Engineering a novel pathway for isoprenoid synthesis

by

Alkiviadis Orfefs Chatzivasileiou
Dipl., National Technical University of Athens (2013)

Submitted to the Department of Chemical Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

Signature redacted

Signature of Author.....
Department of Chemical Engineering
May 20, 2019

Signature redacted

Certified by.....
Gregory Stephanopoulos
Willard Henry Dow Professor of Chemical Engineering and Biotechnology
Thesis Supervisor

Signature redacted

Accepted by.....
Patrick S. Doyle
Robert T. Haslam (1911) Professor of Chemical Engineering
Graduate Officer

Engineering a novel pathway for isoprenoid synthesis

by

Alkiviadis Orfefs Chatzivasileiou

Submitted to the Department of Chemical Engineering
on May 20, 2019 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Chemical Engineering

ABSTRACT

Isoprenoids comprise a large class of chemicals, of significant interest due to their diverse properties. Most isoprenoids are plant secondary metabolites and are of commercial importance due to their varied applications in fields spanning medicine, agriculture, flavors, fragrances, cosmetics and nutrition. Biological production of isoprenoids in microbes is considered to be the most efficient and commercially viable way for their large-scale production. Thus far, isoprenoid biosynthesis has been performed through pathways inextricably linked to glycolysis. Furthermore, these pathways are inherently limited due to their extensive cofactor requirements, complex regulation and large number of steps. In this thesis we present a novel pathway for isoprenoid synthesis, the Isopentenol Utilization Pathway (IUP), which aims to overcome these limitations. This pathway functions through the double phosphorylation of an isopentenol, either isoprenol or prenol, to produce the main precursors to isoprenoid synthesis, isopentenyl diphosphate (IPP) or dimethylallyl diphosphate (DMAPP). This pathway is radically different from naturally-occurring pathways or their engineered variants because it is only two steps long, uses an externally-provided isoprenol as its substrate instead of a glucose-derived catabolite, and uses only a single co-factor, ATP.

We identify suitable enzymes, construct the pathway and proceed to demonstrate an *in vivo* proof of concept. After optimizing the pathway feedstock, we proceed to show that IUP is decoupled from central carbon metabolism. We demonstrate that the IUP can quickly produce copious amounts of IPP & DMAPP and can be used for the production of a variety of isoprenoids. The IUP flux exceeded the capacity of almost all downstream pathways tested, was competitive with the highest isoprenoid fluxes reported as well as against state-of-the art isoprenoid pathways. Furthermore, we elaborate on our progress towards improving the capacity of a downstream farnesene synthesis pathway, to catch up with and fully utilize IUP's production capacity. Finally, we propose a new scheme for the use of the IUP to produce functionalized isoprenoids using functionalized isopentenols to introduce functionalizations in isoprenoid backbones, and we show preliminary results of this application.

Thesis Supervisor: Gregory Stephanopoulos
Title: Willard Henry Dow Professor of Chemical Engineering and Biotechnology

Acknowledgements

There is a large number of people I feel compelled to acknowledge for their invaluable help and support throughout my career at MIT.

First and foremost, I would like to thank Professor Gregory Stephanopoulos, my thesis advisor, for his support and mentorship throughout my PhD. I have always considered him a true ally in all my endeavors, and I am extremely grateful for his mentorship, his trust and belief in me and for all the resources he placed at my disposal. I have found him to be optimistic and patient, even when I have not been, always discerning and focused on impact and the big picture. I am thankful for his mentorship style which has allowed me to grow as a scientist and for the excellent professional relationship which we have maintained throughout the years. I would also like to acknowledge the remaining members of my thesis committee, Professor Heather Kulik and Professor Hadley Sikes, whose support, guidance and insightful advice have benefitted me immensely in the progress of my PhD.

I would like to express my immense gratitude to Steven Edgar, to whom I am immeasurably indebted for his help, his mentorship and his friendship. As an older graduate student in the lab, Steven took it upon himself to mentor me and help me grow to be a more competent scientist. My involvement with the IUP started with Steven; without him none of the work in this thesis would have been possible and for that I cannot thank him enough. I can only hope that I am able to reciprocate even a fraction of all the good Steven has done for me. Whenever I face adversity I am always reminded of his encouragement and his impeccable work ethic, which inspire me to this day.

I have also been fortunate enough to have collaborated with Valerie Ward, during her one-year visit as a postdoctoral researcher in the lab. Her quick mind, strong work ethic and pleasant disposition made collaborating with Valerie a real joy. Working together with Valerie, taking shifts

throughout the day, doubled our productivity helped accomplish things that I would have otherwise been unable to achieve on my own and for that I am grateful.

I have had the pleasure to interact with many lab members. I want to thank Kang Zhou for mentoring me during my first days in the lab and helping me get started. Moreover, I want to thank older grad students and post-docs including Peter Enyeart, Kangjian Qiao, Niju Narayanan, Ben Woolston, Jason King, Devin Currie, Andrew Silverman, Brian Pereira, and Mark Keibler for their help and scientific advice. I would also like to acknowledge younger graduate students Nian Liu, Wentao Dong, Bob Van Hove and Vincent Zu for all the assistance they lent me in my PhD and for all the discussions we had on science, grad school and the life beyond. I would also like to thank Rosangela Dos Santos and Nicholas Pasinella for helping with the lab finances and for enabling the lab to run smoothly and I would like to express my gratitude to Brian Smith for assisting me in my duties as the lab Environment, Health and Safety (EHS) representative.

My feelings towards Lisa Guay are to an extent indescribable. In our long friendship and our work together in the Ashdown House Executive Committee (AHEC) and the Graduate Student Council (GSC) I have developed boundless admiration and respect for her due to her clear vision, dependability and level-headedness. I count myself fortunate to have had the pleasure and honor to work with her for three years in student government and to have had an impeccable personal and professional relationship with her. I want to thank my very good and dear friend Nabeel Dahod, with whom we spent many hours working on coursework and having fun outside the lab. His friendship and support truly means a great deal to me and helped me get through the more difficult times of my PhD. I would also be remiss if I did not acknowledge Yamini Krishnan, who has always been supportive, encouraging and far too kind. I also wish to thank my friends Andrew Fiore, James Kaczmarek, Carlos Pons Siepermann, Daniel Consoli, Harry Watson and Zsigmond Varga for their long friendship and for helping me take my mind off the lab and AHEC/GSC during our regular D&D sessions.

In my career as a student leader, I have had the opportunity to serve together with some truly dedicated and competent individuals. I want to thank fellow AHEC Officers Jess Hebert, Hojun Suk, Alin Tomescu, Andrew Rzeznik, Christopher Foy and Calvin Sun for being a pleasure to work with as we worked towards the smooth running of our dorm. For my career in the GSC, I would like to acknowledge Angie Crews, Krithika Ramchander, Richard Zhang and Nicole Moody, and express my gratitude to all of them for being amazing team players and extraordinarily competent. We would not have achieved nearly as much were it not for all their efforts. Special mention is due to GSC Presidents Arolyn Conwill and Peter Su, under whom I had the distinct privilege of serving, for working in unison with their teams towards the betterment of graduate student life. I wish to express my gratitude towards our GSC admins, Lauren McLean and Elizabeth Granese for their infinite patience and service to the GSC.

Finally, I would like to acknowledge Cynthia Barnhart, Ian Waitz, Suzy Nelson, Blanche Staton, David Friedrich, Peter Cummings, Jag Patel, Kim Haberlin, Judy Robinson, Naomi Carton, Sarah Gallop and Adam Berinsky. Through my multiple interactions from them, I have been influenced by many of their positive traits, inspiring me to keep a cool head during crises, a focused mind during negotiations and allowing me to improve my leadership, management and public speaking skills. I would thus like to thank them for helping me grow as a person as well as for their service to graduate students.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1. Isoprenoids and their industrial relevance

The name isoprenoids or terpenoids is used to refer to a large class of natural products comprising more than 65,000 molecules (1). Isoprenoids are structurally very diverse (2, 3), with most exhibiting complex structures having chiral, cyclic skeletons. Chemically, terpenoids display great structural diversity, with many species being multifunctionalized. Isoprenoids can be found in all living organisms, but they are most abundant in plants, where they play a very important role in plant physiology by being involved in growth regulation, attraction of pollinators, cell defense, photosynthesis, vitamin production and a variety of other biological functions (4). Furthermore, isoprenoids constitute an interesting class of molecules due to their varied applications. Isoprenoids have been used as agrichemicals, such as the diterpenes gibberellins, herbal medicines, such as the triterpenes glycyrrhizin and ginsenosides, nutraceuticals such as the tetraterpenoids astaxanthin and lycopene (5). An important application of isoprenoids is for pharmaceutical purposes, as several terpenoids have been found to have anti-cancer, anti-microbial, anti-allergenic, anti-inflammatory, and immunomodulatory properties (5, 6). Isoprenoids of note in this category include the anticancer agent paclitaxel and the anti-malarial molecule artemisinin (7). Isoprenoids also of interest to the cosmetics industry because they can have interesting coloring, flavoring or olfactory properties. Examples include carotenoids (e.g. lycopene, b-carotene, astaxanthin), citronellol, geraniol, menthol, limonene, pinene and citral (4, 5, 7).

Isoprenoids can be categorized according to the number of isoprene units their backbone is made of. Major classes in this categorization scheme include hemi- (1 isoprene unit-C_5), mono- (2 isoprene units-C_{10}), sesqui- (3 isoprene units-C_{15}), di- (4 isoprene units-C_{20}), sester- (5 isoprene units-C_{25}), tri- (6 isoprene units-C_{30}), tetra- (8 isoprene units-C_{40}) and poly- (more than 8 isoprene units) terpenes (8).

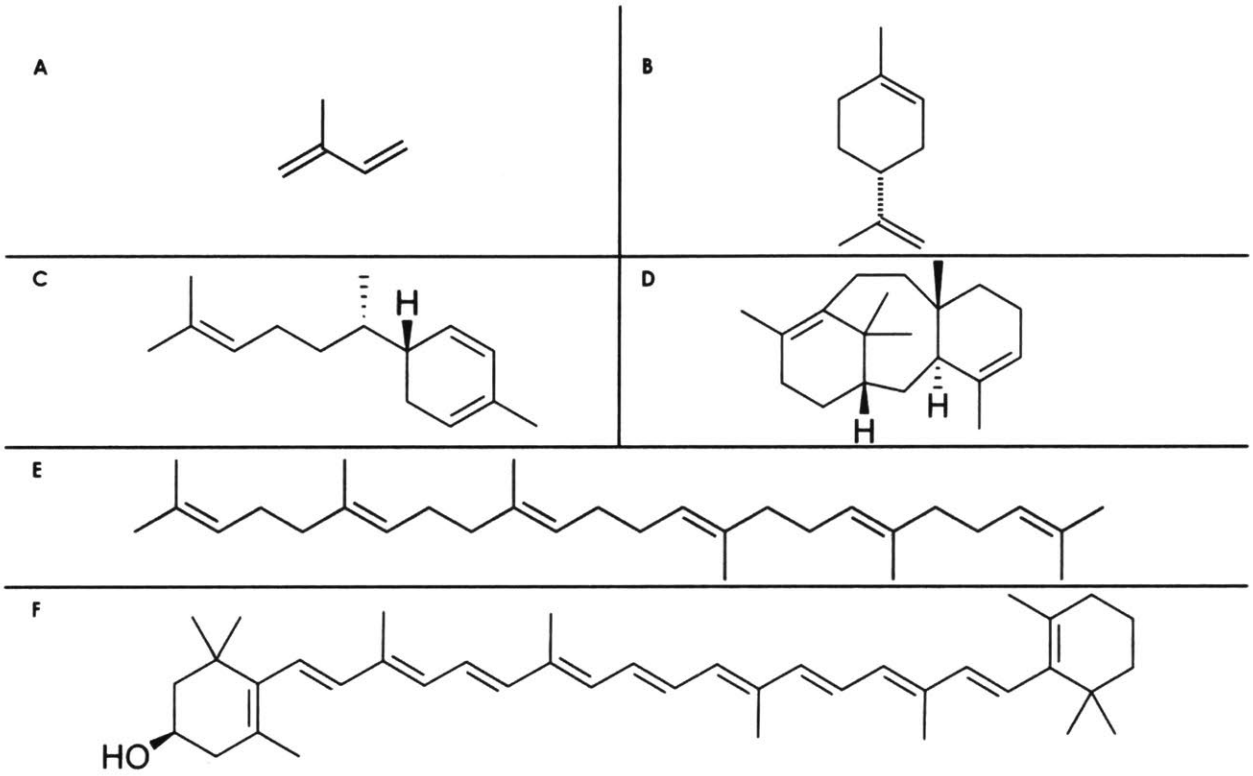


Figure 1-1: Sample isoprenoid structures. (**A**) Isoprene, a hemiterpene, (**B**) Limonene, a monoterpene, (**C**) Zingibernene, a sesquiterpene, (**D**) Taxadiene, a monoterpene, (**E**) Squalene, a triterpene, (**F**) Cryptoxanthin, a tetraterpene

1.2. Production of isoprenoids

One route to commercially produce isoprenoids is by extraction from natural sources (plants). However, since isoprenoids are present in plants in miniscule quantities, this is usually a very costly undertaking (9). Another possible route for the production of isoprenoids is through chemical synthesis, however many of the very properties that make terpenoids excellent therapeutic molecules also make them exceptionally difficult to synthesize. Terpenoids typically

possess complex molecular architectures replete with diverse ring systems and chiral centers, are highly oxygenated, and contain a large number of solvated hydrogen-bond donors and acceptors (10). Consequently, their synthesis has presented particular challenges for synthetic chemistry (11). Pursuing a route for isoprenoid biosynthesis and improving upon it through metabolic engineering has thus emerged as a production method with the capability to overcome the limitations of the plant extraction and chemical synthesis routes.

All isoprenoids are biosynthesized through the condensation of two precursor 5-carbon molecules, dimethylalyl diphosphate (DMAPP) and its isomer isopentenyl diphosphate (IPP). After the two precursor molecules are synthesized they can be further condensed to generate larger precursor molecules, such as geranyl diphosphate (GPP,C_{10}), farnesyl diphosphate (FPP, C_{15}) and geranylgeranyl diphosphate (GGPP,C_{20}), which they themselves can be condensed to form even larger molecules (5). These molecules can then be subjected to further modification by other enzymes, such as cyclization by terpene synthases (TPS) or functionalization by cytochrome P450 monooxygenases (P450s) to produce a wide array of isoprenoids (5, 7).

There exist two metabolic pathways which can produce DMAPP and IPP, the basic building blocks for isoprenoid biosynthesis. These two pathways are the mevalonate pathway (MVA pathway) and the 2-C-methyl-d-erythritol-4-phosphate pathway (MEP pathway) (12). The MVA pathway can be found in animals, some bacteria, in archaea and the plant cytosol, whereas the MEP pathway is present in bacteria, green algae and plant plastids (4, 13). The MEP pathway is the sole pathway for isoprenoid biosynthesis in *Escherichia coli* (13). Both pathways require a supply of NAD(P)H to function (14).

The MVA (mevalonate) pathway (shown in Figure 1-2), in its eukaryotic variant, which is the most prevalent, involves seven reactions for the production of IPP and DMAPP starting from acetyl-CoA. In the first step, two molecules of acetyl-CoA are condensed into aceto-acetyl-CoA by the enzyme acetoacetyl-CoA thiolase (AACT). Then follows further condensation with another molecule of acetyl-CoA to produce 3-hydroxy-3-methylglutaryl CoA (HMG-CoA). This reaction is

catalyzed by HMG-CoA synthase (HMGS). Then, HMG-CoA is reduced by HMG-CoA reductase (HMGCR) to produce mevalonate. In the fourth and fifth reaction steps, mevalonate is phosphorylated into phosphomevalonate by mevalonate kinase (MVK) and then into diphosphomevalonate by phosphomevalonate kinase (PVMK). This step is followed by a decarboxylation step, catalyzed by diphosphomevalonate decarboxylase (MVD), which produces IPP. Finally, DMAPP can be produced through the partial isomerization of IPP, catalyzed by isopentenyl diphosphate isomerase (IDI) (12, 15).

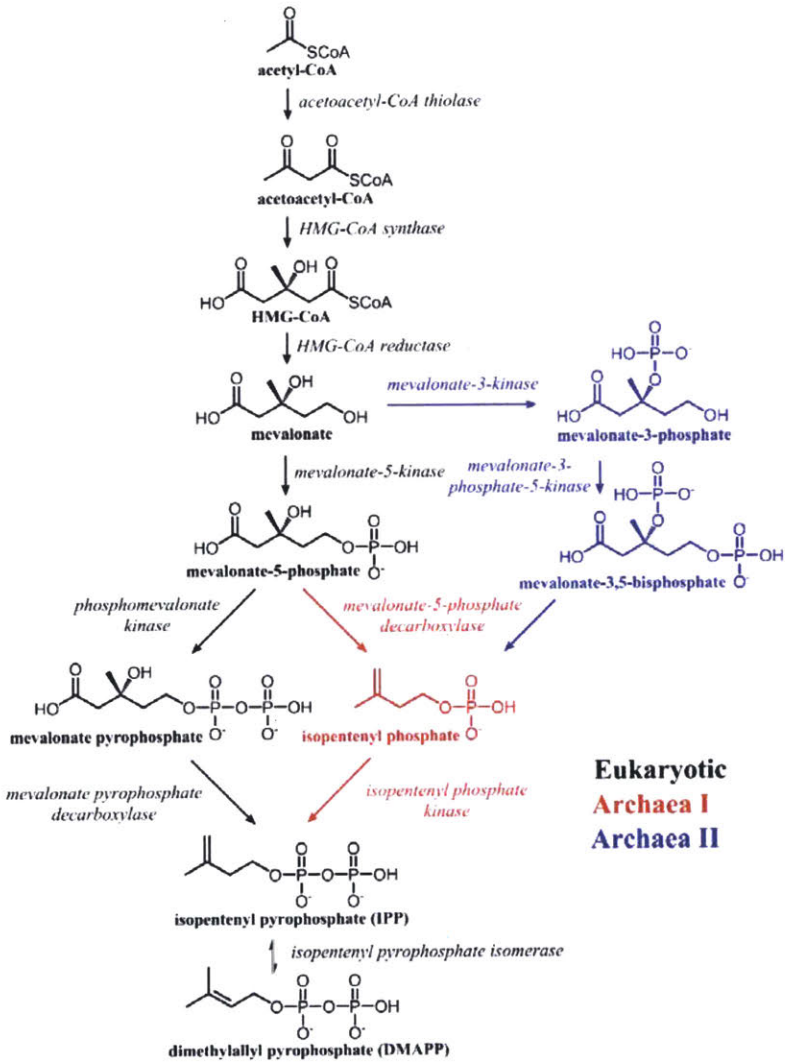


Figure 1-2: The Mevalonate pathway. This pathway, which is utilized by all eukaryotic cells, is one of the two natural pathways for the production of IPP & DMAPP. Highlighted are the variants of the pathway found in archaea (16).

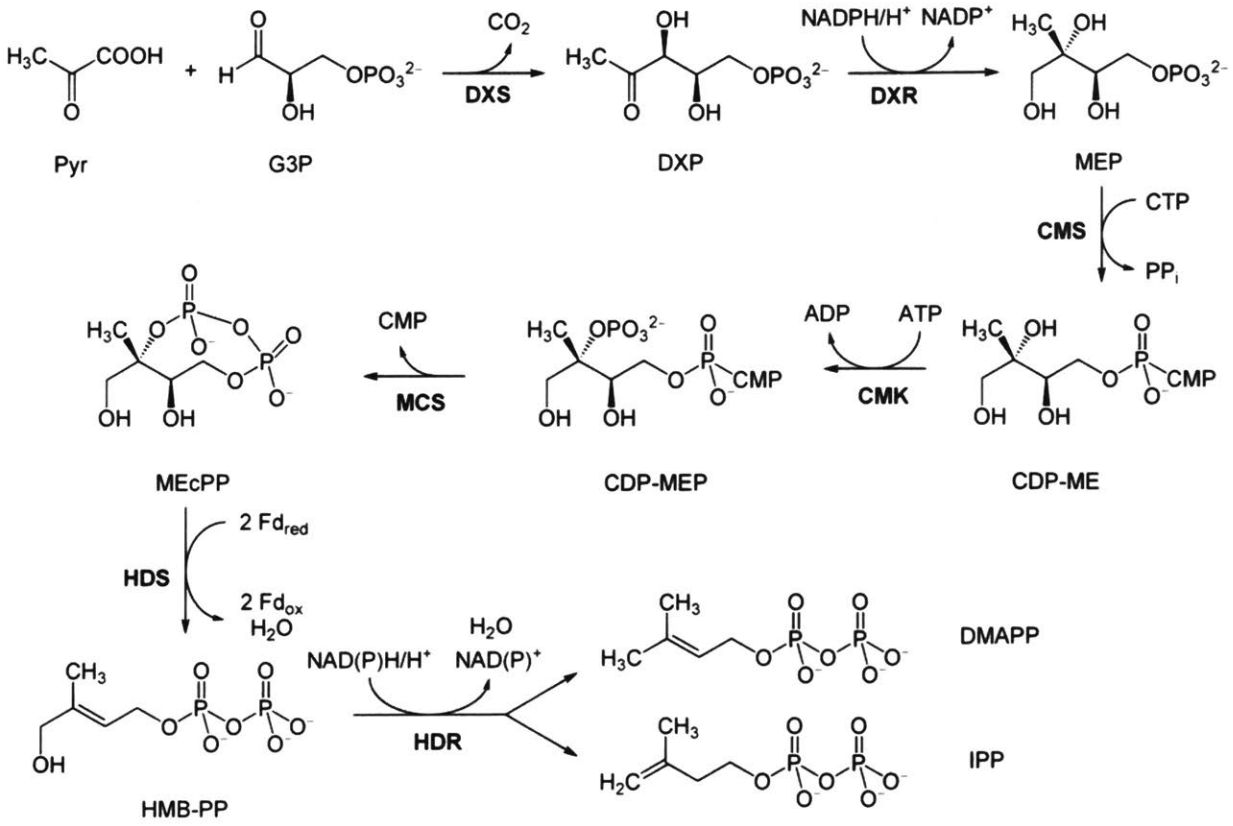


Figure 1-3: The MEP or DXP pathway. This pathway, which is utilized by all prokaryotic cells, is one of the two natural pathways for the production of IPP & DMAPP.

The MEP or DXP pathway utilizes pyruvate and glyceraldehyde 3-phosphate to produce IPP and DMAPP and includes seven reactions. The pathway starts with the condensation of pyruvate and glyceraldehyde 3-phosphate by DXP synthase (Dxs) with thiamine to produce 1-deoxy-D-xylulose 5-phosphate (DXP). Then follows the NADPH-dependent reduction and isomerization of DXP into 2C-methyl-D-erythritol 4-phosphate (MEP), which is catalyzed by DXP reductoisomerase (Dxr/IspC). MEP then reacts with CTP and is converted into 4- diphosphocytidyl-2C-methyl D-erythritol (CDP-ME) through the enzymatic action of 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (YgbP/IspD). CDP-ME undergoes a phosphorylation by the ATP-dependent 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (YchB/IspE) to produce 4-diphosphocytidyl-2C-methyl D-erythritol 2-phosphate (CDP-MEP). Then, CDP-MEP is cyclized by 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (YchB/IspF), with simultaneous elimination of CMP, to

form 2C-methyl-D-erythritol 2,4-cyclodiphosphate (MEC or MEcPP). Then, MEC undergoes a reductive ring opening, catalyzed by the NADPH-dependent 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (GcpE/IspG), which affords 4-hydroxy-3-methylbut-2-en-1-yl diphosphate (HMB-PP). Finally, HMB-PP is reduced by 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase (LytB/IspH) to produce a mixture of IPP and DMAPP (12, 18). The IPP:DMAPP ratio produced by IspH is around 5:1 (7, 19).

1.3. Thesis Objectives and Overview

The main objective of this thesis is to engineer a novel biological pathway for the production of isoprenoids. Initially, we will design a tentative alternative pathway to either of the natural pathways and investigate enzymes for promiscuous activity that will enable this pathway's realization. After establishing *in vivo* and *in vitro* proof of concepts, we will show that the pathway can be used for the production of a variety of isoprenoids, performs with high flux and is thus a viable and competitive alternative to the natural ones. We will then investigate strategies for further improving the pathway, and for using the pathway for the production of novel functionalized isoprenoids.

Chapter 2 opens with an overview of the inherent limitations faced by both the MEP and MVA pathways, as well as some past efforts to overcome them. It then proposes a new upstream pathway for the production of the main isoprenoid intermediate, the Isopentenol Utilization Pathway (IUP) aimed at circumventing these limitations. We subsequently elaborate on our efforts to select the proper enzymes required for this pathway, focusing on the enzyme that catalyzes the first step. After proving that the IUP can work *in vitro*, we proceed to demonstrate an *in vivo* proof of concept in *E. coli*.

In Chapter 3 we conduct an initial assessment of the IUP. After we optimize the feedstock and we investigate the necessity of including a certain enzyme, we focus on measuring the rate at

which the pathway can produce isoprenoid intermediates and estimating the flux going through the pathway. Moreover, we show that the IUP is decoupled from Central Carbon Metabolism, overcoming one of the key limitations faced by the natural pathways.

In Chapter 4, we couple the IUP with downstream modules to produce different isoprenoid final products and particularly focus on the optimization of downstream modules for the production of lycopene and taxadiene. We show that the IUP can be used for the production of a variety of isoprenoids and that, after debottlenecking of the downstream synthesis module, it can do so with flux that competes with some of the best fluxes reported in literature. Moreover, we show that the IUP can favorably compete against strains bearing state-of-the art isoprenoid pathways.

In Chapter 5, we focus on further improving the downstream modules to fully utilize IUP's full potential. After detailing our work on the taxadiene and lycopene downstream, we further discuss our strategy for identifying the maximum productivity of the IUP and then improving it, through coupling the IUP with a downstream module for farnesene production. We then elaborate on our efforts on improving productivity by upregulating the downstream module and finally achieve significant improvements through the simultaneous upregulation of upstream and downstream.

Chapter 6 focuses on the production of novel functionalized isoprenoids. After proposing a scheme for the use of the IUP to introduce functionalizations in isoprenoid backbones, and therefore isoprenoids, we elaborate on our work to produce halogenated amorphadiene as well as halogenated limonene.

In Chapter 7, we shift focus to highlight a number of efforts at improving the performance of the MEP pathway, particularly focusing on overcoming the oxygen sensitivity enzyme IspG suffers from. To achieve this, we investigate strategies for protein engineering IspG, scavenging reactive oxygen species (ROS) by adding antioxidants to the culture broth or producing them *in vivo*; or operating the whole MEP in a species that can naturally modulate ROS levels. This chapter serves to illustrate some of the limitations engineering a natural pathway entails, limitations that the IUP circumvents.

Finally, in Chapter 8, we address the overall implications and impact of the work and propose future directions for study.

1.4. Bibliography

1. Berthelot K, Estevez Y, Deffieux A, Peruch F (2012) Isopentenyl diphosphate isomerase: A checkpoint to isoprenoid biosynthesis. *Biochimie* 94:1621–1634.
2. Ajikumar PK, et al. (2008) Terpenoids: Opportunities for Biosynthesis of Natural Product Drugs Using Engineered Microorganisms. *Mol Pharm* 5(2):167–190.
3. Bohlmann J, Keeling CI (2008) Terpenoid biomaterials. *Plant J* 54(4):656–669.
4. Chandran SS, Kealey JT, Reeves CD (2011) Microbial production of isoprenoids. *Process Biochem* 46(9):1703–1710.
5. Misawa N (2011) Pathway engineering for functional isoprenoids. *Curr Opin Biotechnol* 22(5):627–633.
6. Kirby J, Keasling JD (2009) Biosynthesis of plant isoprenoids: Perspectives for microbial engineering. *Annu Rev Plant Biol* 60(1):335–355.
7. Kirby J, Keasling JD (2009) Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu Rev Plant Biol* 60:335–355.
8. Galata M, Mahmoud S (2012) *Bioactive plant isoprenoids: Biosynthetic and biotechnological approaches* (Elsevier) doi:10.1016/B978-0-444-59514-0.00005-5.
9. Langenheim JH (1994) Higher plant terpenoids: A phytocentric overview of their ecological roles. *J Chem Ecol* 20(6):1223–1280.
10. Koehn FE, Carter GT (2005) The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* 4(3):206–220.
11. Li JW, Vederas JC (2013) Drug Discovery and Natural Products : 161(2009):161–166.
12. Nagegowda D a. (2010) Plant volatile terpenoid metabolism: Biosynthetic genes,

transcriptional regulation and subcellular compartmentation. *FEBS Lett* 584(14):2965–2973.

13. Wang J, Xiong Z, Li S, Wang Y (2015) Exploiting exogenous MEP pathway genes to improve the downstream isoprenoid pathway effects and enhance isoprenoid production in Escherichia coli. *Process Biochem* 50(1):24–32.

14. Partow S, Siewers V, Daviet L, Schalk M, Nielsen J (2012) Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in Saccharomyces cerevisiae. *PLoS One* 7(12):1–12.

15. Matsumi R, Atomi H, Driessen AJM, van der Oost J (2011) Isoprenoid biosynthesis in Archaea - Biochemical and evolutionary implications. *Res Microbiol* 162(1):39–52.

16. Mevalonate pathway diagram Available at: https://commons.wikimedia.org/wiki/File:Wiki_pathway_hi_def_tiff.tif [Accessed February 21, 2019].

17. Non-mevalonate pathway reactions in the biosynthesis of isoprenoids Available at: https://en.wikipedia.org/wiki/Non-mevalonate_pathway#/media/File:Non-mevalonate_pathway.svg [Accessed March 5, 2019].

18. Zhou K, Zou R, Stephanopoulos G, Too H-PP (2012) Metabolite Profiling Identified Methylerythritol Cyclodiphosphate Efflux as a Limiting Step in Microbial Isoprenoid Production. *PLoS One* 7(11):e47513.

19. Wolff M, et al. (2003) Isoprenoid biosynthesis via the methylerythritol phosphate pathway: The (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from Escherichia coli is a [4Fe-4S] protein. *FEBS Lett* 541:115–120.

Chapter 2

Designing a novel isoprenoid biosynthesis pathway

Chapter adapted from (1)

2.1. Introduction

Both of the MVA and MEP pathways, i.e. the natural isoprenoid pathways, are complex, long and suffer from many limitations that must be overcome for them to perform optimally. The numerous attempts made to genetically engineer both natural pathways, highlight that the balancing of precursors, intermediates, catalytic efficiencies and energy demands is critical to achieving high fluxes, but also very complicated.

Precursor supply and energetics are particularly important. At a system level, both the MVA and the MEP pathways require precursors and cofactors for IPP synthesis from central carbon metabolism, therefore competing with other cellular processes for resources, which can complicate attempts to further increase isoprenoid pathway flux. Additionally, in the MEP pathway there is also the need to maintain balance in the supply of G3P and pyruvate, as imbalances can introduce limitations leading to decreased pathway performance (2). Previous research on the matter has tried to address this by engineering the central carbon metabolism, either through the Embden–Meyerhof–Parnas glycolytic pathway to increase G3P supply (3), by overexpressing the Entner-Doudoroff pathway (4) or using it in conjunction with the pentose phosphate pathway (PPP) (2), which can also lead to increases in NADPH regeneration rates and increased isoprenoid titers.

Another key objective for engineering a well-functioning and strong isoprenoid pathway is to prevent the accumulation of toxic intermediates or by-products and avoid each point of pathway

regulation leading to decrease in productivity. Pathway intermediates or downstream products have been shown to inhibit 'gate keeper' enzymes in both the MEP and the MVA pathways, with IPP inhibiting 1-deoxy-D-xylulose 5-phosphate (DXP) synthase (5); Coenzyme A, acetylacetyl-CoA and HMG-CoA inhibiting HMG-CoA synthase (6); HMG, free CoA and NAD(P)+/NADPH inhibiting HMG-CoA reductase (7, 8); and IPP, DMAPP, GPP, FPP inhibiting mevalonate kinase (9). Particular attention can be given to the iron-sulfur enzymes IspG and IspH, which are sensitive to oxygen (10) and whose inactivation leads to carbon loss, due to accumulation and excretion of metabolic intermediates, such as 2-C-methyl-D-erythritol 2,4-cyclodiphosphate (MEC) (11). The complex regulation both pathways face can significantly hinder attempts to up-regulate either.

Finally the maximum theoretical molar yields of both native pathways do not exceed 80% (12, 13), due to loss of carbon for cell growth maintenance requirements, while in experimental practice, observed yields have been much lower than the theoretical maxima (13), thus providing an additional incentive to engineer more efficient pathways.

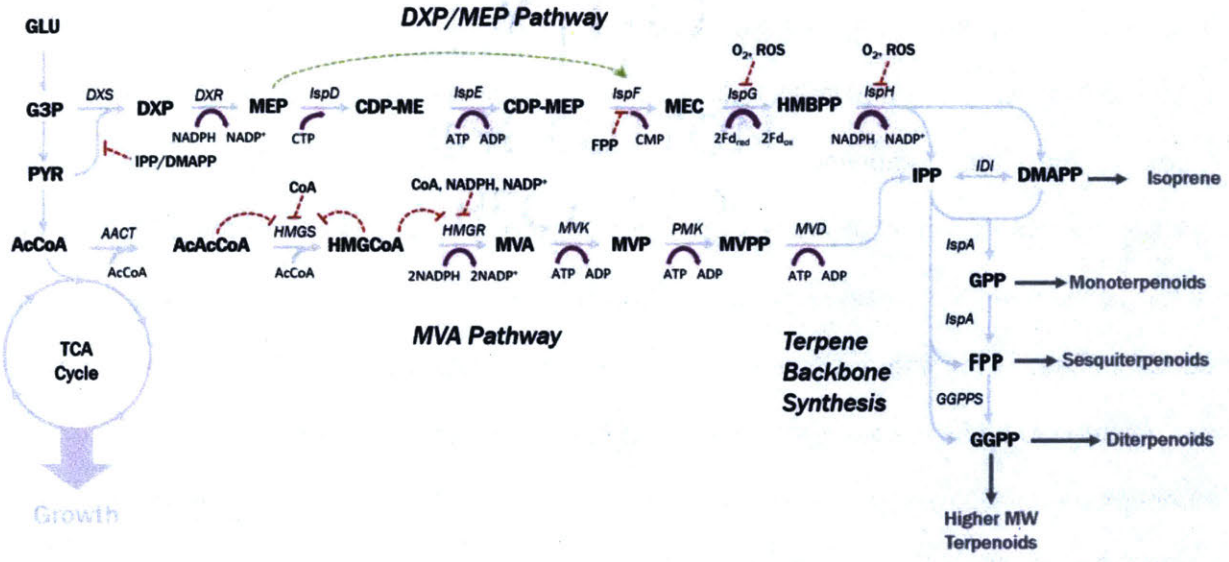


Figure 2-1: Overview of isoprenoid synthesis pathways via the MEP and MVA pathways. The MEP and MVA pathways start with metabolites produced through the central carbon metabolism. Substrates and intermediates are presented in bold and enzymes are presented in italics. Positive regulation is represented by a dashed green arrow, negative regulation by red dashed T lines. Cofactor consumption is represented by purple arrows. Figure adapted from (14).

2.2. Previous efforts at designing alternatives to the natural pathways

As mentioned above, the complex set of limitations the natural pathways face has presented many challenges in optimizing them for isoprenoid production. Thus, a number of attempts have been made to overcome these limitations, by exploring alternative approaches to the use of the endogenous pathways.

In the most straightforward approach, an entire isoprenoid pathway can be heterologously expressed in a host that does not naturally express it (e.g. introducing the MVA pathway in organisms which naturally produce their isoprenoids via the MEP pathway and vice versa), with the rationale being that this pathway will not be subject to the same regulation as the native pathways. This has been accomplished in *S. cerevisiae*, through the transfer of the MEP pathway, with mixed results as *S. cerevisiae* naturally lacks the enzymes required for the maturation of the MEP pathway enzymes IspG and IspH and is thus unable to exhibit performance superior to the native MVA pathway (15, 16). In contrast to the previous, the reverse feat, i.e. transferring the MVA pathway in *E. coli*, which naturally expresses the MEP pathway, has been achieved, and the pathway has been shown to work (17). However, the pathway still suffered from some of the same bottlenecks as when expressed in *S. cerevisiae*, and therefore still needed to be optimized, such as by modulating the levels of enzymes HMG-CoA (18) and MK (19).

In another avenue of approach, the limitations of precursor supply are sought to be alleviated through the exploration of alternative routes of supplying precursors, or different entry points to the MVA and MEP pathways. In the case of the MVA pathway, phosphoketolase (PK) and phosophotransacetylase (PTA) can be used to bypass the TCA cycle and produce Acetyl-CoA from fructose-6-phosphate (F6P) or xylulose-5-phosphate (X5P) (20), or, in another approach, better acetyl-coA production can be achieved through engineering of the pyruvate dehydrogenase bypass (21). In the case of the MEP pathway, alternatives that have been explored include the production of DXP, the first pathway intermediate from 1-deoxy-D-xylulose, which is ultimately

derived from xylulose (22), or the phosphorylation of 2-C-methylerithrytol for the production of MEP, the second intermediate of the pathway (23). Alternative ways for the supply of DXP also include its synthesis from ribulose-5-phosphate in 1 step (24), or from D-arabinose through the promiscuous activity of fructose-6-phosphate aldolase (FSA) and xylulose kinase (XK) (25).

While improving the precursor supply or reaching the entry of an isoprenoid pathway through different precursor routes is interesting, it has to be noted that this does not eliminate all of the downstream bottlenecks of either the MVA or MEP pathways. In order to circumvent the limitations of these downstream bottlenecks, alternative/variant isoprenoid pathways can be employed to bypass them. One such approach, would be the use of the archaeal MVA pathway, a variation on the common eukaryotic pathway, which produces IPP from mevalonate through mevalonate-3-prosphate, thus circumventing the known bottleneck MK (26). Another such approach has been employed by the Keasling group, in which a bypass pathway is used to produce a desired product, isoprenol, in either 1 step starting from mevalonate, or in 2 steps starting from mevalonate-5-phosphate and going through isopentenyl monophosphate (27).

While all the potential alternative/bypass isoprenoid biosynthesis pathways mentioned above manage to overcome some of the limitations of the native pathways, they all still end up suffering from one major limitation, namely the fact that they all rely on precursors that are "leeched" off Central Carbon Metabolism for their carbon needs, or substrates that the cells can also utilize for growth & maintenance, inextricably tying cell growth and product formation. We thus endeavored to design a novel pathway that is simple, efficient and most importantly, decoupled from growth.

2.3. The Isopentenol Utilization Pathway (IUP)

In order to create a novel isoprenoid pathway, we envisioned ways to retrosynthetically produce IPP through biotransformation of simpler molecules. For the first step "back", we looked for inspiration in "alternative" natural isoprenoid pathways. It is known (26) that certain archaea can

produce isoprenoids through a modified version of the MVA pathway, in which IPP is produced from IP (isopentenyl monophosphate) through the enzyme IPK (isopentenyl phosphate kinase). IP itself is produced from mevalonate-5-phosphate through the action of the enzyme mevalonate-5-phosphate decarboxylase (variant MVA pathway shown in Figure 1-2). We thus decided to include the phosphorylation reaction of IP to IPP in our pathway, which could be catalyzed by IPK. Of course, feeding IP in the cell would be infeasible both due to economic considerations, as well as due to the difficulty of getting phosphorylated compounds to pass the cell membrane (28). We therefore needed to take yet another step "back" and figure out a way to produce IP from a simpler compound. We envisioned the preceding step being the phosphorylation of isoprenol, to be catalyzed by a phosphokinase. Isoprenol is a chemical that is inexpensive, being an alcohol would be able to cross the membrane barrier and be utilized by the cell, thus being a feasible entry point for our pathway.

We thus conceived the Isopentenol Utilization Pathway, which would begin from isoprenol and lead to IPP through two consecutive phosphorylations or, it could alternatively start from prenol and lead to DMAPP after two phosphorylations. After IPP or DMAPP are produced, they can be interconverted through IDI, which all cells can naturally express.

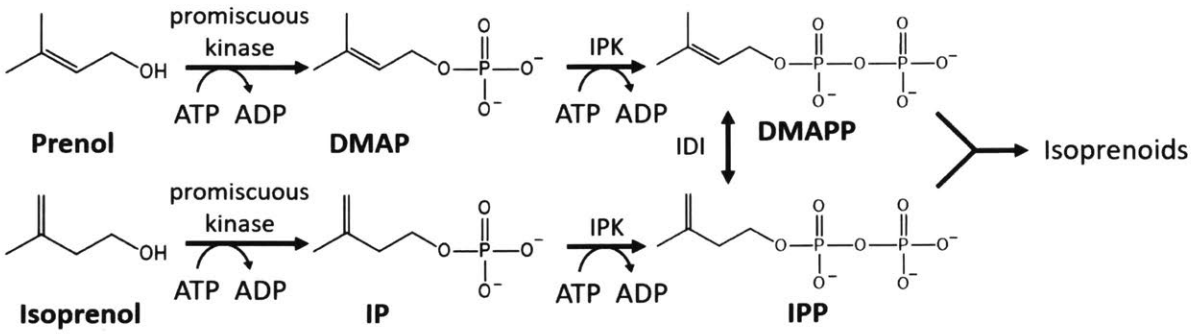


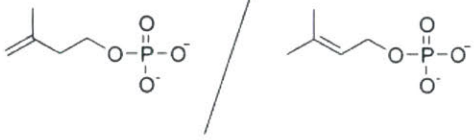
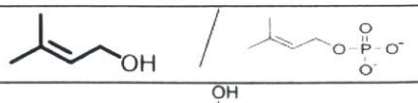
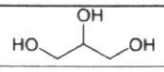
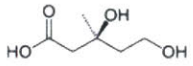
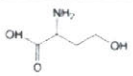
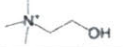
Figure 2-2: The Isopentenol Utilization Pathway (IUP). The IUP can produce the basic isoprenoid metabolic intermediates, IPP and DMAPP in two steps using isoprenol or prenol respectively as feedstock. The steps are catalyzed by a promiscuous kinase and isopentenyl phosphate kinase (IPK). IPP and DMAPP can then be interconverted via isopentenyl-diphosphate isomerase (idi). IPP and DMAPP act as the precursor molecules for larger prenyl diphosphates and eventually isoprenoids.

2.4. Identifying the enzyme catalyzing the first step of the IUP

2.4.1 Selection of candidate enzymes

The main challenge in realizing the IUP lay in the fact that the first phosphorylation step does not occur in nature. However, since it is well known that some phosphokinases can exhibit promiscuous activity (29, 30), we decided to investigate whether we could identify a phosphokinase which could catalyze the reaction of the first step. We selected nine phosphokinases based on the similarity of their natural substrates to our preferred substrate, with the rationale being that the closer their original substrate resembles isoprenol/prenol, the more likely these enzymes are to catalyze the phosphorylation of isoprenol/prenol, through non-specific binding of the molecule to the active site. Our selection of enzymes was also influenced by a recent report, which was published after we had started work in this pathway, that suggested that a triple mutant of the Isopentenyl Phosphate Kinase from *T. acidophilium* (TaIPK-3m) may be able to catalyze the first reaction step (31). The enzymes we selected for screening are shown on Table 2-1:

Table 2-1: Enzymes screened in this work.

Enzyme candidate	Source	Name	Native Substrate
Isopentenyl Phosphate Kinase	*H. volcanii*	HvIPK	
	M. thermoautotrophicus	MtIPK	
	M. janaschii	MjIPK	
	T. acidophilium	TaIPK	
	T. acidophilium – Triple mutant	TaIPK-3m	
Glycerol Kinase	*E. coli*	EcGK	
Mevalonate Kinase	*S. cerevisiae*	ScMK	
Homoserine Kinase	*E. coli*	EcHK	
Choline Kinase	*S. cerevisiae*	ScCK	

Codon-optimized homologs of the genes for the production of these enzymes were expressed in *E. coli* and the genes were subsequently purified, to be used in *in vitro* screens.

2.4.2 Quantification of IP/DMAP and IPP/DMAPP

In order to accurately assay the activity of each enzyme, we decided to measure the activity of each enzyme by directly measuring the concentrations of the product synthesized, via LC-MS/MS, either over time or at some particular timepoint. A possible alternative way of measuring enzyme activity would have been to do it through measuring ATP to ADP turnover, through a coupled assay, as was done in the study (31) that identified the *T. acidophilium* IPK triple mutant as a potential enzyme that can catalyze the first enzymatic step of the IUP. However, this way, although quicker, was not utilized, in order to avoid possible errors that would arise from non-specific hydrolysis of ATP to ADP.

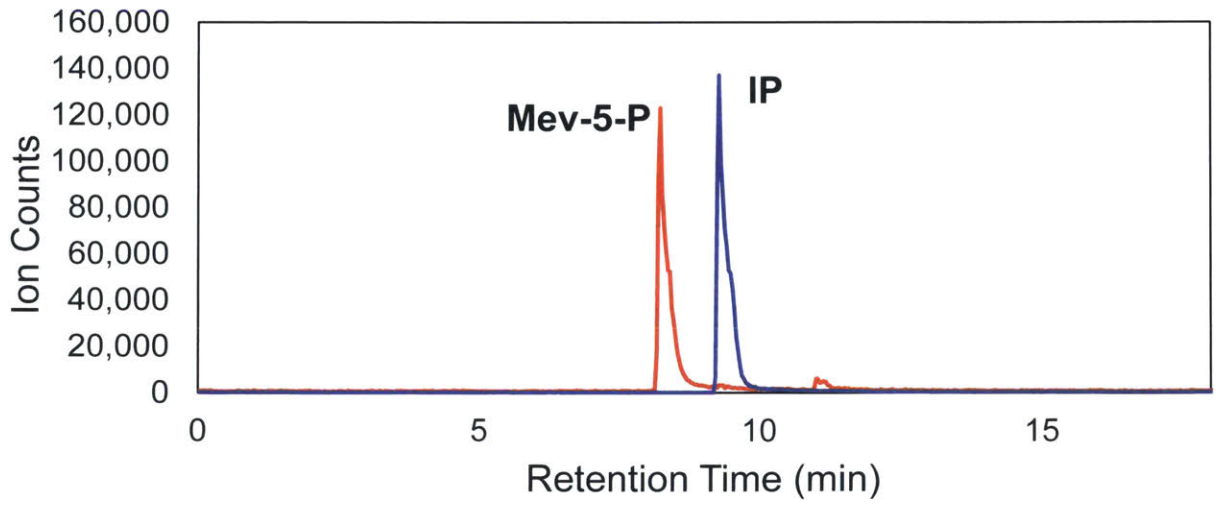


Figure 2-3: Enzymatic production of IP from Mev-5-P. Chromatogram showing the traces of the product (IP) and the substrate (Mev-5-P) of the *in vitro* enzymatic biotransformation catalyzed by mevalonate pyrophosphate decarboxylase.

In order to be able to accurately quantify the products of the two IUP reactions via LC-MS/MS, we needed to have standards for IP/DMAP (for the first step), and IPP/DMAPP (for the second step). While it was easy to acquire the latter two products from commercial vendors, we discovered that

purchasing IP or DMAP would be prohibitively expensive, as they are not commonly available. We thus synthesized our own stock of IP and DMAP.

As a first step, we endeavored to prove our ability to detect IP through LC-MS/MS by producing IP enzymatically from mevalonate-5-phosphate, as is done in archaea (26). In order to catalyze this reaction we used a purified mevalonate pyrophosphate decarboxylase (PMD) from *S. cerevisiae*, which catalyzes the conversion of mevalonate diphosphate to IPP, but can exhibit promiscuous activity. By adapting an assay from (32), we incubated purified PMD together with ATP and mevalonate-5-phosphate for 30min at 40°C and then measured the metabolites in LC-MS/MS. We were thus able to detect a new peak matching the expected IP mass spectrum.

As a second step, we needed to chemically synthesize IP and DMAPP in order to be able to quantify them and generate standard curves for LC-MS/MS. We adapted methods previously reported in literature (33, 34) to chemically synthesize both molecules. We confirmed their synthesis by both LC-MS/MS, by matching their chromatograms against the expected one (i.e. against Figure 2-3), as well as by their ^{1}H and ^{31}P NMR spectra (Figures 2-5, 2-6, 2-7 & 2-8).

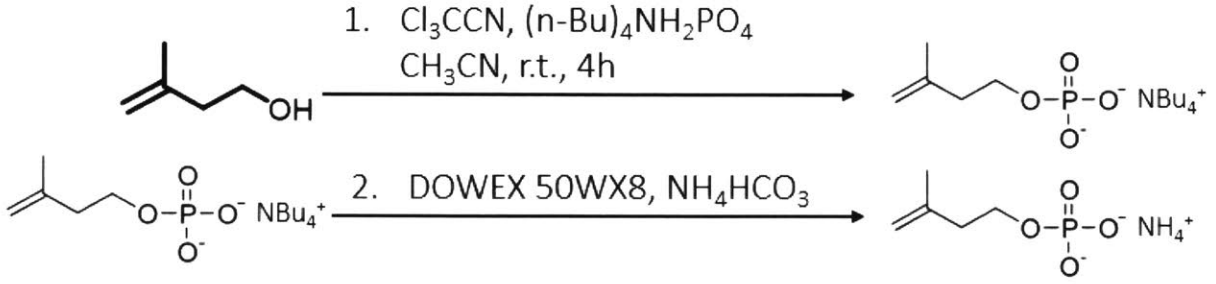


Figure 2-4: Reaction scheme for the chemical production of IP from isoprenol. A similar scheme is used for the production of DMAP from prenol.

Having thus acquired or produced stocks for IP/DMAP and IPP/DMAPP we were able to produce standard curves for the same. It needs to be mentioned that our LC-MS/MS lacks a column able to separate the IP and DMAP peaks, as well as the IPP and DMAPP peaks, therefore any IP/DMAP or IPP/DMAPP measurements in this work refer to the combined concentration of these metabolites.

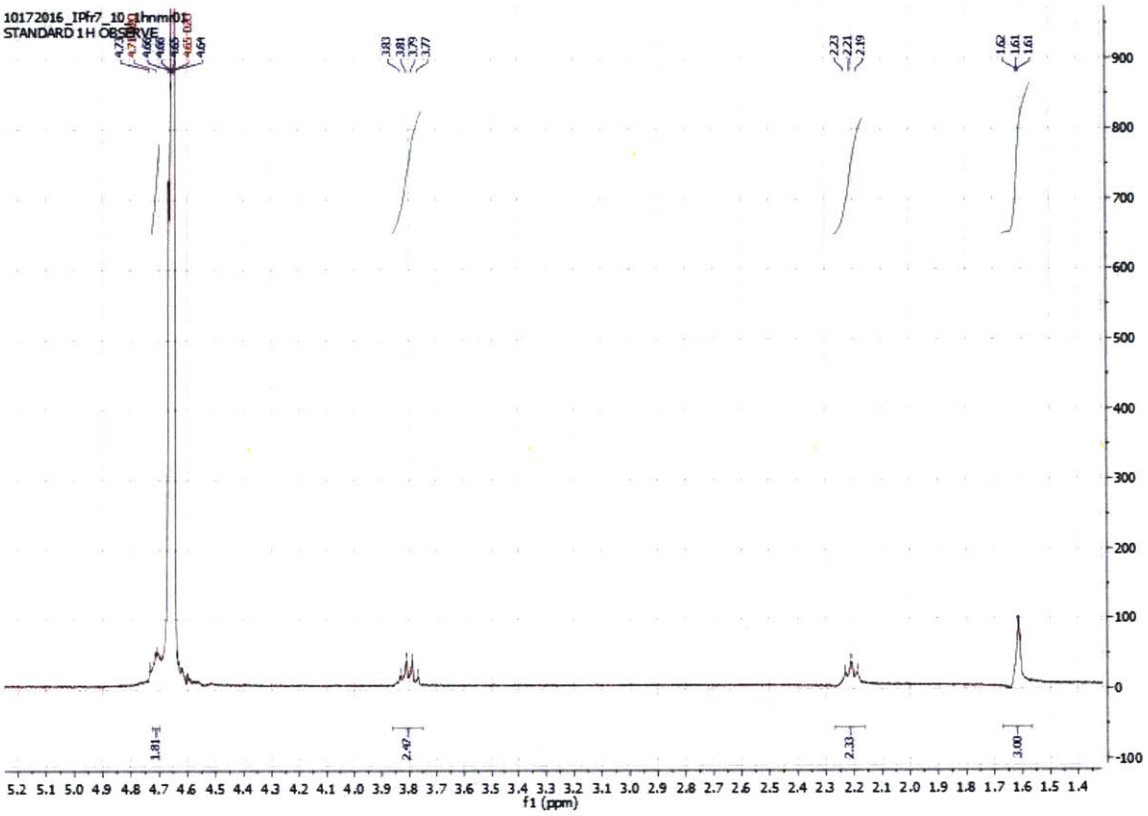


Figure 2-5: ^{1}H NMR of isopentenyl monophosphate (IP)

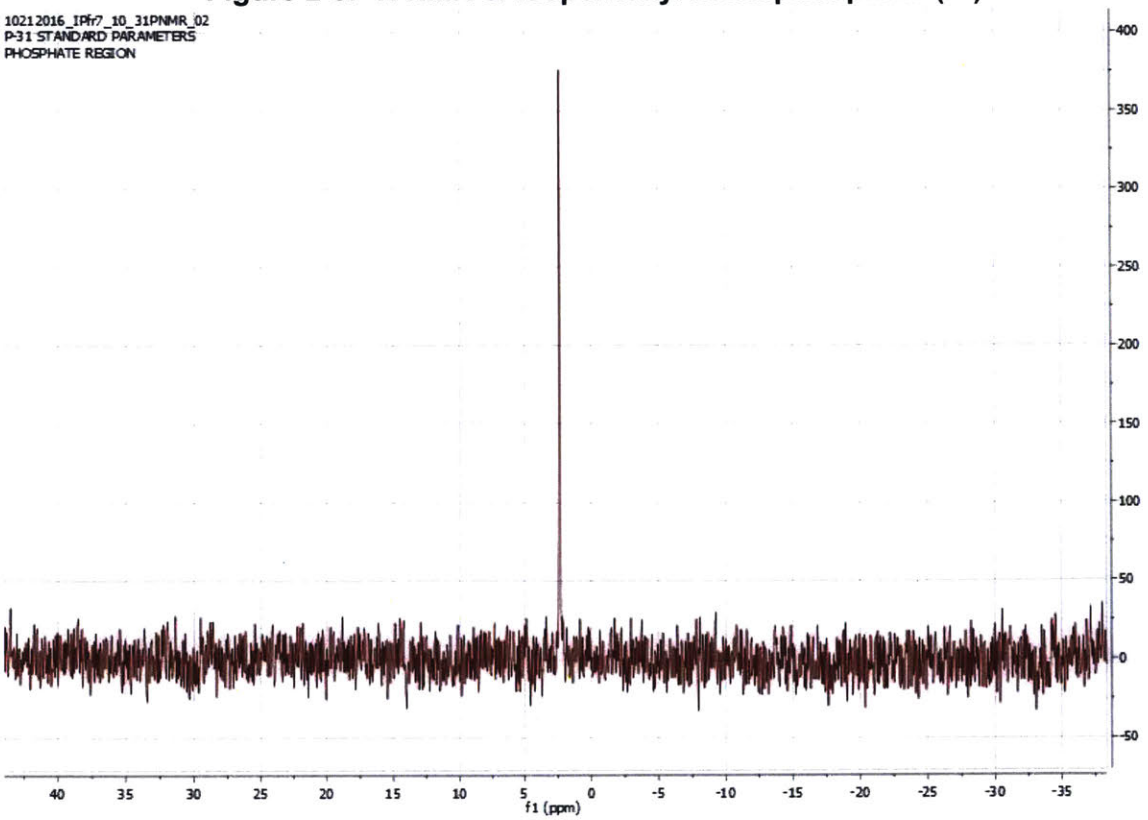


Figure 2-6: ^{31}P NMR of isopentenyl monophosphate (IP)

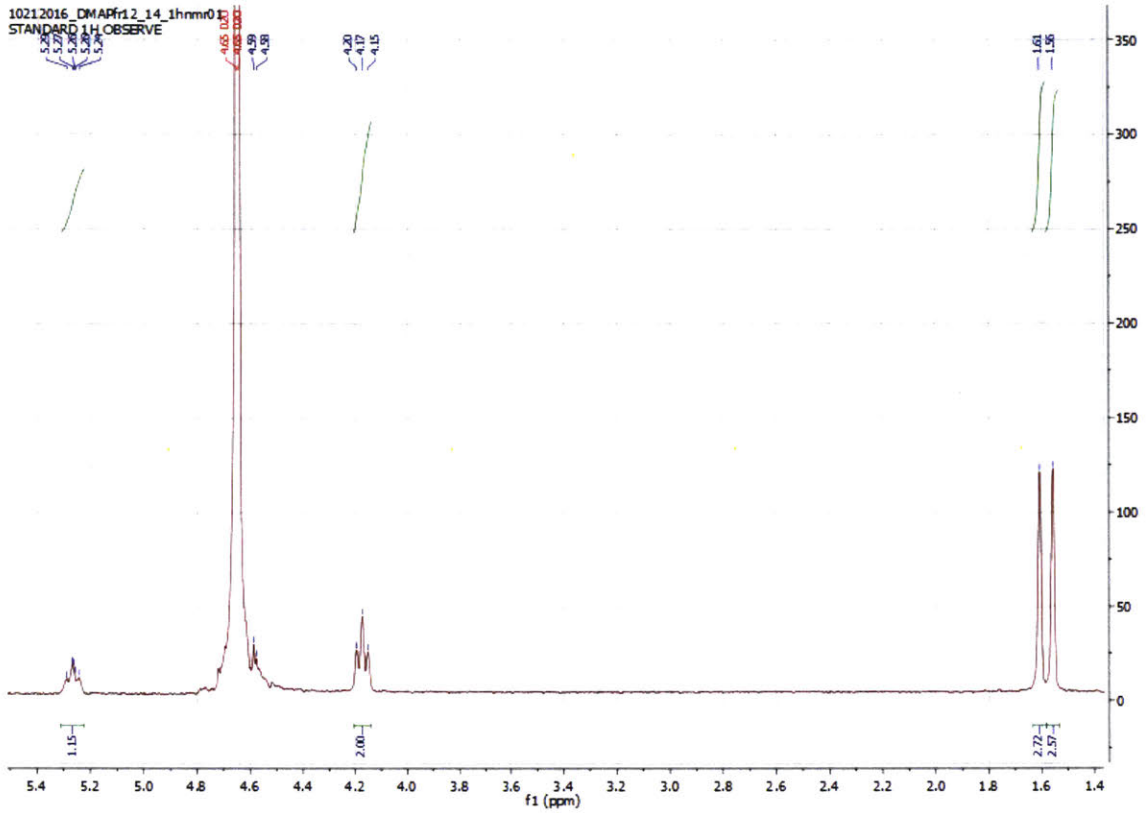


Figure 2-7: ^{1}H NMR of dimethylallyl monophosphate (DMAP)

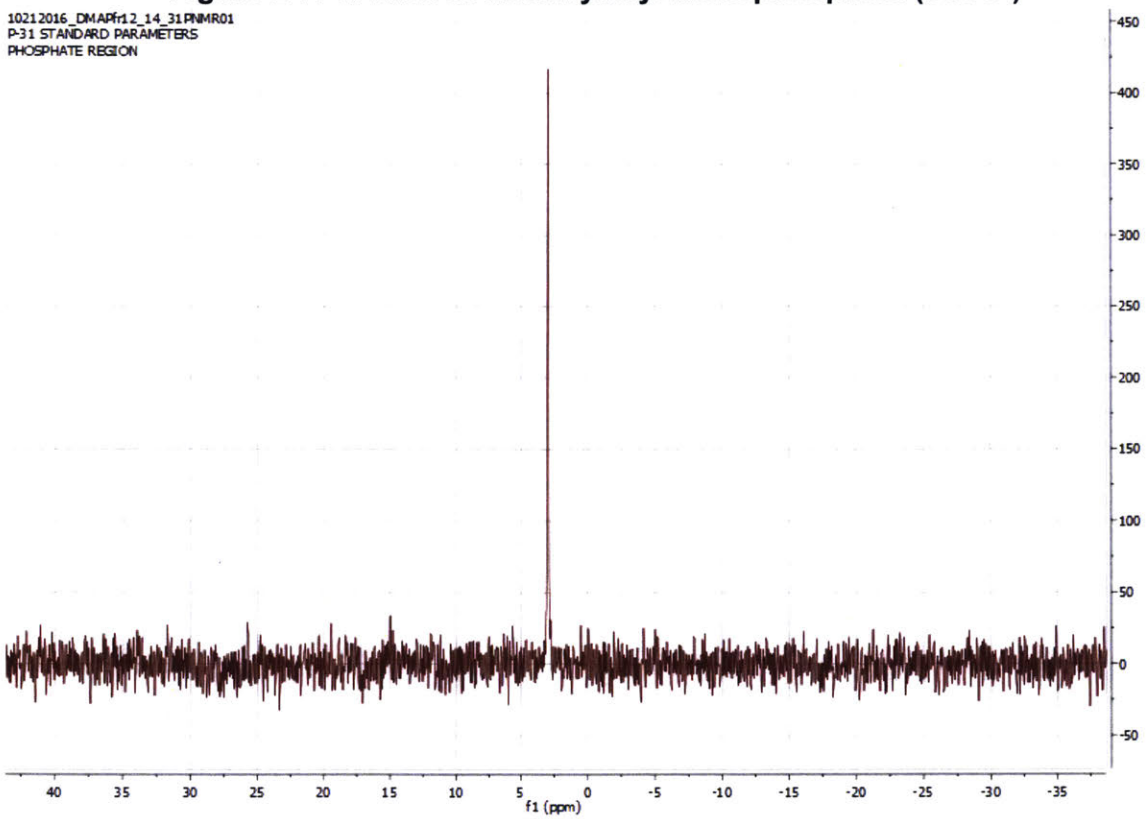


Figure 2-8: ^{31}P NMR of dimethylallyl monophosphate (DMAP)

2.4.3 Identifying the best candidate enzyme through *in vitro* assays

Having established our ability to quantify IP/DMAP, IPP/DMAP we were then able to quantify the activity of candidate enzymes. After incubating the purified enzymes (see "Enzyme expression and purification" and "*In vitro* enzyme assays" in Materials and Methods) overnight together with ATP and isoprenol or prenol, we quantified the products in LC-MS/MS.

