Title
A Synthetic Biochemistry Platform for the Enzymatic Synthesis of Cannabinoids and Other Prenylated Natural Products

Permalink
https://escholarship.org/uc/item/6p2500bg

Author
Valliere, Meaghan

Publication Date
2019

Peer reviewed|Thesis/dissertation
A Synthetic Biochemistry Platform for the Enzymatic Synthesis of Cannabinoids and Other Prenylated Natural Products

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

Meaghan Ann Valliere

2019
ABSTRACT OF THE DISSERTATION

A Synthetic Biochemistry Platform for the Enzymatic Synthesis of Cannabinoids and Other Prenylated Natural Products

by

Meaghan Ann Valliere
Doctor of Philosophy in Molecular Biology
University of California, Los Angeles, 2019
Professor James U. Bowie, Chair

Natural products possess amazing medicinal properties, but in order to capitalize on them we need reproducible and efficient methods to produce them. While this usually means a chemical synthesis approach, natural products can have intricate structures with multiple chiral centers, making synthesis too difficult or expensive. For many complex molecules, bio-based approaches can be advantageous (reviewed in Chapter 1). Cannabinoids are an exciting class of complex natural products that interact with the human endocannabinoid receptors, and have considerable therapeutic potential. Cannabinoids are usually extracted from the cannabis plant and some synthetic routes have been developed. However, due to the high demand for pharmaceutical grade cannabinoids, there are growing efforts to develop a more sustainable and more cost-effective bio-based approach, by metabolic engineering of living organisms. Herein, I test an alternative approach to bio-based cannabinoid production using cell-free enzymatic synthesis, an approach we call synthetic biochemistry.
The dissertation of Meaghan Ann Valliere is approved.

David E. Eisenberg
Joseph A. Loo
Yi Tang
Todd Yeates
James U. Bowie, Committee Chair

University of California, Los Angeles
2019
Table of Contents

Abstract of Dissertation .................................................. ii
Table of Contents ............................................................ iv
Acknowledgements ............................................................ vii
VITA ................................................................................ viii
Publications ......................................................................... ix

Chapter 1: Introduction to Natural Products and How to Make Them 1
  1.1 Introduction to Natural Products in Medicine .................. 2
  1.2 Overview of Methods for the Production of Natural Products 3
  1.3 Engineering Plants for the Production of Natural Products 3
  1.4 Plant Cell Culture for the Production of Natural Products 4
  1.5 Engineering Microbes for the Production of Natural Products 5
    1.5.1 Microbial Production of Biofuels ......................... 6
    1.5.2 Microbial Production of Terpenes and Terpenoids 7
    1.5.3 Microbial Production of Alkaloids ....................... 10
    1.5.4 Microbial Production of Polyketides .................... 12
    1.5.5 Microbial Production of Cannabinoids ................. 14
    1.5.6 Review of microbial engineering for the production of natural products 16
  1.6 Cell-free Approach for the Production of Natural Products 17
  1.7 Synthetic Biochemistry Approach to Natural Product Production 18
  1.8 Synthetic Biochemistry Approach for Cannabinoids .......... 21

Chapter 2: Cell-free platform for the prenylation of natural products with applications to cannabinoid production 38
  2.1 Abstract ................................................................... 39
  2.2 Introduction ............................................................. 39
  2.3 Results and Discussion .............................................. 41
  2.4 Conclusion .............................................................. 46
  2.5 Experimental Section ............................................... 47
  2.6 References .................................................................. 76

Chapter 3: The complete enzymatic synthesis of cannabinoids from simple precursors 81
  3.1 Abstract ................................................................... 82
  3.2 Introduction ............................................................. 82
  3.3 Results and Discussion .............................................. 84
  3.4 Conclusion .............................................................. 90
List of Figures and Tables

Chapter 1: Introduction to Natural Products and How to Make Them

Table 1-1 Summary of microbial engineering projects 23
Figure 1-1 Methods for the production of natural products 24
Figure 1-2 Isobutanol biosynthesis 25
Figure 1-3 Terpene biosynthesis 26
Figure 1-4 Caffeine and psilocybin biosynthesis 27
Figure 1-5 Hydrocodone biosynthesis 28
Figure 1-6 Plant polyketide biosynthesis 29
Figure 1-7 Cannabinoid biosynthesis 30

Chapter 2:

Table 2-1 Enzymes used in the enzymatic platform. 59
Table 2-2 Rosetta suggested mutants varying initial olivetolate position 60
Table 2-3 Evaluation of mutation significance according to the Rosetta energy score function 60
Table 2-4 NphB mutants 61
Table 2-5 Kinetic parameters for NphB mutants 62
Figure 2-1 The synthetic biochemistry platform for the production of prenyl-natural products 63
Figure 2-2 Development of a PDH bypass for the prenylation of aromatic polyketides 64
Figure 2-3 Engineering NphB to improve CBGA production 65
Figure 2-4 Evaluation of the cell-free prenylation system for the production of various cannabinoids. 66
Figure 2-5 Full reaction schematic for the enzymatic synthesis of cannabinoids 67
Figure 2-6 LC-MS chromatogram for prenyl-apigenin 68
Figure 2-7 LC-MS chromatogram for prenyl-daidzein 69
Figure 2-8 LC-MS chromatogram for prenyl-genistein 70
Figure 2-9 LC-MS chromatogram for prenyl-naringenin 71
Figure 2-10 LC-MS chromatogram for prenyl-olivetol  72
Figure 2-11 LC-MS chromatogram for prenyl-resveratrol  73
Figure 2-12 Activity screen of NphB constructs  74
Figure 2-13 SDS-PAGE analysis of CBGA cell-free system precipitate  74
Figure 2-14 Cannabidiolic acid synthase (CBDAS) converts CBGVA into CBDVA.  75
Figure 2-15 Proton NMR Spectrum of CBGVA in CD$_3$OD  76

Chapter 3:

Table 3-1 Summary of published cannabinoid production systems 97
Table 3-2 Summary of all enzymes used in this study 98
Table 3-3 Enzymes used in the full cannabinoid MatB pathway, with final enzyme concentrations 99
Table 3-4 Enzymes used in the full cannabinoid MdcA pathway, with final enzyme concentrations 100
Figure 3-1 Cannabinoid biosynthesis 101
Figure 3-2 Aromatic polyketide production module 102
Figure 3-3 Production of OA and DA over time 103
Figure 3-4 Inhibition of the AP path by OA and DA 103
Figure 3-5 OA and DA inhibition of AP pathway enzymes 104
Figure 3-6 pH dependence of OLS activity 104
Figure 3-7 Effects of additives on OA production 105
Figure 3-8 The effect of additives on AP production and inhibition 106
Figure 3-9 Pathway schematics for the MatB and MdcA (transferase) paths 107
Figure 3-10 Reaction schematics for the enzyme activity assay in this chapter 108
Acknowledgements

This work would not have been possible without the support I received from my family, friends and mentors.
VITA

2014 Bachelor of Science, Chemistry
University of Massachusetts, Amherst

2014-2019 PhD Graduate Student
University of California, Los Angeles

2015-2018 Chemistry Biology Interface Grant Awardee
University of California, Los Angeles
Publications


*Authors contributed equally to this publication


Chapter 1:

Introduction to Natural Products and How to Make Them
1.1 Introduction to Natural Products in Medicine

Nature plays a crucial role in medicine. Natural products, or molecules produced by living organisms, are a rich source of novel molecules that have a myriad of beneficial properties. The idea of using natural products as medicines is not a new concept. In fact, it dates back to ancient civilizations where various plant products were used to manage, treat and prevent disease. Mesopotamia (2600 BC) relied on cedar and cypress oils to cure ailments, Egypt (1500 BC) had a collection of about 700 plant based medicines detailed in “Ebers Papyrus.” Chinese civilizations also cultivated their own list of natural medicines with the first record dating back to 1100 BC.\(^1\) Over time the lists of natural remedies evolved; however, it was not until the 19\(^{th}\) century AD when scientists began to identify the actual molecules that had the medicinal effects.

Morphine was the first natural product to be isolated from its natural source and sold in 1826.\(^2,3\) Since, scientists have continued to identify novel natural products with amazing properties, and while the list of natural product medicines is long, some of the most notable are: penicillin (antibiotic, 1928), paclitaxel (anticancer, 1950s), lovastatin (statin, 1970s), quinine (anti-malarial, 1820) and cannabidiol (analgesic and anti-inflammatory, 1940).\(^1\) Beyond medicinal natural products, scientists have also identified natural products that can be used as fuels, fragrances and dyes. Nature continues to be a rich source for novel molecules, with some predicting only 10% of the biodiversity that exists on the planet has been discovered.\(^4\)

Despite the importance of natural products in everyday life, there has been a paradigm shift from the use of unmodified natural products to semi-synthetic modified natural products, and synthetic compounds.\(^4-6\) This is in part due to the expense associated with identifying and characterizing natural products, and the decision of the Supreme court in *Association for Molecular Pathology v. Myriad Genetics*, 569 U.S. 12 (2013) which decided that natural products cannot be patented, thereby limiting economic motivation for identifying new natural products. However, it is imperative to realize the myriad of applications for natural products in human health.
and other industries, and continue to develop methodologies to identify and synthesize this class of molecules. Herein, we focus on methods to produce natural products.

1.2 Overview of Methods for the Production of Natural Products

Chemical synthesis is the most common method used to produce a molecule, and while this technique can be used to produce almost any molecule, the process is not always economically feasible on an industrial scale. This is often the case for natural products; the chemical synthesis is possible (morphine, paclitaxel, cannabinoids), but the cost of the synthesis process is greater than the cost of extracting the target molecule from a natural source. As a result, the world’s supply of morphine, paclitaxel and cannabinoids as well as an array of other molecules is still dependent on natural sources. However, there are many challenges associated with relying on natural sources to produce these crucial medicines. Plants are slow growing, susceptible to environmental conditions, inconsistent in secondary metabolite production, and secondary metabolites are often present at very low concentrations. A great example is paclitaxel, produced by the pacific yew tree. It requires 3 to 6, 100-year old, pacific yew trees to treat one cancer patient. This is simply not sustainable. Additionally, there can also be political implications. The world’s supply of morphine is predominantly produced in Turkey, India and Afghanistan, and shifts in the political climate could lead to shortages or high prices of the essential drug. So while chemical synthesis and natural product extraction are sufficient for now, in order to capitalize on the properties of natural products, we need new cost-effective, consistent and efficient methods to produce natural products.

Several ideas have been proposed to address these challenges outlined in Figure 1-1, such as engineering plants to improve natural abundance, plant cell culture, engineering microbes, cell-free biosynthesis and synthetic biochemistry. Herein, each method listed above is reviewed for the pros and cons. It is important to realize that natural products are unique and different methods will be suitable for different molecules.

1.3 Engineering Plants for the Production of Natural Products
First, some effort has been dedicated to engineering plants to either improve the abundance of a target molecule, or engineer a plant to make a natural product it would not make normally. Zhang et al were able to engineer tomatoes to improve the production of genistein, a flavonoid with an array of important medicinal properties. Naturally tomatoes produce very low levels of genistein (0.3 mg/g DW), but Zhang et al improved the amount of genistein to 78 mg/g of dry weight, nearly 0.8% of the tomato DW. Additionally, you could engineer a plant to produce a molecule it would not have produced naturally. In addition to boosting genistein levels, Zhang et al were able to engineer tomatoes to produce resveratrol, a compound not naturally produced by tomatoes at levels 100-fold higher than grapes (a good source of natural resveratrol). However genetically engineering plants can be challenging. There are only a few species that have been extensively studied, which limits the number of model systems. Additionally, plant metabolism is complex and it can be difficult to identify molecular components, like transcription factors, that are necessary to boost biosynthesis of the target molecule. This also does not address some of the inherent issues with plant sources, such as cultivation time and susceptibility to mold, pests and environmental conditions.

1.4 Plant cell culture for the production of natural products

Plant cell culture is an exciting alternative to extracting the compound from the natural source. They can be generated from a variety of plant species by first isolating plant tissue (explant) sterilizing the tissue, and plating the sterilized tissue on solid media supplemented with plant hormones and nutrients. The explants proliferate to form a callus (dedifferentiated plant cells). Calluses can then be used to initiate plant cell suspension cultures, which have the ability to produce secondary metabolites. This method has many advantages over extraction from the natural source such as improved sustainability, shorter incubation times, increased production level through metabolic engineering, consistent environmental conditions, and protection from insects and mold.
Cell culture has been particularly fruitful for the production of paclitaxel, a powerful anticancer drug. As previously stated it would take 6, 100 year old trees to produce enough paclitaxel to treat one patient, however the plant cell cultures are able to produce 150 mg/L (~13 liters to treat 1 patient) in 6 weeks, a dramatic improvement. Plant cell culture is well suited for the production of paclitaxel due to the complexity of the biosynthetic pathway, so generating cell lines from cells that naturally produce paclitaxel is a relatively simple process. Some additional examples natural products produced by plant cell culture include: scopolamine (treats motion sickness), protoberberines (antibiotic and anti-inflammatory) rosmarinic acid (anti-inflammatory), shikonin (natural dye) and geraniol (essential oil).

Despite the benefits of plant cell cultures there are still several challenges. The cultivation time is still 2-3 weeks, which is rather long compared to microbial cultivation times. Further genetic instability and physiological heterogeneity leads to variability and unpredictability in secondary metabolite production. The secondary metabolites also remain intracellular which can make it difficult to reach high titers. As this technology develops further it may be possible to address some of these issues. For example, Wilson et al are working to reduce genetic instability and rescue necrotic calluses in order to improve production of paclitaxel in plant cell cultures. However, some challenges are more difficult to address such as improving the titer of toxic compounds. While plant cell culture is an important method for natural product biosynthesis, specifically paclitaxel biosynthesis, the method is not broadly applicable to other natural products due to the complexity of plant metabolism.

1.5 Engineering microbes for the production of natural products

In addition to plants, microbes are a rich source of natural products. Both fungi and bacteria naturally produce an array of useful molecules. In fact, several FDA approved antibiotics are produced via microbial fermentation. However, it is also possible to engineer microbes to produce natural products from other organisms such as plants. Microbes are easy to manipulate, they have relatively short cultivation times (2-10 days), and are easier to culture than plant cells.
As a result, there is an entire field of research dedicated to engineering microbes to produce non-native natural products. Due to the vast number of studies conducted in this field, this review will only focus on several examples of microbial production of biofuels, terpenes, alkaloids and polyketides.

1.5.1 Microbial production of biofuels

In addition to being a rich source of medicines, nature also produces molecules that play an important role in everyday life. Due to the need for energy security and limits on the world’s petroleum supply there is a need for sustainable liquid fuels for transportation. Microbes naturally produce alcohols, isoprenoids, and fatty acids, which can be used as fuels or easily converted into fuels. As a result, a great deal of research has focused on engineering microbes to improve the natural production of biofuels. Microbial production of isobutanol and farnesene are reviewed below.

Alcohol-derived biofuels are the ideal candidate to replace gasoline used in cars. Isobutanol has a high energy density (similar to gasoline) low vapor pressure, and a high octane rating (essential to prevent engine damage). Therefore, it can either be blended into gasoline or replace petrochemicals altogether. There have been several studies published that discuss the microbial production of isobutanol in various organisms, such as: E. coli, C. glutamicum, S. cerevisiae, C. acetobutylicum, R. eutropha, and S. elongatus. The highest titer achieved was an engineered strain of E. coli. Atsumi et al engineered E. coli to produce 22 g/L of isobutanol. First, they overexpressed the enzymes in the isobutanol pathway shown in Figure 1-2. Then, they deleted non-essential genes that would divert precursors and intermediates out of the isobutanol pathway. Finally, they improved the flux of pyruvate into the isobutanol pathway by using a non-native enzyme with better kinetic parameters. A later study by Baez et al using a similar E. coli strain obtained titers of 50 g/L by continuously removing isobutanol from the culture, demonstrating the effects of isobutanol toxicity on E. coli cultures. While alcohol based biofuels are a possible alternative, they are too expensive to compete with the low cost of gasoline. If
titers could be improved further, the cost of the biofuels would decrease, however this is unlikely due to the toxicity of alcohols at high concentrations.

The highest microbial production of a biofuel reported thus far is a sesquiterpene derived biofuel, which can be converted into diesel fuel and jet fuel. Amyris was able to engineer yeast to produce the sesquiterpene farnesene at titers that exceed 130 g/L by re-engineering yeast central metabolism to direct sugar into the isoprenoid pathway.\textsuperscript{24} However, these production levels are still relatively low, when considering the energy demand of US transportation. At this titer, it would require 23.5 L of yeast culture to produce 1 gallon of farnesene. To replace the amount of jet fuel used in the US airlines in 2018 (~17.8 billion gallons) with farnesene, 1.2 billion liters of yeast would need to be cultivated and processed per day a total of 420 billion liters of yeast per year. In addition, while this titer is significantly higher than most reported, the cost of producing farnesene via microbial production is still more expensive than petrochemicals (average $2.15 per gallon). The real challenge with biofuels is not engineering the microbes, but competing with the very low prices of petrochemical products.

While biofuel production in microbes is generally too expensive to compete with fossil fuels, this is not the case with all natural products. Due to the challenges associated with chemical synthesis, a low natural abundance and the variability with plant based production of complex natural products, microbial fermentation is a plausible alternative for the production of high value natural products. Microbes (most commonly \textit{E. coli} and yeast) have been engineered to produce an array of compounds, spanning three of the major classes of natural products: terpenoids, alkaloids and polyketides.

\textbf{1.5.2 Microbial production of terpenes and terpenoids}

Terpenes and terpenoids are a large (55,000+ molecules) diverse class of natural products derived from the mevalonate (MVA) or methylerethritol phosphate pathway (MEP). The products of the MEP and MVA pathways are isopentyl pyrophosphate (IPP) and dimethallyl pyrophosphate (DMAPP), which serve as the core building blocks for terpene biosynthesis (Figure
IPP and DMAPP undergo enzymatically catalyzed condensation reactions to build longer carbon chains. For example, the condensation of DMAPP (C5) and IPP (C5) produces geranyl pyrophosphate (GPP, C10), the precursor to monoterpenes. Condensation of GPP with IPP produces farnesyl pyrophosphate (FPP, C15), which is the precursor to sesquiterpenes, like farnesene mentioned above. The condensation of two FPP molecules yields squalene (30 carbons), which is the precursor to cholesterol. Terpenes are classified by the number of isoprene units which range from monoterpenes (2 isoprene units) to polyterpenes (8+ isoprene units). Terpene synthases catalyze the cyclization of the hydrocarbon moiety of isoprenoids to yield cyclized hydrocarbons. While terpene synthases generally have one major product, they are fairly promiscuous and can produce several different terpenes with one enzyme. For example, the humulene synthase from Abies grandis species mainly produces humulene, but also produces 51 additional sesquiterpenes at lower levels. The benefit of producing 52 different compounds with one enzyme is to maximize the diversity of secondary metabolites made from one biosynthetic pathway, or better yet one enzyme. Due to the myriad of beneficial properties of terpenes, a great deal of research has been dedicated to engineering microbial strains to produce terpenes.

