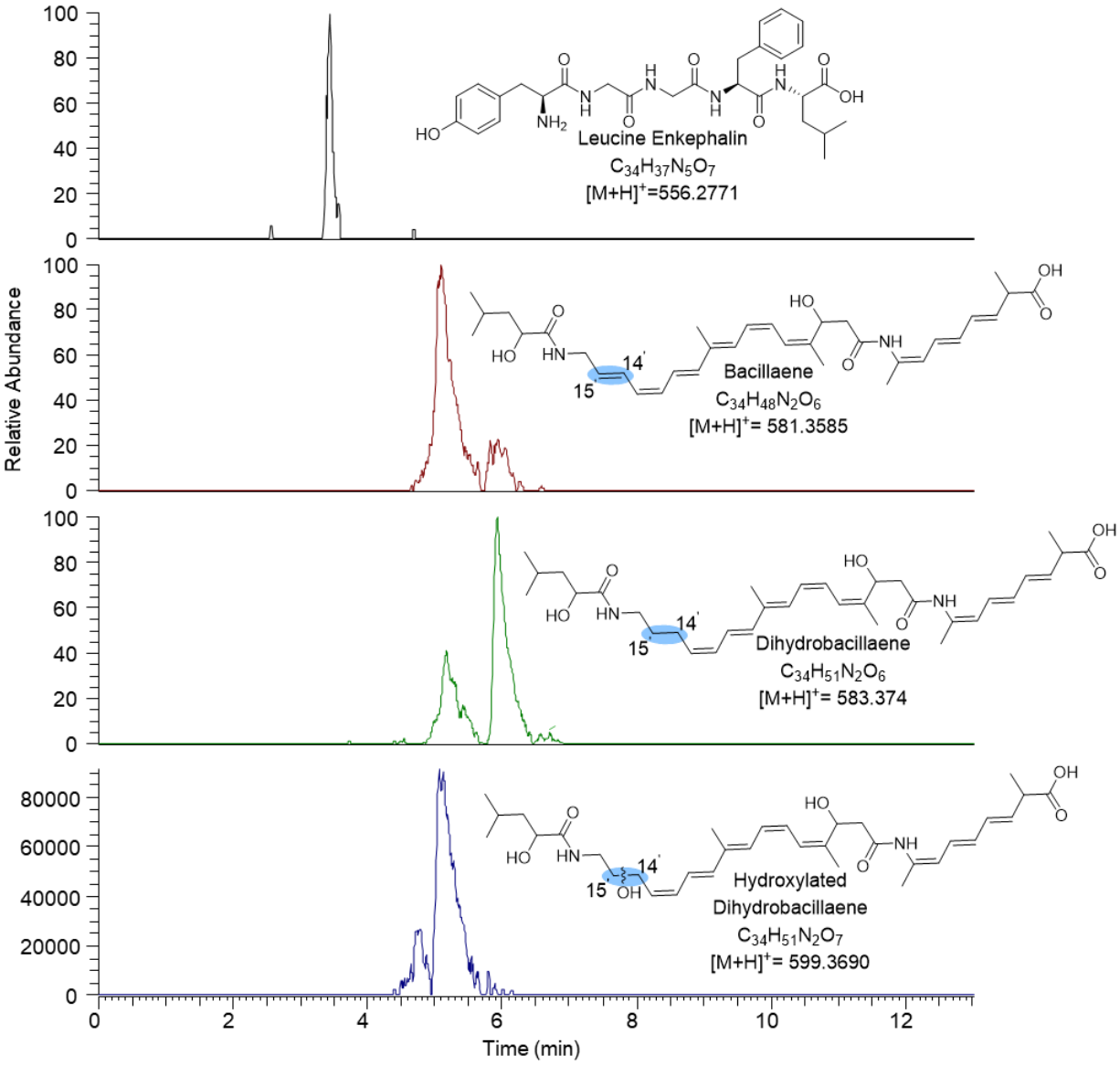
Controls were done that lacked each individual component of the full reaction. This chromatogram shows the control of the natural product extract with the purified enzyme, PksS. The enzyme should need the reductase and NAD(P)H to function. Without those components added to the reaction, no reaction should occur.

Figure 61. Crude natural product extract with purified PksE.



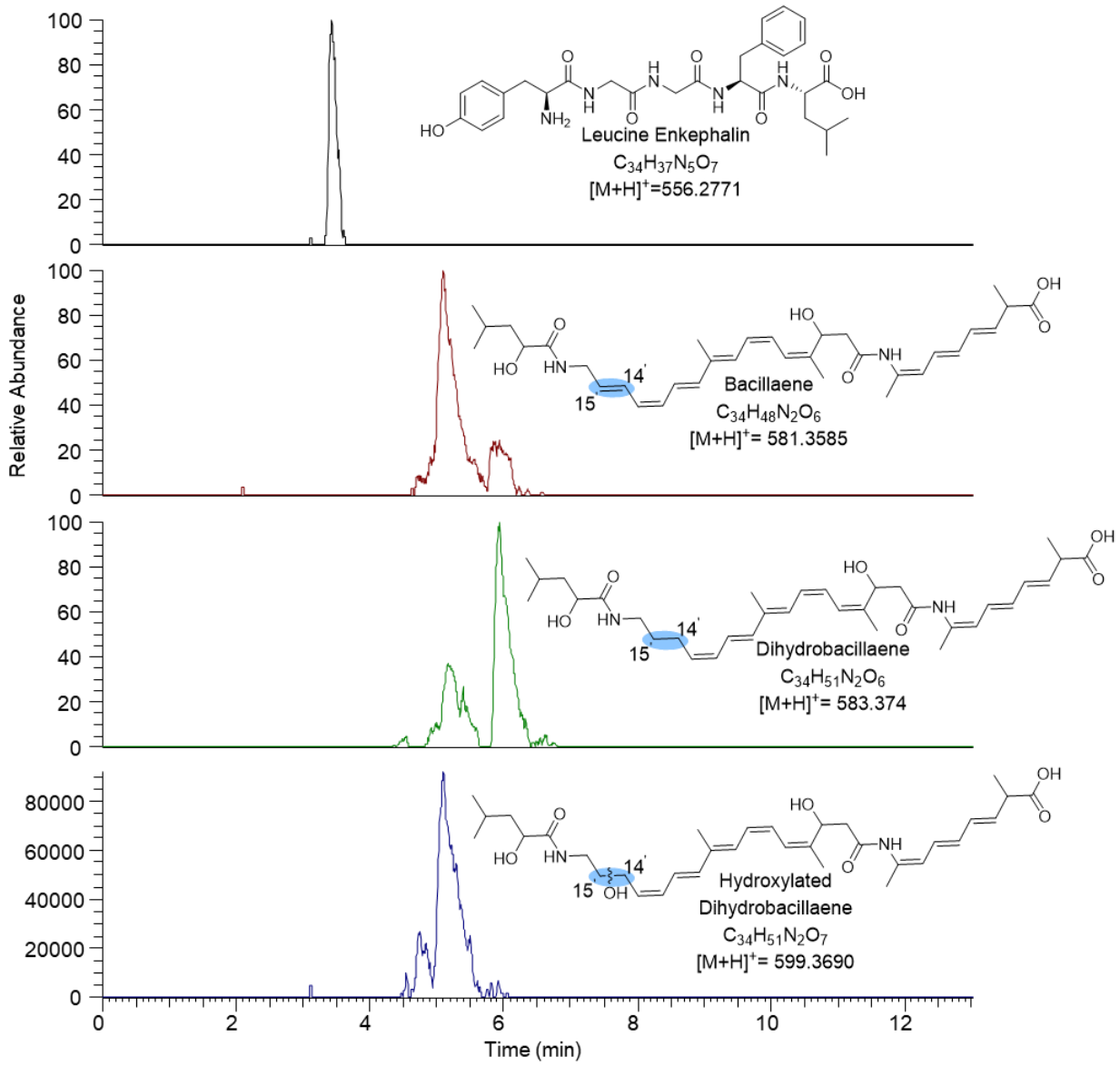
The PksE enzyme should unsaturated the bond between the 14'-15 carbons. The hypothesis was that adding purified PksE would increase the amount of dihydrobacillaene. This would be the case if PksE could act on the full compound and not just the intermediates bound of the enzymes during biosynthesis.