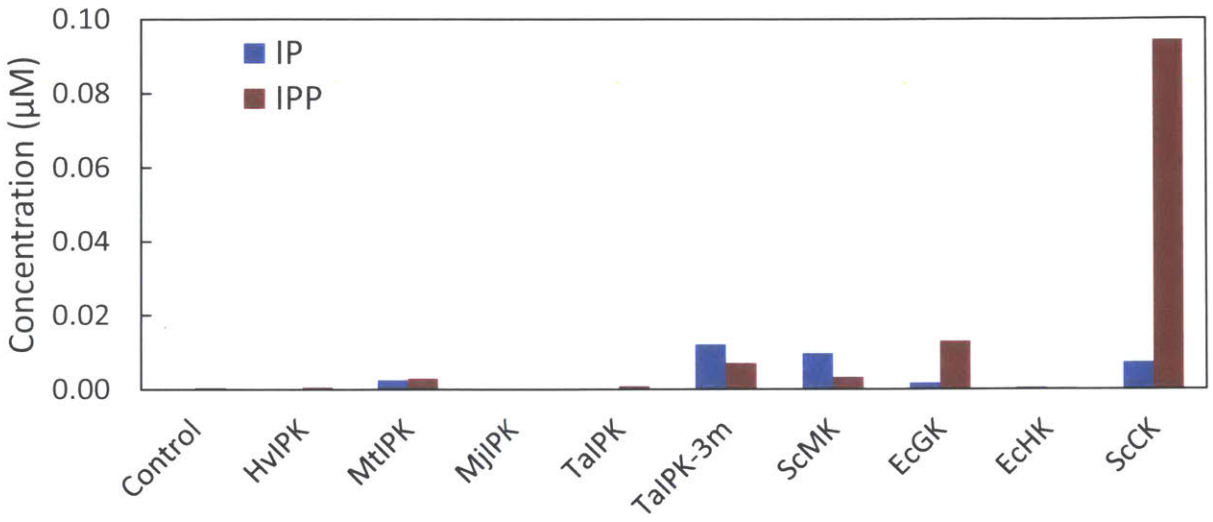


Figure 2-9: Results of *in vitro* overnight enzyme screen to identify a suitable promiscuous kinase converting substrate isoprenol to IP and IPP.

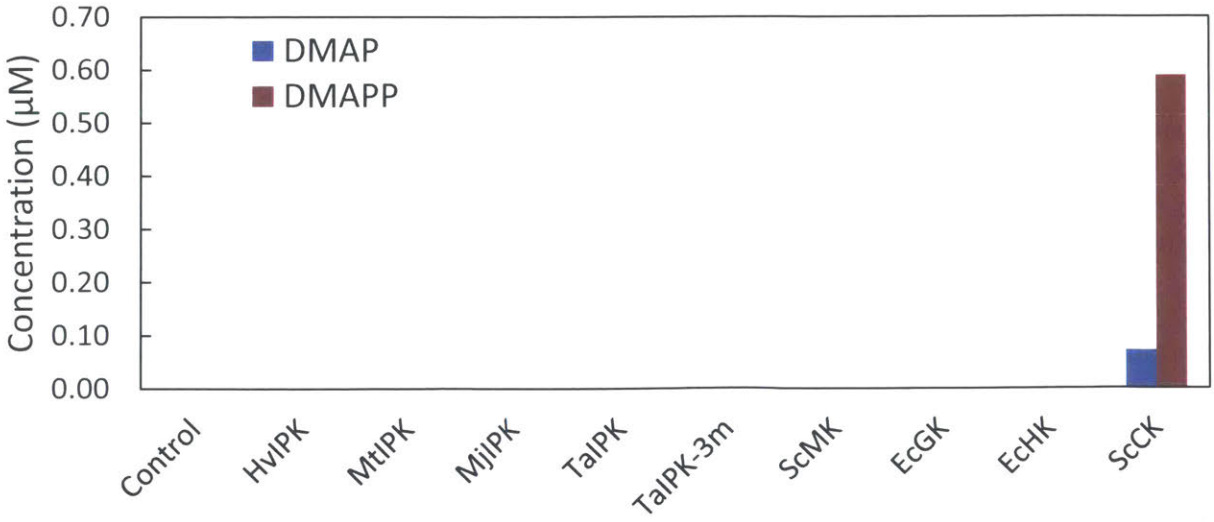


Figure 2-10: Results of *in vitro* overnight enzyme screen to identify a suitable promiscuous kinase converting substrate prenol to DMAP and DMAPP.

The overnight experiments indicated that many of the tested kinases were able to not only produce IP/DMAP, but also IPP/DMAPP, indicating that they could not only catalyze the first phosphorylation step of the IUP, but also the second step as well. While several enzymes could convert isoprenol to IP and even produce IPP as well, choline kinase from *S. cerevisiae* (ScCK) clearly outperformed them by producing considerable amounts of IP and large amounts of IPP, which could only have come from the phosphorylation of IP. In the case of prenol, only ScCK could produce any appreciable amounts of DMAP or DMAPP.

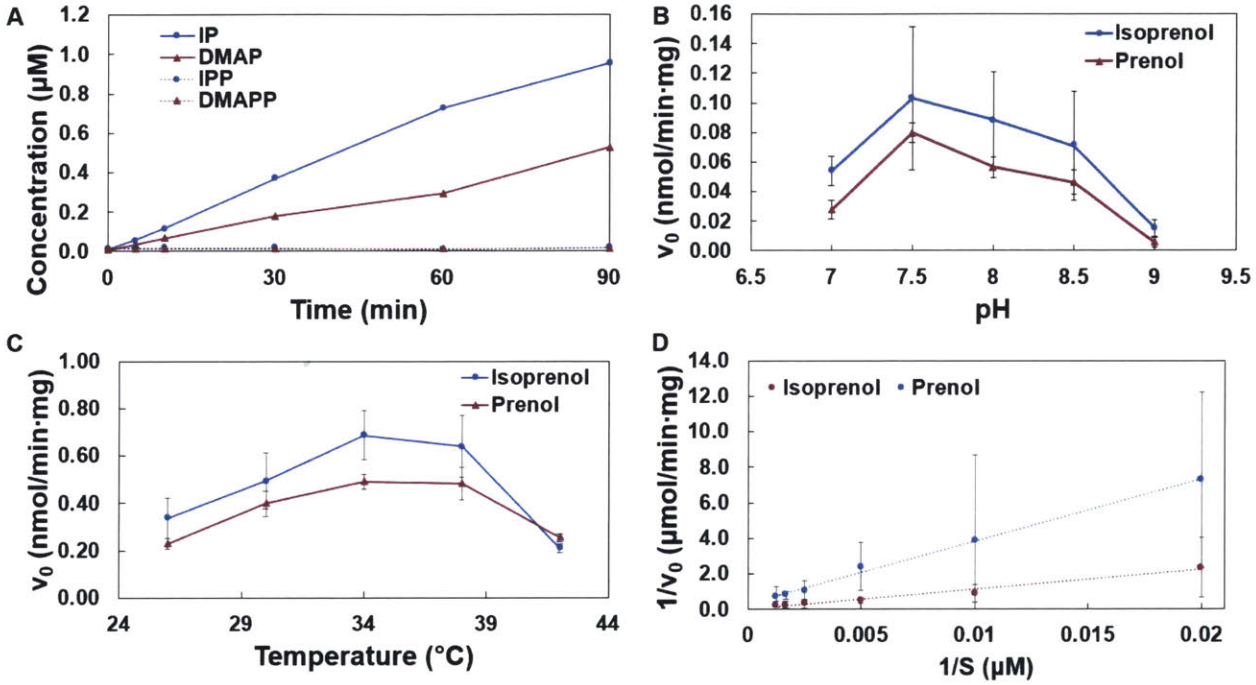


Figure 2-11: Kinetic analysis of choline kinase from *S. Cerevisiae* (ScCK) at a fixed ATP concentration. (**A**) Time-courses to determine the linear region of the enzyme, (**B**) Determination of the optimum pH, (**C**) Determination of the optimum temperature, (**D**) Determination of k_{cat} and K_M with regards to isopentenol substrate

From our data it became clear that the best candidate to catalyze the first step was ScCK. We thus proceeded to study ScCK's kinetics using the purified enzyme. We discovered that, although in longer timescales ScCK can produce more DMAPP (from prenol) than IPP (from isoprenol), in shorter timescales (~90 min), ScCK had a preference for isoprenol, indicated by concentrations of IP that were larger than those of IPP (Figure 2-9 A), unlike what was observed in longer

(overnight) timescales (Figure 2-7). Furthermore, in that timescale, we could detect very little IPP or DMAPP (Figure 2-9 A), suggesting a preference for the first step of the pathway. Our kinetic studies revealed that the optimal pH of ScCK is 7.5 and the optimal temperature is 34-38°C (Figure 2-9 B&C). The Michaelis-Menten constants (K_M) were identified as 4539 or 1113 μM and k_{cat} as 14.7 or 1.1 s^{-1} at 37°C when the substrate is isoprenol or prenol, respectively (Figure 2-9 D).

2.5. Completing the pathway

Since we had identified an enzyme, ScCK, that can catalyze the production of IP/DMAPP, as well as the production of IPP/DMAPP we could have used just one enzyme to catalyze both reactions. However, given that we had observed that in short timescales ScCK does not produce large quantities of IPP, i.e. its enzymatic activity in catalyzing the first phosphorylation step is larger than its activity in catalyzing the second phosphorylation step, it was decided to use two enzymes, with choline kinase catalyzing the first step and an IPK catalyzing the second step. To catalyze the second step of the pathway we chose IPK from *A. thaliana* (AtIPK), as it had the highest reported k_{cat}/K_M (26). Finally, we included an optional IDI, in order to better balance the ratio of IPP and DMAPP. The complete Isopentenol Utilization Pathway is thus composed of the enzymes ScCK, AtIPK and IDI and requires ATP as its sole cofactor.

2.6. *In vivo* proof of concept in *E. coli*

2.6.1 Expression of the pathway *in vivo*

Having identified the components of the pathway, we constructed two plasmids for its expression *in vivo* in *E. coli*. The first plasmid, plasmid pSEVA228-pro4IUPi (herein called "pro4IUP") uses the Standard European Vector Architecture (35) and contains the IUP operon under the control of the constitutive promoter P_{pro4} (36). The second, pTET-IUPi, herein called "pTETIUP", places

control of the IUP under the strong inducible anhydrotetracycline promoter (P_{TET}) (37). The difference in strength of these expression systems has been estimated by green fluorescent protein expression (Figure 2-10).

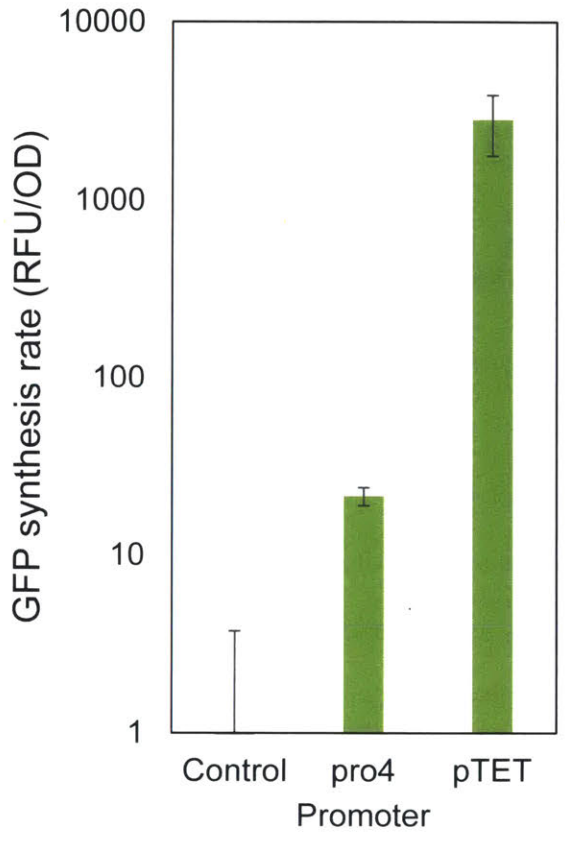


Figure 2-12: Initial characterization of expression system strength. Characterization of expression plasmids using GFP as a reporter protein, with either the pSEVA228pro4-gfp plasmid (pro4), the pTET-gfp plasmid (pTET) induced with 10 ng/mL anhydrotetracycline (aTC) or no plasmid (Control)

2.6.2 Initial proof of concept

In order to determine if the pathway is functioning *in vivo*, we expressed the IUP in *E. coli* using the pTETIUP plasmid, which carries an inducible promoter. Sine the IUP would be producing the same products (i.e. IPP and DMAPP) that the MEP pathway is naturally producing, the need arose to differentiate between the IPP/DMAPP produced by the IUP and the IPP/DMAPP produced by IUP. To this end, we grew *E. coli* on ^{13}C uniformly labeled glucose so that the IPP

produced by the MEP pathway would also be uniformly labeled. Upon reaching the start of the exponential phase we supplemented the medium with unlabeled isoprenol/prenol, which, through the IUP, produced unlabeled IP/DMAP and unlabeled IPP/DMAPP.

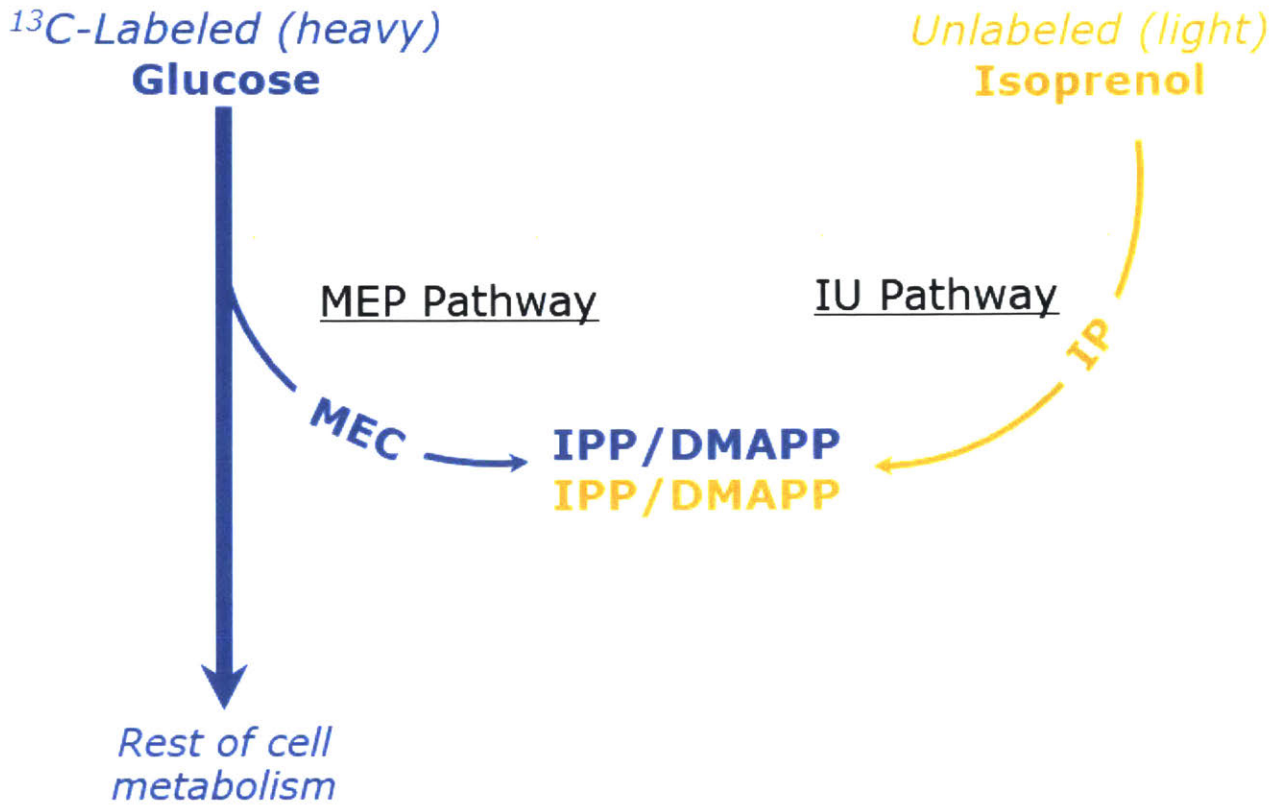


Figure 2-13: Schematic of the labeling experiment. *E. coli* expressing the IUP are grown on ^{13}C labeled glucose and unlabeled isoprenol is spiked in the media during early stationary phase. The glucose is utilized by the MEP pathway to produce labeled IPP/DMAPP, whereas the isoprenol is utilized by the Isopentenol Utilization Pathway to produce unlabeled IPP/DMAPP. We can thus distinguish IPP/DMAPP produced by each pathway.

After supplementing with isoprenol, prenol or no isopentenol and growing our cultures for two days, we extracted the metabolites from the cultures and measured their concentrations through LC-MS/MS (see "Labeling experiment" in Materials and Methods). We were able to detect quantities of unlabeled IPP (and small quantities of unlabeled IP), indicating that the IUP was indeed functional, thus giving us a first proof of concept. We note that in some cases when the IU pathway was not "on" (either due to absence of inducer or isoprenol/prenol) we did detect small

quantities of unlabeled IPP. We theorize that this was due to the presence of (unlabeled) thiamine that was present in the media, that was taken up as a carbon source and converted to IPP though the native MEP pathway. We did not observe any labeled IPP or MEP pathway intermediates.

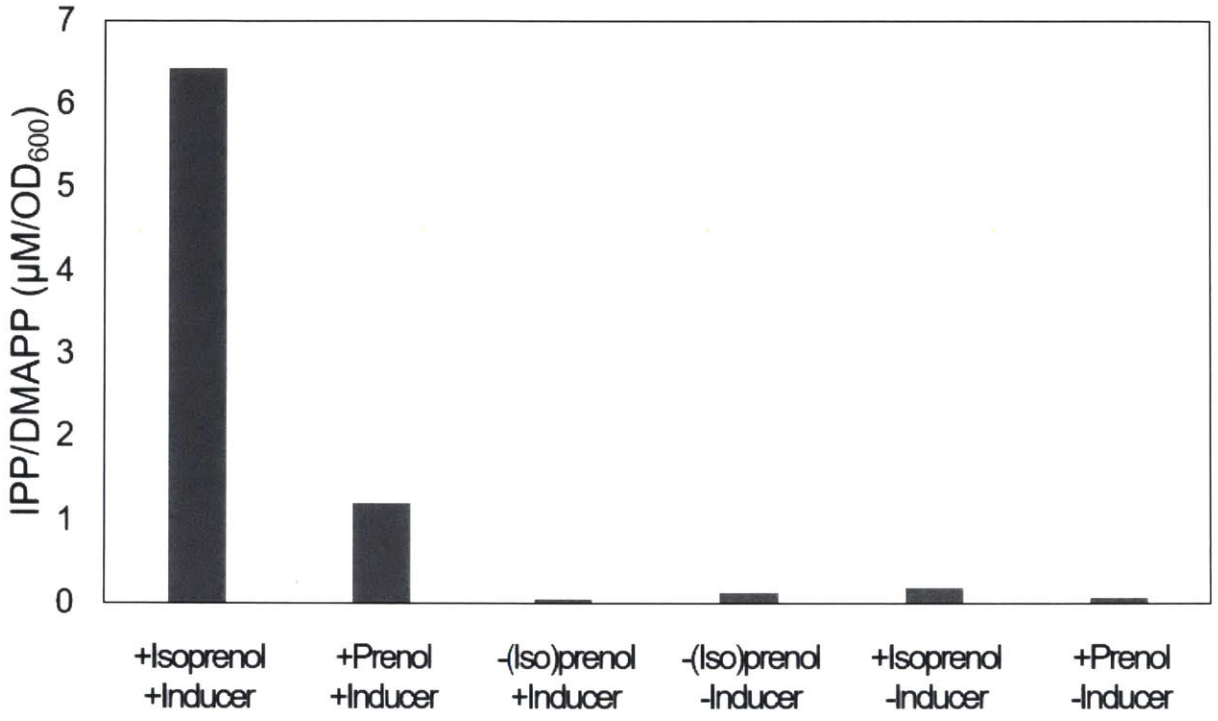


Figure 2-14: *In vivo* proof of concept of the IUP. *E. coli* expressing the IUP are grown on ^{13}C labeled glucose and unlabeled isoprenol is spiked in the media. Shown are the quantities of unlabeled IPP/DMAPP in the presence or absence of isoprenol/prenol and in the presence or absence of the inducer for IUP expression.

2.6.3 IUP is viable as the sole isoprenoid pathway

Having gotten an initial proof of concept, we tested the ability of the IUP to rescue an MEP-knockout strain incapable of producing isoprenoids via its native MEP pathway. This strain would have otherwise been non-viable because to its inability to produce isoprenoids through the MEP pathway, since isoprenoids are necessary for cell survival because they perform essential cellular functions, such as maintenance of membrane fluidity and electron transport (38). Since the MEP pathway is a linear one, knocking out any of the MEP pathway genes would inactivate the entire pathway, and would result in lethality. We decided to knock out the *ispG* gene.

While we were carrying out our knockout experiment, we wanted to make sure that the cells maintained viability regardless of whether the IUP could function *in vivo* and rescue viability. To that end, we would first introduce the lower mevalonate pathway in our strains, which is known to be able to rescue viability in strains in which the MEP has been inactivated (17). The lower mevalonate pathway consists of enzymes MK (mevalonate-5-kinase), PMK (phosphomevalonate kinase) and PMD (mevalonate pyrophosphate decarboxylase), encoded by genes *erg12*, *erg8* and *mvd1* respectively. Thus, before performing our knockout, we transformed our *E. coli* strains with plasmid pBAD33-proA-MEVI through which the lower mevalonate genes are constitutively expressed. While the lower mevalonate pathways genes are always expressed, the pathway is only functioning at when mevalonate is extraneously supplied to the media. The resulting strain therefore carried a "backup" rescue mechanism we could turn on/off at will by supplying (or not supplying) mevalonate.

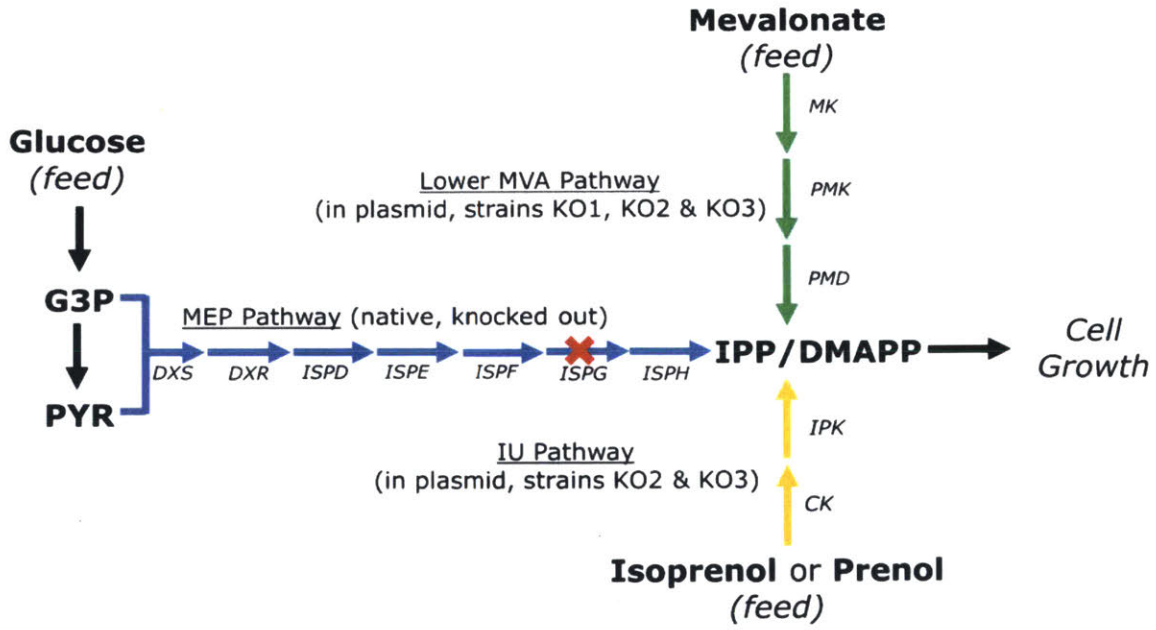


Figure 2-15: Isoprenoid pathways of the MEP knockout strains KO1, KO2 and KO3. In these strains, the native MEP pathway has been knocked out through a deletion in *ISPG*, rendering them unable to produce isoprenoids required for their growth. Growth in strains KO1, KO2 and KO3 can be recovered through the lower MVA pathway, by supplementing the media with mevalonate. In strains KO2 and KO3, growth can be recovered by using the IUP to produce isoprenoids from isoprenol or prenol feed.

Using the CRISPR-Cas9 system (17, 39), we created an MEP-knockout strain (strain **KO1**) carrying the lower mevalonate pathway (see "Knockout of the native MEP pathway" in Materials and Methods). When grown in minimal media, this strain was viable when 1mM mevalonate was supplied to the culture medium, and non-viable when it was not, as we had intended. Furthermore, we noticed that supplying isoprenol or prenol failed to rescue the strain, indicating that neither the lower mevalonate pathway nor some endogenous pathway could convert isoprenol or prenol to IPP.

We then transformed strain KO1 with either plasmid pro4IUP or plasmid pTETIUP, creating strains **KO2** and **KO3** respectively. These strains were able to grow when supplemented with 0.6mM isoprenol (in the absence of mevalonate supplementation) with a lag phase lasting 36 h, compared to the lag phase of 2 h exhibited by wild-type *E. coli*. The exponential growth rates were also slightly lower than that of wild type *E. coli*.

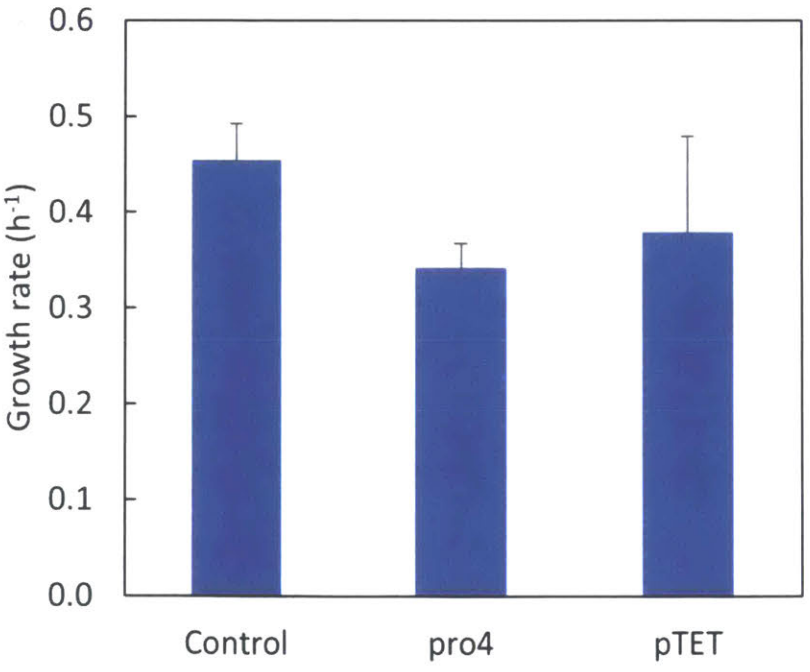


Figure 2-16: Growth rates of the *ΔispG* MEP knockout strains. The control, wild-type MG1655 (DE3), strain **KO2** containing the pro4IUP and **KO3** containing the pTETIUP were grown in M9 minimal media. The culture media was supplemented with 25 mM isoprenol. pTET cultures were induced with 10 ng/mL of aTC.

This was our final *in vivo* proof of concept, which showed that the IUP is a viable pathway in vivo, in that it is strong enough to be able to fully replace the native MEP pathway and can thus act as a full alternative isoprenoid pathway.

2.7. Conclusions

Based on our results, it is evident that a novel pathway for the production of isoprenoids, one that does not rely on simple modifications of the natural isoprenoid pathways is feasible. Our efforts and have resulted in a pathway, the IUP, that can function *in vivo*. Moreover, the IUP was only made possible through the promiscuity of *S. cereviviae* choline kinase (ScCK). As future steps, additional phosphokinases could be screened to identify possible candidate enzymes with higher activity than ScCK.

The overall concept behind this pathway, i.e. the production of a class of chemicals via the biotransformation of a simpler molecule that is externally provided rather than produced by the cell itself, could be expanded to circumvent the limitations of other biosynthetic pathways. The promiscuity of enzymes that are not related to said pathways but act on similar precursors could be exploited to provide for novel alternative biosynthetic pathways. Important factors in choosing whether these alternative pathways are preferable to engineering the native pathways will be the cost of the externally-supplied molecule compared to glucose, or whichever carbon source cells consume, as well as the potential gains in productivity achievable by such bypass pathways.

2.8. Materials and Methods

Strains, plasmids and genes.

E. coli K12 MG1655(DE3) was used as the parent strain for all metabolic pathway expression studies, DH5α (New England Biolabs-NEB) was used for routine cloning purposes and BL21 (DE3) (NEB) was used for the expression of proteins for purification. Strain genotypes and

plasmids used as templates for the construction of the IUP vector and the downstream vectors are listed in **Table A1** in the Appendix. Genes listed in **Table A2** in the Appendix, were custom synthesized, and codon optimized for *E. coli* MG1655 (Integrated DNA Technologies-IDT) where indicated, otherwise they were amplified from an existing plasmid or from genomic DNA. Genomic DNA was purified using the Wizard Genomic DNA Purification Kit (Promega Corporation).

Cloning Protocol.

A standard protocol was used for the cloning of all plasmids described in this work. Primers were designed for Gibson Assembly using the NEBuilder online tool (NEB), and primers were purchased from Sigma-Aldrich. PCR reactions were performed in a Bio-rad C1000 Touch Dual Block thermocycler using 2x Q5 polymerase master mix (NEB) according to manufacturer's recommendations. The products were digested with DpnI (NEB) enzyme for 1h at 37°C to digest the template DNA. The digested PCR products were then run on a 0.8% agarose gel using a Mini or Sub Cell and a Powerpac Basic power supply (Bio-rad). Fragments were gel-extracted using a Zymoclean Gel DNA recovery kit (Zymo Research) according to the manufacturer's recommendation. Fragments were ligated using the Gibson Assembly Master Mix (NEB) for 1h at 50°C and transformed into DH5α (NEB) high efficiency chemical competent cells (NEB) according to standard protocol. Transformants were screened by PCR using colonies boiled in water for 10 min as the template. Two to three positive transformants were cultured overnight in LB media and the plasmid was purified using a Mini-prep kit (Qiagen). Overlapping regions of the new construct were sequenced (Quintara Biosciences, Boston) to confirm the sequence of the new plasmid. Plasmids with confirmed sequences for protein purification were transformed into BL21(DE3) using heat-shock, otherwise plasmids were transformed into MG1655(DE3) for further study by electroporation using a MicroPulser (Bio-rad). Electrocompetent cells were made by a standard glycerol washing of mid-log phase cells (Bio-rad) and stored at -80°C until future use.

For electroporation, 1 µL of purified plasmid in water was added to 50 µL of electrocompetent cells using 1. kV and electroporated in 1 mm path-length cuvettes (Bulldog Bio).

Construction of expression vectors.

All vectors were constructed following the routine cloning protocol described above. Vectors for enzyme expression were constructed by introducing a His-tagged gene (for PMD, ScCK, ScMK, EcGK, EcHK, HvIPK, MtIPK, MjIPK, TaIPK or TaIPK-3m expression) into a pET-28 a (+) vector in the following fashion: Backbone fragments were amplified from pET-28 a (+) vector using the primer pair GB pET28-HisT-vec_f/r. The insert fragments were amplified as follows: the fragments containing the genes for PMD, ScCK or ScMK expression were amplified from S. cerevisiae genomic DNA using the primer pairs GB-pET28-PMD_f/r, GB-pET28-CHOLKIN_f/r or GB-pET28-MEVKIN_f/r respectively, the fragments containing the gene for EcGK or EcHK expression were amplified from *E. coli* genomic DNA using the primer pairs GB-pET28-GLYCKIN_f/r or GB-pET28-HSERKIN_f/r respectively, whereas the fragments the genes for HvIPK, MtIPK, MjIPK, TaIPK or TaIPK-3m expression were amplified from custom synthesized, and codon optimized DNA using the primer pairs GB-pET28-HV_f/r, GB-pET28-MTH_f/r, GB-pET28-MJ_f/r, GB-pET28-THA_f/r, or GB-pET28-THA3m_f/r respectively. The backbone and insert fragments were then assembled to give the respective plasmids. Sequences for the primers used are listed in **Table A3** in the Appendix.

Vectors for the expression of the IUP were constructed by incorporating an IUP operon consisting of the genes *Scck*, *ipk* and *idi* into either a pSEVA228 backbone under the control of the pro4 constitutive promoter (36) for the creation of plasmid pSEVA228-pro4IUPi, or into a pTET backbone (derivative of pET-28a carrying the androtetracycline repressor and promoter region of pBbS2k-rfp) for the creation of plasmid pTET-IUPi. The plasmid pSEVA228-proDIUPi was generated by amplifying the backbone pSEVA228 with the primers GB-SEVA228_f/r and inserting a custom-synthesized promoter sequence, shown in **Table A4**, which incorporates the proD

promoter system (36), which was amplified using the primer pairs GB-proD_f/r, along with the IUP operon, consisting of the genes *Scck*, *ipk* and *idi*. Each of the three operon elements was amplified from custom synthesized DNA fragments (IDT) using the primer pairs GB-chk_f/r, GB-atipk_f/r and GB-iditerm_f/r respectively. In all three cases, the codon-optimized gene coding sequence was preceded by a corresponding optimized ribosomal binding site (RBS) (shown in **Table A4** in the Appendix) and in the case of idi, it was followed by a T7 terminator derived from pET-28(+) (shown in **Table A4** in the Appendix). The RBSs were optimized using the Salis lab RBS optimization tool (40, 41). The PCR fragments were assembled to give pSEVA228-proDIUPi. The plasmid pSEVA228-pro4IUPi was created by replacing the 6 nucleotides in the proD promoter sequence of pSEVA228-proDIUPi with the primer pairs GB-pro4_Mut_f/GB-ProLibrary_Mut_r to amplify the whole plasmid and then subsequently assembling the amplification product. The plasmid pSEVA228pro4-ck-idi, which carries a reduced version of the IUP operon, lacking *ipk*, was created by amplifying pSEVA228-pro4IUPi using the primer pair GB-IUPnoIPK_f/r and then subsequently assembling the amplification product. The pTET-IUPi plasmid was created as follows: The pTET backbone was created by replacing the T7/lac promoter region of pET-28a carrying a methanol utilization operon (pETMeOH500) with the androtetracycline repressor and promoter region of pBbS2k-rfp by Gibson assembly using primers GB-pETMeOH500-f/r and GB-pBbS2k-rfp-f/r. Then, the methanol utilization operon was replaced with the IUP operon from pSEVA228-pro4IUPi by Gibson assembly using the primers pTet-IUP-ins_f/r, and pTet-IUP-ins_f/r, to produce plasmid pTET-IUPi. pSEVA228 was a gift from Jason King. pETMeOH500 was a gift from Benjamin Woolston. pBbS2k-RFP was a gift from Jay Keasling (Addgene plasmid # 35330). Sequences for the primers used are listed in **Table A3** in the Appendix.

Enzyme expression and purification.

BL21 harboring a pET-28 vector for the expression of proteins in Fig. 1B was revived from a glycerol stock by inoculating into LB media and growing at 37°C overnight. Two hundred milliliters

of SOB media (Amresco) in a baffled 1L flask was inoculated at 1% with the overnight culture and grown until an OD of 0.5 at 30°C. The culturere then induced with IPTG at a final concentration of 100 uM. Cultures were incubated for 3-4 h at 30°C for protein synthesis after which they were centrifuged in an Allegra X12R centrifuge (Beckman-Coulter) at 3273 x g for 15 min. The supernatant was removed, and the cell pellets were stored at -20°C until purification. Proteins were purified using the following protocol and at all stages proteins were keep on ice. First, cells were lysed using 5 mL of ice-cold NPI-10 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) using a gas driven high-pressure homogenizer, the EmulsiFlex-C5 (Avestin). After disruption, 100 µM PMSF was added to the lysate. The lysate was centrifuged at 15 000 x *g* for 10 min at 4°C. Ni-NTA resin purchased from Qiagen was equilibrated using 10 column volumes (CV) of NPI-10 buffer in gravity column (Fisher-Scientific). The clarified supernatant was loaded to the column and allowed to drip through by gravity. After all of the lysate was loaded, the column was washed with 10 CV of NPI-20 buffer (20 mM imidazole). Then the protein was eluted from the column using 3 CV of NPI-250 buffer (250 mM imidazole). Protein purification was confirmed by protein gel electrophoresis using a Mini-protean system (Bio-rad) using precast 4-20% acrylamide gels (Bio-rad), Kalidescope Prestained Protein Ladder (Bio-rad), and Tris-glycine, SDS buffer (Bio-rad) at 200V for 20 min. Gels were stained with InstantBlue (Expedeon).

***In vitro* enzyme assays.**

Enzyme assays for screening of isopentenol kinase activity were adapted from (32). First, enzymes were expressed and purified as described above. The purified enzymes were added to the enzyme assay master mix for a final concentration of 2 mM ATP, 10 mM $MgCl_2$, 50 mM NH_4HCO_3 pH 7.5, and 600 µM isoprenol or prenol. They were incubated overnight at 37°C. In the case of the reaction for the enzymatic production of IP through PMD, 625µM of mevalonate-5-phosphate were used instead of isoprenol/prenol and the reaction temperature was 40°C. The reactions were stopped using 5 volumes of ice-cold acetonitrile and centrifuged in an Allegra

X12R centrifuge (Beckman-Coulter) to remove precipitated proteins using a plate adaptor at 3273 x g for 15 min. The supernatant was transferred to a new microplate and frozen at -80°C. The liquid was removed by a 4.5L lyophilizer (Labconco) and the samples were resuspended in an equal volume of water and centrifuged again prior to quantification using LC-MS/MS using a protocol described in the SI Appendix. Kinetic enzyme assays were conducted using the standard assay conditions described above with the following changes. First, the linear range of the assay was determined over a 90 min period. Initial velocity appeared linear over this period. Therefore, subsequent kinetic assays were quenched after 30 min. For assays at different temperatures, the standard reaction mixture was used while the temperature was varied using a water bath. For pH optimum, enzymes were buffer-exchanged into 50mM Tris-HCl at the appropriate pH using 10 kDa nanoseps (Millipore) by exchanging the buffer 5 times, resulting in a dilution factor of over 10000. Afterwards, reactions were performed in the standard reaction mixture with NH_4HCO_3 buffer adjusted to the appropriate pH. For Michaelis-Menten kinetics, only the concentration of isoprenol or prenol was varied between 1.5 – 50 μM.

Synthesis of IP and DMAP.

Isopentenyl monophosphate (IP) and dimethylallyl monophosphate (DMAP) were chemically synthesized using isoprenol or prenol respectively as they are not commercially available. The synthesis process was adapted from (33, 34). All chemicals and solvents were used as supplied without further purification. Both new compounds gave satisfactory spectroscopic and/or analytical data. The structures of the resulting IP and DMAP were confirmed by ^{1}H and ^{31}P-NMR. Spectra were recorded on a Varian Mercury-300 NMR Spectrometer in deuterated water (Sigma-Aldrich) at 300 MHz and chemical shifts (δ) are reported in parts per million (ppm) downfield from the internal standard, tetramethylsilane (TMS).

General Method: Trichloroacetonitrile (2.26 equiv.) and tetrabutylammonium phosphate (1.66 equiv.) were added to a solution of the appropriate isopentenol (1 equiv.) in acetonitrile and the

mixture was stirred for 4h. The solvent was evaporated, and the sample was resuspended in water and cooled for 6h at 4°C. This caused the crystallization and precipitation of the trichloroamide (white crystals), which was removed by filtration. The resulting solution was purified by flash chromatography using a mixture of isopropanol/NH_4OH/H_2O (7:2:1) as the eluent and the product was converted to the correspoinding ammonium salt by using a DOWEX 50WX8 ion-exchange column by percolation using NH_4HCO_3 (0.025 M) as the buffer solution. The resulting solution was lyophilized to recover the product as a solid by lyophilization.

Isopentenyl monophosphate (IP): Prepared following the General Procedure, starting from isoprenol (80 µL, 0.53 mmol), trichloroacetonitrile (120 µL, 1.2 mmol) and tetrabutylammonium phosphate (300 mg, 0.88 mmol) in 2mL acetonitrile. After work-up, IP was obtained as a white solid. The resulting spectra for IP were ^{1}H NMR (300MHz, D_2O): δ: 4.71 (s, 1H), 3.80 (q, 2H), 2.21 (t, 2H), 1.61 (s, 3H) and ^{31}P NMR (300MHz, D_2O): δ 2.38.

Dimethylallyl monophosphate (DMAP): Prepared following the General Procedure, starting from prenol (80 µL, 0.53 mmol), trichloroacetonitrile (120 uL, 1.2 mmol) and tetrabutylammonium phosphate (300 mg, 0.88 mmol) in 2mL acetonitrile. After work-up, DMAP was obtained as a white solid. The resulting spectra for DMAP were ^{1}H NMR (300MHz, D_2O): δ 5.26 (t, 1H), 4.17 (t, 2H), 1.61 (s, 3H), 1.56 (s, 3H) and ^{31}P NMR (300MHz, D_2O): δ 2.96.

Quantification of metabolites.

IP/DMAP and IPP/DMAPP were quantified by LC-MS/MS by comparison to synthetic IP/DMAP made in house according to the procedure described above and IPP, DMAPP standards purchased from Sigma-Aldrich. Liquid Chromatography was conducted using an Agilent 1100 Series HPLC (Agilent Technologies) and the MS/MS was conducted using an API 4000 triple quadrupole mass spectrometer (SCIEX) with ESI running in negative MRM mode. Mobile phases consisted of LCMS grade 10 mM tributylammonium (TBA) (Sigma-Aldrich), 15 mM acetic acid (Sigma) in water (A) and 100% acetonitrile (B). A EC18 column (2.7 µm, 2.1 mm, 50 mm length)

(Agilent) was used to separate 20 µL of sample with a flow rate of 0.3 mL/min and linear gradient program: 0-3 min 0% B, 3-10 min 0-50% B, 10-12 min 50-100% B, 12-18 min 100% B, 18-18.5 min 100-0% B, 0% B until 25 min. Metabolite specific ionization and fragmentation voltages were obtained from a 1 µM standard solution of each metabolite using the Analyst software (v 1.6) and monitored during the chromatography. Peaks were integrated using the Analyst software and compared to a standard curve generated for each metabolite.

Estimation of IUP expression strength.

The expression strength for IUP expression vectors pSEVA228-proDIUPi and pTET-IUPi was approximated through characterization using superfolded GFP (sGFP), as reported in (42), as a reporter gene. Derivatives of both IUP expression vectors, containing the sGFP ORF instead of the IUP operon, were created by first PCR amplifying the vector backbone from pSEVA228-proDIUPi using the primer pair GB-pSEVA-back F/R or from pTET-IUPi using the primer pair GB-pTET-back F/R respectively, then PCR amplifying the insert fragment containing sGFP from plasmid pTrcsGFP (42) using the primer pairs GB-sGFP-pSEVA F/R or GB-sGFP-pTET F/R respectively and finally assembling the respective fragments to give plasmids pSEVA228pro4-gfp and pTET-gfp. To assess the strength of the expression systems, a GFP-based assay, adapted from (36), was used. *E. coli* MG1655 DE3 transformed with either plasmid were grown at 37°C until reaching mid-log phase, at which point GFP fluorescence and OD_{600} was measured (time point 1; tp1). After a further 1.25h of growth (time point 2; tp2), GFP fluorescence and OD_{600} was again assayed and the GFP synthesis rate, which was used as a proxy for promoter strength, was calculated using the equation: Synthesis rate = $(GFP_{tp2}-GFP_{tp1})/OD_{600,average}$.

Knockout of the native MEP pathway.

The MEP pathway was knocked out by deleting *ispG* using the CRISPR-cas9 system, in a procedure adapted from (17, 39). First, pBAD33-proA-MEVI was created by Gibson assembly after PCR amplification of the pBAD33 backbone using primers GB-pBro IAI Vec F/R and

amplification of the lower mevalonate pathway from the pMBIS plasmid using primers GB-proX-Mevi F/R. The resulting plasmid pBAD33-proA-MEV was created to act as a rescue mechanism for the knockout of *ispG* which is normally non-viable. The targeting plasmid, pTargetF-ispG was created by first altering the N20 targeting sequence of the pTargetF plasmid using the primer pair GB-pTargetF-ispGN20_f and GB-ptargetF_N20_r to amplify pTargetF and subsequently circularizing the resulting PCR product with Gibson assembly. The vector was then amplified using the primer pair GB-pTargetF-vec_f/r and homology regions H1 and H2 were inserted. H1 was designed to encompass the 494 base pairs preceding the *ispG* gene and H2 was designed to encompass the 501 base pairs after the *ispG* gene. The homology regions were amplified using GB-ispG-H1_f/r and GB-ispG-H2_f/r respectively. The resulting fragments were then ligated using Gibson Assembly. *E. coli* MG1655 (DE3) was transformed with plasmids pBAD33-proA-MEV and pCas9 and plated on a chloramphenicol and kanamycin LB-agar plate overnight. The resulting double transformant was then grown at 30°C in LB media, which was supplemented with 1 mM D-arabinose at OD_{600} = 0.03. Upon reaching mid-log phase, the cells were harvested and washed with glycerol to make them electrocompetent and were then transformed with plasmid pTargetF-ispG and plated overnight at 30°C on LB-agar plates supplemented with kanamycin, chloramphenicol, spectinomycin, and 1 mM mevalonate. Deletion of *ispG* was confirmed by amplification of the area surrounding *ispG* in the genome using primer pair pCas9-ispG_f/r and sequencing the fragment using primer pCas9-ispG-seq_f. Strain KO1 was obtained by curing the cells of pTargetF-ispG by growing in LB media supplemented with 1 mM IPTG and subsequently curing the cells of pCas9 by growing overnight at 42°C. Strain KO2 was obtained by making KO1 electrocompetent and transforming with plasmid pSEVA-pro4IUPi. pCas and pTargetF were gifts from Sheng Yang (Addgene plasmids # 62225 & 62226). pMBIS was a gift from Jay Keasling (Addgene plasmid # 17817). Mevalonate used in this experiment was produced using the process described in (17) by mixing 1.02 volumes of 2 mM KOH with 1 volume of 2 mM DL-mevalonolactone (Sigma-Aldrich) and incubating at 37 °C for 30 min.

Labeling experiments.

Stains used in the pTET IUP labeling studies were revived in M9 media supplemented with ATCC trace minerals, 0.0045 g/L thiamine and 3.2% w/v U-^{13}C glucose. They were then subcultured in the same media and grown until early stationary phase at 37°C. At this point, 25 mM isoprenol, 25mM prenol or neither was added to cultures, as indicated, along with either 10ng/mL of the inducer anhydrotetracycline or no inducer and the cultures were sampled after approximately 48h. To extract metabolites, samples were by pipetting 5 mL of culture onto a vacuum filter flask with a 25 mm 0.2 um nylon filter. The cells were washed with 10 mL of water and the filter was submerged in ice cold 80% acetonitrile. Times and optical densities for each point were recorded. IPP levels were quantified by LC-MS/MS as described above.

2.9. Bibliography

1. Chatzivasileiou AO, Ward V, Edgar SM, Stephanopoulos G (2019) Two-step pathway for isoprenoid synthesis. *Proc Natl Acad Sci* 116(2):506–511.
2. Liu H, et al. (2013) Combination of Entner-Doudoroff pathway with MEP increases isoprene production in engineered Escherichia coli. *PLoS One* 8(12):e83290.
3. Farmer WR, Liao JC (2001) Precursor Balancing for Metabolic Engineering of Lycopene Production in Escherichia coli. *Biotechnol Prog* 17:57–61.
4. Ng CY, Farasat I, Maranas CD, Salis HM (2015) Rational Design of a Synthetic Entner-Doudoroff Pathway for Improved and Controllable NADPH Regeneration. *Metab Eng* 29:86–96.
5. Banerjee A, et al. (2013) Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J Biol Chem* 288(23):16926–16936.
6. Nagegowda DA, Bach TJ, Chye M-L (2004) Brassica juncea 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase 1: expression and characterization of recombinant wild-type and

mutant enzymes. *Biochem J* 383(Pt. 3):517–27.

7. Brooker JD, Russell DW (1975) Properties of microsomal 3-hydroxy-3-methylglutaryl coenzyme A reductase from Pisum sativum seedlings. *Arch Biochem Biophys* 167:723–729.

8. Bach TJ, Rogers DH, Rudney H (1986) Detergent-solubilization, purification, and characterization of membrane-bound 3-hydroxy-3-methylglutaryl-coenzyme A reductase from radish seedlings. *Eur J Biochem* 154:103–111.

9. Primak YA, et al. (2011) Characterization of a feedback-resistant mevalonate kinase from the archaeon Methanosarcina mazei. *Appl Environ Microbiol* 77(21):7772–7778.

10. Chang W, Song H, Liu H, Liu P (2013) Current development in isoprenoid precursor biosynthesis and regulation. *Curr Opin Chem Biol* 17(4):571–579.

11. Zhou K, Zou R, Stephanopoulos G, Too H-PP (2012) Metabolite Profiling Identified Methylerythritol Cyclodiphosphate Efflux as a Limiting Step in Microbial Isoprenoid Production. *PLoS One* 7(11):e47513.

12. Dugar D, Stephanopoulos G (2011) Relative potential of biosynthetic pathways for biofuels and bio-based products. *Nat Biotechnol* 29(12):1074–1078.

13. Gruchattka E, Hädicke O, Klamt S, Schütz V, Kayser O (2013) In silico profiling of Escherichia coli and Saccharomyces cerevisiae as terpenoid factories. *Microb Cell Fact* 12:84.

14. Ward VCA, Chatzivasileiou AO, Stephanopoulos G (2018) Metabolic engineering of Escherichia coli for the production of isoprenoids. *FEMS Microbiol Lett* 365(10):1–9.

15. Partow S, Siewers V, Daviet L, Schalk M, Nielsen J (2012) Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in Saccharomyces cerevisiae. *PLoS One* 7(12):1–12.

16. Carlsen S, et al. (2013) Heterologous expression and characterization of bacterial 2-C-methyl-d-erythritol-4-phosphate pathway in Saccharomyces cerevisiae. *Appl Microbiol*

Biotechnol 97:5753–5769.

17. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol* 21(7):796–802.

18. Pitera DJ, Paddon CJ, Newman JD, Keasling JD (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metab Eng* 9(2):193–207.

19. Anthony JR, et al. (2009) Optimization of the mevalonate-based isoprenoid biosynthetic pathway in Escherichia coli for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metab Eng* 11(1):13–19.

20. Gardner TS, Hawkins KM, Meadows AL, Tsong E, Tsegaye Y (2013) Production of Acetyl-Coenzyme A Derived Isoprenoids.

21. Shiba Y, Paradise EM, Kirby J, Ro DK, Keasling JD (2007) Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. *Metab Eng* 9(2):160–168.

22. Hemmerlin A, et al. (2006) A Cytosolic Arabidopsis D-Xylulose Kinase Catalyzes the Phosphorylation of 1-Deoxy-D-Xylulose into a Precursor of the Plastidial Isoprenoid Pathway. *Plant Physiol* 142(2):441–457.

23. Kuzuyama T, Takahashi S, Seto H (1999) Construction and characterization of Escherichia coli disruptants defective in the yaeM gene. *Biosci Biotechnol Biochem* 63(4):776–778.

24. Kirby J, et al. (2014) Enhancing Terpene Yield from Sugars via Novel Routes to 1-Deoxy-d-Xylulose 5-Phosphate. *Appl Environ Microbiol* 81(1):130–138.

25. King JR, Woolston BM, Stephanopoulos G (2017) Designing a New Entry Point into Isoprenoid Metabolism by Exploiting Fructose-6-Phosphate Aldolase Side Reactivity of Escherichia coli. *ACS Synth Biol* 6(7):1416–1426.

26. Dellas N, Thomas ST, Manning G, Noel JP (2013) Discovery of a metabolic alternative to the classical mevalonate pathway. *Elife* 2013(2):1–18.

27. Kang A, et al. (2016) Isopentenyl diphosphate (IPP)-bypass mevalonate pathways for isopentenol production. *Metab Eng* 34:25–35.

28. Cooper GM (2000) Cell Membranes. *The Cell: A Molecular Approach* (Sinauer Associates, Sunderland (MA)). 2nd Ed.

29. Gao S, et al. (2008) Substrate promiscuity of pyruvate kinase on various deoxynucleoside diphosphates for synthesis of deoxynucleoside triphosphates. *Enzyme Microb Technol* 43(6):455–459.

30. Li Y, et al. (2011) Substrate promiscuity of n-acetylhexosamine 1-kinases. *Molecules* 16(8):6396–6407.

31. Liu Y, Yan Z, Lu X, Xiao D, Jiang H (2016) Improving the catalytic activity of isopentenyl phosphate kinase through protein coevolution analysis. *Sci Rep* 6(April):24117.

32. VanNice JC, et al. (2014) Identification in haloferax volcanii of phosphomevalonate decarboxylase and isopentenyl phosphate kinase as catalysts of the terminal enzyme reactions in an archaeal alternate mevalonate pathway. *J Bacteriol* 196(5):1055–1063.

33. Lira LM, Vasilev D, Pilli RA, Wessjohann LA (2013) One-pot synthesis of organophosphate monoesters from alcohols. *Tetrahedron Lett* 54(13):1690–1692.

34. Wang Y, Xu H, Jones MK, White RH (2015) Identification of the final two genes functioning in methanofuran biosynthesis in Methanocaldococcus jannaschii. *J Bacteriol* 197(17):2850–2858.

35. Silva-Rocha R, et al. (2013) The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* 41(D1):666–675.

36. Davis JH, Rubin AJ, Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* 39(3):1131–1141.

37. Lee TS, et al. (2011) BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J Biol Eng* 5(1):12.

38. Chemler JA, Yan Y, Koffas MAG (2006) Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in Saccharomyces cerevisiae. *Microb Cell Fact* 5:1–9.

39. Jiang Y, et al. (2015) Multigene editing in the Escherichia coli genome using the CRISPR-Cas9 system. *Appl Environ Microbiol* 81(7):2506–2514.

40. Salis HM, Mirsky EA, Voigt CA (2010) Automated Design of Synthetic Ribosome Binding Sites to Precisely Control Protein Expression. *Nat Biotechnol* 27(10):946–950.

41. Espah Borujeni A, Channarasappa AS, Salis HM (2014) Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res* 42(4):2646–2659.

42. Santos CNS, Koffas M, Stephanopoulos G (2011) Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metab Eng* 13(4):392–400.

Chapter 3

Initial Assessment of the Isopentenol Utilization Pathway

Chapter adapted from (1)

3.1. Introduction

In the previous chapter, we demonstrated the design of a pathway for the synthesis of isoprenoids using isoprenol or prenol as substrate. We identified the appropriate enzymes for the pathway, successfully generated both *in vitro* and *in vivo* proofs of concept. We were able to establish that the pathway is strong enough to act as the sole isoprenoid supply pathway in *E. coli*, thereby completely replacing the native MEP pathway.

Having an isoprenoid pathway that can merely replace the limited isoprenoid needs for cell growth synthesis (2) is does not necessarily mean that our pathway will be able to meet our target, which has been the overproduction of isoprenoids, ideally at levels surpassing those achievable by the engineered MEP and MVA pathways. We thus set out to perform an initial assessment of the IUP's capabilities and limits.

3.2. IUP feedstock optimization

In order to characterize the pathway, we combined the IUP with a downstream module for lycopene synthesis. Lycopene is the C_{40} isoprenoid responsible for the coloration of tomatoes (3), and was chosen because of its red color, which enables it to be readily quantified via UV/Vis spectroscopy. One potential complication with the use of lycopene, however is its propensity to

be rapidly oxidized due to heat, and/or light during storage (4), leading to bleaching of the standard and over-estimation of lycopene titers. Thus, for developing our UV/Vis lycopene quantification standard curves, we produced our own fresh lycopene standard using standard procedure (5). Our fresh lycopene standard exhibited higher absorbance than one commercial standard, as shown by the respective quantification standard curves, indicating that the use of a commercial standard would have led to titer over-estimation.