Limonene is a rather simple, well characterized monoterpene that is commonly used in the fragrance industry and as a food additive. While limonene has a relatively high natural abundance in the rind of citrus fruit, several groups have engineered E. coli to produce this molecule as proof of concept. Additionally, a derivative of limonene, perillyl alcohol, is reported to have anticancer properties, and a low natural abundance making it an interesting target for an engineered microbe. Alonso-Gutierrez et al engineered E. coli to produce limonene at 435 mg/L from glucose, and then introduced a cytochrome P450 to hydroxylate limonene to produce perillyl alcohol at 100 mg/L. However, this level of limonene production still falls short of the theoretical maximum, 3.2 g/L, indicating the system is not an efficient producer of limonene. Willrodt et al improved on the E. coli production system by introducing different mevalonate pathway enzymes, and reached a titer of 2.7 g/L of limonene.
limiting factor in their system, they identify high acetate levels in the high limonene producing strain. While the authors determine this to mean there was sufficient levels of acetyl-CoA, they do not discuss how a competing pathway, like phosphotransacetylase and acetate kinase, could limit the flux of acetyl-CoA into the desired mevalonate pathway. Engineering microbes to produce monoterpenes is further complicated due to the effects monoterpene toxicity. To address this issue both studies utilized an organic overlay to extract the secreted product.

Amorphadiene is a sesquiterpene and a precursor to the essential anti-malarial drug artemisinin. The highest titer of amorphadiene was achieved by Westfall et al. They reached a titer of 40 g/L of amorphadiene in yeast by overexpressing enzymes in the mevalonate pathway and the amorphadiene synthase, and limiting competing pathways. In an attempt to produce artemisinic acid in an engineered microbe, the strain developed by Westfall et al was further engineered to express a cytochrome P450 to produce artemisinic acid. However, despite the high titers of amorphadiene, the artemisinic acid levels remained significantly lower at ~150 mg/L. Amyris improved the titer substantially to 25 g/L of artemisinic acid. The improvement in artemisinic acid was due to the discovery of an aldehyde dehydrogenase (ALDH1) from A. annua that improved the conversion of artemisinic aldehyde into artemisinin. Additionally they increased the viability of the strains by introducing an organic overlay, isopropyl myristate, which helped solubilize the product.

While amorphadiene is a precursor in the biosynthesis of artemisinin, artemisinic acid is not. Therefore, Amyris designed a four step synthetic approach to convert artemisinic acid into artemisinin. The combination of synthetic and bio-based approaches is known as semi-synthesis, and is another approach for the production of natural products. This approach can reduce the difficulty of chemical synthesis, and lead to a more cost effective option. Amyris demonstrated this by chemically converting artemisinic acid derived from microbial fermentation into artemisinin in a 4 step chemical synthesis with an overall yield of 40%. This semi-synthetic process is currently being used by Sanofi to supplement the world’s supply of artemisinin.
Taxadiene is a diterpene and the first intermediate in the biosynthesis of paclitaxel. Several groups have engineered microbes to produce taxadiene and one study produced taxadien-5\(\alpha\)-ol, the second intermediate in the pathway. While taxadiene has been produced at moderate titers in \textit{E. coli} (1 g/L), the production in other organisms, such as \textit{S. cerevisiae}, \textit{B. subtilis} and the fungi \textit{A. alternata TPF6} remains low (< 100 mg/L).\(^{33-36}\) The \textit{E. coli} system that generated 1 g/L taxadiene was also engineered to produce taxadien-5\(\alpha\)-ol, but the titers of taxadiene remain significantly higher (~10 fold) indicating inefficiencies with the hydroxylation of taxadiene.\(^{33}\) This is most likely because cytochrome P450s are challenging to express in \textit{E. coli}. Biggs et al were able to improve hydroxylation of taxanes by optimizing for P450 expression, reductase partner interactions and modifications of the N-terminus of the P450.\(^{37}\) While it may be possible to complete the paclitaxel biosynthetic pathway in a microbe, it would be extremely challenging. Therefore, plant cell culture is still the best route for paclitaxel production.

While there are some success stories of engineering microbes to produce terpenes and terpene derived molecules (artemisinic acid and farnesene), several challenges still remain. Monoterpenes are toxic to both \textit{E. coli} and yeast at relatively low concentrations (<1 g/L).\(^{29,38}\) While cell viability can be improved by using an organic solvent to extract the product, slow diffusion out of the cell can still have an impact. There are also issues with precursor availability. The data from Willrodt et al indicates that competing pathways reduce the availability of the precursor acetyl-CoA.\(^{28}\) Additional challenges, like functional enzyme expression create issues with the production of complex natural products like artemisinin and paclitaxel. While yeast and \textit{E. coli} are relatively easy to engineer, they do not always possess the cellular machinery needed for the functional expression of necessary enzymes.

\subsection*{1.5.3 Microbial production of alkaloids}

Alkaloids are a broad group of natural products that are produced by plants, fungi and bacteria, they are loosely classified as molecules that include one or more nitrogen atoms.\(^{39}\) Some
examples of alkaloids are caffeine (stimulant), opioids (analgesics) and psilocybin (psychotropic). Due to the broad classification of alkaloids, their biosynthesis is rather diverse (Figures 1-4 and 1-5). For example, the molecules listed above are derived from different precursors, caffeine is derived from xanthosine, opioids are derived from tyrosine and is derived from tryptophan. Several alkaloids found in plants have been produced in engineered microbes.

Although caffeine is naturally abundant, Jin et al engineered S. cerevisiae to produce the stimulant. The conversion of xanthosine to caffeine (Figure 1-4) is four enzymatic steps requiring three enzymes, however the nucleosidase is naturally present in yeast, so they only introduced two genes into the strain, a caffeine synthase and a methyl transferase. The strain produced 0.38 mg/L of caffeine, when supplemented with xanthosine. The authors hypothesized the low titer could be due to low nucleosidase activity and screened several non-native nucleosidases, however this did not increase titers. Additionally, Jin et al demonstrate that substrate consumption is not correlated with caffeine production, and that no caffeine is produced if the culture is not supplemented with xanthosine. Because xanthosine is required for primary metabolism of yeast, it is likely that the added xanthosine is being diverted into primary metabolic pathways instead of caffeine biosynthesis, limiting the titers of caffeine.

As mentioned previously, many opioids are still isolated from their natural source. In an attempt to find a better production method yeast was engineered to produce the opioid, hydrocodone. It required introducing over twenty genes into the yeast genome (Figure 1-5). The resulting strain produced 0.3 µg/L of hydrocodone. At this titer it would require nearly 17,000 L of yeast to produce one dose (5 mg) of hydrocodone. The low titer is likely due to the length and complexity of the biosynthetic pathway, but it is a proof of concept suggesting that long, complex exogenous pathways can be functional in yeast. The remaining challenge is identifying the bottlenecks and adjusting flux accordingly.

*Aspergillus nidulans* was engineered to produce psilocybin, a psychotropic molecule under FDA investigation as a treatment for anxiety, depression and substance abuse. Because
psilocybin is naturally produced by fungi, it is logical to use the fungal model organism *A. nidulans* to produce this molecule. The four step pathway, shown in Figure 1-4, was transformed into *A. nidulans*, which resulted in a final titer of 110 mg/L, or 1.5% of mycelium dry weight, similar to the amount naturally produced in mushrooms. However, besides using a fluorescence based assay to identify positive transformants, the authors did not make any additional changes to the fungi. Therefore, by engineering the fungi to limit competing pathways and direct flux into the desired pathway it is possible to reach higher titers.

The microbial production of alkaloids has similar challenges to that of terpenes. It is a challenging to balance competing pathways with a target pathway, like for the production of caffeine. Xanthosine is required for primary metabolism, and so most of the added substrate is diverted into primary metabolism instead of caffeine production. A similar challenge is presented for opioids with tyrosine being an essential metabolite that is required for opioid biosynthesis as well. In addition, longer more complex pathways are generally extremely difficult to engineer into a heterologous host. First, overexpressing twenty enzymes can create a metabolic burden, and it can be difficult to express the enzymes at appropriate levels. Additionally, the enzymes used in the study were sourced from mammals, plants, bacteria in addition to yeast enzymes, which means there could be issues with functional expression. The opioid pathway also produces several non-native intermediates which could have an impact on cellular metabolism. Interestingly, the baseline production of psilocybin in *Aspergillus nidulans* was 110 mg/L, however it may be possible to reach significantly higher titers by engineering the organism further. Previous studies have indicated that altering central metabolism to increase precursor pools can have a dramatic improvement on overall titer.

1.5.4 Microbial production of polyketides

Polyketides are produced by all kingdoms of life, and are a rich source of medicinal compounds. In fact, several are FDA approved antibiotics, immunosuppressants, anticancer and cholesterol reducing medications. The biosynthesis of polyketides is similar to fatty acid
biosynthesis, where small extender units are added to a growing acyl chain; however the subsequent steps (ketoreductase, hydratase and enoyl reductase) are less uniform than FA biosynthesis allowing for more diversity in the final polyketide product. While the chemistry for the biosynthesis of polyketides is highly conserved, the types of polyketide synthases (PKS) involved are very different. There are three general types of PKS reviewed previously by Hertweck in 2009. While several bacteria have been engineered to increase natural production of FDA approved antibiotics, herein the focus will be on producing plant polyketides in microbial hosts.

Stilbenoids and flavonoids are a classes of aromatic polyketides produced by plants, and they have an array of useful bioactivities. These molecules can further modified to yield prenyl-flavonoids and prenyl-stilbenoids (addition of isoprenoid molecule), which has been shown to increase the potency of flavonoids and resveratrol. Due to the low natural abundance of flavonoids, stilbenoids and their prenyl derivatives, they are an interesting target for microbial engineering.

Resveratrol is an aromatic polyketide first isolated from the skin of red grapes. The molecule is potent antioxidant cited to have neuroprotective, cardioprotective and anti-aging properties. The molecule is sold as a supplement, and is also used as an additive in food products and cosmetics. The market for resveratrol in 2017 was valued at $69.1 million based on a report published by Coherent Market Insights. Several groups have engineered microbes to produce this molecule (biosynthesis shown in Figure 1-6). The most successful was Lim et al, reaching a titer of 2.3 g/L. They attribute the high titer to directing carbon flux into the malonyl-CoA precursor, and limiting pathways, like fatty acid biosynthesis that would compete with the resveratrol pathway.

Naringenin is a flavonoid produced by an array of plants. It is a fairly potent estrogen mimic, and is sold as a supplement to ease the symptoms of menopause. In addition it is a potent antioxidant with potential antimicrobial and antiviral properties. There are numerous studies dedicated to engineering E. coli and S. cerevisiae to produce naringenin, however the best titer
of 474 mg/L was reported by Xu et al in 2011. Unlike the number of other studies, Xu et al did more than simply transform the biosynthetic pathway (shown in Figure 1-6) for naringenin into *E. coli*. They sought to increase carbon flux into the pathway by increasing the precursor, malonyl-CoA. To do this they overexpressed key enzymes in the native glycolysis pathway to increase the available pool of acetyl-CoA. Then they introduced acetyl-CoA carboxylase from *Photorhabdus luminescens* to convert the acetyl-CoA into the precursor malonyl-CoA. Additionally, they downregulated genes that would direct carbon into the citric acid cycle. Although Xu et al were able to increase the pool of malonyl-CoA, they did not address the issue of native *E. coli* pathways that siphon off the available malonyl-CoA.

This data along with data for resveratrol production would indicate that one of the limiting factors in the production of polyketides in *E. coli* is the availability of the precursor malonyl-CoA, and that increasing malonyl-CoA levels increases product titers. However, there is potential for other issues that may not be as clear. For example, it is possible that some of intermediates or the product of the target pathway could negatively impact cellular metabolism.

### 1.5.5 Microbial production of cannabinoids

In addition to the various classes of natural products described above, there are some hybrid natural products that span two classes. Cannabinoids for example are derived from the terpenoid pathway and the polyketide pathway (Figure 1-7). Prenyl-flavonoids and prenyl-stilbenoids would also fall under this category. Cannabinoids are a very interesting class of natural products, with some pretty remarkable medicinal properties. For the past 83 years cannabis or marijuana was considered a schedule one drug by the United States, which stunted the research into the molecules that cannabis makes. Although it is still classified as a schedule one drug by the federal government, the perception of the plant and the molecules it makes is starting to change. Thirty-three states now permit the use of medical marijuana and 10 states and Washington DC have legalized it for recreational use.
The shift in perception may be due to recent clinical studies. A non-psychoactive component of cannabis, cannabidiol or CBD was FDA approved to treat severe childhood epilepsy, and has been suggested as a possible treatment for the spasticity associated with diseases like Parkinson’s Disease and Multiple Sclerosis (MS).\textsuperscript{53} THC or $\Delta^9$-tetrahydrocannabinol, the psychoactive component of marijuana is FDA approved as an appetite stimulant for patients going through chemotherapy.\textsuperscript{54} In addition to the conditions these cannabinoids are already approved for, both THC and CBD are cited to have antiemetic, antianxiety, anti-inflammatory, antidepressant and anticonvulsant properties.\textsuperscript{55} While THC and CBD are produced abundantly by cannabis plants, several groups have attempted to engineer microbes to produce cannabinoids and their intermediates in order to create a more sustainable source of these pharmaceutical compounds.

As shown in Figure 1-7 the biosynthesis of cannabinoids can be broken down into three parts, the polyketide biosynthesis (Figure 1-7B), the terpenoid biosynthesis (Figure 1-7A) and the cannabinoid biosynthesis (Figure 1-7C). Gagne et al were the first group to transform the olivetolic acid (OA) biosynthetic pathway into yeast.\textsuperscript{56} Previous work by Taura et al had identified an olivetol synthase, capable of making the decarboxylated form of OA.\textsuperscript{57} Gagne et al posited that there was an additional enzyme in the biosynthetic pathway that would cyclize the polyketide chain to form OA instead of olivetol. In this work they identified olivetolic acid cyclase, which cyclizes tetraketide intermediate to form OA instead of olivetol. Gagne et al transformed the genes necessary to produce OA in S. cerevisiae, and reported a titer of 0.5 mg/L; however this titer is not optimized. Six years later Tan et al engineered the same pathway into E. coli.\textsuperscript{58} They added additional enzymes to increase the malonyl-CoA and hexanoyl-CoA precursors, and were able to obtain 80 mg/L of OA.

At the other end of the biosynthesis (Figure 1-7C), Zirpel et al have engineered several yeast strains to catalyze the final two steps in the biosynthetic pathway. First in 2017, Zirpel engineered P. pastoris to express a promiscuous prenyltransferase, NphB, and THCA synthase
to convert OA into THC. While the yeast expressed both enzymes, they did not detect any THC until they lysed the cells and supplemented with two precursors, OA and geranyl pyrophosphate (GPP). Even then, the production of THC was slow at ~ 10 nmol/L/hr, which was attributed to an inefficient prenylation step, and low levels of the cannabinoid synthase. In later studies, Zirpel et al focused on engineering the yeast strains to improve the expression of the cannabinoid synthases THCAS and CBDAS. They identified five proteins that aid in the expression of the cannabinoid synthases. When they supplemented the yeast overexpressing the cannabinoid synthases with CBGA, they were able to produce over 3 g/L of THCA and 400 mg/L of CBDa, a significant improvement to their first study.

Only one group has reported the entire biosynthesis of cannabinoids in yeast, however the titers are low. They produced the precursor cannabigerolic acid (CBGA) at 8 mg/L, THCA at 1.1 mg/L and CBDa at 4.2 µg/L, when the cultures were supplemented with 1 mM hexanoic acid. These low titers indicate the difficulty associated with engineering a microbe to produce a complex natural product. When the authors supplemented cultures with 1 mM of the precursor OA, they were able to produce higher levels of the CBGA intermediate 200 mg/L a 55% conversion of the OA added. This may indicate that OA biosynthesis is limiting in their strain, probably due to the low malonyl-CoA concentrations.

Similarly to other engineered microbes for opioid, paclitaxel and artemisinic acid production, the cannabinoid pathway is very complex. The root of the pathway is the essential precursor acetyl-CoA, which is required to make both the aromatic polyketide component, and the isoprenoid component of the cannabinoid. Additionally there is always the possibility of product or intermediate toxicity, which can limit titers. Further, expression of the cannabinoid synthases can be challenging. Zirpel et al observed low titers of THC until they co-expressed the synthase with several chaperones, and an enzyme to improve cofactor biosynthesis.

1.5.6 Review of Microbial Engineering for the Production of Natural Products
Listed above are 15 examples of microbes that have been engineered to produce various natural products (summarized in Table 1), and it’s evident that microbial engineering allows for the production of vastly different natural products. Microbial fermentation is an important tool for the production of natural products, and is used in several industrial processes. However, the cases presented above demonstrate the highly variable nature of titers, ranging from µg/L to g/L, and it’s apparent that some studies are more successful than others. The studies that are more successful tend to be shorter metabolic pathways, and they do more than simply add an exogenous biosynthetic pathway, they try to drive carbon flux into the target pathway. This highlights one inherent problem with engineering microbes. The target pathway engineered into the microbe is always competing with native metabolic pathways for cofactors and metabolites. Therefore, essential pathways will always deplete resources from the target pathway. Additionally, metabolites from the background metabolism have the potential to inhibit the target pathway. These interactions are extremely challenging if not impossible to identify in vivo, and can limit the overall titers of any pathway. The final product can also decrease cell viability, like in the case of monoterpenes and alcohols, and limit product titers. Finally, while engineering microbes is fairly simple and efficient compared to other organisms, it still takes a significant amount of time to develop strains, approximately “150 person-years” of work were required to engineer yeast to produce artemisinic acid at 25 g/L.\textsuperscript{62}

There are two alternatives to metabolic engineering, cell-free biosynthesis and synthetic biochemistry. They are built on the same principles, using enzymes to produce a target molecule, but instead of engineering a microbe, the biocatalysts are reconstituted either in a cell lysate or in vitro. The difference between cell-free and synthetic biochemistry is small, with cell-free relying on the cellular machinery in a lysate to recycle essential co-factors, whereas synthetic biochemistry relies on pathways with the ability to recycle the necessary co-factors.

1.6 Cell-free Approach for the Production of Natural Products
Cell-free systems are most commonly used to power protein expression, however there are some cases where lysates have powered the production of a biosynthetic pathway. There's a lot of flexibility within this set up. Kay et al produced 2,3 butanediol at 84 g/L using a lysate derived from an *E. coli* strain expressing all three enzymes in the pathway, but Dudley et al expressed the desired proteins individually, lysed the cells and mixed the various lysates together to produce mevalonate at 17.5 g/L.Interestingly, when Dudley applied the same concept to the production of limonene (6 additional steps from mevalonate), the titers were significantly lower at approximately 100 mg/L. This type of system shares some of the challenges with metabolic engineering. Even though the cells are lysed, metabolic enzymes still remain active; therefore it's possible for carbon flux to be diverted from the target pathway. Dudley et al cite this as a major challenge for the cell-free production of limonene. The native *E. coli* farnesyl pyrophosphate synthase remains active, and siphons flux away from the monoterpene pathway to produce a sesquiterpene alcohol.

Additionally, metabolites present in the lysate can inhibit the pathway of interest and limit titers. However, cell viability is no longer limiting, products can be extracted in real time limiting product inhibition and the target pathway protein levels are easily manipulated.

