INCREASED PRODUCTION OF ISOPRENOL VIA METABOLIC ENGINEERING OF THE MVA PATHWAY IN *E. COLI*

by

SIERRA SKYE MILLIGAN

(Under the Direction of Yajun Yan)

ABSTRACT

Isoprenol is an excellent alternative for fuel sources and a precursor for many commercial chemicals in industrial applications. Microbial production of isoprenol via the mevalonate pathway has been previously optimized in E. coli, resulting in approximately 2.23 g/L titers using the promiscuous activity of known phosphatases, NudB and AphA, and "IPP-Bypass system. However, limitations achieving optimal production still exist. To overcome these limitations, we have developed a more efficient synthetic pathway for enhanced isoprenol production by expanding upon existing metabolic pathways in *E. coli* by incorporating the upper MVA pathway into the genome and regulating lower pathway components using the "IPP-Bypass" and plasmid-based expression systems. In this work we have uncovered a new phosphatase, yigL, capable of isoprenol production with increased titers of 2.92 g/L. This novel finding combined with future work and applications presented herein have the potential to achieve the new industrial standard in isoprenol production.

INDEX WORDS: Isoprenol, mevalonate pathway, metabolic engineering, pathway engineering, *E. coli*

INCREASED PRODUCTION OF ISOPRENOL VIA METABOLIC ENGINEERING OF THE MVA PATHWAY IN *E. COLI*

by

SIERRA MILLIGAN

BS, University of Georgia, 2019

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2021

© 2021

SIERRA SKYE MILLIGAN

All Rights Reserved

INCREASED PRODUCTION OF ISOPRENOL VIA METABOLIC ENGINEERING OF THE MVA PATHWAY IN *E. COLI*

by

SIERRA SKYE MILLIGAN

Major Professor:	Yajun Yan
Committee:	William Kissalita
	Cheryl Gomillion

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2021

DEDICATION

I dedicate this thesis to my family for nursing me with affections and love and their dedication for success in my life. My Daddy for being my voice of sound reasoning and guidance, Mama for being my rock and solid foundation, “Bubba” for being my biggest cheerleader and bodyguard, “Sissy” for being my mini-me and light in the dark, and Grandma for instilling in me a love of science and curiosity.

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who played a role in my academic accomplishments. First and foremost, my parents and siblings, who supported me with love and understanding. Without you, I could have never reached this level of success. My extended family for their continued support throughout this journey.

Secondly, my major professor, Dr. Yajun Yan, for guiding me throughout my academic career and allowing me the opportunity to pursue my dream education. My committee members, each of whom have provided patient advice and guidance throughout my research process. My mentors, Jackie and Ruihua, for teaching me throughout this journey and guiding me in my lab studies.

Thank you all for your unwavering support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Metabolic and Synthetic Biology

As the population continues to grow at exponential rates, and arable land and natural resources are depleted at even greater rates, the imminent need for more sustainable production of high value compounds for chemical, environmental, pharmaceutical, and biological applications is more pressing than ever. According to the World Health Organization, 91% of the global population live in polluted air regions that exceed the WHO's air quality standards. This is predominantly due to the use of nonrenewable resources, such as fossil fuels, in transportation and chemical industries. Continued dependence on such resources will lead to a disruption of environmental equilibrium and places an immense burden on energy security for future generations. Sustainable development in these areas to meet the needs of the current demand without compromising future generations is crucial for population livability. Increased concerns over both sustainability and energy dependence have motivated the development of alternative production methods, including microbial fermentation. However, the demand to meet these growing needs has encountered similarly increasing challenges, mainly owing to the efficiency of desired product synthesis. With such challenges, the applications of metabolic engineering and synthetic biology have come to depend upon each other for the achievement and feasibility of the overreaching goal of increased sustainability in microbial production of valued products. Such strategies have attracted significant attention from both the scientific and engineering community

and from political and environmental advocacy sectors for their potential to mitigate climate change on a global scale, improve energy security for the respective nations that adopt these production methods, decrease production costs, enhance production performance, and reduce the world's dependence on fossil fuels.

Synthetic biology is a multidisciplinary field that combines the disciplines of biology and engineering for the creation of novel biological systems not found in nature (*1, 2*). This includes the design of genetic component libraries, the assembly of genetic circuits within a host organism, and behavioral model prediction of engineered biological systems (*1-4*). Metabolic engineering seeks to optimize the genetic and regulatory processes within cells to improve cellular production and to produce the desired greater product yields (*2, 4, 5*). The coordination of these two fields requires that synthetic biology provide the components and information concerning the biological phenomena while metabolic engineering applies this information towards the creation of an optimized biological synthesis trajectory to create a desired compound (*2, 3*). The combination of both areas allows for the advancement of rational changes in DNA sequence; such changes include specific mutations and gene knockout, the assembly of various parts and components in genetic circuits or biosynthetic pathways, and the integration of DNA fragments within a plasmid or host genome for biological replication (*2-5*). The coordinated and optimized development of both fields in conjunction with one another has led to the production of numerous chemicals and natural products, with valued applicability through systematic engineering of microbial hosts as compared to previous production methods using nonrenewable resources and costly extraction processes of plant materials from non-sustainable sources. With continued coordinated advancement of these two fields, bio-based production and its

corresponding economy could replace the fossil-based chemical industry as a sustainable and feasible alternative.

1.2 Microbial Host and Target Production

A variety of sources share the ability to produce natural products, also known as secondary metabolites, including fungi, plants, and bacteria. Historically, the majority of all medicines originated from plant materials accounting for approximately 80% of all medicines in the early 1900's (*6*). However, natural products often exist in low concentrations in nature. Thus began the search for alternative routes of production. Upon the discovery of penicillin from the species *Penicillium chrysogenum* in 1928 by Alexander Flemming, the production of secondary metabolites underwent a major shift from plants to microorganisms (*6, 7*). Since then, microbial host have served as the central factory for heterologous expression and production of desired targets due to their relative ease of manipulation and efficacy to produce a wide range of products. The combination of metabolic engineering and synthetic biology within such systems serves as a powerful framework to produce chemicals of immediate and practical application whereby we can regulate metabolic networks of microbial host and control the expression of various pathways and genes; this allows for systematic control of desired product formation in a variety of organisms (*8*). Various applications of genetic modification strategies have been commonly employed to enhance microbial production. This includes overexpression of rate-limiting enzymes, eliminating or reducing the burden of competitive pathways, and fermentation optimization. However, with such applications, the efficiency of microbial production can be weakened due to metabolic imbalance. To counteract this, fine-tuning of systematic control may be used such as adjusting promoter strength and gene copy number and optimizing RBS sites.

The overall ability of microbial production is dependent upon the choice of host as the overall production performance, even within identical pathways, may have significant differences. Therein lies the challenge to choosing a suitable model host – a full knowledge of both the microbial background and target biosynthetic pathway is needed to optimize the process.

Over the past century, *Escherichia coli*, a gram negative anaerobic bacterium, has become one of the most ideal platform hosts used for the development and production of industry value target products due to its ease of culturing, its metabolic flexibility, and its rapid growth; the industry possesses a great breadth and depth of physiological and biochemical knowledge of this bacterium which gives us a plethora of tools for genetic and genomic engineering (*9*). Furthermore, commonly used *E. coli* strains such as K-12 and descendants are generally considered harmless as they lack many of the factors associated with disease. These non-pathogenic strains have been widely used in laboratory and industrial applications for the production of pharmaceuticals, food additives, chemicals, and fuels (*9*). *E. coli* offers many benefits compared to that of other hosts including its ability to utilize a wide range of substrates such as glucose, xylose, galactose, lignocellulose, and fatty acids. Many molecular cloning techniques and genetic tools have been developed, demonstrated, and applied in *E. coli* systems such as promoter and ribosome binding site (RBS) engineering, CRISPR-(d)Cas9, and small regulatory RNAs (sRNAs) (*9-12*). *E. coli* has been utilized to produce a multitude of products including nucleotides, recombinant proteins, high-value amino acids, and chemicals (*9*, *12*). While *E. coli* may not be the most versatile of organisms for industrial purposes, extensive progress in engineering new *E. coli* genotypes and phenotypes has led to novel strains that surpass production efforts of traditional native producers and serves as a proof-of-concept model

organism. Continual improvement in the design of non-native biosynthetic pathways is required to produce important classes of molecules.

1.3 Isoprenoids and Isoprenol

Isoprenoids are a class of organic compounds composed of two or more units of hydrocarbons consisting of five carbon atoms arranged in a specific pattern. The five carbon unit that constitutes the backbone of isoprenoids is an unsaturated hydrocarbon known as isoprene (*13*). Isoprenoids contain two or more isoprene units and can have one or more functional units, such as hydroxyl, attached to it contributing to the diversity of compounds. Isoprenoids range in function from pigments and fragrances to vitamins and precursors for high-value chemicals. While the uniquity of the five-carbon isoprene unit had been recognized for some time, the identity of the compounds actually involved in the physiological assembly of the isoprenoids was not known until the 1950's when Konrad E. Bloch and Feodor Lynen discovered that the synthesis of isoprenoids in nature begins with acetyl coenzyme A (*13*). Isoprenoids are classified according to the number of isoprene (C5H8) units they contain. Hemiterpenes consist of a single isoprene unit and oxygen containing derivatives such as prenol are hemiterpenoids.

Isoprenol, also known as 3-methyl-3-butene-1-ol, is a hemiterpene alcohol and a major biosynthetic building block for an assortment of aromatic compounds and high value products. It is a valuable fuel alternative and precursor for commercial chemicals in industry applications including rubber, dyes, solvents, fragrances, polymers, and pesticides. The global isoprenol market was valued at $220.3 million in 2020 and expected to reach $258.8 million by the end of 2026, making it a high value chemical in great demand (77). The major advantage to isoprenol exist in its favorable chemical and physical properties. Historically, ethanol has served as the

main biofuel along with C5 alcohols. However, isoprenol offers a more efficient alternative as the branched-chain unsaturated insopentenol possess higher octane numbers and enhanced low-temperature fluidity as compared to the straight-chain counterpart, making it a better potential gasoline replacement (*14*). Several microbial hosts have been engineered for the biological production of isoprenol, with the most commonly targeted pathways including isoprenoid pathways from both the mevalonate (MVA) pathway and the methylerythritol phosphate (MEP) pathway (*15, 16*).

1.4 Mevalonate Pathway

The mevalonate (MVA) pathway, also known as the isoprenoid pathway, is an essential metabolic pathway in bacteria and one of two major pathways for isoprenoid production. The conventional MVA pathway for isoprenoid production as illustrated in **Figure 1.1**, can be characterized by the upper and lower portions of enzymatic steps.

The upper MVA pathway consist of the first three enzymatic steps allowing for the conversion of acetyl-CoA to MVA. From here, the lower MVA pathway begins whereby MVA is phosphorylated into mevalonate 5-phosphate (Mevalonate-P / MVAP) by mevalonate kinase (MK). This is then subsequently phosphorylated by 5-phosphomevalonare kinase (PMK) into mevalonate 5-diphosphate (MVAPP); and MVAPP is converted to isopentenyl diphosphate (IPP) utilizing diphosphomevalonate decarboxylase (PMD). Such phosphorylation and conversion reactions are completed at the expense of three adenosine triphosphate (ATP) molecules. Finally, desired isoprenoids are produced through the hydrolysis of the IPP pyrophosphate group through various phosphatases, first creating the intermediate isopentenyl monophosphate (IP).