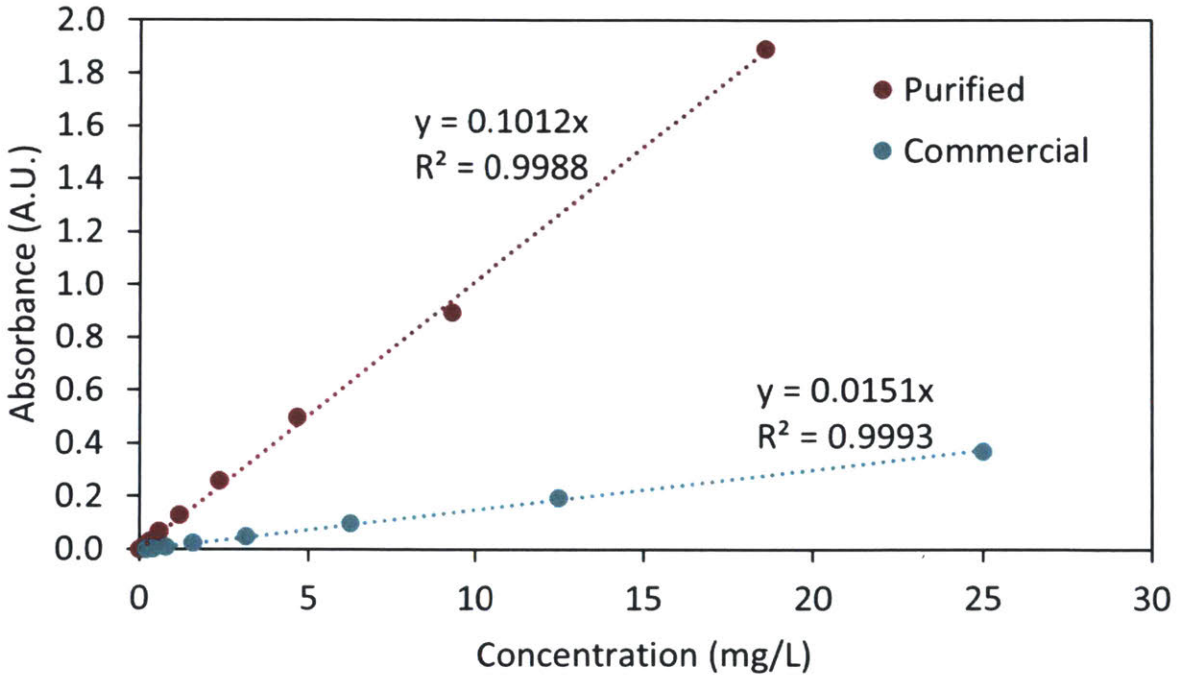


Figure 3-1: Standard Curve for lycopene quantification. Lycopene purchased from a commercial source is compared to lycopene freshly purified in-house.

We partitioned the upstream (IUP) and downstream (lycopene synthesis) genes into two operons carried on separate plasmids. The lycopene plasmid we used, pAC-LYCipi, which we purchased from addgene encoded the genes required for the production of lycopene (*crtE*, *crtB* and *crtI*) as well as a copy of *idi* from *Enterobacter agglomerans* (6). This plasmid was transformed together with either the pro4IUP or the pTET plasmids (for IUP expression) or alone (for control) in *E. coli* MG1655 DE3, which were then cultured in M9 media in 37°C for 48 hours, supplemented with

different quantities of isoprenol or prenol. After extracting and quantifying the lycopene content, we observed that, in general, isoprenol was a more preferred substrate than prenol. Furthermore, we observed that the pro4IUP strain (weaker IUP expression) generally performed better than the pTETIUP strain (stronger IUP expression), probably due to better balancing of the upstream and downstream modules. Since the highest lycopene titer was observed when 25mM of isoprenol was used, we settled on using 25mM isoprenol as the substrate for investigating IUP for all future *E. coli* experiments using this pathway.

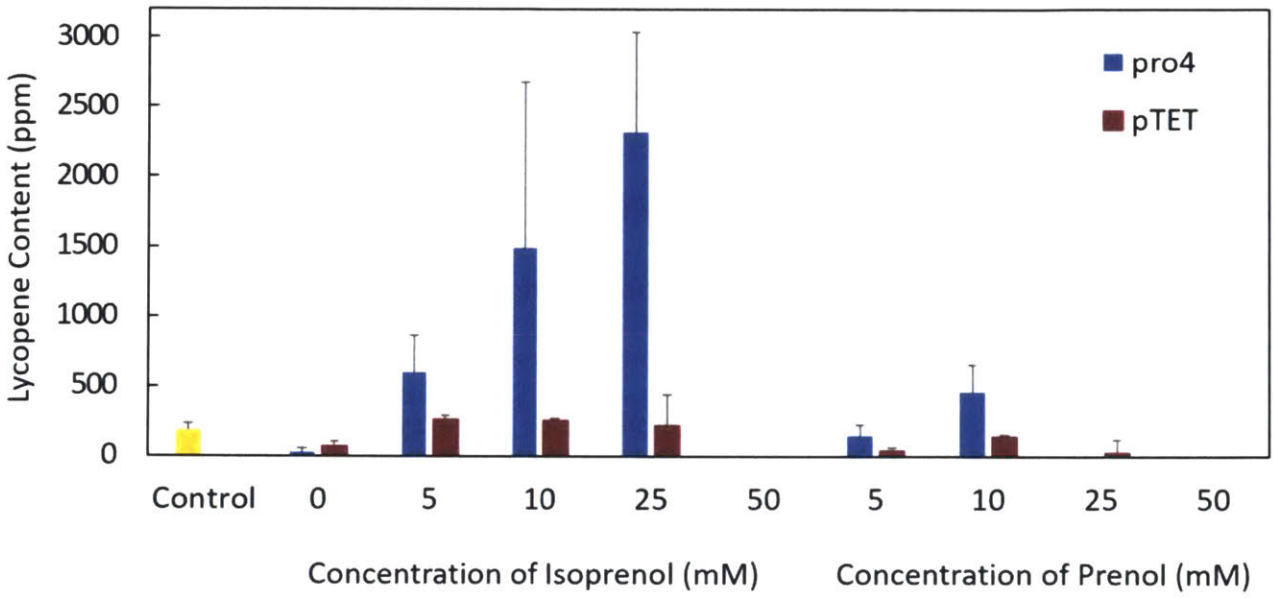


Figure 3-2: Characterization of the IUP using the lycopene pathway. Lycopene content in strains containing the IUP under the control of a low strength constitutive promoter (pro4) or a strong inducible promoter (pTET) induced with 20 ng/mL of aTc under different concentrations of isoprenol or prenol in media. This was compared to the control strain containing only the lycopene production plasmid (pAC-LYCipi), without the IUP plasmid or addition of either isopentenol (Control).

Measuring the concentration of metabolic intermediates, and particularly IPP, revealed that there was a significant pool of IPP/DMAPP in the pro4IUP strain, indicatinf a pathway imbalance, namely that the upstream (IUP) module generated a flux that was higher than the downstream (lycopene) module could consume.

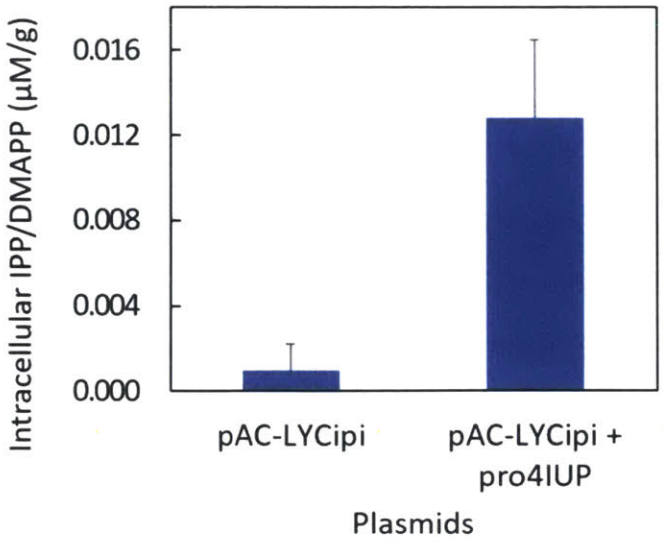


Figure 3-3: The effect of the pro4IUP pathway on intracellular levels of IPP/DMAPP compared to the control strain containing only pAC-LYCipi

3.3. Investigating the necessity of IPK in IUP

We wanted to probe into whether the enzyme that we chose to catalyze the second step of the IUP, AtIPK (Isopentenyl Phosphate Kinase from *A. Thaliana*), was essential for the IUP to function. Based on our experimental results (Figure 2-7), we had observed that ScCK (Choline Kinase from *S. cerevisiae*) was able to produce both IP and IPP when incubated *in vitro* in the presence of isoprenol, therefore indicating that it is able to catalyze both steps of the IUP and potentially negating the need for a second enzyme to catalyze the second step of the pathway.

In order to test whether AtIPK was beneficial or even essential, we removed the *ipk* gene from the IUP operon of the pSEVA228-pro4UPi plasmid, creating plasmid pSEVA228-pro4-ck-idi. This plasmid was also inserted in *E. coli* along with a lycopene expression plasmid and we subsequently assayed the resulting lycopene content. Our results indicated that although the absence of IPK still allowed the IUP to function, as shown by the increase in lycopene titer, the IUP functioned better when an IPK is present. Thus, we continued to include IPK in the IUP operon for all future experiments

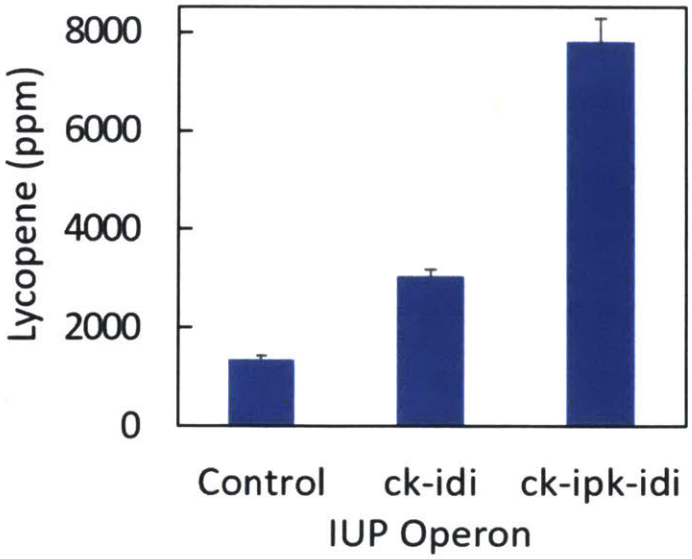


Figure 3-4: The effect of *ipk* gene on the synthesis of lycopene. Lycopene levels are indicative of the performance of the isoprenoid synthesis pathway when IUP is not expressed (Control), when choline kinase and isopentenyl diphosphate isomerase only are expressed (ck-idi) or when the full IUP operon, including isopentenyl monophosphate kinase, is used (ck-ipk-idi)

3.4. Pulse-chase experiment to study metabolite profiles using IUP

The significant IPP/DMAPP pool we noticed in the pro4IUP lycopene strain (Figure 3-3) revealed that the IUP could produce significant amounts of metabolic intermediates. Therefore, we designed a pulse-chase experiment aimed at studying metabolite profiles and flux through the IUP.

3.4.1 Initial estimation of the IUP flux

In this experiment, *E. coli* strains bearing either the pro4IUP or the pTETIUP plasmid, without any downstream cassette were grown at 37°C in M9 media, in shake flasks, using uniformly ^{13}C-labeled glucose as the sole carbon source. During stationary phase (at t=0), 25mM of unlabeled isoprenol were pulsed in and pTETIUP cultures were also induced with 10 ng/mL anhydrotetracycline to start production of the IUP enzymes. We then measured the

concentrations of metabolic intermediates at different time points. As has already been elaborated in Section 2.5.2, Figure 2-12, this scheme would allow us to differentiate the output of either the IUP, which would produce unlabeled IPP/DMAPP from isoprenol, and the native MEP pathway, which would produce uniformly labeled IPP/DMAPP from glucose, through central carbon metabolism. We also chose to measure the concentration of IP, the IUP's only intermediate, as well as MEC, an MEP-pathway metabolite that is known to accumulate in *E. coli* (7).

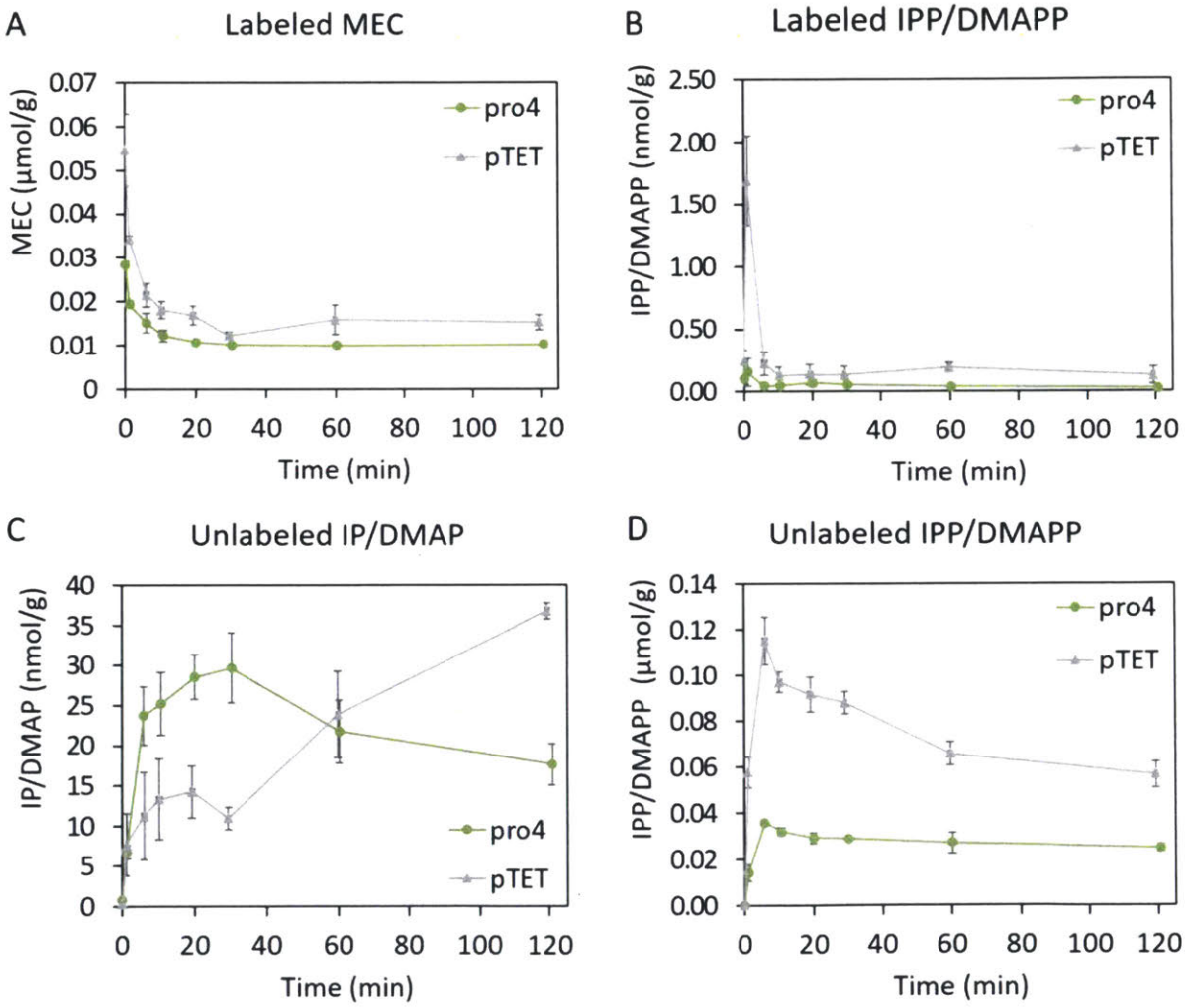


Figure 3-5: Isoprenol pulse-chase experiment for metabolite monitoring. (**A**) Levels of labeled MEC, (**B**) Levels of labeled IPP/DMAPP, (**C**) Levels of unlabeled IP/DMAP, (**D**) Levels of unlabeled IPP/DMAPP respectively in the pro4IUPi strain or pTETIUPi strains over the first 120 min after the pulsing in of isoprenol. pTET cultures were induced with 10 ng/mL of aTC. All experimental values are normalized to dry cell weight and represent the mean ± SD of three biological replicates.

Immediately prior to isoprenol addition, we could only detect fully-labeled MEC, an MEP-pathway metabolite that is known to accumulate in E. coli, as well as a small amount of fully-labeled IPP (Figure 3-5 B), while we could not detect any unlabeled IP or IPP (Figure 3-5 C & D). Within minutes after addition of isoprenol, we noticed that the concentration of unlabeled IP and IPP (Figure 3-5 C & D) rapidly increased to levels significantly higher than the ones the native MEP (Labeled IPP/DMAPP at t=0) produced (Figure 3-5 B). Conversely, the concentration of labeled IPP (Figure 3-5 A) dropped to barely detectable levels, and the levels of labeled MEP (Figure 3-5 A) also decreased significantly. Our interpretation of this data is possible feedback inhibition of the MEP pathways by the high IPP concentrations that the IUP produced, as has previously been reported in literature (8). We can also observe that the IUP is capable of producing IPP in concentrations that the native MEP is incapable of producing, especially in light of its feedback inhibition limitations.

In order to have a first-order estimate of IPP flux through the IUP based on the data of the pulse-chase experiment, a simple model was developed that utilizes the results of the pulse-chase labeling experiment. The basis of the model lies on Eq. 3.1, which models the total concentration of IPP ($[IPP_{TOT}]$) over time. In our experiment, IPP is being produced through either the MEP pathway (at a rate r_{MEP}) or through the IUP (at a rate r_{IUP}) and is consumed at a rate r_C.

$$\frac{d[IPP_{TOT}]}{dt} = r_{MEP} + r_{IUP} - r_C \qquad (3.1)$$

Given the nature of our experiment, we can have either labeled IPP (IPP_L), which is produced through the MEP pathway, or unlabeled IPP (IPP_{UL}), produced through the IUP. Therefore:

$$\frac{d[IPP_L]}{dt} = r_{MEP} - \lambda r_C \qquad (3.2)$$

$$\frac{d[IPP_{UL}]}{dt} = r_{IUP} - (1-\lambda) r_C \qquad (3.3)$$

In the above equations λ indicates the fraction of IPP that is labeled, i.e:

$$\lambda = \frac{[\mathrm{IPP_L}]}{[\mathrm{IPP_{TOT}}]} \xrightarrow[\mathrm{IPP_L+IPP_{UL}=IPP_{TOT}}]{} 1-\lambda = \frac{[\mathrm{IPP_{UL}}]}{[\mathrm{IPP_{TOT}}]} \quad (3.4)$$

Using the definition of λ, Eq. 3.3 is reworked as follows:

$$(1-\lambda)\frac{\mathrm{d}[\mathrm{IPP_{TOT}}]}{\mathrm{d}t} = \frac{\mathrm{d}\lambda}{\mathrm{d}t}[\mathrm{IPP_{TOT}}] + r_{\mathrm{IUP}} - (1-\lambda)r_{\mathrm{C}} \quad (3.5)$$

At this stage, the following assumptions are made: First it is assumed IPP consumption follows a 1st order rate law, in aggregate, in the cell. Secondly, it is assumed that that the fraction of labeled IPP is very small, something that is corroborated by our data. Therefore:

$$r_{\mathrm{C}} = k[\mathrm{IPP_{TOT}}]$$

$$1-\lambda \approx 1$$

In light of the above assumptions Eq. 3.5 is therefore transformed as follows:

$$\frac{\mathrm{d}[\mathrm{IPP_{TOT}}]}{\mathrm{d}t} = \left(\frac{\mathrm{d}\lambda}{\mathrm{d}t} - k\right)[\mathrm{IPP_{TOT}}] + r_{\mathrm{IUP}} \quad (6) \quad (3.6)$$

We further assume that for at least the first 30 mins of the experiment there exists a quasi-steady state, meaning that the terms $\left(\frac{\mathrm{d}\lambda}{\mathrm{d}t} - k\right)$ and r_{IUP} will remain relatively constant, rendering Eq. 3.6 a first order ordinary differential equation. If we solve Eq. 6 using the initial condition $\mathrm{IPP_{TOT,0}}=\mathrm{IPP_0}$ (IPP at t=0, which is measured), we get the solution:

$$[\mathrm{IPP_{TOT}}] = \frac{\mathrm{e}^{\mathrm{A}t}\left(\mathrm{B} + \mathrm{A}\,\mathrm{IPP_0}\right) - \mathrm{B}}{\mathrm{A}} \quad (7) \quad (3.7)$$

Therefore, we calculated the total IPP concentration from our data by summing up the measured values for labeled and unlabeled IPP and then performing a least-square fit to Eq. 7 (with $\left(\frac{\mathrm{d}\lambda}{\mathrm{d}t} - k\right)$ and r_{IUP} as the fitting parameters) in order to estimate the value of r_{IUP}, i.e. the flux through the IUP. We thus estimated that the IUP flux when using the pro4IUP plasmid was 0.379 μM/(OD·h),

or 1.63 µmol/(g_{dcw}·h) and the flux under the control of the pTETIUP plasmid was 1.145 µM/(OD·h) or 4.93 µmol/(g_{dcw}·h). This result revealed that our pathway's ability to generate copious amounts of IPP translates to a high flux, that is comparable (about 33%) to the best flux achieved by the MEP pathway (9), showing the potential of the IUP.

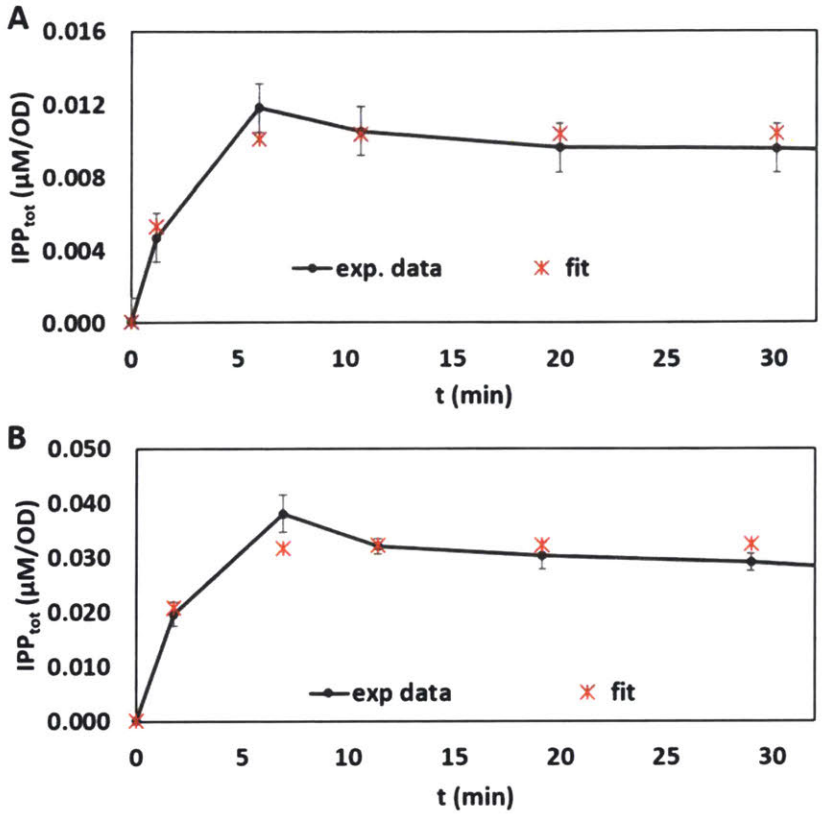


Figure 3-6: Curve fitting to isoprenol pulse-chase experiment results for flux estimation. A curve was fitted to the experimental results (0-30min) in the case of the (**A**) pro4IUPi strain or the (**B**) pTETIUPi strain to estimate the flux through the IUP. pTET cultures were induced with 10 ng/mL of aTC.

3.4.2 IUP is orthogonal to central carbon metabolism

As part of the above experiment, we also monitored the metabolic profiles of the glycolytic intermediates phosphoenolpyruvate (PEP) and 3-phosphoglyceric acid (3PG), two metabolites which participate in central carbon metabolism. We observed that their labeling patterns remained unchanged after the isoprenol pulse and both metabolites were almost uniformly ^{13}C-labeled,

meaning that they were produced through catabolism of uniformly ^{13}C-labeled glucose. The fact that the labeling pattern did not become lighter after isoprenol was pulsed in the media indicated that carbons from isoprenol, which entered the cell through the IUP, did not "leech" into the central carbon metabolism, thereby leading us to the conclusion that the IUP is uncoupled from main glycolysis and orthogonal to it. This is therefore a clear indication that through the IUP we can uncouple growth and isoprenoid production, and theoretically have the IUP give us 100% yield on isoprenol, since carbon from isoprenol cannot be "lost" to non-isoprenoid products, features that cannot be achieved with natural pathways.

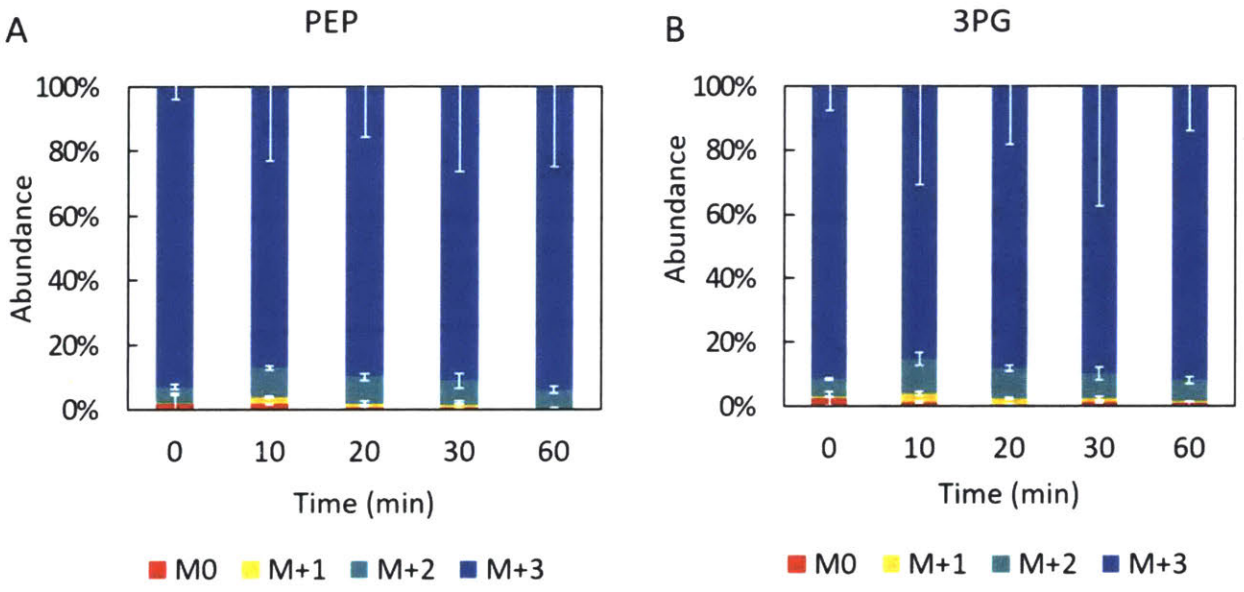


Figure 3-7: Isoprenol pulse-chase experiment for metabolite monitoring. Labeling patterns for (**A**) phosphoenolpyruvate (PEP) and (**B**) 3-phosphoglycerate (3PG) respectively in the pro4IUPi strain over the first 60 min. The labeling pattern remained unchanged even after 24h.

3.5. Investigating the effect of the IUP on cell growth

In light of the IUP's ability to produce large quantities of IPP/DMAPP, we wanted to explore the possibility of the IUP having a detrimental effect on growth. Literature suggests that high concentrations of IPP can be toxic to the cell in a multitude of ways (10, 11), and high IPP concentrations have been known to hinder previous isoprenoid pathway engineering efforts (12–

14). Additionally, in our experiments we had noticed a small decrease in cell concentration in cultures in which the IUP was expressed, compared to control cultures not expressing the IUP or supplemented with isoprenol. In order to assay the effect of the IUP on cell growth, we grew *E. coli* in which the IUP was expressed via the pro4IUP or the pTET plasmids or was not expressed (control) in the presence of either no isoprenol or 25mM isoprenol in the media and measured the exponential growth rate. Our experiments revealed that while there is indeed a growth defect in cultures in which the IUP is active, this can be attributed to the presence of isoprenol in the media, rather than the IUP itself, as the control strain also suffers the same growth defect when grown in 25mM isoprenol.

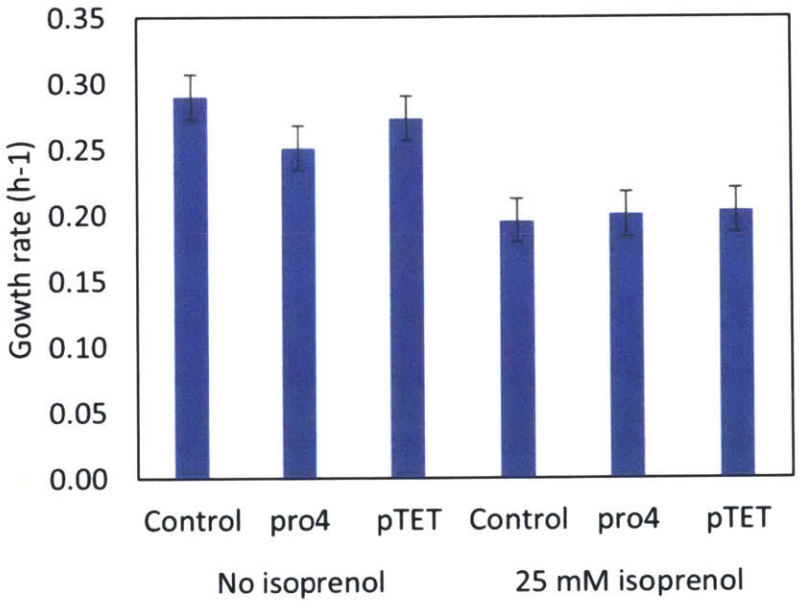


Figure 3-8: Growth Rate Comparison. Exponential growth rates of strains containing the pSEVA228pro4-IUP plasmid (pro4), the pTET-IUP plasmid (pTET) or no plasmid (control), grown in M9 minimal media supplemented with either 0 or 25 mM isoprenol. All values represent the means ± SD of 3 biological replicates.

3.6. Conclusions

In summary, we identified the optimal substrate for IUP as isoprenol and its optimal concentration as 25mM. Furthermore, we showed that although IPK is not necessary for the function of the IUP,

its expression boosts the performance of the IUP, validating our approach of not relying on ScCK only for the catalysis of both steps. Our pulse chase experiments demonstrated that the IUP can function with high flux and can achieve a flux value that is roughly one third that achieved by our lab's best result (9), namely the flux sustained by the engineered MEP pathway that led to the production of 1g/L taxadiene over 5 days. Furthermore, we observed that the IUP is decoupled from Central Carbon Metabolism, rendering the pathway's theoretical yield to 100%.

3.7. Materials and Methods

Strains, plasmids and genes; cloning protocol

Materials and methods were the same as the ones as described in the relevant sections of Chapter 2.

Construction of expression vectors.

All vectors were constructed following the routine cloning protocol described in Chapter 2. Vectors pSEVA228-pro4IUPi and pTET-IUPi for the expression of the IUP pathway were constructed by the method described in Chapter 2. The plasmid pSEVA228pro4-ck-idi, which carries a reduced version of the IUP operon, lacking *ipk*, was created by amplifying pSEVA228-pro4IUPi using the primer pair GB-IUPnoIPK_f/r and then subsequently assembling the amplification product. Sequences for the primers used are listed in **Table A3** in the Appendix. Codon-optimized sequences are listed in **Table A5** in the Appendix. The lycopene plasmid pAC-LYCipi was a gift from Francis X Cunningham Jr (Addgene plasmid # 53279) (6).

Quantification of metabolites.

IP/DMAP, IPP/DMAPP, MEC, PEP and 3PG were quantified by LC-MS/MS by comparison to synthetic IP/DMAP made in house according to the procedure described in Chapter X and IPP,

DMAPP, MEP, PEP and 3PG standards purchased from Sigma-Aldrich using the method described in Chapter 2.

Quantification of lycopene.

Lycopene content was quantified by UV-Vis spectroscopy. First, 1 mL of cells was transferred to an amber microtube and centrifuged at 16 000 x g for 2 min. The cell pellet was then resuspended in 1mL of a 50% ethanol, 50% acetone solution and vortexed for 30 min (VWR). The solution was centrifuged to remove particulates and 200 µL was transferred to a microplate and the absorbance at 475 nm was recorded. This was compared to a standard curve generated using a standard freshly purified in-house as the commercial standards purchased from three different companies (Indofine Chemical Company, Inc., Carbosynth, Santa Cruz Biotechnology) were found to be overestimating the lycopene content by about 10-fold, presumably due to degradation and bleaching of these standards during storage (4). Therefore, a fresh standard was prepared according to a standard protocol (5) from *E. coli* biomass expressing the lycopene synthesis genes. First, *E. coli* biomass was centrifuged, then resuspended in acetone in the dark and left to stir for 1h. This solution was then filtered, and a small amount of acetone was added to wash residual lycopene from the cells. The filtrate was then chilled at -20°C to induce crystallization of the lycopene. Crystals were recovered by filtration and the crystallization process was repeated twice to purify the lycopene. The resulting lycopene was dried, weighed and subjected to UV/VIS spectroscopy to confirm its authenticity. It was compared to a commercially available standard from Indofine in Fig S 5. It was then resuspended in 50% ethanol, 50% acetone solution to create a standard curve for quantification. Lycopene content was calculated using the cell density of the culture calculated from the optical density at 600 nm using a correlation of 0.33 g/A.U.

Quantification of isoprenol.

Isoprenol was quantified using GC-FID using ultra-pure helium as the carrier gas. First, 1mL of aqueous sample was centrifuged at 1000 x g to pellet any solids (e.g. cells) and 100µL of the

supernatant were transferred to a vial for sampling via GC. Isoprenol was quantified from 1 µL of the above on a HP-INNOWAX capillary column (30m, 250 µm, 0.15µm) (Agilent Technologies) using a 7890B Series GC and with a flame ionization detector. Chromatography was performed under the following conditions: splitless injection, inlet temperature 260°C, constant inlet pressure 15.707psi. An oven program of 40°C, hold 1min, 20°C/min until 200°C, hold 2 min, 40°C/min until 250°C, hold 5 min was used. The FID was operated at temperature of 260°C with a He flow of 30mL/min, an air flow of 300mL/min and a makeup flow of 25mL/min. Isoprenol was quantified using a standard curve based on commercial isoprenol standard (Sigma-Aldrich).

Cultivation in serum bottles.

All media and media additives were prepared according to manufacturers' recommendations and autoclaved or sterile-filtered (when casamino acids were supplemented) prior to use. Antibiotics and inducers were filter sterilized and stored as 1000x solutions at -20°C until use. Strains were revived from glycerol stocks in LB media (BD) containing the appropriate antibiotic by culturing overnight at 37°C. Overnight cultures were inoculated at 1% (v/v) into 20 mL of M9 media (US Biological Life Sciences) containing 0.32% w/v glucose, 0.5% w/v casamino acids (Tecknova) and ATCC trace minerals. When they reached an OD_{600} of 0.5, if necessary, 25 mM (or the specified concentration) isoprenol was added as substrate for the IUP and/or 10 ng/mL or 20 ng/mL of anhydrotetracycline, as indicated, was added to induce the pTET IUP operon and/or IPTG was added to a final concentration of 100 µM to induce downstream plasmid expression. Strains were cultured in 110 mL serum bottles with rubber stoppers to prevent the evaporation of isoprenol. Strains for lycopene production were grown as 37°C. Metabolites and lycopene were quantified using protocols described above.

Labeling experiments.

Stains used in the pro4 and pTET IUP labeling studies were revived in M9 media with 3.2% w/v U-^{13}C glucose. They were then subcultured in the same media and grown until early stationary

phase at 37°C. Samples were taken prior to the start of the pulse by pipetting 5 mL of culture onto a vacuum filter flask with a 25 mm 0.2 um nylon filter. The cells were washed with 10 mL of water and the filter was submerged in ice cold 80% acetonitrile. At this point, 25 mM isoprenol was added to each culture and the cultures were sampled at approximately 1, 5, 10, 15, 30, 60 and 120 min. Times and optical densities for each point were recorded. IP, IPP, MEP, 3PG and PEP levels were quantified by LC-MS/MS as described above. All trials were performed using three biological replicates. A method described in the in the main text was used to estimate the IUP flux. For taxadiene labeling experiments, the cultures were prepared similarly except they were incubated at 30°C for 48 h after induction and a C18 flash resin was added. At 48 h, the metabolites were extracted and determined by LC-MS/MS as described in the above. Taxadiene was eluted from the resin and quantified using the GCMS method described in the above.

3.8. Bibliography

1. Chatzivasileiou AO, Ward V, Edgar SM, Stephanopoulos G (2019) Two-step pathway for isoprenoid synthesis. *Proc Natl Acad Sci* 116(2):506–511.
2. Siperstein MD (1984) Role of cholesterogenesis and isoprenoid synthesis n DNA replication and cell growth. *J Lipid Res* 25:1462–1468.
3. Takehara M, et al. (2014) Characterization and Thermal Isomerization of (all-E)-Lycopene. *J Agric Food Chem* 62(1):264–269.
4. Srivastava S, Srivastava AK (2015) Lycopene; chemistry, biosynthesis, metabolism and degradation under various abiotic parameters. *J Food Sci Technol* 52(1):41–53.
5. Davis WB (1949) Preparation of Lycopene from Tomato Paste for Use as a Spectrophotometric Standard. *Anal Chem* 21(10):1226–1228.
6. Cunningham FX, Lee H, Gantt E (2007) Carotenoid biosynthesis in the primitive red alga Cyanidioschyzon merolae. *Eukaryot Cell* 6(3):533–545.

7. Zhou K, Zou R, Stephanopoulos G, Too H-PP (2012) Metabolite Profiling Identified Methylerythritol Cyclodiphosphate Efflux as a Limiting Step in Microbial Isoprenoid Production. *PLoS One* 7(11):e47513.

8. Banerjee A, et al. (2013) Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J Biol Chem* 288(23):16926–16936.

9. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* 330(2010):70–74.

10. Kang A, et al. (2016) Isopentenyl diphosphate (IPP)-bypass mevalonate pathways for isopentenol production. *Metab Eng* 34:25–35.

11. George KW, et al. (2018) Integrated analysis of isopentenyl pyrophosphate (IPP) toxicity in isoprenoid-producing Escherichia coli. *Metab Eng* 47(January):60–72.

12. Withers ST, Gottlieb SS, Lieu B, Newman JD, Keasling JD (2007) Identification of isopentenol biosynthetic genes from Bacillus subtilis by a screening method based on isoprenoid precursor toxicity. *Appl Environ Microbiol* 73(19):6277–6283.

13. George KW, et al. (2014) Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production. *Biotechnol Bioeng* 111(8):1648–1658.

14. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol* 21(7):796–802.

Chapter 4

Coupling the IUP with downstream modules for the synthesis of isoprenoids

Sections of this chapter adapted from (1)

4.1. Introduction

As described in Chapter 3, we observed that the IUP can function well *in vivo* and produce lycopene, a trial product, at levels higher that the wild-type control which utilized the MEP pathway. We also observed that the IUP can produce IPP, the key isoprenoid intermediate, with high flux and at high concentrations, indicating the pathway's potential. In order to utilize this abundance of IPP, we proceeded to couple the two-step IUP (i.e. upstream module) with pathways for the production of different isoprenoids (downstream modules).

4.2. Tools for measuring IUP products and yield

Having already proven (see Chapter 3) that the IUP can produce lycopene (a C_{40}, nonvolatile product) at high titer, we were interested in establishing whether the IUP can be used for the production of lighter, volatile isoprenoids. Furthermore, we wanted to explore ways for measuring the pathway yield on isoprenol.

4.2.1 Method for extraction and quantification of volatile isoprenoids

Unlike lycopene, which can be easily quantified by drying the cells, extracting lycopene and subsequently using UV/Vis spectrophotometry, volatile isoprenoids need to be extracted/captured

in situ and then be quantified via GC-MS. The usual means of extracting isoprenoids *in situ* are either a dodecane overlay over the culture broth or a C18 resin suspension, which is subsequently collected and has volatile isoprenoids extracted from it. Before we could embark on the production of volatile isoprenoids through the IUP, we wanted to establish that isoprenol would not get captured by either of the extraction media, thereby lowering the potential yield and altering the optimal isoprenol concentration in the culture media.

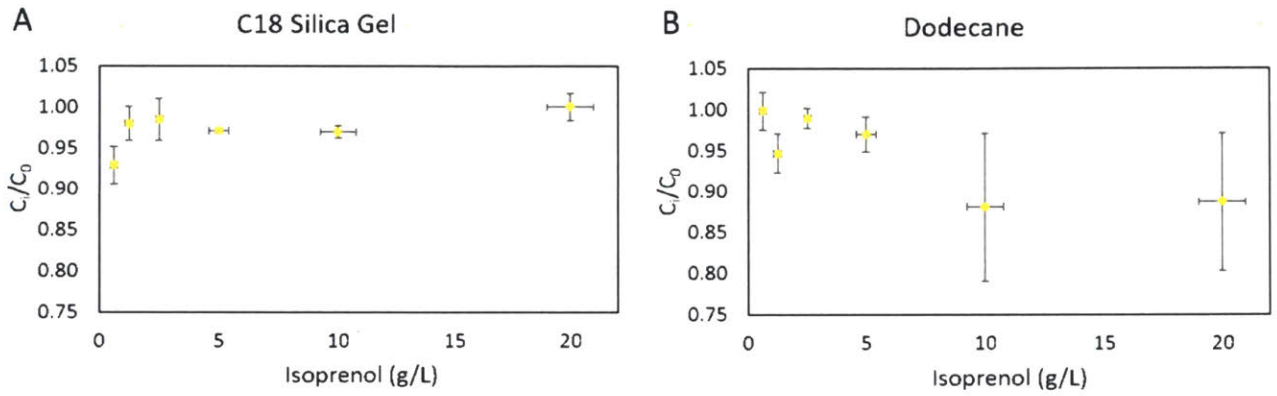


Figure 4-1: Isoprenol adsorption isotherms (**A**) for C18 silica gel, (**B**) for dodecane. All values represent the means ± SD of 3 replicates.

To investigate isoprenol absorbance by C18 silica gel resin and its partition in the water/dodecane system, we incubated cell culture media which contained isoprenol in concentrations of ~0.5 – 20 g/L and into which we either added C18 resin or overlaid by dodecane for 2 days in 37°C in order to mimic culture conditions. We then measured the concentration of isoprenol in the aqueous culture media and compared it to the initial isoprenol concentrations. We observed (Figure 4-1) that isoprenol absorbance by either C18 silica gel or dodecane was very low and that we could therefore use either, as they does not strip isoprenol from the media.

4.2.2 Method for estimating yield

An issue that we were interested in addressing was the yield of the IUP pathway, i.e. the percent conversion of isoprenol to useful products. To that end, we performed an experiment wherein we

produced lycopene via the IUP (using the pSEVA228-pro4IUPi plasmid), with the objective of quantifying the produced lycopene and then comparing it against the amount of isoprenol consumed. As a control, we also had a culture which produced lycopene without the use of the IUP, that is, lycopene was produced via the MEP pathway. Isoprenol (25mM) was added in both cultures, with the objective of accounting for evaporation, which would happen in both cases, whereas consumption and conversion to lycopene would only happen in the case of the IUP-expressing culture.

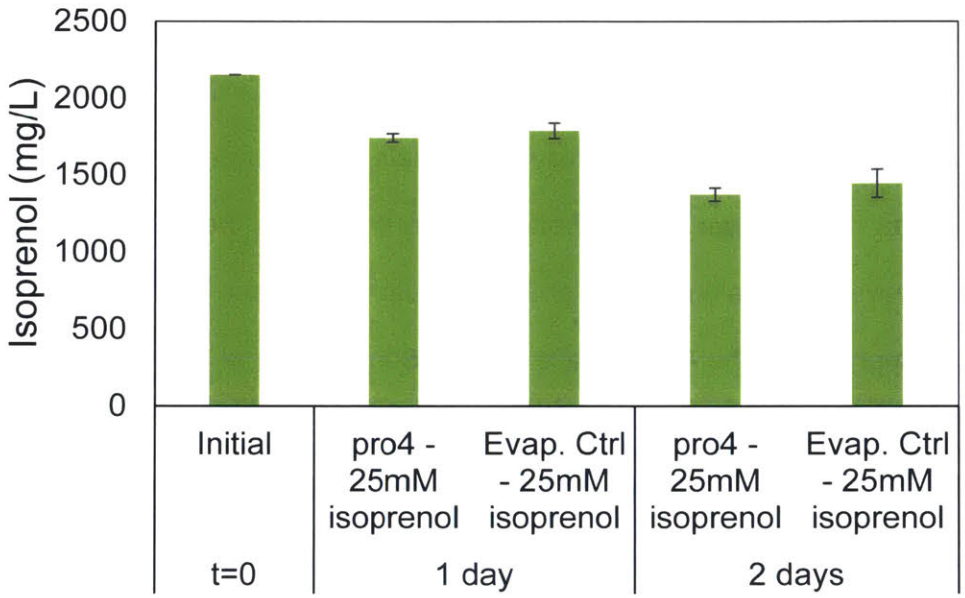


Figure 4-2: Isoprenol levels in cultures with lycopene-producing strains. Strains were cultured at M9 media and the isoprenol content was measured after 1 or 2 days. Cultures differ on whether they express the IU pathway (pro4) or they do not (Evap. Ctrl.). Concentrations are reported as means ± SD of three biological replicates.

Unfortunately, we noted that the experimental error in measuring the concentration of isoprenol in the culture media was high enough that it made the comparison between the "consumption + evaporation" lines and "evaporation" control impossible, thus not allowing us to estimate the yield on isoprenol. This issue mirrors a similar difficulty our lab has experienced with the estimation of CH_4 incorporation by methanotrophic cultures (1), a problem that still has not been fully resolved.

4.3. Optimization and use of downstream modules

4.3.1 Lycopene

As a first step in the lycopene culture optimization, we altered the culture conditions, specifically the media composition, the timing of isoprenol addition and the temperature. As a baseline, we used the strain bearing our lycopene plasmid, along with the pSEVA228-pro4IUP plasmid for IUP expression (or absent that in the case of the no IUP control). Cells would be cultured at 37°C in M9 media with 3.2 g/L glucose and Wolfe's Trace Minerals. Isoprenol (25mM) would be added at OD_{600}=0.5, i.e. early exponential phase.

In varying the media composition, we chose to try using 6.4 g/L glucose as the carbon source, to allow cells to grow for longer and hopefully accumulate more lycopene, or use 3.2 g/L glycerol, since glycerol has been known to give better results with either the MEP (2) or the MVA (3) isoprenoid pathways. Besides the carbon source, we varied the trace element solution to replace Wolfe's Trace Minerals with the ATCC Trace Mineral Supplement, which is similar, or completely change the media to biotin-aneurin-folic acid (BAF) media (4), which have a high concentration of glucose and are also supplemented with yeast extract and thiamine. In terms of isoprenol addition we attempted to add isoprenol at OD_{600}=0.5, i.e. late exponential/early stationary phase. We also attempted culturing the cells at 30°C.

We were not able to conclusively identify one of these different culture conditions as being superior to the proIUP Control. We noticed that the use of glycerol instead of glucose did not increase isoprenoid productivity unlike what has been observed when producing isoprenoids through the MEP pathway (2), which is consistent with our observation that isoprenoid production is decoupled from central carbon metabolism. Furthermore, the thiamine and yeast extract addition featuring in BAF media did not noticeably improve lycopene production.

Unfortunately, in many cases, we observed that the use of pAC-LYCipi did not lead to robust lycopene production; e.g. in certain experiments some of the replicates would not produce

lycopene, whereas others would, and this phenomenon would repeat even if the replicates originated from the same colony. Thus, in order to improve stability of the plasmid, as well as boost the downstream, since we had already observed high IPP concentrations, we proceeded to engineer the plasmid itself.

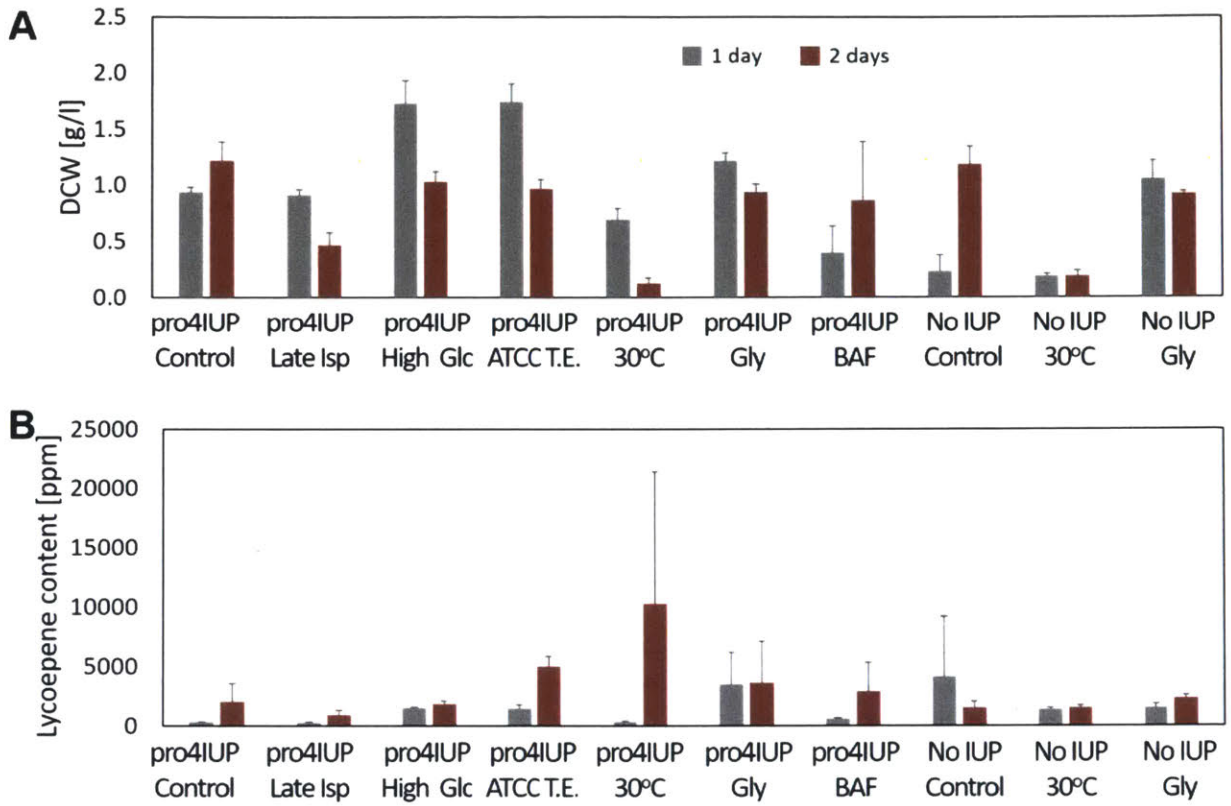


Figure 4-3: Optimization of lycopene strain culture conditions (**A**) Culture dry cell weight (**B**) Lycopene concentration expressed as ppm=μg/g_{dcw}. Cultures differ on whether they express the IU pathway (pro4) or they do not (No IUP). The control cultures are cultured at 37°C in M9 media with 3.2 g/L glucose and Wolfe's Trace Minerals and have isoprenol added to them at OD_{600}=0.5. Non-control cultures either use glycerol instead of glucose (*Gly*), higher concentrations of the latter (*High Glc*), different trace mineral composition (*ATCC T.E.*) or altogether different culture media (*BAF*). They can also vary in that isoprenol is added at higher OD_{600} (Late Isp), or they are cultured at 30°C. Values are reported as means ± SD of three biological replicates.

We first started by varying the plasmid origin or replication, exchanging the low copy number pAC origin with the p15A (copy number of approximately 10), pBR322 (copy number of approximately 20) and pUC (copy number of approximately 200). We observed that changing the origin made

the plasmid much more robust in terms of lycopene productivity replicability and also increased the lycopene titers.

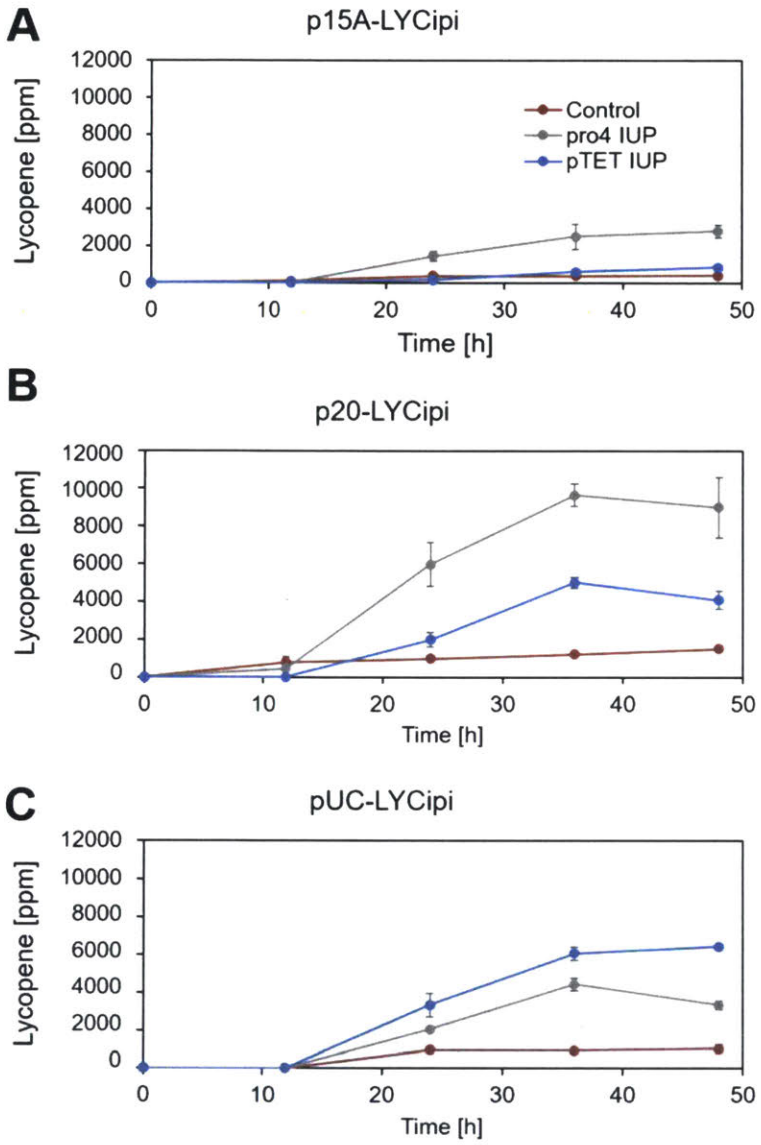


Figure 4-4: Effect of lycopene plasmid origin of replication on lycopene content (**A**) Origin p15A (Copy # ~10) (**B**) Origin pBR322 (Copy # ~20) (**C**) Origin pUC (Copy # ~200). Lycopene concentration. Isoprenol was added when each culture reached OD_{600}=0.5 Cultures bearing the above plasmids expessed the IUP though either the pSEVA228-pro4IUPi plasmid (pro4-IUP) or the pTET-IUPi plasmid (pTET-IUP) or did not express it at all (Control). Concentrations are expressed as ppm=μg/g_{dcw} and are reported as means ± SD of three biological replicates.

For the next step, we changed the origin of replication to pSC101 (copy number of approximately 5) but also the promoter, replacing the original, constitutive promoter with a T7 promoter, inducible

by IPTG (Isopropyl β-D-1-thiogalactopyranoside). The resulting plasmid p5T7-LYCipi was able to produce, through the IUP, similar lycopene titers (~10,000ppm) to the previously constructed plasmids. However, on closer inspection, we noticed that the majority of lycopene content was produced within the first 8h, indicating that the downstream module, as expressed through p5T7-LYCipi was highest-producing fastest yet.

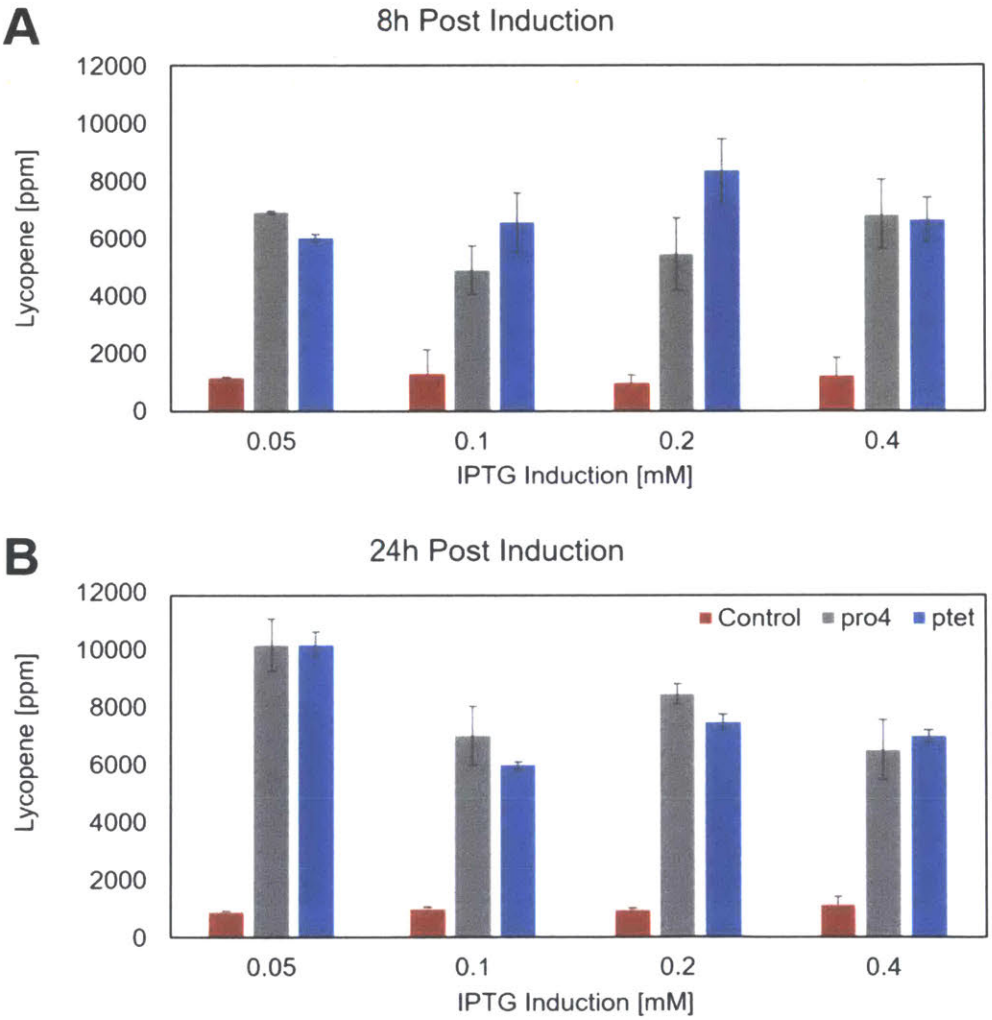


Figure 4-5: Change in origin of replication and lycopene operon promoter results in higher lycopene productivity. Shown is the lycopene concentration achievable through the use of the p5T7-LYCipi plasmid (T7 promoter, origin of replication pSC101, copy no. ~5) within (**A**) 8 hours or (**B**) 24 hours after isoprenol and inducer were added to the cultures, when they reached OD_{600}=0.5. Cultures either expressed the IU pathway though the pSEVA228-pro4IUPi plasmid (pro4) or the pTET-IUPi plasmid (pTET), or did not express it at all (Control). Concentrations are expressed as ppm=μg/g_{dcw} and are reported as means ± SD of three biological replicates.