1.7 Synthetic Biochemistry Approach to Natural Product Production

Synthetic biochemistry provides several advantages over metabolic engineering. One of the major advantages is the carbon flux and co-factors are not diverted from the desired pathway. Many of the successful studies mentioned above required altering central metabolism to account for the target pathway, but in every instance it was a balancing act between essential pathways and the target pathway. A synthetic biochemistry approach also allows for rapid design, build, test cycles, which makes it easier and faster to identify pathway bottlenecks. Once bottlenecks are identified, it is very easy to tailor the enzyme activity to alleviate the problem, something that is very challenging to do *in vivo*. Additionally, because enzymes can be expressed and purified from a range of expression platforms (bacteria, yeast, plant and mammalian), it increases the pool of
potential enzymes that can be used. Finally, like the cell-free approach the products can be easily extracted, limiting problems associated with product toxicity. However there is a challenge with this approach. Without the cell, the \textit{in vitro} pathways need to incorporate components that will balance and regenerate co-factors.

In their simplest form, \textit{in vitro} enzymatic systems can be broken down into two modules, the sugar breakdown module and the build module. The sugar breakdown module catabolizes the sugar into 2-3 carbon building blocks and generates high energy co-factors like ATP and reducing equivalents in the process. Those components are then assembled in the build phase to generate the final product. The challenge is the high energy cofactors produced in the sugar breakdown module are not always balanced with the build module. This is the case for the \textit{in vitro} biosynthesis of monoterpenes and isobutanol.

Korman et al engineered a system with 27 enzymes to biosynthesize monoterpenes.\textsuperscript{66} To produce a monoterpene, the feedstock, glucose, was converted into pyruvate via glycolysis, pyruvate was converted into acetyl-CoA with pyruvate dehydrogenase (PDH), acetyl-CoA fed into the mevalonate pathway to yield geranyl pyrophosphate (GPP), which was converted into a monoterpene via a monoterpene synthase. The pathway originally was not stoichiometrically balanced for the reducing equivalents, NAD(P)H. The sugar breakdown module originally produced 6 moles of NADH, and the build module required 2 moles of NADPH. Not only were the reducing equivalents stoichiometrically unbalanced, they were not the type the build module needed. To counteract this imbalance Korman et al employed a molecular purge valve. The purge valve consists of three components. First, an NAD+ specific glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is responsible for maintaining carbon flux through glycolysis. An NADH oxidase (NoxE) was used to burn excess reducing equivalents, and an NADP+ specific GAPDH was used to generate the reducing equivalents required for the mevalonate pathway. All three components were required for the \textit{in vitro} system to run efficiently. Using the NADP+ specific GAPDH alone leads to a buildup of NADPH, which prevents the conversion of glyceraldehyde-3-
phosphate (G3P) into 1,3-bisphosphoglycerate (BPG). This would eliminate ATP regeneration via the glycolysis pathway and eventually stop the system. The pathway yielded 12.5 g/L of limonene and 14.9 g/L of pinene, which is significantly higher than titers achieved in microbes.

Opgenorth et al engineered an in vitro enzymatic system to produce isobutanol. As previously mentioned alcohols are fairly toxic to microorganisms. Therefore, an in vitro, cell-free approach might lead to higher titers. For isobutanol, the sugar breakdown module is glycolysis which nets 2 reducing equivalents and 2 ATP per cycle. The isobutanol build module requires 2 moles of NADPH, but 0 ATP. Therefore, in order to balance the system Opgenorth et al utilized GapN, an enzyme that converts G3P and NADP+ into 3-phosphoglycerate and NADPH, thereby eliminating a step that regenerates an ATP. This leads to a system that is stoichiometrically balanced, however, this may not be optimal in a cell-free system. Opgenorth et al found that ATPase activity from purified enzymes or spontaneous ATP hydrolysis can deplete ATP stores and limit product titers. So, they engineered a component to generate excess ATP under high phosphate conditions (when ADP levels are high), called the molecular rheostat. The rheostat is a branchpoint in glycolysis providing two paths that convert G3P into 3PG, with one path produces ATP and the other does not. The GapN path directly converts G3P into 3PG yielding no ATP. The phosphorylation path converts G3P into BPG with an NADP+ specific GAPDH, and BPG is converted into 3PG using phosphoglycerate kinase, which regenerates an ADP into ATP. Because the activity of the GAPDH enzyme is dependent on the phosphate concentration, the rheostat ATP production is also dependent on the phosphate concentration. This system yielded 24 g/L of isobutanol, close to titers reported in microbes. However, the limiting factor of the system was identified as enzyme instability in high isobutanol, and the authors were able to determine precisely which enzymes were the least stable. Since, the lab has stabilized many of the enzymes in the pathway by either engineering the enzyme or using a more thermostable variant. Current titers in a 15 mL bioreactor have reached nearly 300 g/L isobutanol orders of magnitude higher than microbial production (data not yet published).
While the monoterpane and isobutanol pathways required sophisticated methods for cofactor regeneration, that may not always be the case. You et al utilized an 5 step enzymatic system to convert starch into myo-inositol (vitamin B8), which is co-factor independent. The enzymes were expressed in *E. coli* and purified using a simple heating step. The resulting biosynthesis of starch to myo-inositol yielded 95 g/L in 48 hours on an 18,000 L scale. This was a remarkable study because of the scale and final titers achieved. It demonstrates that it is possible to translate the enzymatic systems to an industrial scale.

Synthetic biochemistry is still in its infancy, but it has the potential to be a powerful tool for the production of natural products. There are a few obstacles that need to be addressed before more complex systems are industrially relevant. First, since synthetic biochemistry does not use cell lysates or cells, so cofactors like ATP, coenzyme A and NAD(P)+ need to be added to start the system. While recycling these cofactors and using them for several iterations would reduce the associated cost, it is important to find an inexpensive source for these molecules. Additionally, the enzyme catalysts can also be an expensive component of these systems, however this cost can be overcome by recycling the enzymes and using them for several iterations. The use of thermostable enzymes is a key factor, demonstrated by You et al. The thermostable enzymes are able to be purified by a simple heating step, reducing the cost associated with protein purification. Additionally, they are able to recycle the enzymes and use them in subsequent bioreactor runs. While the obstacles might take time to solve, they are definitely solvable problems, making synthetic biochemistry an alternative approach for the biosynthesis of natural products.

1.8 Synthetic biochemistry approach for cannabinoids

Described above are several bio-based approaches for the production of natural products. When seeking a bio-based approach for natural product production, it is important to recognize that different systems will work better for different molecules. In the case of paclitaxel it is clear that plant cell culture is currently the best method, however for monoterpenes a synthetic biochemistry approach may be a better option based on titer.
The focus of this thesis is to evaluate a synthetic biochemistry approach for the production of cannabinoids and other prenylated aromatic polyketides. Due to the low cannabinoid titers seen with metabolic engineering it is possible that an alternative approach may be more successful. The low titers are most likely due to competition for the precursor acetyl-CoA. It requires 9 acetyl-CoA molecules to produce 1 molecule of THC or CBD. In addition to the 9 molecules of acetyl-CoA needed for cannabinoid biosynthesis, there are other essential competing pathways like fatty acid biosynthesis. Additionally the 9 acetyl-CoA molecules are split between the isoprenoid pathway (6) and the polyketide pathway (3), and engineering the yeast to express each enzyme at the level needed to balance the acetyl-CoA flux is extremely difficult. A synthetic biochemistry approach for this pathway may be the better option due to better control over reaction components, flux, and a sufficient supply of the precursor acetyl-CoA. Due to these advantages, synthetic biochemistry could be a useful tool for the sustainable production of cannabinoids.
<table>
<thead>
<tr>
<th>Molecule</th>
<th>Organism</th>
<th>Titer</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isobutanol</td>
<td><em>E. coli</em></td>
<td>22 g/L</td>
<td>Baez</td>
</tr>
<tr>
<td>Farnesene</td>
<td><em>S. cerevisiae</em></td>
<td>130 g/L</td>
<td>Meadows</td>
</tr>
<tr>
<td>Limonene</td>
<td><em>E. coli</em></td>
<td>435 g/L</td>
<td>Alonso-Gutierrez</td>
</tr>
<tr>
<td>Perillyl Alcohol</td>
<td><em>E. coli</em></td>
<td>100 mg/L</td>
<td>Alonso-Gutierrez</td>
</tr>
<tr>
<td>Amorphadiene</td>
<td><em>S. cerevisiae</em></td>
<td>40 g/L</td>
<td>Westfall</td>
</tr>
<tr>
<td>Artemisinic Acid</td>
<td><em>S. cerevisiae</em></td>
<td>150 mg/L</td>
<td>Westfall</td>
</tr>
<tr>
<td></td>
<td><em>S. cerevisiae</em></td>
<td>25 g/L</td>
<td>Paddon</td>
</tr>
<tr>
<td>Taxadiene</td>
<td><em>E. coli</em></td>
<td>1 g/L</td>
<td>Ajikumar</td>
</tr>
<tr>
<td>Taxadien-5a-ol</td>
<td><em>E. coli</em></td>
<td>~ 60 mg/L</td>
<td>Ajikumar</td>
</tr>
<tr>
<td>Caffeine</td>
<td><em>S. cerevisiae</em></td>
<td>0.38 mg/L</td>
<td>Jin</td>
</tr>
<tr>
<td>Hydrocodone</td>
<td><em>S. cerevisiae</em></td>
<td>0.3 µg/mL</td>
<td>Galanie</td>
</tr>
<tr>
<td>Psilocybin</td>
<td><em>A. nidulans</em></td>
<td>110 mg/L</td>
<td>Hoefgen</td>
</tr>
<tr>
<td>Resveratrol</td>
<td><em>E. coli</em></td>
<td>2.3 g/L</td>
<td>Lim</td>
</tr>
<tr>
<td>Naringenin</td>
<td><em>E. coli</em></td>
<td>474 mg/L</td>
<td>Xu</td>
</tr>
<tr>
<td>THCA</td>
<td><em>S. cerevisiae</em></td>
<td>8 mg/mL</td>
<td>Luo</td>
</tr>
<tr>
<td>CBDA</td>
<td><em>S. cerevisiae</em></td>
<td>4.2 µg/mL</td>
<td>Luo</td>
</tr>
<tr>
<td>OA</td>
<td><em>S. cerevisiae</em></td>
<td>0.48 mg/L</td>
<td>Gagne</td>
</tr>
<tr>
<td></td>
<td><em>E. coli</em></td>
<td>80 mg/L</td>
<td>Tan</td>
</tr>
</tbody>
</table>

**Table 1-1:** Summary of microbial engineering projects
Figure 1-1: Methods for the production of natural products. Listed above are the various bio-based approaches for the production of natural products. Underneath the method is the approximate amount of time it takes to produce the target product.
Figure 1-2: Isobutanol Biosynthesis. The enzyme abbreviations are shown in blue, and the intermediate names are shown under each molecule.
Figure 1-3: Terpene biosynthesis. Terpenes are derived from the isoprenoid molecules produced by the MEP pathway or the mevalonate pathway. The isoprenoids, isopentyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), undergo a series of condensation reactions to generate longer chain isoprenoids. Terpene synthases cyclize the hydrocarbon moiety and cleave the pyrophosphate. Shown here are some examples of monoterpenes, sesquiterpenes and diterpenes.
Figure 1-4: Caffeine and Psilocybin Biosynthesis. Shown above is the biosynthesis of xanthene derived alkaloids (caffeine) and tryptophan derived alkaloids (psilocybin). The biosynthesis of these molecules is rather simple with 3 and 4 enzymatic steps respectively. The enzyme names are shown to the right of the reaction arrows.
Figure 1-5: Hydrocodone Biosynthesis. Hydrocodone is an alkaloid derived from tyrosine. The biosynthesis is rather complicated with 13 enzymatic steps, that require high energy cofactors.
**Figure 1-6: Plant polyketide biosynthesis.** Shown in this figure is the biosynthetic pathway for the production of two aromatic polyketide naringenin and resveratrol. They both require coumaryl-CoA as the starter unit, and then malonyl-CoA is added to the growing polyketide chain. Finally, the chains are cyclized to produce the final aromatic polyketides.
**Figure 1-7: Cannabinoid Biosynthesis.** Cannabinoid biosynthesis is derived from the fatty acid, isoprenoid and aromatic polyketide metabolic pathways. **(A)** Isoprenoids are generated through the MEP pathway (not shown) or the mevalonate pathway, each molecule of GPP produced requires 6 acetyl-CoA and 6 ATP. **(B)** Aromatic polyketide biosynthesis is dependent on small chain fatty acids, and acetyl-CoA. This pathway requires 2 ATP and 3 acetyl-CoA molecules to produce the aromatic polyketides (olivetolic acid and divarinic acid). **(C)** The cannabinoid path starts by transferring the prenyl group from GPP onto OA, and cyclizing the geranyl moiety, to form the final cannabinoid.
References


doi:10.1007/128_2011_133


Chapter 2: A cell-free platform for the prenylation of cannabinoids with applications for cannabinoid production
2.1 Abstract

Prenylation of natural compounds adds structural diversity, alters biological activity, and enhances therapeutic potential. Because prenylated compounds often have a low natural abundance, alternative production methods are needed. Metabolic engineering enables natural product biosynthesis from inexpensive biomass, but is limited by complexity of secondary metabolite pathways, intermediate and product toxicities, and substrate accessibility. Alternatively, enzyme catalyzed prenyl transfer provides excellent regio- and stereo-specificity, but requires expensive isoprenyl pyrophosphate substrates. Here we develop a flexible cell-free enzymatic prenylating system that generates isoprenyl pyrophosphate substrates from glucose to prenylate an array of natural products. The system provides an efficient route to cannabinoid precursors cannabigerolic acid (CBGA) and cannabigerovarinic acid (CBGVA) at > 1 g/L, and a single enzymatic step converts the precursors into cannabidiolic acid (CBDA) and cannabidivarinic acid (CBDVA) respectively. Cell-free methods may provide a powerful alternative to metabolic engineering for chemicals that are hard to produce in living organisms.

2.2 Introduction

Prenylated natural products are a large class of bioactive molecules with demonstrated medicinal properties. Examples include prenyl-flavanoids, prenyl-stilbenoids and cannabinoids (see Figure 2-1). Cannabinoids in particular show immense therapeutic potential with over 100 ongoing clinical trials as antiemetics, anticonvulsants, antidepressants and analgesics. Nevertheless, despite the therapeutic potential of prenyl-natural products, their study and use is limited by the lack of cost-effective production methods. Plant-derived prenyl-compounds are difficult to isolate due to the structural similarity of contaminating molecules, and the variable composition between crops. These challenges are further exacerbated when attempting to isolate low abundance compounds. Many chemical syntheses have been developed to address
the challenges associated with making prenylated natural products, but they are generally impractical for drug manufacturing due to the degree of complexity and low yields.

Microbial production is a useful alternative to natural extraction for prenylated natural products, but comes with many challenges such as the need to divert carbon flux from central metabolism and product toxicity to name a few. For example, prenyl natural products like prenylnaringenin, prenyl-resveratrol and cannabidiol (CBD) are derived from a combination of the metabolic pathways for fatty acid, isoprenoid, and polyketide biosynthesis. So high-level production requires efficient re-routing of long, essential and highly regulated pathways. Despite the challenges, many groups have engineered microbes to produce unprenylated polyketides, like naringenin, resveratrol and olivetolate, but at relatively low levels (110 mg/L, 391 mg/L and 80 mg/L respectively). Obtaining prenylated products is even more challenging because GPP is an essential metabolite that is toxic to cells at moderate concentrations, creating a significant barrier for high level microbial production. So, in spite of intense interest, to our knowledge there are no published reports of the complete biosynthesis of prenyl-flavonoids, prenyl-stilbenoids or cannabinoids in recombinant microbes.

Much recent effort has focused on alternative methods for cannabinoid production. Two groups have produced the polyketide cannabinoid intermediate, olivetolic acid (OA) at low levels in yeast (0.5 mg L\(^{-1}\)) or \(E. \ coli\) (80 mg L\(^{-1}\)), but did not prenylate OA or produce a cannabinoid from the biosynthesized OA. In other work, tetrahydrocannabinolic acid THCA was produced in cell extracts from either exogenously added geranyl-pyrophosphate (GPP) and OA in a two enzyme pathway or from cannabigerolic acid (CBGA) using a single enzyme. However, it is unclear how GPP or CBGA could be obtained at sufficient levels for economical production due to the high cost of these molecules.

Here we propose an alternative biological approach to prenylated natural product biosynthesis using a cell-free enzymatic platform we call synthetic biochemistry, which has shown great promise for the production of bio-based molecules. The synthetic biochemistry
approach frees us from worrying about the toxicity of products and intermediates, affords rapid design-build-test cycles, precise control of all system components, and complete flexibility in pathway design. Nevertheless, building highly complex systems involving dozens of enzymes, associated cofactors and myriad metabolites on a large scale outside the context of the cell is an enormous challenge. One of the keys to making commercially viable cell-free systems is reducing enzyme costs by employing stable enzymes that can last for long periods of time. Recently Zhang and co-workers converted maltodextrin into inositol at a 20,000 L scale in a 5 enzyme system using thermophilic enzymes purified by simple heating step\textsuperscript{25}, demonstrating that at least simple cell-free systems can reach industrial scale. Another key requirement is designing systems that effectively generate and recycle high energy cofactors (ATP, NAD(P)H) so that they can be used many times. We have previously reported a flexible enzymatic purge valve and rheostats for the regulating the supply of reducing equivalents and ATP\textsuperscript{26-28}, allowing us to build systems that run for many days and produce high titers of isobutanol and terpenes. Here we employ these concepts to develop cell-free production of a variety of prenylated compounds. We use glucose as a feedstock to produce GPP and optimize the system for the high-titer production of the cannabinoid compounds CBGA and cannabigerovarinic acid (CBGVA).

2.3 Results and Discussion

Construction of the cell-free prenylation pathway

Our synthetic biochemistry approach is outlined in Figure 2-1 and Figure 2-5 and expands on a system we developed previously for terpene production\textsuperscript{26}. First, glucose is broken down via a modified glycolysis pathway to produce high energy cofactors ATP and NADPH in addition to the carbon building block, acetyl-CoA using an alternative pyruvate oxidation pathway.\textsuperscript{26} The acetyl-CoA is then assembled into the prenyl-donor compound, GPP, via the mevalonate pathway using the ATP and NADPH produced from glycolysis. Importantly, a purge valve\textsuperscript{26} introduced into the glycolysis pathway balances NADPH production and consumption while maintaining carbon flux. The prenylation module then uses the GPP to prenylate exogenously added substrate to
yield the desired prenylated product. To expand the capabilities of our synthetic biochemistry platform we developed a prenylating system that employs a non-specific prenylating enzyme such as NphB, AtaPT, or NovQ to produce an array of prenyl-compounds derived from glucose\textsuperscript{29–31}. We then further engineered NphB using Rosetta to specifically prenylate OA.

As a first test of the system, we built the full cell-free system (23 enzymes) to generate GPP from glucose and employed wild-type NphB to prenylate its preferred substrate 1,6 dihydroxynapthalene (1,6 DHN; added exogenously). 1,6 DHN was added at the beginning of the reaction along with glucose. Up to \( \sim 400 \text{ mg/L (1.3 mM)} \) of prenylated product was obtained from 2.5 mM 1,6 DHN. However, increasing the 1,6 DHN concentration from 2.5 to 5 mM, decreased final titers 2-fold suggesting that 1,6 DHN inhibited one or more enzymes (Figure 2-2A). Enzyme assays revealed that pyruvate dehydrogenase (PDH) was inhibited by 1,6 DHN, as well as olivetol, resveratrol, and olivetolate (Figure 2-2B). Therefore, to engineer a general prenylation system, we sought to eliminate PDH.