Figure 62. Full reaction with NADH.



The extracted ion chromatograms of the full reaction with all the components are shown above looking at each bacillaene compound. This reaction contained the crude natural product extract, crude natural product reductase, purified enzyme, and NADH as the cofactor. Without knowing the reductase for PksS, we cannot know which cofactor is used, so reactions were ran using different ones to determine which cofactor.

Figure 63. Full reaction with NADPH.



The extracted ion chromatograms of the full reaction with all the components are shown above looking at each bacillaene compound. This reaction contained the crude natural product extract, crude natural product reductase, purified enzyme, and NADPH as the cofactor. Without knowing the reductase for PksS, we cannot know which cofactor is used, so reactions were ran using different ones to determine which cofactor.

Figure 64. Mass spectrum of bacillaene.

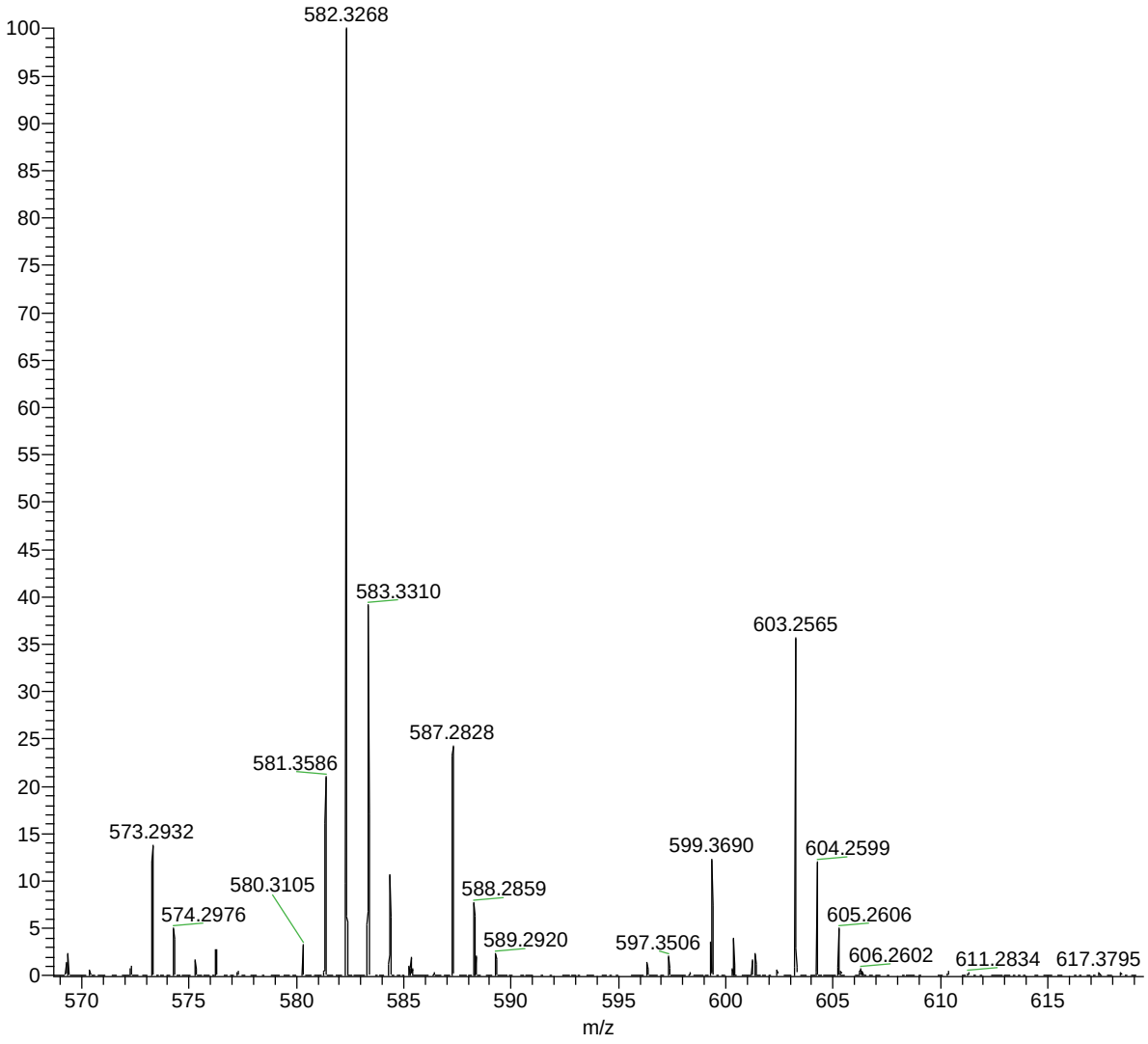


Figure 65. Dihydrobacillaene mass spectrum.

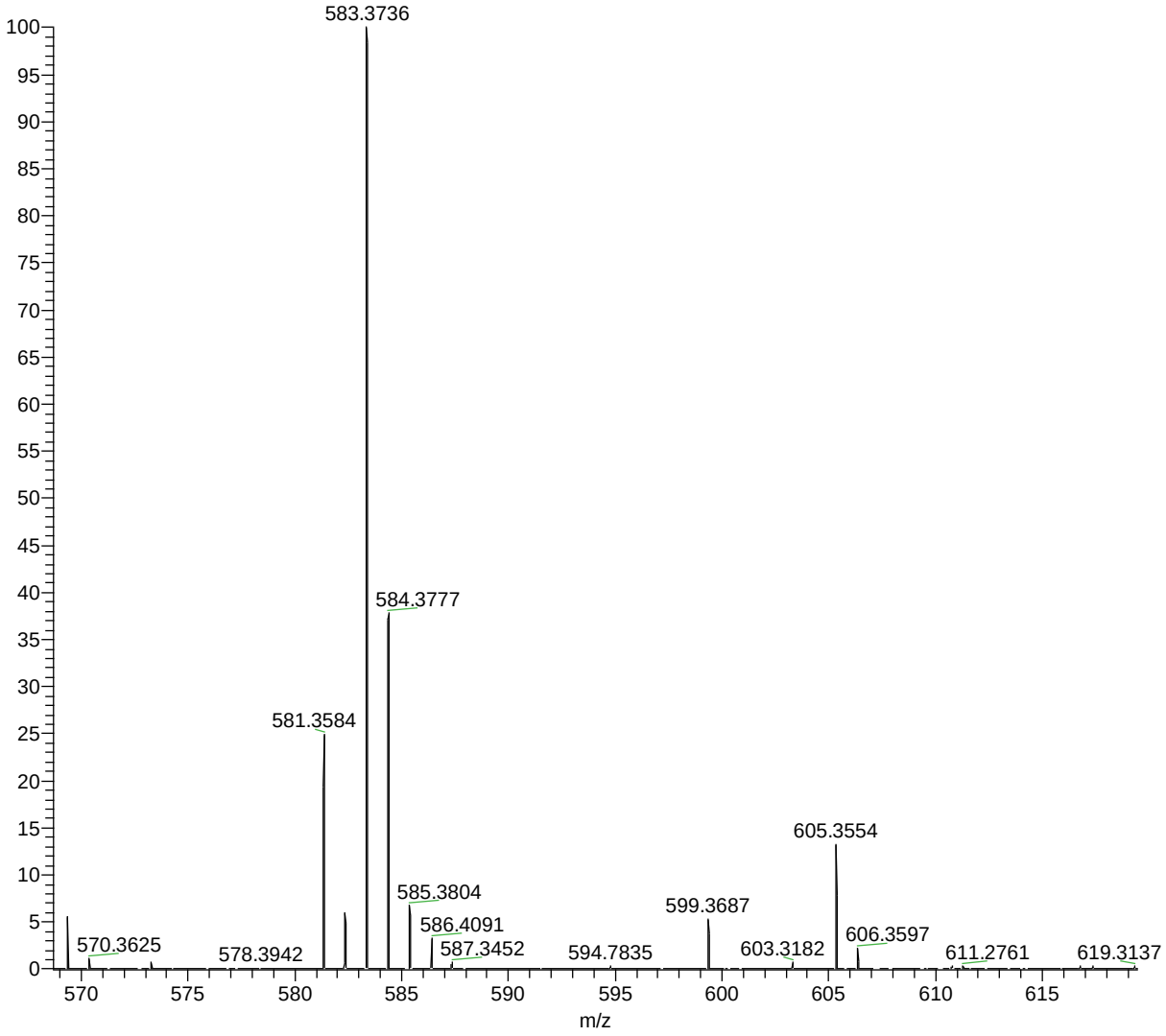


Figure 66. Hydroxylated dihydrobacillaene mass spectrum.