4.3.2 Volatile isoprenoids

In order to prove IUP's usefulness as a general isoprenoid synthesis pathway, we decided to attempt to combine it with modules for the production of volatile isoprenoids, which constitute a large class of isoprenoid products. We primarily focused on the production of the volatile isoprenoid taxadiene, a precursor to the anticancer drug paclitaxel we have previously produced in the lab with high titer (2).

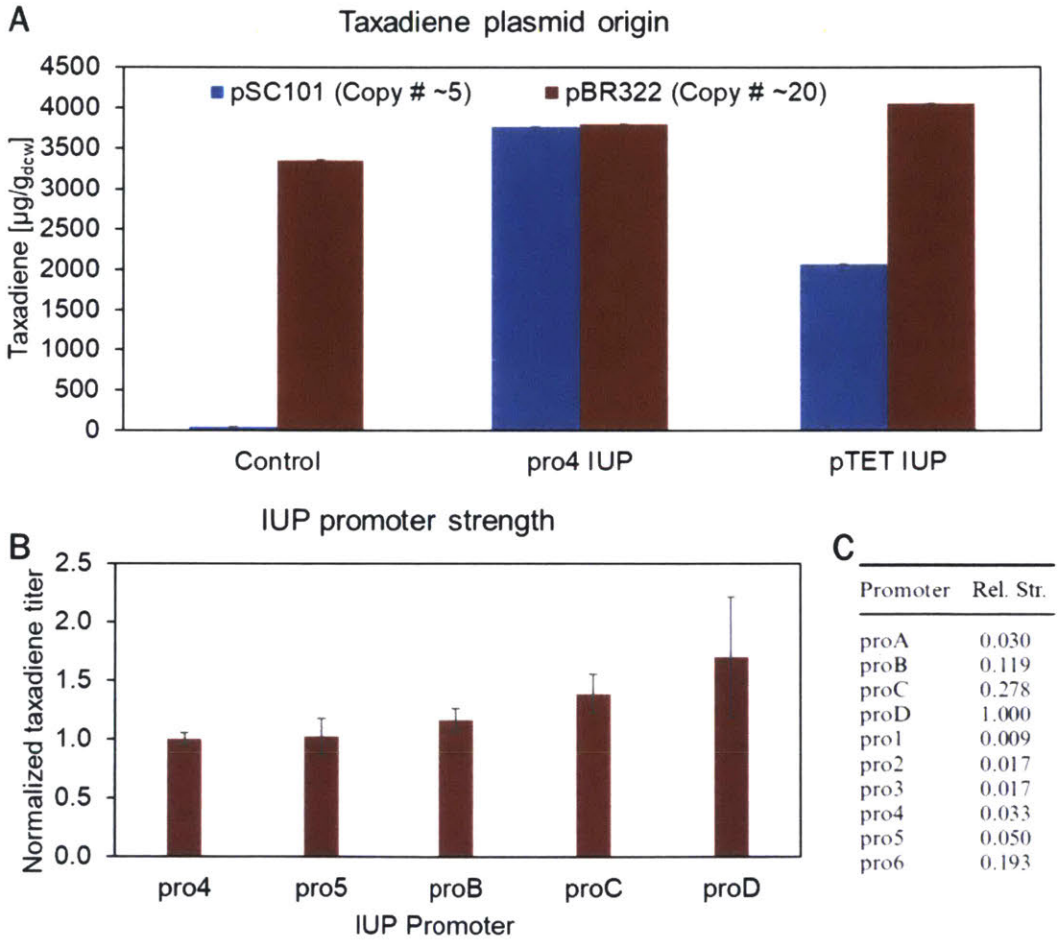


Figure 4-6: Effect of taxadiene plasmid origin of replication and IUP promoter strength on taxadiene content. Taxadiene titers after culturing for 48h in M9 media supplemented with 1% tryptone at 22°C. (**A**) Strains contained either taxadiene plasmid p5T7-tds-ggpps (origin pSC101, copy # ~5) or p20T7-tds-ggpps (origin pBR322, copy # ~20) and expressed the IUP though either the pSEVA228-pro4IUPi plasmid (pro4 IUP) or pTET-IUPi plasmid (pTET IUP), or not at all. (**B**) Strains expressed the taxadiene pathway through the plasmid p5T7-tds-ggpps and the IUP though plasmids of the series pSEVA228-proXIUPi (proXIUP), where (**C**) "proX" are constitutive promoters of varying strengths. Concentrations are reported as means ± SD of three biological replicates.

We attempted improve taxadiene titer by varying the copy number of the taxadiene plasmid, p5T7-tds-ggpps, by changing its origin of replication, pSC101 (copy number of approximately 5) to pBR322 (copy number of approximately 20), thus creating p20T7tds-ggpps (Figure 4-6 A). We noticed that the new plasmid backbone improved taxadiene for the wild type strain but not our pathway. We also tried varying the strength of the IUP plasmid, by using progressively stronger constitutive promoters (5) to create plasmids pSEVA228-proXIUPi (where X=4, 5, B, C or D), transforming the IUP plasmids together with the p5T7tds-ggpps plasmid and culturing them at the optimal culture conditions previously identified. We observed (Figure 4-6 B) that, although the taxadiene content was increased with increasing IUP strength, the increases were not statistically significant.

The above cultures were run under culture conditions which had been optimized through rounds of experiments designed through Design of Experiments (DOE), the full results of which are available in Tables A6-i-iv in the Appendix to optimize the culture conditions. To identify the optimal culture conditions, we varied the culture temperature (22, 26 or 30 °C), the IPTG induction for the taxadiene plasmid (0.05, 0.125 or 0.2 mM), the OD_{600} at which induction of the upstream or the downstream modules occurs (~0.4, ~0.6 or ~0.8) and the concentration of isoprenol supplementation (2.5, 5.0, 10, 12.5, 20 or 30 mM). Other parameters that were varied were the carbon source carbon source (glucose or glycerol), its concentration (5.5, 7.5, 12.5, 17.5 or 19.5 g/L), the concentration of isoprenol supplementation (2.5, 5.0, 12.5 or 25 mM) and an additional nitrogen source (1% of either tryptone, casamino acids and yeast extract). The IUP was expressed via pSEVA228-pro4IUPi (pro4IUP), or the pSEVA228-pro5IUPi (pro5IUP) plasmids, both of which use constitutive promoters, with the pro5 promoter being 50% stronger than the pro4 promoter (5), or the pTETIUP plasmid, which uses an inducible promoter for which we varied the concentration of its inducer, anhydrotetracycline (10, 15 or 20 ng/mL).

Out of these factors, the ones significantly affecting taxadiene productivity were deemed to be culture temperature, which was optimized at 22°C, carbon source concentration, which was

optimized at 20g/L glucose and isoprenol supplementation, which was optimized at 10mM and the nitrogen source, the optimal being 1% tryptone supplementation. Of the above optimizations only the supplementation with a nitrogen source led to fold-increases in taxadiene titer.

As we looked towards the synthesis of volatile isoprenoids beyond taxadiene, we tried to find culture media that would work for the synthesis of a wide range of isoprenoids. In addition to un-supplemented M9 media, we investigated M9 media supplemented by 5% casamino acids and KM defined media (without coconut milk) (6). We also investigated the use of rich LB media (7), which could help us avoid the generally poor protein synthesis associated with growing in minimal media. Our product in this case was miltiradiene a molecule with applications in medicine (8).

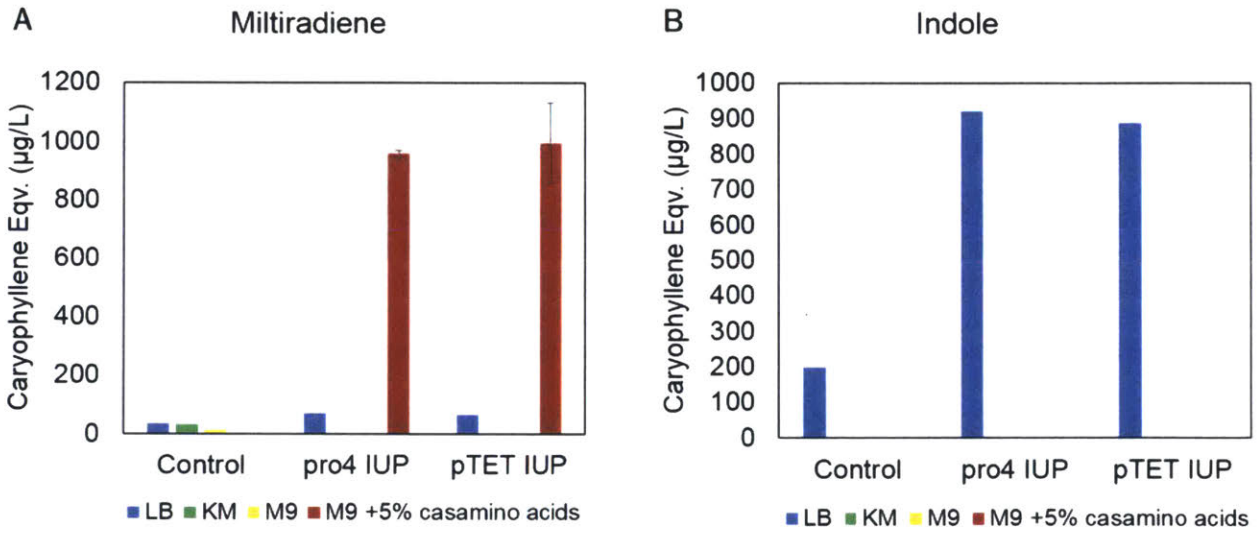


Figure 4-7: Media optimization for miltiradiene synthesis through the IUP. (**A**) Miltiradiene titers after culturing for 48h in LB, KM, M9 media or M9 media supplemented with 5% casamino acids. (**B**) Corresponding indole levels produced by the cultures. The IUP is expressed under the control of the pro4 or pTET promoters (10 ng/mL aTC). Results are compared with those achieved by a control expressing only the downstream cassette or those achievable through an upregulated MEP pathway (trcMEP). Concentrations are expressed as equivalents of the internal standard caryophyllene and are reported as means ± SD of three biological replicates.

We only observed appreciable amounts of miltiradiene produced through the IUP in the case of culture in LB or when using M9 media supplemented with 5% casamino acids. Interestingly enough, when culturing in LB we observed the formation of indole, a compound known to

accumulate in unbalanced pathways and to be detrimental to isoprenoid production (2, 10). When using any form of defined media, no indole would be observed. Given the large, 10-fold increase in titer miltiradiene observed when using M9 media with 5% casamino acids, we elected to use these media for all subsequent runs.

4.4. The IUP can produce a variety of volatile isoprenoids

To further assay the IUP's ability to produce volatile isoprenoids, we transformed our *E. coli* with an IUP plasmid (either pro4IUP or pTETIUP) and a plasmid containing downstream operons for the production of valencene, miltiradiene, and taxadiene, which we had already worked on optimizing, as well as amorphadiene and limonene.

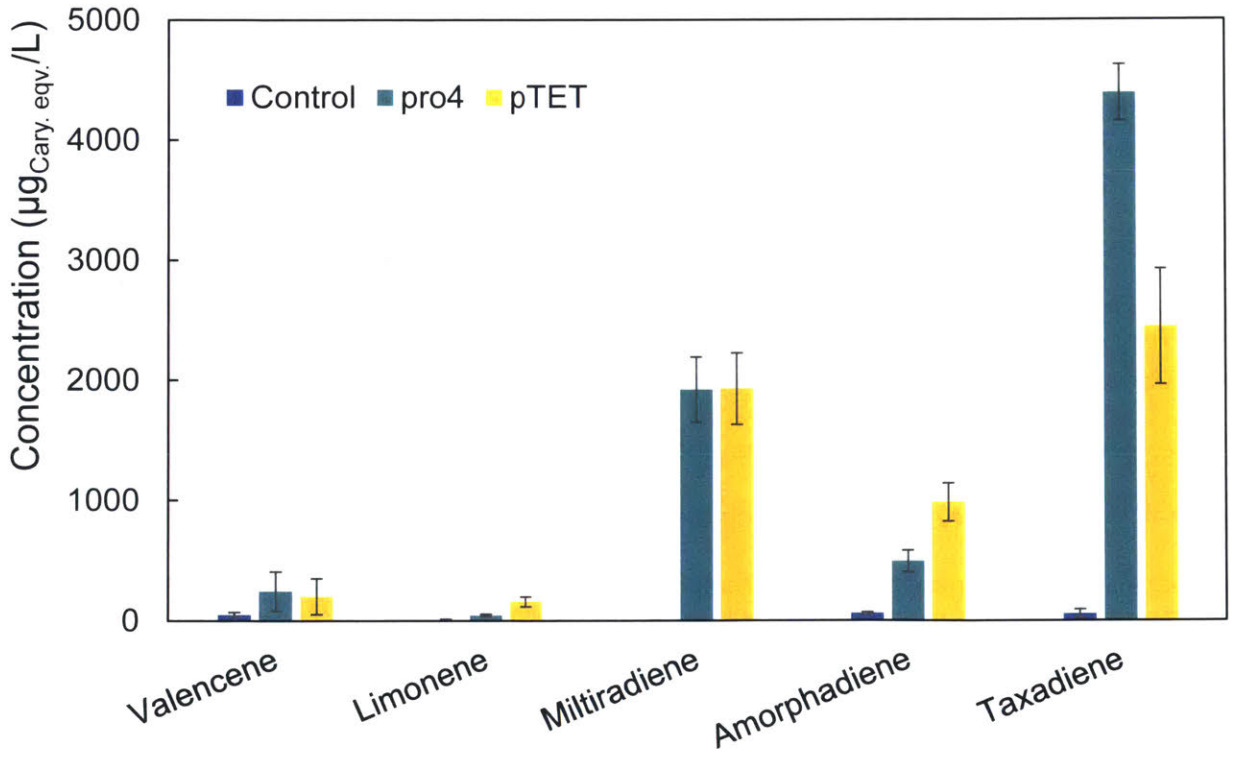


Figure 4-8: Use of the Isopentenol Utilization Pathway for the production of volatile isoprenoids. Isoprenoid product titers after culturing for 48h expressing the IUP under the control of the pro4 or pTET promoters (10 ng/mL aTC), along with a control expressing only the downstream cassette. Concentrations are expressed as equivalents of the internal standard caryophyllene.

After culturing the cells in tubes using our optimized M9 media and measuring the isoprenoid concentrations, we established that in almost all cases, valencene being the exception, expression of the IUP led to higher isoprenoid titers. We also observed that the pro4IUP plasmid produced better results, presumably due to better balancing of the upstream (up to IPP) and downstream (IPP until final product) pathway modules. We are thus able to show through our experiments that the IUP can be used as a general platform for the production of multiple isoprenoid products (Figure 4-8).

In order to quantify the contributions of both the IUP and MEP pathways and unequivocally establish that it is the IUP's function which leads to higher titers, we looked at the case of the pro4IUP taxadiene-producing strain, which was the best producer, and performed a labeling experiment similar to the ones performed before (see Sections 2.5.2 and 3.4.1).

This strain was grown in M9 media using ^{13}C uniformly labeled glucose as the sole carbon source. Upon reaching OD_{600} of 0.5, taxadiene production was induced with IPTG, and, if applicable, isoprenol was added to the media. In these cultures, the pro4IUP plasmid was present or not present (+IUP or –IUP, respectively) and either no isoprenol or 25 mM isoprenol was added (-ISO or +ISO, respectively). After 48h, the metabolic intermediates (IP/DMAP, IPP/DMAPP, GPP, FPP & GGPP) were extracted and quantified using a new LC-MS/MS we developed for this purpose, while taxadiene produced was quantified by GC-MS. As expected, in cultures without isoprenol, no IP and very low or undetectable levels of pathway intermediates IPP/DMAPP, GPP, FPP, were found. When isoprenol was supplied to the IUP strain, there was a marked increase in all pathway metabolites, with GGPP accumulating to extremely high levels (600 $\mu mol/g_{dcw} \pm 1.89$) after 48h of growth.

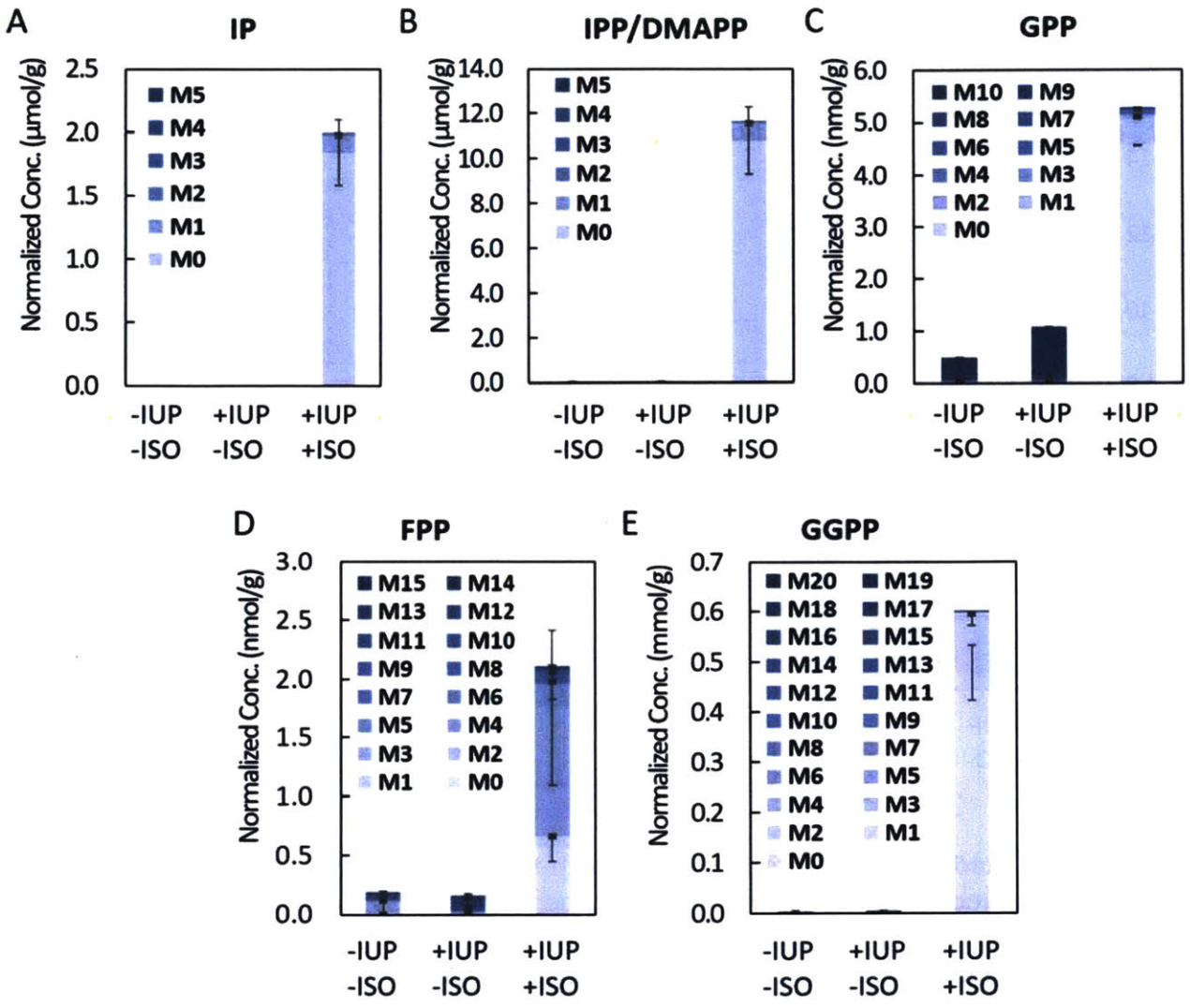


Figure 4-9: Metabolite levels from cultures with taxadiene-producing strains growing in U-^{13}C labeled glucose. The cultures differ on whether they express the IU pathway (+/- IUP) and on whether the culture media was supplemented with unlabeled isoprenol at t=0 (+/- ISO). Metabolic intermediate pools are analyzed after 48h of culture. Concentrations and labeling patterns for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively. All metabolite concentrations are reported as means ± SD of three biological replicates.

Furthermore, the taxadiene produced by the pro4IUP strain supplemented with isoprenol was 96.3% unlabeled, as confirmed by its mass spectrum. Since taxadiene and all the metabolites leading to it were close to 100% unlabeled, we could deduce that taxadiene was almost entirely produced by the IUP from unlabeled isoprenol and that the IUP handily outcompeted the MEP pathway, as had also been indicated by previous experiments (see Chapters 2.5.2 and 3.4.1).

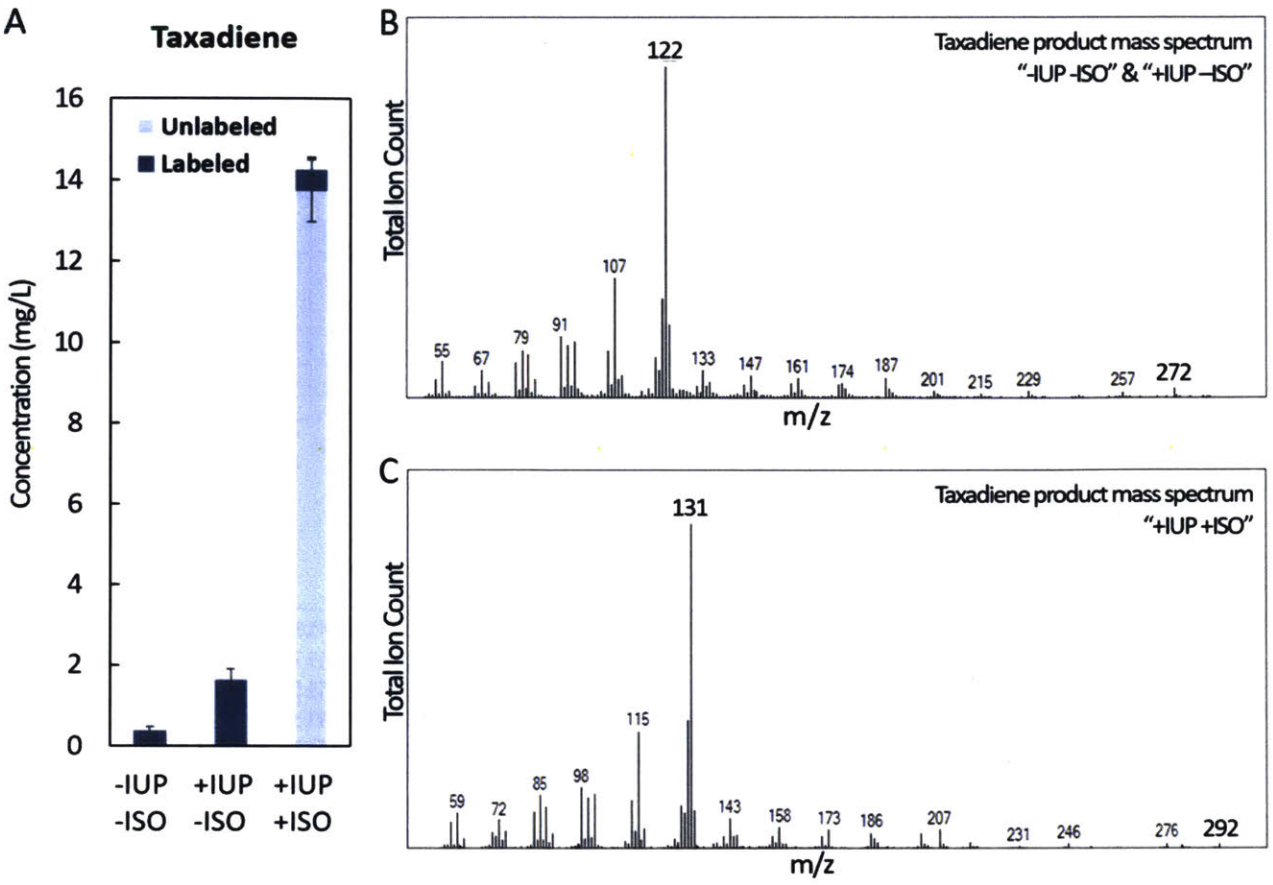


Figure 4-10: Taxadiene levels from cultures with taxadiene-producing strains growing in U-^{13}C labeled glucose. The cultures differ on whether they express the IU pathway (+/- IUP) and on whether the culture media was supplemented with unlabeled isoprenol at t=0 (+/- ISO). Taxadiene and metabolic intermediate pools are analyzed after 48h of culture. (**A**) Concentrations and labeling pattern for taxadiene. Concentrations are reported as means ± SD of three biological replicates. (**B**) Labeled taxadiene mass spectrum, (**C**) Unlabeled taxadiene mass spectrum.

4.5. Identifying bottlenecks in downstream synthesis

In order to investigate further optimization avenues, we decided to focus on the lycopene-producing strain, as the candidate for nonvolatile isoprenoids, and the taxadiene-producing strain as the best-producing candidate for volatile isoprenoids. To establish whether the upstream operon (IUP) or the downstream operons were the limiting factor in the production of these isoprenoids, we assessed the intracellular metabolites IP, IPP, GPP, FPP, and GGPP for the pro4IUP p5T7tds-ggpps (taxadiene) strain and the pro4IUP p5T7-LYCipi (lycopene) strain, i.e. all

the intermediates between isoprenol and final products, which are common in both cases. Our experiment used the same type of pulse-chase technique described earlier (Section 2.5.2).

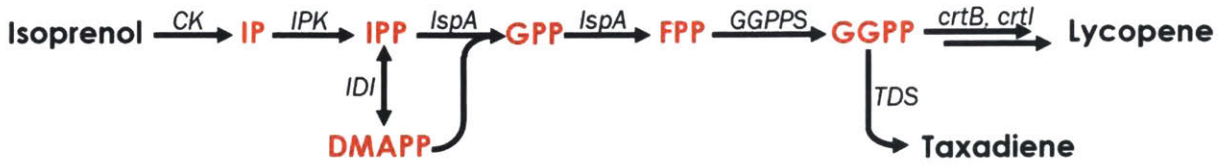


Figure 4-11: Metabolic intermediates between Isoprenol and Lycopene or Taxadiene. The metabolic intermediates are shown in bold red, enzymes catalyzing the reactions are shown in italics.

In the case of lycopene, the results (Figure 4-12A) indicated that the downstream lycopene flux was still limiting, as evidenced by the high IPP accumulation and further optimization of downstream isoprenoid production was still necessary to achieve higher titers using the IUP. The high degree of IPP accumulation suggested that enzymes downstream of this intermediate, i.e IspA, CrtE (GGPP synthase), CrtB (phytoene synthase), or CrtI (phytoene desaturase) may be limiting enzymes in lycopene synthesis.

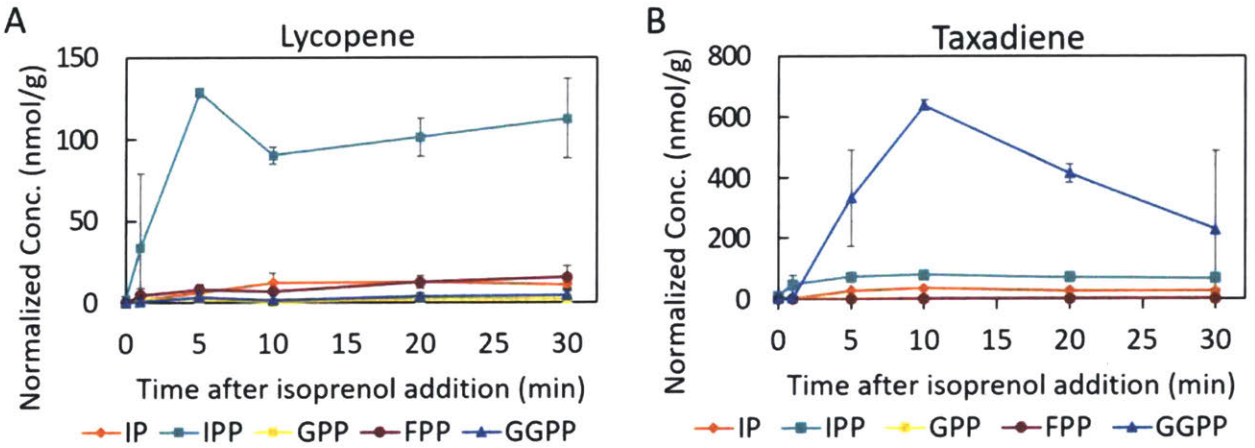


Figure 4-12: Concentrations of metabolic intermediates: Strains expressing the IU pathway along with a plasmid for the production of (**A**) Lycopene (p5T7-LYCipi) or (**B**) taxadiene (p5T7tds-ggpps) respectively are cultured and 25 mM isoprenol is added at OD_{600}=0.5 (t=0). All metabolite and product concentrations are reported as means ± SD of three biological replicates.

In the case of the taxadiene cultures (Figure 4-12B) it was easier to draw conclusions, as very high levels of GGPP were found to accumulate in the cell, suggesting taxadiene synthase (TDS) activity was insufficient to accommodate the high flux generated by the IUP and terpenoid backbone synthesis. This might also serve to explain why our optimization efforts were not able to achieve fold-increases in taxadiene productivity similar to the ones we achieved in the case of lycopene.

4.6. Debottlenecking the pathway to improve lycopene production

As is apparent from Figure 4-15, the lycopene and taxadiene strains, which were identical except for their GGPP synthases and downstream product-synthesis cassettes, exhibited different metabolite concentration profiles. For the synthesis of GGPP, whereas lycopene cultures used CrtE derived from *Enterobacter agglomerans*, taxadiene cultures employed the homolog GGPPS from *Taxus canadensis*. Based on the metabolite profiles, we came to the conclusion that the very high GGPP accumulation in the taxadiene-producing strain could likely be attributed to the *T. canadensis* GGPPS, which appeared to be a quite potent enzyme. In order to take advantage of this potency to debottleneck lycopene production, we created a new lycopene vector replacing *crtE* with *ggpps* from *T. canadensis*, to create plasmid p5T7-LYCipi-ggpps. After our initial 20mL culture results indicated that this new plasmid was superior to the original lycopene plasmid, we cultured the original strain (using *E. agglomerans* CrtE to synthesize GGPP) and the new strain (using *T. canadensis* GGPPS to synthesize GGPP) in batch bioreactors at 37°C, with the IUP being expressed through the pro4IUP plasmid and the results were compared.

We observed that glucose was depleted within 6h for the CrtE reactors and 9h in the GGPPS reactors but we did note that the GGPPS reactors started with slightly higher glucose at the time of induction (i.e. when the culture reached OD_{600}=0.5). In both cases, the optical densities were similar.

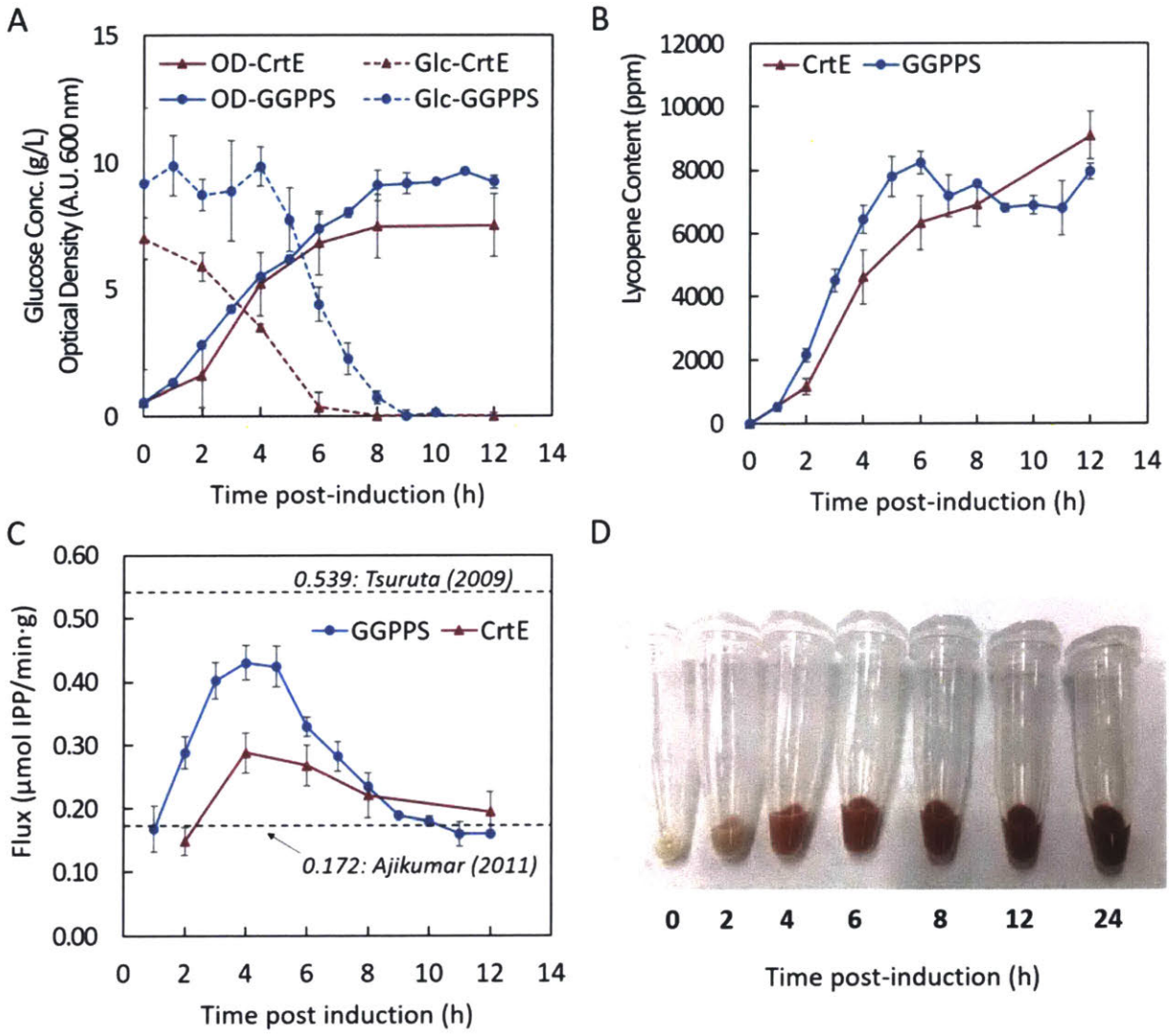


Figure 4-13: Batch bioreactor cultivation of lycopene production utilizing the IUP. The IUP was expressed under the control of the pro4 promoter along with a p5T7-LYC vector containing either *crtE* or *ggpps*. (**A**) Glucose concentration and optical density over time. (**B**) Lycopene content over time. (**C**) Cumulative IPP flux calculated from lycopene productivity and comparison to some of the highest reported isoprenoid fluxes in the literature. (**D**) Cell pellets taken from one CrtE bioreactor at different time points. All values represent the mean ± SD based on samples taken from 3 bioreactor runs.

In the case of the lycopene content, we observed that it increased until glucose was depleted, suggesting that lycopene flux is tied to active growth, with the rate of lycopene production being significantly higher in the case of the GGPPS culture compared to the CrtE culture. However, we also observed that the maximum lycopene content reached (~10,000 ppm) was similar for both

cultures and consistent with the lycopene content observed in previous cultivations in 20 mL serum bottles. This would suggest that lycopene production from the IUP is limited by the capacity of *E. coli* to store this hydrophobic molecule, which is thought to accumulate in the cell membrane (11). This is also reflected in the rate of lycopene production, which appears to slow down or plateau as the lycopene content approaches ~10,000 ppm.

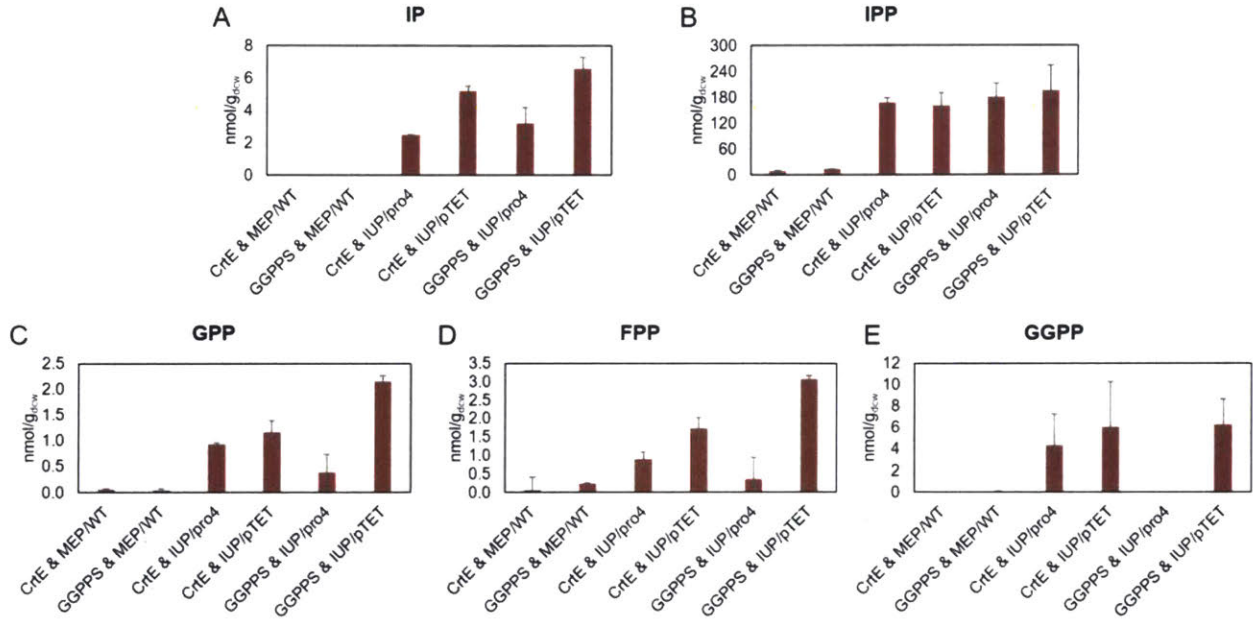


Figure 4-14: Metabolite levels from cultures with lycopene-producing strains. The cultures differ on whether they express the IU pathway via the pro4 or the pTET plasmids, or not at all (WT) and on whether the lycopene plasmid is using the *E. agglomerans* CrtE or the *T. canadiensis* GGPPS to synthesize GGPP. Concentrations for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively. All metabolite concentrations are reported as means ± SD of three biological replicates.

As a follow-up experiment, we also investigated the effect substituting CrtE from *E. agglomerans* with GGPPS from *T. canadensis* had on the metabolite profile. We observed that while the IPP levels remained relatively constant, switching from CrtE to GGPPS led to a decrease in the concentrations of GPP, FPP and GGPP in the case of the weaker expression of the IUP (i.e. the GGPPS & IUP/pro4 strain, which was the one we used in the bioreactor experiment). We also noted that precursor concentrations (especially IPP) were still substantial, indicating underutilized potential and room for further improvements and debottlenecking.

4.7. Comparison of IUP with established pathways

In order for the IUP to become a credible alternative pathway to the established engineered pathways, the IUP should be able to match or outperform said pathways. Thus, we conducted a series of comparisons to prove IUP's competitiveness.

4.7.1 Comparisons based on pathway flux

Using the data acquired in our latest bioreactor experiment, we calculated the IUP flux, expressed in IPP equivalents, based on the amount of isoprenoid produced. We noticed that the maximum flux we achieved, 0.430 μmol IPP/min·g_{dcw} in the GGPPS bioreactor run (Figure 4-13 C), far exceeded the flux achieved by our lab's best result with the MEP pathway (2) in which 1 g/L of taxadiene was produced in 5 days (flux of 0.172 μmol IPP/min·g_{dcw}), and can compete with one of the most impressive results for the MVA pathway (12), in which 25 g/L of amorphadiene was produced in 6 days (flux of 0.593 μmol IPP/min·g_{dcw}).

What is notable about our results is that they show that IUP is able to compete with either pathway, flux-wise, even though IUP itself has received little optimization, whereas in contrast the MVA and MEP pathways have been heavily engineered over decades (13). Another important point that must be made is that while in both benchmarks the cultures were conducted in tightly controlled, fed-batch bioreactors over multiple days, our experiment was much simpler in that it consisted of batch bioreactor runs that lasted less than 1 day. Therefore, we believe that further engineering of the IUP itself, as well as more sophisticated fermentation control would lead to even better IUP performances. Unfortunately, the finite lycopene storage capacity the *E. coli* membrane exhibits (11) does not allow us to properly compare isoprenoid product titers, due to the maximum capacity being reached in ~6 h, significantly inhibiting further lycopene production.

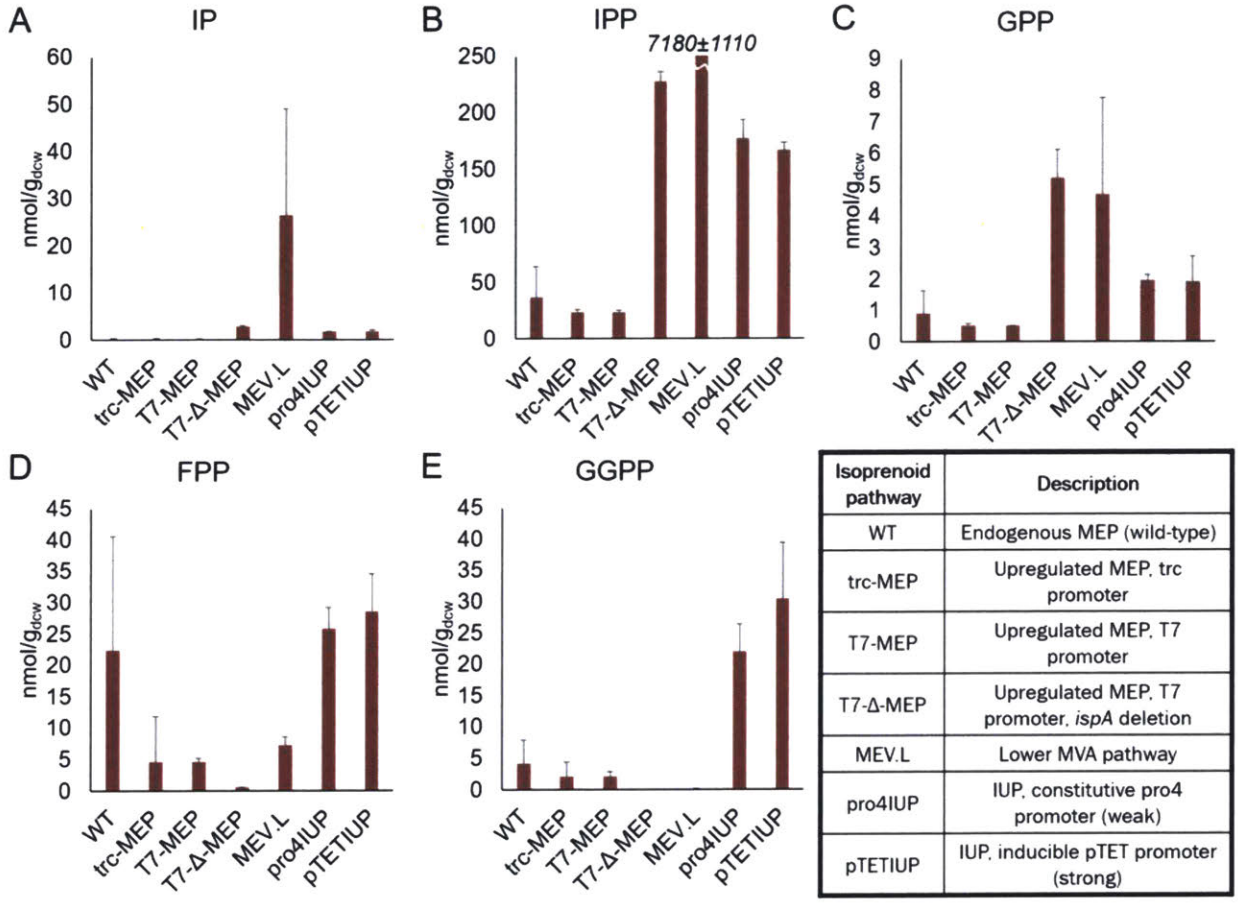
4.7.2 Comparisons based on metabolic intermediate accumulation

One way in which the potential of an isoprenoid pathway can be estimated is through directly measuring the "output" of the pathway. As a first approach, we decided to express different isoprenoid pathways in the absence of downstream module in *E. coli*, culture them in the same media and investigate the accumulation of the metabolic intermediates (IP, IPP, GPP, FPP, GGPP) produced by each pathway over time. The metabolite accumulation over a short time could then be used as a metric for the isoprenoid pathway flux, making this experiment a simplified version of the pulse-chase experiment of Section 3.4. Our key assumption in this experiment is that the rate of natural metabolic intermediate consumption is low and similar in all strains, as it only relates to essential cellular functions (e.g. maintaining membrane fluidity), which should be the same regardless of which isoprenoid pathway the metabolic intermediates are being produced from. The experiment was conducted by culturing *E. coli* strains which express different isoprenoid pathways in M9 media at 37°C. At early exponential phase (OD_{600}=0.5), cultures are supplemented with inducers and/or isoprenol or mevalonate, in the cases of the IUP or the lower MVA pathway respectively. Metabolic intermediate levels were then quantified after 30 min.

E. coli strains expressing one of the following isoprenoid pathways were cultured:

1. Endogenous MEP pathway (WT: wild type *E. coli*)
2. Upregulated MEP, integrated in genome under the trc promoter (trc-MEP) (2)
3. State-of-the-art, upregulated MEP, integrated in genome under the T7 promoter, developed by Manus Bio (T7-MEP)
4. State-of-the-art, upregulated MEP, integrated in genome under the T7 promoter, with an *ispA* deletion, developed by Manus Bio (T7-Δ-MEP)
5. Lower mevalonate pathway, constitutively expressed through use of the plasmid pBAD33-proD-MEVI in which genes *erg12*, *erg8* and *mvd1* are under the proD promoter (MEV.L)
6. IUP, constitutively expressed through use of the plasmid pSEVA228-pro4IUPi (pro4-IUP)
7. IUP, inducibly expressed through use of the plasmid pTET-IUPi (pTET-IUP)

At t=0 strains 2-4 are induced with 0.1 mM IPTG, strain 5 is supplemented with 1 mM mevalonate, strains 6&7 are supplemented with 25 mM isoprenol and strain 7 is supplemented with 20 ng/mL anhydrotetracycline.



Isoprenoid pathway	Description
WT	Endogenous MEP (wild-type)
trc-MEP	Upregulated MEP, trc promoter
T7-MEP	Upregulated MEP, T7 promoter
T7-Δ-MEP	Upregulated MEP, T7 promoter, *ispA* deletion
MEV.L	Lower MVA pathway
pro4IUP	IUP, constitutive pro4 promoter (weak)
pTETIUP	IUP, inducible pTET promoter (strong)

Figure 4-15: Metabolite levels from cultures expressing different isoprenoid pathways. Insert provides details on the 7 different isoprenoid pathways assayed. Concentrations for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively at 30min after pathway induction. All metabolite concentrations are reported as means ± SD of three biological replicates.

Our results showed that the lower mevalonate pathway (MEV.L) can clearly outperform all the other tested pathways, including the IUP, in terms of raw output, as observed by the extremely high IPP concentrations produced. However, it needs to be noted that these very high IPP concentrations led to a significant growth defect in the MEV.L strains. Furthermore, the observed concentrations of GPP, FPP and GGPP were small or nonexistent compared to the IUP.

Moreover, in contrast to the full mevalonate pathway, the lower mevalonate pathway bypasses enzyme 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGR), which catalyzes the production of mevalonate from 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA), and is known to be the main bottleneck in the pathway, as it is subject to strict regulation and can be inhibited in many ways (13). Therefore, the lower mevalonate pathway's function is dependent on mevalonate supplementation, which is expensive, rendering the pathway's industrial use not realistic.

Concerning the MEP pathways, the strains expressing the wild type MEP (WT) or the upregulated MEP pathway under the trc or T7 promoters (trc-MEP and T7-MEP) did not produce high quantities of metabolic intermediates that were higher than the ones the IUP produced. In the strain in which the upregulated MEP under the T7 promoter is combined with an *ispA* deletion (T7-Δ-MEP), significant accumulation of IPP and GPP were observed, signaling that this is indeed a strong isoprenoid pathway.

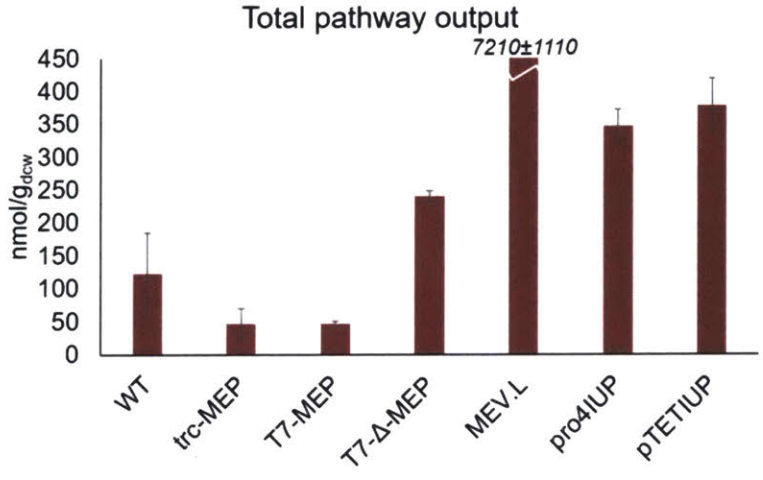
Concerning the IU pathways (pro4IUP & pTETIUP), both pathways produced significant quantities of IPP, GPP, FPP and GGPP, which were comparable to or higher than the respective metabolite concentrations produced by the other pathways tested. In particular, when compared to its closest competitor T7-Δ-MEP on productivity, IUP produced similar metabolite concentrations after 30 min. One distinguishing feature of the IUP is that its expression led to the production of significant quantities of GGPP. This is surprising, as *E. coli* is not known to produce appreciable quantities of GGPP (14) without the expression of an exogenous GGPP synthase. We hypothesize that in the case of the IUP, the high pathway flux and the high precursor (particularly FPP) concentrations lead to the production of appreciable quantities of GGPP, through Le Chatelier's principle, via some unknown mechanism that is endogenous to *E. coli*. One potential explanation for this phenomenon is that *E. coli* may naturally be able to produce GGPP in minute amounts through the activity of its native IspA, which synthesizes GPP and FPP, the GGPP precursors. This possible explanation is supported by the fact that IspA can be mutated

(via mutation Y80D) to selectively produce GGPP, thus raising the possibility of promiscuous IspA activity.

We also calculated the total output of each pathway, which we calculated by the following formula:

$$\textbf{Total pathway output} = C_{IPP} + 2 \cdot C_{GPP} + 3 \cdot C_{FPP} + 4 \cdot C_{GGPP} \quad (9.1)$$

This formula sums up all the isoprenoid metabolites produced though a pathway and normalizes them by C_5 isoprene unit/IPP. For example, since GPP is a C10 molecule produced from condensation of IPP and DMAPP, the concentration of GPP is multiplied by 2. Thus, we have an aggregate metric of all the isoprene units produced by a pathway. We can see that IUP expressed by either system (pro4IUP or pTETIUP) clearly outperforms all pathways assayed, with the exception of the lower mevalonate pathway which produces huge quantities of IPP, thus allowing us to claim that IUP is a competitive pathway.



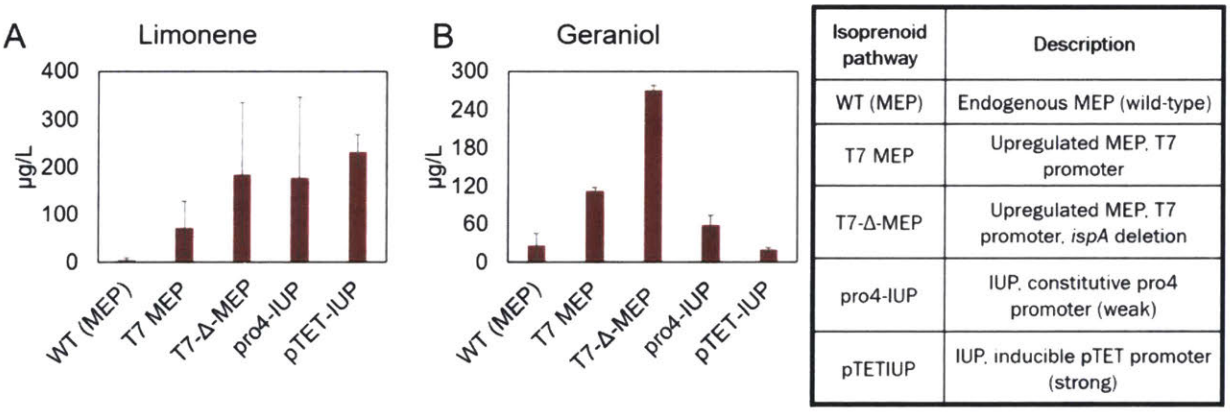
Isoprenoid pathway	Description
WT	Endogenous MEP (wild-type)
trc-MEP	Upregulated MEP, trc promoter
T7-MEP	Upregulated MEP, T7 promoter
T7-Δ-MEP	Upregulated MEP, T7 promoter, *ispA* deletion
MEV.L	Lower MVA pathway
pro4IUP	IUP, constitutive pro4 promoter (weak)
pTETIUP	IUP, inducible, pTET promoter (strong)

Figure 4-16: Total pathway output of different isoprenoid pathways. Insert provides details on the 7 different isoprenoid pathways assayed. Total pathway output is expressed as the sum of pathway metabolite concentrations normalized by number of isoprene units at 30min after pathway induction. All normalized concentrations are reported as means ± SD of three biological replicates.

4.7.3 Comparisons based on final product titer

While measuring the output of an isoprenoid pathway in terms of metabolic intermediates can be useful, measurements in the absence of a strong downstream sink tend to underestimate pathway flux. This can also be exacerbated due to the fact that high metabolic intermediate concentrations can be toxic, inhibit growth and feedback-inhibit the natural MEP and MVA pathways (15).

Given the significant metabolic intermediate accumulations generated by the IUP (and other pathways we assayed) signifying untapped potential, we decided to also compare pathway performance using final product titer as a metric. This had already been attempted before, in assays where the IUP outperformed trc-MEP in the production of valencene (Figure 4-6 A) and miltiradiene (Figure 4-7). We thus decided to also couple the isoprenoid pathways to two more downstream modules, one for the production of limonene and one for the production of geraniol and assay the pathway output in terms of product titer.



Isoprenoid pathway	Description
WT (MEP)	Endogenous MEP (wild-type)
T7 MEP	Upregulated MEP, T7 promoter
T7-Δ-MEP	Upregulated MEP, T7 promoter, *ispA* deletion
pro4-IUP	IUP, constitutive pro4 promoter (weak)
pTETIUP	IUP, inducible pTET promoter (strong)

Figure 4-17: Production of Limonene and Geraniol through different isoprenoid pathways. Insert provides details on the 5 different isoprenoid pathways assayed. (**A**) Limonene or (**B**) geraniol (C) titers after culturing for 48h in M9 media at 37°C. Concentrations are reported as means ± SD of three biological replicates.

Results differed in the case of limonene and geraniol. While in the case of limonene, the IUP strains produced titers of comparable concentration to the ones produced by the T7-Δ-MEP strains, in the case of geraniol the T7-Δ-MEP strain clearly outperformed the IUP ones. The

relative differences in pathway performance can probably be traced to the differences in the downstream module performance (limonene vs geraniol synthases).

These inconclusive results highlight the importance of proper pathway balancing to avoid bottlenecks, and maximize titer. In order to properly assay the pathway potential, the IUP (or any isoprenoid pathway) should be tied to a strong downstream module which can convert pathway potential to isoprenoid product. We therefore recognized that further work was required to identify the true limits of the IUP and attempt to surpass them.

4.8. Conclusions and future directions

In this chapter we have investigated how the IUP can be coupled with downstream pathways for the production of isoprenoids. In order to improve final product titer through the IUP, we looked into a number of factors; a summary of the results is given in Table 4-1.

After optimization, demonstrated that the IUP is a pathway that can be used for the production of a variety of isoprenoids, both volatile and non-volatile and can do so at significantly higher titers than the *E. coli* endogenous MEP pathway. We also demonstrated, through labeling experiments, that titer increases are wholly attributable to the IUP.

Furthermore, we demonstrated that the IUP can sustain fluxes that approach or exceed some of the best reported in literature, compete with engineered state of the art MEP or MVA pathways, and sometimes compare favorably against them by producing metabolic intermediates at high concentrations and rates.

As was demonstrated, due to the high flux of the IUP makes it possible to improve product titer and productivity merely by improving the downstream pathway, something which we successfully achieved in the case of lycopene. Further work is required in the case of taxadiene, as our efforts did not achieve significant improvements. Coupling the IUP with a high-flux

downstream could be the ideal way of identifying the IUP's limits, as a first step of further improving it.

Table 4-1: Factors tested for the optimization of isoprenoid synthesis via the IUP

Factor Tested	Effect
Synthesis of lycopene via the IUP	
Change in carbon source (glucose vs glycerol) Change in glucose concentration Change in trace element solution Culturing in BAF vs M9 media Change in timing of isoprenol addition Change in culture temperature (30°C vs 37°C)	No significant effect
Increase in downstream plasmid copy number	>5-fold titer increase
Change in downstream plasmid promoter	~6-fold productivity increase
Synthesis of taxadiene via the IUP	
Increase in downstream plasmid copy number	No significant effect
Increase in IUP expression (via stronger promoter)	No significant effect
Decrease in culture temperature (to 20°C) Increase in carbon source concentration (20g/L) Decrease in isoprenol concentration (to 10mM)	<2-fold increase in titer
Changes in concentration for inducers for downstream (IPTG) or upstream (andhydrotetracycline) modules Timing for IPTG or anhydrotetracycline induction Change in carbon source (glucose vs glycerol)	No significant effect
Supplementation of M9 media with 1% tryptone	~2-fold increase in titer
Synthesis of miltiradiene via the IUP	
Supplementation of M9 media with 5% casamino acids	~10-fold increase in titer

Further work is also required on establishing the yield (on isoprenol) for the pathway, especially in minimizing the error of the measurement. This could be achieved by running the IUP cultures in sealed serum bottles to avoid isopentenol evaporation and sampling the isoprenol concentration directly using a headspace sampler and sending the sample directly to a GC.

4.9. Materials and Methods

Strains, plasmids and genes; cloning protocol

Materials and methods were the same as the ones as described in the relevant sections of Chapter 2. Strains T7-MEP and T7-Δ-MEP were a gift from Manus Biosynthesis (MIT MTA 8234).

Construction of expression vectors.