The synthetic pathway to produce five-carbon alcohols, isopreniods, from isopentenyl diphosphate was optimized by Chou's group in 2012 through the construction of targeted libraries from evolutionarily related proteins able to catalyze a specific chemistry on various substrates, also known as enzyme families, by screening for those that facilitated the conversion of isoprenoids from IPP, allowing for the removal of the intermediate IP in the pathways and decreasing the number of enzymatic steps (**Figure 1.1**). Two different superfamilies were found to be capable of such reactions, HAD phosphatases and Nudix hydrolases, and the overexpression of NudB was found to have the highest impact on isoprenol conversion resulting in approximately 55 mg/L production titers. Numerous studies have worked to increase isoprenol production via pathway optimization in *E. coli* where 70% of apparent theoretical yield has been reached resulting in the highest titers of 2.23 g/L thus far (*15, 17, 18*). However, there still exist limitations in the MVA pathway which will be discussed in greater depth to come.

1.5 Goals of This Work

The primary goal of this work is to develop a more efficient synthetic pathway for enhanced isoprenol production by expanding upon existing metabolic pathways in *E. coli* as discussed here. Chapter 2 illustrates the continued systematic regulation of the mevalonate pathway through optimization and overexpression of rate limiting enzymes. Its primary focus is to determine if production can be increased through the integration of the upper MVA pathway into the *E. coli* genome, previously completed by our lab, in conjunction with plasmid-based overexpression expression of lower pathway enzymatic steps. Chapter 3 is a continuation based on the results found in Chapter 2 that describes the application of alternative enzymes, specifically the final phosphatase responsible for the conversion of IP to isoprenol, and

comparison of final production titers for optimal conversion. Chapter 4 discusses the future applications to be applied to this work including the immediate extension of the work based on results found for further strain development and the inhibition of competitive pathways using various metabolic engineering techniques.

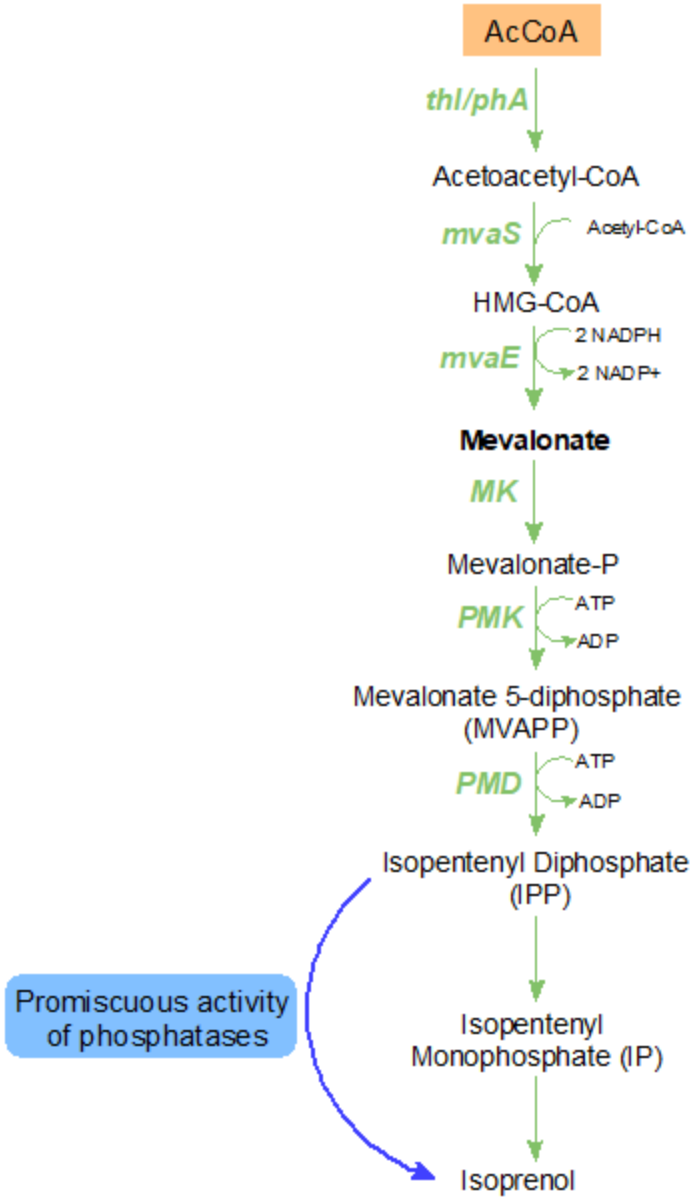

Figure 1.1. Two pathways for isoprenol production via mevalonate pathway in *E. coli*. The original pathway (Pathway I) consists of 8 enzymatic steps with the accumulation of intermediates IPP and IP. The second pathway (Pathways II) uses promiscuous phosphatase activity and converts IPP directly into isoprenol.

CHAPTER 2

MEVALONATE PATHWAY CHROMOSOMAL INTEGRATION, REGULATION, AND OPTIMIZATION FOR ENHANCED BIOSYNTHESIS OF ISOPRENOL

2.1 Introduction and Literature Review

The improvement of isoprenoid production relies on the efficiency of the MVA pathway and the enzymatic steps within. Much work has been done with the aim of increasing metabolic flux throughout the system to increase desired isoprenoid titers. Traditional approaches for controlled expression over the MVA pathway in E.coli rely on plasmid-based expression systems (*12*). The advantage to such systems is their ease of manipulation and gene expression control using a variety of copy numbers and promoter systems. However, despite the ability of these systems to achieve high protein expression levels, they ultimately result in genetic instability and low mean productivity. This is due to the increased metabolic burden and unequal distribution of plasmids to daughter cells leading to segregation instability (*19*). Furthermore, antibiotics must be employed for the maintenance of bacterial plasmids, increasing the overall cost of production and potential for antibiotic resistance. For this reason, chromosomal integration, expression, and optimization of heterologous pathways directly into the genome has become of great interest in regard to generating stable and reproducible strains preferred for production. Although chromosomal integration cannot decrease the burden of transcriptional and translational tasks, it can reduce the burden of metabolites from plasmid-replication equipment (*20*).

As with any application, there are challenges to overcome including the copy number of the desired gene, chromosomal location of insert, and expression cassette optimization. Numerous advancements have been made in the last few decades with the creation and application of molecular tools capable of integrating foreign DNA into bacterial chromosomes with the promise of alleviating such challenges (*20*). Based on previous studies, our lab has designed an *E coli* strain with the upper portion of the MVA pathway integrated into the genome (BW25113(F')ΔEP2 asl::MVA) that will be employed for the duration of this work. This strain is a derivative of our lab's previously created ΔEP2 strain that contains knockouts of *edd* and *pfkB*, enzymes involved in the EDP and EMPP glycolytic pathways, respectively. Disruption of these was found to increase MVA production by redistributing glycolytic flux (*21*). As glycolysis is the breakdown of glucose into essential building blocks and energy for the cell, it occupies a central role in *E. coli* growth and bioproduction. By redirecting carbon flux through desired pathways, we can increase titers by reducing burden of competitive pathways and ultimately increase efficiency in cell growth and production. BW25113(F')ΔEP2 asl::MVA integrates the upper MVA pathway onto the asl locus of ΔEP2. As chromosomal integration position impacts the performance of the pathway based on the effects of foreign genes, it was therefore taken into consideration upon producing this strain. Based on previous findings and work completed by our lab, the asl locus was found to be the most successful in MVA accumulation. This *E. coli* strain also contains the catalytically inactive (d)-Cas9 protein, which will be discussed in greater depth in the following chapter.

The primary focus of this chapter is to determine if chromosomal integration of the upper MVA pathway, in conjunction with optimized plasmid-based overexpression of lower pathway enzymatic steps, is more efficient in isoprenol production than previous findings. With this, we

also sought to optimize lower pathway rate-limiting steps by reducing the burden of toxic intermediates and energy consuming steps. Reviewing available literature, we found various studies working toward increased production of isoprenoids, and specifically isoprenol. While numerous advancements have been made in the optimization of the conventional MVA pathway in *E. coli* resulting in various levels of isoprenol production, there still exist limitations (*14, 15, 17, 18, 22-25*). The most prevalent of these obstacles to greater production is the accumulation of IPP, which can be toxic for bacterial growth. To overcome such limitations of the conventional MVA pathway, an alternative IPP-bypass MVA pathway was developed by taking advantage of the promiscuous activity of PMD towards the non-native substrate MVAP to produce IP, as illustrated in **Figure 2.1** (*15, 18, 23*) . This prevents the potential accumulation of toxic IPP and decreases the total energy input required for isoprenol production. Kang's group successfully screened and utilized *in vitro* enzyme kinetics of PMD mutant libraries and identified those that significantly increased isoprenol production up to 1130 $\pm$ 5 mg/L. The mutant R74H-R147K-M212Q was found to have the highest positive influence on production titers. Based on these findings, we employed the optimized PMD mutant within our work (PMD*). Furthermore, the characterization of an assortment of phosphatases has emerged in the aims of finding those which possess the ability to efficiently convert IP to isoprenol. Both NudB and AphA have been previously observed with an ability to produce isoprenol from both IPP and IP. The purpose of this chapter is to illustrate the findings of chromosomal integration for the upper MVA pathway in conjunction with lower pathway plasmid expression systems utilizing the "IPP-Bypass" and most relevant phosphatases as illustrated in **Figure 2.1**. Because both NudB and AphA have been previously used in production efforts, they will serve as a control compared to previous findings for the findings of this chapter as the combination of chromosomal integration and

plasmid-based expression of the MVA pathway utilizing the "IPP-Bypass" has not been tested before, to the best of our knowledge.

As previously stated, plasmids have long been used in biotechnology applications due to their proclivity to be easily manipulated and are readily transferable to host cells. Plasmids replicate autonomously from the bacterial chromosome in a predictable and controlled manner and are usually present in more than one copy per cell, which leads to higher recombinant gene dosages (*26*). To co-exist with the host in a stable fashion and to minimize metabolic burden, plasmids must control their replication so that the copy number does not overload the cellular machinery and demand an expenditure of excess energy. Thus, plasmid copy number plays a central role in the expression of genes. For the purpose of this study, we also tested which copy number or combination of copy numbers for the expression of the lower MVA pathway performed the best in final conversion of isoprenol. We focused specifically on the expression and replication of MK, because of the nature of the pathway. We hypothesized that MK's functionality may not rely on as high of expression levels compared to other genes. Results from these experiments will lead us to further optimization efforts based on enzymatic limitations.

2.2 Materials and Methods

2.2.1 Bacterial Strains, Plasmids, and Experimental Materials

E. coli cells were grown for plasmid construction, propagation, and inoculum preparation using Luria-Bertani (LB) medium. Shake flask fermentation for isoprenol production and analysis was conducted using the biosynthesis medium M9Y. M9Y contains (per liter): M9 salts (11.28 g), yeast extract (5 g), glucose (20 g), 1M MgSO4 (1 mL), 0.1M CaCl2 (1 mL), and VB1 (100uL). Kanamycin and ampicillin (50 mg/L) antibiotics were added to cultures when

necessary. *E. coli* strain XL1-Blue was used for plasmid propagation and gene cloning, and the high and medium copy number plasmids, pZE12 and pCS27 were employed for gene cloning, expression, and assembly in this work. *E. coli* BW25113(F')ΔEP2 asl::MVA was used for production. **Table 2.1** lists the strains and plasmids used in this study.

Phusion High-Fidelity DNA polymerase, restriction endonucleases, and Quick Ligation Kit were purchased from New England Biolabs (Ipswich, MA, USA). Zyppy™ Plasmid Miniprep Kit and Zymoclean™ Gel DNA Recovery Kit were purchased from Zymo Research (Irvine, CA, USA). Standard chemicals were purchased from Sigma-Aldrich unless otherwise specified.