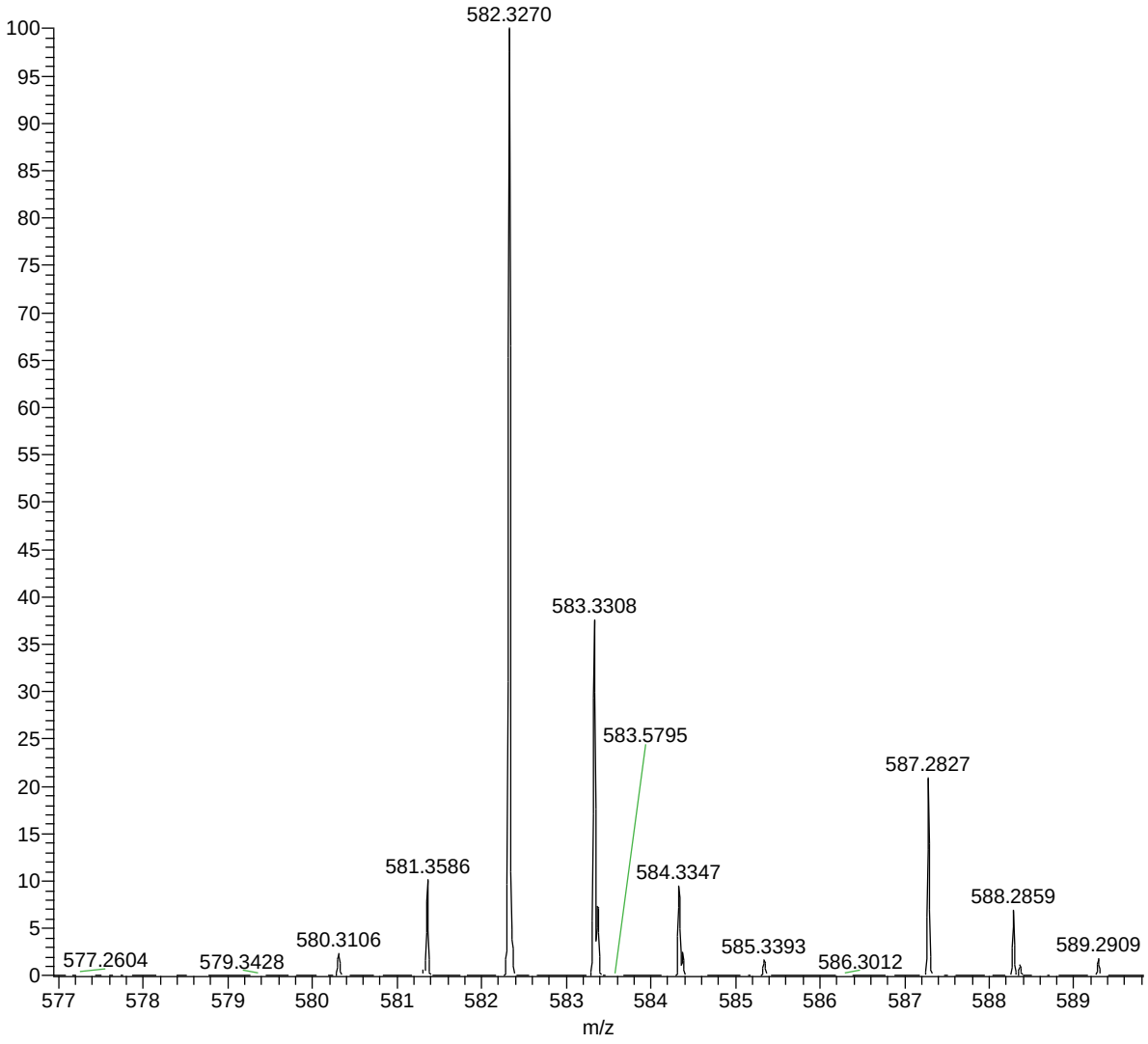
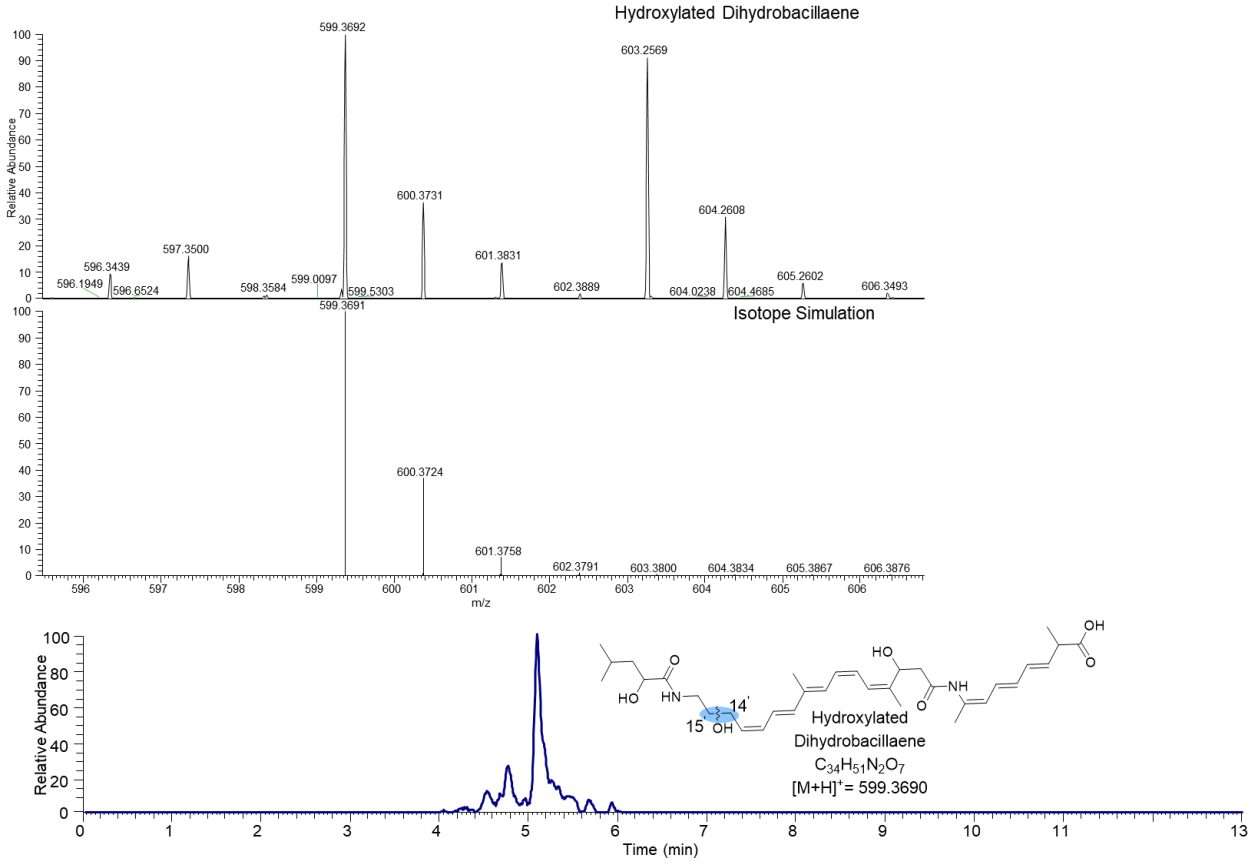
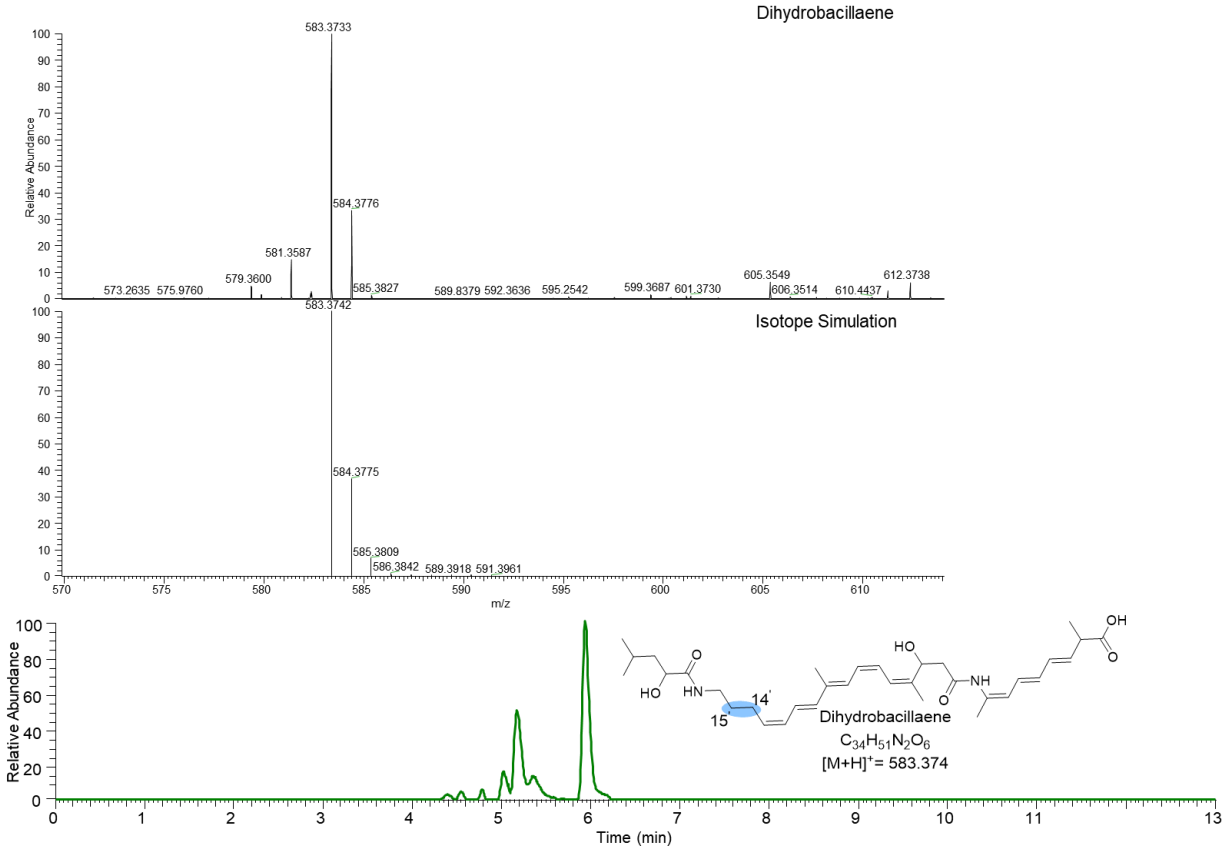