To remove the need for PDH, we implemented an PDH bypass (Figure 2-1). In the PDH bypass, pyruvate is converted to acetyl-CoA using a pyruvate oxidase (PyOx) to produce acetyl-phosphate followed by the action of acetyl-phosphate transferase (PTA). The PDH bypass had two advantages. First, PDH is a large enzyme complex that is difficult to work with, so bypassing PDH streamlines enzyme production. More importantly, initial experiments revealed that the bypass is not subject to the inhibition seen at higher concentrations of 1,6 DHN. Once we confirmed the PDH bypass improved 1,6 DHN titers, we began to optimize the system as a general prenylation system. We varied co-factor concentrations, protein levels, and environmental conditions such as temperature and pH to identify the ideal set of conditions. Throughout this process we found that ATP, NADP\textsuperscript{+}, phosphate and NphB concentrations had the greatest impact on the final titer. As shown in Figure 2-2A, when we employed the PDH bypass, we found a 4-fold increase in titers of 5-prenyl-1,6 DHN when starting with 5 mM 1,6 DHN (Figure 2-2A). When
utilizing the PDH bypass system, approximately 50% of 1,6 DHN was converted in 24 hours, reaching a final titer of 705 ± 12 mg/L (Figure 2-2B).

We then tested the ability of the PDH bypass cell-free system to prenylate a variety of aromatic substrates (apigenin, diadzein, genistein, naringenin, olivetol, OA and resveratrol). All the aromatic substrates were prenylated using one or more of the tested prenyl transferases (Figure 2-2C and Figures 2-6 to 2-11). Thus, it is possible to produce a variety of prenylated natural products using a cell-free enzymatic system to generate the expensive co-substrates GPP and DMAPP. Further, the ease with which an exogenous substrate can be added to a synthetic biochemistry system is a great advantage because it is often not possible to add co-substrates exogenously to microbes since they cannot enter the cell.16

To test whether we could use synthetic biochemistry to produce high levels of therapeutically relevant prenylated products, we focused optimization efforts on cannabinoids due to the growing interest in new ways to make these medically important compounds. As shown in Figure 2-2D the initial system produced the cannabinoid precursor CBGA at a constant rate of 2.1 mg L⁻¹ hr⁻¹ over 72 hours and reached a final titer of only 132 mg L⁻¹ (Figure 2-2D).

Although the system produced CBGA, there were two problems. First, the turnover rate of the prenyl transferase NphB for CBGA production is extremely poor ($k_{cat} = 0.0021 ± 0.00008$ min⁻¹, Table 2-5). Second, prenylation of OA by NphB is highly non-specific, generating a major side-product, 2-O-geranyl olivetololate16. We therefore sought to improve CBGA production by enhancing the activity and specificity of NphB by design.

**Redesign of NphB to improve CBGA synthesis**

Briefly, OA was docked into the active site of the NphB crystal structure 32, then Rosetta was used to predict mutations that would improve OA binding. We narrowed the Rosetta results to a 22 construct library (see Tables 2-2 and 2-3), and screened for CBGA production (Table 2-4). We made several key observations during the initial screen, Figure 2-12: (1) Y288A (M1) and Y288N (M2) by themselves dramatically enhanced activity, as predicted by computation; (2) The
presence of Y288N in any construct decreased the enzyme yield suggesting Y288N may be a destabilizing mutation (making Y288A the preferred mutation); (3) The addition of G286S in the Y288N (M10) background appeared to improve activity further over Y288N (M2), suggesting that G286S could be another favorable mutation; (4) We noted an activity improvement of Y288A/F213N/A232S (M15) over Y288A/F213N (M5) suggesting that A232S may also be a favorable mutation. From these initial observations we constructed a focused library with all but one of the constructs in the second library exhibiting activity at least 100-fold higher than WT NphB in an endpoint assay (Figure 2-3B).

The best two constructs, M23 and M31, exhibited dramatically improved activity and specificity. Both had $k_{cat}$ values 1000-fold higher than WT NphB and both produce only the correct prenylated isomer, CBGA. As shown in Figure 2-3C, WT NphB produces CBGA, but the dominant product is a prenylated side-product, 2-O-geranyl olivetolate, whereas M23 makes CBGA almost exclusively. Overall, the designed enzyme is a much more active and specific CBGA synthase than WT NphB, and is easier to work with than the natural cannabis prenyltransferase, which is an integral membrane protein.\textsuperscript{16,33} Our soluble, CBGA synthase (M23) could potentially be applied in both cell-free and \textit{in vivo} systems to improve cannabinoid production.

\textit{Production of cannabinoid precursors}

With our designed CBGA synthase in hand (M23), we tested the ability to produce CBGA directly from glucose and OA using the full synthetic biochemistry system, including the PDH bypass (Figure 2-1). The initial productivity of the system using M23 was 67 mg L\textsuperscript{-1}hr\textsuperscript{-1} with a final titer of 744 ± 34 mg L\textsuperscript{-1} CBGA—100-fold faster and 21-fold higher titer than CBGA production using WT NphB (Fig 2-4A). We noted that with the mutant NphB enzyme, maximum titers were reached within 24 hours, after which production spontaneously stopped. In contrast, the system with the wild-type enzyme ran continuously for up to 4 days, suggesting enzymes and cofactors remain active and viable for longer periods of time, consistent with prior work\textsuperscript{26}. So what is stopping the reaction at the higher titers? We observed that reactions turned cloudy once ~500
mg L⁻¹ CBGA was produced. We collected the precipitate and identified a mix of enzymes in the precipitate by SDS-PAGE analysis (Figure 2-13), indicating high-levels of CBGA in solution may be causing enzymes to precipitate. We therefore sought to continually remove the product in situ during the reaction (a capability difficult to implement in living systems).

Initially a fixed volume nonane overlay was used for each reaction to extract CBGA. Unfortunately, CBGA is more soluble in water than nonane, limiting the amount of CBGA that can be extracted with a simple overlay. We therefore designed a flow system that would capture CBGA from the nonane layer and trap it in a separate buffered reservoir (Fig. 2-4B). By implementing this flow system we hoped to maintain a lower concentration of CBGA in the reaction vessel to mitigate enzyme precipitation. The flow system indeed improved the final titers to 1.25 ± 0.07 g/L, however enzyme precipitation still occurred at about 24 hours.

We next evaluated the system flexibility by replacing OA with divarinic acid (DA) to produce the precursor of many rare cannabinoids, CBGVA. We first tested whether our designed enzymes would be active with DA as the substrate. Kinetic analysis (Table 2-5) indicated that M31 effectively prenylates DA, with catalytic efficiencies 15-fold higher than M23 and 650-fold higher than WT NphB. We therefore utilized M31 to produce CBGVA from glucose and DA. As shown in Figure 2-4A, CBGVA was produced at a maximum productivity of ~107 mg L⁻¹ hr⁻¹, and reached a final titer of 1.74 ± 0.09 g L⁻¹, converting 92% of the added DA to CBGVA. The nonane flow system was not needed for the production of CBGVA because CBGVA was less potent in precipitating enzymes.

Production of the final cannabinoids

To illustrate the production of other cannabinoids from the central cannabinoids CBGA and CBGVA, we employed CBDA synthase to convert CBGA into CBDA and CBGVA into CBDVA. Conversion of CBGA into CBDA has been demonstrated by several groups¹⁷,34–36. In our case, we simply transferred the nonane overlay containing CBGA to an aqueous solution containing CBDA synthase, and indeed we were able to convert CBGA into CBDA at a constant
rate of 14.4 ± 0.8 mg L⁻¹ hr⁻¹ mg total protein⁻¹ over the course of 4 days converting 25% of the CBGA added to CBDA (Figure 2-4C). To our knowledge it is not known whether CBGVA can be converted into the rare cannabinoid CBDVA using the CBDA synthase. So, we added CBGVA, extracted from the cell-free system, to a reaction containing CBDA synthase. CBDVA was produced (Figure 2-14) by CBDA synthase at a rate of 7.1 ± 0.1 mg L⁻¹ hr⁻¹ mg total protein⁻¹ for 24 hours. We note that the cannabinoid acids can undergo spontaneous decarboxylation or heat induced decarboxylation to ultimately form additional bioactive cannabinoids cannabidiol (CBD) and cannabidivarin (CBDV). Thus, our system provides opportunities for ultimately producing a wide-variety of cannabinoids.

2.4 Conclusions

Our results demonstrate the power and flexibility of a cell-free approach, not only for the production of pure, therapeutically relevant cannabinoids and other prenylated natural products, but for bio-derived chemicals in general. Freedom from worries about cell viability allowed us to focus on pathway optimization rather than minimizing GPP toxicity, while the lack of a cell membrane barrier freed us to design a system with added aromatic molecules, which would not be possible in cells. Moreover, we could flexibly change the input from OA to DA to target rare cannabinoids without redesigning an entire pathway. Finally, it was straightforward to identify and focus our efforts on fixing the bottleneck steps. When we started this project we were only able to produce 9 mg/L of CBGA using the monoterpen pathway developed by Korman et al.²⁶ By introducing the PDH bypass and optimizing for co-factors, enzymes and environmental factors we were able to increase those titers to 132 mg/L. To improve titers further we engineered the NphB prenyltransferase, which further increased titers to 600 mg/L of CBGA. The final bottleneck was enzyme stability in the presence of CBGA, so by limiting the CBGA in the reaction vessel, we increased the titer to 1.25 g/L of CBGA, nearly a 140-fold improvement. Solutions were quickly implemented due to speedy design-build-test cycles, rapidly yielding results that far exceed published results using living cells. Like all new technology, the current system will need
additional technical developments to become commercially viable, but our results suggest that synthetic biochemistry can become a realistic option for producing bio-based chemicals.

2.5 Experimental Section

Chemicals and Reagents

Yeast hexokinase and Corynebacterium glutamicum catalase were purchased from Sigma Aldrich. Aerococcus viridians pyruvate oxidase was purchased from A.G. scientific. All cofactors and reagents were purchased from either Sigma Aldrich or Thermo Fisher Scientific, with the exception of olivetolic acid, which was purchased from Santa Cruz Biotechnology and divarinic acid, which was purchased from Toronto Research Chemicals.

Cloning and purification of enzymes

The NphB gene was purchased as a gene block from IDT DNA, and cloned into a pET 28(+) vector using the Gibson Assembly method. The remaining enzymes were amplified from genomic DNA or a plasmid, and cloned into pET28(+) using the same Gibson assembly method. All plasmids were transformed into BL21(DE3) Gold, and enzymes expressed in LB media with 50 µg/mL kanamycin. 1 L cultures were inoculated with 2 mL of a saturated culture in the same media, and grown to an OD_{600} of 0.5 – 0.8 at 37 °C. The cultures were induced with 1 mM IPTG, and expressed at 18 ºC for 16 hours. The cells were harvested by centrifugation at 2,500 x g, and resuspended in ~ 20 mL lysis buffer: 50 mM Tris [pH 8.0], 150 mM NaCl, and 10 mM imidazole. The cells were lysed using an Emulsiflex instrument. The lysate was clarified by centrifugation at 20,000 x g, and the supernatant was batch bound to 1 mL NiNTA resin for 30 mins at 4 ºC. The resin was transferred to a gravity flow column. The resin was washed with 10 column volumes of wash buffer: 50 mM Tris [pH 8.0], 150 mM NaCl, and 10 mM imidazole. The protein was then eluted with 2 column volumes of elution buffer: 50 mM Tris [pH 8.0], 150 mM NaCl, 250 mM imidazole and 30% (v/v) glycerol. Enzymes were flash frozen in elution buffer using liquid N₂, and the enzyme stocks were stored at -80 ºC.
**PDH Cell-free Reactions**

The PDH reactions were assembled in two parts. First the co-factors and substrates were combined in one tube, and the enzymes were combined in another. The reactions were initiated by mixing the co-factors and enzymes in a final volume of 200 µL. The final substrate and co-factor concentrations were as follows: 500 mM glucose, 1 mM 1,6 fructose bisphosphate, 4 mM ATP, 0.5 mM 2,3 bisphosphoglycerate, 0.5 mM NAD⁺, 1.5 mM CoA, 1.5 mM NADP⁺, 0.5 mM TPP, 6 mM MgCl₂, 10 mM KCl, 50 mM Tris pH 8.0 and 20 mM phosphate buffer pH 8.0, 5 mM glutathione and 0.5 – 5 mM 1,6 DHN. The enzyme amounts added to the reaction can be found in Table 2-1 (PyOx and PTA were not added to these reactions). The reactions were quenched at 24 hours.

**PDH Activity Assays**

PDH was assayed for activity in the presence of several aromatic polyketides. The vehicle control was 1% ethanol, and the activity was compared to an assay without the aromatic polyketides. The final reaction volume was 200 µL, and contained 2 mM NAD⁺, 2 mM CoA, 1 mM TPP, 5 mM MgCl₂, 5 mM KCl, 50 mM Tris pH 8.0, and 5 µL of 1.25 mg/mL PDH. The reactions were set up in a 96-well plate. The aromatic polyketides (dissolved in ethanol) were added to a final concentration of 1 mM and the ethanol control was added to a final concentration of 1% (v/v). The plate was incubated at room temperature for 10 minutes, and the reactions were initiated with 10 µL of 100 mM pyruvate. The absorbance at 340 nm was monitored for 10 minutes using an M200 spectrometer. Because the aromatic molecules had a background absorbance at 340 nm, the reactions were blanked using the reaction mixture and aromatic molecule, but instead of initiating the reaction with pyruvate, water was added. The initial rates were determined using the initial slope of a linear fit. The amount of NADH produced per unit time was calculated using Beer’s law, and the extinction coefficient of 6.22 x 10³ M⁻¹cm⁻¹. Reactions were performed in triplicate, and the average value and standard error were calculated.

**PyOx/PTA Cell-free Reactions**
The PyOx/PTA reactions were assembled in two pieces. First the co-factors and substrates were combined in one tube, and the enzymes were combined in another. The final co-factor and substrate concentrations in the 200 µL reaction were as follows: 500 mM glucose, 1 mM 1,6 fructose bisphosphate, 4 mM ATP, 0.5 mM 2,3 bisphosphoglycerate, 0.5 mM NAD+, 1.5 mM CoA, 3 mM mM NADP+, 0.5 mM TPP, 6 mM MgCl2, 10 mM KCl, 50 mM Tris pH 8.0 and 50 mM phosphate buffer [pH 8.0]. The amount of enzyme added to each reaction is detailed in Table 2-1. The co-factors and enzymes were mixed to initiate the reaction, and a 500 µL nonane overlay was added to the top. The reactions were incubated at room temperature shaking gently on a gel shaker.

When the aromatic substrate was the varied component 0.5 to 5 mM of the aromatic substrate was added to the reaction, and the reactions were quenched at 24 hours. When time was the varied component, 5 mM of 1,6 DHN was added, and separate reactions were quenched at ~12, 24, 48 and 72 hours.

Conditions for the olivetolate and divarinic acid reactions (produced CBGA and CBGVA respectively) were altered slightly. Optimization of the cannabinoid pathway showed that the same titers could be achieved with less glucose, so we reduced the glucose concentration to 150 mM (we did not test lower glucose concentrations). Additionally, increasing the NADP+ concentration to 6 mM and decreasing the ATP concentration to 1 mM led to higher titers of CBGA. The olivetolate concentration was set at 5 mM. The amount of NphB added to the reaction was variable. The data shown in Figure 2-2C utilized 1.5 mg/mL NphB, and the reactions were quenched at ~4, 8, 14, 24, 48, 72 and 96 hours. The data shown in Figure 2-4A was achieved with 0.5 mg/mL of WT NphB and M23 and M31 (for divarinic acid), and reactions were quenched at ~ 6, 9, 12, 24, 48, 72 and 96 hours.

The conditions were identical to the method above with the following exceptions, the final concentration of the aromatic substrates was 1 mM and the initial glucose concentration was 150 mM. Additionally, the final concentration of the prenyl-transferase was 1 mg/mL, and we tested
AtaPT, NovQ and NphB with apigenin, daidzein, genistein, naringenin and resveratrol. We also tested NphB with olivetol, olivetolate and 1,6 DHN. The reactions were quenched at 24 hours.

**Quenching reactions**

To quench the reactions, the aqueous and organic layer were transferred to a 1.5 mL microcentrifuge tube. The reaction vial was washed with 200 µL of ethyl acetate, which was then pooled with the reaction in the microcentrifuge tube. The samples were vortexed for 5-10 seconds and then centrifuged for 3 minutes at 16,060 x g. The organic layer was removed, and the remaining aqueous layer was extracted 2 additional times with 200 µL of ethyl acetate. For each sample the organic extract was pooled, and then evaporated using a vacuum centrifuge. The samples were re-dissolved in methanol for HPLC analysis.

*For olivetolate / CBGA*

Due to the observed protein precipitation, the CBGA reactions shown in Figure 2-4A were extracted in the presence of 0.12 g of urea (solid), to facilitate the extraction of CBGA. This was unnecessary for the WT NphB CBGA data in Figure 2-2C because the proteins did not precipitate.

**Quantification of products**

The reactions were fractionated by reverse phase chromatography on a C18 column (4.6 x 100 mm) using a Thermo Ultimate 3000 HPLC. The column compartment temperature was set to 40 °C, and the flow rate was 1 mL/min. The compounds were separated using a gradient elution with water + 0.1% TFA (solvent A) and acetonitrile + 0.1% TFA (solvent B) as the mobile phase. Solvent B was held at 20% for the first minute. Then solvent B was increased to 95% B over 4 minutes, and 95% B was then held for 3 minutes. The column was then re-equilibrated to 20% B for three minutes, for a total run time of 11 minutes.

The cannabinoids (CBGA, CBDA, CBDVA) were quantified using an external calibration curve derived from an analytical standard purchased from Sigma Aldrich. The 5-p-1,6-DHN and CBGVA NMR samples were used to generate an external calibration curve because authentic
standards were not available (see below). A known concentration of the standard was dissolved in water, and then extracted using the method detailed above.

**Quantify prenyl-products without authentic standards**

Due to the lack of authentic standards for the prenyl-products prenyl-apigenin, prenyldaidzein, prenyl-naringenin, prenyl-genistein, prenyl-resveratrol and prenyl-olivetol, we quantified the prenyl-products based on substrate consumption. To generate a standard curve, serial dilutions of each aromatic substrate were subjected to the reaction mix, but to prevent product formation the prenyl-transferase was left out. We used LC-MS to quantify the amount of substrate consumed by the reaction compared to the standard curve.

ESI-TOF measurements were carried out on a Waters LCT-Premier XE Time of Flight Instrument controlled by MassLynx 4.1 software (Waters Corporation, Milford MA). The instrument was equipped with the Multi Mode Ionization source operated in the electrospray mode. A solution of Leucine Enkephalin (Sigma Chemical, L9133) was used in the Lock-Spray to obtain accurate mass measurements. Samples were infused using direct loop injection on a Waters Acquity UPLC system. Samples were separated on a Waters Acquity UPLC system using an Acquity BEH C18 1.7 µm column (50 x 2.1 mm) and were eluted with a gradient of 30 - 95% solvent B over 10 min (solvent A: water, solvent B: acetonitrile, both with 0.2% formic acid (vol/vol)). Mass spectra were recorded from a mass of 300 - 2000 daltons.