All vectors were constructed following the routine cloning protocol described in Chapter 2. Vectors pSEVA228-pro4IUPi and pTET-IUPi for the expression of the IUP pathway were constructed by the method described in Chapter 2. Plasmid pBAD33-proD-MEV for the expression of the lower mevalonate pathway was created by inserting a point mutation in plasmid pBAD33-proA-MEV through the use of the primer pair ProD Mut_f /ProLibrary_Mut_r. The vectors for the expression of the downstream isoprenoid modules were constructed as follows: The copy number of the lycopene plasmids were varied by first amplification of the genes *crtE crtI*, *crtB* and *ipi* as well as the endogenous lycopene promoter using primers GB_pAC-LYCipi_f/r from the pAC-LYCipi plasmid. The fragment containing the gene *aadA1* (spectinomycin resistance) was amplified using GB_aadA1_f/r from p5T7tds-ggpps and the origins pUC19 and pBR322 were amplified using GB_pUC19_f/r and GB_pBR322_f/r from pUC19 and pET28a respectively. These fragments were assembled with the appropriate origin to create pUC-LYCipi and p20-LYCipi. To create p5T7-LYCipi, the backbone of p5T7tds-ggpps was amplified using GB_p5T7_f/r and the lycopene synthesis genes were amplified from p20-LYCipi using the primers GB_p20-LYCipi_f/r and then assembled. To make p5T7-LYCipi-ggpps, the backbone of p5T7-LYCipi was amplified using primers p5T7Lyc-back_f/r and the *ggpps* was amplified from p5T7tds-ggpps using p5T7Lyc-ggpps_f/r and then assembled. p5T7gpps-ls and p5T7ispA-ads were created by PCR amplification of the p5T7tds-ggpps vector using primers GB_p5t7ggppstds_f/r to create the backbone from the T7 terminator to the T7 promoter. The primers GB_gpps_ls_f/r were used to amplify the *gpps-ls* operon from JBEI-6409 for Gibson Assembly in this backbone created the p5T7-gpps-ls vector. Primers GB_ispA_f/r and GB_ads_f/r (with RBS encoded on the primer) were used to amplify *ispA* from p5T7vs-ispA and *ads* from pADS respectively. These two fragments were assembled into the same backbone as *gpps-ls* to create the p5T7-ispA-ads vector. Plasmid p20T7tds-ggpps was created by amplifying the origin pBR322 from pET28a, amplifying the backbone, promoter and expression cassette of p5T7tds-ggpps using primer pair

GB_p5t7tds-ggpps_f/r and assembling the resulting amplification products. pAC-LYCipi was a gift from Francis X Cunningham Jr ((16), Addgene plasmid # 53279) pADS was a gift from Jay Keasling ((17), Addgene plasmid # 19040). JBEI-6409 was a gift from Taek Soon Lee ((18), Addgene plasmid # 47048). Plasmid p5-T7-LS-L-GPPS for limonene production and plasmid p5-T7-GES-L-GPPS for geraniol production were gifts from Manus Biosynthesis (MIT MTA 8234). Vector sequences for the primers used are listed in **Table A3** in the Appendix. Codon-optimized sequences are listed in **Table A5** in the Appendix.

Quantification of lycopene.

Lycopene content was quantified using the protocol described in Chapter 3.

Quantification of volatile isoprenoids.

Volatile isoprenoids were quantified using GC-MS using ultra-pure helium as the carrier gas. First, the C18 resin used to capture the isoprenoids was vacuum-filtered from the cells and culture media using BioSpin columns (Bio-rad). The resin was then spun at 1000 x g to remove residual water, then eluted in ethyl acetate containing 36 mg/L caryophyllene as an internal standard which allowed a 20-fold concentration of the isoprenoid for quantification. The 1 µL of the eluted isoprenoid was quantified on a HP-5 MS UI capillary column (30m, 250 µm, 0.25µm) (Agilent Technologies) using a 7890B Series GC and a 5977B MS. Chromatography was performed under the following conditions: splitless injection, inlet temperature 280°C, constant inlet pressure 115.8 kPa, valve temperature 300°C, and MS transfer line 300°C. An oven program of 100°C, hold 1min, 15°C/min until 200°C, hold 2 min, 30°C/min until 250°C, hold 1 min, and 30°C/min until 290°C, hold 2 min was used for determination of taxadiene, miltiradiene, valencene, and amorphadiene. Limonene was separated using an oven program of 80°C, hold 3min, 10°C/min until 140°C, hold 2 min, 45°C/min until 290°C, hold 1 min. The MS was operated at an ion source temperature of 280°C, and a quadrupole temperature of 180°C. Ions were scanned between a mass of 40 to 400 at 1.562 u/s. Taxadiene was quantified using a standard curve based on the

m/z 122 ion which has the greatest abundance in unlabeled taxadiene. The 131 m/z ion was used to quantify labeled taxadiene using the same standard curve generated from purified unlabeled taxadiene. Taxadiene was purified using a semi-preparative HPLC using a Supelco Discovery C18 (25 cm, 10 mm, 5 um) column under isocratic conditions, 89% acetonitrile in water at 8 mL/min on a Shimadzu LC-2AD HPLC (19) equipped with a SPD-M20A diode array set at 210 nm. The fractions containing taxadiene as confirmed by GC-MS were collected using a fraction collector, pooled and recovered by rotary evaporation on a Buchi Rotavapor R-210. The purified taxadiene was weighed and resuspended for generation of the standard curve.

Quantification of metabolites.

IP/DMAP and IPP/DMAPP were quantified by LC-MS/MS by comparison to synthetic IP/DMAP made in-house according to the procedure described in Chapter 2 and IPP & DMAPP standards purchased from Sigma-Aldrich. An alternative method using the same instrumentation as the method described in Chapter 2 was used for the simultaneous detection and quantification of IP/DMAP, IPP/DMAPP, GPP, FPP and GGPP: An Xbridge C18 column (150mm, 3.5 µm, 2.1 mm) (Waters) was operated with a mobile phase of 0.1% v/v TBA, 0.12% v/v acetic acid, and titrated with ~0.5% v/v 5N NH_4OH until a pH of 8.5 was reached (A). The elutant was 100% acetonitrile (B). A series of linear gradients: 0-5 min 0% B, 5-20 min 0-65% B, 20-25 min 65% B, 25-30 min 100% B, 30-35 min 100% B, 35-36 min 100-0% B, 0% B until 45 min, was used to separate these analytes which were then compared to standard curves generated using standards purchased from Sigma-Aldrich and/or Cayman Chemicals. Samples were resuspended in the aqueous mobile phases described above.

Cultivation in serum bottles.

Cultivation in serum bottles was conducted under the same conditions described in Chapter 3, with the following changes, listed below: M9, rich LB or KM defined media was used, as described in the main text. M9 media was prepared as in Chapter 3, rich LB was prepared as in (7) and KM

defined media was prepared as in (6) omitting coconut milk from the recipe. When strains containing downstream operons for volatile isoprenoid (limonene, amorphadiene, valencene, miltiradiene, and taxadiene) production were used, 100 µL of C18 flash resin (VWR) was added to the cultures at induction time to capture these products. Strains for lycopene and amorphadiene production were grown as 37°C, otherwise all cultures were performed at 30°C. Metabolites, lycopene and volatile isoprenoids were quantified using protocols described above.

Cultivation in bioreactors.

The strains p5T7-LYCipi, p5T7-LYCipi-ggpps, and p5T7tds-ggpps with pro4IUP were cultivated in a 3-L Bioflo 110 bioreactor (New Brunswick) with aeration, agitation, and pH control. One and a half liters of defined media (M9 salts, 5 g/L casamino acids, ATCC trace elements solution, 100µL antifoam 204, and 50 µg/mL spectinomycin and kanamycin) was inoculated at 1% v/v with an overnight culture (12 h) grown in LB media. Aeration (0.3-1 vvm) and agitation (250-1250 rpm) was controlled by a cascade to maintain dissolved oxygen at 40% of saturation. pH was controlled by addition of 25% v/v NH_4OH. Temperature was controlled at 37°C for lycopene cultures and 30°C for taxadiene cultures. When an OD of 0.5 was reached, 1.5mL of 0.1 M IPTG and 3.75 mL of isoprenol were added. For taxadiene cultures, temperature was reduced to 22°C after induction. Cell density was monitored by UV/Vis spectroscopy at 600 nm, while glucose consumption was determined by HPLC using a Aminex HPX-87H column (300 x 7.8 mm) (Bio-rad) on an Infinity 1260 series HPLC (Agilent) at a flow rate of 0.7 mL/min with 14 mM H_2SO_4, at room temperature using a refractive index detector set at 50°C. C18 flash resin was added to taxadiene bioreactors to capture taxadiene and eluted in acetonitrile for purification by semi-preparative HPLC as previously described (19).

4.10. Bibliography

1. Woolston BM, King JR, Reiter M, Van Hove B, Stephanopoulos G (2018) Improving

formaldehyde consumption drives methanol assimilation in engineered E. coli. *Nat Commun* 9(1). doi:10.1038/s41467-018-04795-4.

2. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* 330(2010):70–74.
3. Yoon SH, et al. (2009) Combinatorial expression of bacterial whole mevalonate pathway for the production of β-carotene in E. coli. *J Biotechnol* 140(3–4):218–226.
4. DSMZ BAF Medium Composition. Available at: https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium392.pdf.
5. Davis JH, Rubin AJ, Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* 39(3):1131–1141.
6. Kao KN, Constabel F, Michayluk MR, Gamborg OL (1974) Plant protoplast fusion and growth of intergeneric hybrid cells. *Planta* 120(3):215–227.
7. Bertrani G (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *J Bacteriol* 62(3):293–300.
8. Zhou YJ, et al. (2012) Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J Am Chem Soc* 134(6):3234–3241.
9. Satoh K, et al. (2019) Valencene as a naturally occurring sesquiterpene monomer for radical copolymerization with maleimide to induce concurrent 1:1 and 1:2 propagation. *Polym Degrad Stab* 161:183–190.
10. Yadav VG, De Mey M, Giaw Lim C, Kumaran Ajikumar P, Stephanopoulos G (2012) The future of metabolic engineering and synthetic biology: Towards a systematic practice. *Metab Eng* 14(3):233–241.
11. Wang GS, Grammel H, Abou-Aisha K, Sägesser R, Ghosh R (2012) High-level production of the industrial product lycopene by the photosynthetic Bacterium rhodospirillum rubrum. *Appl Environ Microbiol* 78(20):7205–7215.
12. Tsuruta H, et al. (2009) High-level production of amorpha-4, 11-diene, a precursor of the

antimalarial agent artemisinin, in Escherichia coli. *PLoS One* 4(2):e4489.

13. Ward VCA, Chatzivasileiou AO, Stephanopoulos G (2018) Metabolic engineering of Escherichia coli for the production of isoprenoids. *FEMS Microbiol Lett* 365(10):1–9.
14. Reiling KK, et al. (2004) Mono and diterpene production in Escherichia coli. *Biotechnol Bioeng* 87(2):200–212.
15. Kang A, Meadows CW, Canu N, Keasling JD, Lee TS (2017) High-throughput enzyme screening platform for the IPP-bypass mevalonate pathway for isopentenol production. *Metab Eng* 41(March):125–134.
16. Cunningham FX, Lee H, Gantt E (2007) Carotenoid biosynthesis in the primitive red alga Cyanidioschyzon merolae. *Eukaryot Cell* 6(3):533–545.
17. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol* 21(7):796–802.
18. Alonso-Gutierrez J, et al. (2013) Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production. *Metab Eng* 19:33–41.
19. Edgar S, et al. (2016) Mechanistic Insights into Taxadiene Epoxidation by Taxadiene-5α-Hydroxylase. *ACS Chem Biol* 11(2):460–469.

Chapter 5

Identifying the current maximum productivity of the IUP and improving upon it

5.1. Introduction

As suggested by our pathway results so far (Chapters 3 & 4) IUP is capable of producing the main isoprenoid intermediates IPP & DMAPP with high flux, as evidenced by the results of the pulse-in experiments, as well as the very encouraging lycopene production results, which show that the IUP is competitive with some of the best reported pathways. Furthermore, our results on metabolic intermediate accumulation show that when coupling the IUP with a downstream isoprenoid synthesis pathway, there is significant accumulation of IPP/DMAPP or metabolic intermediates after IPP/DMAPP (e.g. GGPP), meaning that the bottlenecks lie downstream of the two IUP steps. This hints that we have not yet reached the limits of the IUP. We therefore began work on overcoming limitations in the downstream modules we used so far, as well as exploring the maximum limits of IUP, as it currently stands, with the intention of pushing beyond them.

5.2. Overcoming limitations in the downstream modules

As mentioned earlier, our experiments revealed two major limitations in IUP performance in both the case of lycopene and taxadiene synthesis, i.e. the two downstream modules we explored the most. In the case of lycopene, we were not able to achieve lycopene contents above 10,000ppm ($\mu g/g_{dcw}$) (Section 4.5), whereas in the case of taxadiene, we had identified taxadiene synthase

as the clear bottleneck (Section 4.4). In both cases, we pursued strategies aimed at overcoming these limitations.

5.2.1 Improving maximum lycopene titer

Our hypothesis for our inability to achieve lycopene content above 10,000 ppm hinged around the concept that *E. coli* membrane has a finite capacity to store lycopene (1) and it was possible that the membranes in our strains were reaching saturation.

In order to overcome this capacity issue, we looked into utilizing different *E. coli* strains, which had mutations/knockouts in genes connected with membrane function and could exhibit hypervesiculation phenotypes. The thinking behind this decision lay in the hypothesis that cells that produce membrane vesicles may be capable of storing additional lycopene in said vesicles, thereby increasing the effective membrane capacity for lycopene.

The strains we looked into were the following:

1. JW0728-1 (2): This single-gene knockout strain cannot produce TolR, a protein which plays a role, through the Tol-Pal system, in outer membrane invagination during cell division and is important for maintaining outer membrane integrity (3).
2. JW1667-5 (2): This single-gene knockout strain is missing Lpp, a lipoprotein which interacts with the peptidoglycan in an interaction that contributes to the maintenance of the structural and functional integrity of the cell membrane (4).
3. JE5505 (5): This strain has a mutation in the gene responsible for the production of Lpp, effectively making the cell unable to produce it.
4. TPS30 (6): This strain has mutations in the *fii* gene, the product of which interacts with the bacterial membrane and is associated with the TolA & TolB proteins (of the Tol-Pal system required for E. coli membrane integrity (7)).

Each of these strains, was transformed with p5T7-LYCipi-ggpps (the improved lycopene plasmid) and cultured for 24h in M9 media supplemented with 5% casamino acids, the optimized minimal

media we settled on using (Chapter 4). Unfortunately, we failed to see any improvement in lycopene when compared to our control, strain MG1655-DE3, which we have been using in all our experiments. Conversely, we observed significantly lower titers, as directly was evident by both the color of the colonies.

Figure 5-1: Evaluation of different *E. coli* strains for lycopene production: Vials with cell pellets of different knockout E. coli strains which have been transformed with a plasmid for lycopene production (red color) and cultured in M9 media with casamino acids for 24h at 37°C. The four knockout strains (JW0728-1, JW1667-5, JE5505 and TPS30) were compared against wild-type E. coli (MG1566) as the control.

5.2.2 Improving the taxadiene synthase enzyme

As has already been mentioned (Section 4.4), the IUP produces isoprenoid intermediates at rates greater than what the taxadiene synthase we have used can consume, indicating a clear bottleneck at this enzyme (Figure 4-15). This suggests improving taxadiene synthase to increase its activity and make use of the existing IUP potential.

Our previous efforts of increasing taxadiene productivity by increasing taxadiene synthase copy number or the IUP expression strength had not been successful (Figure 4-6). We created a different strategy for mutagenizing taxadiene synthase and screening for more active mutants based on a concept previously explored in literature to create a colorimetric screening assay (8). This approach involves the construction of an expression vector for the simultaneous co-expression of two different operons for synthesis of two different isoprenoid products. The first operon is used for the production of a colored product, which in our case we chose the red-colored lycopene, given our prior experience with it. The second operon is responsible for the production

of a colorless product, which in our case is taxadiene. Since both isoprenoid synthesis pathways draw from the same pools (i.e. the isoprenoid precursors), each operon is competing with the other for resources. Therefore, the ratio of lycopene/taxadiene product flux should be directly proportional to the relative activities of the operons. Since one product is colored (red) and the other is not, we should theoretically be able to visually distinguish which of the two isoprenoids, lycopene (red) or taxadiene (colorless), is preferentially produced in a strain by observing changes in the color of each colony.

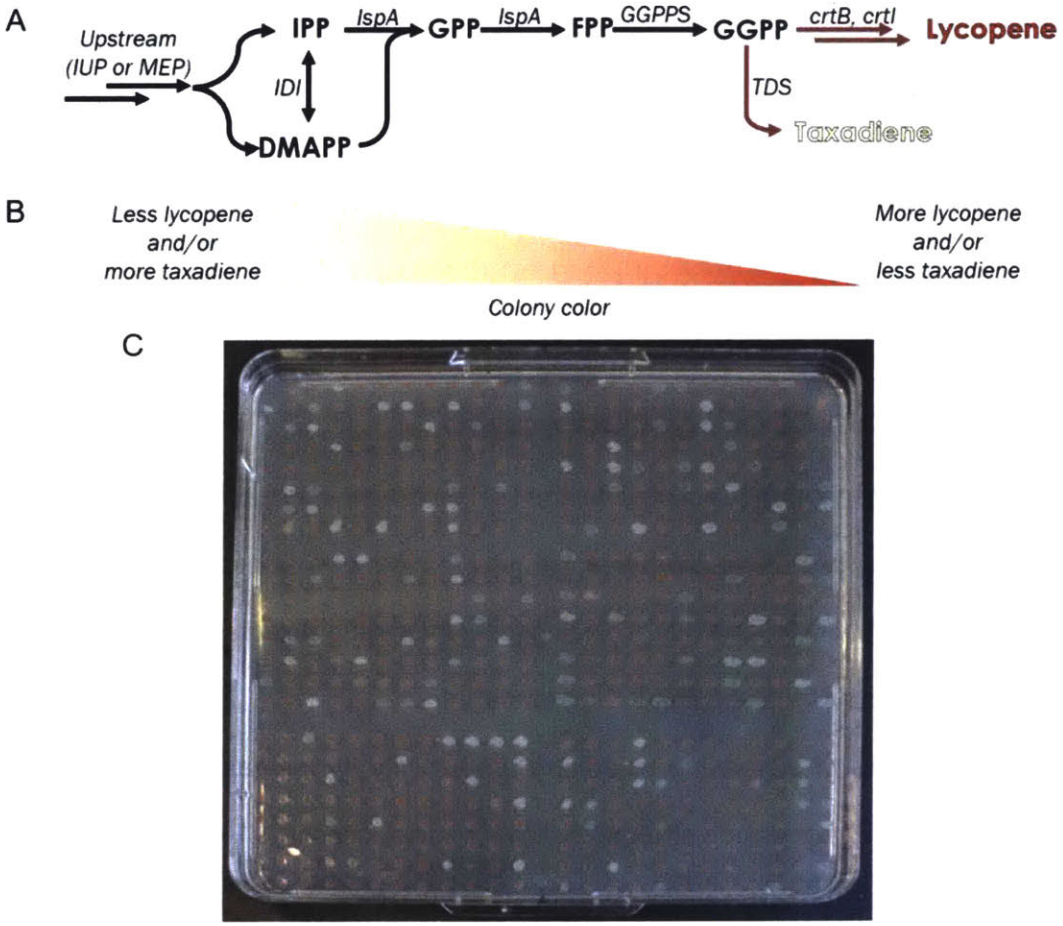


Figure 5-2: Colorimetric screen for the improvement of taxadiene synthase: (**A**) A strain can be made to simultaneously produce both the red-colored lycopene and the colorless taxadiene through the expression of an upstream isoprenoid pathway, genes *IDI*, *IspA* and *GGPPS*, which produce the common metabolic intermediates and genes *crtB* & *crtI* for lycopene production and *TDS* for taxadiene production. (**B**) Strains that produce both lycopene and taxadiene produce colonies whose color is indicative of the taxadiene and/or lycopene content. (C) Example of a plate with colonies simultaneously expressing lycopene synthesis genes and mutated taxadiene synthase.

The experimental strategy followed involves mutating taxadiene synthase through random mutagenesis, while leaving the lycopene expression operon intact on the expression vector. Thus a library of vectors which only vary on the taxadiene synthase gene sequence is created. By transforming *E. coli* strains with this vector library, a library of strains which display varying intensity of taxadiene production and therefore different relative fluxes towards taxadiene synthesis vs lycopene synthesis is created. By screening for white/less red colonies, which should be producing more taxadiene and less lycopene, potential candidates for more potent taxadiene synthases can be identified. After culturing said colonies and quantifying their taxadiene titers, the plasmids of the best producers can be extracted and sequenced to find the most active taxadiene synthase. Multiple rounds of mutagenesis can be carried out as required in order to get sufficient improvements in activity.

In our experimental strategy, we chose to transform the library of plasmids in a trc-MEP strain, i.e. one that carries genes for the upregulation of the MEP pathway in its genome under the trc promoter (9) in order to ensure an increased IPP flux that would lead to appreciable amounts of taxadiene and/or lycopene. We chose the trc-MEP strain instead of a strain that expressed the IUP due to the difficulty of providing proper isoprenol supplementation in solid media, given isoprenol's volatility. By transforming the trc-MEP strains with the plasmid library, we generated ~400 unique colonies. Out of those, we selectively chose the white colonies, cultured them in M9 media, measured their taxadiene titers and compared them to the taxadiene titer achievable using the vector bearing the non-mutated taxadiene operon (control).

Many of the white colonies, when cultured, produced taxadiene levels significantly higher that the taxadiene produced by the control strain, even 130-fold higher. After extracting the plasmids from the 8 best producing strains and sequencing the taxadiene synthase gene, we found that some of the plasmids had accrued mutations in the areas adjacent to the taxadiene synthesis operon, namely in the beginning of the lycopene synthesis operon, potentially impacting lycopene synthesis instead of enhancing taxadiene synthesis.

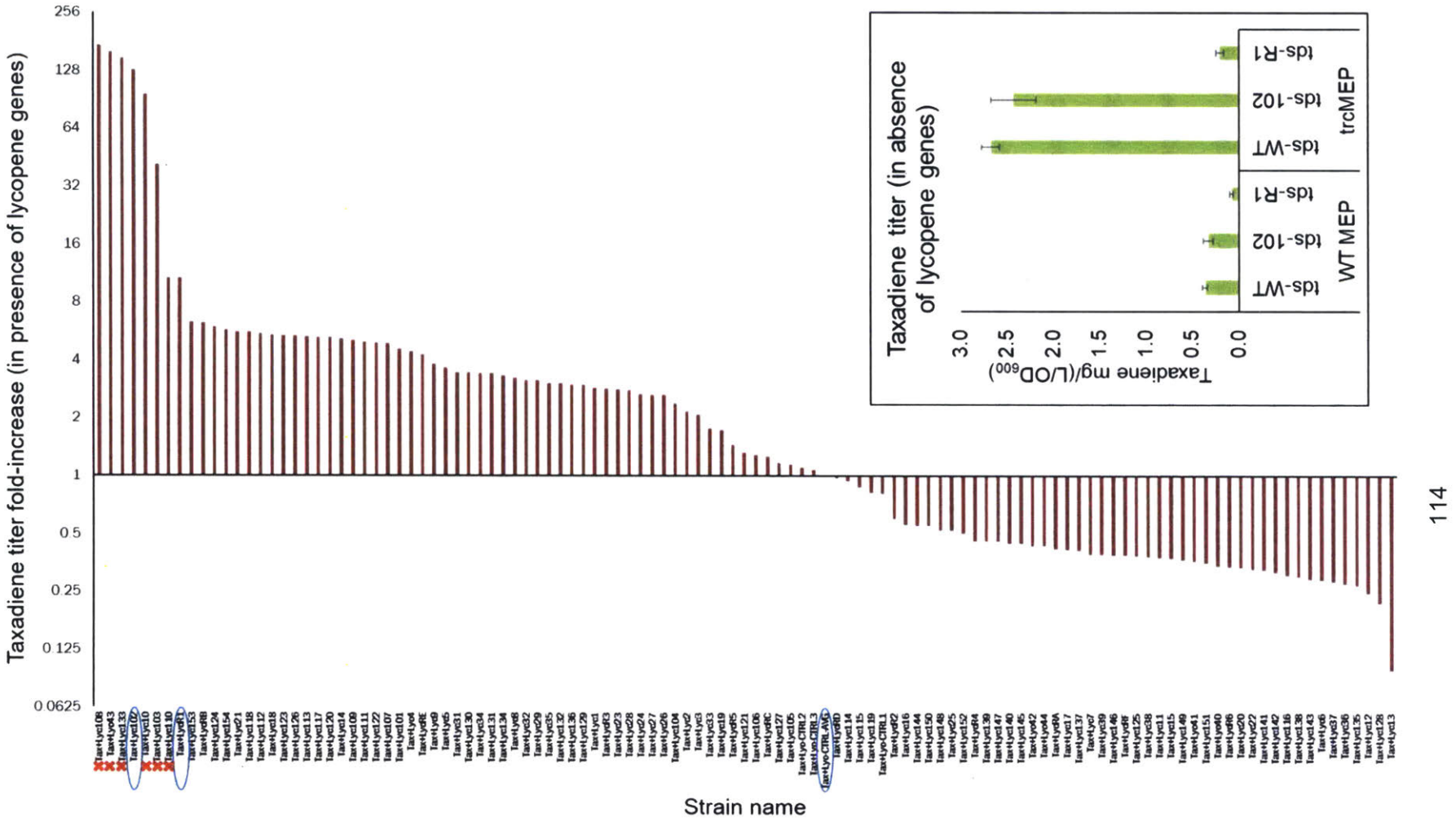


Figure 5-3: Taxadiene titers achievable by strains with mutated taxadiene synthases: Strains were grown in M9 media with +5% casamino acids at 37°C for two days and had the levels of taxadiene produced quantified. (**Main picture**) Strains bore a plasmid expressing two operons, one for lycopene synthesis and one for taxadiene synthesis. The taxadiene synthase gene has been mutated to produce different strains. All titers are normalized against the titer produced by the non-mutated version (Tax+Lyc CTRL AVG). Some strains were rejected (red X), circled are the strains bearing taxadiene synthase variants chosen for the next stage. (**Insert**) Next stage: Strains expressed either a non-mutated taxadiene synthase (tds-WT) or a mutated version of the gene (tds-102 and tds-R1), and produced isoprenoid precursors through either the non-upregulated MEP pathay (WT MEP) or an upregulated version of the MEP pathway (trcMEP).

We identified two promising taxadiene synthase mutants tds-102 and tds-R1, in which we detected a small number of mutations in taxadiene synthase namely A64R in tds-102 and Q661L in tds-R1. The mutated taxadiene synthase genes were amplified and assembled in production vectors (without lycopene synthesis operon) for the purpose of comparing them with the non-mutated taxadiene synthase (tds-WT). These vectors were then transformed in either wild-type *E. coli,* expressing the endogenous non-upregulated MEP pathway or in the trc-MEP strain expressing the upregulated MEP pathway.

We found that neither mutant managed to produce taxadiene at higher titer than the original non-mutated taxadiene synthase in either case, with tds-102 producing similar titers to the non-mutated version of taxadiene synthase and tds-R1 consistently underperforming. Results from a subsequent experiment with the IUP as the upstream were similar to the trc-MEP case in that neither taxadiene synthase mutant outperformed the non-mutated version.

5.2.3 Possible alternative approaches

Concerning the improvement of lycopene accumulation, it is suggested that more work is carried out to examine potential *E. coli* strains which could exhibit hypervesiculating phenotypes. Another way to overcome this limitation would be to express the IUP and the lycopene synthesis pathway in a different organism with higher storage capacity, something which we have already achieved in *Y. lipolytica* (10), which is able to store lycopene in its hydrophobic lipid bodies. An avenue that is being pursued by our collaborators involves increasing *E. coli* capacity by simultaneously producing hydrophobic molecules in tandem with lycopene, thereby increasing its storage capacity.

In the case of our work on improving taxadiene synthase, one potential explanation of the discrepancies exhibited between the results given by simultaneous lycopene & taxadiene production vs the production of taxadiene alone could be that in the former case there might have been some unidentified mutations in the lycopene operon in areas that we did not sequence, i.e.

far from the taxadiene operon, thus decreasing lycopene production (and shifting more resources towards taxadiene production), thereby confounding our results during our initial screen and misidentifying tds-102 and tds-R1 as potential candidates. It is possible that taxadiene synthase candidates that appeared less active than tds-102 and tds-R1 may actually be good candidates, so testing more candidates may be a worthwhile effort. Furthermore, conducting a new round of mutagenesis on taxadiene synthase, while taking greater care to minimize off-target mutations may yield better results.

5.3. Establishing the maximum limit of the pathway

As has already been mentioned earlier, even though the IUP has been subjected to very little optimization, namely optimization of feeding conditions (25mM isoprenol) and the realization that absence of IPK adversely affects the pathway, we have so far been unable to reach the pathway's limits, as evidenced by the high IPP accumulation.

In most experiments reported in Chapters 3 & 4, the IUP was expressed using the pro4IUP plasmid, which has a low copy number (~2-3) and uses a relatively weak constitutive promoter. We also noticed that even if a stronger expression system is used, such as pTET-IUP plasmid, only modest improvements (if at all) in flux or final titer are observed, even though the pTET-IUP expression system has been assayed, through GFP expression, as being at least an order of magnitude stronger than pro4IUP (Figure 2-11). This may partly be due to the fact that high IPP accumulation can lead to toxicity effects (11) preventing us from achieving proportionally higher flux increases. Thus, the most tractable way of exploring the limits of the IUP would be by coupling the pathway with a very active downstream module.

5.3.1 Using farnesene synthesis as a highly expressed downstream module

We identified the ideal downstream module to be the one for the synthesis of farnesene, a C_{15} linear isoprenoid. Unlike lycopene, farnesene is a volatile compound which the cell excretes and which we can easily remove/capture from the culture media, e.g. through a dodecane overlay or C18 resin, avoiding capacity problems. Farnesene synthases (FS) can be expressed well in *E. coli* to give impressive (g/L) titers of farnesene with high flux (12–14). As such, recent literature suggests that farnesene synthase is one of the most active terpenoid synthases known. Because of the enzyme's high activity, it is not expected to act as a bottleneck, therefore making it a good candidate for downstream coupling with the IUP, either alone or in tandem with the enzyme IspA, which produces GPP and FPP.

The experimental plan would see the creation of a downstream expression vector for the overexpression of either FS alone or FS together with IspA. This vector would then be used together with an (upstream) IUP vector for the production of farnesene in culture experiments in which both the farnesene titers as well as the levels of metabolic intermediates would be monitored (using GC-MS and LC-MS/MS respectively). The initial expectation would be to observe high IPP accumulation, due to the IPP supply (from upstream-IUP) being greater than the IPP consumption (from downstream-farnesene synthase).

In order to achieve the balancing of upstream and downstream, the expression levels of the downstream module would be increased by either increasing promoter induction levels assuming an inducible promoter is used, by using a stronger promoter or a stronger RBS, by changing the order of genes (FS, IspA) in the farnesene synthesis operon or finally by changing the origin of the expression vector to change the plasmid copy number. Such increases would be predicted to lead to an increase in farnesene titer, but more importantly, a decrease in IPP levels. The current limits of the IUP would be discovered by continually strengthening the downstream until no further increase in farnesene titer and very low IPP levels are observed, signaling that the downstream and upstream modules are balanced. As soon as this happens, an iterative process can be

employed to push past IUP's current limits. In short, the upstream (IUP) can be expressed higher, using any number of methods which will be described later. This would lead to an imbalance in the full synthesis pathway, predicted to manifest itself as accumulation of metabolic intermediates, mainly IPP. The expression levels of the downstream (farnesene synthesis) can then be increased until the IPP concentrations return to very low levels, signaling that the whole pathway is now again balanced but at an overall higher flux level, which can be corroborated by a higher overall farnesene titer/productivity.

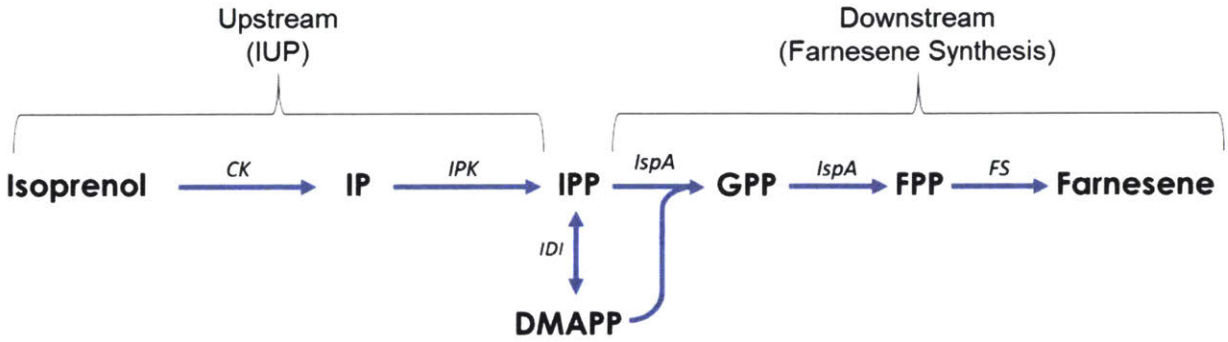
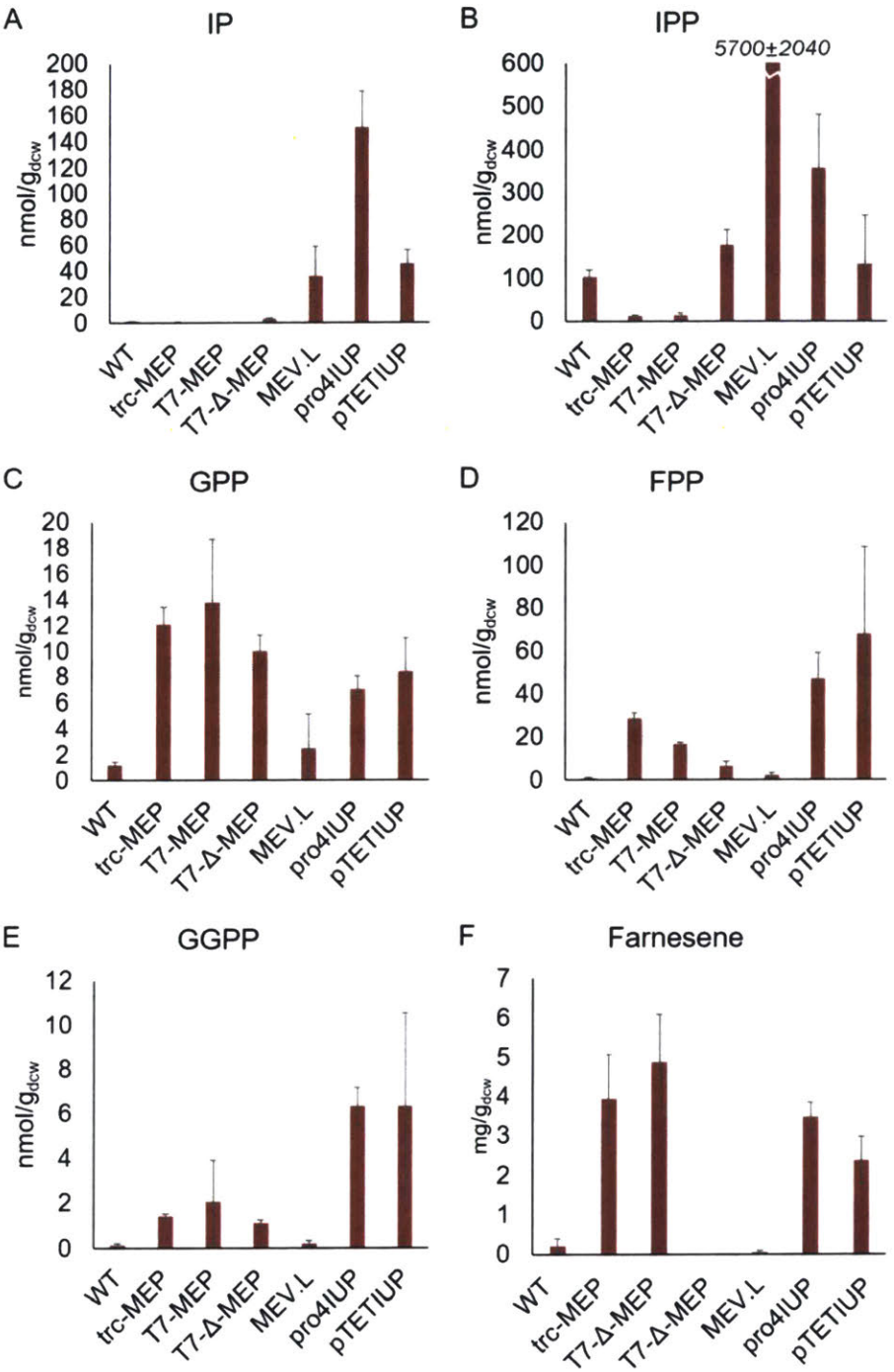


Figure 5-4: The farnesene synthesis pathway: The IUP produces IPP & DMAPP, the main isoprenoid intermediates, which then get converted to farnesene through the downstream pathway. When the upstream and downstream are balanced, there should be very little IPP/DMAPP accumulation.

5.3.2 Iterations on farnesene production optimization

We acquired a codon-optimized farnesene synthase gene (*aFS*) from Manus Biosynthesis, which we inserted, along with *ispA* into an operon in which *ispA* precedes *aFS* in the operon sequence. Said operon was inserted in a plasmid with origin of replication pSC101 (copy no. ~5) and a T7 promoter, to create plasmid p5T7-ispA-aFS through which the downstream farnesene synthesis module can be expressed. This largely mirrors the expression systems we used for the expression of other isoprenoids produced *in vivo* via the IUP (Chapters 3 & 4).



Isoprenoid pathway	Description
WT	Endogenous MEP (wild-type)
trc-MEP	Upregulated MEP, trc promoter
T7-MEP	Upregulated MEP, T7 promoter
T7-Δ-MEP	Upregulated MEP, T7 promoter, *ispA* deletion
MEV.L	Lower MVA pathway
pro4IUP	IUP, constitutive pro4 promoter (weak)
pTETIUP	IUP, inducible pTET promoter (strong)

Figure 5-5: Metabolite levels and farnesene titers from cultures expressing different isoprenoid pathways. Cells were cultured for 24h in M9 media at 37°C. Insert provides details on the 7 different isoprenoid pathways assayed. Concentrations for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively and (**F**) farnesene titers. All metabolite concentrations and farnesene titers are reported as means ± SD of three biological replicates.

As a first step, we chose to compare the farnesene titer achievable by combining this plasmid with different isoprenoid pathways (see Section 4.6.2), including the MEP pathway (native or upregulated), the lower mevalonate pathway and the IUP, as well as the intermediate accumulations. One remarkable observation was that even though the lower mevalonate pathway produced orders of magnitude more IPP than any of the other pathways, this did not translate to a farnesene titer, presumably due to the severe imbalance in upstream and downstream. Concerning the final farnesene titer, what we observed was that even though the IUP did not produce higher titers than those achieved through the use of different upstream pathways (e.g. the upregulated MEP pathway under either the trc or T7 promoters), the titers were comparable. Furthermore, the intermediate concentrations achieved through the IUP were higher than those in other tested pathways, with the exception of the lower mevalonate pathway. Since the IUP produced similar amounts of farnesene as upregulated MEP, but displayed more potential (higher IPP, FPP accumulation), this led to the conclusion that the downstream module needed to be upregulated to catch up to the IUP upstream.

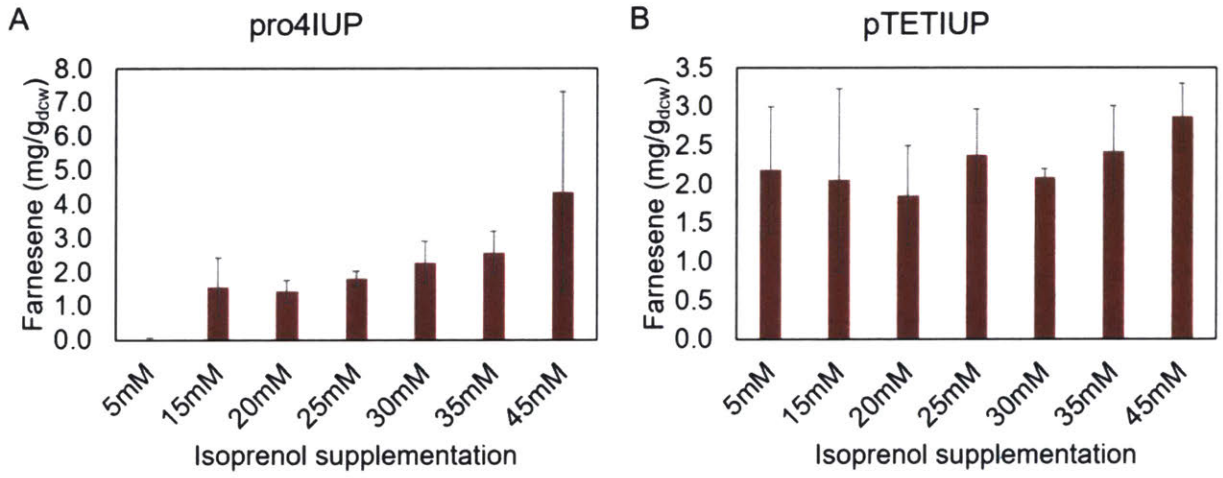
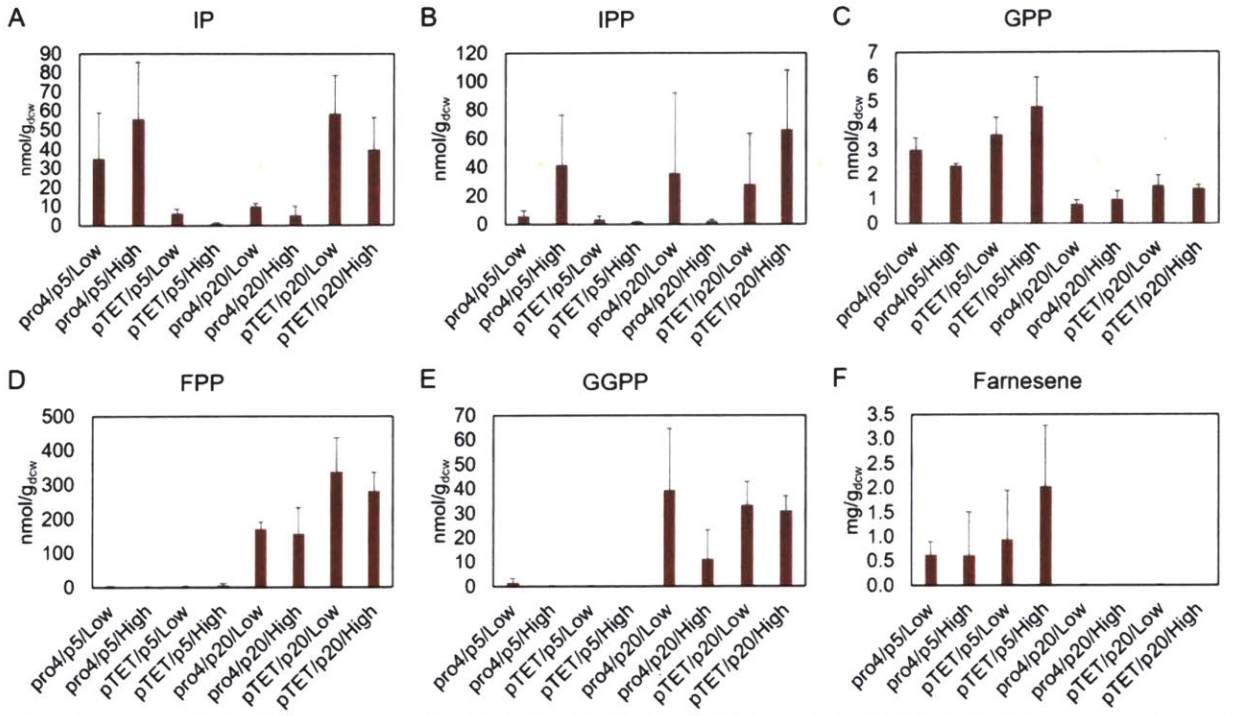


Figure 5-6: Farnesene titers from IUP cultures supplemented with different isoprenol levels. Strains are expressing the IUP under the control of a low strength constitutive promoter (pro4IUP) or a strong inducible promoter (pTET) induced with 20 ng/mL of aTc under different concentrations of isoprenol. Cells were cultured for 24h in M9 media at 37°C. Farnesene titers are reported as means ± SD of three biological replicates.

As a follow-up experiment, to ensure that we were operating the IUP at the optimal isoprenol levels, we varied said levels from 5mM to 45mM (Figure 5-6). Since we did not observe any significant changes in farnesene titer, we continued to use 25mM of isoprenol as IUP supplementation.

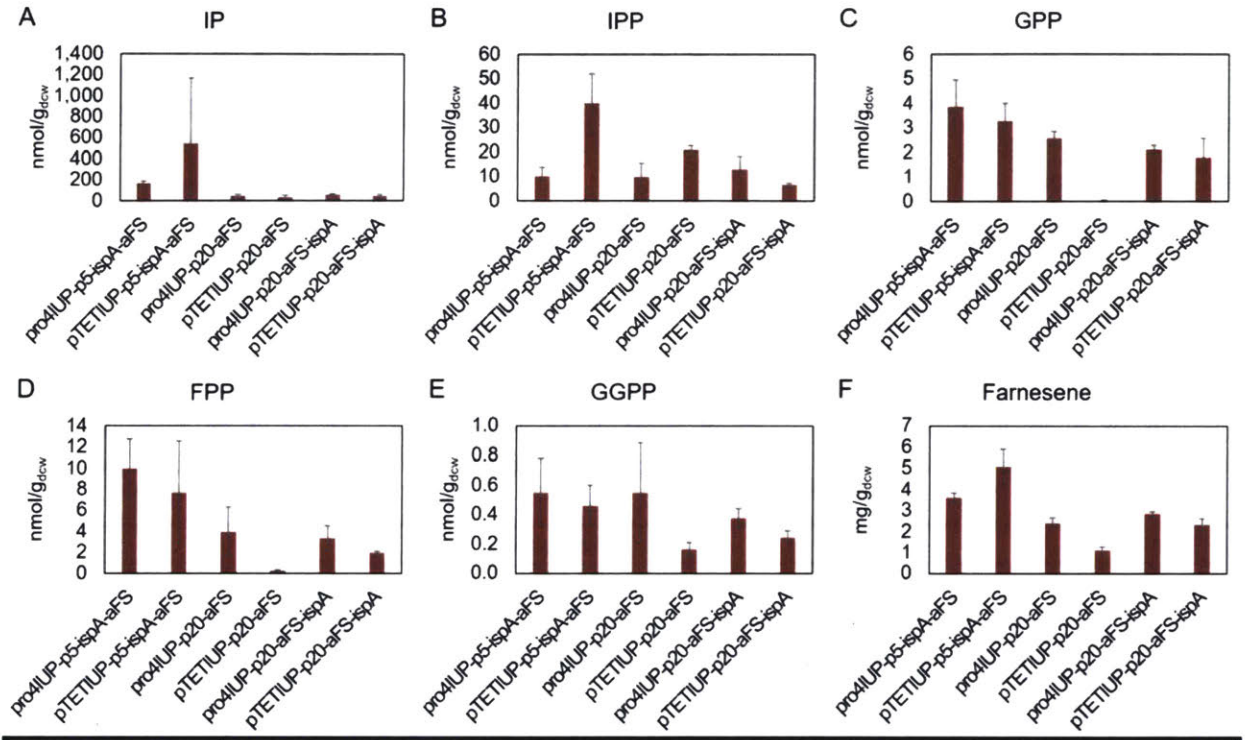


Strain	Description
pro4 vs. pTET	IUP expression mode: pro4: constitutive pro4 promoter (weak), pTET: inducible pTET promoter (strong)
p5 vs p20	Downstream plasmid origin of replication: p5: pSC101 (copy no. ~5), p20: pBR322 (copy no. ~20)
Low vs High	Downstream plasmid induction level: Low: 0.1mM IPTG, High: 0.2mM IPTG

Figure 5-7: Metabolite levels and farnesene titers under different strengths for upstream (IUP) and downstream (farnesene synthesis) modules. Cells were cultured for 24h in M9 media at 37°C. Strains differ on IUP expression mode, downstream plasmid origin of replication and induction levels (insert provides further details). Concentrations for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively and (**F**) farnesene titers. All metabolite concentrations and farnesene titers are reported as means ± SD of three biological replicates.

For the first round of optimization for the downstream module, we varied the IPTG induction levels for the downstream plasmid from 0.1mM to 0.2mM; however, our results (Figure 5-7) showed no significant difference in isoprenoid precursor concentrations or farnesene titers related to this

change. We also varied the copy number of the farnesene synthesis plasmid (downstream) from ~5 to ~20 by changing the plasmid origin of replication from pSC1010 to pBR322. We observed (Figure 5-7) that by increasing the downstream plasmid copy number the farnesene titer was significantly decreased (to almost 0) whereas the FPP concentration was increased. Thus, we believe that the simultaneous upregulation of *ispA* and *aFS* led to FPP accumulation (which, through pathway imbalance led to decreased farnesene titer). This led us to the conclusion that the overall imbalance resulted only in shifting the bottleneck to FPP, right before FS.

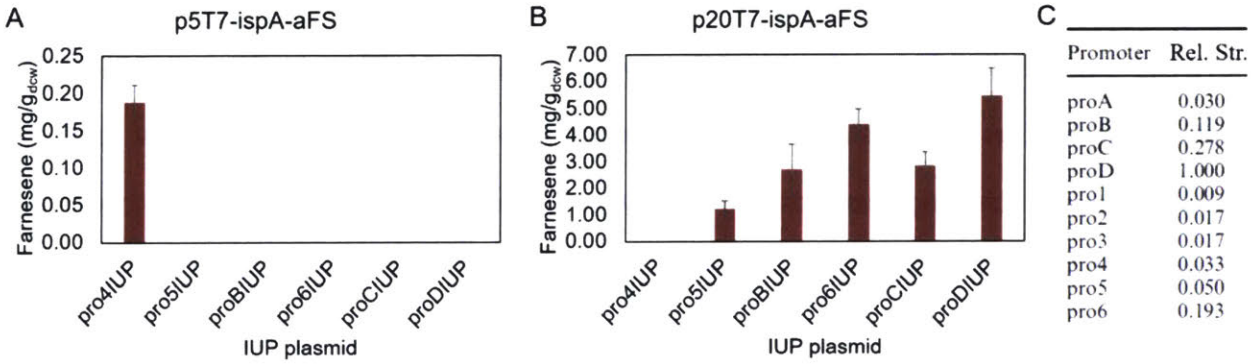


Plasmids	Description
pro4IUP vs. pTETIUP	IUP plasmid: pro4IUP: constitutive pro4 promoter (weak), pTETIUP: inducible pTET promoter (strong)
p5T7-ispA-aFS	Farnesene synthesis plasmid: Copy #:~5, ispA (GPP & FPP synth.) preceding aFS (farn. synth.) in operon
p20T7-aFS	Farnesene synthesis plasmid: Copy #:~20, aFS (farn. synth.), no ispA (GPP & FPP synth.) in operon
p20T7-aFS-ispA	Farnesene synthesis plasmid: Copy #:~20, aFS (farn. synth.) preceding ispA (GPP & FPP synth.) in operon

Figure 5-8: Metabolite levels and farnesene titers under different downstream (farnesene synthesis) plasmids. Cells were cultured for 24h in M9 media at 37°C. Strains differ on IUP expression plasmids, downstream plasmid origin of replication and operon (insert provides further details). Concentrations for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively and (**F**) farnesene titers. All metabolite concentrations and farnesene titers are reported as means ± SD of three biological replicates.

For the next round of optimization for the downstream module, we attempted to improve aFS activity without simultaneously upregulating ispA, which could exacerbate FPP accumulation. We thus created a plasmid (p20T7-aFS-ispA) with high copy number (origin of replication pBR322, copy no. ~20), in which the order of genes in the operon was flipped, so that the *aFS* coding sequence was to the 5' of the *ispA* coding sequence in the operon, in order to improve *aFS* expression and decrease *ispA* expression as it is known (15) that that the expression of a gene increases as its position moves closer to the beginning of the operon. We furthermore created a high copy number plasmid (p20T7-aFS) in which only expressed *aFS* and not *ispA*, relying on the *E. coli*'s native ispA to produce GPP & FPP.

We observed (Figure 5-8) that neither of the two new downstream plasmids could outperform the original downstream plasmid (p20T7-ispA-aFS). However, one interesting observation was that the lack of *ispA* upregulation, combined with a strong IUP expression (i.e. the pTETIUP-p20aFS strain) led to zero GPP and FPP levels, indicating that there is no bottlenecking in these nodes, but also lower farnesene titers, highlighting the importance of *ispA* upregulation, albeit in lower levels.



Promoter	Rel. Str.
proA	0.030
proB	0.119
proC	0.278
proD	1.000
pro1	0.009
pro2	0.017
pro3	0.017
pro4	0.033
pro5	0.050
pro6	0.193

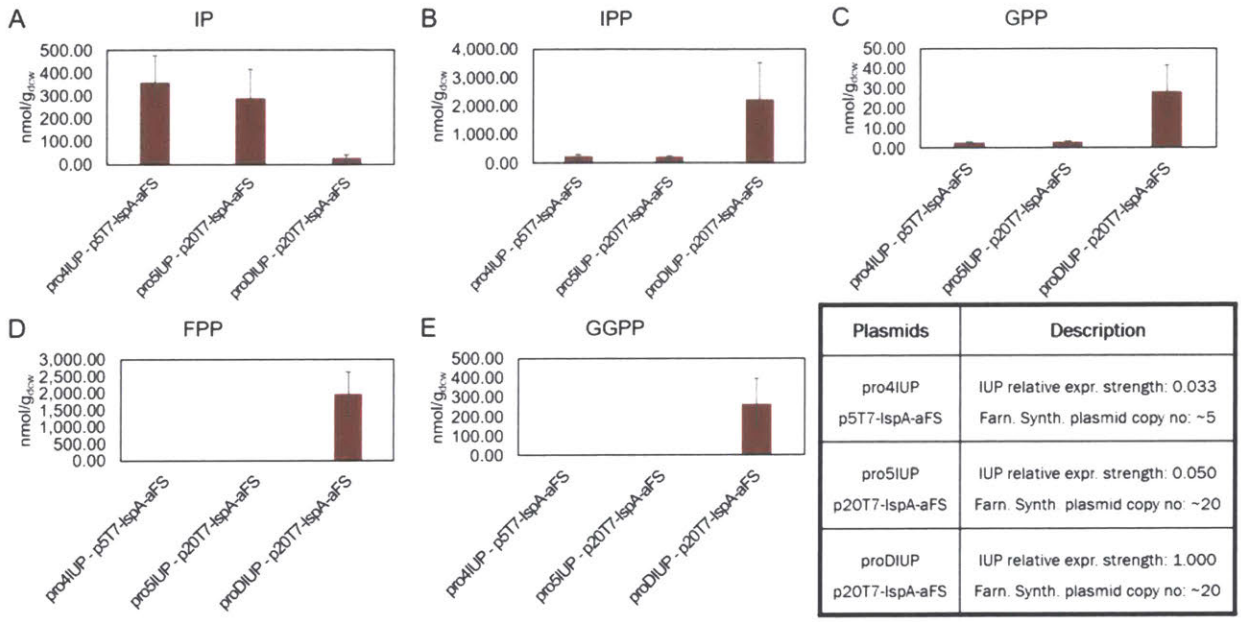
Figure 5-9: Improved farnesene production through the IUP. (**A&B**) Farnesene titers after culturing for 24h in M9 media at 37°C. Strains expressed the farnesene pathway through the plasmid (**A**) p5T7-ispA-aFS (copy no. ~5) or (**B**) p20T7-ispA-aFS (copy no. ~5) and the IUP though plasmids of the series pSEVA228-proXIUPi (proXIUP), where (**C**) "proX" are constitutive promoters of varying strengths. Concentrations are reported as means ± SD of three biological replicates.

For the next step of optimization/debottlenecking we took a different approach. Instead of focusing solely on upregulating the downstream, while keeping the upstream modules unchanged, we decided to try the simultaneous upregulation of both upstream and downstream. With regards to the IUP, we varied the strength of the pathway by using plasmids of the series pSEVA228-proXIUPi (where X=4, 5, B, 6, C or D), which utilize progressively stronger constitutive promoters (16).

We observed that, in the case of low expression (Figure 5-9 A) for the downstream module, any IUP expression above the pro4 level led to complete abolition of farnesene production, presumably due to pathway imbalance. We were pleasantly surprised, however, to observe that the simultaneous upregulation of both the downstream and upstream modules (IUP promoters stronger than pro4) (Figure 5-9 B) led to higher farnesene levels, which were higher than those achieved by the control strain (with plasmids p5-T7-ispA-aFS and pSEVA228-pro4IUP) we had so far been using, in which both upstream and downstream were expressed at lower levels, a very positive result.

We further measured the metabolic intermediate levels produced by the control strain (carrying plasmids p5-T7-ispA-aFS and pSEVA228-pro4IUP), the best producing of the strains with a high copy number farnesene plasmid (carrying plasmids p20-T7-ispA-aFS and pSEVA228-proDIUP) as well as one strain with the high copy farnesene plasmid that did not produce farnesene as well (carrying plasmids p20-T7-ispA-aFS and pSEVA228-pro5IUP). We observed (Figure 5-10) that in the cases of the control strain (pro4IUP – p5T7-IspA-aFS) and the poorly producing strain (pro5IUP – p20T7-IspA-aFS) metabolite profiles looked similar, with significant accumulations of both IP and IPP (in the order of ~350 nmol/mg_{dcw} and ~250 nmol/mg_{dcw}), whereas the levels of GPP and FPP remained low. On the other hand, the best producing strain (proDIUP – p20T7-IspA-aFS) exhibited low levels of IP, but high, or very high levels of IPP, GPP. FPP and even GGPP. This is interesting for two reasons. First, the IPP accumulation in itself was very high (in the order of 2200 nmol/mg_{dcw}), in fact higher than ever before achieved via the IUP.

This brings the IPP accumulation in the same order of magnitude as that achieved by the lower mevalonate pathway (5700 nmol/mg_{dcw}, Figure 5-5 B), showing that the IUP, through a simple upregulation, can produce IPP fluxes competing with the most highly producing pathway so far. Secondly, we can observe that, unlike the case of the lower mevalonate pathway, where the pathway imbalance resulted in high IPP accumulation but no farnesene production (Figure 5-5 F), in this case the simultaneous upregulation of downstream and upstream could lead to better farnesene titers.



Plasmids	Description
pro4IUP p5T7-IspA-aFS	IUP relative expr. strength: 0.033 Farn. Synth. plasmid copy no: ~5
pro5IUP p20T7-IspA-aFS	IUP relative expr. strength: 0.050 Farn. Synth. plasmid copy no: ~20
proDIUP p20T7-IspA-aFS	IUP relative expr. strength: 1.000 Farn. Synth. plasmid copy no: ~20

Figure 5-10: Metabolite levels of select farnesene-producing strains. Cells were cultured for 24h in M9 media at 37°C. Strains differ on IUP expression plasmids and farnesene plasmids (insert provides further details). Concentrations for metabolic intermediates (**A**) IP, (**B**) IPP/DMAPP, (**C**) GPP, (**D**) FPP and (**E**) GGPP respectively and (**F**) farnesene titers. All metabolite concentrations are reported as means ± SD of three biological replicates.

5.4 Conclusions and future directions

In this chapter, we investigated strategies for the improvement of lycopene storage capacity and taxadiene synthase. More importantly, we investigated the possibility of identifying the IUP's limits and pushing past them via coupling the IUP with a strong downstream, for which we selected farnesene. We observed that the IUP is a promising pathway, as it can produce farnesene in titers

that compete with those achievable by state-of-the art strains and pathways. However, the large accumulation of metabolic intermediates observed still indicate that effort is still required to fully utilize the IUP's potential and translate these intermediates to final products.