Figure 67. Predicted isotope of hydroxylated dihydrobacillaene.



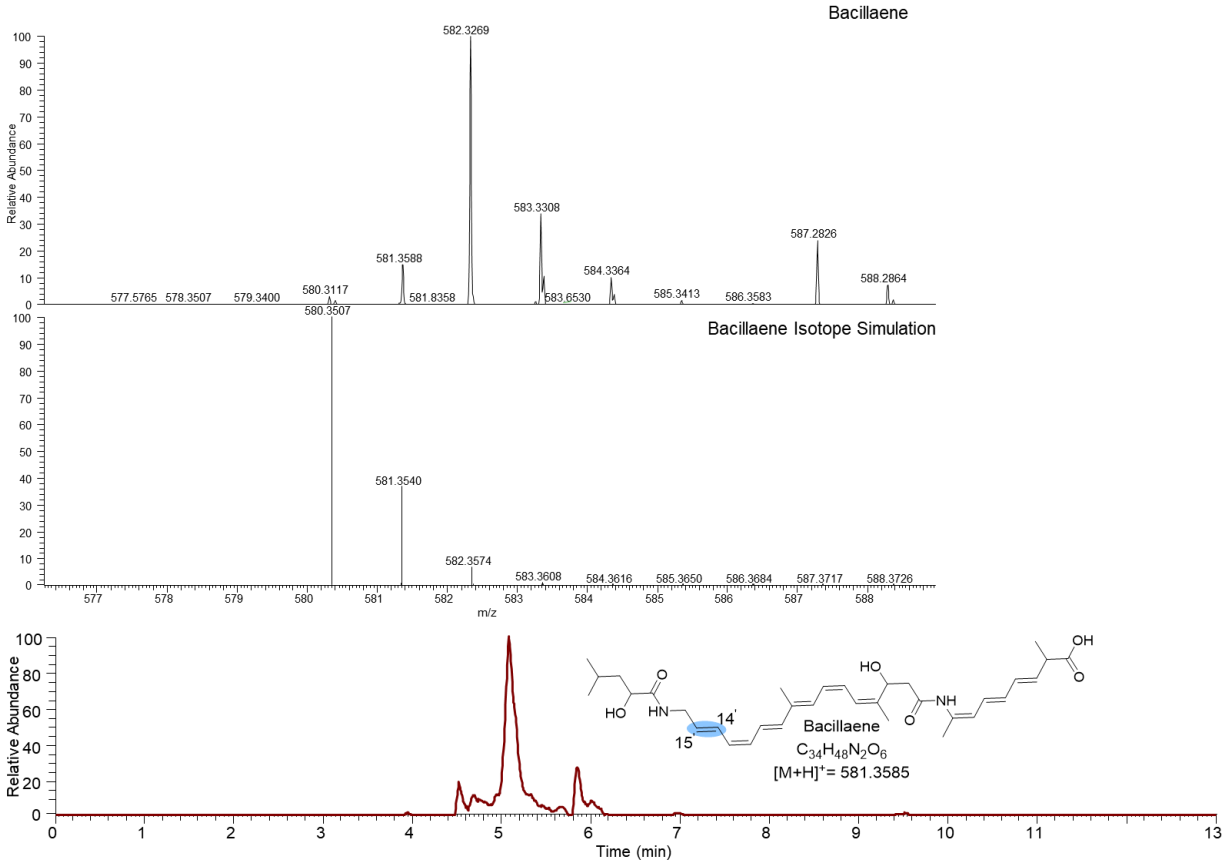
To determine if hydroxylated dihydrobacillaene was in fact being produced by the bacteria, without commercially available standards, isotope simulations were performed. The isotope pattern of the extract along with the simulation is shown with the chromatogram. The isotope simulation confirmed the predicted peak to be hydroxylated dihydrobacillaene.

Figure 68. Predicted isotope of dihydrobacillaene.



To determine if dihydrobacillaene was in fact being produced by the bacteria, without commercially available standards, isotope simulations were performed. The isotope pattern of the extract along with the simulation is shown with the chromatogram. The isotope simulation confirmed the predicted peak to be dihydrobacillaene.

Figure 69. Predicted isotope of bacillaene.