2.2.2 Construction of Plasmids

All genes were amplified using polymerase chain reaction (PCR) using Phusion High-Fidelity DNA polymerase, and plasmid construction and DNA manipulations were performed following the standard molecular cloning techniques and protocols. pZE12-PMD* was obtained by inserting the optimized *PMD** from *E.* coli into the high copy number plasmid pZE-12-luc between *XhoI* and *SphI* under the control of the pLacO1 promoter and Ti terminator. pCS-MK was constructed by cloning MK from *S. cerevisiae* into medium copy number plasmid pCS27. pZE12-PMD*-NudB was designed by inserting NudB from *E. coli* into previously created pZE12-PMD* plasmid, between *SphI* and *SalI*. Similarly, pZE12-PMD*-Apha was constructed by inserting *Apha* from *E. coli* into pZE12-PMD* using the same method. pZE12-PMD*-ScMK was obtained by cloning MK into pZE12-PMD* between *SpeI* and *SalI* under the control of the pLacO1 promoter and T1 terminator system. pZE12-PMD*-NudB-ScMK was further obtained by inserting *NudB* into pZE12-PMD*-ScMK between *SphI* and *AvrII*. Likewise, pZE12-PMD*-

AphA-ScMK was obtained by inserting *AphA* into pZE12-PMD*-ScMK between using the same method.

Table 2.1 Strains and plasmids used in this study.

Strain	**Properties**	**Source**
E.coli XL1-Blue		This Lab
E. coli BW25113 (F')		This Lab
E. coli BW25113 (F') ΔEP2::MVA	*E. coli* BW25113 (F') with *edd and pfkB* knockout and upper MVA pathway genome integration	This Lab
Plasmid	**Properties**	**Source**
pZE12-luc	P_LLac01, colE ori, Amp^r	This Lab
pCS27	P_LLac01, P15A ori, Kan^r	This Lab
pZE12-PMD*	pZE12-luc harboring *PMD* mutant from *E. coli*	This Study
pCS27-MK	pCS27 harboring *MK* from *S. cerevisiae*	This Study
pZE12-PMD*-ScMK	pZE12-luc harboring *PMD* from *E. coli* and *MK* from *S. cerevisiae* in one operon	This Study
pZE12-PMD*-NudB	pZE12-PMD* harboring *NudB* from *E. coli*	This Study
pZE12-PMD*-AphA	pZE12-PMD* harboring *Apha* from *E. coli*	This Study
pZE12-PMD*-NudB-ScMK	pZE12-PMD*-ScMK harboring *NudB* from *E. coli*	This Study
pZE12-PMD*-AphA-ScMK	pZE12-PMD*-ScMK harboring *AphA* from *E. coli*	This Study

2.2.3. Culture Conditions

LB medium was used for inoculation and cell propagation, and M9Y was used for *de novo* production shake flask experiments. Triplicate seed cultures were inoculated at 37℃ and 270 rpm overnight in either 20mL liquid LB medium, based on the specific experiment, with appropriate antibiotics on a gyratory shaper (New Brunswick). For *de novo* production, 2% of inoculum from overnight seeding cultures was transferred into 20mL M9Y medium with appropriate antibiotics and grown at 37℃ and 270 rpm for 3 hours. Cultures were then taken and induced with 1mM isopropyl-β-D-thiogalactoside (IPTG) and samples (1mL) were harvested by centrifugation every 12 hours for 48 hours and supernatants were subjected to subsequent HPLC analysis.

2.2.4 HPLC-quantitative Analysis

Culture samples were centrifuged at 13,000 rpm for 10 minutes and supernatants were quantitatively analyzed using HPLC analysis along with isoprenol standards and (Dionex Ultimate 3000) with Coregel-64H column (Transgenomic) and Ultimate 3000 Photoiode Array Detector. Mobile phases for peak detection were 4mM H2SO4 (solvent A). Analytes were separated following a program with 100% solvent A for 60 min at a flow rate of 0.4 mL/min. Isoprenol peaks were observed at approximately 55.4 minutes, while MVA peaks were observed around 33.7 minutes. Quantification was based on the peak area in reference to the commercial standards of each chemical.

2.3 Results and Discussion

2.3.1 Plasmid Copy-Number

For isoprenol production via the IPP-bypass MVA pathway and phosphatase overexpression, we used various gene expression combinations on high and medium plasmids transformed into the upper MVA pathway, genome integrated host strain BW25113(F')ΔEP2 asl::MVA. Previous studies obtained 2.23 g/L isoprenol production via optimization of rate-limiting steps. Through the combination of chromosomal integration and plasmid expression, out efforts reached 2.5 g/L after 48 hours with the employment of pZE-PMD*-NudB-ScMK, based on HPLC analysis of samples in reference to commercial standards. This confirms that chromosomal integration in conjunction with plasmid expression is more efficient than previous methods and should be used as the foundation of further research.

Comparing the combination of medium and high copy-number plasmids, the results show that the high copy number plasmid pZE12 expressing the various gene clusters has a higher efficiency compared to the dual expression system of pZE12 and pSC27 for isoprenol production (**Figure 2.2**). This establishes the need for higher expression of mevalonate kinase (MK) in production efforts and will be applied from here out.

pZE12-luc, a control group lacking all lower enzymatic steps of the MVA pathway, was used as a conversion comparison for constructs. When comparing the MVA accumulation to isoprenol production of constructs, we found that only partial conversion occurs leaving large amounts of MVA unaccounted for (**Figure 2.2**). This suggest that the bottleneck of isoprenol production resides in the lower portion of the pathway, encompassing the MK, PMD, and phosphatase genes. Because we have already observed an increase in MK performance based on plasmids expression systems and PMD is already optimized, we can make the inference that the

further limitations extend from the expressed phosphatase. While the performance of the two pZE12 phosphatase constructs, pZE-PMD*-NudB-ScMK and pZE-PMD*-Apha-ScMK, show comparable amounts of isoprenol production, the most interesting find was that in the absence of an overexpressed phosphatase, responsible for the conversion of IP to isoprenol, the plasmid pZE-PMD*-ScMK still showed comparable amounts of isoprenol production. This suggest that there is an endogenous phosphatase(s) at work in the genome responsible for these production efforts, thus solidifying our assumption of product synthesis limitation based on the conversion performed by the phosphatases.

Another interesting result from our experiments was the observation of residual MVA in the pZE-PMD*-NudB-MK construct, which was not present in the other pZE12 constructs. For the dual expression systems, we could potentially expect to see some residual MVA as MK is placed on medium copy number plasmid and thus the conversion of MVA to downstream products might not be as efficient; however, this phenomena on high copy number plasmid and only in the presence of *nudB* is unique and further work should be done to determine the reason for the phenomena in these cases.

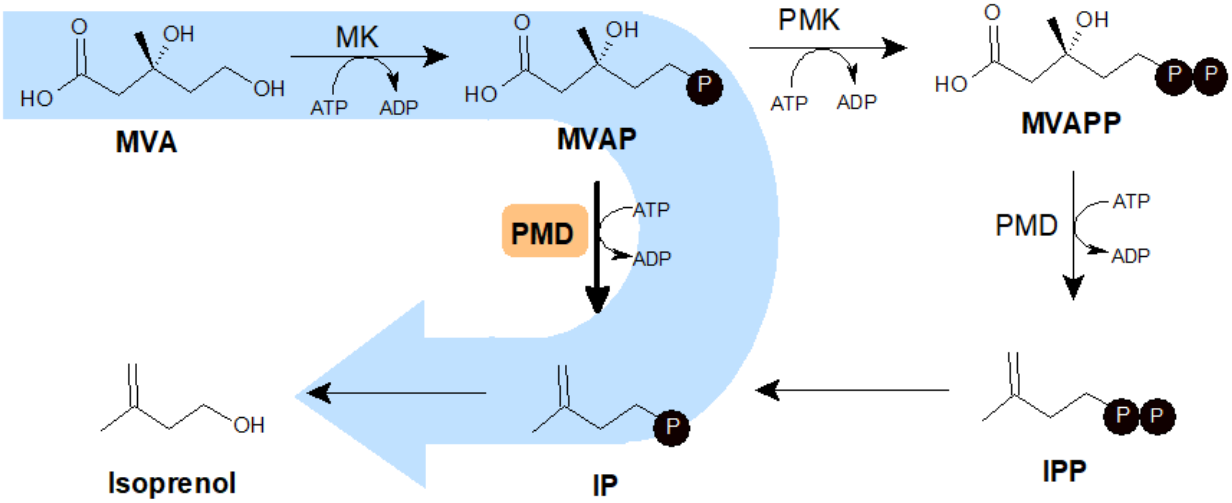

Figure 2.1 Original and IPP-Bypass mevalonate pathways for isoprenol production. The original mevalonate pathway (indicated by thin arrows) produces isopentenyl diphosphate (IPP) and isopentenyl monophosphate (IP) as intermediates in isoprenol production. The IPP-Bypass MVA pathway developed by Kang (indicated by thick arrow and mapped in blue) exploits the promiscuous activity of PMD (indicated by bold lettering and colored box) for energy efficient production. IPP-Bypass excludes the accumulation of IPP that can be toxic for cell growth. Metabolites referred: MVA, mevalonate; MVAP, mevalonate monophosphate; MVAPP, mevalonate diphosphate; IPP, isopentenyl diphosphate; IP, isopentenyl monophosphate.

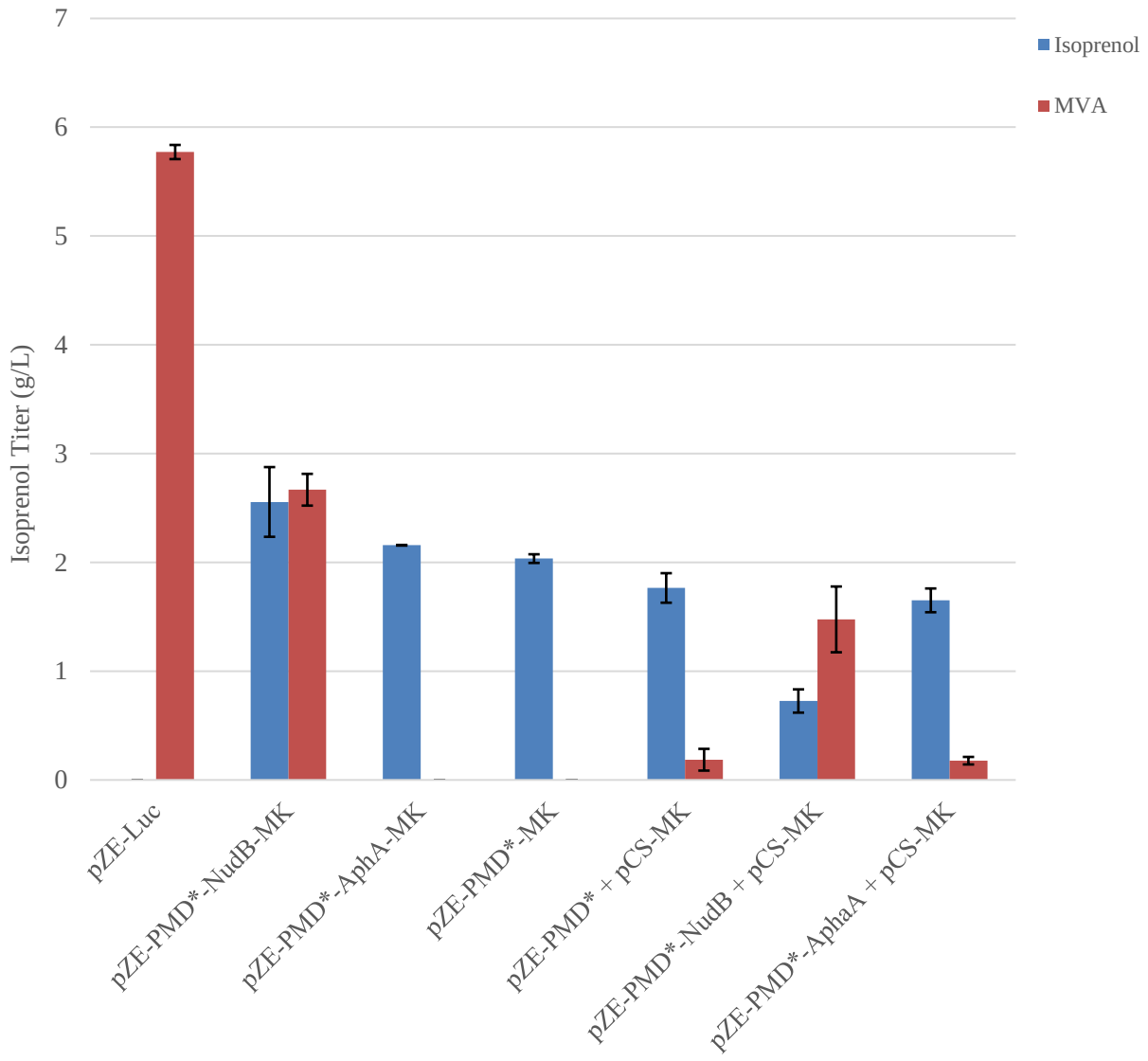

Figure 2.2. *De novo* production of isoprenol and MVA production via dual chromosomal integration and plasmid expression. BW25113(F')ΔEP2 asl::MVA was transformed with various combinations of plasmid expression systems where mevalonate kinase (MK) activity is tested via high (pZE) and medium (pCS) copy-number plasmids. Optimized PMD* and targeted phosphatases are expressed on pZE12 plasmids.