One interesting observation was that when *ispA* and *aFS* were upregulated, FPP accumulated to significant levels (200-400 nmol/g_{dcw}) (Figure 5-7) and farnesene production was abolished. This observation could be explained by literature reports that IspA, the enzyme responsible for the synthesis of FPP can be allosterically inhibited by its own product (17), which could thereby deleteriously impact downstream farnesene synthesis. Moreover, *E. coli* endogenously expresses endogenous phosphatases, which can catalyze farnesol formation (18). It is very likely that these phosphatases compete with farnesene synthase for the FPP pool and divert production towards farnesene, similar to previous literature reports (12). Poor farnesene titers could also be explained by the fact that FPP is also used for the synthesis of quinones (19) and lipid carriers (20).

When the IUP is upregulated together with the farnesene synthesis we observed that (Figures 5-9 & 5-10) even though FPP concentrations were ~2000 nmol/g_{dcw} farnesene production was not abolished. One possible explanation could be that IUP upregulation resulted in FPP accumulation so high that the above issues were partly overcome, although the fact that farnesene titer only increased by an order of magnitude between the pro4IUP-p5T7-IspA-aFS and the proDIUP-p20T7-IspA-aFS cases whereas FPP accumulation increased by 3 orders of magnitude highlights inefficiencies in this approach.

Based on our latest results, it is obvious that a new avenue of inquiry lies in the potential simultaneous upregulation of upstream and downstream, as well as further investigation of enzymes that compete with farnesene synthase to ensure the overall pathway is balanced. In the case of the downstream, quick ways in which it can be improved include a change in plasmid copy number. The ways in which the upstream IUP could be improved range from the simple ones described above, i.e. improving expression by changing induction levels, promoter strength,

plasmid copy number or using different operon structures, to more complex ones such as protein engineering or using different homologs for the IUP enzymes.

5.5. Materials and Methods

Strains, plasmids and genes; cloning protocol

Materials and methods were the same as the ones as described in the relevant sections of Chapter 2. Strains T7-MEP and T7-Δ-MEP were a gift from Manus Biosynthesis (MIT MTA 8234). Strains JW0728-1, JW1667-5, JE5505 and TPS30 were purchased from the Yale CGSC (Coli Genetic Stock Center).

Construction of expression vectors.

All vectors were constructed following the routine cloning protocol described in Chapter 2. Vectors pSEVA228-pro4IUPi and pTET-IUPi for the expression of the IUP pathway were constructed by the method described in Chapter 3. The plasmids pSEVA228-proXIUPi (where X= 5, B, 6 or C) were created by replacing the 6 nucleotides in the proD promoter sequence of pSEVA228-proDIUPi with the primer pairs GB-proX_Mut_f/GB-ProLibrary_Mut_r (where X= 5, B, 6 or C) to amplify the whole plasmid and then subsequently assembling the amplification product. Plasmid p5T7-ispA-aFS was created by PCR-amplifying the backbone, promoter and *ispA* gene of plasmid p5T7ispA-ads using the primer pair GB-p5T7ispA-F/R and the aFS plasmid from the provided construct using primer pair GB-aFS-F/R and assembling the amplification product. Plasmid p20T7-ispA-aFS was created by PCR-amplifying the pBR322 origin of replication from plasmid p20-LYCipi using the primer pair GB-p20ori-F, amplifying the backbone, antibiotic gene and expression cassette of plasmid p5T7-ispA-aFS using the primer pair GB-T7ispAaFS-F/R. Plasmid p20T7-aFS-ispA was created by PCR-amplifying the following three elements from plasmid p20T7-ispA-aFS: i) the backbone using the primer pair GB-p20T7-F/R, ii)*ispA* using primer pair GB-ispA_in_p20T7-IspA-aFS_F/R and iii)*aFS* using the primer pair GB-aFS_in_p20T7-IspA-

aFS_F/R and assembling the fragments. To create plasmid p20T7-aFS the following two elements were PCR-amplified from plasmid p20T7-ispA-aFS and subsequently assembled: i) the backbone using the primer pair GB-p20T7-F/R and ii). The codon-optimized farnesene synthase gene (*aFS*) was a gift from Manus Biosynthesis (MIT MTA 8234). Vector sequences for the primers used are listed in **Table A3** in the Appendix. Codon-optimized sequences are listed in **Table A5** in the Appendix.

Quantification of metabolites, lycopene, taxadiene and farnesene.

IP/DMAP, IPP/DMAPP, GPP, FPP and GGPP were detected and quantified by LC-MS/MS using the means described in Chapter 4. Lycopene content was quantified using the protocol described in Chapter 3. Taxadiene was quantified as described in Chapter 4. Farnesene was quantified using GC-MS using the same protocol as the one for the quantification of limonene (Chapter 4). A standard purchased from Sigma-Aldrich was used for quantification.

Cultivation in serum bottles.

Cultivation in serum bottles was conducted under the same conditions described in Chapter 4, with the following changes, listed below: When strains containing downstream operons for farnesene production were used, 2mL of dodecane overlay was added to the cultures at induction time to capture it. Strains were grown as 37°C. Metabolites, lycopene and farnesene were quantified using protocols described above.

Mutagenesis and screen for improving taxadiene synthesis.

A vector for the co-production of lycopene and taxadiene (p5T7-lyc-tds-kan) was constructed via amplifying taxadiene synthase from p5T7tds-ggpps using primers GB_p5t7ggppstds_f/r, amplifying the lycopene synthesis operon and the backbone (minus the spectinomycin resistance gene) from p5T7-LYCipi using primers GB-p5T7Lyc-back_f/r, amplifying a kanamycin resistance gene from pTET-IUP using primers GB-KanRes_f/r and assembling. Variants of the vector with

mutated taxadiene synthase were created as above, with the sole difference that taxadiene synthase was amplified using the GeneMorph II Random Mutagenesis Kit (Agilent) following the instructions of the manufacturer. Vectors bearing mutants were transformed in electrocompetent trc-MEP *E. coli* strains, which were plated in kanamycin-selective LB plates and incubated overnight. Individual colonies were picked by toothpicks and grown in 96-well plates in kanamycin-selective LB media overnight, upon which time they were picked, stamped on a large kanamycin-selective LB plate, supplemented with 0.1 mM IPTG (576 colonies per plate) and the plates were incubated overnight. All the colonies that were not visibly red and ~10 random red ones were cultured in serum bottles (as described above) for the production of taxadiene. The plasmids of the best producers were extracted and the taxadiene synthases sequenced for mutations. To create taxadiene synthesis vectors without the lycopene production operon, the mutated taxadiene syntheses of the best producers were amplified from their respective plasmids via primer pair GB-TDS_in_p5T7_f/r, the backbone of p5T7tds-ggpps was amplified via primer pair GB-p5T7-backbone_f/r and the resulting fragments assembled.

5.6. Bibliography

1. Wang GS, Grammel H, Abou-Aisha K, Sägesser R, Ghosh R (2012) High-level production of the industrial product lycopene by the photosynthetic Bacterium rhodospirillum rubrum. *Appl Environ Microbiol* 78(20):7205–7215.
2. Baba T, et al. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* 2. doi:10.1038/msb4100050.
3. Gerding MA, Ogata Y, Pecora ND, Niki H, De Boer PAJ (2007) The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli. *Mol Microbiol* 63(4):1008–1025.
4. Narita S ichiro, Tokuda H (2017) Bacterial lipoproteins; biogenesis, sorting and quality

control. *Biochim Biophys Acta - Mol Cell Biol Lipids* 1862(11):1414–1423.

5. Hirota Y, Suzuki H, Nishimura Y, Yasuda S (1977) On the process of cellular division in Escherichia coli: a mutant of E. coli lacking a murein-lipoprotein. *Proc Natl Acad Sci U S A* 74(4):1417–1420.
6. Sun TP, Webster RE (1986) fii, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant tolA and tolB. *J Bacteriol* 165(1):107–115.
7. Lloubès R, et al. (2002) The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity? *Res Microbiol* 152(6):523–529.
8. Furubayashi M, et al. (2014) A high-throughput colorimetric screening assay for terpene synthase activity based on substrate consumption. *PLoS One* 9(3). doi:10.1371/journal.pone.0093317.
9. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* 330(2010):70–74.
10. Luo Z, et al. Enhancing lycopene synthesis by bypassing the native pathways and modulating intracellular hydrophobicity in Yarrowia lipolytica. *Prep*.
11. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol* 21(7):796–802.
12. Wang C, et al. (2011) Metabolic engineering of Escherichia coli for α-farnesene production. *Metab Eng* 13(6):648–655.
13. You S, et al. (2017) Utilization of biodiesel by-product as substrate for high-production of B-farnesene via relatively balanced mevalonate pathway in Escherichia coli. *Bioresour Technol* 243:228–236.
14. Wang C, Zada B, Wei G, Kim S-WW (2017) Metabolic engineering and synthetic biology approaches driving isoprenoid production in Escherichia coli. *Bioresour Technol* 241:430–438.

15. Lim HN, Lee Y, Hussein R (2011) Fundamental relationship between operon organization and gene expression. *Proc Natl Acad Sci* 108(26):10626–10631.

16. Davis JH, Rubin AJ, Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* 39(3):1131–1141.

17. Park J, Zielinski M, Magder A, Tsantrizos YS, Berghuis AM (2017) Human farnesyl pyrophosphate synthase is allosterically inhibited by its own product. *Nat Commun* 8:1–8.

18. Wang C, et al. (2010) Farnesol production from Escherichia coli by harnessing the exogenous mevalonate pathway. *Biotechnol Bioeng* 107(3):421–429.

19. Asai KI, et al. (1994) The identification of Escherichia Coli ispB (CEL) gene encoding the octaprenyl Diphosphate synthase. *Biochem Biophys Res Commun* 202(1):340–345.

20. Bouhss A, Trunkfield AE, Bugg TDH, Mengin-Lecreulx D (2008) The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* 32(2):208–233.

Chapter 6

Use of the IUP for the production of functionalized isoprenoids

6.1. Introduction

6.1.1 Central Concept

As has been mentioned earlier (Chapter 1), all isoprenoids are derived from the C_5 molecules IPP and DMAPP, which are then condensed to form GPP, FPP and GGPP, forming isoprenoid backbone molecules, which can be further functionalized. In nature, such functionalization steps are catalyzed by enzymes such cytochrome P_{450}, however, in many cases, specific chemistries such as halogenations are not known to occur in nature or could occur in very limited fashion. An opportunity for innovation opened by the IUP's simplicity is the possibility for using this pathway to introduce hereto unique and untested ways of functionalization into isoprenoid backbones, leading to the production of novel isoprenoid molecules. This new, short pathway offers a unique opportunity for synthesizing new isoprenoid backbone structures that would otherwise be very difficult or impossible to produce by biosynthesis though the traditional isoprenoid pathways (MEP/MVA) or total chemical synthesis, while allowing for specificity in stereochemistry.

The main idea behind this line of inquiry lies in the fact that through the IUP it is possible, in only two steps, to produce an isoprenoid intermediate, a C_5 diphosphate, through double phosphorylation of a C_5 alcohol. While in its "mundane" embodiment, this would translate to the production of IPP or DMAPP from isoprenol or prenol, what is being considered that instead of using (non-functionalized) isoprenol or prenol as the IUP's entry point, the isopentenol utilization

pathway is used for the double phosphorylation of a *functionalized* C_5 alcohol. Thus, it is hypothesized that functionalized IPP or DMAPP could be produced. These functionalized IPP/DMAPP molecules would then be used as substitutes for IPP/DMAPP to produce functionalized GPP, FPP, GGPP and ultimately functionalized isoprenoids. Such a use of the IUP would be the first real opportunity for easy, one-pot synthesis of a great array of functionalized isoprenoids, greatly expanding the accessible chemical space that can be investigated for drug discovery and testing.

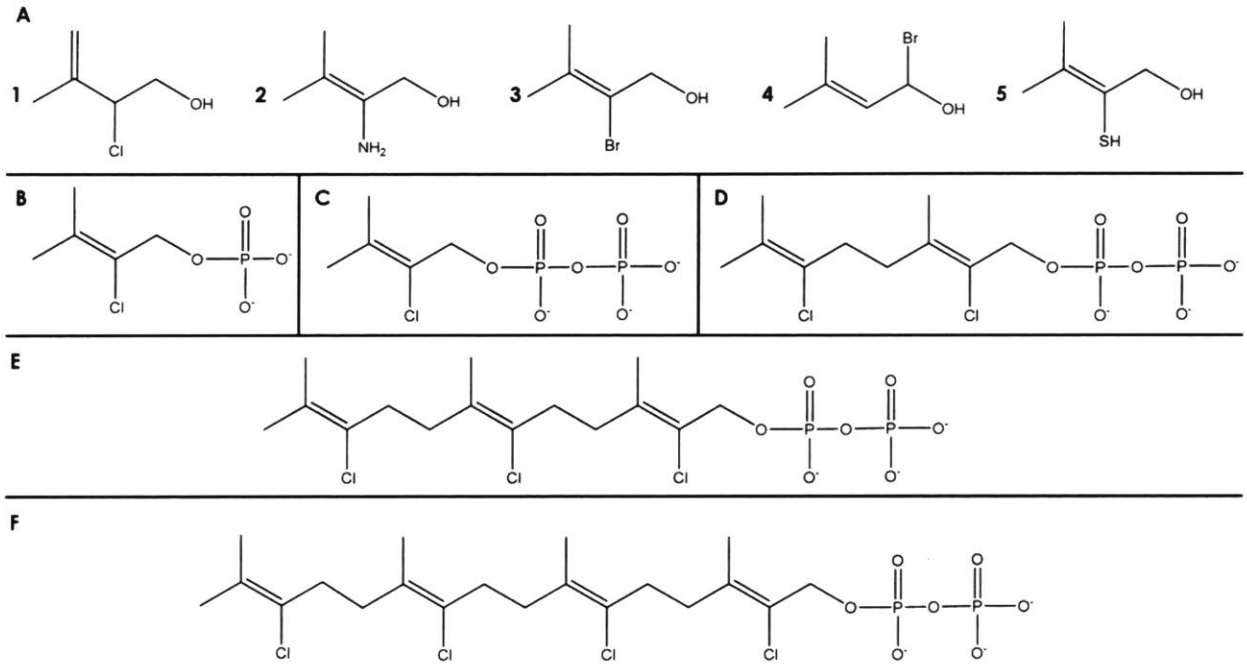


Figure 6-1: Functionalized isopentenols and functionalized isoprenoid metabolic intermediates: (**A**) Examples of functionalized isopentenols through the addition of a (*1*) chlorine group, (*2*) amine group, (*3&4*) bromine group, (*5*) thiol group. (**B-F**) metabolic intermediates that could hypothetically be derived from 2-chloro-3-methylbut-3-en-1-ol [chlorinated isoprenol, A1], which are the chlorinated equivalents of (**B**) IP, (**C**) IPP, (**D**) GPP, (**E**) FPP and (**F**) GGPP

The working hypothesis is that all relevant pathway enzymes, will be able to act upon functionalized (e.g. halogenated) substrates through promiscuous activity. It is known from literature (1, 2) that many enzyme classes, especially phosphokinases, an enzymatic class to which both of the IUP enzymes belong, can indeed act promiscuously. As a case in point, when

we screened many phosphokinases for promiscuous activity on isoprenol/prenol, we observed that most of them could promiscuously phosphorylate both isoprenol and prenol, which were not the enzymes' native substrates, with choline kinase being the best performer (Chapter 2.3.3). This hypothesis for promiscuity is an extension of the above.

Using the IUP for this task, instead of a traditional isoprenoid pathway, is particularly appealing due to its multiple advantages. It would be exceedingly difficult if not impossible to adopt this strategy (use of functionalized substrates for the production of functionalized isoprenoids) for use with the natural isoprenoid pathways since their inextricable linkage of with central carbon metabolism would functionalize most all metabolic intermediates inactivating critical enzymes and leading to cell death. The IUP's complete decoupling from the cell's main metabolism circumvents this problem. This is advantageous because one could have cells producing the non-functionalized isoprenoid precursors needed for their survival from glucose, though the MVA/MEP pathways, and also express the IUP and feed the appropriate functionalized isopentenol for the production of functionalized isoprenoid precursors pathways without any fear of cross-talk between the two pathways. Furthermore, the whole issue of possible toxicity of functionalized isopentenols or isoprenoid precursors could be circumvented via implementing the pathway in a fully *in-vitro* approach using a metabolically reconstituted IUP in a cell-free system, thus bypassing the need for a cell culture system.

6.1.2 Cell-free isoprenoid production systems

The use of a cell-free system can be used to circumvent many of the challenges that an *in vivo* isoprenoid production can pose, including regulation (3, 4), especially if they originate from regulation systems other than that of the isoprenoid pathway itself, as well as avoid toxicity issues caused by over-accumulation of pathway intermediates (5). Though the adoption of a cell-free biocatalysis approach, great simplifications compared to the *in vivo* approach can be achieved, as there is no longer any need to perform strain engineering to remove any competing metabolic

pathways, manage pathways that are essential for cell survival, or address potential transcriptional and translational regulation.

Another advantage of cell-free systems lies in that they can be systematically used to study pathway kinetics and elucidate rate-limiting steps, which in turn can produce insights, such as identifying the most significant factors controlling the overall flux of the pathway or identifying pathway regulatory mechanisms at the protein level (6, 7), that can further guide *in vivo* optimization of metabolic pathways (6, 8). This has already been attempted in the case of the MVA pathway, which has been studied and optimized for the production of farnesene (7) and amorphadiene (9), by identifying previously unknown regulatory interactions and rate controlling steps and then using them to improve the pathway flux *in vivo*. Furthermore, balancing of the pathway is a far simpler endeavor *in vitro*, as the experimenter can alter the relative concentrations of relative enzymes at-will and with precision, something that is not as simple in an *in vivo* system with a complicate transcription and translation system. However, potential hurdles to using a cell free system at scale can be the high cost of enzyme production and purification, but also the need to use expensive labile cofactors which many pathways rely on. There is therefore a need to develop a simple and efficient platform.

The most interesting and complete example of a cell free isoprenoid production systems has been a recently reported cell-free system for the production of monoterpenes (10) through an *in vitro* reconstituted MVA pathway. While this system was able to achieve high titers (>11 g/L over 7 days) of pinene, sabinene, and limonene, and mitigate the toxicity effects which limit monoterpene production *in vivo*, the system was an exceedingly complex one, requiring 27 enzymes to sustainably synthesize monoterpenes from glucose. The major complicating factor for isoprenoid production through either natural pathway is the fact that the direct precursors to these pathways, acetyl-CoA, pyruvate, and glyceraldehyde-3-phosphate are not bulk chemicals making them unsuitable for cell-free synthesis on their own (11). Thus additional enzymes, such as glycolysis enzymes and enzymes for cofactor generation, utilization and balance are required

if either isoprenoid pathway is to be used with a cheap feedstock such as glucose. This is the reason that this system required such a large number of enzymes, complicating its implementation.

6.1.3 *In vitro* reconstitution of the IUP

Since the IUP is a pathway that is far shorter than either the MVA or MEP and uses a single cofactor, ATP, the pathway can be more easily engineered compared to the natural pathways, as well as reconstituted *in vitro* in a cell-free system. In fact, using the IUP for cell-free monoterpene biosynthesis, to mirror the case of (10), would dramatically shorten the pathway to only 4 enzymes (CK, IPK, IspA, and a monoterpene synthase) and even the synthesis of high-value cyclic diterpenes like taxadiene would only require 5 enzymes. Moreover, the fact that IUP's precursor is a bulk chemical that can be directly supplied to the system, thereby eliminating the need to also reconstitute additional enzymes to handle the pathway input, provides a definite edge over the natural pathways. A cell-free IUP could have significant potential for commercial production of isoprenoids particularly for high value compounds.

In our recent efforts (12), we metabolically reconstituted the IUP *in vitro*, characterized it and demonstrate its use for cell-free synthesis of mono-,sesquit-, and diterpenoids. After performing kinetic modelling and sensitivity analysis to identify the most significant parameters for productivity, we used the insights gathered to demonstrate an *in vitro* IUP system that can produce 220 mg/L of the diterpene taxadiene, in 9h, almost 3-fold faster than any system reported thus far. We have therefore established that it is possible to reconstitute the IUP *in vitro* and use it for the production of a variety of isoprenoids with high flux. We also note that the very good performance in taxadiene production gives us an alternative approach for taxadiene production, which has so far been difficult to optimize *in vivo* (Chapter 5.2.2).

6.2 Establishing a proof of concept

We elected to work on establishing a proof of concept by using the *in vitro* cell-free system using the enzymatically reconstituted IUP as our platform. The reasons for using this system instead of an *in vivo* one relate to the advantages of a cell-free system, which have been elaborated above and the major ones of which have to do with the degree of control we would have over the system and a desire for overcoming possible toxicity issues from the functionalized C_5-alcohol we use as substrate. Although the IUP is decoupled from central carbon metabolism, meaning that we can implement an *in vivo* approach without fear of halogenating compounds besides IPP/DMAPP and their derivatives, inadvertently poisoning the cell, there still remains the issue of unknown toxic effects on growth arising from the fact that the cell has to grow in the presence of an unknown alcohol. While this latter issue could potentially be addressed in an *in vivo* system by either evolving the cells themselves to achieve higher tolerance to the functionalized alcohol, adjusting or titrating said alcohol, or even replacing it with a differently functionalized, less toxic alcohol, this would create yet another confounding factor.

Concerning the type of functionalization we would introduce, we decided to aim for the production of halogenated isoprenoids. The halogen group is a relatively simple functional group which is expected to only have steric or polar interactions with residues in the enzyme active pockets but cannot form any hydrogen/S-S bonds, unlike amino- or thiol- groups respectively. Furthermore, the inclusion of halogens in the structure of organic molecules has been known to confer therapeutic properties or enhance existing ones (13). This project could therefore potentially open the route for the creation of novel pharmaceutical molecules.

6.2.1 Production of halogenated amorphadiene

As the main test case, we selected amorphadiene as the main target molecule. Amorphadiene is a C_{15} precursor to artemisinin, a potent antimalarial drug (14). As part of their initiative to eliminate

malaria, the Bill and Melinda Gates Foundation has significantly invested in research around the bioproduction of artemisinin, by funding research conducted by Amyris (15) and Sanofi (16). However, given reports of rising resistance to traditional antimalarial drugs, such as artemisinin (17–19), being able to produce halogenated versions of artemisinin or other isoprenoid drugs, would certainly aid in new drug testing and discovery. The success of this project would therefore present a very high-impact improvement, especially considering the renewed efforts by organizations such as the Gates Foundation for the discovery of new anti-malarial drugs.

As a first step, we chose to focus on the production of triply chlorinated amorphadiene, the main precursor to artemisinin. This amorphadiene variant would be produced by feeding 2-chloro-3-methylbut-3-en-1-ol, a chlorinated variant of isoprenol. These compounds would be converted into a chlorinated variant of IPP, a double chlorinated variant of GPP, a triply chlorinated variant of FPP (Figure 6-1 A1, B-E), and ultimately triply chlorinated amorphadiene.

This substrate was selected due to a number of factors:

1. The position of the halogenation in the second carbon is important, as it can lead to the production of a functionalized FPP that can be cyclized into triply chlorinated amorphadiene using the same mechanism as regular amorphadiene without exceeding the maximum valence of the carbons to which the halogens are bound. Positioning the halogen in other carbons could render proper cyclization impossible. Furthermore, the positioning of the halogens in the resulting amorphadiene is such that it is conceivable that the amorphadiene can be converted to artemisinin without issues on carbon valence.
2. Our experimental results for both the *in vivo* (Chapters 2-4) and *in vitro* (12) embodiments of the IUP have shown that the pathway functions better when using isoprenol as the substrate as compared to its performance when prenol is used. We therefore expect better performance when using a functionalized variant of isoprenol versus one of prenol.
3. The substrate is available via commercial vendors, although it has to be synthesized on-demand. The same applies to 2-bromo-3-methylbut-2-en-1-ol (brominated isoprenol) and

2-bromo-3-methylbut-2-en-1-ol (brominated prenol) but not any other halogenated isopentenols with the halogen in the second carbon.

4. There is readily available literature on the organic synthesis of 2-chloro-3-methylbut-3-en-1-ol using a simple two-step method (20), meaning that it could potentially be synthesized in the lab, allowing for a decrease in cost and possible production in higher quantities by the Stephanopoulos lab.

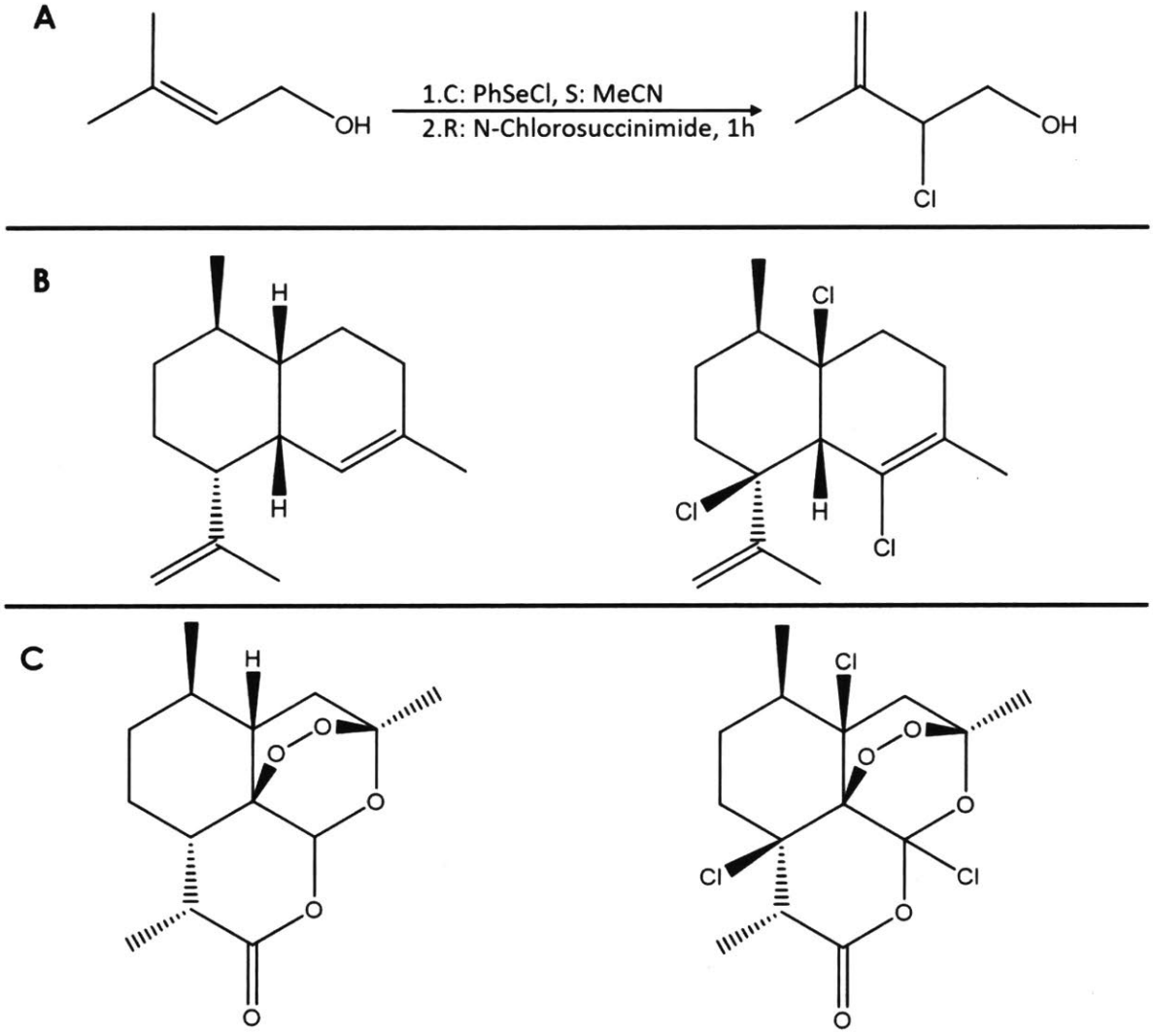


Figure 6-2: Chlorinated isoprenol and its expected derivatives: (**A**) 2-chloro-3-methylbut-3-en-1-ol (chlorinated isoprenol) and a reaction scheme for its synthesis. (**B**) *Left:* Structure of amorphadiene, *Right:* Hypothesized structure of triply chlorinated amorphadiene derived from 2-chloro-3-methylbut-3-en-1-ol, (**C**) *Left:* Structure of artemisinin, *Right:* Hypothesized structure of triply chlorinated artemisinin, derived from 2-chloro-3-methylbut-3-en-1-ol. Structures of amorphadiene or artemisinin would

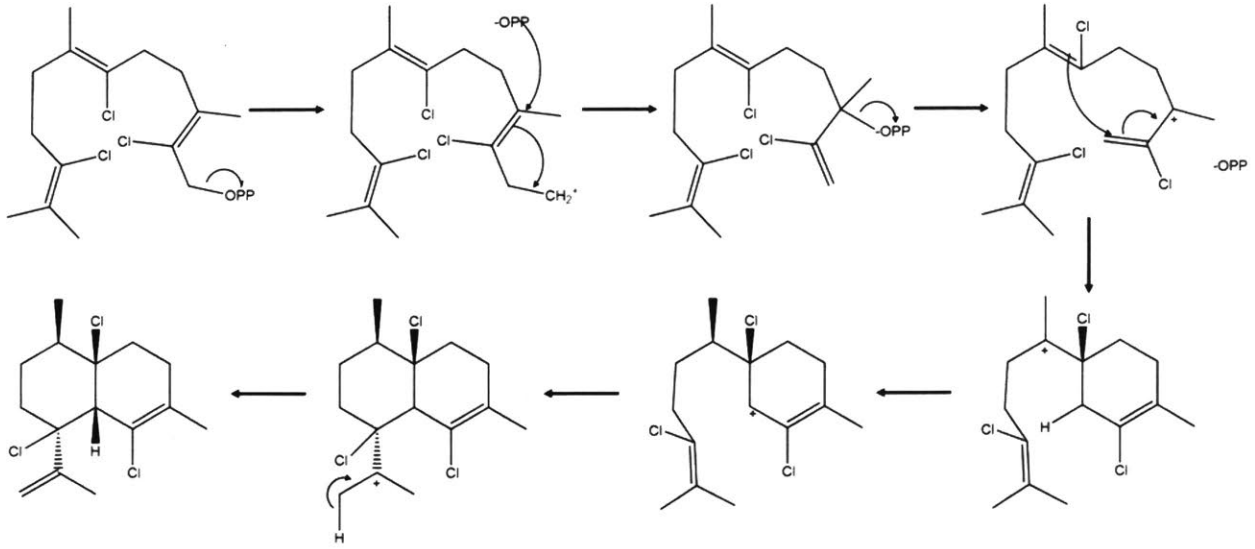


Figure 6-3: Hypothesized mechanism for the cyclization of triply chlorinated FPP to triply chlorinated amorphadiene: It is hypothesized that triply chlorinated FPP (top left) can be converted to triply chlorinated amorphadiene (bottom left) through the activity of the enzyme amorphadiene synthase, following a mechanism analogous to that of the cyclization of FPP to amorphadiene (21).

Although 2-chloro-3-methylbut-3-en-1-ol (chlorinated isoprenol) was our preferred substrate, the first experiments were conducted using 2-bromo-3-methylbut-2-en-1-ol (brominated prenol) as substrate, owing to the significantly lower purchase cost of the latter over the former, as well as earlier availability. We thus expected to produce the triply brominated amorphadiene analog, with expected molecular weight of 441.05 and expected m/Z ratios (and relative abundances) for the parent ions as 439.92 (100%), 441.92 (97.3%), and 437.92 (34.3%), as predicted by ChemDraw, which we tried to identify via GC-MS, as well as the hypothetical brominated intermediates, shown below, for which we scanned via LC-MS/MS.

Table 6-1: Expected metabolic intermediates derived from brominated prenol

ID	Formula	Exp. Parent Ion m/Z(1)	Exp. Daughter Ion m/Z(2)
Brom. IP	$C_5H_8BrO_4P^{2-}$	243/245	79/81
Brom. IPP	$C_5H_8BrO_7P_2^{3-}$	323/325	79/81
Twice Brom. GPP	$C_{10}H_{15}Br_2O_7P_2^{3-}$	469/471/473	79/81
Thrice Brom. FPP	$C_{15}H_{22}Br_3O_7P_2^{3-}$	615/617/619/621	79/81

(1): Multiple different ions scanned to account for the two isotopes ^{79}Br and ^{81}Br, with isotopic abundances of 51% and 49% respectively.
(2): 79 is the m/Z ratio of a phosphate group, mirroring the fragmentation patterns of non-halogenated metabolic intermediates. 79 & 81 also happen to be the m/Z ratios expected from ^{79}Br and ^{81}Br ions

All experiments were conducted in the already demonstrated fashion for producing amorphadiene *in vitro* (12), with the only difference being the substrate used. In short, our method involved first His-tag purifying choline kinase (ScCK) and isoprentenyl monophosphate kinase from *A. thaliana* (AtIPK) (i.e. the IUP enzymes), as well as isopentenyl-diphosphate isomerase (IDI), geranyl diphosphate/farnesyl diphosphate synthase (IspA) (the required isomerase and backbone enzymes) and amorphadiene synthase (ADS). These enzymes were then resuspended in ammonium bicarbonate buffer, supplied with ATP and brominated prenol and incubated overnight at 37°C. Samples for LC-MS/MS analysis (intermediates) and samples for GC-MS analysis (final product) were prepared, which were identical, but for the fact that the latter also had a dodecane overlay for concentrating the final product.

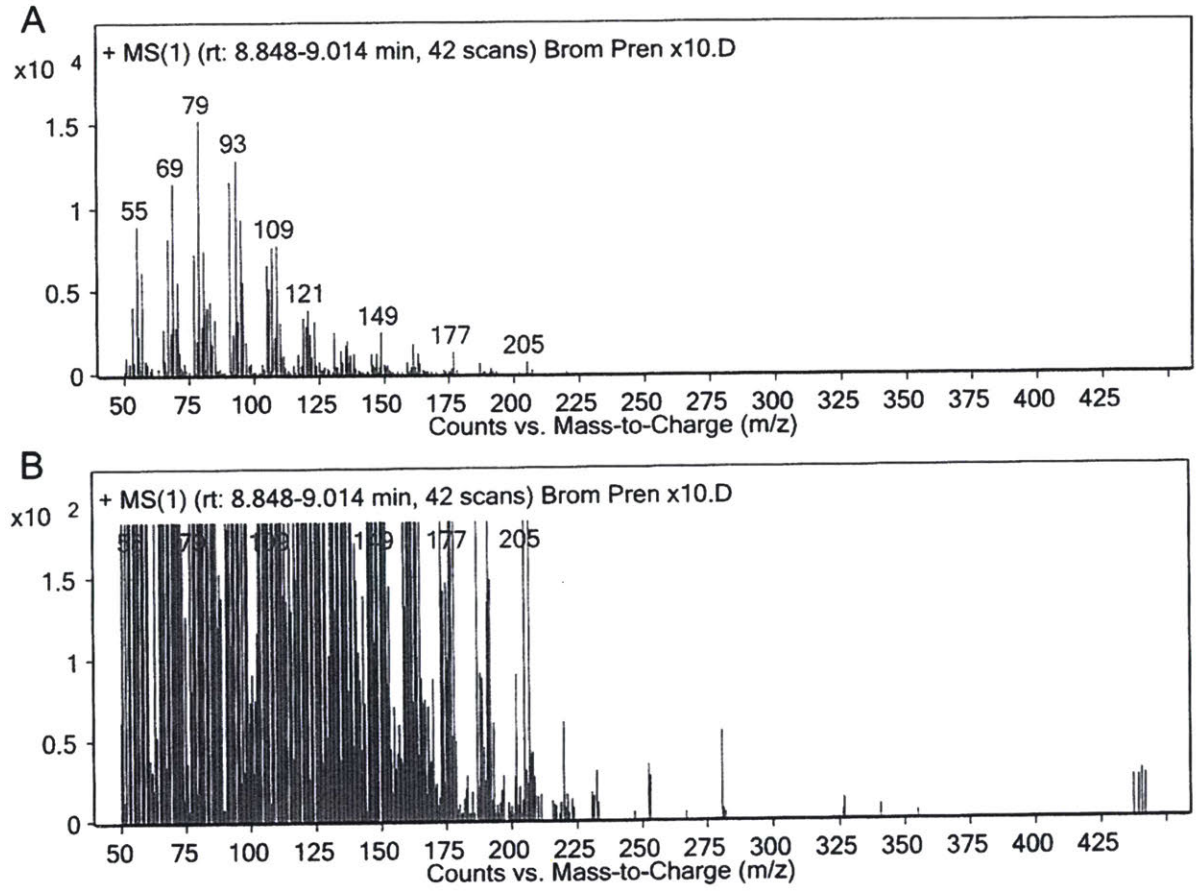


Figure 6-4: Spectrum of peak that could possibly be attributed to triply brominated amorphadiene: (**A**) The full mass spectrum, (B) Previous spectrum zoomed in to highlight the largest observed ions (m/Z 440, 442 and 438), which are consistent with the parent ions expected for triply brominated amorphadiene.

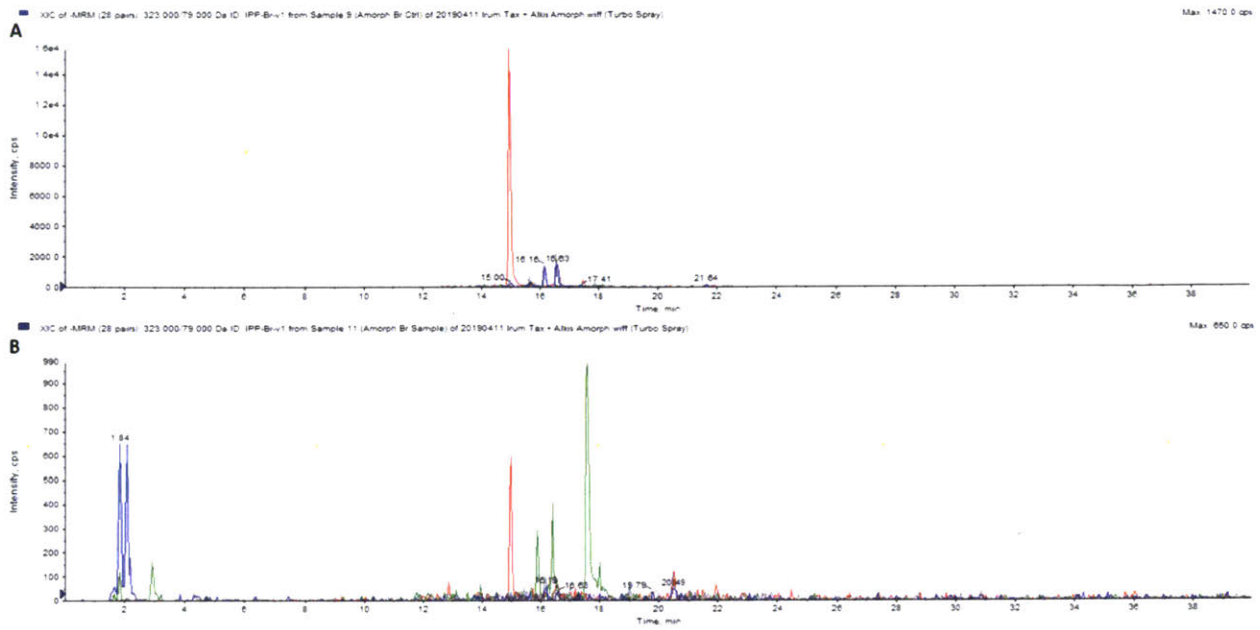


Figure 6-5: LC-MS/MS chromatograms for brominated IPP: (**A**) No enzyme control. (**B**) Regular run. In addition to the peaks appearing in the no enzyme control, new peaks appear at ~2min (blue, when scanning for IPP brominated with ^{79}Br) and at ~16min (green, when scanning for IPP brominated with ^{81}Br)

We were unfortunately unable to get conclusive proof of the formation of brominated amorphadiene from our LC-MS/MS or GC-MS results, even though we were able to observe non-functionalized amorphadiene when we used non-functionalized isoprenol as the substrate. When we increased enzyme concentration ten-fold, we were able to observe a GC-MS peak that could possibly be triply brominated amorphadiene, based on the m/Z of the largest observed ions, which did align with the predicted m/Z of the parent ions. However, given the very small chromatogram peak observed, further isolation, purification and definitive identification via NMR was not possible. Furthermore, our LC-MS/MS results were not able to yield any conclusive insights. While in our LC-MS/MS results, noticed the appearance of new peaks, that is new compared to the peaks that appear in the "no enzyme" control, in the chromatogram for brominated IPP, which could indicate the appearance of the brominated intermediate, the fact that these new peaks, which were for different "versions" of brominated IPP, either carrying ^{79}Br or ^{81}Br, eluted at wildly different times (~2 min, near where non-functionalized IPP would appear or ~16 min) indicated

that at least one of them must be a false peak, making identification difficult.

The appearance of unexplained peaks in the chromatogram also confounded our LC-MS/MS search. For example, although it initially appeared that we could observe the brominated equivalent of IP in LC-MS/MS, based on expected parent and daughter ions, this particular peak also appeared in the "no enzyme" control as well, appearing as if brominated prenol is spontaneously phosphorylated or that a second bromine is present in the molecule. This latter possibility arose because the m/Z of the phosphate group (i.e. 79) happens to be the same as that of $^{79}Br^-$, a fact which acts as a confounding factor in identifying the peak. The appearance of this peak could also be explained by either a contamination/impurity or degradation in our feedstock or some potential rearranging & combination of ions while the brominated prenol is being ionized prior to entering the MS chamber. It ultimately was not possible to identify any new peaks not appearing in the "no enzyme" control that could potentially be attributed to brominated variants of IP/IPP/GPP/FPP. Thus, our experiments with brominated prenol were inconclusive.

As a next step, we chose to proceed with the purchase of 2-chloro-3-methylbut-3-en-1-ol (chlorinated isoprenol), for which we identified a new vendor that could provide it at a lower cost. It was hoped that the fact that it was a variant of isoprenol instead of prenol would make its conversion into product by the IUP more likely, given IUP's preference for isoprenol over prenol. Furthermore, it was hoped that possible steric inhibitions caused by the bromine atom's large size could be alleviated by switching to chlorine, whose atom is smaller.

In this case, we expected to produce the triply chlorinated amorphadiene analog, with expected molecular weight of 307.68 and expected m/Z ratios (and relative abundances) for the parent ions as 306.07 (100%), 308.07 (95.9%), and 310.06 (30.6%), as predicted by ChemDraw, which we tried to identify via GC-MS, as well as the hypothetical chlorinated intermediates, shown in Table 6-2, for which we scanned via LC-MS/MS.

We were unfortunately unable to identify any peaks that would hint at the production of chlorinated amorphadiene in the GC-MS. In our LC-MS/MS results, we did however notice a new

peak appear in the chromatogram for chlorinated GPP (compared to the peaks that appear in the "no enzyme" control), which might indicate the appearance of the chlorinated intermediate. However, further experiments are needed to replicate this single result with a better signal/noise ratio and identify ways for producing subsequent metabolites.

Table 6-2: Expected metabolic intermediates derived from chlorinated isoprenol

ID	Formula	Exp. Parent Ion m/Z(1)	Exp. Daughter Ion m/Z(2)
Chlor. IP	$C_5H_8ClO_4P^{2-}$	199/201	79
Chlor. IPP	$C_5H_8ClO_7P_2^{3-}$	279/281	79
Twice Chlor. GPP	$C_{10}H_{15}Cl_2O_7P_2^{3-}$	381	79
Thrice Chlor. FPP	$C_{15}H_{22}Cl_3O_7P_2^{3-}$	483	79

(1): Multiple different ions scanned to account for the two isotopes ^{35}Cl and ^{37}Cl, with isotopic abundances of 76% and 24% respectively. In the case of Chlor. GPP & FPP only the ions containing ^{35}Cl and not ^{37}Cl were scanned for.
(2): 79 is the m/Z ratio of a phosphate group, mirroring the fragmentation patterns of non-halogenated metabolic intermediates

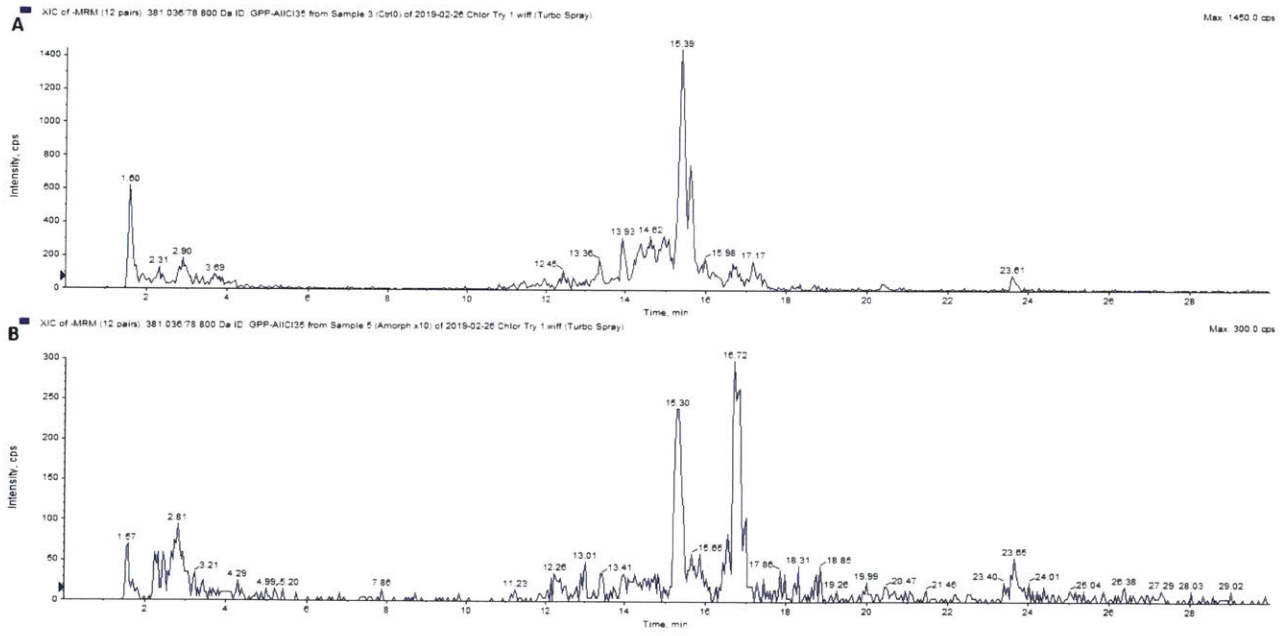


Figure 6-5: LC-MS/MS chromatograms for doubly chlorinated GPP: (**A**) No enzyme control. The only significant peak appears at ~15.4 min (**B**) Regular run. In addition to the peak at ~15.3 min, which also appears in the no enzyme control, a new peak appears at ~16.7 min

6.2.2 Production of chlorinated limonene

As a way of widening our search into different final molecules, we also investigated combining the IUP and isoprenoid backbone synthesis enzymes with limonene synthase (LS) to look into

the possibility of producing a chlorinated analog. Limonene was chosen as a molecule that is simpler than amorphadiene (monoterpenoid vs sesquiterpernoid) and because we are already capable of producing non-chlorinated limonene *in vitro* (12). We expected the doubly chlorinated limonene analog to have a molecular weight of 205.12 and expected m/Z ratios (and relative abundances) for the parent ions as 204.05 (100%), 206.04 (63.9%), and 205.05 (10.8%), as predicted by ChemDraw. Experiments were run in a similar way as the experiments for the production of amorphadiene, but the experiment for limonene production substituted LS for ADS.

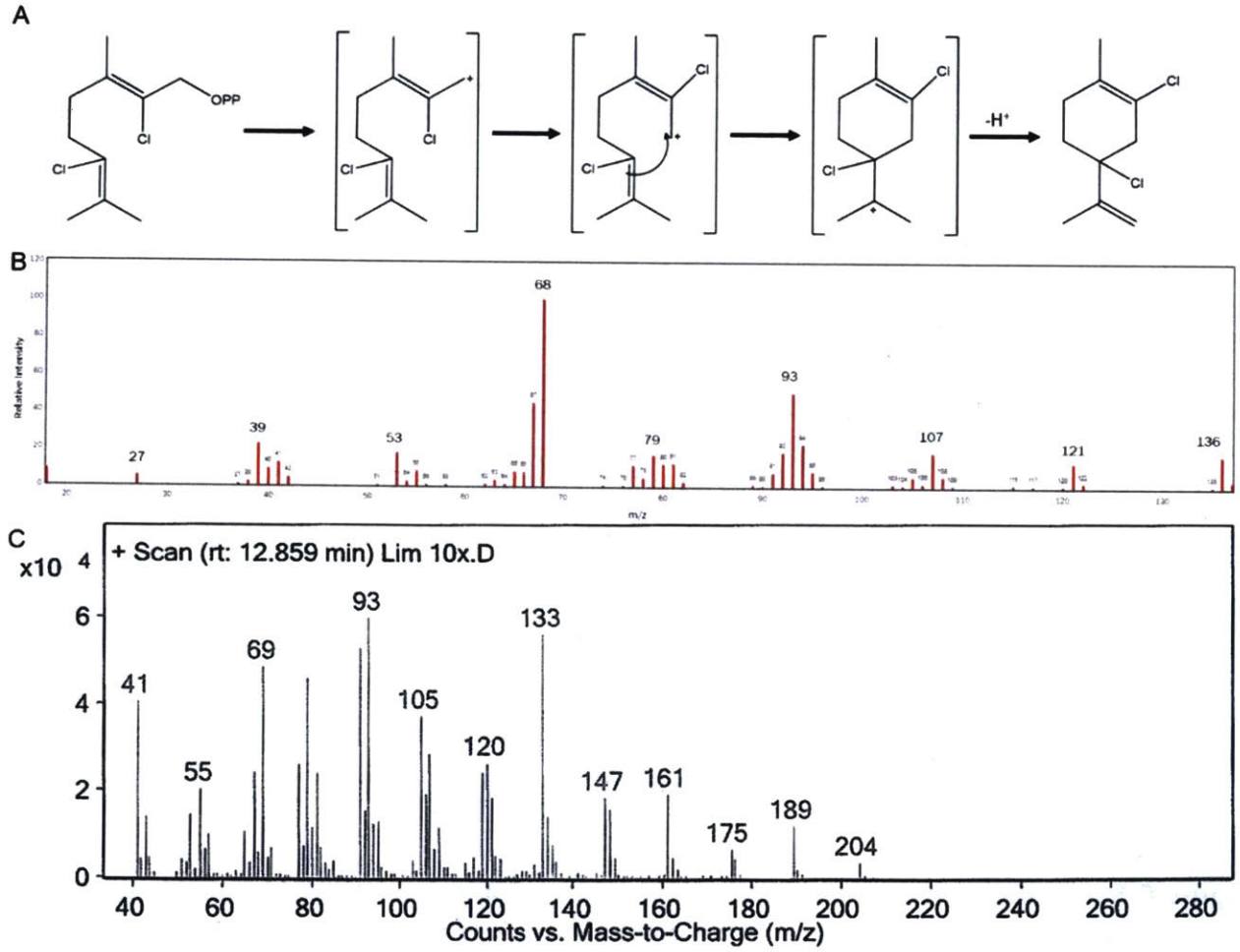


Figure 6-7: Doubly chlorinated limonene: (**A**) It is hypothesized that doubly chlorinated GPP (left) can be converted to doubly chlorinated limonene (right) through the activity of the enzyme limonene synthase, following a mechanism analogous to that of the cyclization of GPP to limonene (22), (**B**) Mass spectrum of limonene (**C**) Mass spectrum of peak initially identified as doubly chlorinated limonene

An interesting peak we observed in the GC-MS chromatogram appeared, on first glance, to potentially be a peak for chlorinated limonene (Figure 6-5 C). The mass spectrum assigned to this peak appeared to have a molecular ion with m/Z ratio of 204, which was identical to the m/Z ratio expected for most abundant of the parent ions of twice chlorinated limonene. Furthermore, the fragmentation pattern exhibited in that peak was quite similar to the fragmentation pattern of natural limonene, with key ions, or ions very close to them appearing in both spectra, such as ions with m/Z such as 55 & 53 in our spectrum vs 53 in literature spectrum for limonene, 67 & 68 vs 69, 107 vs 105 &107 and 93 which appears both in our spectrum and the literature one. This experiment was replicated under different enzyme concentrations, going up to 10-fold increase, and several control experiments were also performed.

Unfortunately, the follow-up experiments revealed that this peak is could be attributed to a potential degradation product of caryophyllene, an internal standard we used in our GC-MS runs for quantification purposes. We have therefore also been unsuccessful at producing chlorinated limonene, but have acquired insights on experiment controls and challenges we will encounter on peak identification.

6.3. Conclusions and future directions

As mentioned earlier, although we do have some encouraging hints that our approach could work, we have yet been unsuccessful in producing a workable proof-of-concept for our idea of using the IUP for the production of functionalized isoprenoids.

One major confounding factor for our experimental efforts has been the fact that our search our search for the products in GC-MS and intermediates in LC-MS/MS was happening "in the blind", i.e. we are searching for ions based on their expected m/Z ratios and/or fragmentation patterns rather than by comparison to literature values or to standards. Furthermore, we are not certain whether these new molecules will be separable by our gas/liquid chromatography columns

and whether the same chromatographic methods that we have been using for non-functionalized isoprenoids and the intermediates leading up to them can also be used for their functionalized analogs.

These two issues could be addressed if we were to first set up and calibrate our methods using external standards Since the molecules we want to produce are completely novel, commercial standards for us to use unfortunately do not exist. Future work on this field could potentially include synthesizing said intermediates in our lab, or in collaboration with a different lab.

Additional substrates can be tried in order to try overcoming possible issues with the polarity and/or steric inhibitions large bromine or chlorine molecules could cause. One avenue that could be explored would be to procure or produce fluorinated isoprenol, which is expected to be easier to process through the IUP, being less bulky than brominated or chlorinated isopentenols. Other substrates that could be considered would be ones that have functional groups other than halogens, such as methyl, thiol, amino or carbonyl groups.

6.4. Materials and Methods

Strains, plasmids and genes; cloning protocol

Materials and methods were the same as the ones as described in the relevant sections of Chapter 2. Refer to **Table A1** in the Appendix for a listing of strains and to **Table A2** in the Appendix for a listing of genes.

Construction of expression vectors.

Vectors for enzyme expression were constructed by introducing a His-tagged gene (for ScCK, AtIPK, IDI, IspA, ADS or LS expression) into a pET-28 a (+) vector in the following fashion: Backbone fragments were amplified from pET-28 a (+) vector using the primer pair GB pET28-HisT-vec_f/r. The insert fragments were amplified as follows: the fragment containing the genes

ScCK expression was amplified from *S. cerevisiae* genomic DNA using the primer pair GB-pET28-ScCK_f/r, the fragments containing AtIPK and IDI were amplified from plasmid pSEVA228-pro4IUPI using the primer pairs GB-pET28-atIPK_f/r and GB-pET28-IDI_f/r, the fragments containing IspA and ADS were amplified from plasmid p5T7-ispA-ads using the primer pairs GB-pET28-IspA_f/r and GB-pET28-ADS_f/r and the fragment for LS expression was amplified from plasmid p5T7gpps-ls using the primer pair GB-pET28-LS_f/r. Vector sequences for the primers used are listed in **Table A3** in the Appendix. Codon-optimized sequences are listed in **Table A5** in the Appendix.

Multi-enzyme *in vitro* reactions.

The method for multi-enzyme reactions was adapted from (12). In short, all enzymes were suspended in 50 mM ammonium bicarbonate (pH 7.4) with 10 mM $MgCl_2$, 2 mM $MnCl_2$ and 0.05% (w/v) Tween 20. At t=0 10 mM ATP, and 5 mM chlorinated isoprenol or brominated prenol were added and the solutions were incubated at 37°C (for functionalized amorphadiene) or 30°C (for functionalized limonene). Samples that were to be used for GC-MS analysis were overlaid with 1/10 volume of dodecane. Enzyme concentrations used were 25 µg/mL ScCK, 15 µg/mL atIPK, 25.4 µg/mL IDI, 37.15 µg/mL IspA, and 25 µg/mL for ADS or LS.

Detection of functionalized metabolites, amorphadiene and limonene.

Functionalized IP/DMAP, IPP/DMAPP, GPP, FPP were scanned for via LC-MS/MS using the means and method described in Chapter 4, with the only difference that the ions scanned for were the ones listed in Tables 6-1 or 6-2 if brominated prenol or chlorinated isoprenol was used as substrate. Functionalized limonene and amorphadiene were scanned for quantified using GC-MS using the same protocol as the one for the quantification of limonene and amorphadiene or taxadiene (Chapter 4), edited so that the MS would scan at m/Z ratios high enough to be able to detect the expected m/Z ratios for the respective hypothesized parent ions.

6.5. Bibliography

1. Gao S, et al. (2008) Substrate promiscuity of pyruvate kinase on various deoxynucleoside diphosphates for synthesis of deoxynucleoside triphosphates. *Enzyme Microb Technol* 43(6):455–459.
2. Li Y, et al. (2011) Substrate promiscuity of n-acetylhexosamine 1-kinases. *Molecules* 16(8):6396–6407.
3. Chen Y, Zhou YJ, Siewers V, Nielsen J (2015) Enabling Technologies to Advance Microbial Isoprenoid Production. *Adv Biochem Eng Biotechnol* 148:143–160.
4. Ward VCA, Chatzivasileiou AO, Stephanopoulos G (2018) Metabolic engineering of Escherichia coli for the production of isoprenoids. *FEMS Microbiol Lett* 365(10):1–9.
5. George KW, et al. (2018) Integrated analysis of isopentenyl pyrophosphate (IPP) toxicity in isoprenoid-producing Escherichia coli. *Metab Eng* 47(March):60–72.
6. Guo W, Sheng J, Feng X (2017) Mini-review: In vitro Metabolic Engineering for Biomanufacturing of High-value Products. *Comput Struct Biotechnol J* 15:161–167.
7. Zhu F, et al. (2014) In vitro reconstitution of mevalonate pathway and targeted engineering of farnesene overproduction in Escherichia coli. *Biotechnol Bioeng* 111(7):1396–1405.
8. Galloway DA, Laimins LA, Division B, Hutchinson F (2015) In Vitro Reconsitution of Metabolic Pathways: Insights into Nature's Chemical Logic. *Synlett* 26(8):87–92.
9. Chen X, Zhang C, Zou R, Stephanopoulos G, Too HP (2017) In Vitro Metabolic Engineering of Amorpha-4,11-diene Biosynthesis at Enhanced Rate and Specific Yield of Production. *ACS Synth Biol* 6(9):1691–1700.
10. Korman TP, Opgenorth PH, Bowie JU (2017) A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nat Commun* 8(May):1–8.
11. Boronat A, Rodriguex-Concepcion M (2015) Terpenoid Biosynthesis in Prokaryotes. *Adv*

Biochem Eng Biotechnol 148:3–18.

12. Ward VCA, Chatzivasileiou AO, Stephanopoulos G Cell free biosynthesis of isoprenoids from isopentenol. *Prep*.

13. Kosjek T, Heath E (2012) Halogenated Heterocycles as Pharmaceuticals. *Halogenated Heterocycles: Synthesis, Application and Environment*, ed Iskra J (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 219–246.