To determine if bacillaene was in fact being produced by the bacteria, without commercially available standards, isotope simulations were performed. The isotope pattern of the extract along with the simulation is shown with the chromatogram. The isotope simulation did not confirm the predicted peak to be bacillaene.

REFERENCES

(1) Schallmey, M.; Singh, A.; Ward, O. P.; Schallmey, M.; Singh, A.; Ward, O. P. Developments in the Use of Bacillus Species for Industrial Production. *Canadian Journal of Microbiology* **2004**, *50* (1), 1–17.

(2) Woldemariamyohannes, K.; Wan, Z.; Yu, Q.; Li, H.; Wei, X.; Liu, Y.; Wang, J.; Sun, B. Prebiotic, Probiotic, Antimicrobial, and Functional Food Applications of Bacillus Amyloliquefaciens. *Journal of Agricultural and Food Chemistry* **2020**, *68* (50), 14709–14727.

(3) Earl, A. M.; Losick, R.; Kolter, R. Ecology and Genomics of Bacillus Subtilis. *Trends Microbiol* **2008**, *16* (6), 269–275.

(4) Crits-Christoph, A.; Diamond, S.; Butterfield, C. N.; Thomas, B. C.; Banfield, J. F. Novel Soil Bacteria Possess Diverse Genes for Secondary Metabolite Biosynthesis. *Nature* **2018**, *558* (7710), 440–444.

(5) Riley, E. P.; Schwarz, C.; Derman, A. I.; Lopez-Garrido, J. Milestones in Bacillus Subtilis Sporulation Research. *Microbial Cell* **2021**, *8* (1), 1–16.

(6) Parte, A. C.; Carbasse, J. S.; Meier-Kolthoff, J. P.; Reimer, L. C.; Göker, M. List of Prokaryotic Names with Standing in Nomenclature (LPSN) Moves to the DSMZ. *Int. J. Syst. Evol. Microbiol* **2020**, *70*, 5607–5612.

(7) Stragier, P.; Losick, R. MOLECULAR GENETICS OF SPORULATION IN BACILLUS SUBTILIS. *Annu Rev Genet* **1996**, *30* (1), 297–341.

(8) Higgins, D.; Dworkin, J. Recent Progress in Bacillus Subtilis Sporulation. *FEMS Microbiol Rev* **2012**, *36*, 131–148.

(9) de Hoon, M. J. L.; Eichenberger, P.; Vitkup, D. Hierarchical Evolution of the Bacterial Sporulation Network. *Current Biology* **2010**, *20* (17), R735–R745.

(10) Phillips, Z. E. v; Strauch, M. A. Bacillus Subtilis Sporulation and Stationary Phase Gene Expression. *Cellular and Molecular Life Sciences* **2002**, *59*, 392–402.

(11) Setlow, P. Spores of Bacillus Subtilis: Their Resistance to and Killing by Radiation, Heat and Chemicals. *J Appl Microbiol* **2006**, *101*, 514–525.

(12) Koburger, T.; Weibezahn, J.; Bernhardt, J.; Homuth, G.; Hecker, M. Genome-Wide MRNA Profiling in Glucose Starved Bacillus Subtilis Cells. *Molecular Genetics and Genomics* **2005**, *274*, 1–12.

(13) Zheng, C.; Yu, Z.; Du, C.; Gong, Y.; Yin, W.; Li, X.; Li, Z.; Römling, U.; Chou, S. H.; He, J. 2-Methylcitrate Cycle: A Well-Regulated Controller of Bacillus Sporulation. *Environ Microbiol* **2020**, *22* (3), 1125–1140.

(14) Ireton, K.; Jin, S.; Grossman, A. D.; Sonenshein, A. L. Krebs Cycle Function Is Required for Activation of the Spo0A Transcription Factor in Bacillus Subtilis. *Proc Natl Acad Sci U S A* **1995**, *92* (7), 2845.

(15) Reddick, J. J.; Sirkisoon, S.; Dahal, R. A.; Hardesty, G.; Hage, N. E.; Booth, W. T.; Quattlebaum, A. L.; Mills, S. N.; Meadows, V. G.; Adams, S. L. H.; Doyle, J. S.; Kiel, B. E. First Biochemical Characterization of a Methylcitric Acid Cycle from Bacillus Subtilis Strain 168. *Biochemistry* **2017**, *56* (42), 5698–5711.

(16) Textor, S.; Wendisch, V. F.; de Graaf, A. A.; Muller, U.; Linder, M. I.; Linder, D.; Buckel, W. Propionate Oxidation in Escherichia Coli: Evidence for Operation of a Methylcitrate Cycle in Bacteria. *Arch Microbiol* **1997**, *168*, 428–436.

(17) Reddick, J. J.; Williams, J. K. The MmgA Gene from Bacillus Subtilis Encodes a Degradative Acetoacetyl-CoA Thiolase. *Biotechnol Lett* **2008**, *30* (6), 1045–1050.

(18) Vegunta, Y. Biochemical Characterization of MmgB, a Gene Encoding a 3-Hydroxybutyryl-CoA Dehydrogenase from Bacillus Subtilis168 and Genetic Evidence for the Methylcitric Acid Cycle in Bacillus Subtilis 168. Thesis, UNCG, Greensboro, NC, 2011.

(19) Russell, S. A. OVEREXPRESSION, PURIFICATION, AND CHARACTERIZATION OF MmgB AND MmgC FROM BACILLUS SUBTILIS STRAIN 168. Thesis, UNCG, Greensboro, NC, 2008.

(20) Dubey, M. K.; Broberg, A.; Funck Jensen, D.; Karlsson, M.; Mukesh Dubey, C. K. Role of the Methylcitrate Cycle in Growth, Antagonism and Induction of Systemic Defence Responses in the Fungal Biocontrol Agent Trichoderma Atroviride. *Microbiology (N Y)* **2013**, *159*, 2492–2500.

(21) Acharya, R. OVEREXPRESSION, PURIFICATION, AND CHARACTERIZATION OF MmgD. Thesis, UNCG, Greensboro, NC, 2009.

(22) Hardesty, G. The Characterization of MmgE from Bacillus Subtilis. Thesis, UNCG, Greensboro, NC, 2012.

(23) Sirkisoon, S. R. Characterization of CitB in Methylcitric Acid Cycle of Bacillus Subtilis 168 and Characterization of Antimicrobial Activity in the Mucosal Epithelial Layer and Gill Tissue of Largmouth Bass (Micropterus Salmoides). Thesis, UNCG, Greensboro, NC, 2014.

(24) Booth, II, W. T. Characterization of the Biochemical Activity of the Open Reading Frame "YQIQ" of Bacillus Subtilis. Thesis, UNCG, Greensboro, NC, 2011.

(25) Clardy, J.; Walsh, C. Lessons from Natural Molecules. *Nature* **2004**, *432* (7019), 829–837.

(26) Cragg, G. M.; Newman, D. J. Natural Products: A Continuing Source of Novel Drug Leads. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830* (6), 3670–3695.

(27) Chen, H.; Du, L. Iterative Polyketide Biosynthesis by Modular Polyketide Synthases in Bacteria. *Appl Microbiol Biotechnol* **2016**, *100* (2), 541–557.

(28) Shen, B. Polyketide Biosynthesis beyond the Type I, II and III Polyketide Synthase Paradigms. *Curr Opin Chem Biol* **2003**, *7* (2), 285–295.

(29) Walker, P. D.; Weir, A. N. M.; Willis, C. L.; Crump, M. P. Polyketide β-Branching: Diversity, Mechanism and Selectivity. *Nat Prod Rep* **2021**, *38*, 723–756.

(30) N Helfrich, E. J.; Ueoka, R.; Chevrette, M. G.; Hemmerling, F.; Lu, X.; Leopold-Messer, S.; Minas, H. A.; Burch, A. Y.; Lindow, S. E.; Piel, J.; Medema, M. H. Evolution of Combinatorial Diversity in Trans-Acyltransferase Polyketide Synthase Assembly Lines across Bacteria. *Nature Communications* **2021**, *12* (1), 1422.