CHAPTER 3

ENZYME SCREENING FOR ALTERED ISOPRENOL PRODUCTION

3.1 Introduction and Literature Review

Enzymes are proteins that act as biological catalysts for reactions dependent on the nature and specificity of the enzyme itself. They aid in reaction acceleration and are needed for almost all cellular metabolic processes. They act upon mostly specific substrates for the conversion of desired products and are categorized into families where members are related by sequence and similarity in reaction catalyzation and function. However, research has shown that enzymes have the ability to act upon many different substrates outside of their normal biological pathway, which gives us cause to believe that this innate ability might be exploited in novel ways to increase the production efficiency of our desired product by means of using these atypical substrates as a production pathway. Based on the results from experiments conducted in Chapter 2, we concluded that there must be native enzymes at work, aside from those previously tested, that account for the isoprenol production in pZE12-PMD*-ScMK construct, lacking overexpression of any single phosphatase. In this chapter, we will build upon this hypothesis and determine if there is a native enzyme more capable of dephosphorylating IP into isoprenol than those previously observed.

Dephosphorylation is the removal of a phosphate moiety carried out by hydrolysis, which means the enzyme uses water to carry out the reaction. Phosphatases are the hydrolytic enzymes capable of this reaction mechanism. In order to determine if there is a native enzyme at work, we

must screen for those capable of such. To do so, we ran a BLAST analysis of the NudB genetic sequence using BioCyc and MG1655 as our reference genome. BLAST is a tool used to find regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. Because the phosphatase NudB performed the best in the initial production fermentation, we used it as our reference enzymatic sequence. The original findings from this BLAST resulted in 17 different enzymes. However, upon looking into the reaction mechanisms associated with these enzymes we noticed that many of them would not be viable candidates for overexpression as their modes of action were not primarily dephosphorylation. For instance, *kefC,* which is the gene associated with the KefC potassium CPA2 transporter, was a result from our BLAST that does not match the criteria for this study.

Because NudB has promiscuous activity towards IPP and IP, and based on the results gathered from the BLAST, we knew we had to broaden our search and find more alternative enzymes that held a closer similarity to the reaction mechanism of NudB rather than the protein sequence primarily. To do so, we searched for phosphatases in the EcoCyc databased under the reference genome MG1655. This resulted in approximately 77 matches, some matching those previously found in BLAST results. Due to time and efficiency, we could not test all possible candidates and thus chose different enzymes from the results that acted on a variety of substrates in order to allow for vast testing of substrate promiscuity. As NudB belongs to large family of phosphatase, multiple chosen alternatives were of the same family. It is important to note that further testing of the remaining options can be conducted in the same manner as fulfilled here and alternative results may be found. With that being said, we narrowed down the testing population to 17 alternative phosphatases for production screening. They are as follows: yihX,

apg, ybiV, yidA, yigL, phoA, yigB, ybjI, yieH, nudJ, nudF, hxpB, hxpA, nudI, otsB, nudL, and yggV. Their information can be found in **Supplementary Table 3.2**. Because NudB has promiscuous activity towards IP, it is hypothesized that these will also. While there are multiple ways we can determine which enzyme, if any, are comparable to NudB, the simplest and most efficient method of doing so would be to utilize targeted gene regulation techniques such as knockouts and gene interference.

We can probe the function of a gene or genomic feature by mutating or deleting a locus of interest and observing the resulting genotype (*27*). The adaptive immunity system in bacteria and archea, also known as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated proteins (Cas proteins), has been employed as a powerful gene editing tool due to its cost-effectiveness and ease of use (*10*). Naturally, CRISPR is an array of short, conserved repeat sequences interspaced by unique DNA sequences of similar size called spacers, which often originated from virus or plasmid DNA (*28*). The system provides immunity against invaders by acquiring these short pieces of the foreign DNA (spacers) and incorporating them into the CRISPR region of the genome. These sequences allow the host to recognize the invaders in subsequent exposures and interrupt their function (*29*). Three different types of CRISPR-Cas systems have been identified based on their respective mechanisms to generate crRNA and Cas proteins that catalyze the nucleic acid cleavage (*10, 27, 29, 30*). The Type II CRISPR system, which is the most commonly employed for genome editing due to its simplicity, requires just one Cas protein, Cas9, and two RNA components (*27, 29*). The two RNA components, crRNA and tracrRNA, base pair to form a functional guide RNA. crRNA, also known as CRISPR RNA, is comprised of the DNA sequences from the invaders directing the complex to the matching sequence; and tracrRNA, also known as transactivating CRISPR RNA, contains a region that is

complementary to the repeat region of the CRISPR locus and acts as a handle for recognition of by Cas9 DNA nuclease. Once the crRNA and tracrRNA bind, they form a complex with DNA nuclease, Cas9, and direct it to the target sequence via the crRNA. Together this complex is able to identify and cleave invading DNA, rendering its functionality. Identification of the invading DNA for sequences complementary to the crRNA occurs through the binding of Cas9 to specific sequences in the invading genome, termed Proto-spacer Adjacent Motifs or PAMs (*27*). PAM is a short DNA sequence, usually 2-6 base pairs in length, that follows the DNA region targeted for cleavage of the CRISPR-Cas9 system (*31*). Thus, the genomic locations that can be targeted for editing by CRISPR are limited by the location and presence of PAM sequences. Different Cas9 proteins recognize different PAM sites, the most common being 5'-NGG-3' recognized by Cas9 from *Streptococcus pyogenes* is used for genome editing. Two critical arginine residues in the SpCas9, Arg 1333 and Arg1335, interact with the guanines of the PAM, positioning the phosphate of the DNA backbone 5' to the PAM sequence to initiate interaction with a phosphate-lock loop within the Cas9 system and facilitate DNA unwinding (*27, 28*). Should the DNA be complementary to the guide RNA, a RNA:DNA hybrid forms (R loop) and cleavage occurs. Such DNA cleavage is a result of two different Cas9 nuclease domain actions, HNH and RuvC, leading to induced double strand breaks. In response to such damage, cells employ one of two repair pathways, non-homologous end joining (NHEJ) or homology-directed repair (HDR). Because this repair happens after DNA replication but before cell division, the break can be repaired by a donor template without any mutations. This can be exploited for precise genome editing by introducing a desired donor template for repair and taking advantage of the cellular DNA repair pathways, assuming the sequence of your target gene is known and there is a PAM sequence within appropriate distance.

Since its discovery, the CRISPR-Cas9 system has been studied extensively and optimized for genome editing purposes. Researchers found that crRNA and tracrRNA can be combined into a single guide RNA (sgRNA) for ease of manipulation. Furthermore, the CRISPR interference (CRISPRi) system was derived for targeted silencing of transcription in bacteria from the CRISPR system. First reported by Lei S. Qi and colleagues, the technique offers a complementary approach to RNA interference but primarily regulates gene expression on the transcriptional level and can be used to repress the transcription of any gene (*32*). The system uses a catalytically inactive version of the Cas9, also known as (d)Cas9 which contains two point mutations in both the RuvC-like (D10A) and HNH nuclease (H840A) domains, eliminating all endonucleolytic activity in the protein and inhibiting transcription (*32*). The group was successful in proving that when paired with sgRNA designed to complement 20 base pairs of any target gene, it was able to efficiently silence with up to 99.9% repression by hindering transcription elongation or initiation. The sgRNA is constructed with a 20 nucleotide target (nt) specific complementary region, a Cas9-binding RNA structure, and a transcription terminator derived from *S. pyogenes* (*32*). Because of its predictability, simplicity, and cost-efficiency, CRISPRi offers many advantages over other gene regulation techniques such as RNA interference and engineering DNA-binding proteins and will thus be used for the purpose of gene silencing in this work.

The purpose of this chapter is to determine the affinity of alternative endogenous enzymes towards isoprenol production using CRISPRi to systematically target and silence each enzyme separately, the sgRNA constructs and plasmids are listed in **Table 3.1**. Once constructed, the sgRNA constructs are combined with the plasmid pZE12-PMD*-ScMK for dual transformation into host strain *E. coli* BW25113(F')ΔEP2 asl::MVA, containing the (d)-Cas9

protein responsible for CRISPRi performance. This allows us to express the MVA pathway up to the point of IP dephosphorylation in hopes that one of the targeted native phosphatases, within the *E. coli* genome, will proceed to isoprenol production. Because we are silencing the expression of the targeted gene, then if the enzyme is a viable candidate there will be decrease isoprenol production compared to the positive control pZE12-PMD*-ScMK, meaning it has a positive impact on isoprenol conversion. Based on these results, gene clusters of viable candidates can then be made overexpressing in pZE-PMD*-ScMK the respective enzymes and tested against the previously studied NudB and Apha in fermentation production analysis.