**NMR Spectroscopy**

NMR spectroscopy was used to identify prenyl-products, and quantify 5-p-1,6-DHN. The PyOx/PTA cell-free system was used to produce prenyl-DHN. 200 µL reactions were pooled, and extracted 3 times with an equivalent amount of nonane and then the nonane was evaporated. The product of the reactions was suspended in 500 µL of deuterated methanol (CD$_3$OD), with 2 mM 1,3,5-trimethoxybenzene (TMB) as an internal standard. Spectra were collected on an AV400 Bruker NMR spectrometer. The amount of the prenylated compound in the sample was determined with reference to the internal TMB standard. We compared the proton signal from
TMB (3H, singlet) at 6.05 ppm with an aromatic proton corresponding to 5-p-1,6-DHN (1H, doublet) at 7.27 ppm.

NMR was also used to identify the product of the enzymatic system with divarinic acid as the aromatic substrate. The PyOx/PTA system was set up as detailed above, and the reactions were quenched at 24 hours. The reactions were extracted as detailed above, and analyzed on the HPLC. There was a new major peak at 6.7 minutes that we predicted to be the prenylated divarinic acid. We HPLC purified the peak, removed the solvent, and re-dissolved the pure component in 600 µL of CD$_3$OD. A proton spectrum collected with an AV500 Bruker NMR spectrometer was compared to a proton spectrum published by Shoyama et al for CBGVA to confirm that CBGVA was the main product. Based on the report by Shoyama et al the study by Bohlmann et al, we conclude that the prenylation of divarinic acid occurs at the C3 carbon of divarinic acid. Shoyoma et al published the chemical shifts of CBGVA in CD$_3$OD, so by direct comparison of our NMR spectra to the published chemical shifts we conclude that we produced CBGVA. This is further supported by the work conducted by Bohlmann, which suggests that if the prenylation occurred at the C5 site, we would observe a proton with a chemical shift around 5.8 ppm, which we did not observe, Figure 2-15.

**Rosetta Design to modify the binding pocket of NphB**

We placed olivetolate in the active site of NphB in six different starting positions denoted as Olivetolate P1-6 in Table 2-2. We ran ROSETTA 5 times for each olivetolate position for a total of 30 designs. The mutations predicted in each design are listed in Table 2-2. For each olivetolate position we chose a consensus set of mutations (i.e. the most frequently chosen residue) to evaluate further: Consensus Group A through F (Table 2-2). We then sought to evaluate the relative importance of each ROSSETTA suggested mutation. For each Consensus Group, we set the mutations back to WT residue, one at a time, and used ROSETTA to calculate the change in energy score (see Table 2-3). Those that caused the largest change in
energy were deemed to be the most important mutants to include in the library for experimental testing.

To model the olivetolic acid, we took the 4MX.sdf 3-D structure of olivetolate from the 5B09 crystal structure and added hydrogen atoms to the structure assuming pH 7 using Open Babel 2.3.1.\(^{39}\) A rotamer library was generated for olivetolic acid using the Bio Chemical Library (BCL) molecule: Conformer Generator 3.5 using the PDB library.\(^{40}\) Finally, the aromatic bonds were manually annotated into the file before generating the parameter file read by Rosetta using the script main/source/python/public/molfile_to_params.py in the Rosetta 3.7 release. The parameter file for geranyl s-thioldiphosphate (GST) was generated without a rotamer library using the GST.sdf file from the 1ZB6 crystal structure. The olivetolic acid molecule was then manually placed into the co-crystal structure of NphB with GST and DHN (1ZB6) with the DHN and crystallographic waters removed using pymol. The olivetolic acid was placed in 6 different positions in the active site with the plane of the olivetolate aromatic ring parallel to the GST alkyl tail and the desired prenylation site 3.7 angstroms away from the eventual carbocation mirroring the placement of DHN in the 1ZB6 crystal structure. Residues 49, 162, 213, 224, 232, 233, 234, 271, 286, and 288 were allowed to be any amino acid during the Rosetta design with other sidechains held in a fixed position and the backbone fixed. The designed residues were in direct contact with the olivetolate and not in direct contact with GST. The fixed backbone script main/source/bin/fixbb.static.linuxgccrelease from the Rosetta 3.7 release was run with the all possible rotamers (-ex4), using the input sidechains (-use_input_sc), sidechains minimized after design (minimize_sidechains), the linear memnode interaction graph (-linmem_ig 10), and both with and without the ligand weighted score function (-score:weights ligand). From the identical starting point each design was run 5 times using the -nstruct input.

**Initial NphB mutant library screening**

For screening of the initial library, we performed small scale expression and purifications. 25 mL of LB media with 50 µg/mL of kanamycin was inoculated with 25 uL of a saturated culture
of BL21 DE3 Gold harboring the NphB expression plasmid. The cultures were incubated at 37°C until the OD$_{600}$ reached 0.4 – 0.6. The expression of the NphB constructs were induced with the addition of 1 mM IPTG, followed by incubation for 18 hours at 18°C. Cells were harvested by centrifugation at 2500 x g. The pellets were re-suspended in 500 µL of lysis buffer: 50 mM [Tris pH 8.0], 150 mM NaCl, and 5 mM imidazole and lysed by sonication. The cell lysate was clarified by centrifugation at 20,000 x g for 10 minutes at 4°C, and the supernatant was incubated at 4°C with 50 µL of NiNTA resin. A 96-well spin column plate was used to purify the NphB constructs. The supernatant/resin was applied to the column and centrifuged for 2 mins at 500 x g. 500 µL of lysis buffer was then added, and the plate was centrifuged again for 1 minute at 500 x g. The protein was eluted using 200 µL of elution buffer (50 mM Tris [pH 8.0], 150 mM NaCl, 250 mM imidazole and 30% (v/v) glycerol).

The enzymes were assayed under the following conditions: 2.5 mM geranyl pyrophosphate, 5 mM olivetolate, 5 mM MgCl$_2$, 50 mM Tris pH 8.0, ~0.1 mg/mL NphB mutant in a final volume of 100 µL. All enzymes were first diluted to 0.5 mg/mL using elution buffer so the final concentration of imidazole was the same in each reaction. The reactions were incubated for 12 hours at room temperature, then extracted 3 times with 100 µL of ethyl acetate. The organic extract was pooled for each reaction and the solvent was removed using a vacuum centrifuge. The samples were redissolved in 100 µL of methanol and subjected to HPLC analysis.

**Focused NphB mutant library screening**

For the focused library, we performed 1 L scale expression and purification of the NphB constructs as described above. The enzymes were assayed under the following conditions: 2.5 mM GPP, 5 mM olivetolate, 5 mM MgCl$_2$, 50 mM Tris pH 8.0 and ~ 1 mg/mL of NphB enzyme in a final volume of 100 µL. The reactions were incubated at room temperature for 1 hour. 40 µL of each reaction was quenched in 80 µL of acetonitrile. The samples were centrifuged for 5 minutes at 16,060 x g, to remove precipitated proteins. The supematant was analyzed using HPLC as described above.
Enzyme Kinetic Parameters

The reactions were set up under the following conditions: 50 mM Tris [pH 8.0], 2.5 mM GPP, 5 mM MgCl\(_2\), ~27 µM enzyme, and olivetolate or divarinic acid was varied from 0.1 mM to 6 mM in a final volume of 200 µL. 40 µL of the reaction was quenched in 80 µl acetonitrile + 0.1% TFA, at the time intervals detailed below. The reactions were centrifuged for 5 minutes at 13,000 16,060 x g to pellet the protein, and the supernatant was analyzed using the HPLC method detailed above. The initial rate was plotted vs the concentration of substrate, and fit with the Michaelis-Menten equation to determine the kinetic parameters \(k_{cat}\) and \(K_M\) (OriginPro). Each Michaelis-Menten curve was performed in triplicate. The average and standard deviation of the kinetic parameters are reported. The time courses with olivetolate as the substrate were as follows: for WT, M1, M10 and M30 the time course was 3, 6, 9, and 12 minutes. For M25 the reactions were quenched at 1, 2, 4 and 8 minutes, and for M31 the reactions were quenched at 1, 2, 4 and 6 minutes.

The conditions were altered slightly to characterize the constructs with divarinic acid as the substrate. For M31, the time course was 0.5, 1, 1.5 and 2 minutes. For M23, the time course was 5, 10, 15 and 20 minutes, and for WT NphB the time course was 8, 16, 24 and 32 minutes. The enzyme concentration for the mutants was ~27 µM, and the concentration of WT NphB was ~ 35 µM.

GC-MS characterization of isomer profile for WT NphB and M23

Samples were dissolved in 200 µL of ethyl acetate. GC-MS measurements were carried out using an Agilent Model 7693 Autosampler, 7890B Gas Chromatograph, and 7250 Q-TOF Mass Selective Detector in the Electron Ionization mode. Sample injection was carried out in split mode with inlet temperature set to 280°C. Separation was carried out on an Agilent HP5-MS column with dimensions 30m x 250 µm x 0.25 µm. Ultra High Purity Grade He (Airgas) was used as carrier gas with the flow set to 1.1 mL/min in constant flow mode. The initial oven temperature was set to 120°C for 1 min followed by a 20°C/min ramp to a final temperature of 300°C which was
maintained for 4 min. A 3.0 min solvent delay was used. EI energy was set to 15 eV. The MSD was set to scan the 50 - 500 m/z range. Data collection and analysis were performed using Mass Hunter Acquisition and Qualitative Analysis software (Agilent).

Due to the increased temperature of the GC inlet, CBGA undergoes spontaneous decarboxylation as described by Radwan et al.,\textsuperscript{41} resulting in an M+ ion at 316 m/z. The retention time corresponding to the 316 m/z ion for the CBGA standard was 10.48 minutes.

**Nonane-flow system for the extraction of CBGA from solution**

A PyOx/PTA reaction was set up as detailed above. A 500 µL nonane overlay was added to the reaction in a 2 ml glass vial which was covered with 2 layers of breathable cell culture film. 2 18-gauge needles were inserted into a 15 mL falcon tube at the ~750 µL mark and the 3.5 mL mark. Luer locks to tubing connectors were connected to the needles and Viton tubing was connected to the other end of the luer lock. 18-gauge needles were connected to the other end of the tubing via a luer lock connector and inserted through the mesh covering so they were only touching the nonane layer and not the reaction. 2 mL of Tris buffer [pH 8.5] was added to the 15 mL conical tube, and 6 mL of nonane was added. The nonane was pumped through the system using a peristaltic pump (~ 1 mL/min) such that the nonane flowed from the top of the reaction, through the buffered solution (~18 cm tubing). The nonane pumped into the reservoir separated into the top layer of the 15 mL conical tube. The nonane from the top of the 15 mL conical tube was pumped into the top of the reaction vial (~55 cm tubing). This essentially diluted the CBGA throughout the system driving the diffusion of CBGA into the nonane layer and out of the reaction.

**Cloning CBDAS**

A gene block of CBDAS with the signaling peptide was ordered from IDT codon optimized for *Pichia pastoris*. The signal sequence was removed by PCR amplifying from the 28\textsuperscript{th} residue of the protein sequence (NPREN…) through the end of the protein, with overhangs compatible with the pPICZ\textalpha vector. The PCR product was cloned into the pPICZ\textalpha vector digested with EcoRI and XbaI using the Gibson cloning method. The product of the assembly reaction was
transformed into BL21 Gold (DE3) cells a clone with the correct sequence isolated. The plasmid was digested with Pmel for 2 hours, and then purified using the Qiagen PCR purification protocol. The plasmid was transformed into Pichia pastoris X33 using electroporation. Immediately following electroporation, the cells were incubated in 1 mL of cold 1 M sorbitol and 1 mL of YPD media without shaking for 2 hours. The cells were plated on YPDS plates with 500 µg/mL of zeocin. Colonies were screened using PCR for the presence of the CBDAS gene between the AOX1 promoter and terminator. For screening, the colonies were re-suspended in 15 µL of sterile water and 5 µL of the resuspended colony was transferred into a PCR tube with 0.2% SDS. The samples were heated for 10 minutes at 99ºC, and then 1 µL was used as the template for PCR. Six colonies with positive colony PCR hits were screened for the expression of CBDAS.

**CBDAS Expression Test**

The six colonies were grown overnight at 30ºC in 25 mL of BMGY to obtain a saturated culture. The overnight cultures were used to inoculate a 25 mL culture in BMGY media and grown to an OD of ~2. The cells were harvested by centrifugation at 2,000 x g for 10 minutes. The cell pellet was re-suspended in 90 mL of BMMY media, and incubated at 30ºC for 5 days. Each day, 1 mL of the culture was removed for SDS-PAGE analysis, and 500 µL of methanol was added to the remaining culture. On day 3 the cultures were screened for CBDAS activity. The 1 mL culture samples were centrifuged to pellet the cells (16,060 x g, 5 min). 50 µL of the media was used in a subsequent activity assay, and the remainder of the media was stored at -80ºC in addition to the cell pellet. The assay conditions were as follows: 100 µL of 200 mM citrate buffer, 100 µM CBGA, 5 mM MgCl₂, 5 mM KCl, 1 mM FAD and 50 µL of the expression media. in a final volume of 200 µL. The reactions were incubated overnight at room temperature and then extracted 3 times with 200 µL of ethyl acetate. The ethyl acetate extractions were pooled for each sample, and removed using a vacuum centrifuge. The samples were re-suspended in 200 µL of methanol and analyzed by HPLC. All clones produced active CBDAS.
The culture from three clones (~300 mL total), was collected to obtain CBDAS activity. The cells were pelleted by centrifuging at ~3,000 x g for 20 minutes at 4°C. Then the supernatant was passed through a 0.22 µm filter. The media was concentrated and buffer exchanged into 100 mM citrate buffer pH 5.0 using a 50,000 MWCO protein concentrator from Millipore.

**Production of CBDVA and CBDA**

To convert the precursors CBGA and CBGVA into CBDA and CBGVA respectively, a secondary reaction was set up with cannabidiolic acid synthase (CBDAS). To produce CBDA, a PyOx/PTA enzymatic system was set up as detailed above to produce CBGA. After 24 hours 200 µL of the nonane overlay from the CBGA reaction was transferred to a CBDAS reaction vessel. In the aqueous layer: 50 mM Hepes [pH 7.0], 5 mM MgCl₂, 5 mM KCl, 25 µM FAD, 0.1 mg/mL CBDAS concentrate. The reaction was incubated at 30°C with gentle shaking. Reactions were quenched at 12, 24, 48, 72 and 96 hours.

To produce CBDVA, HPLC purified CBGVA was converted to CBDVA using CBDAS. The final reaction volume was 200 µL, with 50 mM Hepes [pH 7.0], 5 mM MgCl₂, 5 mM KCl, 25 µM FAD and 0.1 mg/mL (total protein) of CBDAS concentrate. A 200 µL nonane overlay was added, and the reactions were incubated at 30 °C with gentle shaking. The reactions were quenched at ~ 24, 48, 72 and 96 hours.
<table>
<thead>
<tr>
<th>Enzyme Abb.</th>
<th>Full Name</th>
<th>Organism</th>
<th>Amount Added to Rxn (mg/mL)</th>
<th>Acquisitio n Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hex</td>
<td>Hexokinase</td>
<td>S. cerevisiae</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>Pgi</td>
<td>Glucose-6-phosphate isomerase</td>
<td>G. thermodenitrificans</td>
<td>0.48</td>
</tr>
<tr>
<td>3</td>
<td>PfkA</td>
<td>Phosphofructokinase</td>
<td>G. stearothermophilus</td>
<td>0.18</td>
</tr>
<tr>
<td>4</td>
<td>Fba</td>
<td>Fructose-1,6-bisphosphate aldolase</td>
<td>S. aureus</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>TpiA</td>
<td>Triose phosphate isomerase</td>
<td>G. stearothermophilus</td>
<td>0.16</td>
</tr>
<tr>
<td>6</td>
<td>Gap</td>
<td>Gald-3-P dehydrogenase</td>
<td>E. coli K12</td>
<td>0.07</td>
</tr>
<tr>
<td>7</td>
<td>mGap</td>
<td>Gald-3-P dehydrogenase D34A/L35R/T36K</td>
<td>G. stearothermophilus</td>
<td>0.18</td>
</tr>
<tr>
<td>8</td>
<td>NoxE</td>
<td>NADH Oxidase</td>
<td>L. lactis</td>
<td>0.25</td>
</tr>
<tr>
<td>9</td>
<td>Pgk</td>
<td>Phosphoglycerate Kinase</td>
<td>G. stearothermophilus</td>
<td>0.06</td>
</tr>
<tr>
<td>10</td>
<td>dPgm</td>
<td>Phosphoglycerate Mutase (2,3 BPG dependent)</td>
<td>E. coli K12</td>
<td>0.29</td>
</tr>
<tr>
<td>11</td>
<td>Eno</td>
<td>Enolase</td>
<td>E. coli K12</td>
<td>0.08</td>
</tr>
<tr>
<td>12</td>
<td>PykF</td>
<td>Pyruvate Kinase (FBP dependent)</td>
<td>E. coli K12</td>
<td>0.37</td>
</tr>
<tr>
<td>PDH</td>
<td>Pyruvate Dehydrogenase</td>
<td>AceE</td>
<td>E. coli K12</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AceF</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lpd</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>PyOx</td>
<td>Pyruvate Oxidase</td>
<td>A. viridans</td>
<td>1 U</td>
</tr>
<tr>
<td>14</td>
<td>PTA</td>
<td>Acetyl-phosphate transferase</td>
<td>G. stearothermophilus</td>
<td>0.06</td>
</tr>
<tr>
<td>15</td>
<td>PhaA</td>
<td>Acetyl-CoA acetyltransferase</td>
<td>R. eutropha</td>
<td>0.12</td>
</tr>
<tr>
<td>16</td>
<td>HMGS A110G</td>
<td>HMG-CoA Synthase</td>
<td>E. faecalis</td>
<td>0.18</td>
</tr>
<tr>
<td>17</td>
<td>HMGR</td>
<td>HMG-CoA Reductase</td>
<td>E. faecalis</td>
<td>0.16</td>
</tr>
<tr>
<td>18</td>
<td>MVK</td>
<td>Mevalonate Kinase</td>
<td>M. mazi</td>
<td>0.14</td>
</tr>
<tr>
<td>19</td>
<td>PMVK</td>
<td>Phosphomevalonate Kinase</td>
<td>S. pneumonia</td>
<td>0.2</td>
</tr>
<tr>
<td>20</td>
<td>MDC</td>
<td>Diphosphomevalonate Kinase</td>
<td>S. pneumonia</td>
<td>0.19</td>
</tr>
<tr>
<td>21</td>
<td>IDI</td>
<td>Isopentyl-PP Isomerase</td>
<td>E. coli K12</td>
<td>0.3</td>
</tr>
<tr>
<td>22</td>
<td>FPPS S82F</td>
<td>Farnesyl-PP synthase</td>
<td>G. stearothermophilus</td>
<td>0.09</td>
</tr>
<tr>
<td>23</td>
<td>NphB</td>
<td>Aromatic prenyltransferase</td>
<td>Streptomyces sp. CL190</td>
<td>Variable</td>
</tr>
<tr>
<td>24</td>
<td>CBDAS</td>
<td>Cannabidiolic Acid Synthase</td>
<td>C. sativa</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Ppase</td>
<td>Pyrophosphatase</td>
<td>G. stearothermophilus</td>
<td>0.11</td>
</tr>
<tr>
<td>26</td>
<td>Cat</td>
<td>Catalase</td>
<td>C. glutamicum</td>
<td>0.1 U</td>
</tr>
</tbody>
</table>

Table 2-1: Enzymes used in the enzymatic platform. The cells highlighted in yellow indicate proteins that were only added to the PDH reactions and not to the PyOx/PTA reactions.
Table 2-2: Rosetta suggested mutants varying initial olivetolate position. Olivetolate (OA) was placed into the active site at NphB based off the location of 1,6 DHN with the C3 carbon 3.7 angstroms above the C1 carbon of GPP. We tested 6 different OA starting positions, and ran each scenario 5 times. The resulting NphB mutants are shown in the table above.