(31) Donadio, S.; Katz, L. Organization of the Enzymatic Domains in the Multifunctional Polyketide Synthase Involved in Erythromycin Formation in Saccharopolyspora Erythraea. *Gene* **1992**, *111* (1), 51–60.

(32) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of Saccharopolyspora Erythraea. *Nature* **1990**, *348* (6297), 176–178.

(33) Keatinge-Clay, A. Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *J Mol Biol* **2008**, *384* (4), 941–953.

(34) Wagner, D. T.; Zeng, J.; Bailey, C. B.; Gay, D. C.; Yuan, F.; Manion, H. R.; Keatinge-Clay, A. T. Structural and Functional Trends in Dehydrating Bimodules from Trans-Acyltransferase Polyketide Synthases. *Structure* **2017**, *25* (7), 1045-1055.e2.

(35) Helfrich, E. J. N.; Piel, J. Biosynthesis of Polyketides by Trans-AT Polyketide Synthases. *Nat Prod Rep* **2016**, *33* (2), 231–316.

(36) Lamb, D. C.; Waterman, M. R. Unusual Properties of the Cytochrome P450 Superfamily. *Philosophical Transactions of the Royal Society* **2013**, *368* (1612).

(37) Prier, C. K.; Arnold, F. H. Chemomimetic Biocatalysis: Exploiting the Synthetic Potential of Cofactor-Dependent Enzymes to Create New Catalysts. *J Am Chem Soc* **2015**, *137* (44), 13992–14006.

(38) Jung, S. T.; Lauchli, R.; Arnold, F. H. Cytochrome P450: Taming a Wild Type Enzyme. *Current Opinion in Biotechnology* **2011**, *22*, 809–817.

(39) Omura, T.; Sato, R. The Carbon Monoxide-Binding Pigment of Liver Microsomes I. EVIDENCE FOR ITS HEMOPROTEIN NATURE. *J Biol Chem* **1964**, *239* (7).

(40) Vargas-Bautista, C.; Rahlwes, K.; Straight, P. Bacterial Competition Reveals Differential Regulation of the Pks Genes by Bacillus Subtilis. *J Bacteriol* **2014**, *196* (4), 717–728.

(41) Müller, S.; Strack, S. N.; Ryan, S. E.; Kearns, D. B.; Kirby, J. R. Predation by Myxococcus Xanthus Induces Bacillus Subtilis to Form Spore-Filled Megastructures. *Appl Environ Microbiol* **2015**, *81* (1), 203–210.

(42) Müller, S.; Strack, S. N.; Hoefler, B. C.; Straight, P. D.; Kearns, D. B.; Kirby, J. R. Bacillaene and Sporulation Protect Bacillus Subtilis from Predation by Myxococcus Xanthus. *Appl Environ Microbiol* **2014**, *80* (18), 5603–5610.

(43) Chen, X. H.; Scholz, R.; Borriss, M.; Junge, H.; Mögel, G.; Kunz, S.; Borriss, R. Difficidin and Bacilysin Produced by Plant-Associated Bacillus Amyloliquefaciens Are Efficient in Controlling Fire Blight Disease. *J Biotechnol* **2009**, *140* (1–2), 38–44.

(44) Stein, T. Bacillus Subtilis Antibiotics: Structures, Syntheses and Specific Functions. *Mol Microbiol* **2005**, *56* (4), 845–857.

(45) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus Subtilis Phosphopantetheinyl Transferase for Peptidyl Carder Protein Domains in Peptide Synthetases. *Biochemistry* **1998**, *37* (6), 1585–1595.

(46) Kunst, F. The Complete Genome Sequence of the Gram-Positive Bacterium Bacillus Subtilis. *Nature* **1997**, *390*, 249–255.

(47) al Bertini, A. M.; Caramori, T.; Scoffone, F.; Scottit, C.; Galizzi, A.; Albertini, A. M. Sequence around the 159 Region of the Bacillus Subtilis Genome: The PksX Locus Spans 33.6 Kb. *Microbiology (N Y)* **1995**, *141*, 299–309.

(48) Mayerl, F.; Fisher, S.; Pirnik, D.; Aklonis, C.; Dean, L.; Meyers, E.; Fernandes, P. Bacillaene, a Novel Inhibitor of Procaryotic Protein Synthesis Produced by Bacillus Subtilis: Production, Taxonomy, Isolation, Physico-Chemical Characterization and Biological Activity. *J Antibiot (Tokyo)* **1995**, *48* (9), 997–1003.

(49) Butcher, R. A.; Schroeder, F. C.; Fischbach, M. A.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Clardy, J. The Identification of Bacillaene, the Product of the PksX Megacomplex in Bacillus Subtilis. *Proc Natl Acad Sci U S A* **2007**, *104* (5), 1506–1509.

(50) Reddick, J. J.; Antolak, S. A.; Raner, G. M. PksS from Bacillus Subtilis Is a Cytochrome P450 Involved in Bacillaene Metabolism. *Biochem Biophys Res Commun* **2007**, *358* (1), 363–367.

(51) Moldenhauer, J.; Chen, X. H.; Borriss, R.; Piel, J. Biosynthesis of the Antibiotic Bacillaene, the Product of a Giant Polyketide Synthase Complex of the Trans-AT Family. *Angewandte Chemie - International Edition* **2007**, *46* (43), 8195–8197.

(52) Bumpus, S. B.; Magarvey, N. A.; Kelleher, N. L.; Walsh, C. T.; Calderone, C. T. Polyunsaturated Fatty-Acid-like Trans-Enoyl Reductases Utilized in Polyketide Biosynthesis. *J Am Chem Soc* **2008**, *130* (35), 11614–11616.

(53) Blake, C.; Christensen, M. N.; Kovács, T.; Kovács, K. Molecular Aspects of Plant Growth Promotion and Protection by Bacillus Subtilis. **2021**, *34* (1).

(54) Steil, L.; Serrano, M.; Henriques, A. O.; Völker, U. Genome-Wide Analysis of Temporally Regulated and Compartment-Specific Gene Expression in Sporulating Cells of Bacillus Subtilis. *Microbiology (N Y)* **2005**, *151* (2), 399–420.

(55) Bryan, E. M.; Beall, B. W.; Moran, C. P. A Sigma E -Dependent Operon Subject to Catabolite Repression during Sporulation in Bacillus Subtilis. *J Bacteriol* **1996**, *178* (16), 4778–4786.

(56) Ireton, K.; Jint, S.; Grossman, A. D.; Sonensheintt, A. L. Krebs Cycle Finction Is Required for Activation of the SpoOA Transcription Factor in Bacillus Subtilis. *Dev Biol* **1995**, *92*, 2845–2849.

(57) Blank, L.; Green, J.; Guest, J. R. AcnC of Escherichia Coli Is a 2-Methylcitrate Dehydratase (PrpD) That Can Use Citrate and Isocitrate as Substrates. *Microbiology (N Y)* **2002**, *148* (1), 133–146.

(58) Tabuchi, T.; Aoki, H.; Uchiyama, H.; Nakahara, T. Agricultural and Biological Chemistry 2-Methylcitrate Dehydratase, a New Enzyme Functioning at the Methylcitric Acid Cycle

of Propionate Metabolism. *Agricultural and Biological Chemistry* **1981**, *45* (12), 2823–2829.

(59) Serio, A. W.; Pechter, K. B.; Sonenshein, A. L. Bacillus Subtilis Aconitase Is Required for Efficient Late-Sporulation Gene Expression. *J Bacteriol* **2006**, *188* (17), 6396–6405.

(60) Krawczyk, H.; Martyniuk, T. Characterisation of the 1H and 13C NMR Spectra of Methylcitric Acid. *Spectrochim Acta A Mol Biomol Spectrosc* **2007**, *67* (2), 298–305.

(61) Adams, S. L. H. Investigation and Comparison of Stereoisomer Products of 2-Methylcitrate Synthase and 2-Methylcitrate Dehydratase in Escherichia Coli Strain K12 and Bacillus Subtilis Strain 168. Thesis, UNCG, Greensboro, NC, 2017.