Table 3.1. sgRNA constructs and plasmids

Plasmid	Properties	Source
pCS27-ApaL1-eGFP	P_LLac01, P15A ori, Kan^r, eGFP	This Study
pCS-ΔyihX	pCS27-ApaL1-eGFP harboring *yihX* sgRNA template from *E.coli*	This Study
pCS-Δapg	pCS27-ApaL1-eGFP harboring *apg* sgRNA template from *E.coli*	This Study
pCS-ΔybiV	pCS27-ApaL1-eGFP harboring *ybiV* sgRNA template from *E.coli*	This Study
pCS-ΔyidA	pCS27-ApaL1-eGFP harboring *yidA* sgRNA template from *E.coli*	This Study
pCS-ΔyigL	pCS27-ApaL1-eGFP harboring *yigL* sgRNA template from *E.coli*	This Study
pCS-ΔphoA	pCS27-ApaL1-eGFP harboring *phoA* sgRNA template from *E.coli*	This Study
pCS-ΔyigB	pCS27-ApaL1-eGFP harboring *yigB* sgRNA template from *E.coli*	This Study
pCS-ΔybjI	pCS27-ApaL1-eGFP harboring *ybjI* sgRNA template from *E.coli*	This Study
pCS-ΔyieH	pCS27-ApaL1-eGFP harboring *yieH* sgRNA template from *E.coli*	This Study
pCS-ΔnudJ	pCS27-ApaL1-eGFP harboring *nudI* sgRNA template from *E.coli*	This Study

pCS-ΔnudF	pCS27-ApaL1-eGFP harboring *nudF* sgRNA template from *E.coli*	This Study
pCS-ΔhxpB	pCS27-ApaL1-eGFP harboring *hxpB* sgRNA template from *E.coli*	This Study
pCS-ΔhxpA	pCS27-ApaL1-eGFP harboring *hxpA* sgRNA template from *E.coli*	This Study
pCS-ΔnudI	pCS27-ApaL1-eGFP harboring *nudL* sgRNA template from *E.coli*	This Study
pCS-ΔotsB	pCS27-ApaL1-eGFP harboring *otsB* sgRNA template from *E.coli*	This Study
pCS-ΔnudL	pCS27-ApaL1-eGFP harboring *nudL* sgRNA template from *E.coli*	This Study
pCS-ΔyggV	pCS27-ApaL1-eGFP harboring *yggV* sgRNA template from *E.coli*	This Study
pZE-PMD*-ScMK-ybiV	pZE-PMD*-MK harboring amplified *ybiV* from *E.coli*	This Study
pZE-PMD*-ScMK-yigL	pZE-PMD*-MK harboring amplified *yigL* from *E.coli*	This Study
pZE-PMD*-MK-yieH	pZE-PMD*-ScMK harboring amplified *yieH* from *E.coli*	This Study
pZE-PMD*-MK-nudL	pZE-PMD*-ScMK harboring amplified *nudL* from *E.coli*	This Study

3.2 Materials and Methods

3.2.1 Bacterial Strains, Plasmids, and Experimental Materials

Bacterial strains and plasmids used in this study are listed in **Table 3.1.** *E.coli* XL1-Blue was used for plasmid construction and propagation. *E. coli* BW25113(F')ΔEP2 asl::MVA was used for production as it also contains the catalytically inactive (d)Cas9 protein on the second locus, which is needed for CRISPRi function. High and medium copy number plasmids, pZE12 and pCS27, were used as backbones plasmids for construction. *E. coli* cells were grown for plasmid construction, propagation, and inoculum preparation using Luria-Bertani (LB) medium. Shake flask fermentation for isoprenol production and analysis was conducted using the biosynthesis medium M9Y. Kanamycin and ampicillin (50 mg/L) antibiotics were added to cultures when necessary.

Phusion High-Fidelity DNA polymerase, restriction endonucleases, and Quick Ligation Kit were purchased from New England Biolabs (Ipswich, MA, USA). Zyppy™ Plasmid Miniprep Kit and Zymoclean™ Gel DNA Recovery Kit were purchased from Zymo Research (Irvine, CA, USA). Standard chemicals were purchased from Sigma-Aldrich unless otherwise specified.

3.2.2 Construction of Plasmids

To improve IP hydrolysis to isoprenol, we tested 17 phosphatases for their promiscuous activity, all from the *E. coli* genome. These include yihX, apg, ybiV, yidA, yigL, phoA, yigB, ybjI, yieH, nudJ, nudF, hxpB, hxpA, nudI, otsB, nudL, and yggV. Plasmid construction and DNA manipulations were performed following the standard molecular cloning techniques and protocols. For the construction of the phosphatase interference vectors, the phosphatase genes

were individually cloned on the pCS27-ApaL1-eGFP plasmid with synthesized sgRNA between *ApaL1* and *BamHI* under the control of P_L*LacO1* promoter and T1 terminator. Cloning of sgRNA would replace eGFP expression and provide additional screening for correct constructs.

3.2.3 Culture Conditions

LB medium was used for inoculation and cell propagation, and M9Y was used for *de novo* production shake flask experiments. Triplicate seed cultures were inoculated at 37℃ and 270 rpm overnight in liquid LB medium with appropriate antibiotics on a gyratory shaper (New Brunswick). For *de novo* production of enzyme screening, 2% of inoculum from overnight seeding cultures was transferred into 3mL M9Y medium with appropriate antibiotics and grown at 37℃ and 270 rpm for 3 hours and shake flask overexpression experiments were transferred into 20 mL M9Y medium under the same conditions. Cultures were then taken and induced with 1mM isopropyl-β-D-thiogalactoside (IPTG) and samples (1mL) were harvested by centrifugation at 48 hours and supernatants were subjected to subsequent HPLC analysis. The initial experiments for screening operated on a smaller scale due to the population size of samples.

3.2.4 HPLC-quantitative Analysis

Culture samples were centrifuged at 13,000 rpm for 10 minutes and supernatants were quantitatively analyzed using HPLC analysis along with isoprenol standards and (Dionex Ultimate 3000) with Coregel-64H column (Transgenomic) and Ultimate 3000 Photoiode Array Detector. Mobile phases for peak detection were 4mM H2SO4 (solvent A). Analytes were separated following a program with 100% solvent A for 60 min at a flow rate of 0.4 mL/min. Isoprenol peaks was observed at approximately 55.4 minutes, while MVA peaks were observed

around 33.7 minutes. Quantification was based on the peak area in reference to the commercial standards of each chemical.

3.3 Results and Discussion

Through the employment of CRISPRi using sgRNA targeting desired native enzymes based on narrowed BLAST results for similar sequencing and functionality, we were able to determine which enzymes had an affinity for the conversion of IP to isoprenol compared to NudB and AphA previously characterized. Intriguingly, based on preliminary screening, *nudB* and *aphA* actually had increased percent production compared to the control when silenced inferring that their inhibition may be more beneficial than their overexpression. Out of the seventeen possible alternative candidates, we found four that were applicable to the work. Based upon HPLC data analysis, *ybiV, yigL, yieH,* and *nudL* all had decreased isoprenol conversion percentage compared to the control group as illustrated in **Figure 3.1**. This suggest that each of these enzymes have the ability to act in a positive manner towards the production of isoprenol and could be employed in the creation of gene clusters for overexpression and production analysis. Although *nudL* was not greatly decreased compared to the control, we proceeded to move forward with comparisons as *nudL* is part of the same family of *nudB* which has proven to be successful in isoprenol conversion. It is important to note that the preliminary study conducted to analyze the alternative enzymes was completed in smaller fermentation working volumes compared to traditional approaches due to the size of the population being tested. Thus, the production titers are compared as percentages to the control set to 100%.

Interestingly, in this preliminary study, we also observed increased isoprenol conversion percentages in constructs targeting *yihX, apg, ybjI,* and *otsB* compared to the control group. This

suggest that these targeted endogenous enzymes have a negative influence on isoprenol production and their continued inhibition in future work could be beneficial towards increased isoprenol titers. The most important information gathered from this experimental data is that *NudB* and *AphA* are not the most efficient phosphatases for isoprenol conversion, resulting in another production bottleneck. Based on the results gathered *ybiV*, *yigL,* and *yieH* have a higher affinity for isoprenol production compared to *nudB* and *aphA*, as their specific production based on targeted silencing was decreased more significantly (**Figure 3.2**). Consequently, efforts should focus on incorporation of these specific enzymes.

Based on the results of the enzyme screening, we conducted extended testing in shake flask to determine if the previous claim that nudB and aphA are not the most efficient phosphatases for isoprenol production. To test this, we amplified the desired genes that had a been found to have positive influence on isoprenol production (ybiV, yigL, yieH, and nudL) and cloned them separately into the expression vector pZE-PMD*-ScMK. This allowed us to overexpress each enzyme respectively and compare production efforts to NudB and AphA, which had been previously observed in Chapter 1. We also chose to look at overexpression of *nudJ*, despite it depicting a negative influence on isoprenol conversion, due its relation to *nudB*. Our expected results were that *yieH* would have highest isoprenol titers based on initial enzyme screening and *nudJ* would produce less than the control. HPLC analysis of samples concluded that *yigL* was the best performing phosphatase system resulting in approximately 2.95 g/L, surpassing that of both *nudB* and *aphA* (**Figure 3.3**). While this is different than our expected results, it is noted that both *yigL* and *yieH* performed better than the control and both serve as viable candidates for future production efforts. Results also confirmed that *nudJ* resulted in decreased titers compared to the control as expected from preliminary screening. It is noted that

the results for *ybiV* and *nudL* are not illustrated as plasmid construction for these two strains was unsuccessful. In conclusion, based on preliminary screening and fermentation results, repeat studies should be completed to ensure consistency in results and observe findings of those constructs not successful in this study. Furthermore, enzyme characterization should be completed to determine the kinetic factors associated with these results.

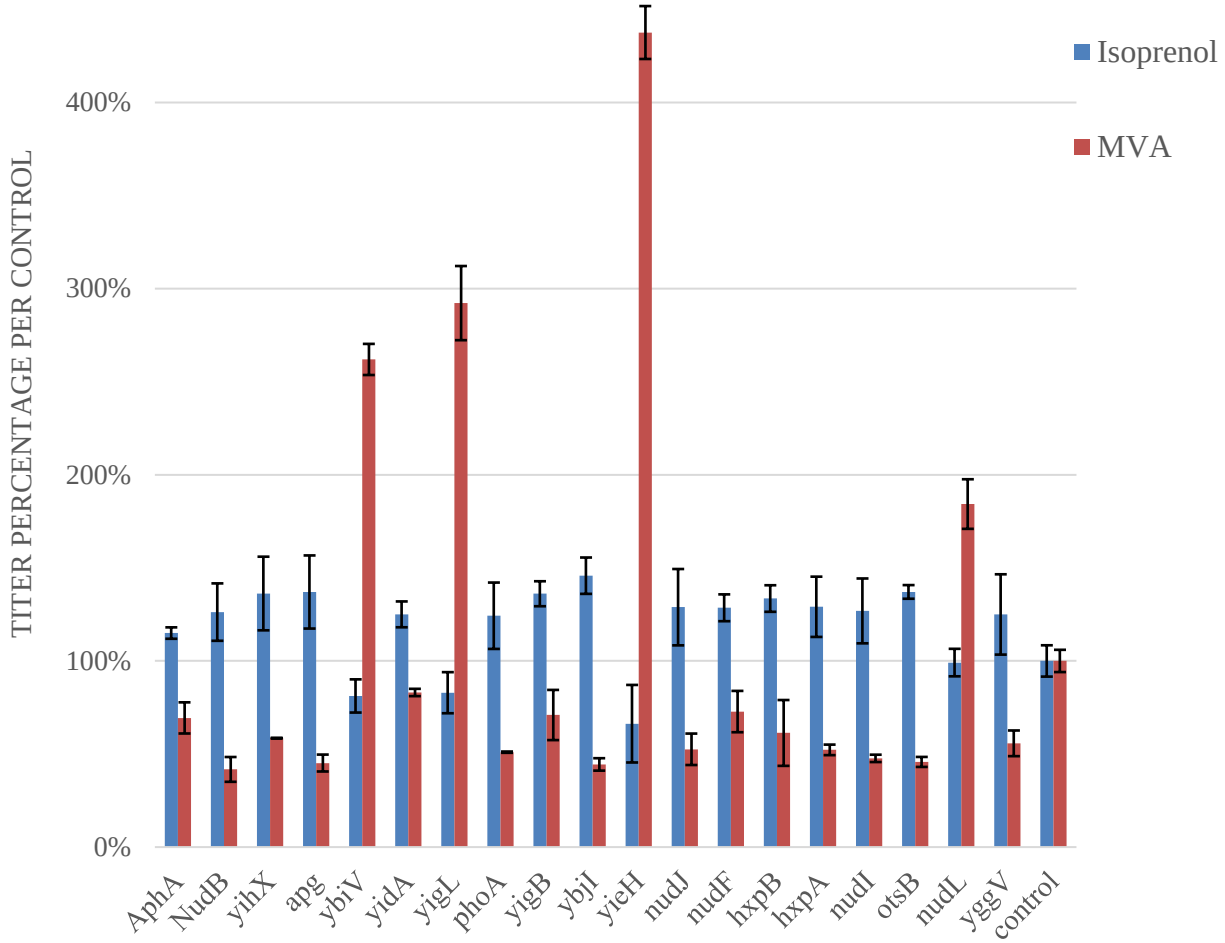

Figure 3.1. **Enzyme screening for altered isoprenol production.** *De novo* production of isoprenol in *E. coli* via targeted sgRNA integration on expression vector pCS27-ApaLI-eGFP in combination with pZE-PMD*-ScMK. The production of isoprenol was measured 48 hours post transfer inoculation and induction by 0.5 mM IPTG. The control for this experiment was pZE-PMD*-ScMK + pCS27.