14. Ro DK, et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086):940–943.

15. Amyris GN (2016) Amyris Creates Program to Reduce the Cost and Increase Access to Leading Malaria Treatment. 2016.

16. Peplow M (2016) Synthetic biology's first malaria drug meets market resistance. *Nature* 530(7591):389–390.

17. Lu F, Culleton R, Zhang M, Ramaprasad A, von Seidlein L (2017) Emergence of Indigenous Artemisinin-Resistant Plasmodium falciparum in Africa. *N Engl J Med* 376(10):991–993.

18. Ashley EA, et al. (2014) Spread of Artemisinin Resistance in *Plasmodium falciparum* Malaria. *N Engl J Med* 371(5):411–423.

19. Fairhurst RM, Dondorp AM (2016) Artemisinin-resistant Plasmodium falciparum malaria. *PLoS Pathog* 11(4):1–14.

20. Mellegaard-Waetzig SR, Wang C, Tunge JA (2006) Selenium-catalyzed oxidative halogenation. *Tetrahedron* 62(30):7191–7198.

21. Tang X, Allemann RK, Wirth T (2017) Optimising Terpene Synthesis with Flow Biocatalysis. *European J Org Chem* 2017(2):414–418.

22. Mann JC, Hobbs JB, Banthorpe D V., Harborne JB (1994) *Natural Products: Their Chemistry and Biological Significance*.

Chapter 7

Improving the productivity of the MEP pathway

7.1. Introduction

In previous chapters we focused on the creation of a novel isoprenoid pathway, the Isopentenol Utilization Pathway (IUP) in order to circumvent the limitations posed by the native MEP and MVA pathways. In this chapter, we will outline aspects of the work which we performed chronologically prior to the work on the IUP. This work will help highlight the complexity of working with one of the native pathways and will illustrate some of the bottlenecks that the IUP overcomes.

In choosing with which of the two native pathways we should work, we investigated the current productivity, as well as the long-term potential of either pathway. While the MVA pathway has been heterologously expressed in *E. coli* with success, outperforming the native MEP pathway(1), studies comparing the efficiency of the MEP and MVA pathways through *in silico* analysis have come to the conclusion that the MEP pathway is more efficient than the MVA pathway with regards to theoretical yield, carbon utilization and productivity in the conversion of glucose to isoprenoids (2–5). This promise of higher theoretical performance and the great potential for improvement is what lead us to occupy ourselves with the attempt to increase the productivity of the MEP pathway for the production of isoprenoids.

7.2. Two important enzymes in the MEP pathway: IspG & IspH

7.2.1 Fe-S clusters are essential to IspG & IspH activity

IspG and its downstream enzyme IspH catalyze the two last steps of the MEP pathway. These two enzymes are similar in several ways. Both enzymes have been found to contain [4Fe-4S] clusters, which are linked to their catalytic activity(6). UV and Mößbauer spectra of both IspG and IspH show the presence of an $[4Fe\text{-}4S]^{2+}$ cluster in both enzymes (7, 8). In both cases, the [4Fe-4S] cluster is coordinated by three cysteine residues, which are conserved among species. A mutation in any of the conserved cysteine residues leads to a significant drop in enzyme catalytic activity, roughly in the order of 10^5 (9, 10). Furthermore, both enzymes catalyze a reductive dehydroxylation of their substrates. However, IspG must also catalyze the ring opening of MEC (10).

The most important factor in maintaining the catalytic activity of IspG and IspH is ensuring that their [4Fe-4S] centers do not get damaged. It has been reported that IspG/IspH catalytic activity in aerobic conditions is reduced due to the clusters being oxygen sensitive and damaged by oxygen (6), other reactive oxygen species such as H_2O_2 (11) and other univalent oxidants. This oxidative inactivation is a phenomenon not unique to IspG/IspH but rather a common phenomenon in Fe-S proteins such as dehydratases (11–13), aconitases (11, 14) and others. This phenomenon has been linked to iron oxidation which then leads to destabilization of the Fe-S cluster (13). Thus, the inactivation of the Fe-S clusters is directly related to the oxidation states of their iron atoms. In literature, there is no information about the oxidation state of the iron atoms in the IspG Fe-S cluster. However, in the case of IspH, Mößbauer spectroscopy has uncovered a $[4Fe\text{-}4S]^{2+}$ cluster in which one pair of iron atoms is valence-trapped (Fe^{3+}-Fe^{2+}) while the other shares a delocalized excess electron ($Fe^{2.5+}$-$Fe^{2.5+}$), but it has also been hypothesized that *in vivo* these pairs might change so that all four iron atoms are in the $Fe^{2.5+}$ state (8). Dehydratases,

which also have a [4Fe-4S] cluster have been reported to contain two Fe(II) and two Fe(III) atoms (13).

Univalent oxidants affect the $[4Fe\text{-}4S]^{2+}$ clusters by oxidizing them into $[4Fe\text{-}4S]^{3+}$. The next step is the loss of one Fe^{2+} to produce the catalytically inactive $[3Fe\text{-}4S]^{+}$ cluster (12). The exact details of the cluster oxidation vary by the oxidant used and by the organism. Examples of the proposed mechanism for the oxidation of a $[4Fe\text{-}4S]^{2+}$ cluster by reactive oxygen species are (11, 15):

$$[4Fe\text{-}4S]^{2+} + H_2O_2 \rightarrow [4Fe\text{-}4S]^{3+} + OH^- + HO^{\bullet} \textit{ or } [4Fe\text{-}4S]^{2+} + O^{-2} + 2H^+ \rightarrow [4Fe\text{-}4S]^{3+} + H_2O_2$$

$$[4Fe\text{-}4S]^{3+} \rightarrow [3Fe\text{-}4S]^{+} + Fe^{2+}$$

It is therefore apparent that IspG and IspH activity can be diminished in the presence of reactive oxygen species and this can create issues in the aerobic production of isoprenoids.

7.2.2 Attempting to overcome the bottleneck posed by IspG

One of our primary goals was the retention of IspG/IspH activity in aerobic conditions. As has already been mentioned, these enzymes show diminished activity aerobically due to the oxidation of their Fe-S clusters. This has also been observed in literature (16) and in our lab, since past experimental data showed increased MEC accumulation in aerobic conditions, while the levels of the same remained very small in anaerobic conditions. Furthermore, past research in our lab has indicated that the overexpression of IspG can lead to better performance of the MEP pathway (16), especially in aerobic conditions, which further supports the notion that IspG performance is a bottleneck.

We conceived of three broad ways in which the aerobic activity of IspG could potentially be retained.

1. Continually replace or repair the Fe-S clusters as they are being destroyed, through the cell's innate Fe-loading pathways, such as those expressed by the *isc* and *suf* operons

(12, 17), along with an array of putative repair enzymes such as *yftE* (18), *ftnA* (19), *ftnB* (20) and *bfr* (19).

2. Attempt to prevent ROS from reaching the Fe-S cluster. This could be achieved through the addition of antioxidant molecules in the culture media that will screen the ROS, following an example found in literature (14). Other possibilities would include the production of said molecules in the cell itself or the overexpression of enzymes known to act as ROS scavengers such as catalase or superoxide dismutase (21).
3. The most intriguing way of dealing with this problem, however, would be to attempt to make IspG itself be naturally tolerant to attacks by ROS. While this would be a very challenging task, the high impact of such a possibility made it an avenue worth pursuing.

7.2.3 Strategy for engineering an oxygen tolerant IspG enzyme

Although many Fe-S enzymes in the cell are inhibited by oxygen, there have been reported cases of Fe-S containing enzymes where exposure to oxygen does not lead to significant decline of their catalytic activity. A preliminary, non-exhaustive search of enzymes which, like IspG/IspH exhibit [4Fe-4S] clusters uncovered three examples of oxygen-tolerant iron-sulfur enzymes. The first two examples were NADH dehydrogenase (Complex I) and succinate-coenzyme Q reductase (SQR) (Complex II) which are enzymes that participate in the respiratory chain and Krebs cycle respectively (22, 23). Since they participate in pathways that are naturally aerobic, their oxygen tolerance has not been investigated, since it is taken for granted.

The third case, however, concerns hydrogenases, which are enzymes that catalyze the reaction $H_2 \leftrightarrow 2H^+ + 2e^-$ (24). There are two variants of hydrogenases, oxygen sensitive (usually cytosolic) and oxygen tolerant (usually membrane-bound), both variants having similar overall structure. They are composed of large subunit carrying the Ni-Fe active site, a small subunit containing three Fe-S clusters and a cytochrome, which, in the case of oxygen sensitive hydrogenases is cytochrome B and in the case of oxygen tolerant hydrogenases is cytochrome c

(25). Although the structure of the active site is not different between oxygen tolerant/sensitive hydrogenases, the oxygen-tolerant hydrogenase variant features in its small subunit a unique [4Fe-3S] cluster proximal to the active site coordinated by six conserved cysteine residues, two of which are consecutive, instead of the more common [4Fe-4S] cluster coordinated by four cysteines that can be found in corresponding, anaerobic (oxygen-sensitive) hydrogenases (24, 26, 27).

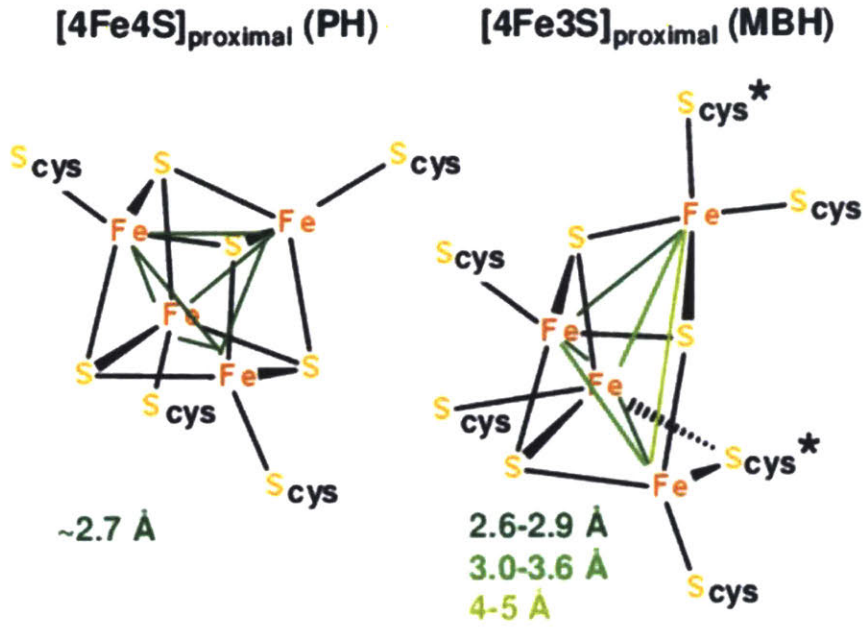


Figure 7-1: Differences between the proximal Fe-S clusters of oxygen-sensitive (PH) and oxygen-tolerant (MBH) hydrogenases: Supernumerary cysteines shown with a *. Figure from (28)

Such consecutive cysteines have also been found coordinating Fe-S clusters in Complex I. In the case of Complex I, it was found that consecutive cysteine residues confer structural flexibility since both or either of the consecutive cysteines can coordinate with the Fe-S cluster (26). In oxygen-sensitive hydrogenases, the extra two cysteine residues are replaced by glycine residues (29). When an oxygen-tolerant hydrogenase was mutagenized to replace the two extra cysteines with glycines, it became sensitive to inhibition by oxygen (26, 29). and the proximal cluster's configuration was found to have converted to [4Fe-4S] (29). The two different cluster variants

behave differently when oxidized. Whereas in the oxygen-sensitive cluster the series of oxidations is $[4Fe\text{-}4S]^{+} \leftrightarrow [4Fe\text{-}4S]^{2+} \rightarrow [4Fe\text{-}4S]^{3+}$, with the last step being irreversible, contributing to the inactivation of the cluster, the oxygen-tolerant cluster can be oxidized $[4Fe\text{-}3S]^{3+} \leftrightarrow [4Fe\text{-}3S]^{4+} \leftrightarrow [4Fe\text{-}3S]^{5+}$, with all the steps being reversible (30). This would mean that the [4Fe-3S] cluster is not as easily inactivated by oxidation.

7.3. Work on engineering an oxygen-tolerant IspG

7.3.1 Background and approach

As has been stated previously, an intriguing way of addressing the problem of IspG's sensitivity to oxygen could be to adopt a protein engineering approach to create a better performing IspG, hopefully increasing the overall yield of the MEP pathway. This approach would involve mutagenizing IspG to have it incorporate traits that are believed to lead to oxygen sensitivity in enzymes with similar Fe-S clusters.

The approach that we chose to take is to attempt to mimic the features of oxygen-tolerant hydrogenases. It has been reported for membrane-bound (oxygen tolerant) hydrogenases that oxygen resistance may be less due to O_2 exclusion from the active site and more due to the structure of the electron carrier system (26), and the unique [4Fe-3S] cluster has been identified as being a crucial factor for the operation of the electron system and therefore this tolerance. Our chosen approach revolved around mutagenizing residues near the active site into cysteines with the objective or trying to recreate the kind of oxygen-tolerant [4Fe-3S] cluster coordinated by six cysteines that has been reported for hydrogenases. Mutation of an oxygen-sensitive hydrogenase has been shown in literature to abolish oxygen tolerance when the additional two cysteines found only in oxygen-tolerant hydrogenase variants are mutated into glycines to make the enzyme resemble an oxygen-sensitive variant (26, 29). Furthermore, it has been found that the greater reduction in oxygen tolerance is effected by replacing one of the two consecutive cysteines with

a glycine, whereas replacing the other supernumerary cysteine also lead to a drop in oxygen sensitivity, but not as significant (26, 29).

In our research plan, we would attempt to do the opposite of what has been attempted in literature for hydrogenases, that is to introduce cysteines in the place of glycines and determine if that would lead to increased oxygen tolerance. Furthermore, we would take an incremental approach, under which we would first introduce a single supernumerary cysteine next to a pre-existing cysteine (preferably in the place of a glycine, in order to mimic the reverse approach taken in literature with hydrogenases), in order to create the Cys-Cys tandem system which has been tentatively linked to oxygen tolerance in hydrogenases. Assuming incremental oxygen tolerance was achieved the second cysteine would then be introduced, bringing the total number of cysteines to 6, same as in the hydrogenase.

If rational mutagenesis is to be carried out, knowledge of the protein structures is required, as this would lead to the identification of the active site as well as the identification of the specific cysteine residues coordinating the [4Fe-4S] cluster, so that residues adjacent to them can be mutagenized. A search through the Protein Data Bank (PDB) only yielded a small number of structires we could use in our work. The structures that we found useful were IspG structures from *Thermus Thermophilus HB8* (ID: 4S38 (31)), *Thermus Thermophilus HB27* (ID:4G9P (31)) and an IspG structure from *Aquifex Aeolicus* (ID: 3NOY(10)). Fortunately, the *Thermus Thermophilus HB8* variant (4S38) has been studied *in vitro* in the past (32) and has been found to be a rather active variant. Interestingly enough, an IspG structure from *E. coli* was not found. We chose to order codon optimized genes for the *Thermus Thermophilus HB8*, and *Aquifex Aeolicus* in our study. Furthermore, we also elected to use the *E. coli* IspG variant in our study, in which case we chose mutagenesis targets after developing a simple homology model to determine the probable location of the active site.

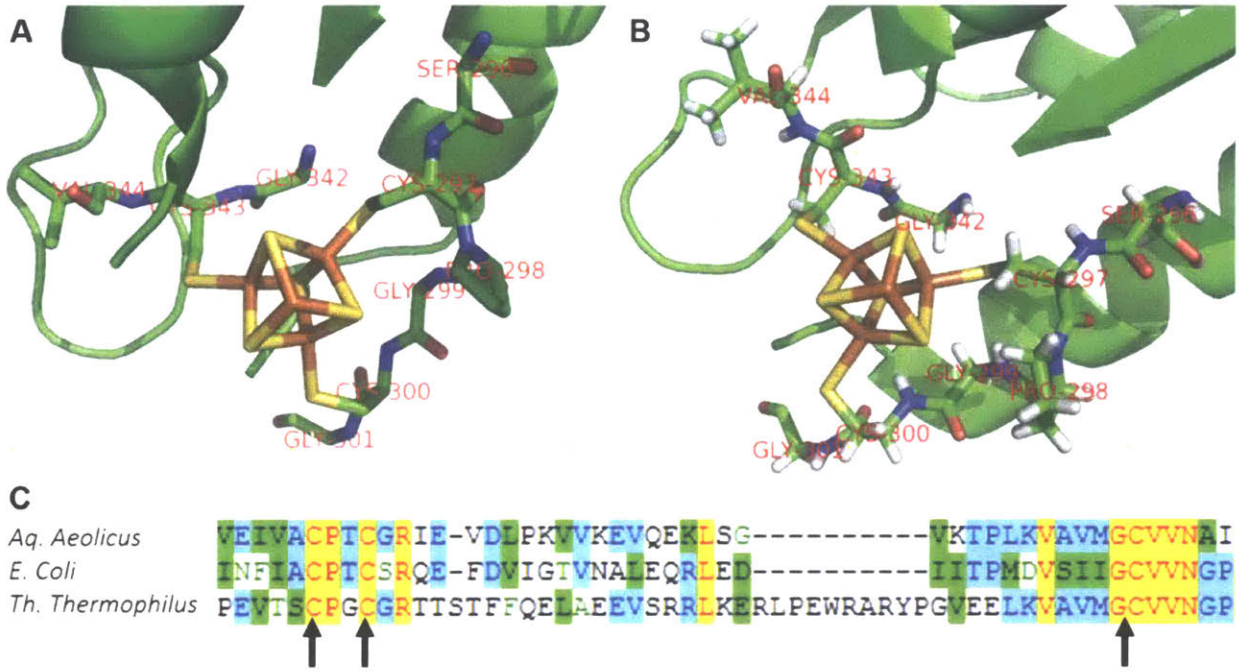


Figure 7-2: Identifying residues to mutagenize in IspG: Crystal structures of IspG from *A. Aeolicus* (**A**) and *T. Thermophilus* HB8 (**B**) focusing on the [4Fe-4S] cluster and its immediate environment. (C): Homology sequence between the *A. aeolicus*, *T. Thermophilus* HB8 and *E. coli* IspG homologs. Arrows highlight the positions of cysteines

Based on the above, we created a list of residues that we believed would be good targets for mutagenesis. This list is shown below:

Table 7-1: List of cysteine residues that coordinate with the [4Fe-4S] cluster in IspG variants and residues adjacent to them

A. Aeolicus IspG			*T. Thermophilus* IspG			*E. Coli* IspG		
Ala-264	Cys-265	Pro-266	Ser-296	Cys-297	Pro-298	Ala-268	Cys-269	Pro-270
Thr-267	Cys-268	**Gly-269**	**Gly-299**	Cys-300	**Gly-301**	**Thr-271**	Cys-272	**Ser-273**
Gly-299	Cys-300	Val-301	**Gly-342**	Cys-343	Val-344	**Gly-303**	Cys-304	Val-305
Residues to be mutagenized to Cys shown in **bold**								

For the first round of mutations, wherein we would introduce the fifth cysteine, we designed 11 possible variants for *ispG* genes. Four variants were from from *E. coli* (EC-WT, EC-271, EC-273 & EC-303), four from *T. Thermophilus* (TT-WT, TT-299, TT-301 & TT-342) and three from *A.*

Aeolicus (AA-WT, AA-269 & AA-299). The suffix WT indicates a wild type gene (not mutated), whereas a numerical suffix indicates a residue that has been mutated into a cysteine.

In order to assay the efficacy of our mutations we based our assay on the output of the pathway, as measured by a cell's ability to produce taxadiene To that effect, we would employ a strain previously developed in our lab for taxadiene overproduction (MG1655 DE3 ΔaraA::T7MEP ΔlacY::T7TG), referred to as the T7-MEP strain (4). We would then introduced the mutant *ispG* genes via a plasmid and expressed them. Our thinking was that if the IspG produced by the *ispG* mutant genes on plasmids has been made more oxygen-tolerant due to the mutations we have introduced, then flux would be shifted to go through the oxygen-tolerant IspG, leading to increased taxadiene output and decreased concentrations of MEC, its substrate, as previous results had indicated that there were no major bottlenecks downstream of IspG.

We then created mutated versions of the *ispG* genes and integrated them in a pSEVA228 (33) plasmid. The advantages of this plasmid are its low copy number and the fact that it has a titratable promoter, which is induced using m-toluic acid. This would allow us to tune the expression levels of the mutant *ispG* genes independently of the MEP pathway an inducible and upregaluated version of which had been integrated in the genome of the T7-MEP strain and could be induced by IPTG. Having created the plasmids, we then transformed them into the T7-MEP strain, along with a plasmid for taxadiene overproduction and proceeded to run our assays. The setup of the assay was that of an aerobic shake-flask culture, so as to have oxygen enter our system and partially inactivate the native IspG, with the hope that our mutants, if oxygen-tolerant, would not be inactivated and enhance titer.

7.3.2 Experimental results

We proceeded to run shake-flask experiments for all our strains, i.e. strains expressing one of the aforementioned 11 variants of *ispG* through a plasmid as well as no vector (denoted as TAX) and

empty vector (denoted as X) strains. The results of the shake flask experiments can be summarized as follows:

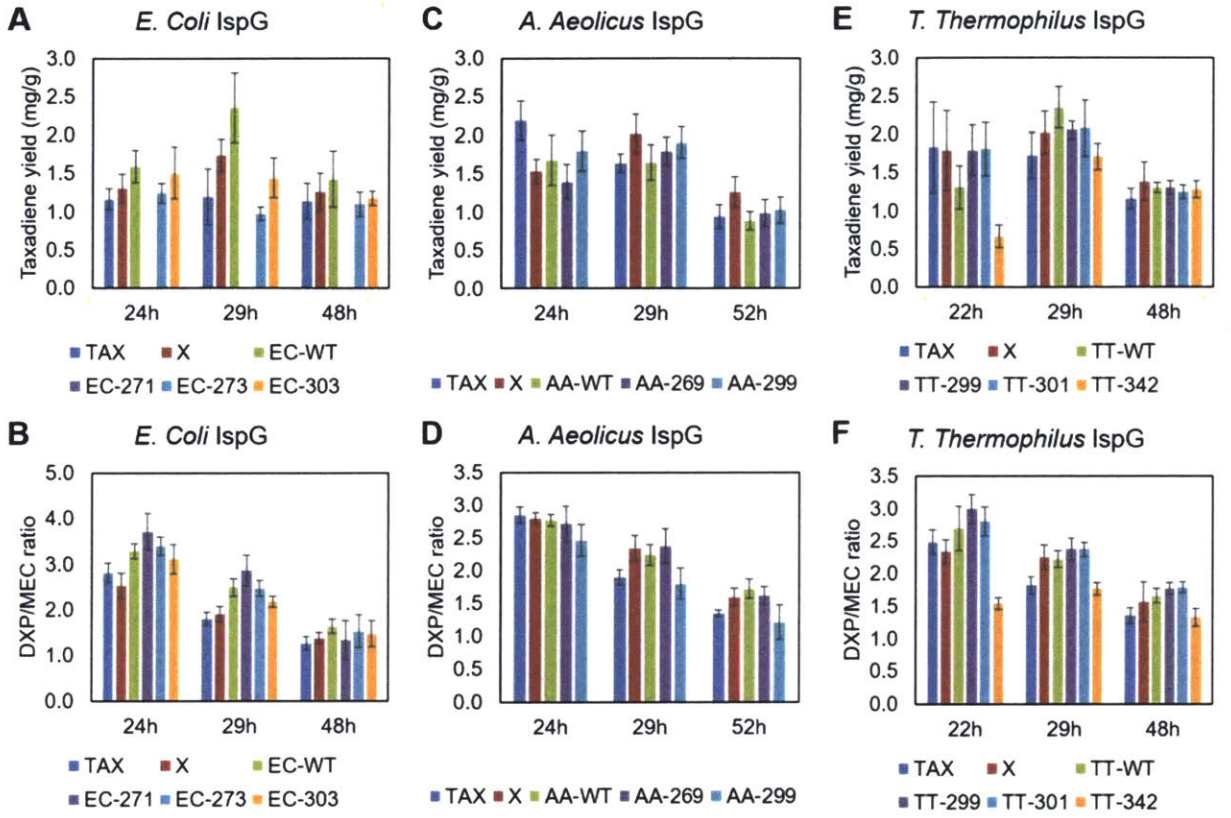


Figure 7-3: Taxadiene yields and DXP/MEC ratios achieved in shake-flask runs utilizing different in IspG mutants: *E. coli* strains bearing a vector for the expression of an *ispG* mutant from either *E. coli* (EC) (**A&B**), *A. aeolicus* (AA) (**C&D**) or *T. thermophilus* (TT) (**E&F**) are cultured in shake flasks. Runs are identified by the type of *ispG* included in mutant, denoted in the format (Origin)-(residue mutated to Cys). No vector is denoted as TAX and empty vector is denoted as X. (**A**,**C**,**E**): Taxadiene yield (in mg taxadiene produced per g glucose consumed), (**B**,**D**,**F**): Molar ratios of DXP (1st intermediate in MEP pathway) over MEC (5th intermediate in MEP pathway and substrate to ispG).

For the experiments involving the use of an *ispG* gene from *E. Coli* (EC-WT, EC-271, EC-273 & EC-303), it was found that the expression of an additional *ispG* gene did indeed lead to lowering of the MEC levels compared to the no vector and empty vector strains. Furthermore, we discovered that the non-mutated version of IspG outperformed the mutants with regards to taxadiene production, in line with previous findings in the literature about the enhancement of the

MEP pathway (16). The expression of EC-273 and EC-303 did lead to decrease of MEC levels, but did not increase taxadiene production significantly. Uncharacteristically, when the mutant *ispG* EC-271 was expressed, it led to lowering of the levels of both DXP and MEC, increase in the DXP/MEC ratio, and abolition of taxadiene production. We speculated that this could be due to some fundamental difference between EC-271 and the other E. coli *ispG* variants which leads the cells to behave in this radically different way. It could be that the expression of EC-271 could lead to high metabolic burden, leading to the abolition of taxadiene productivity in the same way that high expression levels of the wild type E. coli *ispG* led to small or zero taxadiene titers. It might be useful to look more into the EC-271 to discover what led it to behave so much more differently than the other *ispG* variants.

For the experiments involving the use of an *ispG* gene from *T. Thermophilus* (TT-WT, TT-299, TT-301 & TT-342), it was found that in most cases the expression of an additional *ispG* gene did not lead to great differences to taxadiene production or metabolic intermediate levels. TT-301 slightly outperformed the other *ispG* variants with regards to taxadiene production at certain time points. The most interesting results were gathered from the strain with the TT-342 *ispG*, whose expression led to a decrease in cell growth rate and to overall lower DXP levels and higher MEC levels than the other variants and an overall yield that was slightly higher than that any other case, however at the cost of increased culture time.

For the experiments involving the use of an *ispG* gene from *A. aeolicus* (AA-WT, AA-269 & AA-299), it was found that DXP, MEC ratios were similar in most cases, and taxadiene production was only marginally improved with the expression of an additional *A. aeolicus ispG* gene. Furthermore, the expression of AA-299 led to higher taxadiene yield and productivity compared to the other *A. Aeolicus ispG* variants and thr no vector strain. However, counterintuitively, it exhibited lower DXP/MEC ratios than the strains it was compared to.

7.3.3 Discussion and potential future avenues of inquiry

It was found that lowering MEC levels was shown to not necessarily lead to higher taxadiene titers, contrary to expectations. It needs to be noted that in all our runs we did not see great shifts in productivity or intermediate levels by expressing additional *ispG* variants using plasmids. Possible reasons for this could be the fact that our primary output (taxadiene) is produced downstream of IspG and downstream processes might "drown out" any possible changes we might have effected. It is also possible that each individual strain might have its own optimal induction levels for the plasmid-bound *ispG*, so using the same induction level for all *ispG* variants could lead to non-optimal results.

A point that also needs to be stressed, however, is that while the measurements from our *in vivo* biological assay show whether MEC levels are decreased and taxadiene production is enhanced, assaying whether the ultimate goal, enhancing the productivity of the MEP pathway and producing higher isoprenoid titers has been reached, they cannot tell us whether we have actually created an oxygen tolerant version of IspG *per se*. In order to gain greater understanding of the effects on the enzyme itself, an *in vitro* assay that could measure the levels of HMBPP produced through the reaction of MEC catalyzed through purified IspG variants would be required, in the spirit of similar experiments performed in literature to show IspG activity *in vitro* (10, 34). The main idea would be that if the Fe-S complex has indeed been turned oxygen-tolerant into one of the IspG variants we created, the activity of the enzyme itself would retained after exposure to oxygen, or could be regained after exposure to a suitable reducing environment, since the Fe-S cluster would not be irreparably damaged. While we have been successful in purifying all IspG proteins, as can be shown by SDS-PAGE gels, but despite numerous attempts we have so far been unable to replicate literature experiments in which the anaerobic activity of IspG is assayed. This could probably be due to the imperfect techniques we had been using to maintain anaerobicity (which is probably broken, thus oxidizing the Fe-S clusters, leading to inactivation of IspG) and the sensitivity of our LC-MS/MS, which we used to measure HMBPP levels.

Another reason for this inconclusive data could be that the presence of two IspGs, one produced through the expression of the genomic *ispG* and the other through the plasmids we transform in the cells, might also "drown out" the effect of the additional *ispG* we added. It would theoretically be much simpler if we could have only one *ispG* variant and even more preferably have this *ispG* variant be in the genome. Our efforts to replace the native genomic *ispG* strain with our variants in "one go" through homologous recombination using the λ-Red system were unsuccessful. However, as has been shown, we were eventually able to use a technique involving the CRISPR-Cas9 system and the lower mevalonate pathway to produce strain strain KO1 (Chapter 2), in which the genomic ispG is knocked out. In the future, using could be one way to address this issue, in which case the *ispG* mutants could be expressed in a "clean" background, hence lead to more conclusive results concerning the efficacy of the mutants.

7.4. Examining the effect of antioxidants

7.4.1 Background and approach

As had been mentioned above, one other avenue of potentially preventing the inactivation of IspG by ROS would be to diminish the levels of ROS themselves in the environment that IspG is in. One way to decrease the levels of reactive oxygen species in the cytosol would be through the use of antioxidants to scavenge ROS. An example of such a use for antioxidants can be found in literature, where the inactivation of IRP (iron regulatory protein), a protein with a [4Fe-4S] cluster, was reportedly diminished through the addition of 30mM of antioxidant NAC in the culture media (14). Another way would be through boosting the expression of ROS-scavenging enzymes.

7.4.2 Supply of antioxidants in culture media

Shake-flask experiments were conducted to ascertain what effect the addition of certain antioxidants in culture media would have in isoprenoid production. We used antioxidant chemical

compounds already present in the lab. We ran shake flask cultures of the same taxadiene-overproducing strain we used for our *ispG* assays, in which the culture media were supplemented with a small quantity of antioxidant, with the intent of seeing whether the levels of MEC drop and taxadiene titer would increased, indicating that IspG is being protected from. Out of the elements of our chemical inventory, we singled out three chemicals. One was L-glutathione, which in its reduced form has been stated to be a major endogenous antioxidant, given that glutathione reductase is one of the antioxidant defense systems in the cell (35). A second one was taurine, whose use as an antioxidant has also been proposed (36). The third one was L-ascorbic acid, a well-known antioxidant and mild reducing agent.

The conditions of the shake-flask runs were the same as before, with the level of the antioxidant supplementation was set at 3mM. Initially, we tried to have 30mM of supplementation in order to copy literature precedent (14), but quickly found that at such high levels L-ascorbic acid was toxic (no cell growth), whereas we could not make a stock solution of L-glutathione that was concentrated enough for our needs. Shake flask experiments lasted until glucose was fully consumed, with an average of 36h.

Table 7-2: Effect of antioxidant supplementation in media on taxadiene production and select MEP pathway intermediates

Assay	OD_{600}	Taxadiene titer	Taxadiene/OD	Productivity	Yield	DXP/OD	MEC/OD
	(AU)	$\frac{mg_{tax.}}{L}$	$\frac{mg_{tax.}}{OD_{600}\ L}$	$\frac{mg_{tax}}{OD_{600}\ h}$	$\frac{mg_{tax.}}{g_{glu.}}$	$\frac{\mu M}{OD_{600}}$	$\frac{\mu M}{OD_{600}}$
No antiox.	8.14	15.68±0.81	1.92±0.15	0.037	0.85	51.86	104.64
L-glutathione 3mM	13.48	24.77±7.69	1.84±0.60	0.034	1.31	38.40	29.41
Taurine 3mM	12.5	16.30±4.90	1.31±0.41	0.026	0.86	34.66	24.19
L-asc. Acid 3mM	6.88	18.52±3.68	2.69±0.61	0.049	0.98	24.14	17.00

We observed that the final taxadiene titer per cell density (mg/L/OD_{600}) did improve slightly when L-ascorbic acid was supplemented in the media. Even though cell growth rate was slowed,

higher productivity ($mg_{tax}/OD_{600}/h$) was observed. The yield ($mg_{tax}/g_{glu.}$) increased slightly. In the case of L-glutathione supplementation, higher taxadiene titers were achieved and cell growth and final OD_{600} was also enhanced, leading to an overall lower productivity, even though the yield increased. Taurine supplementation led to similar yield and titer as the non-supplementation case and lower productivity due to higher final cell density. In all cases of supplementation with antioxidants, MEC levels were lowered and the DXP/MEC ratio was decreased compared to the non-supplementation cases.

These results show that there could be potential in utilizing antioxidants to improve the productivity of the MEP pathway. However, it needs to be stated that the levels of antioxidant supplementation in the media would need to be optimized before moving forward. Furthermore, supplementing the culture media with antioxidants could potentially be expensive and therefore not industrially relevant.

7.4.3 Engineering cells to maintain low ROS levels

An approach more industrially relevant than supplying antioxidants to the media could be to engineer a cell to keep intracellular ROS levels low. One such approach would be through the expression of genes that lead to the production of antioxidant molecules, which would act as ROS scavengers, in the cell. An alternative approach would be to overexpress superoxide dismutase (SOD), which acts as an ROS scavenger in cells (along with catalase) and has been reported to prevent decline in the activity of α,β-Dihydroxyisovalerate Dehydratase, an enzyme that is also sensitive to O_2^- (21), or attempt the overexpression of catalase which is also linked with protection from ROS.

In our work, we investigated the effects of the expression of four different genes. The first gene that was tested was Pyrroloquinoline quinone synthase (PQQs), which leads to the production of PQQ, a chemical that has been reported to increase *E. coli* resistance to superoxide radicals and singlet oxygen species (37), stimulate catalase & SOD (38) and act as a growth

factor (39). PQQs was cloned on the pSC101 backbone (approximately 5 copies per cell) under the trc promoter, as was the AhpC, the gene for the production of (alkyl hyperoxide reductase), an enzyme that has been linked to protection from oxidative stress and reduction of ROS (40). The up-regulation of SOD, an enzyme that catalyzes the breakup of superoxide radicals was also investigated. Two variants of SOD were tested, an SOD_B, which is one of the *E. coli* Fe-SOD genes (41) and was expressed in a pET vector (copy number 20) under the araC (arabinose inducible) promoter, as well as another Fe-SOD variant from *Synechocystis* sp. PCC 6803, which has also been reported as conferring oxidative stress tolerance in cells (42). This gene was cloned on the pSC101 backbone, under the trc promoter. The aforementioned plasmids were then transformed into the taxadiene-overproducing strain and in the case of these strains, we ran bioreactors in order for us to have as controlled conditions as possible. Bioreactor runs lasted until glucose was depleted with the bioreactor run with the base strain (no vector) lasting about 32h, whereas runs using strains in which an antioxidant enzyme was also expressed ran longer, for an average of 45h.

Unfortunately, none of the strains in which antioxidant genes were overexpressed managed to outperform the base strain, as they all had lower final taxadiene titers and productivities. The only interesting result was the case of the Fe-SOD from *Synechocystis*, which had higher taxadiene titer per optical density than the base strain, since it managed to produce around 2/3 of the base strain's titer with only 1/3 of the final OD_{600}, albeit at the cost of a longer run (~50h vs. 32h). In all strains that expressed an antioxidant gene, the DXP/MEC ratio was higher than the base strain, but this did not lead to a corresponding increase in taxadiene titer. In the case of the strain in which *Ahpc* was introduced, it was found that taxadiene production was abolished. Like in similar situations in which taxadiene production was abolished, we speculate that the high expression levels of the gene may be partially responsible.

Table 7-3: Effect of overexpression of antioxidant genes on taxadiene production and select MEP pathway intermediates

Antioxidant gene overexpressed	OD_{600}	Taxadiene titer	Taxadiene/dcw	Productivity	Yield	MEC/dcw	DXP/MEC
	(AU)	$\frac{mg_{tax.}}{L}$	$\frac{mg_{tax.}}{g_{dcw}L}$	$\frac{mg_{tax}}{g_{dcw}h}$	$\frac{mg_{tax.}}{g_{glu.}}$	$\frac{\mu mol}{g_{dcw}}$	$\mu M/\mu M$
None (Wild Type)	21.00	36.50	5.79	0.21	2.14	4.91	0.68
PQQs (*D. radiodurans*)	8.00	12.20	5.08	0.11	0.75	19.26	0.80
SOD (*Synechocystis*)	7.12	22.79	10.67	0.15	0.86	18.51	0.71
SOD_B (*E. Coli*)	7.10	11.25	5.28	0.17	0.71	28.67	0.80
AhpC (*E. Coli*)	8.23	0	0	0	0	*Did not measure*	*Did not measure*

These results show the need for further optimization with regards to the expression of genes that have antioxidant activity and regulate the O_2 levels in the cell. However, the levels of O_2 and ROS in the cell are probably controlled by more than any single gene that we might want to express, thereby making our goal of engineering *E. coli* to tune the intracellular O_2 levels very difficult.

7.5. Engineering *Azotobacter vinelandii* into an isoprenoid-producing strain

Work on A. vinelandii was performed in collaboration with Steven Edgar.

7.5.1 Background and approach

Seeing that it may be too difficult and time consuming to engineer E. coli to keep intracellular O_2 levels very low, we took a step back and considered whether we could possibly find examples of finely tuned intracellular levels of O_2 nature itself. Our research showed that a bacterial species, *Azotobacter vinelandii* might be the answer to our problem. *Azotobacter vinelandii* is a bacterium that is an obligate aerobe (43). One of *A. vinelandii*'s main features is the fact that it can fix nitrogen under anaerobic conditions. It needs to be noted that nitrogenases (nitrogen-fixation

enzymes) are oxygen-sensitive and yet *A. vinelandii* is able to use nitrogen metabolism with great efficiency. Furthermore, in addition to nitrogenases, *A. vinelandii* also has a host of genes that encode other oxygen-sensitive enzymes (43).

The fact that *A. vinelandii* can efficiently use anaerobic pathways has been studied by several groups. It is now known that *A. vinelandii* displays a trait known as respiratory protection, under which it can modulate the levels of intracellular O_2, thereby protecting its oxygen sensitive enzymes from damage. This is known to happen in a multitude of ways. First, *A. vinelandii* has one of the highest respiratory rates of any known bacterium, allowing it to quickly "burn through" oxygen and keep its levels low in the cell. Additionally, through its respiratory regulation machinery (protein CydR and other respiratory proteins), *A. vinelandii* can adjust oxygen consumption rates and shift between using oxygen for energy or merely consuming it for the sole purpose of keeping O_2 levels low (43). Furthermore, *A. vinelandii* can produce alginate, regulate its composition and surround itself with alginate "walls" that can protect the cell from exogenous oxygen (43). Other ways in which respiratory protection in *A. vinelandii* work involve its ROS-protecting enzymes, such as SOD, and its regulation of the supply of ATP & reducing equivalents (44).

Given the fact that *A. vinelandii* has been shown to be able to utilize pathways involving oxygen-sensitive enzymes with great efficiency and employ a number of respiratory protection methods to ensure that said enzymes do not get destroyed by ROS, we believe that *A. vinelandii* could be used as a platform for the overexpression of the MEP pathway, for the production of isoprenoids. One advantages of *A. vinelandii* would be the aforementioned respiratory protection, which could protect IspG/IspH. Utilizing a strain that has been optimized by nature to protect oxygen-sensitive enzymes would probably be easier than attempting to engineer our *E. coli* strain to protect its ROS sensitive enzymes as well. The fact that it can naturally produce C_{20} isoprenoids, unlike E. coli, which needs to have an additional gene be heterologously expressed for that purpose also makes *A. vinelandii* an interesting host.

7.5.2 Establishing tools to engineer *A. vinelandii*

Initially, our studies of *A. vinelandii* were focused on introducing and expressing genes in the cell, so as to make sure we can introduce requisite genes for MEP pathway overexpression in the future. Our initial approach involved using a plasmid to introduce any genes of interest. Concerning selection techniques, we looked into the use of many antibiotics before settling at Kanamycin and determining the minimum inhibitory concentration (on an agar plate) at 0.125μg/L. Concerning the vector, we elected to use the pRK origin of replication as it has been transformed in *A. vinelandii* previously in literature (45), is well characterized and its copy number can be altered by mutagenesis. We chose the pM promoter, which uses m-toluic acid for its induction. It is titratable, its strength can be altered via mutagenesis and has been demonstrated to be functional in a wide range of organisms. Based on those characteristics for a vector, we chose the plasmid pSEVA228. Unfortunately, however, we did not succeed in transforming *A. vinelandii* cells via electroporation, putting an end to our plans for introducing genes via plasmid electroporation. We then looked into employing a different strategy for introducing a plasmid in *A. vinelandii* and we elected to use conjugation, using *E. coli* MFDpir as the donor, which lacked any antibiotic resistance and was auxotrophic in DAP, making it easier for us to remove it from our cultures/plates after conjugation. Ultimately, however, we were unable to introduce plasmids in strains, as we could not confirm that the plasmid remained in the strain after multiple rounds of culturing.

To overcome the deficiencies in the introduction of genes through plasmids we looked into ways of introducing our genes of preference into the genome, by exploiting the native *recA* that *A. vinelandii* expresses to recombine gene cassettes in genomic loci of our choice, which we would target by introducing large homology sequences flanking the integration cassettes. We identified 3 possible integration sites, based on literature precedent and possible impact to the cell. The first possible spot is the locus of for the gene that encodes AlgD (46) (GDP-mannose-6-dehydratase) which plays a role in the formation of alginate layers around *A. vinelandii*. There is

also literature precedent for integration near the loci for AcxA (47), which is a protein that regulates nitrogenase expression, and ScrX (48), which regulates sucrose catabolism and gets activated in the presence of sucrose. Integration of genes there could also exploit its native promoter and express gene by using sucrose as an inducer.

We developed a two-step strategy for genomic integration of our desired genes. In the first step, we would replace a native gene (*algD*, *scrX*, *acxA*) in the genome with a cassette that contained *sacB* and *neoR*. The latter gene conferred resistance to kanamycin, which we would use to select for *A. vinelandii* colonies that have successfully had this cassette integrated in their genomes. The former gene conferred sensitivity to sucrose and is vital to the second step of the process. In the second step, the cassette that was inserted in the first step is itself replaced by a cassette bearing one or more genes of our choice. Since the genes that we want to introduce (e.g. MEP pathway genes) need not necessarily confer some phenotypic change we can easily select for, for the second step we employ the sucrose sensitivity conferred by sacB to make sure that we have managed to excise the original insert and replaced with our insert of choice by ensuring that *A. vinelandii* colonies tin which this has not happened will not grow in the presence of sucrose. Both steps are to be carried out by first cloning the cassette (flanked by targeting homology regions) in a pSEVA228 plasmid, which we transform into an *E. coli* MFDpir. Then, we would perform conjugation between the donor (*E. coli* MFDpir) and the recipient (*A. vinelandii*) and plate the A. vinelandii in media without DAP. Only *A. vinelandii* should survive on the agar plates, and we can furthermore select the correct (i.e. the ones that have introduced the cassette in their genome) colonies by selecting for/against kanamycin resistance/sucrose sensitivity.

So far we have achieved integration of the *sacB-neoR* cassette on all three loci (*algD*, *scrX* and *axcA*) and have produced strains that exhibit phenotypic traits consistent with such an integration, that is, sensitivity to sucrose and resistance. Furthermore, the strains are stable and continue to exhibit the phenotype after many rounds of subculturing, also an indication of a successful genomic integration.

However, rather than rely solely on the phenotype, which indicated that our strains have successfully taken up the cassette, we also tried to verify the genotype. In the case of the *A. vinelandii* strains in which the cassette was inserted in the *scrX* locus, we got mixed results. We discovered that after sequencing the product of the PCR amplification of the insertion locus and the area surrounding it, the sequence would give us mixed peaks which were partly those of *sacB*, which would be in that genomic location if the cassette integrated successfully, and partly those of *scrX*, the wild-type gene originally in that location, which would only be in that genomic location if the cassette failed to integrate. This was a seemingly self-contradictory result which needed to be investigated before moving forward. Further literature research revealed that *A. vinelandii* has multiple chromosomes (49), which can explain our result, in that it is possible we are only getting integration on only a fraction of the many chromosomes that *A. vinelandii* has, thus explaining the mixed peaks. Indeed, when PCR amplifying with primers that are internal and unique to either *sacB* or *scrX* we are able to get both bands from the same colony, further supporting the hypothesis.

A number of possible avenues could be used to address this problem. The first avenue involves following literature precedent by sequentially subculturing *A. vinelandii*, in order for genome segregation to occur and for a "clean" strain with only one type of genome to emerge. Another way would be to consider integrating the cassette in the locus of a counterselectable marker, so as to ensure that only "clean" strains can survive after integration. A preliminary look into the literature (50) search and comparison with the *A. vinelandii* genome (43) has led us to identify three possible genes that could be used as counterselectable markers: *lacY*, *pheS* and *thyA*.

7.6. Conclusions

In this chapter we detailed some of the avenues we explored in order to overcome the deficiencies of one of the MEP pathway enzymes, IspG. The avenues we investigated have not so far yielded impressive results, for although in some cases, such as when antioxidants are considered, we are able to achieve modest improvements, we were overall not able to make significant breakthroughs. It also needs to be noted that these efforts have all been focused on fixing a single enzyme, out of 7 in the MEP pathway. Although it is conceivable that some approach may be able to debottleneck the pathway at IspG, the pathway will still suffer from its fundamental limitations, such as its connection to Central Carbon Metabolism, its long length and its need for multiple cofactors.

Ultimately, by way of juxtaposition, this chapter, along with a large body of literature on the engineering of the MVA and MEP pathways highlights many of the limitations the engineering of either pathway entails, limitations that the Isopentenol Utilization Pathway manages to successfully circumvent.

7.7. Materials and Methods

Strains, plasmids and genes; cloning protocol

Materials and methods were the same as the ones as described in the relevant sections of Chapter 2. pET-araC-SOD_B was a gift from Kang Zhou. Refer to **Table A1** in the Appendix for a listing of strains and to **Table A2** in the Appendix for a listing of genes.

Construction of expression vectors.

All the vectors for expressing additional copies of *ispG* were constructed by PCR-amplifying the backbone of p5T7vs-ispA using primer pair GP-pSEVAispGBB F/R and inserting the appropriate *ispG*. The *E. coli ispG* was amplified from the E. coli genome using primer pair GB-pSEVA-ECispG F/R. The *A. aeoliculs ispG* was amplified from a custom-synthesized, codon-optimized

fragment using the primer pair GB-pSEVA-AAispG F/R and the *T. thermophilus ispG* was amplified from a custom-synthesized, codon-optimized fragment using the primer pair GB-pSEVA-TTispG F/R. The backbone and appropriate *ispG* were ligated to produce the appropriate plasmid. To create plasmids with the appropriate *ispG* mutants the above plasmids were PCR amplified and then cyclized using the following primer pairs: For the EC-271, EC-273 & EC-303 mutants the primer pairs used were respectively GB-EC271-F/R, GB-EC273-F/R and GB-EC303-F/R. For the TT-299, TT-301 & TT-342 mutants the primer pairs used were respectively GB-TT-299-F/R, GB-TT-301-F/R and GB-TT-342-F/R. For the AA-269 & AA-299 mutants the primer pairs used were respectively GB-AA269-F/R and GB-AA299-F/R. To create the plasmids for the expression/production of ROS scavengers, we PCR amplified the backbone of p5TrcMEP using the primer pair GB-P5TRC-BB-F/R. *AhpC* was amplified from the E. coli genome using the primer pair GB-P5TRC-Ahpc-F/R. PQQ and SOD were amplified from custom-ordered codon optimized gene fragments using primer pairs GB-P5TRC-PQQ-F/R and GB-P5TRC-SOD-F/R respectively. Vector sequences for the primers used are listed in **Table A3** in the Appendix. Codon-optimized sequences are listed in **Table A5** in the Appendix.

Detection and quantification of DXP, MEC and taxadiene.

DXP and MEC were scanned for via LC-MS/MS using the means and method described in Chapter 3. Taxadiene was quantified using GC-MS using the method described in Chapter 4.

Cultivation in bioreactors.

Strains were cultivated in a 3-L Bioflo 110 bioreactor (New Brunswick) with aeration, agitation, and pH control. 1.25L of defined media (the media composition was that developed for bioreactor runs for the taxadiene overproducing strain (33), without addition of yeast extract, plus 100µL antifoam 204 and 50 µg/mL spectinomycin) was inoculated at 1% v/v with an overnight culture (12 h) grown in LB media. Aeration (0.3-1 vvm) and agitation (250-1250 rpm) was controlled by a cascade to maintain dissolved oxygen at 20% of saturation. pH was controlled by addition of 25%

v/v NH_4OH. Temperature was controlled at 30°C. When an OD of 0.5 was reached, IPTG and of m-toluic acid were added to final concentrations of 0.1mM and 0.4μM respectively. Cell density and glucose consumption were monitored as described in Chapter 4.

Cultivation in shake flasks.

Cultivation in shake flasks was conducted in a 50mL volume using under the same media used for the bioreactor runs (described above). 0.3mL C18 flash resin was added to cultures at induction time to capture taxadiene. Metabolites, and taxadiene were quantified using protocols described above.

Conjugation in *A. vinelandii*.

A. vinelandii strain to be transformed (recipient) is precultured in 10 mL of Burk's media for 3 days. On day 3, the *E. coli* MFDPir (donor), which has already been transformed with the payload is precultured in 10 mL LB media, supplemented with Diaminopimelic acid (DAP). The next day the donor is inoculated in 3 mL LB media supplemented with DAP and the appropriate antibiotic (as determined by the payload), and the recipient is subcultured in 12 mL Burk's media. Once the recipient achieves OD_{600} of about 0.1-0.3, the donor is again subcultured into 10 mL LB media, supplemented with DAP and antibiotic. When both donor and recipient simultaneously reach an OD_{600} of 0.4, then 1 mL of recipient culture is spun down (4000 rcf, 5 min) to remove the supernatant and resuspended in 100 mL LB supplemented with DAP and 1mL of donor culture is spun down (3000 rpm, 4 min) to remove the supernatant and very gently resuspended in 100 mL LB supplemented with DAP. The above two are suspensions are mixed and are pipetted on top of a filter paper (0.22 μm) resting on an LB-agar plate supplemented with DAP. The plate with the filter paper is incubated for 3 h at 30°C and then the filter paper is placed in a 50 mL tube with 10 mL Burk's media and vortexed vigorously. The filter is removed from the tube and the tube is centrifuged (3000 rpm, 12 min). Then the 9.5 mL of supernatant is removed and the pellet is resuspended in the remainder of the liquid. 100 mL of suspension is then plated on a Burk's

media—agar plate supplemented with the appropriate antibiotic (as determined by the payload). The resulting colonies have been transformed with the payload. In all the above cultures, the recipient is grown at 30°C and the donor at 37°C.

7.8. Bibliography

1. Yang J, et al. (2012) Bio-isoprene production using exogenous MVA pathway and isoprene synthase in Escherichia coli. *Bioresour Technol* 104:642–647.
2. Wang J, Xiong Z, Li S, Wang Y (2015) Exploiting exogenous MEP pathway genes to improve the downstream isoprenoid pathway effects and enhance isoprenoid production in Escherichia coli. *Process Biochem* 50(1):24–32.
3. Partow S, Siewers V, Daviet L, Schalk M, Nielsen J (2012) Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in Saccharomyces cerevisiae. *PLoS One* 7(12):1–12.
4. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* 330(2010):70–74.
5. Meng H, Wang Y, Hua Q, Zhang S, Wang X (2011) In silico analysis and experimental improvement of taxadiene heterologous biosynthesis in Escherichia coli. *Biotechnol Bioprocess Eng* 16:205–215.
6. Chang WC, Song H, Liu HW, Liu P (2013) Current development in isoprenoid precursor biosynthesis and regulation. *Curr Opin Chem Biol* 17(4):571–579.
7. Seemann M, et al. (2005) Isoprenoid biosynthesis in chloroplasts via the methylerythritol phosphate pathway: The (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) from Arabidopsis thaliana is a [4Fe-4S] protein. *J Biol Inorg Chem* 10:131–137.
8. Seemann M, et al. (2009) Isoprenoid biosynthesis via the MEP pathway : In-vivo Mössbauer Spectroscopy identifies a [4Fe-4S] 2 + center with unusual coordination

sphere in the LytB protein. *J Am Chem Soc* 131(37):13184–13185.

9. Gräwert T, et al. (2004) IspH protein of Escherichia coli: studies on iron-sulfur cluster implementation and catalysis. *J Am Chem Soc* 126(2):12847–12855.
10. Lee M, et al. (2010) Biosynthesis of Isoprenoids: Crystal Structure of the [4Fe-4S] Cluster Protein IspG. *J Mol Biol* 404(4):600–610.
11. Jang S, Imlay J a. (2007) Micromolar intracellular hydrogen peroxide disrupts metabolism by damaging iron-sulfur enzymes. *J Biol Chem* 282(2):929–937.
12. Py B, Barras F (2010) Building Fe-S proteins: bacterial strategies. *Nat Rev Microbiol* 8(juNE):436–446.
13. Liochev S, Fridovich I (1994) The role of O2.- in the production of HO.: In vitro and in vivo. *Free Radic Biol Med* 16(1):29–33.
14. Pantopoulos K, Hentze MW (1995) Rapid responses to oxidative stress mediated by iron regulatory protein. *EMBO J* 14(12):2917–2924.
15. Imlay J a (2003) Pathways of oxidative damage. *Annu Rev Microbiol* 57:395–418.
16. Zhou K, Zou R, Stephanopoulos G, Too H-PP (2012) Metabolite Profiling Identified Methylerythritol Cyclodiphosphate Efflux as a Limiting Step in Microbial Isoprenoid Production. *PLoS One* 7(11):e47513.
17. Tokumoto U, et al. (2002) Network of protein-protein interactions among iron-sulfur cluster assembly proteins in Escherichia coli. *J Biochem* 131:713–719.
18. Justino MC, Almeida CC, Teixeira M, Saraiva LM (2007) Escherichia coli Di-iron YtfE protein is necessary for the repair of stress-damaged iron-sulfur clusters. *J Biol Chem* 282(14):10352–10359.
19. Expert D, Boughammoura A, Franza T (2008) Siderophore-controlled iron assimilation in the enterobacterium Erwinia chrysanthemi: Evidence for the involvement of bacterioferritin and the suf iron-sulfur cluster assembly machinery. *J Biol Chem* 283(52):36564–36572.

20. Velayudhan J, Castor M, Richardson A, Main-Hester KL, Fang FC (2007) The role of ferritins in the physiology of Salmonella enterica sv. Typhimurium: A unique role for ferritin B in iron-sulphur cluster repair and virulence. *Mol Microbiol* 63(5):1495–1507.

21. Kuo CF, Mashino T, Fridovich I (1987) α,β-Dihydroxyisovalerate Dehydratase, A superoxide sensitive enzyme. *J Biol Chem* 262:4724–4727.

22. Ohnishi T (1998) Iron – sulfur clustersrsemiquinones in Complex I. *I Can*:186–206.

23. Yankovskaya V, et al. (2003) Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science* 299(January):700–704.

24. Volbeda A, et al. (2013) Crystal structure of the O2-Tolerant membrane-bound hydrogenase 1 from escherichia coli in complex with its cognate cytochrome b. *Structure* 21(1):184–190.

25. Sigfridsson KG V, et al. (2014) Structural differences of oxidized iron-sulfur and nickel-iron cofactors in O2-tolerant and O2-sensitive hydrogenases studied by X-ray absorption spectroscopy. *Biochim Biophys Acta - Bioenerg* 1847(2):162–170.

26. Goris T, et al. (2011) A unique iron-sulfur cluster is crucial for oxygen tolerance of a [NiFe]-hydrogenase. *Nat Chem Biol* 7(march):310–318.

27. Fritsch J, et al. (2011) The crystal structure of an oxygen-tolerant hydrogenase uncovers a novel iron-sulphur centre. *Nature* 479:249–252.

28. Sigfridsson KGV, et al. (2015) Structural differences of oxidized iron–sulfur and nickel–iron cofactors in O2-tolerant and O2-sensitive hydrogenases studied by X-ray absorption spectroscopy. *Biochim Biophys Acta - Bioenerg* 1847(2):162–170.

29. Lenz O, et al. (2010) H2 conversion in the presence of O2 as performed by the membrane-bound [NiFe]-Hydrogenase of Ralstonia eutropha. *ChemPhysChem* 11:1107–1119.

30. Wulff P, Day CC, Sargent F, Armstrong FA (2014) How oxygen reacts with oxygen-tolerant respiratory [NiFe]-hydrogenases. *Proc Natl Acad Sci U S A* 111(18):6606–11.

31. Quitterer F, et al. (2015) Atomic-Resolution Structures of Discrete Stages on the Reaction Coordinate of the [Fe4S4] Enzyme IspG (GcpE). *J Mol Biol* 427(12):2220–2228.

32. Zhao L, Dietzel, Kevin L, Wichman, Gale A (2012) Intl. Patent WO 2012/135591 A2.

33. Silva-Rocha R, et al. (2013) The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* 41(D1):666–675.

34. Xiao Y, Zahariou G, Sanakis Y, Liu P (2009) IspG Enzyme activity in the deoxyxylulose phosphate pathway: Roles of the iron-sulfur cluster. *Biochemistry* 48:10483–10485.

35. Blehert DS, Fox BG, Chambliss GH (1999) Cloning and Sequence Analysis of Two Pseudomonas Flavoprotein Xenobiotic Reductases Cloning and Sequence Analysis of Two Pseudomonas Flavoprotein Xenobiotic Reductases. 181(20).

36. Sirdah MM (2015) Protective and therapeutic effectiveness of taurine in diabetes mellitus: A rationale for antioxidant supplementation. *Diabetes Metab Syndr Clin Res Rev* 9(1):55–64.

37. Khairnar NP, Misra HS, Apte SK (2003) Pyrroloquinoline-quinone synthesized in Escherichia coli by pyrroloquinoline-quinone synthase of Deinococcus radiodurans plays a role beyond mineral phosphate solubilization. *Biochem Biophys Res Commun* 312:303–308.

38. Misra HS, et al. (2004) Pyrroloquinoline-quinone: A reactive oxygen species scavenger in bacteria. *FEBS Lett* 578:26–30.

39. Yang XP, Zhong GF, Lin JP, Mao DB, Wei DZ (2010) Pyrroloquinoline quinone biosynthesis in Escherichia coli through expression of the Gluconobacter oxydans pqqABCDE gene cluster. *J Ind Microbiol Biotechnol* 37:575–580.

40. De Spiegeleer P, et al. (2005) Investigation into the resistance of lactoperoxidase tolerant Escherichia coli mutants to different forms of oxidative stress. *FEMS Microbiol Lett* 252:315–319.

41. Chiang SM, Schellhorn HE (2012) Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Arch Biochem Biophys* 525(2):161–169.