(62) Horswill, A. R.; Escalante-Semerena, J. C. In Vitro Conversion of Propionate to Pyruvate by Salmonella Enterica Enzymes: 2-Methylcitrate Dehydratase (PrpD) and Aconitase Enzymes Catalyze the Conversion of 2-Methylcitrate to 2-Methylisocitrate. *Biochemistry* **2001**, *40*, 4703–4713.

(63) Reddick, J. J.; Sirkisoon, S.; Dahal, R. A.; Hardesty, G.; Hage, N. E.; Booth, W. T.; Quattlebaum, A. L.; Mills, S. N.; Meadows, V. G.; Adams, S. L. H.; Doyle, J. S.; Kiel, B. E. First Biochemical Characterization of a Methylcitric Acid Cycle from Bacillus Subtilis Strain 168. *Biochemistry* **2017**, *56* (42), 5698–5711.

(64) Traxler, M. F.; Kolter, R. Natural Products in Soil Microbe Interactions and Evolution. *Nat Prod Rep* **2015**, *32* (7), 956–970.

(65) Romero, D.; Traxler, M. F.; López, D.; Kolter, R. Antibiotics as Signal Molecules. *Chem Rev* **2011**, *111* (9), 5492–5505.

(66) Smith, S.; Tsai, S.-C. The Type I Fatty Acid and Polyketide Synthases: A Tale of Two Megasynthases. *Natural Product Reports* **2007**, *24* (5), 1041–1072.

(67) Herbst, D. A.; Townsend, C. A.; Maier, T. The Architectures of Iterative Type I PKS and FAS. *Nat Prod Rep* **2018**, *35*, 1046.

(68) Kornfuehrer, T.; Eustáquio, A. S. Diversification of Polyketide Structures: Via Synthase Engineering. *Medchemcomm* **2019**, *10* (8), 1256–1272.

(69) Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structure of a Modular Polyketide Synthase. *Nature* **2014**, *510* (7506), 512–517.

(70) Klaus, M.; Buyachuihan, L.; Grininger, M. Ketosynthase Domain Constrains the Design of Polyketide Synthases. *ACS Chem Biol* **2020**, *15* (9), 2422–2432.

(71) Keatinge-Clay, A. T. The Structures of Type I Polyketide Synthases. *Nat Prod Rep* **2012**, *29* (10), 1050–1073.

(72) Fischbach, M. A.; Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chemical Reviews* **2006**, *106* (8), 3468–3496.

(73) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angewandte Chemie-International Edition* **2009**, *48* (26), 4688–4716.

(74) Jenner, M. Predicting the Peculiar. *Nat Chem Biol* **2019**, *15* (8), 760–761.

(75) Kosol, S.; Jenner, M.; Lewandowski, J. R.; Challis, G. L. Protein-Protein Interactions in Trans-AT Polyketide Synthases. *Nat Prod Rep* **2018**, *35* (10), 1097–1109.

(76) Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary Implications of Bacterial Polyketide Synthases. *Mol Biol Evol* **2005**, *22* (10), 2027–2039.

(77) Du, L.; Lou, L. PKS and NRPS Release Mechanisms. *Nat Prod Rep* **2010**, *27* (2), 255–278.

(78) Marahiel, M. A. A Structural Model for Multimodular NRPS Assembly Lines. *Natural Product Reports* **2016**, *33*, 136–140.

(79) Beck, C.; Garzón, J. F. G.; Weber, T. Recent Advances in Re-Engineering Modular PKS and NRPS Assembly Lines. *Biotechnology and Bioprocess Engineering* **2020**, *25* (6), 886–894.

(80) Robertsen, H. L.; Musiol-Kroll, E. M. Actinomycete-Derived Polyketides as a Source of Antibiotics and Lead Structures for the Development of New Antimicrobial Drugs. *Antibiotics* **2019**, *8* (4), 157.

(81) Caffrey, P.; Bevitt, D. J.; Staunton, J.; Leadlay, P. F. Identification of DEBS 1, DEBS 2 and DEBS 3, the Multienzyme Polypeptides of the Erythromyein-Producing Polyketide Synthase from Saccharopolyspora Erythraea. *FEBS Letters* **1115**, *304* (2), 225–228.

(82) Rawlings, B. J. Type I Polyketide Biosynthesis in Bacteria (Part A — Erythromycin Biosynthesis) (1994 to 2000). *Nat Prod Rep* **2001**, *18* (2), 190–227.

(83) Staunton, J.; Wilkinson, B. Biosynthesis of Erythromycin and Rapamycin. *Chemical Reviews* **1997**, *97*, 2611–2629.

(84) Cane, D. E. Programming of Erythromycin Biosynthesis by a Modular Polyketide Synthase. *Journal of Biological Chemistry* **2010**, *285* (36), 27517–27523.

(85) Zhang, H.; Wang, Y.; Wu, J.; Skalina, K.; Pfeifer, B. A. Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. Coli as a Heterologous Host. *Chem Biol* **2010**, *17* (11), 1232–1240.

(86) Rix, U.; Fischer, C.; Remsing, L. L.; Rohr, J. Modification of Post-PKS Tailoring Steps through Combinatorial Biosynthesis. *Nat Prod Rep* **2002**, *19* (5), 542–580.

(87) Liu, X.-J.; Kong, R.-X.; Niu, M.-S.; Qiu, R.-G.; Tang, L. Identification of the Post-Polyketide Synthase Modification Enzymes for Fostriecin Biosynthesis in Streptomyces Pulveraceus. *J. Nat. Prod* **2013**, *76*, 529.

(88) Gober, R.; Wheeler, R.; Rohr, J. Post-PKS Enzyme Complexes. *Medchemcomm* **2019**, *10* (11), 1855–1866.

(89) Heinrich, S.; Grote, M.; Sievers, S.; Kushnir, S.; Schulz, F. Polyether Cyclization Cascade Alterations in Response to Monensin Polyketide Synthase Mutations. *ChemBioChem* **2022**, *23* (2), 1–8.

(90) Guengerich, F. P. Mechanisms of Cytochrome P450-Catalyzed Oxidations. *ACS Catal* **2018**, *8*, 10964–10976.

(91) Guengerich, F. P. A History of the Roles of Cytochrome P450 Enzymes in the Toxicity of Drugs. *Toxicol Res* **2021**, *37* (1), 1–23.

(92) Peter Guengerich, F. Cytochrome P450 Research and The Journal of Biological Chemistry. *Journal of Biological Chemistry* **2019**, *294* (5), 1671–1680.

(93) Guengerich, F. P.; Waterman, M. R.; Egli, M. Recent Structural Insights into Cytochrome P450 Function. *Trends Pharmacol Sci* **2016**, *37* (8), 625–640.

(94) Chuo, S.-W.; Liou, S.-H.; Wang, L.-P.; David Britt, R.; L. Poulos, T.; F. Sevrioukova, I.; B. Goodin, D. Conformational Response of N-Terminally Truncated Cytochrome P450 3A4 to Ligand Binding in Solution. *Biochemistry* **2019**, *58* (37), 3903–3910.

(95) Sellner, M.; Fischer, A.; Don, C. G.; Smieško, M. Supporting Information: Conformational Landscape of Cytochrome P450 Reductase Interactions. *International Journal of Molecular Sciences* **2021**, *22* (3), 1–14.

(96) Dutta Dubey, K.; Shaik, S. Cytochrome P450□The Wonderful Nanomachine Revealed through Dynamic Simulations of the Catalytic Cycle. *Acc. Chem. Res.* **2019**, *52*, 389–399.

(97) Raju, D. R.; Kumar, A.; BK, N.; Shetty, A.; PS, A.; Kumar, R. P.; Lalitha, R.; Sigamani, G. Extensive Modelling and Quantum Chemical Study of Sterol C-22 Desaturase Mechanism: A Commercially Important Cytochrome P450 Family. *Catal Today* **2022**, *397–399*, 50–62.