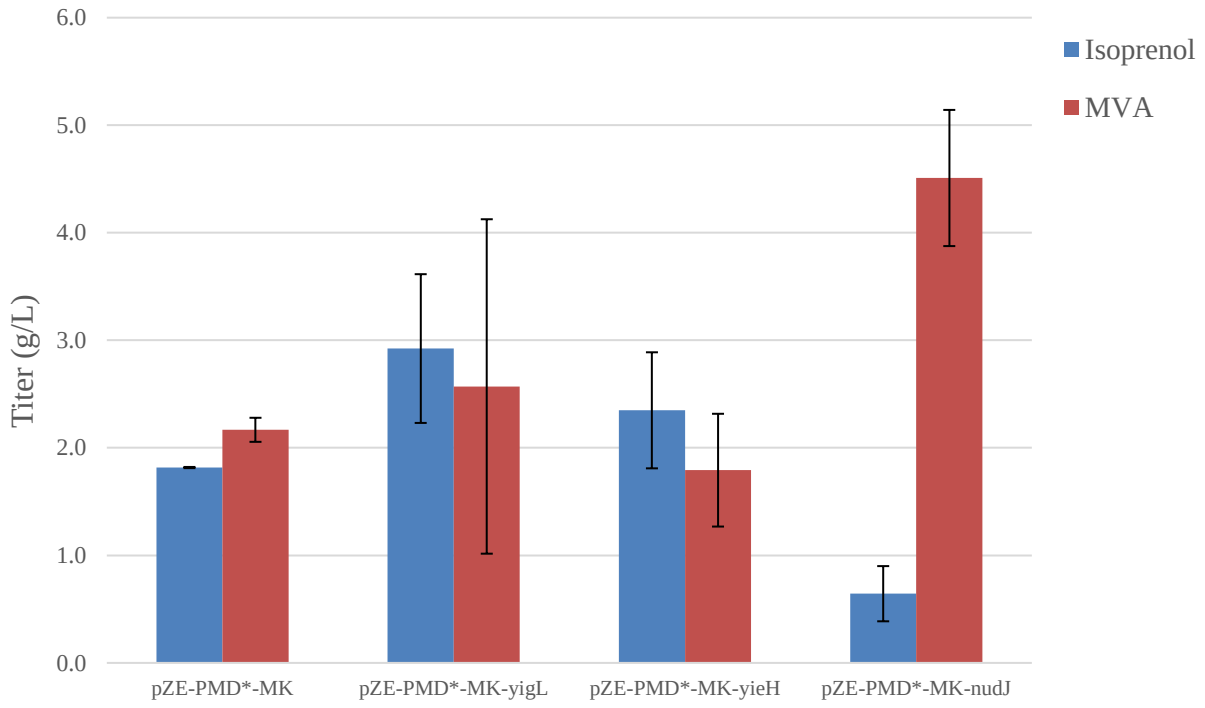

Figure 3.2. ***De novo*** **production of isoprenol via gene clusters from enzyme screening results**. Targeted genes for alternative phosphatases were transformed into pZE-PMD*-ScMK in triplicates. Isoprenol production was measured 48 hours post transfer inoculation and induction by 0.5 mM IPTG. The control for this experiment was pZE-PMD*-ScMK.

Supplementary Table 3.2 Alternative Enzyme Information.

Gene	Protein	Molecular Function	Common Substrate
aphA	Class B acid phosphatase	Hydrolase	Phosphate monoesters
nudB	Dihydroneopterin triphosphate diphosphatase	Hydrolase	Dihydroneopterin triphosphate
yihX	Alpha-D-glucose 1-phosphate phosphatase	Hydrolase	Alpha-D-glucose 1-phosphate (Glc1P)
apg	Glucose-1-phosphatase	Hydrolase	Glucose 1-phosphate
ybiV	Sugar phosphatase YbiV	Hydrolase	Sugar phosphates
yidA	Sugar phosphatase YidA	Hydrolase	Sugar phosphates
yigL	Pyridoxal phosphate phosphatase YigL	Hydrolase	Sugar phosphates
phoA	Alkaline phosphatase	Hydrolase	Phosphate monoesters
yigB	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YigB	Hydrolase	5-amino-6-(5-phospho-D-ribitylamino)uracil
ybjI	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YbjI	Hydrolase	5-amino-6-(5-phospho-D-ribitylamino)uracil
yieH	6-phosphogluconate phosphatase	Hydrolase	D-gluconate 6-phosphate
nudJ	Phosphatase NudJ	Hydrolase	4-amino-2-methyl-5-hydroxymethylpyrimidine pyrophosphate, and thiamine pyrophosphate
nudF	ADP-ribose pyrophosphatase	Hydrolase	ADP-mannose, ADP-glucose, and ADP-ribose
hxpB	Hexitol phosphatase B	Hydrolase	Sugar phosphates
hxpA	Hexitol phosphatase A	Hydrolase	Sugar phosphates
nudI	Nucleoside triphosphatase NudI	Hydrolase	Nucleoside triphosphates
otsB	Trehalose-6-phosphate phosphatase	Hydrolase	Trehalose 6-phosphate, 2-deoxyglucose-6-phosphate, and glucose-6-phosphate
nudL	Uncharacterized Nudix hydrolase NudL	Hydrolase	Nucleoside diphosphate derivatives (*Predicted*)
yggV	dITP/XTP pyrophosphatase	Hydrolase	Non-canonical purine nucleotides

CHAPTER 4

FUTURE WORK

As with any practice, thinking outside the immediate realm of optimization is a must in order to meet the supply of and demand for high value product for future generations. While the focus of this work revolved around the production of isoprenol via metabolic regulation and engineering of the MVA pathway, it would be foolhardy to assume that stopping here and accepting the results of this experiment would suffice for industrial scale production. For this reason, future efforts must build upon the scope of this work to further increase isoprenol titers. The purpose of this chapter is to discuss the various applications to be done as an extension of this work and the possibilities for both increasing volume and for increasing efficiency of yield. This includes using the gathered data to further expand upon the phosphatases found by utilizing the combination of knockouts found to increase isoprenol production in conjunction with overexpressed phosphatases. Furthermore, the assumption for this study centered on the hypothesis that the bottleneck of isoprenol production resided in the lower portion of the MVA pathway. While we have proven that this is acceptable based on the results gathered, this does not consider the metabolic processes and imposed burdens outside this pathway that might hinder desired chemical production. This would include taking into account the broader aspects of metabolic flux throughout the system, rate limiting intermediates of production, and inhibition of competitive pathways.

4.1 Extension of Current Work

Future work, applied in conjunction with results gathered in this study, could lead to potentially even greater yields of isoprenol by further optimizing the process outlined above. We have proven that one possible bottleneck of isoprenol production lies in the lower portion of the MVA pathway, but this can be further exploited. From the enzyme screening we have found multiple enzymes that have either positive or negative influence on isoprenol production; careful application of this knowledge has great potential to further optimize this process. Based on the results of the enzyme screening in Chapter 3, we can continue the construction of strains that possess both overexpression and knockout of desired genes. This would begin with determining which enzyme naturally inhibits isoprenol production the most. From results of preliminary sgRNA screening, future work can focus on fermentation of yihX, apg, ybjI, and ots as they increased isoprenol titers the most compared to the control, meaning they have a negative native influence on production. Based on results gathered from this, future work to further streamline this process can be conducted by performing dual transformation of the optimal sgRNA construct with *yigL,* which performed the best in overexpression experiments. Furthermore, the combination of multi-gene overexpression utilizing both *yigL* and *nudB* could be employed as well as in amalgamation with sgRNA constructs. While these efforts should result in increased titers, it is important to keep in mind the metabolic burden placed on the cell during multi-gene plasmid expression and efforts should be made to reduce this burden. This limitation to the process could potentially include the use of multi-plasmid systems using different copy-numbers.

4.2 Metabolic Flux in *E. coli* and Inhibition of Competitive Pathways

Metabolic flux in microbial cell factories is represented by a large number of metabolic reactions involving the conversion of a chosen carbon source into the building blocks needed for biosynthesis, and the rate of turnover of molecules through metabolic networks. As the complexity of microbial pathways plays a pivotal role in the final accumulation of any product and because these metabolic networks are highly branched, desired chemical production is dependent on the metabolic flux throughout the system. These networks are connected by shared metabolites and cofactors such as ATP, ADP, NADH, and NADPH (*33*). Flux is regulated by the enzymes involved in each respective pathway and is a function of gene expression, translation, modifications, and interactions (*33*). As production efforts are dependent on the flux through desired pathways and network branching, directed and increased flux through targeted pathways could result in increased production efforts. Because multiple pathways may share common metabolites, or precursors, it is important that we regulate the pathways so that biochemical reactions are being driven towards desired product formation. For the purpose of this section, we will look directly at the production and consumption of acetyl-CoA by multiple pathways and the regulation of such. However, unbalanced distribution of metabolic flux in microbial cell factories can result in poor cell growth and inefficient production. Therefore, it is important to find the proper balance between cellular metabolism and chemical synthesis.

Acetyl-CoA is an important metabolic intermediate shared by multiple pathways within the metabolic flux network and is responsible for many reactions involved in protein, carbohydrate, and lipid metabolism. Acetyl-CoA is also a platform chemical for the production of high-value products, including isoprenol (*34*). In *E. coli*, acetyl-CoA can be produced from glucose during glycolysis, acetate, or fatty acids via β-oxidation (**Figure 4.1**). Because acetyl-

CoA is a precursor for a variety of products, overexpression of the pathways that form these products can result in the depletion of acetyl-CoA, potentially compromising cellular function (*35*). In the presence of sufficient oxygen, acetyl-CoA is mainly oxidized via the TCA cycle to generate energy; however, acetyl-CoA is also consumed for the biosynthesis of fatty acids which is later returned to the system. A common metabolic engineering technique to improve desired product yield is to overexpress bottleneck enzymes and knockout competing pathways to increase carbon flux towards the pathway of interest. In this section we will discuss the various pathways that possess this ability.

4.2.1 Fatty Acid Synthetic Pathway

The primary role of bacterial fatty acids is to act as the hydrophobic component of the membrane and storage lipids (*36*). The coordination of growth rate and membrane synthesis is vital for organism longevity and production. Fatty acids are synthesized in *E. coli* by a dissociated-enzyme (type II) system. The system begins with the precursors derived from acetyl-CoA pool and there are two known mechanisms for the initiation of fatty acid biosynthesis. The first pathway consists of the conversion of acetyl-CoA to malonyl-CoA followed by the conversion of malonyl-CoA to malonyl-ACP. The second pathway condenses acetyl-CoA with malonyl-ACP to form acetoacetyl-ACP. Each pathway is dictated by specific enzymes and cofactors and inhibition of any could result in redirected flux through the system.