Table 2-3: Evaluation of mutation significance according to the Rosetta energy score function. For each group of consensus mutants from Supplementary Table 3, we returned one mutation back to WT, and the remainder were not changed. We used Rosetta to calculate the difference in the energy score when the residue was set back to WT. We repeated that for each
Rosetta mutation, and the results are shown in the table above. The mutations contributing to the largest change in energy are highlighted in yellow.

<table>
<thead>
<tr>
<th>CONSTRUCT NAME</th>
<th>MUTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Y288A</td>
</tr>
<tr>
<td>M2</td>
<td>Y288N</td>
</tr>
<tr>
<td>M3</td>
<td>Y288A, F213H</td>
</tr>
<tr>
<td>M4</td>
<td>Y288A, F213N</td>
</tr>
<tr>
<td>M5</td>
<td>Y288N, V49S</td>
</tr>
<tr>
<td>M6</td>
<td>Y288S, V49N</td>
</tr>
<tr>
<td>M7</td>
<td>Y288A, V49S</td>
</tr>
<tr>
<td>M8</td>
<td>Y288N, V49T</td>
</tr>
<tr>
<td>M9</td>
<td>Y288N, I234T</td>
</tr>
<tr>
<td>M10</td>
<td>Y288N, G286S</td>
</tr>
<tr>
<td>M11</td>
<td>Y288N, F213N, V49G</td>
</tr>
<tr>
<td>M12</td>
<td>Y288A, F213N, I234T</td>
</tr>
<tr>
<td>M13</td>
<td>Y288S, F213N, V49N</td>
</tr>
<tr>
<td>M14</td>
<td>Y288N, F213G, I234T</td>
</tr>
<tr>
<td>M15</td>
<td>Y288A, F213N, A232S</td>
</tr>
<tr>
<td>M16</td>
<td>Y288N, F213N, A232S</td>
</tr>
<tr>
<td>M17</td>
<td>Y288N, F213G, V49T</td>
</tr>
<tr>
<td>M18</td>
<td>Y288N, V49S, V271N</td>
</tr>
<tr>
<td>M19</td>
<td>Y288N, F213N, V49S, V271N</td>
</tr>
<tr>
<td>M20</td>
<td>Y288N, F213G, V49T, V271H</td>
</tr>
<tr>
<td>M23*</td>
<td>Y288A, A232S</td>
</tr>
<tr>
<td>M24*</td>
<td>Y288A, G286S, A232S</td>
</tr>
<tr>
<td>M25*</td>
<td>Y288A, G286S, A232S, F213H</td>
</tr>
<tr>
<td>M27*</td>
<td>Y288V, G286S</td>
</tr>
<tr>
<td>M28*</td>
<td>Y288V, G286S, A232S</td>
</tr>
<tr>
<td>M30*</td>
<td>Y288A, A232S</td>
</tr>
<tr>
<td>M31*</td>
<td>Y288V, A232S</td>
</tr>
</tbody>
</table>

Table 2-4: NphB mutants. The NphB construct name, and the amino acid mutations are shown above. The asterisk (*) denotes constructs that were added after the first round was screened.
Table 2-5: Kinetic parameters for NphB mutants. Each construct was evaluated for its CBGA production with olivetolate (OA) as a substrate. WT, M23 and M31 were evaluated with divarinic acid as well. The symbol \((^b)\) indicates that divarinic acid was the substrate.
Figure 2-1: The synthetic biochemistry platform for the production of prenyl-natural products. First, glucose is broken down into pyruvate through a glycolysis pathway modified to regulate NADPH levels (22) (12 enzymatic steps). Then, either PDH or the PDH bypass converts pyruvate into acetyl-CoA. Acetyl-CoA is converted into GPP via the mevalonate pathway (8
enzymatic steps). By varying the aromatic prenyltransferase (aPT) and aromatic substrate we are able to produce various prenyl-flavonoids and prenyl-stilbenoids using the same central pathway. We developed variants of the prenyltransferase NphB (dNphB) to produce CBGA or CBGVA. CBGA is converted to cannabidiol and CBGVA is converted to cannabidivaric acid via cannabidiolic acid synthase (CBDAS). It is possible to produce other cannabinoids by using different cannabinoid synthases (THCAS and CBCAS).

Figure 2-2: Development of a PDH bypass for the prenylation of aromatic polyketides. (A) The comparison of the final titers achieved with the full pathway utilizing PDH (PDH system - gray trace) and the PDH bypass system (blue trace) at different concentrations of 1,6 DHN. (biological replicates, n=3). (B) The activity of pyruvate dehydrogenase (E. coli PDH) measured in the
presence of various aromatic polyketides and 2% ethanol (vehicle) (biological replicates, n = 3). (C) Various aromatic substrates were added to the pathway with either NphB, AtaPT or NovQ prenyltransferase (biological replicates, n=3). The result is a variety of C5 and C10 prenyl-natural products. (* indicates titer not determined) (D) Production of 5-prenyl-1,6 DHN (blue trace) over time compared to a separate reaction to produce CBGA (green trace). Both reactions utilized the PDH bypass and WT NphB (biological replicates, n=3).

Figure 2-3: Engineering NphB to improve CBGA production. (A) A model of olivetolate in the active site of WT NphB (1ZB6). Residues highlighted in yellow and purple were allowed to vary during the design process. The residues in yellow had the largest effects on activity with OA and were the positions targeted in the focused library. B) The results of an activity assay to determine the approximate activity of NphB mutants with olivetolate as the substrate. The fold improvement is an average of triplicate reactions with GPP (2.5 mM), olivetolate (5 mM), MgCl2(5 mM), and 1 mg/mL of WT or mutant NphB (biological replicates, n=3). (C) GC-MS chromatograms of the full pathway reaction products using M23 and WT NphB compared to a CBGA standard.
Figure 2-4: Evaluation of the cell-free prenylation system for the production of various cannabinoids. (A) The cell-free enzymatic production (from glucose) of cannabinoid precursors over time. CBGA production using M23 is shown in the dark aqua trace and WT NphB in the light green trace. The production of CBGVA using M31 is shown in the dark blue trace. The concentration of WT, M23 and M31 NphB was fixed at 0.5 mg/mL (biological replicates, n=3). (B)
Using a nonane flow CBGA capture system, we were able to obtain a higher titer of CBGA (1.2 g/L). The nonane layer was exchanged using a peristaltic pump, which circulated the nonane in the direction indicated by the arrows. (C) Production of CBDA over time using CBDAS (biological replicates, n=3).

**Figure 2-5: Full reaction schematic for the enzymatic synthesis of cannabinoids.** Glucose is broken down to pyruvate through a modified glycolysis pathway (dark blue) that includes a purge valve system. The purge valve (boxed) allows carbon flux to continue through the glycolysis pathway without building up excess NADPH. Pyruvate is converted to acetyl-CoA through the PDH bypass outlined in light blue. Acetyl-CoA is then converted into GPP via the mevalonate pathway (aqua). Finally, the GPP from the mevalonate pathway is used to prenylate aromatic...
polyketide. Shown here is the prenylation of olivetolate to produce CBGA; however, olivetolate could be replaced with a wide range of aromatic substrates to generate various prenylated products. A prenylated product like CBGA can be converted into a variety of cannabinoids. An example of the conversion of CBGA into CBD by the action of CBDAS and a spontaneous decarboxylation is shown.

Figure 2-6: LC-MS chromatogram for prenysl-apigeninLC-MS chromatogram with extracted for masses corresponding to prenyl-apigenin products. Each panel is labeled for the enzyme that produced the prenyl-products, and the m/z that corresponds to the prenyl-products. AtaPT and NovQ demonstrate both mono and di-prenylation.
Figure 2-7: LC-MS chromatogram for prenyl-daidzein. LC-MS chromatogram with extracted for masses corresponding to prenyl-daidzein products. Each panel is labeled for the enzyme that produced the prenyl-products, and the m/z that corresponds to the prenyl-products. NovQ demonstrate both mono and di-prenylation, whereas AtaPT demonstrates only diprenylation.
Figure 2-8: LC-MS chromatogram for prenyl-genistein LC-MS chromatogram with extracted for masses corresponding to prenyl-genistein products. Each panel is labeled for the enzyme that produced the prenyl-products, and the m/z that corresponds to the prenyl-products. AtaPT and NovQ demonstrate both mono and di-prenylation.
**Figure 2-9: LC-MS chromatogram for prenyl-naringenin.** LC-MS chromatogram with extracted masses corresponding to prenyl-naringenin products. Each panel is labeled for the enzyme that produced the prenyl-products, and the m/z that corresponds to the prenyl-products. AtaPT and NovQ demonstrate both mono and di-prenylation, NphB demonstrated both mono and di-geranylation.
Figure 2-10: LC-MS chromatogram for prenyl-olivetol. LC-MS chromatogram with extracted for masses corresponding to prenyl-olivetol products. This panel is labeled with the enzyme that produced the prenyl-products, and the m/z that corresponds to the prenyl-products.
Figure 2-11: LC-MS chromatogram for prenyl-resveratrol. LC-MS chromatogram with extracted for masses corresponding to prenyl-resveratrol products. Each panel is labeled for the enzyme that produced the prenyl-products, and the m/z that corresponds to the prenyl-products. AtaPT and NovQ demonstrate both mono and di-prenylation, and NphB demonstrates both mono and di-geranylation.
Figure 2-12: Activity screen of NphB constructs. NphB constructs from the initial round were expressed and purified and assayed for CBGA production. The constructs are shown on the x-axis, and their activity relative to WT activity is shown on the y-axis.

Figure 2-13: SDS-PAGE analysis of CBGA cell-free system precipitate. Lane 1: 1 kb DNA ladder from BioPioneer. Lane 2: Enzyme mastermix before it was added to the reaction. Lane 3: The washed precipitate from the reaction at 24 hours.
Figure 2-14: Cannabidiolic acid synthase (CBDAS) converts CBGVA into CBDVA. HPLC chromatogram of reaction extracts of CBGVA in the presence (red) and absence (pink) of CBDAS compared to a CBDVA standard (black).
Figure 2-15: Proton NMR Spectrum of CBGVA in CD$_3$OD. The entire spectra is displayed, and the inlet zooms in on the aromatic proton at ~6.2 ppm.

2.6 References


Chapter 3: An enzymatic synthesis of aromatic polyketides, and integration into the enzymatic synthesis of cannabinoids
3.1 Abstract

Due to the excitement surrounding the medicinal properties of cannabis, there has been a push to identify a more sustainable method for the production of the bioactive components. While most groups are focused on engineering microbes for the production of these molecules, we present a synthetic biochemistry alternative. Herein, we constructed a platform for the production of cannabinoids. We started with a simple system for the production of the aromatic polyketide intermediates, olivetolic acid (OA) and divarinic acid (DA) from inexpensive inputs. The synthetic biochemistry platform allowed us to identify that OA and DA were inhibiting several enzymes in the pathway. By including additives that reduced the inhibitory effect of OA and DA, we reached a titer of nearly 220 mg/L (0.98 mM) of OA and 180 mg/L (0.92 mM) of DA, nearly 3-fold more than previous reports. To build on the work from Chapter 1, we constructed a pathway with 32 enzymatic steps to convert glucose, malonate and hexanoate into CBGA, and reached a titer of 42 mg/L (~120 µM), 5-fold higher than previous work.

3.2 Introduction

Cannabinoids are an exciting class of secondary metabolites produced by the cannabis plant.\textsuperscript{1} Two cannabinoids, tetrahydrocannabinol (THC) and cannabidiol (CBD) are already FDA approved drugs. THC is approved to manage side effects associated with chemotherapy, like nausea and decreased appetite, and CBD is approved to manage seizures for patients with Lennox-Gastaut Syndrome and Dravet Syndrome, severe forms of epilepsy.\textsuperscript{2} However, with over one hundred ongoing clinical trials these two molecules are still being evaluated for many other medicinal properties. In addition to these well-known naturally abundant cannabinoids there are over 110 unique cannabinoids produced at lower levels in cannabis, just waiting to be studied.\textsuperscript{3} Unfortunately their low natural abundance makes them difficult to isolate.

Currently, the main source of cannabinoids is the cannabis plant. The extraction of THC and CBD from a plant source is economically feasible because they are naturally abundant; however this plant-based extraction may not be the best approach for some of the cannabinoids
with a low natural abundance. Secondary metabolite production is highly dependent on environmental conditions like temperature, light, moisture, insects, mold and other species in the plant environment. These factors can be extremely difficult to control when cultivating plants, which can lead to inconsistent production of secondary metabolites, like cannabinoids. Additionally, to produce pharmaceutical grade cannabinoids, the plant extracts must undergo a rigorous purification to remove all other cannabinoids and terpenes. This process can be time consuming, expensive, and lead to product loss. For the more abundant cannabinoids, these issues are not so dire, but for low abundance cannabinoids it can be a very expensive and inefficient process.

Due to the high demand for pharmaceutical grade cannabinoids and challenges with plant-based production several companies and research groups have been searching for alternatives. The majority have turned to engineered microbes, but this approach can be very challenging due to the complexity of the biosynthetic pathway. The full cannabinoid biosynthesis is shown in Figure 3-1, which can be broken down into three modules. The first is the isoprenoid module, which produces the isoprenoid geranyl pyrophosphate (GPP) from acetyl-CoA via the mevalonate pathway. The second module is the aromatic polyketide biosynthesis, which utilizes a Type III polyketide synthase (PKS) to convert hexanoyl-CoA and malonyl-CoA (derived from acetyl-CoA) into olivetolic acid (OA). The third module is the cannabinoid biosynthesis, in which the products from module 1 and module 2 are combined to yield cannabigerolic acid, which is then converted into the final cannabinoid by a cannabinoid synthase. Despite the difficulty, there are several reports of cannabinoid production in microbes. Most groups focused on engineering a single module, but one group was able to recapitulate the entire pathway in yeast. The results are summarized in Table 3-1.

While difficult, metabolic engineering could make it easier to produce rare cannabinoids and cannabinoid analogs by adding different fatty-acid substrates. For example, the substrate hexanoate is converted into OA, the precursor to the common cannabinoids (cannabidiol,
tetrahydrocannabinol), but butyrate is converted into DA, which is the precursor to the rare divarin cannabinoids that have a lower natural abundance. The enzymes in the cannabinoid biosynthesis are promiscuous, so a simple change in substrate can produce a different cannabinoid.^[10]

Despite considerable effort, cannabinoid titers in microbes have been low, most likely due to the competition between the cannabinoid path and central metabolism. It requires 9 equivalents of acetyl-CoA to produce just 1 cannabinoid equivalent. Acetyl-CoA is an essential molecule for fatty acid biosynthesis and sterol biosynthesis, and therefore it is difficult to divert flux away from those pathways. This problem is further exacerbated because the acetyl-CoA needs to be split between the aromatic polyketide module and the isoprenoid module, which is difficult to do efficiently.[^11] In addition to competing with primary metabolism, the pathway is energy intensive, requiring 7 ATP to produce a single cannabinoid. Nevertheless, due to the complexity of microbial metabolism it is difficult to know exactly where the pathway bottlenecks are located.

Some of the major challenges associated with metabolic engineering are simplified in a cell free synthetic biochemistry platform: (1) there are no endogenous pathways to compete with, so all of the flux is directed into the target pathway, (2) enzyme levels are easily adjusted to alleviate bottlenecks and balance flux through pathway branches (i.e. isoprenoid branch and OA branch) (3) since there is no background metabolism it is easier to identify why systems stop, and implement solutions to address those problems. A synthetic biochemistry approach has been successful for other classes of natural products, and therefore may be amenable for cannabinoid production.[^12]^1[^13] The focus of Chapter 1 was to develop a synthetic biochemistry platform for the prenylation of OA or DA (isoprenoid module and cannabinoid module). Herein, we develop an aromatic polyketide module, and integrate that with the work from Chapter 1 to design a complete in vitro enzymatic synthesis of cannabinoids from glucose, malonate and hexanoate.

### 3.3 Results and Discussion

*Constructing the enzymatic pathway for aromatic polyketide (AP) production*
Previous reports define the enzymatic steps from acetyl-CoA and hexanoate.\textsuperscript{8,9} We constructed a synthetic biochemistry system with 6 enzymes to produce OA and DA (Figure 3-2). Due to challenges expressing large complexes like acetyl-CoA carboxylase (ACC), we opted to use a malonyl-CoA synthetase (MatB) from \textit{Rhodopseudomonas palustris} to generate malonyl-CoA (MalCoA) from malonate. To produce hexanoyl-CoA, we chose the acyl activating enzyme 3 (AAE3) from \textit{Cannabis sativa}. Previous studies indicate that AAE3 is active for both hexanoate and butyrate, and is faster than acyl activating enzyme 1 (AAE1), which has also been used \textit{in vivo} to produce hexanoyl-CoA.\textsuperscript{9,10} To finish the pathway we used olivetol synthase (OLS), identified by Taura et al, and the olivetolic acid cyclase (OAC) identified by Gagne et al, both from \textit{Cannabis sativa}.\textsuperscript{8,14} While OLS is predicted to be the Type III polyketide synthase responsible for the production of OA and DA in cannabis plants, in \textit{E. coli}, \textit{S. cerevisiae} and \textit{in vitro}, this enzyme synthesizes several byproducts shown in Figure 3-2. Unfortunately, there are no other polyketide synthases reported to produce OA or DA, so in addition to the production of OA and DA, our system also generates pentyl diacetic acid lactone (PDAL), hexanoyl triacetic acid lactone (HTAL), and olivetol. Due to the promiscuity of the enzymes it is possible to produce both OA and DA from this pathway. To produce OA, hexanoic acid is added, but to produce DA, butyric acid is added.

To provide ATP in the system, we included an ATP recycle module comprised of creatine kinase (CPK) and adenylate kinase (ADK). The AMP generated from AAE3 and MatB was converted back into ADP using ADK, and then the ADP was converted back into ATP using creatine kinase and creatine phosphate. To prevent the build-up of pyrophosphate, we included a pyrophosphatase (Ppase) from \textit{Geobacillus stearothermophilus}. To make this aromatic polyketide (AP) module compatible with the system reported in Chapter 1, the pathway was optimized at pH 8.0.

\textit{Production of OA and DA over time}
We determined the productivity of the system with hexanoate and butyrate as substrates. The reactions were quenched at 5, 10, 20, 40, 60, 90, 120, 150, 180, 210, 240, 270 and 300 minutes, and the samples were analyzed by HPLC. The results are shown in Figure 3-3. OA is produced at nearly $0.8 \pm 0.3$ mg/L/min (270 µM/L/hr) for the first 150 minutes, reaching a final titer of $148 \pm 34$ mg/L. We observed a lag in DA production for the first 40 minutes, with the highest productivity, $0.62 \pm 0.1$ mg/L/min (250 µM/L/hr) occurring between 40 minutes and 120 minutes. The production of DA eventually stops at approximately 240 minutes, reaching a final titer of $78 \pm 12$ mg/L.