(98) Guengerich, F. P.; Isin, E. M. Mechanisms of Cytochrome P450 Reactions. *Acta Chim Slov* **2008**, *55* (1), 7–19.

(99) Shaik, S.; Dubey, K. D. The Catalytic Cycle of Cytochrome P450: A Fascinating Choreography. *Trends Chem* **2021**, *3* (12), 1027–1044.

(100) Ji, L.; Faponle, A. S.; Quesne, M. G.; Sainna, M. A.; Zhang, J.; Franke, A.; Kumar, D.; van Eldik, R.; Liu, W.; de Visser, S. P. Drug Metabolism by Cytochrome P450 Enzymes: What Distinguishes the Pathways Leading to Substrate Hydroxylation over Desaturation? *Chemistry - A European Journal* **2015**, *21* (25), 9083–9092.

(101) Tian, S.; Jiang, L.; Gao, • Qiang; Zhang, J.; Zong, • Mei; Zhang, H.; Ren, Y.; Guo, S.; Gong, • Guoyi; Liu, • Fan; Xu, Y. Efficient CRISPR/Cas9-Based Gene Knockout in Watermelon. *Plant Cell Rep* **2017**, *36*, 399–406.

(102) Liu, P.; Jenkins, N. A.; Copeland, N. G. A Highly Efficient Recombineering-Based Method for Generating Conditional Knockout Mutations. *Genome Res* **2003**, *13* (3), 476–484.

(103) Vagner, V.; Dervyn, E.; Ehrlich, S. D. A Vector for Systematic Gene Inactivation in Bacillus Subtilis. *Microbiology (N Y)* **1998**, *144* (11), 3097–3104.

(104) Mortensen, R. M. Double Knockouts: Production of Mutant Cell Lines in Cardiovascular Research. *Hypertension* **1993**, *22* (4), 646–651.

(105) Bec, S.; Yulfo-Soto, G. E.; Vaillancourt, L. J. Relative Efficiency of Split-Marker versus Double-Crossover Relative Efficiency of Split-Marker versus Double-Crossover Replacement Protocols for Production of Deletion Mutants in Strain Replacement Protocols for Production of Deletion Mutants in Strain PH-1 of Fusarium Graminearum PH-1 of Fusarium Graminearum. *Fungal Genetics Reports* **2021**, *65* (1), 1–18.

(106) Jiang, F.; Doudna, J. A. CRISPR-Cas9 Structures and Mechanisms. *Annu Rev Biophys* **2017**, *46*, 505–529.

(107) Khanzadi, M. N.; Khan, A. A. CRISPR/Cas9: Nature's Gift to Prokaryotes and an Auspicious Tool in Genome Editing. *J Basic Microbiol* **2020**, *60* (2), 91–102.

(108) Bortesi, L.; Fischer, R. The CRISPR/Cas9 System for Plant Genome Editing and Beyond. *Biotechnol Adv* **2015**, *33* (1), 41–52.

(109) Kaspar, F.; Neubauer, P.; Gimpel, M. Bioactive Secondary Metabolites from Bacillus Subtilis: A Comprehensive Review. **2019**.

(110) Li, H.; Han, X.; Dong, Y.; Xu, S.; Chen, C.; Feng, Y.; Cui, Q.; Li, W. Bacillaenes: Decomposition Trigger Point and Biofilm Enhancement in Bacillus. *ACS Omega* **2021**, *6*, 1093–1098.

(111) Nguyen, T. A.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. Exploiting the Mosaic Structure of Trans-Acyltransferase Polyketide Synthases for Natural Product Discovery and Pathway Dissection. *Nat Biotechnol* **2008**, *26* (2), 225–233.

(112) Calderone, C. T.; Kowtoniuk, W. E.; Kelleher, N. L.; Walsh, C. T.; Dorrestein, P. C. Convergence of Isoprene and Polyketide Biosynthetic Machinery: Isoprenyl-S-Carrier Proteins in the PksX Pathway of Bacillus Subtilis. *Proc Natl Acad Sci U S A* **2006**, *103* (24), 8977–8982.

(113) Reddick, J. J.; Antolak, S. A.; Raner, G. M. PksS from Bacillus Subtilis Is a Cytochrome P450 Involved in Bacillaene Metabolism. *Biochem Biophys Res Commun* **2007**, *358* (1), 363–367.

(114) Moldenhauer, J.; Chen, X. H.; Borriss, R.; Piel, J. Biosynthesis of the Antibiotic Bacillaene, the Product of a Giant Polyketide Synthase Complex of the Trans-AT Family. *Angewandte Chemie - International Edition* **2007**, *46* (43), 8195–8197.

(115) Miller, B. R.; Gulick, A. M. Structural Biology of Non-Ribosomal Peptide Synthetases. *Methods Mol Biol* **2016**, *1401*, 3.

(116) Fisch, K. M. Biosynthesis of Natural Products by Microbial Iterative Hybrid PKS-NRPS. *RSC Adv* **2013**, *3* (40), 18228–18247.

(117) Ansari, M. Z.; Yadav, G.; Gokhale, R. S.; Mohanty, D. NRPS-PKS: A Knowledge-Based Resource for Analysis of NRPS-PKS Megasynthases. *Nucleic Acids Res* **2004**, *32*.

(118) Um, S.; Fraimout, A.; Sapountzis, P.; Oh, D. C.; Poulsen, M. The Fungus-Growing Termite Macrotermes Natalensis Harbors Bacillaene-Producing Bacillus Sp. That Inhibit Potentially Antagonistic Fungi. *Sci Rep* **2013**, *3* (1), 1–7.

(119) Erega, A.; Stefanic, P.; Danevčič, T.; Možina, S. S.; Mulec, I. M. Impact of Bacillus Subtilis Antibiotic Bacilysin and Campylobacter Jejuni Efflux Pumps on Pathogen Survival in Mixed Biofilms. *Microbiol Spectr* **2022**, *10* (4), 1–14.

(120) Qin, W.; Liu, Y.; Ren, P.; Zhang, J.; Li, H.; Tian, L.; Li, W. Uncovering a Glycosyltransferase Provides Insights into the Glycosylation Step during Macrolactin and Bacillaene Biosynthesis. *ChemBioChem* **2014**, *15* (18), 2747–2753.

(121) Li, H.; Han, X.; Zhang, J.; Dong, Y.; Xu, S.; Bao, Y.; Chen, C.; Feng, Y.; Cui, Q.; Li, W. An Effective Strategy for Identification of Highly Unstable Bacillaenes. *J Nat Prod* **2019**, *82*, 3346.

(122) Landy, M.; Warren, G. H.; RosenmanM, S. B.; Coli, L. G. Bacillomycin: An Antibiotic from Bacillus Subtilis Active against Pathogenic Fungi. *Proceedings of the Society for Experimental Biology and Medicine* **1948**, *67* (4), 539–541.

(123) Otte, K.; Kühne, N. M.; Furrer, A. D.; Baena Lozada, L. P.; Lutz, V. T.; Schilling, T.; Hertel, R. A CRISPR-Cas9 Tool to Explore the Genetics of Bacillus Subtilis Phages. *Lett Appl Microbiol* **2020**, *71* (6), 588–595.

(124) Wicke, N.; Radford, D.; Verrone, V.; Wipat, A.; French, C. BacilloFlex: A Modular DNA Assembly Toolkit for Bacillus Subtilis Synthetic Biology. *bioRxiv* **2017**, 185108.

(125) Stuchbury, G.; Münch, G. Optimizing the Generation of Stable Neuronal Cell Lines via Pre-Transfection Restriction Enzyme Digestion of Plasmid DNA. *Cytotechnology* **2010**, *65*, 189–194.

(126) Turgeon, N.; Laflamme, C.; Ho, J.; Duchaine, C. Elaboration of an Electroporation Protocol for Bacillus Cereus ATCC 14579. *J Microbiol Methods* **2006**, *67* (3), 543–548.