The first pathway for fatty acid synthesis begins with carboxylation of acetyl-CoA to malonyl-CoA by acetyl-CoaA carboxylase (ACC) in the presence of bicarbonate. ACC is composed of four subunits encoded by four genes: *accA, accB, accC,* and *accD* (*37*). AccA and AccD are proteins that catalyze the carboxyltransferase partial reaction, AccB is the biotin

carboxyl carrier protein (BCCP), and AccC is the biotin carboxylase (*38*). ACC has previously been established as a rate-limiting step as increased levels of ACC increased the rate of fatty acid synthesis (*39*). Malonyl-CoA, produced from the ACC reaction, is required for all elongation steps and is utilized for fatty acid synthesis only once it has been converted in malonyl-ACP by malonyl-CoA/ACP transacylase (FabD) (*36*, *40*). The second pathway for fatty acid synthesis involves the condensation of acetyl-CoA and malonyl-ACP by FabH, β-ketoacyl-ACP synthase III. FabH is thought to play a role in determining the type of fatty acid made and thus has major influence on the total biosynthesis. With this, the ACC gene cluster, FabD, and FabH seem appropriate targets for the regulation of the fatty acid synthetic pathway.

4.2.2 TCA Cycle

Another competitive pathway for acetyl-CoA consumption is the citric acid cycle, also known as the tricarboxylic acid cycle (TCA) cycle. This is a catabolic pathway for energy sources for cells and an important part of aerobic respiration. The cycle also provides precursors of certain amino acids and of reducing agent NADH that are both used in many other reactions throughout cellular metabolism. While the TCA cycle is important for cell growth, previous studies have already found *E. coli* strains lacking TCA cycle-regulated genes that exhibit longer stationary phase survival compared to wild-type (*41*). The condensation of acetyl-CoA with oxaloacetate to form citrate in *E. coli* is catalyzed by the *gtlA* gene product (*35*). Because of its key position as the first enzymatic step of the TCA cycle and there is no alternative enzyme, *gltA* has been assumed to play an important role as a control point in determining the metabolic rate of the cell and metabolic flux through the system (*42*).

4.2.2 Conclusion

Each of these pathway processes have the potential to increase isoprenol production by inhibition and downregulation of key genes thereby redistributing flux through the MVA pathway. However, because the network of biochemical reactions is so complex and is an important part of cell growth, differentiation, and biosynthesis, multiple strategies may be appropriate to achieve optimal results for production of desired product. Future work to be done in this area would consist of systematically targeting each enzyme separately to determine the impact on both cell growth and isoprenol production. This could include the employment of CRISPRi/sgRNA or fine tuning of expression vectors and promotors for downregulation of genes based on gene importance. For instance, complete inhibition of *gltA* will likely not result in viable cell growth; however, down regulation of gene expression could provide adequate metabolic tradeoff for both central metabolism and biosynthesis. Another potential method to achieve proper results would involve having genes involved in fatty acid synthesis inhibited by CRISPRi/sgRNA with the application of proper and controlled induction time. This would allow ample time for cells to accumulate necessary membrane products before the process starts. From here, multi-gene regulation could be employed to further the investigation of isoprenol production.

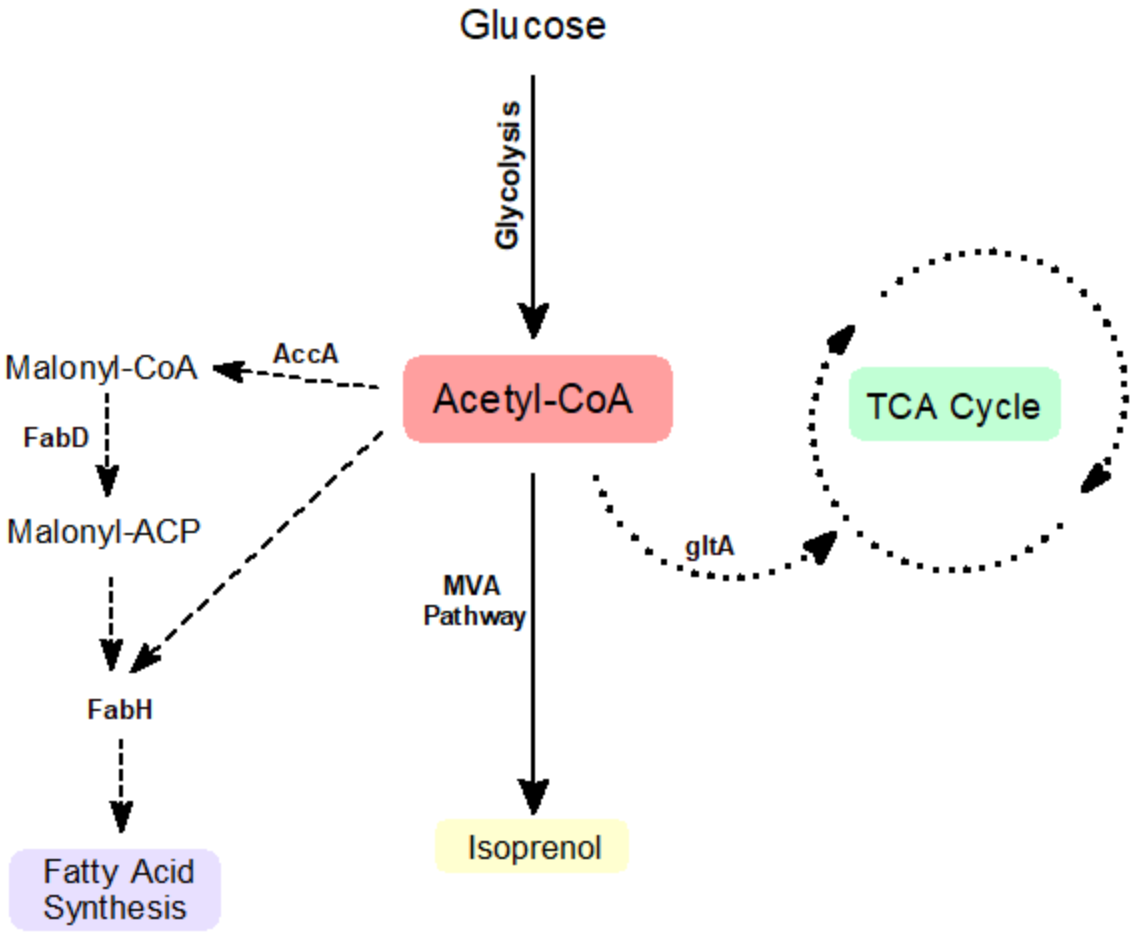

Figure 4.1. Metabolic flux of *E. coli*. Acetyl-CoA is formed through glycolysis of the central carbon metabolism from glucose. Multiple pathways branch from this as part of the biosynthetic network. Multiple pathways utilize Acetyl-CoA for product synthesis. Dashed lines represent the two pathways of fatty acid synthesis, dotted lines represent flux through the TCA cycle, and strict lines represent desired pathway flux.

CHAPTER 5

CONCLUSION

Industrial microbes such as E. coli harbor multiple biosynthetic pathways and give us the ability to synthesize industrially valuable chemicals from cost-efficient and readily available carbon feedstocks. With increasing understanding of these biosynthetic components and their machinery, it is possible to engineer microbial host for industrial production of myriad desired products. Isoprenol serves as an excellent alternative for fuel sources and as a precursor for many commercial chemicals in industrial applications including rubber which could be a viable alternative to petroleum-based applications. Unfortunately, however, current production methods are insufficient to meet future market demand. The utilization of metabolic engineering tools such as CRISPRi for increased production has made extraordinary inroads in the last decade toward bridging the gap between supply and demand and continues to be at the forefront of novel creation. Considering the structural diversity and ease of manipulation in microbial host, they remain an ideal host for chemical synthesis. The work here describes the establishment of a more efficient synthetic pathway for enhanced isoprenol production by expanding upon existing metabolic pathways in E. coli as discussed. The continued systematic regulation of the mevalonate pathway through the optimization and overexpression of rate limiting enzymes based on previous studies and chromosomal integration of upper MVA pathway into E.coli and the application of alternative enzymes responsible for the conversion of IP to isoprenol for optimal conversion. These results have demonstrated the ability of expanding the production of isoprenol

and have laid the foundation in laboratory proceedings for further improvement of products and processes. Furthermore, we have discussed the future applications to this work including enzyme characterization and optimization, and the inhibition of competitive pathways using various metabolic engineering applications for increased final isoprenol titers.

In conclusion, we have successfully established a more efficient biosynthetic pathway for isoprenol production by identifying a bottleneck in the lower MVA pathway encompassing the hydrolysis of IP by various phosphatases and the employment of different copy-number plasmids expressing mevalonate kinase (MK). We have uncovered a new phosphatase, yigL, capable of the acting upon IP to produce isoprenol, allowing us to achieve 2.92 g/L isoprenol titers. In this work, we mainly focused on new pathway design and validation conducted on shake flask proportions. To test the industrial potentials, we do believe that bioreactor-based fermentation is necessary. Prior to mass application of this method, further investigation is warranted to ensure the most efficient production pathways and strains can be used for large-scale production, and each facet of this process still needs to be further optimized as discussed in Chapter 4. The synthetic pathways established in this work would undergo this optimization and work as a basis for the future to scale-up production. Quantities of isoprenol necessary for industrial and energy use could potentially be achieved using these methods, ensuring a decreased dependence on fossil fuels and non-renewable petroleum products for manufacturing, energy production, and transportation. This will allow greater energy security for nations that historically have had to depend on others for their resources, and will, with judicious application of the method described here, create a path to sustainable energy production.

REFERENCES

1. F. Xin *et al.*, in *Current Developments in Biotechnology and Bioengineering*, S. P. Singh, A. Pandey, G. Du, S. Kumar, Eds. (Elsevier, 2019), pp. 207-232.
2. G. Stephanopoulos, Synthetic Biology and Metabolic Engineering. *ACS Synthetic Biology* **1**, 514-525 (2012).
3. R. Garcia-Granados, J. A. Lerma-Escalera, J. R. Morones-Ramirez, Metabolic Engineering and Synthetic Biology: Synergies, Future, and Challenges. *Front Bioeng Biotechnol* **7**, 36 (2019).
4. A. A. Cheng, T. K. Lu, Synthetic Biology: An Emerging Engineering Discipline. *Annual Review of Biomedical Engineering* **14**, 155-178 (2012).
5. B. M. Woolston, S. Edgar, G. Stephanopoulos, Metabolic engineering: past and future. *Annu Rev Chem Biomol Eng* **4**, 259-288 (2013).
6. J. V. Pham *et al.*, A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front Microbiol* **10**, (2019).
7. A. Fleming, The discovery of penicillin. *British Medical Bulletin* **2**, 4-5 (1944).
8. V. Chubukov, A. Mukhopadhyay, C. J. Petzold, J. D. Keasling, H. G. Martín, Synthetic and systems biology for microbial production of commodity chemicals. *NPJ Syst Biol Appl* **2**, 16009 (2016).
9. S. Pontrelli *et al.*, Escherichia coli as a host for metabolic engineering. *Metabolic Engineering* **50**, 16-46 (2018).

10. J. D. Sander, J. K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347-355 (2014).

11. D. Na *et al.*, Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nature Biotechnology* **31**, 170-174 (2013).

12. G. L. Rosano, E. A. Ceccarelli, Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol* **5**, 172-172 (2014).

13. R. H. Kluger, in *Encyclopedia Britannica*. (2018).

14. Z. Yanning *et al.*, Metabolic engineering of Escherichia coli for high-specificity production of isoprenol and prenol as next generation of biofuels. *Biotechnology for Biofuels* **6**, 1-13 (2013).

15. A. Kang *et al.*, Isopentenyl diphosphate (IPP)-bypass mevalonate pathways for isopentenol production. *Metabolic Engineering* **34**, 25-35 (2016).

16. S. Atsumi, T. Hanai, J. C. Liao, Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89 (2008).

17. K. W. George *et al.*, Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production. *Biotechnology and Bioengineering* **111**, 1648 (2014).

18. A. Kang, C. W. Meadows, N. Canu, J. D. Keasling, T. S. Lee, High-throughput enzyme screening platform for the IPP-bypass mevalonate pathway for isopentenol production. *Metabolic Engineering* **41**, 125-134 (2017).