**Product inhibition of the AP module**

To determine why OA production stopped, we screened the effect of various metabolites on the pathway (data not shown). The screen identified OA as an inhibitor of the DA pathway. To further explore the inhibition, we titrated 0.25 to 5 mM OA into reactions producing DA, and we found that reactions with even 500 µM OA reduced DA production 90% (Figure 3-4), indicating that OA is a potent inhibitor of one or more enzymes in the pathway. We also tested the effect of 0.25 mM to 5 mM DA on OA production. DA was a less potent inhibitor of the AP production pathway, requiring 5 mM DA to reduce OA production 75%. It is important to note that AAE3, OLS and OAC are promiscuous enzymes, and the precursors to divarinic acid are not ideal substrates for these enzymes. Therefore, it is possible that the OA inhibition effect is exaggerated in this dataset because even a slight decrease in enzyme activity could dramatically reduce precursor concentrations and alter the kinetics of subsequent steps.

**Identification of enzymes inhibited by OA and DA**

To further elucidate the OA and DA inhibition, we assayed MatB, AAE3, ADK and CPK in the presence of 5 mM OA or 5 mM DA (Fig 3-5A). OA is a fairly potent inhibitor of MatB, AAE3 and ADK, reducing activity to 24%, 8% and 12% respectively in the presence of 5 mM OA compared to the control. To determine if any of these enzymes were the cause for the steep drop
in DA production, we also tested OA inhibition at 250 µM (Figure 3-5B). While we observe a 35% decrease in ADK activity at 250 µM OA, it is unlikely the sole cause of the decrease in DA production (Figure 3-5B). So, we tested OLS for OA and DA inhibition at 250 µM and 1 mM respectively. At 250 µM OA, OLS activity decreased 80%, which is most likely the cause for the significant drop in DA production at 250 µM OA. DA was a less potent inhibitor of OLS, requiring 1 mM DA to reduce activity 50%. The product inhibition observed in the AP production pathway will be a major limitation moving forward.

**OLS and OLS/OAC activity assays**

While testing conditions for OLS inhibition assays, we identified some interesting characteristics. We observed that OLS exclusively produces PDAL at pH 8.0. To ensure the enzyme was active and functioning properly, we tested the conditions reported by Taura et al.\textsuperscript{15} The malonyl-CoA, hexanoyl-CoA, and OLS concentrations were the same, but we used citrate buffer to maintain a pH of 5.5. At pH 5.5 the OLS assay appeared normal, producing olivetol as well as HTAL and PDAL (Figure 3-6A and 3-6B). We also tested the OLS/OAC coupled assay for OA production at pH 5.5 and pH 8.0 and observed a similar trend. At pH 8.0 the OLS/OAC assay produced little to no OA. This result was odd because the OA production pathway produced 40 mg/L of OA in 20 minutes at pH 8.0, indicating that OLS and OAC are active at pH 8.0 and capable of producing OA.

There are several potential reasons why the OA production platform works better than the individual activity assay. The most likely reason is the concentration of substrates is significantly higher in the production platform than in the activity assay. The production platform could yield up to 5 mM hexanoyl-CoA and 15 mM malonyl-CoA, whereas the OLS activity assay only had 100 µM and 200 µM respectively. It is possible that the higher concentration of substrate improved the kinetics of the second OLS step, allowing for olivetol or OA formation. Another possibility is that there was a metabolite or buffer component that activated the OLS enzyme at pH 8.0. We therefore tested various metabolites and buffer components and observed that the addition of 20
mM acetyl-phosphate (AcP) dramatically improved the OLS/OAC activity at pH 8.0, shown in Figure (3-6B). The addition of 20 mM AcP produced the same amount of OA as the OLS/OAC assay at pH 5.5.

**AP module with additives**

After identifying AcP as a potential activator of OLS, we checked the literature for other known activators of Type III polyketide synthases. While we did not find any other instances of Type III PKS being activated by AcP or molecules like it, several studies suggested that Type III PKSs are activated by bovine serum albumin (BSA).\(^\text{16,17}\) So, we tested the effects of AcP and BSA on our OA production platform. Figure 3-7 shows that 5 mM AcP does not have an effect on the products made, and BSA increases the production of PDAL, with a small increase in OA production. However, when BSA and AcP are added together there is a dramatic increase in OA production, and the ratio of the side products to OA is also improved compared to the BSA only reaction.

We repeated the time course experiments (Figure 3-8A and B) and the OA and DA inhibition experiments (3-8C and D) in the presence of the additives, 20 mg/mL BSA and 5 mM AcP. Not only did we observe a 2-fold increase in rate of OA and DA production and a ~2-fold increase in final titer, but we observed a decrease in the product toxicity of OA and DA. Although the mechanism of activation for AcP and BSA in this system is not known, we offer the following hypothesis. BSA is able to stabilize enzymes in the pathway, so they are less susceptible to OA and DA inhibition, while AcP either stabilizes the triketide intermediate, or activates the second OLS step.

**Initial add-in experiments for OA production path**

While the additives were able to reduce the inhibition effects of OA and DA on the AP production pathway, the pathway still stopped with only 20% of the substrates converted into OA or DA. To investigate why, we set up add-in experiments. Several identical reactions were set up, and at 2 hours (when the reactions normally stopped) we added in various enzymes, co-factors
and substrates. The reactions were incubated for another hour, and analyzed using the HPLC. The addition of malonate and hexanoate, ATP, creatine phosphate, or any of the enzymes did not have a dramatic effect (no change over 1.25-fold). This data may suggest the issue is the accumulation of a toxic or inhibitory metabolite because adding in the various components had little to no effect. However, this is only a single replicate, so it is difficult to know for sure what increase is significant.

**Design of the full cannabinoid pathway in vitro**

Based on the data from Chapter 1, we already had an isoprenoid system that was able to generate the precursor GPP. From the experiments above, we have a system capable of producing the AP precursors, OA and DA. Thus, as a simple initial test for building a full pathway, we simply combined the two systems to determine if CBGA could be produced. We note that the combined system is unsustainable because insufficient ATP is generated. In particular, the system requires 14 mol ATP (6 for Mev path, 8 for OA path) to make 1 mol CBGA, but glycolysis only produces 6 mol of ATP. Since MatB is used to produce malonyl-CoA, we called this the MatB path (Figure 3-9A). In spite of the ATP imbalance, combining the OA production pathway with the isoprenoid pathway from Chapter 1 yielded sufficient ATP to generate 12 mg/L of CBGA.

To build a more sustainable pathway, we needed to reduce the amount of ATP required per cycle. One option, commonly used *in vivo* is acetyl-CoA carboxylase (ACC), which converts acetyl-CoA into malonyl-CoA. However, large enzyme complexes, like ACC, can be difficult to express and purify. Buried deep in the literature was a second option, malonyl-CoA decarboxylase. As a complex this enzyme decarboxylates malonyl-CoA, however the alpha subunit (MdcA) possess transferase capabilities. So the alpha subunit alone could transfer the CoA from acetyl-CoA to malonate to yield acetate and malonyl-CoA. This was an exciting option because the isoprenoid pathway already produced acetyl-CoA, and the enzyme did not require ATP. Using MdcA changes the ATP stoichiometry dramatically, instead of being net -8 ATP per cycle, using the transferase yields +1 ATP per cycle (Figure 3-10B). The remaining challenge was
to balance the flux through the mevalonate pathway with the flux through the OA pathway. Reducing the amount of thiolase in the reaction led to a final titer of 42 mg/L of CBGA. Initial experiments suggest that AMP build up is the reason the reaction stops.

3.4 Conclusions

Herein, we designed an AP module to convert hexanoate and malonate into the cannabinoid precursors OA and DA, and then added the AP module to the cannabinoid and isoprenoid modules from Chapter 1 to construct a complete cannabinoid biosynthesis in vitro. Throughout this process we identified several factors that would impact the production of cannabinoids in engineered microbes as well as in vitro. First, both OA and DA are able to inhibit an assortment of enzymes indicating that increased concentrations of OA and DA in vivo could be toxic to cells by inhibiting essential pathways. However, the major hurdle for in vivo or in vitro production of cannabinoids is the inhibition of OLS at relatively low OA and DA concentrations. Without alleviating the inhibition, it will not be possible to reach high titers. We demonstrate that BSA and AcP are able to reduce the effects of inhibition and improve AP titers in vitro, but this solution would be challenging to implement in vivo. The byproducts also pose a major challenge, diverting a substantial amount of the precursors away from the desired product. Unfortunately there is currently no replacement for the OLS enzyme to my knowledge.

After alleviating some of the inhibitory effects of OA and DA, the AP system still stopped prematurely. While it would be difficult to identify the cause in vivo, our in vitro system is much simpler, and so we should be able to identify the faulty component. Preliminary data suggests that a toxic component builds up, but we are still working to confirm those results and identify the toxic component. Overall, the synthetic biochemistry approach yielded 3-fold more OA than previous reports at $237 \pm 17.9$ mg/L, and produced $181.5 \pm 3.7$ mg/L DA.

In addition to designing the AP pathway, we also incorporated it in a full cannabinoid biosynthetic pathway. The pathway was designed to convert glucose, malonate and hexanoate into cannabigerolic acid (CBGA). While this data is preliminary, it demonstrates that we can
assemble long and complex enzymatic pathways (32 steps) to produce natural products. Our titers are 5-fold higher than those reported in vivo, and there is still a lot of optimization to be done with potential for significant improvement.

3.5 Future Directions

There is a growing interest in cannabinoids and their medicinal properties, but current methods for production still depend on the natural source. While this is sufficient for high abundance cannabinoids, it is inefficient and expensive to isolate low abundance cannabinoids. Therefore, we need alternative production methods for these molecules. A bio-based approach poses some advantages over chemical synthesis. For example, the approach does not rely on toxic chemicals, and the biosynthesis is modular so it is easy to produce a variety of products from the same pathway.

The work in this thesis indicates that a synthetic biochemistry platform, if optimized further, may be an alternative production method. Future work would require identifying pathway bottlenecks, improving enzyme stability, and implementing solutions to address both the inhibition of OLS and the production of byproducts (protein engineering or homologues), but previous studies indicate that high level production of natural products is possible. As of right now it is difficult to know which bio-based approach will be the most successful, however the data in this thesis suggests that the synthetic biochemistry platform is definitely a contender.

3.6 Experimental Section

Chemicals and reagents

Divarinic acid and olivetolic acid standards were purchased from Enamine and Toronto Research Chemicals respectively, and cannabigerolic acid (CBGA) standard was purchased from Sigma Aldrich. Co-factors were purchased from either Thermo Fisher Scientific, Sigma Aldrich or Santa Cruz Biotechnology. *Aerococcus viridians* pyruvate oxidase was purchased from AG Scientific, and *Corynebacterium glutamicum* catalase, Bovine Serum Albumin (BSA), S.
*cerevisiae* hexokinase and pyruvate kinase with lactate dehydrogenase (PKLDH) were purchased from Sigma Aldrich.

**Cloning and purification of enzymes**

Several enzymes used in this study were cloned and expressed in prior studies (XX citation). The new enzymes MatB, AAE3, OLS, OAC and NphB M31 Pross 10 were synthesized by Twist Bioscience, and were cloned into pET28 (+) at the Nde1, Xho1 restriction sites using the Gibson Assembly method described previously (XX citation). MdcA was synthesized and cloned into the pET28(+) vector with Nde1/Xho1 restriction sites by Twist Bioscience.

The majority of the enzymes were expressed in *E. coli* BL21 (DE3) Gold, with the exception of OLS and MdcA which were expressed in the C43 Walker strain. Autoclaved LB media with 50 ug/mL kanamycin was inoculated with 1 mL of saturated culture, and grown to an OD$_{600}$ of 0.6-0.8. Protein expression was induced by adding 1 mM of IPTG, and the cultures were incubated overnight at 18 ºC. The cells were harvested by centrifugation at 2,500 x g, and resuspended in 20 mL of binding buffer (50 mM Tris [pH 8.0], 150 mM NaCl and 10 mM imidazole). The cells were lysed using an Emulsiflex instrument, and the lysate was clarified by centrifugation at 20,000 x g. NiNTA resin (20% ethanol) was added to the clarified lysate (2 mL/1 L culture), and incubated at 4 ºC for 30 minutes. The clarified lysate was transferred to a gravity flow column. The flow through was discarded, and the column was washed with 5-10 column volumes of binding buffer. The wash was discarded, and the enzyme was eluted with 2-3 column volumes of elution buffer (50 mM Tris [pH 8.0], 150 mM NaCl, 250 mM imidazole, 25 % (v/v) glycerol). Enzymes were frozen using liquid nitrogen and stored a -80 ºC until needed.

**HPLC Method**

Samples were analyzed by reverse phase chromatography on a C8 or C18 column (4.6 x 100 mm) using a Thermo Ultimate 3000 HPLC. The column compartment temperature was set to 40 ºC, and the flow rate was 1 mL/min. The compounds were separated using a gradient elution with water + 0.1% TFA (solvent A) and acetonitrile + 0.1 % TFA (solvent B) as the mobile phase.
Solvent B was held at 20% for the first minute. Then solvent B was increased to 95% B over 4 minutes, and held at 95% B for 3 minutes. The column was then re-equilibrated to 20% B for three minutes, for a total run time of 11 minutes. Standards were used to identify the retention time, and to produce an external standard curve for quantification.

**Aromatic polyketide time course +/- Additives**

The time course experiments were set up under the following conditions. First, the substrates and co-factors were combined so that the final concentrations were: 15 mM malonate, 5 mM butyrate (to make DA) or hexanoate (to make OA), 1 mM CoA, 2 mM ATP, 20 mM Creatine Phosphate, 5 mM MgCl$_2$, 5 mM KCl and 50 mM Tris pH 8.0 and water added so that the final volume after the addition of enzymes would be 700 µL. The final enzyme concentrations were 0.075 mg/mL MatB, 0.6 mg/mL AAE3, 0.13 mg/mL OLS to make DA and 0.065 mg/mL OLS to make OA, 0.66 mg/mL OAC, 0.16 mg/mL Ppase, 2 units of creatine phosphokinase, and 0.07 mg/mL of ADK. The reactions were initiated by mixing the cofactors with the enzymes. 50 µL aliquots were quenched in 150 µL of methanol at 5, 10, 20, 40, 60, 90, 120, 150, 180, 210, 240 and 300 minutes. The precipitated enzymes were removed by centrifugation (16,000 x g for 2 minutes), and the supernatant was analyzed by HPLC.

The time course reactions in the presence of additives were carried out under the same conditions specified above, but the reactions also contained 20 mg/mL BSA and 5 mM acetyl-phosphate.

**OA/DA inhibition of AP path**

To determine the effect of product inhibition on the biosynthetic pathway, OA was added to reactions biosynthesizing DA, and DA was added to reactions biosynthesizing OA. Because OA and DA are well separated on the HPLC we were able to observe the effect of one molecule on the biosynthesis of the other. The reaction conditions were identical to the time course conditions, however the final reaction volume was 200 µL. The cofactors, buffer and water were
mixed, and either OA or DA was added to the reaction (0 to 5 mM). The reactions were initiated by adding the enzyme mix. Because the OA and DA stocks were in 100% ethanol, the 0 mM control contained 1% ethanol to account for the solvent. All reactions were quenched at 90 minutes and analyzed using the HPLC. The amount of OA or DA produced was normalized to the control, and plotted against the OA or DA concentration.

**MatB Activity Assay**

A coupled enzymatic assay was used to determine the activity of MatB in the presence of OA and DA. The assay schematic is shown in Figure XX. The reaction conditions were: 2.5 mM malonate, 2 mM ATP, 1 mM CoA, 2.5 mM phosphoenolpyruvate (PEP), 1 mM NADH, 5 mM MgCl$_2$, 10 mM KCl, 0.35 mg/mL ADK, 0.75 µg/mL MatB, 1.6 units of PK and 2.5 units of LDH, and 50 mM Tris [pH 8.0]. Background ATPase activity was controlled for by leaving out the substrate (malonate), and either 1% ethanol, 250 µM or 5 mM OA or 5 mM DA was added to the remaining reactions. The activity of MatB was determined by monitoring decreasing absorbance at 340 nm due to NADH consumption using an M2 SpectraMax. To ensure that MatB was limiting at 5 mM OA or DA, MatB was doubled to 1.5 µg/mL. The rate of the reaction doubled indicating that MatB was the limiting component in the system. The rate of NADH consumption at 5 mM OA and 5 mM DA was normalized to the 1% ethanol control.

**AAE3 Activity Assay**

A coupled enzymatic assay, similar to the one above was used to determine the activity of AAE3 in the presence of OA and DA. The assay schematic is shown in Figure XX. The conditions were the same as the MatB assay with the following modifications: 2.5 mM hexanoate was added in lieu of malonate, and 15 µg/mL of AAE3 was added in lieu of MatB. To ensure that AAE3 was limiting, AAE3 was doubled in the presence of 5 mM OA or DA. The rate of the reaction doubled indicating AAE3 is limiting.

**ADK Activity Assay**
A coupled enzymatic assay was used to determine the activity of ADK in the presence of OA and DA. The conditions were similar to the MatB assay, with the following modifications: 2 mM AMP was added in lieu of malonate, CoA was not added, and 0.001 mg/mL of ADK was added. To ensure that ADK was the limiting reagent at 5 mM OA and DA, the amount of ADK was doubled. The 2-fold increase in rate suggested that ADK was the limiting factor.

**CPK Activity Assay**

A coupled enzymatic assay was used to determine the activity of CPK in the presence of OA or DA. The reaction schematic is shown in Figure XX. The reaction conditions were: 5 mM Creatine Phosphate, 2 mM ADP, 5 mM glucose, 2 mM NADP+, 5 mM MgCl₂, 5 mM KCl, 0.3 mg/mL Zwf, 0.1 mg/mL Sc Hex and 0.08 units CPK. The positive control reaction contained 1% ethanol, and either 5 mM of OA or DA was added to the remaining reactions. The absorbance of NADPH at 340 nm was monitored. To ensure that CPK was limiting, the CPK addition was doubled at 5 mM OA and 5 mM DA. The resulting rate doubled, which indicates CPK is limiting even at high OA and DA.

**OLS Activity Assay**

OLS was assayed by setting up the following conditions: 200 µM malonyl CoA, 100 µM hexanoyl-CoA, 0.65 mg/mL OAS, in either 50 mM citrate buffer pH 5.5 or 50 mM Tris buffer pH 8.0. The reactions were initiated by the addition of OAS, and then they were quenched at 30 minutes by adding 150 µL of methanol to the 50 µL reaction. The samples were centrifuged at ~16,000 x g for 2 minutes to pellet the proteins. The supernatant was analyzed using the HPLC. Due to lack of authentic standards for PDAL and HTAL, only olivetol could be quantified.

For the inhibition experiments the conditions were altered to: 1 mM malonyl-CoA, 400 µM hexanoyl-CoA in 50 mM citrate buffer, pH 5.5 in a final volume of 200 µL. Either 1% ethanol, 250 µM OA or 1 mM DA was added to the reaction, and then the reactions were initiated by adding 0.65 mg/mL OLS. 50 µL aliquots were quenched at 2, 4, 6 and 8 minutes in 150 µL of methanol. The reactions were vortexed briefly and centrifuged at 16,000 x g for 2 minutes to pellet the
proteins. The supernatant was analyzed by HPLC. The raw peak areas of HTAL, PDAL and olivetol were summed and plotted against time to determine the rate. The rate of the OA supplemented reaction and the DA supplemented reaction were normalized to the ethanol control.

**OLS/OAC Activity Assay**

To produce OA, the same OLS conditions specified above were used, but olivetolic acid cyclase (OAC) was added to the reaction at 0.6 mg/mL. The reactions were quenched and analyzed in the same manner as the OLS assay. Due to the inconsistencies in this assay compared to OA path data, we investigated potential activators. Acetyl-phosphate and BSA were added to the assays individually at 5mM – 40 mM AcP and 10 – 30 mg/mL BSA final concentration.

**Full pathway set up**

The enzymes used in this study and the final concentrations can be found in Table XX for the MatB path and Table XX for the MdcA path. For the MatB path, the cofactors were added at the following concentrations: 150 mM glucose, 1 mM fructose bisphosphate, 2 mM ATP, 0.25 mM NAD+, 3 mM NADP+, 2 mM CoA, 0.25 mM 2,3-bisphosphoglycerate, 6 mM MgCl2, 10 mM KCl, 0.5 mM thiamine pyrophosphate, 50 mM phosphate pH 8.0, 5 mM hexanoate, 15 mM malonate, 5 mM creatine phosphate, and 50 mM Tris, pH 8.0. The reactions were initiated by the addition of the enzyme master mix detailed in Table XX. The reaction was incubated overnight at room temperature, and the reaction was quenched and extracted 3 times with 200 µL of ethyl acetate. The ethyl acetate was evaporated using a vacuum centrifuge. The sample was dissolved in 200 µL of methanol and analyzed using HPLC.

The enzymes for the MdcA path can be found in Table XX. The MdcA reaction was set up under the same cofactor conditions specified above with the following changes: 3 mM ATP, 0.25 mM AMP, 25 mM creatine phosphate and no Tris buffer.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Inputs</th>
<th>Organism</th>
<th>Module(s)</th>
<th>Molecule</th>
<th>Titer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagne et al 8</td>
<td>Galactose and hexanoate</td>
<td><em>S. cerevisiae</em></td>
<td>AP</td>
<td>OA</td>
<td>0.5 mg/L</td>
</tr>
<tr>
<td>Tan et al 9</td>
<td>Glycerol and hexanoate</td>
<td><em>E. coli</em></td>
<td>AP</td>
<td>OA</td>
<td>80 mg/L</td>
</tr>
<tr>
<td>Luo et al 10</td>
<td>Galactose and hexanoate</td>
<td><em>S. cerevisiae</em></td>
<td>IP, AP, CB</td>
<td>CBGA, THCA, CBDA</td>
<td>7.2 mg/L, 8 mg/L, 4.2 µg/L</td>
</tr>
<tr>
<td>Zirpel et al 6</td>
<td>OA and GPP</td>
<td><em>P. pastoris</em></td>
<td>CB</td>
<td>THCA</td>
<td>0.5 µg/L</td>
</tr>
<tr>
<td>Zirpel et al 7</td>
<td>CBGA</td>
<td><em>P. pastoris</em></td>
<td>CB</td>
<td>THCA</td>
<td>3 g/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CBDA</td>
<td>400 mg/L</td>
</tr>
<tr>
<td>Valliere et al 18</td>
<td>Glucose and OA</td>
<td><em>SynBio</em></td>
<td>IP and CB</td>
<td>CBGA, CBGVA, CBDA</td>
<td>1.25 g/L, 1.7 g/L, 25 mg/L</td>
</tr>
</tbody>
</table>

**Table 3-1: Summary of published cannabinoid production systems.** The inputs column specifies the molecules added to produce the cannabinoid. The module(s) column specifies which part of cannabinoid synthesis the authors were engineering. Abbreviations: AP is aromatic polyketide, IP is isoprenoid, CB is cannabinoid, OA is olivetolic acid, CBGA is cannabigerolic acid, THCA is tetrahydrocannabinolic acid, CBDA is cannabidiolic acid and CBGVA is cannabigerovarinic acid.
<table>
<thead>
<tr>
<th>Enzyme Abbreviation</th>
<th>Full Name</th>
<th>Organism</th>
<th>Acquisition Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hex</td>
<td>Hexokinase</td>
<td><em>S. cerevisiae</em></td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>Pgi</td>
<td>Glucose-6-phosphate isomerase</td>
<td><em>G. thermodenitrificans</em></td>
<td>ABO6822</td>
</tr>
<tr>
<td>PfkA</td>
<td>Phosphofructokinase</td>
<td><em>G. stearothermophilus</em></td>
<td>KOR92562</td>
</tr>
<tr>
<td>Fba</td>
<td>Fructose-1,6-bisphosphate aldolase</td>
<td><em>S. aureus</em></td>
<td>BAR10119</td>
</tr>
<tr>
<td>TpiA</td>
<td>Triose phosphate isomerase</td>
<td><em>G. stearothermophilus</em></td>
<td>KOR95273</td>
</tr>
<tr>
<td>Gap</td>
<td>Gald-3-P dehydrogenase</td>
<td><em>E. coli K12</em></td>
<td>NP_416293</td>
</tr>
<tr>
<td>mGap</td>
<td>Gald-3-P dehydrogenase D34A/L35R/T36K</td>
<td><em>G. stearothermophilus</em></td>
<td>NP_416293</td>
</tr>
<tr>
<td>NoxE</td>
<td>NADH Oxidase</td>
<td><em>L. lactis</em></td>
<td>WP_015425842</td>
</tr>
<tr>
<td>Pgk</td>
<td>Phosphoglycerate Kinase</td>
<td><em>G. stearothermophilus</em></td>
<td>WP_415276</td>
</tr>
<tr>
<td>dPgm</td>
<td>Phosphoglycerate Mutase (2,3 BPG dependent)</td>
<td><em>E. coli K12</em></td>
<td>NP_417259</td>
</tr>
<tr>
<td>Eno</td>
<td>Enolase</td>
<td><em>E. coli K12</em></td>
<td>KOR95272</td>
</tr>
<tr>
<td>PykF</td>
<td>Pyruvate Kinase (FBP dependent)</td>
<td><em>E. coli K12</em></td>
<td>NP_416191</td>
</tr>
<tr>
<td>PyOx</td>
<td>Pyruvate Oxidase</td>
<td><em>A. viridans</em></td>
<td>AG Scientific</td>
</tr>
<tr>
<td>PTA</td>
<td>Acetyl-phosphate transferase</td>
<td><em>G. stearothermophilus</em></td>
<td>WP_05353264</td>
</tr>
<tr>
<td>PhaA</td>
<td>Acetyl-CoA acetyltransferase</td>
<td><em>R. eutropha</em></td>
<td>CAJ92573</td>
</tr>
<tr>
<td>HMGS A110G</td>
<td>HMG-CoA Synthase A110G</td>
<td><em>E. faecalis</em></td>
<td>WP_010785222</td>
</tr>
<tr>
<td>HMGR</td>
<td>HMG-CoA Reductase</td>
<td><em>E. faecalis</em></td>
<td>AAG02439</td>
</tr>
<tr>
<td>MVK</td>
<td>Mevalonate Kinase</td>
<td><em>M. mazei</em></td>
<td>AAM31458</td>
</tr>
<tr>
<td>PMVK</td>
<td>Phosphomevalonate Kinase</td>
<td><em>S. pneumonia</em></td>
<td>WP_000562411</td>
</tr>
<tr>
<td>MDC</td>
<td>Diphosphomevalonate Kinase</td>
<td><em>S. pneumonia</em></td>
<td>NP_357933</td>
</tr>
<tr>
<td>IDI</td>
<td>Isopentyl-PP Isomerase</td>
<td><em>E. coli K12</em></td>
<td>NP_417365</td>
</tr>
<tr>
<td>FPPS S82F</td>
<td>Farnesyl-PP synthase S82F</td>
<td><em>G. stearothermophilus</em></td>
<td>KOR95521</td>
</tr>
<tr>
<td>NphB M31</td>
<td>Aromatic prenyltransferase M31</td>
<td><em>Streptomyces sp. CL190</em></td>
<td>BAE00106.1</td>
</tr>
<tr>
<td>Ppase</td>
<td>Pyrophosphatase</td>
<td><em>G. stearothermophilus</em></td>
<td>O05724</td>
</tr>
<tr>
<td>Cat</td>
<td>Catalase</td>
<td><em>C. glutamicum</em></td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>AAE3</td>
<td>Acyl Activating Enzyme 3</td>
<td><em>C. sativa</em></td>
<td>AFD33347.1</td>
</tr>
<tr>
<td>MatB</td>
<td>Malonyl-CoA Synthetase</td>
<td><em>R. plaustris</em></td>
<td>CAE25665.1</td>
</tr>
<tr>
<td>OLS</td>
<td>Olivetol Synthase</td>
<td><em>C. sativa</em></td>
<td>BAG14339.1</td>
</tr>
<tr>
<td>OAC</td>
<td>Olivetolic Acid Cyclase</td>
<td><em>C. sativa</em></td>
<td>AFN42527.1</td>
</tr>
<tr>
<td>ADK</td>
<td>Adenylate Kinase</td>
<td><em>G. thermodenitrificans</em></td>
<td>ATO37397.1</td>
</tr>
<tr>
<td>CPK</td>
<td>Creatine Kinase</td>
<td>Rabbit Muscle</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>MdcA</td>
<td>Alpha Subunit Malonate Decarboxylase</td>
<td><em>G. stearothermophilus</em></td>
<td>OQO99201.1</td>
</tr>
</tbody>
</table>

**Table 3-2: Summary of all enzymes used in this study.**
## MatB Pathway

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>mg/mL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Glycolysis</strong></td>
<td></td>
</tr>
<tr>
<td>Sc Hex</td>
<td>0.02</td>
</tr>
<tr>
<td>Gs PfkA</td>
<td>0.32</td>
</tr>
<tr>
<td>Sa Fba</td>
<td>0.18</td>
</tr>
<tr>
<td>Gs TpiA</td>
<td>0.04</td>
</tr>
<tr>
<td>Gs Pgi</td>
<td>0.17</td>
</tr>
<tr>
<td>Ec GapA</td>
<td>0.05</td>
</tr>
<tr>
<td>Gs GapM6</td>
<td>0.18</td>
</tr>
<tr>
<td>Gs Pgk</td>
<td>0.03</td>
</tr>
<tr>
<td>Ec dPgm</td>
<td>0.38</td>
</tr>
<tr>
<td>Ec Eno</td>
<td>0.08</td>
</tr>
<tr>
<td>Ec PykF</td>
<td>0.56</td>
</tr>
<tr>
<td>Av PyOx</td>
<td>1 unit</td>
</tr>
<tr>
<td>Gs PTA</td>
<td>0.06</td>
</tr>
<tr>
<td>Li NoxE</td>
<td>0.25</td>
</tr>
<tr>
<td>Ca Catalase</td>
<td>125 units</td>
</tr>
<tr>
<td><strong>Mevalonate</strong></td>
<td></td>
</tr>
<tr>
<td>Re PhaA</td>
<td>0.12</td>
</tr>
<tr>
<td>Ef HMGS A110G</td>
<td>0.22</td>
</tr>
<tr>
<td>Ef HMGR</td>
<td>0.58</td>
</tr>
<tr>
<td>Mm MVK</td>
<td>0.16</td>
</tr>
<tr>
<td>Spne PMVK</td>
<td>0.23</td>
</tr>
<tr>
<td>Spne MDC</td>
<td>0.22</td>
</tr>
<tr>
<td>Ec IDI</td>
<td>0.23</td>
</tr>
<tr>
<td>Gs FPPS S82F</td>
<td>0.04</td>
</tr>
<tr>
<td>SCL190 NphB</td>
<td>0.45</td>
</tr>
<tr>
<td>Gs Ppase</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>Olivetolate</strong></td>
<td></td>
</tr>
<tr>
<td>Rp MatB</td>
<td>0.03</td>
</tr>
<tr>
<td>Cs AAE3</td>
<td>0.18</td>
</tr>
<tr>
<td>Cs OLS</td>
<td>0.25</td>
</tr>
<tr>
<td>Cs OAC</td>
<td>0.87</td>
</tr>
<tr>
<td>Gt ADK</td>
<td>0.07</td>
</tr>
<tr>
<td>Creatine Kinase</td>
<td>2 units</td>
</tr>
</tbody>
</table>

Table 3-3: Enzymes used in the full cannabinoid MatB pathway, with final enzyme concentrations
<table>
<thead>
<tr>
<th>Transferase Pathway</th>
<th>Enzyme</th>
<th>mg/mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycolysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sc Hex</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Gs PfkA</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Sa Fba</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Gs TpiA</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Gs Pgi</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Ec GapDH</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Gs GapM6</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Gs Pgk</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Ttg dPgm</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Ec EnoNH</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Gs PykA</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Av PyOx</td>
<td>1 unit</td>
<td></td>
</tr>
<tr>
<td>Gs PTA</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Ll NoxE</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Ca Catalase</td>
<td>125 units</td>
<td></td>
</tr>
<tr>
<td>Mevalonate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re PhaA</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Ef HMGS-A110G</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Ef HMGR</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>Mm MVK</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Spne PMVK</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Spne MDC</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Ec IDI-CH</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Gs FPPS S82F</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>SCL190 NphB 31</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Gs PPase</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Olivetolate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gs MdcA</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Cs AAE3</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Cs OAS</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Cs OAC</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Gt ADK</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Creatine Kinase</td>
<td>2 units</td>
<td></td>
</tr>
</tbody>
</table>

Table 3-4: Enzymes used in the full cannabinoid MdcA pathway with final enzyme concentrations
Figure 3-1: Cannabinoid Biosynthesis. Cannabinoid biosynthesis is derived from the fatty acid, isoprenoid and aromatic polyketide metabolic pathways. (A) Isoprenoids are generated through the MEP pathway (not shown) or the mevalonate pathway, each molecule of GPP produced requires 6 acetyl-CoA and 6 ATP. (B) Aromatic polyketide biosynthesis is dependent on small chain fatty acids, and acetyl-CoA. This pathway requires 2 ATP and 3 acetyl-CoA molecules to produce the aromatic polyketides (olivetolic acid and divarinic acid). (C) The cannabinoid path starts by transferring the prenyl group from GPP onto OA, and cyclizing the geranyl moiety, to form the final cannabinoid.
Figure 3-2: Aromatic Polyketide Production Module. Malonyl-CoA is generated from malonate using MatB, hexanoyl-CoA or butyryl-CoA is generated from hexanoate or butyrate by acyl activating enzyme 3 (AAE3). OLS adds malonyl-CoA to the growing polyketide chain, which is eventually cyclized by olivetolic acid cyclase to produce either olivetolic acid or divarinic acid. The polyketide intermediates are unstable, and can spontaneously cyclize to form the products highlighted in the blue box. The AMP produced by AAE3 was recycled into ATP using adenylate kinase (ADK) and creatine kinase (CPK).
Figure 3-3: Production of OA and DA over time. The dark blue squares represent OA titers from 5 to 300 minutes, and the green squares represent DA titers from 10 to 300 minutes. This data is the average of 3 replicates, and the error bars are the standard deviation.

Figure 3-4: Inhibition of the AP path by OA and DA. The OA inhibition of the DA path was tested by adding exogenous OA to the system producing DA. The amount of DA produced was normalized to the control, and plotted against the OA concentration (blue squares). The amount of OA produced was normalized to the control and plotted against the DA concentration (green squares). This data is the average of 3 replicates, and the error bars are the standard deviation.
Figure 3-5: OA and DA inhibition of AP pathway enzymes. (A) MatB, AAE3, ADK and CPK enzyme activities were tested at 5 mM OA (blue) and 5 mM DA (green). The percent activity remaining in the presence of 5mM of the inhibitor is shown. (B) MatB, AAE3, ADK and OLS were tested for inhibition at 250 µM OA, and OLS was tested for DA inhibition at 1 mM. The percentage of remaining activity for each enzyme is shown. This data is the average of 3 replicates, and the error bars are the standard deviation.

Figure 3-6: pH dependence of OLS activity. (A) An activity assay for OLS was performed at pH 8.0 (green trace) and pH 5.5 (purple trace), and the resulting chromatogram is shown. Relevant
peaks are highlighted with the dashed grey line. (B) An OLS/OAC coupled assay at pH 5.5 (purple trace), pH 8.0 (green trace) and at pH 8.0 with 5 mM acetyl phosphate, AcP (blue trace). Relevant peaks are highlighted with the dashed grey line. HTAL is hexanoyl triacetic acid lactone, PDAL is pentyldiacetic acid lactone, Ool is olivetol and OA is olivetolic acid.

**Figure 3-7: Effects of additives on OA production.** The product profile of four reactions are shown: no activator (red trace), 5 mM acetyl phosphate, AcP (orange trace), 20 mg/mL BSA (blue trace), and 20 mg/mL BSA with 5 mM AcP (purple trace). The relevant peaks are highlighted with a dashed grey line. HTAL is hexanoyl triacetic acid lactone, PDAL is pentyldiacetic acid lactone and OA is olivetolic acid.
Figure 3-8: The effect of additives on AP production and inhibition. (A) A comparison of the time dependent production of OA in the presence (light blue squares) and absence (dark blue squares) of additives (5 mM AcP and 20 mg/mL BSA). (B) A comparison of the time dependent production of DA in the presence (dark green squares) and absence (light green squares) of additives (5 mM AcP and 20 mg/mL BSA). (C) The relative amount of DA produced in the presence (light blue squares) and absence (dark blue squares) of additives (5 mM AcP and 20 mg/mL BSA) with 0 to 5 mM OA added. (D) The relative amount of OA produced in the presence (dark green squares) and the absence (light green squares) of the additives (5 mM AcP and 20
mg/mL BSA) with 0 to 5 mM DA added. This data is the average of 3 replicates, and the error bars are the standard deviation.

Figure 3-9: Pathway schematics for the MatB and MdcA (transferase) paths. (A) Schematic for the MatB path. The malonyl-CoA production is ATP dependent, but otherwise not connected with the pathway. The purple line represents glycolysis, the blue line represents pyruvate oxidation, the green line represents the mevalonate pathway, the pink line represents the cannabinoid pathway, and the yellow lines represent the aromatic polyketide pathway. The current titer using this pathway is 12 mg/L. (B) Schematic for the MdcA transferase path. The malonyl-CoA production is no longer ATP dependent, and is tied in to the pyruvate oxidation path, and the mevalonate path. The purple line represents glycolysis, the blue line represents pyruvate oxidation,
oxidation, the green line represents the mevalonate pathway, the pink line represents the cannabinoid pathway, and the yellow lines represent the aromatic polyketide pathway. The current titer for the system is 42 mg/L.

Figure 3-10: Reaction schematics for the enzyme activity assays used in this chapter. The reporter enzymes are shown in green, which were present in excess and the limiting enzyme is shown in purple. (A) The activity assay for MatB. (B) The activity assay for AAE3. (C) The activity assay for ADK. (D) The activity assay for CPK.
3.7 References


18. Valliere, M. A. A cell-free biosynthetic platform for the prenylation of natural products with applications to cannabinoid production.