19. J. Wang *et al.*, Engineering of a Highly Efficient Escherichia coli Strain for Mevalonate Fermentation through Chromosomal Integration. *Appl Environ Microbiol* **82**, 7176-7184 (2016).

20. B. Ou, C. Garcia, Y. Wang, W. Zhang, G. Zhu, Techniques for chromosomal integration and expression optimization in Escherichia coli. *Biotechnology and Bioengineering* **115**, 2467-2478 (2018).

21. W. D. Hollinshead *et al.*, Examining Escherichia coli glycolytic pathways, catabolite repression, and metabolite channeling using Δpfk mutants. *Biotechnology for Biofuels* **9**, 212 (2016).

22. K. W. George *et al.*, Metabolic engineering for the high-yield production of isoprenoid-based C-5 alcohols in E-coli. *SCIENTIFIC REPORTS* **5**, (2015).

23. A. Kang *et al.*, Optimization of the IPP-bypass mevalonate pathway and fed-batch fermentation for the production of isoprenol in Escherichia coli. *Metabolic Engineering* **56**, 85-96 (2019).

24. C.-L. Liu, H.-R. Bi, Z. Bai, L.-H. Fan, T.-W. Tan, Engineering and manipulation of a mevalonate pathway in Escherichia coli for isoprene production. *Applied Microbiology & Biotechnology* **103**, 239-250 (2019).

25. C. Wang, B. Zada, G. Wei, S.-W. Kim, Metabolic engineering and synthetic biology approaches driving isoprenoid production in Escherichia coli. *Bioresource Technology* **241**, 430-438 (2017).

26. M. Jahn, C. Vorpahl, T. Hübschmann, H. Harms, S. Müller, Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. *Microbial Cell Factories* **15**, 211 (2016).

27. D. M. Thurtle-Schmidt, T. W. Lo, Molecular biology at the cutting edge: a review on CRISPR/CAS9 gene editing for undergraduates. *Biochemistry and Molecular Biology Education* **46**, 195-205 (2018).

28. T. Karvelis *et al.*, crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. *RNA Biol* **10**, 841-851 (2013).

29. Y. Cui, J. Xu, M. Cheng, X. Liao, S. Peng, Review of CRISPR/Cas9 sgRNA design tools. *Interdisciplinary Sciences: Computational Life Sciences* **10**, 455-465 (2018).

30. Y. Ma, L. Zhang, X. Huang, Genome modification by CRISPR/Cas9. *The FEBS journal* **281**, 5186-5193 (2014).

31. D. Gleditzsch *et al.*, PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA Biol* **16**, 504-517 (2019).

32. M. H. Larson *et al.*, CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* **8**, 2180-2196 (2013).

33. J. Nielsen, It Is All about MetabolicFluxes. *Journal of Bacteriology* **185**, 7031-7035 (2003).

34. S. Zhang *et al.*, Metabolic engineering for efficient supply of acetyl-CoA from different carbon sources in Escherichia coli. *Microbial Cell Factories* **18**, 130 (2019).

35. A. Krivoruchko, Y. Zhang, V. Siewers, Y. Chen, J. Nielsen, Microbial acetyl-CoA metabolism and metabolic engineering. *Metabolic engineering* **28**, (2014).

36. J. E. Cronan, J. Thomas, Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *Methods Enzymol* **459**, 395-433 (2009).

37. S. Lee, E. Jeon, H. S. Yun, J. Lee, Improvement of fatty acid biosynthesis by engineered recombinant Escherichia coli. *Biotechnology and Bioprocess Engineering* **16**, 706 (2011).

38. E. S. James, J. E. Cronan, Expression of two Escherichia coli acetyl-CoA carboxylase subunits is autoregulated. *J Biol Chem* **279**, 2520-2527 (2004).

39. M. S. Davis, J. Solbiati, J. E. Cronan, Jr., Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in Escherichia coli. *J Biol Chem* **275**, 28593-28598 (2000).

40. S. Lee, E. Jeon, Y. Jung, J. Lee, Heterologous Co-expression of accA, fabD, and Thioesterase Genes for Improving Long-Chain Fatty Acid Production in Pseudomonas aeruginosa and Escherichia coli. *Applied Biochemistry and Biotechnology* **167**, 24-38 (2012).

41. S. Gonidakis, S. E. Finkel, V. D. Longo, Genome-wide screen identifies Escherichia coli TCA-cycle-related mutants with extended chronological lifespan dependent on acetate metabolism and the hypoxia-inducible transcription factor ArcA. *Aging Cell* **9**, 868-881 (2010).

42. S. Park, J. McCabe, J. Turna, R. Gunsalus, Regulation of the citrate synthase (gltA) gene of Escherichia coli in response to anaerobiosis and carbon supply: role of the arcA gene product. *Journal of bacteriology* **176**, 5086-5092 (1994).

43. High-throughput enzyme screening platform for the IPP-bypass mevalonate pathway for isopentenol production. (2017).

44. Metabolic engineering for the high-yield production of isoprenoid-based C5 alcohols in E. coli. (2015).

45. T. M. M. Abdellatief *et al.*, Recent trends for introducing promising fuel components to enhance the anti-knock quality of gasoline: A systematic review. *Fuel* **291**, 120112 (2021).

46. T. Baba *et al.*, Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006.0008-2006.0008 (2006).

47. K. H. Choi, R. J. Heath, C. O. Rock, beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *Journal of bacteriology* **182**, 365-370 (2000).

48. H. H. Chou, J. D. Keasling. (AMERICAN SOCIETY FOR MICROBIOLOGY, United States, 2012), pp. 7849.

49. J. M. Clomburg, S. Qian, Z. Tan, S. Cheong, R. Gonzalez, The isoprenoid alcohol pathway, a synthetic route for isoprenoid biosynthesis. *Proceedings of the National Academy of Sciences of the United States* **116**, 12810 (2019).

50. J. E. Cronan, G. L. Waldrop, Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research* **41**, 407-435 (2002).

51. Y. Dang *et al.*, Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome biology* **16**, 1-10 (2015).

52. G. del Solar, M. Espinosa, Plasmid copy number control: an ever-growing story. *Mol Microbiol* **37**, 492-500 (2000).

53. D. R. Ede, N. Farhang, J. D. Stover, R. D. Bowles, in *Comprehensive Biomaterials II*, P. Ducheyne, Ed. (Elsevier, Oxford, 2017), pp. 589-599.

54. S. Fahs, P. Lujan, M. Köhn, Approaches to Study Phosphatases. *ACS Chemical Biology* **11**, 2944-2961 (2016).

55. Y. Fujita, H. Matsuoka, K. Hirooka, Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology* **66**, 829-839 (2007).

56. R. J. Heath, C. O. Rock, Inhibition of β-Ketoacyl-Acyl Carrier Protein Synthase III (FabH) by Acyl-Acyl Carrier Protein in Escherichia coli(∗). *Journal of Biological Chemistry* **271**, 10996-11000 (1996).

57. M.-J. Heo, H.-M. Jung, J. Um, S.-W. Lee, M.-K. Oh, Controlling Citrate Synthase Expression by CRISPR/Cas9 Genome Editing for n-Butanol Production in Escherichia coli. *ACS Synthetic Biology* **6**, 182-189 (2017).

58. J. Hou *et al.*, Rewiring carbon flux in Escherichia coli using a bifunctional molecular switch. *Metabolic Engineering* **61**, 47-57 (2020).

59. H. J. Janßen, A. Steinbüchel, Fatty acid synthesis in Escherichia coli and its applications towards the production of fatty acid based biofuels. *Biotechnology for Biofuels* **7**, 7 (2014).

60. Y. Jiang *et al.*, Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* **81**, 2506-2514 (2015).

61. J. Kim *et al.*, Engineering Saccharomyces cerevisiae for isoprenol production. *Metabolic Engineering* **64**, 154-166 (2021).

62. A. Krivoruchko, Y. Zhang, V. Siewers, Y. Chen, J. Nielsen, Microbial acetyl-CoA metabolism and metabolic engineering. *Metabolic Engineering* **28**, 28-42 (2015).

63. E. Kuznetsova *et al.*, Genome-wide Analysis of Substrate Specificities of the Escherichia coli Haloacid Dehalogenase-like Phosphatase Family*. *Journal of Biological Chemistry* **281**, 36149-36161 (2006).

64. S. J. Li, J. E. Cronan, Jr., Growth rate regulation of Escherichia coli acetyl coenzyme A carboxylase, which catalyzes the first committed step of lipid biosynthesis. *J Bacteriol* **175**, 332-340 (1993).

65. Z. Luo *et al.*, Enhancing isoprenoid synthesis in Yarrowia lipolytica by expressing the isopentenol utilization pathway and modulating intracellular hydrophobicity. *Metabolic Engineering* **61**, 344-351 (2020).

66. J. Mairhofer, T. Scharl, K. Marisch, M. Cserjan-Puschmann, G. Striedner, Comparative Transcription Profiling and In-Depth Characterization of Plasmid-Based and Plasmid-Free Escherichia coli Expression Systems under Production Conditions. *Appl Environ Microbiol* **79**, 3802-3812 (2013).

67. S. C. Phulara, P. Chaturvedi, P. Gupta, Isoprenoid-Based Biofuels: Homologous Expression and Heterologous Expression in Prokaryotes. *Appl Environ Microbiol* **82**, 5730-5740 (2016).

68. S. M. Podkovyrov, T. J. Larson, Identification of Promoter and Stringent Regulation of Transcription of the fabH, fabD and fabG Genes Encoding Fatty Acid Biosynthetic Enzymes of Escherichia Coli. *Nucleic Acids Research* **24**, 1747-1752 (1996).

69. J. P. Torella *et al.*, Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proceedings of the National Academy of Sciences* **110**, 11290-11295 (2013).

70. G. Tyler, S. S. Potter, D. Somnath, Integrating gene regulatory pathways into differential network analysis of gene expression data. *Scientific Reports* **9**, 1-12 (2019).

71. M. Wade, High-throughput silencing using the CRISPR-Cas9 system: a review of the benefits and challenges. *Journal of biomolecular screening* **20**, 1027-1039 (2015).

72. Y. Wei, A. Mohsin, Q. Hong, M. Guo, H. Fang, Enhanced production of biosynthesized lycopene via heterogenous MVA pathway based on chromosomal multiple position integration strategy plus plasmid systems in Escherichia coli. *Bioresource Technology* **250**, 382-389 (2018).

73. P. Xu *et al.*, Modular optimization of multi-gene pathways for fatty acids production in E. coli. *Nature Communications* **4**, 1409 (2013).

74. X. Zhang, A. Agrawal, K. Y. San, Improving fatty acid production in Escherichia coli through the overexpression of malonyl coA-acyl carrier protein transacylase. *Biotechnol Prog* **28**, 60-65 (2012).

75. Z. Zuo, J. Liu, Structure and Dynamics of Cas9 HNH Domain Catalytic State. *Scientific Reports* **7**, 17271 (2017).

76. World Health Organization. (2018, May 2). *Ambient air pollution*. World Health Organization. https://www.who.int/airpollution/ambient/en/#:~:text=Around%2091%25%20of%20the%20world's,and%20South%2DEast%20Asia%20regions

77. MarketWatch. (2021, May 3). *Isoprenol Market Growth, Size 2021, Development Plans, Growing At A CAGR of 2.3%, Covid-19 Impact Analysis, Recent Trends, Top Countries Data, and 2026 Forecast Research Report.* MarketWatch. https://www.marketwatch.com/press-release/isoprenol-market-growth-size-2021-development-plans-growing-at-a-cagr-of-23-covid-19-impact-analysis-recent-trends-top-countries-data-and-2026-forecast-research-report-2021-05-03.