42. Bhattacharya J, GhoshDastidar K, Chatterjee A, Majee M, Majumder AL (2004) Synechocystis Fe superoxide dismutase gene confers oxidative stress tolerance to Escherichia coli. *Biochem Biophys Res Commun* 316:540–544.

43. Setubal JC, et al. (2009) Genome sequence of Azotobacter vinelandii, an obligate aerobe specialized to support diverse anaerobic metabolic processes. *J Bacteriol* 191(14):4534–4545.

44. Oelze J (2000) Respiratory protection of nitrogenase in Azotobacter species: Is a widely held hypothesis unequivocally supported by experimental evidence? *FEMS Microbiol Rev* 24(4):321–333.

45. Phadnis SH, Das HK (1987) Use of the plasmid pRK 2013 as a vehicle for transposition in Azotobacter vinelandii. *J Biosci* 12(2):131–135.

46. Campos A, Marti JM (1996) Characterization of the gene coding for GDP-mannose dehydrogenase (algD) from Azotobacter vinelandii . Characterization of the Gene Coding for GDP-Mannose Dehydrogenase (algD) from Azotobacter vinelandii. 178(7):1793–1799.

47. Dos Santos PC, Johnson DC, Ragle BE, Unciuleac MC, Dean DR (2007) Controlled expression of nif and isc iron-sulfur protein maturation components reveals target specificity and limited functional replacement between the two systems. *J Bacteriol* 189(7):2854–2862.

48. Johnson DC, Unciuleac MC, Dean DR (2006) Controlled expression and functional analysis of iron-sulfur cluster biosynthetic components within Azotobacter vinelandii. *J Bacteriol* 188(21):7551–7561.

49. Punita, Jafri S, Reddy MA, Das HK (1989) Multiple chromosomes of Azotobacter

vinelandii. *J Bacteriol* 171(6):3133–3138.

50. Reyrat J, Pelicic V, Gicquel B (1998) Counterselectabe Markers: Untapped Tools for Bacterial Genetics and Pathogenesis. 66(9):4011–4017.

Chapter 8

Conclusions

8.1. Thesis Summary

In this thesis, the engineering of a new pathway for the production of isoprenoids, the Isopentenol Utilization Pathway (IUP), was presented. After identifying suitable enzymes and constructing the pathway, we attempted to probe the limits of the IUP for producing various isoprenoid downstream products. The IUP flux exceeded the capacity of almost all downstream pathways tested and was competitive with the highest isoprenoid fluxes reported. We then proceeded to work on improving the downstream pathways in order to utilize the full potential of the IUP. Furthermore, we proposed a novel way for the production of functionalized isoprenoids through the use of the IUP.

In Chapter 2 we proposed the design of the Isopentenol Utilization Pathway (IUP), a two-step pathway that can produce isopentenyl diphosphate or dimethylallyl diphosphate, the main precursors to isoprenoid synthesis, through sequential phosphorylation of isopentenol isomers isoprenol or prenol. We identified the *S. cerevisiae* choline kinase (ScCK) as the appropriate enzyme to catalyze the first phosphorylation step through its promiscuous activity and selected the *A. thaliana* isopentenyl phosphate kinase (AtIPK) to catalyze the second step. We then showed that the IUP can be used to produce isoprenoids and therefore rescue viability in an *E. coli* knockout strain that had its ability to produce isoprenoids knocked out. In Chapter 3, we determined that the ideal feedstock for the IUP is 25 mM isoprenol, we observed that, although not essential to the pathway, AtIPK enables the pathway to function well, that the IUP can produce IPP/DMAPP very quickly and in quantities orders of magnitude higher than the native MEP pathway. Furthermore, we show that the IUP is decoupled from central carbon metabolism,

bringing the pathway's theoretical yield to 100%. In Chapter 4 we discuss methods for extracting and product and measuring its concentration, while ensuring that isoprenol is not removed from the media and discuss our difficulties on measuring yield. We couple the IUP with a downstream module for lycopene production, which we proceed to improve, by first investigating the proper culture conditions and then by altering the origin of replication of the plasmid through which it is expressed, as well as the operon's promoter. We also couple the IUP with a downstream module for the production of volatile isoprenoids, including valencene, miltiradiene, valencene, amorphadiene and taxadiene, proving that the IUP can be used to produce a variety of products. Furthermore, we prove that increases in product titer and intermediate concentrations when the IUP is expressed can be wholly attributed to the IUP. After not succeeding in improving production of taxadiene, even though we use similar techniques to as in the case of lycopene production improvement, we investigate both downstream pathways to identify bottlenecks. We identify that, in the taxadiene production module the GGPP synthase used is very potent, while taxadiene synthase is underperforming. By introducing said taxadiene synthase in the lycopene production strain we achieve further improvements in lycopene productivity. We proceed to show that the IUP can achieve fluxes comparable with some of the best reported so far in the literature and that the IUP can compete with some of the state-of-the-art strains in terms of raw output of isoprenoid intermediates. In Chapter 5 we go one step further by investigating ways for improving the maximum lycopene titer, by increasing *E. coli* membrane capacity or ways of improving taxadiene productivity by mutating the taxadiene synthase gene. We propose an iterative method of establishing the maximum limit of the IUP and pushing past by coupling the IUP with a strong downstream module for farnesene synthesis. We investigate the effect increasing the downstream module strength has on farnesene titers and intermediate accumulation. When we simultaneously upregulate both the upstream and the downstream, we observe the largest increase in farnesene titer and provide further proof of the IUP's remarkable ability to produce large quantities of isoprenoid intermediates. In Chapter 6, we propose a way of using the IUP to

introduce functionalizations to isoprenoids, by using functionalized isopentenols as feedstock, instead of isoprenol. Establish our proof of concept, we attempt to synthesize brominated or amorphadiene and chlorinated through an *in vitro* enzymatic reconstitution of the IUP. While we are unable to conclusively prove the formation of a functionalized isoprenoid product, we do detect LC-MS/MS peaks that could potentially attributed to functionalized metabolic intermediates, encouraging hints that our approach could work. In Chapter 7 we change scope and describe work we conducted prior to the conception of the IUP which focused improving the function of the enzyme IspG to enhance the productivity of the MEP pathway. We first focus on engineering IspG for greater oxygen tolerance by introducing additional cysteines in order to mimic oxygen-tolerant hydrogenases and we then investigate the possibility of scavenging reactive oxygen species to improve performance by either supplementing the media with antioxidants, or producing ROS scavengers *in vivo*. Both efforts give marginal improvements. Finally, we investigate the possibility of using *A. vinelandii* as the host strain for overexpressing the MEP pathway.

8.2 Future directions

The creation of the IUP opens a new field in isoprenoid synthesis, by offering a novel way for the production of isoprenoids. We propose that this thesis be followed by a more rigorous investigation of the IUP and a concerted effort to further evaluate its capabilities and establish it as a commercially viable biosynthetic pathway. Further work could focus on the establishment of a high-titer & productivity production pipeline for particularly impactful isoprenoids through the IUP, such as amorphadiene for the cure of malaria, for which the Gates Foundation is seeking alternative routes for better and more cost-effective production, since current production still remains too costly for use in developing countries. In order to achieve efficient production, additional work is required in identifying the upper limit of IUP's current production capacity, and balancing any downstream pathway with the upstream (IUP), work which could follow the iterative

and incremental method we laid out in Chapter 5, or could focus on the simultaneous upregulation of both upstream and downstream. As has already been elaborated in Chapter 5, upregulation of both downstream and upstream could be achieved by a variety of genetic tools. Once the IUP's current limits have been reached, special attention should be given to the possibility of screening for more potent choline kinases, or employing protein engineering methods to produce such. In a related field, enzymatic assays should be considered to measure any potential inhibitory effects on choline kinase or IPK from isoprenoid intermediates, such as IP, IPP, GPP, FPP and GGPP and/or definitively establish that the UP is free from any regulation.

Further effort is required in the field of accurately measuring the yield of the pathway, which is a particularly important cost consideration. As we have elaborated in Chapter 4, development of methods focused at precisely measuring the isoprenol concentration in cultures could potentially be effected via GC-MS through headspace sampling from sealed vials or from a bioreactor gas exhaust.

As we have briefly mentioned in Chapter 6, the IUP has been successfully been enzymatically reconstituted *in vitro* and successfully used for the production of taxadiene and other volatile isoprenoids. Avenues for further exploration include investigating whether the in vitro IUP could be used for the production of lipophilic, non-volatile isoprenoids, such as lycopene. One possible suggested experiments for such investigation could be to run the enzymatic reaction in a water-oil emulsion, with the enzymes operating in the aqueous phase, while the oil phase would sequester the lipophilic products. Further research could also be directed towards enzyme immobilization. Concerning the production of functionalized products, the most important element inhibiting future research is the ability to definitively detect and identify the hypothetical products and intermediates. It is therefore suggested that efforts are taken to either chemically synthesize or purchase some of the functionalized molecules we wish to biosynthetically produce and analyze their spectra with GC-MS and/or LC-MS/MS, to establish exactly what we future experiments should be looking for. It is also suggested that the possibility of using different

functionalized analogs of isoprenol and prenol as pathway feedstocks, such as fluorinated analogs, be considered, with the main considerations being their ability to dock in enzyme cavities and procurement/synthesis cost. A direction that can be considered if factors such as steric or polar effects lead to low reaction rates for functionalized intermediates would be protein engineering of the IUP enzymes and downstream pathways enzymes to achieve improvements.

8.3 Concluding remarks

The IUP is the only alternative ever proposed to either the MVA or MEP pathways for isoprenoid synthesis that is radically different from them, does not merely provide a different entry point or bypass parts of either pathway, but bypasses MEP and MVA in their entirety. As an alternative isoprenoid pathway, the IUP is completely orthogonal to the natural MEP and non-archaeal MVA pathways and able to circumvent their previously mentioned limitations, including their tight regulation and oxygen dependence (MVA pathway) or sensitivity (MEP).

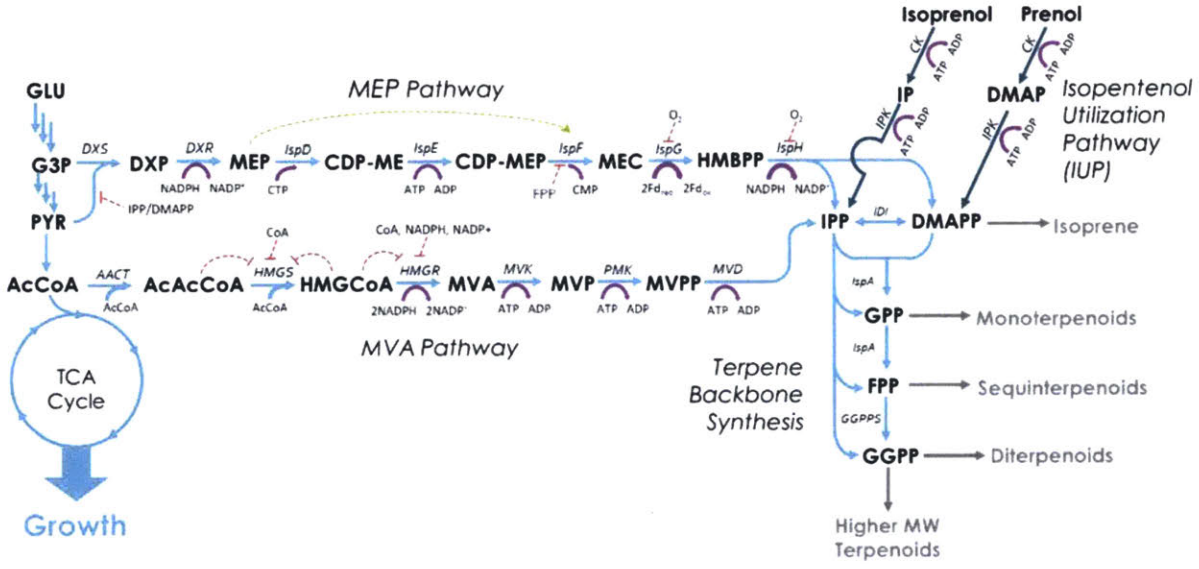


Figure 8-1: Overview of isoprenoid synthesis pathways via the MEP and MVA pathways and the novel Isopentenol Utilization Pathway (IUP). The MEP and MVA pathways start with metabolites produced through the central carbon metabolism, whereas IUP functions through extraneously supplied isoprenol or prenol. Substrates and intermediates are presented in bold and enzymes are presented in italics. Positive regulation is represented by a dashed green arrow, negative regulation by red dashed T lines. Cofactor consumption is represented by purple arrows.

The IUP also has many advantages over previously reported bypass/alternative isoprenoid pathways. Whereas the natural isoprenoid pathways and all their engineered derivatives function by diverting precursor metabolites from central carbon metabolism, inextricably linking isoprenoid production and glucose metabolism, with most of the substrate used for growth and only a small fraction for isoprenoid production, the IUP is completely independent from central carbon metabolism. This means that the IUP does not compete for resources/precursors with other cellular processes, since its substrate (isoprenol or prenol) are externally supplied. Moreover, this also means that no carbon is diverted from the IUP towards central carbon metabolism, where it could be "wasted" to supply cellular processes not leading towards isoprenoid production. This brings the IUP's theoretical yield close to 100%, since isoprenol or prenol can only be used towards isoprenoid production, or otherwise recycled, a significant improvement over the theoretical yields of the natural pathways (1), something infeasible for any pathway starting from a sugar. This advantage also allows us to envision a process wherein isoprenol/prenol is added to a bioreactor with either cells expressing the IUP or an *in vitro* metabolically enzymatically reconstituted IUP. Any residual isoprenol/prenol that the cells/enzymes did not convert to product could then be recovered from the broth, e.g. through distillation, and recycled back into the reactor. We could envision such process functioning in either batch or continuous mode and having a 100% theoretical yield on isoprenol.

This decoupling of central carbon metabolism and isoprenoid production gives us unparalleled flexibility in culturing cells for isoprenoid production, because we can now effectively partition cellular functions into two distinct independent modules. One such module would consume glucose (or another carbon-energy source) for cell growth and maintenance, whereas the other, consisting of the IUP and isoprenoid-production pathways, would function by consuming externally provided isoprenol/prenol. The ability that IUP affords us to separate growth from isoprenoid production can be leveraged in ways that are impossible/impractical for the natural

pathways. By employing strategies for the co-utilization of two different substrates, we can separate the phases of growth and production and independently tune each phase through the feeding of either substrate. So far, control of the start of the terpenoid production phase has been achieved using inducible promoters, however such systems can be leaky, ineffective and unsuitable for industrial practice. In our pathway, this is reduced to simply deciding when to start feeding a certain substrate. We can easily optimize glucose feeding for controlling cell growth, and subsequently, at a time of our choosing, initiate isoprenoid production through the IUP by adding isoprenol/prenol to the media while simultaneously restricting glucose feeding to limit its use to cell maintenance and ATP production. Thus, the IUP allows for feeding strategies that simply cannot be implemented in current isoprenoid pathways, which rely on glucose.

Another opportunity for innovation opened by the IUP is the very real possibility for using this pathway to introduce hereto unique and untested ways of functionalization into isoprenoid backbones, leading to the production of novel isoprenoid molecules. Such a use of the IUP would be the first real opportunity for easy, one-pot synthesis of a great array of functionalized isoprenoids, greatly expanding the accessible chemical space that can be investigated for drug discovery and testing. We note that it would be exceedingly difficult if not impossible to adopt this strategy (use of functionalized substrates for the production of functionalized isoprenoids) for use with the natural isoprenoid pathways, as the inextricable linkage of the natural isoprenoid pathways with central carbon metabolism would functionalize most or all metabolic intermediates, inactivating critical enzymes and leading to cell death.

The IUP's major advantage is that it is much simpler than either natural isoprenoid pathway and their engineered variants, because it comprises only two steps (vs. 7 steps), uses only ATP as a cofactor (vs. the natural and engineered pathways, which utilize both ATP & NADH) and, according to our preliminary results, there is no significant regulatory inhibition of its enzymatic steps. These advantages translate to a pathway significantly easier to optimize than either the MEP or the MVA pathway and can easily be reconstituted *in vitro*, from start to finish. Moreover,

cofactor balancing should be easier since the single required cofactor's (ATP) availability could be tweaked through oxygen availability/aeration.

Our data shows that, with minimal optimization, the IUP can produce isoprenoids with high flux, approaching the isoprenoid fluxes of pathways that have been optimized for decades (2, 3). Furthermore, we know from our results that we have not yet reached the full potential of the pathway. The IUP thus opens new dimensions in isoprenoid synthesis and the development of low-cost isoprenoid pharmaceuticals. We believe this research to be significant because the IUP can be used to produce numerous isoprenoid products with beneficial pharmaceutical uses, as well as isoprenoid products with uses in other high-impact fields.

Finally, although this thesis and our published work so far (4) focuses on the IUP as used in *E. coli*, we note that the IUP is a simple two-step pathway that can be introduced in almost any microbe. We and our collaborators have successfully expressed the IUP in *Y. lipolytica* (5), *C. glutamicum* and *S. cerevisiae*. Furthermore, the pathway can also be *in vitro* reconstituted and used in a cell-free system for the biosynthesis of isoprenoids (6). Thus, the IUP is a simple, strong and versatile pathway with the potential to revolutionize the field.

8.4. Bibliography

1. Gruchattka E, Hädicke O, Klamt S, Schütz V, Kayser O (2013) In silico profiling of Escherichia coli and Saccharomyces cerevisiae as terpenoid factories. *Microb Cell Fact* 12:84.
2. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* 330(2010):70–74.
3. Tsuruta H, et al. (2009) High-level production of amorpha-4, 11-diene, a precursor of the antimalarial agent artemisinin, in Escherichia coli. *PLoS One* 4(2):e4489.
4. Chatzivasileiou AO, Ward V, Edgar SM, Stephanopoulos G (2019) Two-step pathway for

isoprenoid synthesis. *Proc Natl Acad Sci* 116(2):506–511.

5. Luo Z, et al. Enhancing lycopene synthesis by bypassing the native pathways and modulating intracellular hydrophobicity in Yarrowia lipolytica. *Prep*.

6. Ward VCA, Chatzivasileiou AO, Stephanopoulos G Cell free biosynthesis of isoprenoids from isopentenol. *Prep*.

APPENDIX

Table A1: List of strains and plasmids

Host/Strain	Description	Reference
MG1655 (DE3)	*ΔendA ΔrecA (λ DE3)*	
DH5α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ*(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	NEB
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	NEB
KO1	MG1655 (DE3), *ΔispG,* pBAD33-proA-MEVI	This study
KO2	MG1655 (DE3), *ΔispG,* pBAD33-proA-MEVI, pSEVA228pro4IUPi	This study
KO3	MG1655 (DE3), *ΔispG,* pBAD33-proA-MEVI, pTETIUPi	This study
MFDpir	MG1655 RP4-2-Tc::[ΔMu1::aac(3)IV-ΔaphA-Δnic35-ΔMu2::zeo] ΔdapA::(erm-pir) ΔrecA	(1)
trc-MEP	MG1655 (DE3)::[P_{trc} *dxs-idi-ispD-ispF*]	(2)
T7-MEP	MG1655 (DE3)::[$P_{T7lacUV}$ *dxs-idi-ispD-ispF*]	Manus Bio
T7-Δ-MEP	MG1655 (DE3)::[$P_{T7lacUV}$ *dxs-idi-ispD-ispF*] *ΔispA*	Manus Bio
JW0728-1	*Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *ΔtolR787::kan λ⁻ rph-1 Δ(rhaD-rhaB)568 hsdR514*	(3)
JW1667-5	*Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *λ⁻ Δlpp-752::kan rph-1 Δ(rhaD-rhaB)568 hsdR514*	(3)
JE5505	*Δ(gpt-proA)62 lacY1 tsx-29 glnX44*(AS) *galK2*(Oc) *λ⁻ Δlpp-254 pps-6 hisG4*(Oc) *xylA5 mtl-1 argE3*(Oc) *thiE1*	(4)
TPS30	*ara-600 Δ(gpt-lac)5 tolR8::Cm λ⁻ relA1 spoT1 thiE1*	(5)

Plasmids	Description (origin, antibiotic marker, promoter, operon)	Reference
pET28a(+)	pBR322, Kn^R, $P_{T7lacUV}$, enzymes from Table 2-1 or one of {*ls, ads, ispA, Atipk*} with N terminal 6 x his tag	Novagen
pADS	pTrc99A derivative containing the ADS gene; Ap^R	(6)

pJBEI-6409	p15A, Cm^R, P_{lacUV5}, *atoB*, *hmgs*, *hmgr*, P_{lacUV5}, *mvk*, *pmk*, *pmd*, *idi*, P_{trc}, trGPPS, ls	(7)
pAC-LYCipi	p15A, Cm^R, *crtE*, *ipi*, *crtI*, *crtB*, endogenous promoter	(8)
pSEVA228	RK2, Kn^R, *xlyS-Pm*	(9)
pBbS2k-RFP	SC101, Kn^R, P_{TET} , *rfp*	(10)
pETMEOH500	pBR322, Kn^R, $P_{T7lacUV}$, *mdh*	This study
pTETmdh	pBR322, Kn^R, P_{TET}, *mdh*	This study
pMBIS	RK2, Tc^R, P_{trc}, *erg12*, *erg8*, *mvd1*, *idi*, *ispA*	(6)
pBAD33-proA-MEVI	p15A, Cm^R, P_{proA}, *erg12*, *erg8*, *mvd1*	This study
pCas9	pSC101 ori, RepA101ts, Kn^R, P_{araC}, *cas9*	(11)
pTargetF	pij23119, pMB1, Sp^R	(11)
pTargetF-ispG	pij23119, pMB1, Sp^R	This study
p20-LYCipi	pBR322, Sp^R, *crtE*, *ipi*, *crtI*, *crtB*, endogenous promoter	This study
pUC-LYCipi	pUC19, Sp^R, *crtE*, *ipi*, *crtI*, *crtB*, endogenous promoter	This study
p5T7-LYCipi	pSC101, Sp^R, $P_{T7lacUV}$, *crtE*, *ipi*, *crtI*, *crtB*	This study
p5T7-LYCipi-ggpps	pSC101, Sp^R, $P_{T7lacUV}$, *ggpps*, *ipi*, *crtI*, *crtB*	This study
p5T7tds-ggpps	pSC101, Sp^R, $P_{T7lacUV}$, *tds*, *ggpps*	(2)
p5T7ksl-ggpps	pSC101, Sp^R, $P_{T7lacUV}$, *ksl*, *ggpps*	(12)
p5T7vs-ispA	pSC101, Sp^R, $P_{T7lacUV}$, *vs*, *ispA*	(12)
p5T7ggps-ls	pSC101, Sp^R, $P_{T7lacUV}$, *gpps*, *ls*	This study
p5-T7-LS-L-GPPS	pSC101, Sp^R, $P_{T7lacUV}$, *gpps*, *ls*	Manus Bio
p5-T7-GES-L-GPPS	pSC101, Sp^R, $P_{T7lacUV}$, *gpps*, *ges*	Manus Bio
p5T7ispA-ads	pSC101, Sp^R, $P_{T7lacUV}$, *ispA*, *ads*	This study
P5T7-lyc-tds-kan	pSC101, Kn^R, $P_{T7lacUV}$, *ggpps*, *ipi*, *crtI*, *crtB*, *tds*	This study
p5T7-ispA-aFS	pSC101, Sp^R, $P_{T7lacUV}$, *ispA*, *aFS*	This study
p5207-ispA-aFS	pBR322, Sp^R, $P_{T7lacUV}$, *ispA*, *aFS*	This study
p20T7-aFS-ispA	pBR322, Sp^R, $P_{T7lacUV}$, *ispA*, *aFS*	This study
p20T7-aFS	pBR322, Sp^R, $P_{T7lacUV}$, *aFS*	This study
pSEVA228-pro4IUPi	RK2, Kn^R, P_{pro4}, *ck*, *ipk*, *idi*	This study
pSEVA228-pro5IUPi	RK2, Kn^R, P_{pro5}, *ck*, *ipk*, *idi*	This study
pSEVA228-pro6IUPi	RK2, Kn^R, P_{pro6}, *ck*, *ipk*, *idi*	This study

pSEVA228-proBIUPi	RK2, Kn^R, P_{proB}, *ck, ipk, idi*	This study
pSEVA228-proCIUPi	RK2, Kn^R, P_{proC}, *ck, ipk, idi*	This study
pSEVA228-proDIUPi	RK2, Kn^R, P_{proD}, *ck, ipk, idi*	This study
pTET-IUPi	pBR322, Kn^R, P_{TET}, *ck, ipk, idi*	This study
pTrcsGFP	pBR322, Amp^R, P_{trc}, *sgfp*	(13)
pSEVA228pro4-gfp	RK2, Kn^R, P_{pro4}, *sgfp*	This study
pTET-gfp	pBR322, Kn^R, P_{TET}, *sgfp*	This study
pSEVA228pro4-ck-idi	RK2, Kn^R, P_{pro4}, *ck, idi*	This study
pSEVA228-EC-WT	RK2, Kn^R, *xlyS-Pm, EcispG*	This study
pSEVA228-EC-271	RK2, Kn^R, *xlyS-Pm, EcispG* (T271C)	This study
pSEVA228-EC-273	RK2, Kn^R, *xlyS-Pm, EcispG* (S273C)	This study
pSEVA228-EC-303	RK2, Kn^R, *xlyS-Pm, EcispG* (G303C)	This study
pSEVA228-AA-WT	RK2, Kn^R, *xlyS-Pm, AaispG*	This study
pSEVA228-AA-269	RK2, Kn^R, *xlyS-Pm, AaispG* (G269C)	This study
pSEVA228-AA-299	RK2, Kn^R, *xlyS-Pm, AaispG* (G299C)	This study
pSEVA228-TT-WT	RK2, Kn^R, *xlyS-Pm, TtispG*	This study
pSEVA228-TT-299	RK2, Kn^R, *xlyS-Pm, TtispG* (G299C)	This study
pSEVA228-TT-301	RK2, Kn^R, *xlyS-Pm, TtispG* (G301C)	This study
pSEVA228-TT-342	RK2, Kn^R, *xlyS-Pm, TtispG* (G342C)	This study
p5TrcMEP	pSC101, Sp^R, P_{trc}, *dxs, ipi, ispD, ispF*	(2)
P5TrcPQQs	pSC101, Sp^R, P_{trc}, *pqqs*	This study
P5TrcFeSOD	pSC101, Sp^R, P_{trc}, *fe-sod*	This study
P5TrcAhpC	pSC101, Sp^R, P_{trc}, *AhpC*	This study
pET-araC-SOD_B	pBR322, Kn^R, P_{araC}, *SOD_B*	This study

Ap^R = ampicillin, Kn^R = kanamycin, Tc^R = Tetracyclin, Sp^R = Spectinomycin

Table A2: List of genes used in this study and their origins

Genes	Origin (Accession Number)
tds	*Taxus brevifolia* (AAC49310.1), codon optimized, truncated first 60 amino acids, methionine added
ggpps	*Taxus canadensis* (AAD16018.1), codon optimized, truncated first 98 amino acids, methionine added
crtE, crtI, crtB, ipi	*Pantoea agglomerans, crtE* (AAA21260.1), *crtB* (AFZ89043.1), *crtI* (AFZ89042.1), *ipi* (AAA64978.1)
ksl	*Salvia miltiorrhiza*, codon optimized, methionine added, (ABV08817.1)
vs	*Callitropsis nootkatensis*, codon optimized, methionine added (AFN21429.1)
ls	*Mentha spicata* (AAC37366.1), codon optimized
gpps	*Abies grandis* (AAN01134.1), codon optimized
ads	*Artemisia annua* (AEQ63683.1), codon optimized
ispA	*E. coli* (WP_097750737.1)
ipk	*Arabidopsis thaliana* (AAN12957.1), codon optimized
pmd	*S. cerevisiae* (AAC49252.1), codon optimized
Scck	*S. cerevisiae* (AAA34499.1), codon optimized
Hvipk	*Haloferax volcanii* (ADE04091.1), codon optimized
Mtipk	*Methanothermobacter thermautotrophicus* (AAB84554.1), codon optimized
Mjipk	*Methanocaldococcus jannaschii* (AAB98024.1), codon optimized
Taipk	*Thermoplasma acidophilum* (CAC11251.1), codon optimized
Taipk-3m	*Thermoplasma acidophilum* (CAC11251.1), codon optimized, three mutations (V72I , Y140V, K203G)
Ecgk	*E. coli* (AAA23913.1)
erg12/Scmk	*S. cerevisiae* (CAA29487.1), codon optimized
Echk	*E. coli* (AAC73114.1)
idi	*E. coli* (AAD26812.1)
EcispG	*E. coli* (AAC07467.1)
AaispG	*A. aeolicus* (CAA45783.1), codon optimized
TtispG	*T. thermophilus* (BAD70128.1), codon optimized

pqqs	*D. radiodurans* (AAF12672.1), codon optimized
SOD_B	*E. coli* (CAA27580.1)
fe-sod	*Synechocystis sp. PCC 6803* (BAA18027.1), codon optimized
AhpC	*E. coli* (BAA02485.1)
aFS	*A. annua* (CAC12731.1), codon optimized
ges	*O. basilicum* (AAR11765.1), codon optimized

Table A3: List of primers used in this study

Name	Sequence
GB_p5t7tds-ggpps_r	ATGGTATATCTCCTTATTAAAGTTAAAC
GB_p5t7tds-ggpps_f	TATTAGTTAAGTATAAGAAGGAGATATAC
GB_gpps_ls_f	TAATAAGGAGATATACCATATGGAATTTGACTTCAACAAATAC
GB_gpps_ls_r	CTTCTTATACTTAACTAATACGAGGAAGCGGAATATATC
GB_ispA_f	TAATAAGGAGATATACCATATGGACTTTCCGCAGCAAC
GB_ispA_r	CTCCTTCTTAAAAGATCCTTTATTTATTACGCTGGATGATGTAGTC
GB_ads_f	GTAATAAATAAAGGATCTTTTAAGAAGGAGATATACATGGCCCTGAC CGAAGAG
GB_ads_r	CTTCTTATACTTAACTAATATCAGATGGACATCGGGTAAAC
GB_pAC-LYCipi_r	CAGTTATTGGTGCCCTTAAACG
GB_pAC-LYCipi_f	TAAGCTTTAATGCGGTAGTTTATCAC
GB_aadA1_f	AGGGCACCAATAACTGGGTGAACACTATCCCATATC
GB_aadA1_r	TAACCGTATAATCATGGCAATTCTGGAAG
GB_pUC19_f	GCCATGATTATACGGTTATCCACAGAATC
GB_pUC19_r	CTACCGCATTAAAGCTTAAGGATCTAGGTGAAGATC
GB_pBR322_f	ATTGCCATGATTCCCCTTGTATTACTGTTTATG
GB_pBR322_r	CTACCGCATTAAAGCTTAACTCAAAGGCGGTAATAC
GB_p5T7_r	ATGGTATATCTCCTTATTAAAGTTAAACAAAATTATTTCTACAGGG
GB_ p5T7 _f	TTAATAAGGAGATATACCATATGGTGAGTGGCAGTAAAGC
GB_p20-LYCipi_f	CTCCTTCTTATACTTAACTAATACTGCGTGAACGTCATGGC
GB_p20-LYCipi_r	TATTAGTTAAGTATAAGAAGGAGATATAC
GB-pET28-HisT-vec f	CACCACCACCACCACCAC
GB-pET28-HisT-vec r	CGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTC
GB-pET28-PMD_f	CTTTAAGAAGGAGATATACCATGTACCATCGTCCACTCG
GB-pET28-PMD_r	CAGTGGTGGTGGTGGTGGTGCTCTGCAACGCCAAGTTCAC
GB-pET28-ScCK_f	AAGAAGGAGATATACCGATGGTACAAGAATCACGTC
GB-pET28-ScCK_r	TCAGTGGTGGTGGTGGTGGTGCAAATAACTAGTATCGAGGAAC
GB-pET28-EcGK f	AAGAAGGAGATATACCGATGACTGAAAAAAAATATATCGTTGC

GB-pET28-EcGK r	TCAGTGGTGGTGGTGGTGGTGTTCGTCGTGTTCTTCCCAC
GB-pET28-EcHK f	AAGAAGGAGATATACCGATGGTTAAAGTTTATGCCCC
GB-pET28-EcHK r	TCAGTGGTGGTGGTGGTGGTGGTTTTCCAGTACTCGTGC
GB-pET28-TaIPK-3m f	AAGAAGGAGATATACCGATGATGATTCTGAAAATCGGAG
GB-pET28-TaIPK-3m r	TCAGTGGTGGTGGTGGTGGTGTCGAATGACAGTACCGATG
GB-pET28-MjIPK_f	AAGAAGGAGATATACCGATGCTGACCATCCTGAAATTAG
GB-pET28-MjIPK_r	TCAGTGGTGGTGGTGGTGGTGTTCGCTAAAGTCGATCTC
GB-pET28-TaIPK f	AAGAAGGAGATATACCGATGATGATTCTTAAGATAGGGGG
GB-pET28-TaIPK r	TCAGTGGTGGTGGTGGTGGTGACGAATGACGGTTCCGATG
GB-pET28-Mtipk_f	AAGAAGGAGATATACCGATGATCATTCTGAAACTGGG
GB-pET28-Mtipk_r	TCAGTGGTGGTGGTGGTGGTGATGTTTTCCTGTGATACGC
GB-pET28-HvIPK_f	AAGAAGGAGATATACCGATGTCCCTGGTGGTCCTTAAA
GB-pET28-HvIPK_r	TCAGTGGTGGTGGTGGTGGTGTTCCCCGCGAATGACTGT
GB-pET28-ScMK_f	TTTAAGAAGGAGATATACCGATGTCATTACCGTTCTTAAC
GB-pET28-ScMK_r	CAGTGGTGGTGGTGGTGGTGCTATGAAGTCCATGGTAAATTC
GB-pETMeOH500_f	ATGACCCACCTGAACATC
GB-pETMeOH500_r	GCGCAACGCAATTAATGTAAG
GB-pBBS2k-rfp_f	TTACATTAATTGCGTTGCGCTTAAGACCCACTTTCACATTTAAG
GB-pBBS2k-rfp _r	GCGATGTTCAGGTGGGTCATATGTATATCTCCTTCTTAAAAGATC
GB-pET28-atIPK_f	AAGAAGGAGATATACCGATGGAACTCAATATCAGCG
GB-pET28-atIPK_r	GTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGTTTGCTGAA GCGGATGATG
GB-pET28-IDI_f	AAGAAGGAGATATACCGATGCAAACGGAACACGTC
GB-pET28-IDI_r	GTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGTTTAAGCTG GGTAAATGCAG
GB-pET28-IspA_f	AAGAAGGAGATATACCGATGGACTTTCCGCAGCAAC
GB-pET28-IspA_r	GTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGTTTATTACGC TGGATGATGTAGTC
GB-pET28-ADS_f	AAGAAGGAGATATACCGATGGCCCTGACCGAAGAG

GB-pET28-ADS_r	GTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGGATGGACATCGGGTAAACC
GB-pET28-LS_f	AAGAAGGAGATATACCGATGCGTCGCAGTGGTAATTAC
GB-pET28-LS_r	GTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGGGCGAAAGGTGCAAACAG
GB-pTet-IUP-Ins_f	TTTAAGAAGGAGATATACATATGGTGCAGGAGTCCCGC
GB-pTet-IUP-Ins_r	GTCGACGGAGCTCGAATTCGTTATTTGCTGAAGCGGATGATGGTC
GB-pTet-Vec_f	CGAATTCGAGCTCCGTCG
GB-pTet-Vec_r	ATGTATATCTCCTTCTTAAAAGATCTTTTGAATTC
Pro4 Mut_f	GGGCATGCATAAGGCTCGGATGATATATTCAGGGAGACC
Pro5 Mut_f	GGGCATGCATAAGGCTCGTAGGATATATTCAGGGAGACC
Pro6 Mut_f	GGGCATGCATAAGGCTCGTAAAATATATTCAGGGAGACC
ProB Mut_f	GGGCATGCATAAGGCTCGTAATATATATTCAGGGAGACC
ProC Mut_f	GGGCATGCATAAGGCTCGTATGATATATTCAGGGAGACC
ProD Mut_f	GGGCATGCATAAGGCTCGTATAATATATTCAGGGAGACC
ProLibrary_Mut_r	CGAGCCTTATGCATGCCC
GB-SEVA228_f	GGGTCCCCAATAATTACG
GB-SEVA228_r	CAGCTGGGCGCGCCGTAG
GB-proD_f	TTCTACGGCGCGCCCAGCTGTTCTAGAGCACAGCTAACAC
GB-proD_r	TCCTTGCGTTGAAACCGTTGTGGTCTCC
GB-chk_f	CAACGGTTTCAACGCAAGGAAACACATTAAG
GB-chk_r	TTTCTTGTACTTACAGGTAGCTGGTGTC
GB-atipk_f	CTACCTGTAAGTACAAGAAAAGTCAGTAGTC
GB-atipk_r	CTCCTTAGTTTTATTTGCTGAAGCGGATG
GB-iditerm_f	CAGCAAATAAAACTAAGGAGGTCTATATGC
GB-iditerm_r	ATCGTAATTATTGGGGACCCGATATAGTTCCTCCTTTCAG
GB-IUPnoIPK_f	CTACCTGTAAAACTAAGGAGGTCTATATGC
GB-IUPnoIPK_r	CTCCTTAGTTTTACAGGTAGCTGGTGTC
pCas9-ispG_f	GCGACATTGAAGAAGATAAGG
pCas9-ispG_r	GTTTACGGTGTAAGCGATCC

pCas9-ispG-seq_f	GATTGCTGGCTGGAGGTCAC
GB-pTargetF-ispGN20_f	GTCCTAGGTATAATACTAGTCGCTGCGTATCCGTTCGCGAGTTTTAGAGCTAGAAATAGC
GB-ptargetF-N20_r	ACTAGTATTATACCTAGGACTGAG
GB-pTargetF-vec_f	CACCACCGACTATTTGCAAC
GB-pTargetF-vec_r	CTCGAGTAGGGATAACAGGGTA
GB-ispG-H1_f	CCCTGTTATCCCTACTCGAGCCAGCGTCTGTGGATACTACC
GB-ispG-H1_r	TCCCATCACGTCTCCCGCGTTACCCGTC
GB-ispG-H2_f	ACGCGGGAGACGTGATGGGAAGCGCCTC
GB-ispG-H2_r	GTTGCAAATAGTCGGTGGTGCTTCGCAGCCCAACTGATG
p5T7Lyc-ggpps_f	TTAATAAGGAGATATACCATATGTTCGACTTCAACGAG
p5T7Lyc-ggpps_r	TTGAACCCAAAAGGGCGGTATTAGTTTTGACGAAAGGC
p5T7Lyc-back_f	TACCGCCCTTTTGGGTTC
p5T7Lyc-back_r	ATGGTATATCTCCTTATTAAAGTTAAAC
GB-sGFP-pSEVA F	GAAAGAGGAGAAATACTAGTATGAGCAAGGGCGAAGAG
GB-sGFP-pSEVA R	CAAGCTTGTCGACGGAGCTCTTACTTATAGAGTTCATCCATGCC
GB-pSEVA-back F	GAGCTCCGTCGACAAGCTTG
GB-pSEVA-back R	ACTAGTATTTCTCCTCTTTCTCTAGTAAAAGTTAAAC
GB-sGFP-pTET F	TTTAAGAAGGAGATATACATATGAGCAAGGGCGAAGAG
GB-sGFP-pTET R	GTCGACGGAGCTCGAATTCGTTACTTATAGAGTTCATCCATGCC
GB-pTET-back F	CGAATTCGAGCTCCGTCG
GB-pTET-back R	ATGTATATCTCCTTCTTAAAAGATCTTTTGAATTC
GB-pBro IAI Vec F	GAGCTCCGTCGACAAGCT
GB-pBro IAI Vec R	ACTAGTATTTCTCCTCTTTCTCTAGTAAAAG
GB-proX-Mevi Ins F	CTAGAGAAAGAGGAGAAATACTAGTATGTCATTACCGTTCTTAACTTC
GB-proX-Mevi Ins R	CAAGCTTGTCGACGGAGCTCTTATTCCTTTGGTAGACCAG
GP-pSEVAispGBB F	TAACGTGATGGGAAGCGC
GP-pSEVAispGBB R	GAAAAATCTCCCGCGTTAC
GB-pSEVA-ECispG F	ggtaacgcgggagatttttcATGCATAACCAGGCTCCAATTC

GB-pSEVA-ECispG R	aggcgcttcccatcacgttaTTATTTTTCAACCTGCTGAACG
GB-pSEVA-AAispG F	ggtaacgcgggagatttttcATGATCCAGAAACGTAAAACCC
GB-pSEVA-AAispG R	aggcgcttcccatcacgttaTTAGTTGGTGCCACCGTC
GB-pSEVA-TTispG F	ggtaacgcgggagatttttcATGGAAGGTATGCGTCGTC
GB-pSEVA-TTispG R	aggcgcttcccatcacgttaTTAGGCTTTCGGTGCAAAAC
GB-P5TRC-BB-F	GTCGACCATCATCATCATC
GB-P5TRC-BB-R	ATGTTATTCCTCCTTATTTAATCG
GB-P5TRC-Ahpc-F	TAAATAAGGAGGAATAACATATGTCCTTGATTAACACC
GB-P5TRC-Ahpc-R	TGATGATGATGATGGTCGACTTAGATTTTACCAACCAGG
GB-P5TRC-PQQ-F	TAAATAAGGAGGAATAACATATGGTGGCATTTCTCCGTG
GB-P5TRC-PQQ-R	TGATGATGATGATGGTCGACTCATGCGTGACTTACCAATG
GB-P5TRC-SOD-F	TAAATAAGGAGGAATAACATATGGCATATGCACTGCCG
GB-P5TRC- SOD-R	TGATGATGATGATGGTCGACTTATGCTGCTGCCAGATTTG
GB_p5t7ggppstds_f	ACGCAGTATTAGTTAAGTATGAAATAATTTTGTTTAACTTTAAGAAGGA
GB_p5t7ggppstds_r	GTGCCACCTGCATTCGCCAATCCGGATATAG
GB-KanRes_f	TTGGCGAATGCAGGTGGCACTTTTCGGG
GB-KanRes_r	ATGAACAGCATGTAACACCTTTTCTACGGGGTCTGACG
GB-p5T7Lyc-back_f	AGGTGTTACATGCTGTTC
GB-p5T7Lyc-back_r	ATACTTAACTAATACTGCGTGAAC
GB-TDS_in_p5T7_f	ATAAGGAGATATACCATATGTCTAGCTCTACGGGTACG
GB-TDS_in_p5T7_r	AACAAAATTATTTCTAGTTTAGACCTGGATTGGATCGATG
GB-p5T7-backbone_f	TAAACTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
GB-p5T7-backbone_r	CATATGGTATATCTCCTTATTAAAGTTAAAC
GB-p5T7ispA-F	TTAAGTATAAGAAGGAGATATACATATG
GB-p5T7ispA-R	GATCCTTTATTTATTACGCTG
GB-aFS-F	AGCGTAATAAATAAAGGATCTTTTAAGAAGGAGATATACATGG
GB-aFS-R	TATCTCCTTCTTATACTTAACTAATATTAAACCACCATCGG
GB-p20ori-F	CGTGTTGTCGGCAGCTTTGATTCGATAAGCAGCATCGC

GB-p20ori-R	TGAATCGGCCAACGCGAATTCCCGACAGTAAGACGGGTAAG
GB-T7ispAaFS-F	AATTCGCGTTGGCCGATTC
GB-T7ispAaFS-R	TCAAAGCTGCCGACAACAC
GB-p20T7-F	TTAAGTATAAGAAGGAGATATACATATG
GB-p20T7-R	TTAAACAAAATTATTTCTACAGGG
GB-ispA_in_p20T7-IspA-aFS_F	TGTAGAAATAATTTTGTTTAACTTTAATAAGGAGATATACCATATGG
GB-ispA_in_p20T7-IspA-aFS_R	CTTCTTAAAAGATCCTTTATTTATTACGCTG
GB-aFS_in_p20T7-IspA-aFS_F	ATAAAGGATCTTTTAAGAAGGAGATATACATGG
GB-aFS_in_p20T7-IspA-aFS_R	TATCTCCTTCTTATACTTAACTAATATTAAACCACCATCGG
GB-aFS_in_p20T7-aFS_F	TGTAGAAATAATTTTGTTTAATTTTAAGAAGGAGATATACATGG
GB-aFS_in_p20T7-aFS_R	TATCTCCTTCTTATACTTAACTAATATTAAACCACCATCGG
GB-EC271-F	TGTTCGCGTCAGGAATTTG
GB-EC271-R	ATTCCTGACGCGAACAACACGGGCAGGCGATGAAG
GB-EC273-F	CGTCAGGAATTTGATGTTA
GB-EC273-R	CATCAAATTCCTGACGACAACAGGTCGGGCAGGCG
GB-EC303-F	TGCGTGGTGAATGGCCCAG
GB-EC303-R	GGCCATTCACCACGCAACAGATAATCGAAACGTCC
GB-TT-299-F	TGTGGTCGTACCACCAGTA
GB-TT-299-R	TGGTGGTACGACCACAACACGGACAGCTGGTAACT
GB-TT-301-F	CGTACCACCAGTACCTTTT
GB-TT-301-R	AGGTACTGGTGGTACGACAACAACCCGGACAGCTG
GB-TT-342-F	TGCGTTGTTAATGGTCCGG
GB-TT-342-R	GACCATTAACAACGCAACACATAACTGCAACTTTC
GB-AA269-F	CGTATTGAAGTGGATCTGC
GB-AA269-R	GATCCACTTCAATACGACAACAGGTCGGACATGCA

GB-AA299-F	TGTGTTGTTAATGCAATTG
GB-AA299-R	TTGCATTAACAACACAACACATAACTGCAACTTTC

Table A4: Custom elements used for the creation of vector pSEVA228-proDIUPi

Promoter sequence
TTCTAGAGCACAGCTAACACCACGTCGTCCCTATCTGCTGCCCTAGGTCTATGAGTGGTT GCTGGATAACTTTACGGGCATGCATAAGGCTCGTATAATATATTCAGGGAGACCACAACG GTTTC
RBS for *ck*
AACGCAAGGAAACACATTAAGGAGGTTTAA
RBS for *ipk*
GTACAAGAAAAGTCAGTAGTCTAAGGAGGTAAGC
RBS for *idi*
AACTAAGGAGGTCTAT
T7 terminator region
GCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGC ATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTA TATC

Table A5: Custom-synthesized, codon optimized gene sequences

AaispG

ATGATCCAGAAACGTAAAACCCGTCAGATTCGTGTTGGCAATGTTAAAATTGGTGGTGATG
CACCGATTGTTGTTCAGAGCATGACCAGCACCAAAACCCATGATGTTGAAGCAACCCTGAA
TCAGATTAAACGTCTGTATGAAGCCGGTTGCGAAATTGTTCGTGTTGCAGTTCCGCATAAA
GAAGATGTGGAAGCACTGGAAGAAATCGTGAAAAAAAGCCCGATGCCGGTTATTGCCGAT
ATTCATTTTGCACCGAGCTATGCATTTCTGAGCATGGAAAAAGGTGTTCATGGCATTCGTAT
TAATCCGGGTAATATTGGCAAAGAAGAGATCGTTCGCGAAATCGTTGAAGAGGCAAAACGT
CGTGGTGTTGCCGTTCGTATTGGTGTTAATAGCGGTAGCCTGGAAAAAGATCTGCTGGAAA
AATATGGTTATCCGAGCGCAGAAGCCCTGGCAGAAAGCGCACTGCGTTGGAGCGAAAAAT
TTGAAAAATGGGGCTTCACCAACTATAAAGTGAGCATTAAAGGTAGTGATGTGCTGCAGAA
TGTTCGTGCCAATCTGATTTTTGCAGAACGTACCGATGTTCCGCTGCATATTGGTATTACCG
AAGCAGGTATGGGCACCAAAGGTATTATCAAAAGCAGCGTTGGTATTGGCATTCTGCTGTA
TATGGGTATTGGTGATACCGTTCGTGTTAGCCTGACCGATGATCCGGTTGTTGAAGTTGAA
ACCGCATATGAAATTCTGAAAAGCCTGGGTCTGCGTCGTCGCGGTGTTGAAATTGTTGCAT
GTCCGACCTGTGGTCGTATTGAAGTGGATCTGCCGAAAGTTGTTAAAGAGGTTCAAGAAAA
ACTGAGCGGTGTGAAAACACCGCTGAAAGTTGCAGTTATGGGTTGTGTTGTTAATGCAATT
GGCGAAGCACGTGAAGCAGATATTGGTCTGGCATGTGGTCGTGGTTTTGCATGGCTGTTC
AAACATGGTAAACCGATCAAAAAAGTGGATGAAAGCGAAATGGTTGACGAGCTGCTGAAAG
AAATTCAGAATATGGAAAAAGACGGTGGCACCAACTAA

TtispG

ATGGAAGGTATGCGTCGTCCGACCCCGACCGTTTATGTTGGTCGTGTTCCGATTGGTGGT
GCACATCCGATTGCAGTTCAGAGCATGACCAATACCCCGACCCGTGATGTTGAAGCAACC
ACCGCACAGGTTCTGGAACTGCATCGTGCAGGTAGCGAAATTGTTCGTCTGACCGTTAATG
ATGAAGAAGCAGCAAAAGCAGTTCCGGAAATTAAACGTCGTCTGCTGGCAGAAGGTGCCG
AAGTTCCGCTGGTTGGTGATTTTCATTTCAATGGTCATCTGCTGCTGCGCAAATATCCGAAA
ATGGCAGAAGCACTGGATAAATTTCGTATTAATCCGGGTACACTGGGTCGTGGTCGTCATA
AAGATGAACATTTTGCAGAAATGATCCGCATTGCCATGGATCTGGGTAAACCGGTTCGTAT
TGGTGCAAATTGGGGTAGCCTGGATCCGGCACTGCTGACCGAACTGATGGATCGTAATGC
ACGTCGTCCGGAACCGAAAAGCGCACATGAAGTTGTTCTGGAAGCCCTGGTTGAAAGCGC
AGTTCGTGCATATGAAGCAGCACTGGAAATGGGTCTGGGTGAAGATAAACTGGTTCTGAG
CGCAAAAGTTAGCAAAGCACGTGATCTGGTTTGGGTTTATCGTGAACTGGCACGTCGTACC
CAGGCACCGCTGCATCTGGGTCTGACCGAAGCAGGTATGGGTGTTAAAGGTATTGTTGCA
AGCGCAGCCGCACTGGCTCCGCTGCTGCTGGAAGGCATTGGTGATACCATTCGTGTTAGC
CTGACACCGGCACCGGGTGAACCGCGTACCAAAGAAGTTGAAGTTGCACAAGAAATTCTG
CAGGCACTGGGTCTGCGTGCATTTGCACCGGAAGTTACCAGCTGTCCGGGTTGTGGTCGT
ACCACCAGTACCTTTTTTCAAGAACTGGCCGAAGAAGTTAGCCGTCGTCTGAAAGAACGTC
TGCCGGAATGGCGTGCACGTTATCCGGGTGTTGAAGAACTGAAAGTTGCAGTTATGGGTT
GCGTTGTTAATGGTCCGGGTGAAAGCAAACATGCACATATTGGTATTAGCCTGCCTGGTGC
GGGTGAAGAACCTAAAGCACCGGTTTATGCAGATGGTAAACTGCTGACAATTCTGAAAGGT
GAAGGTATCGCAGAAGAATTTCTGCGTCTGGTGGAAGATTATGTGAAAACCCGTTTTGCAC
CGAAAGCCTAA

fe-sod

ATGGCATATGCACTGCCGAATCTGCCGTATGATTATACCGCACTGGAACCGTGTATTAGCA
AAAGCACCCTGGAATTTCATCATGATAAACACCATGCAGCCTACGTGAACAATTTTAACAAT
GCAGTTGCAGGCACCGATCTGGATAATCAGAGCATTGAAGATGTTATTAAAGCCGTTGCCG
GTGATGCAAGCAAAGCAGGTATTTTCAATAATGCAGCACAGGCATGGAACCACAGCTTTTA
TTGGAATTGTATGAAACCGGGTGGTGGTGGTCAGCCGAGCGGTGCACTGGCAGATAAAAT
CAATGCAGATTTTGGTAGCTTTGACGCCTTTGTGGAAGCATTTAAACAGGCAGGCGCAACC
CAGTTTGGTAGCGGTTGGGCATGGCTGGTGCTGGATAATGGCACCCTGAAAGTTACCAAA
ACCGGTAATGCAGAAAATCCGATGACCGCAGGTCAGACACCGCTGCTGACCATGGATGTT
TGGGAACATGCATATTATCTGGATTATCAAAATCGCCGTCCGGATTACATTGCCGATTTTCT
GGGTAAACTGGTGAATTGGGATTTTGTTGCAGCAAATCTGGCAGCAGCATAA

pqqs

ATGGTGGCATTTCTCCGTGGCCCAACACATGTAGATTTCAATCTAACGAACGGTTGTAATTT
GGCTTGCTCTCACTGTCACAGTGCTTCGGGACCGAAACTGGATAACGAGCTTAAAACTGAA
GAGATTCTTCAGACAATCGATGCCCTCCATATCATAGGTGCTTTAAAAATTGCTTTTGCAGG
CGGGGAACCTTTTATCAGACGTGACATCTTTAATATTCTGAGTCATGCCTGTAGTCTTCCTG
GCTGGGGTATTTCAGTGATCACCAACGGTTTTTACCTCAACAGTCTAACAGTAGAAAAACTT
AAGGCTCAATGTCCAAATTTAAGCATCAACATTAGCGTGGATGGTTCGACACCAGCCGGTT
ACAGCACACTCCGCAAGCAATTGAATCGCCCGGATGCCGATCCTCAACCTCTGTTTGAACG
TGTTCTATCCGGCATCGACAACGTGGTGAGATCTGGCATGAGCAACTCGGTCAACTTCACC
ATTACCAAAGCCACTCTGCACGACATTGAAGCTACATATGAACTTGTTGTGGATCGCATAG
GTGCCGACAACATGGTGGCGATTAAATTCTTCCCTGGCGGATACGGGAAAGAACATCTGG
ATCTGTACGAAATTCCCTATGACATGTGGGACGAACATTTCGTCGCCTTGACCCGCGCGAA
ACTTGGAGGGAAACTTGAGCGGCTACAAATCTCGGTTCCGGCAGCCTGGGAGTTCTACCT
GCCGTTGATCAATGCGGGTATTCCGATCGAGGAAGCGGAAGAAGTATGGGGTTACCGCAG
CCCACTTAGGGAAGACATGTACGCTCGCATGCGGGAAGTGGGCGACGTGGCGGGCATTT
CGGAACTGTGTATCTCTTCAGACGGCGAGGTGTACCCGTCAGTGCTATTGGTCGGGGAAA
AAACGATGAGCTGCGGGAACCTGCGCGAGCACTCATTGCAGCGTATCTGGCAGGAATCTT
CATGCCTGATGGCTCTGCGAGAGTTGAAGCTCTTCGATCTAAATGGCAATTGCACAAAGTG
CGGGGTAAGGGAGGTATGCGGAGGCGGCTCCCGCTCACGGGCCTTCTCCCGCGAAGGC
GACTTCCGTGATATGGACTACGTGTGCCCGATCGTCGAGGTTCCATTGGTAAGTCACGCAT
GA

Pmd

ATGTACCATCGTCCACTCGAACGCATGCCTGGCCTGGCGACTGCCATTCCGGCCCCATAC
GCAGATCTTGTCGAGCGCATGGCCACGGCCGATGCGGACCTGCGCGCTGCTCTGCGTGC
GCATGGCCTGACTTGGGAAGAGTATCCGGATCTGACTGGCGCTGCACGCGAACGTGGGG
CGGCTGCGGCACTTGCGTACCCGATGCAGGGCGTTTTGAAGTATCATGGACTGTCTGACT
GGGATTATCGCATTGCATTTTTACCGTCGGTATCTTTGTGTAACGATGCGGGCCACACCTT
GACGCTGGTCGAATTTGACCCGGACCTGGCGACGGATTGTGCCACCATCAACGGCCATGT
TGCGCGCGGGCGTGAACTGGAACGCGTCCGCCAGAGCCTGGACGCTATCCGTGCTGCTA
GCGGGGCGACAGTTCGTGCGCGCGTAATGTCGCGCAACGTGACGCGCGGAACCCGTATG
GGAAAGGGCTTAGGTTCCTCGGCTGCGGCAAGCGCGGCGCTCGCGTTGGCAGCCATCGC

CGCCCTGTATGGGGATGAAGCGGCCGCGAATCGCCGCCTGGTCTCTTGCATGGCACGCC
TTCTCGCGGGCTCGGGTTGCCGCAGCGCCGCGGGTGGGTGTTCTATTTGGTTCAGTAGTC
CGGGTATGCCTCATGAGGATTCATTTGCGGTGCGTCTGGACGATGCGGGGCAGTTGGATG
ATGTTCGCCTCATTACGGTCCCATTAGATAGTCGCATTGGTTTGAAAACCGAACAAGCACA
CCTCGATGCGCCGGGGAGCGCTCTTTTTCGCTGCTGGATGTTAAGTCGTCGTGACGAGGC
TCTGGCATGCATCGCAGCGGCCCGTACGGGGGATTGGCGCACCTTAGGTCAATGGGCCG
AGCTGGACAGCATGCGCCTGCACGGTATTACCATGTCTGGTTCCCTGGAGAACAAGCTTAT
TGGCTGGGAACCAGAAAATATTGTTTTATTCCGCATGTGCAACGATCTGCGTTCAAGCGGC
GTGCCCGTATATTGCTCCACCGACACTGGCCCTACGGCCGTGTTCATTACCCATCGCGAC
TATGAAGACGCAGTTGTATCTGCTATCGAAGGGCTGGGGCTCAGCTTGGAAGCTATCCGC
GGCCGCGTAGCCGGTCCGGCCCGCCTGGTTGACATCGCTTGGGCGGGCGGTGAACTTGG
CGTTGCAGA

Hvipk

ATGTCCCTGGTGGTCCTTAAACTGGGTGGATCCGTCGTCACCGATAAAGATGAACCTGAGA
CAGTAGACGAGGCGGGACTGGCTGCCGCGGCAGACGCGGTTGCCCCTCTGGCAGAGTC
GCGGCGGGTAGTCGTTGTCCATGGAGGGGGATCATTTGGCCACCACCATGCAGCAGAAC
ATGGGGTTAGTTCCGAATCAGGCTCCCATGATGCCAGAGGCGTCCGTGCAATCCATGACG
CGATGAAACGTTTGAACGACGCCGTGCTGGATGCACTCGAAGAGCGGGGTGTGGCGGCT
CTGCCGGTACATCCGCTGTCAGCGGGTGCAAGAGAAGCAGACGGCTCGCTGAGCCTGCC
GCTGGCTGCCACGGAAACGATGTTGGATGAAGGATTTGTGCCTGTGCTCCATGGTGACGT
CATCTCTCATGCAGGAAAGGGGGCTACAATAGTGTCTGGGGATGATTTGGTGGTAAGCCTT
GCCTCAGGCCTAGGCGCCGACCGCGTCGGTTTATGTAGCACTGTTCCGGGCGTTCTGGAT
GCAGACGGCGATGTAATTCCAGAAATAACGGCGTTTGCCGACGCAGCAGATGCGCTGGGA
GGTTCAGATAGTACCGACGTGACCGGTGGGATGGCCGCCAAAGTGAGAAAACTGCTTGCC
TTGGGAGCGCCAGCCCATGTATTTGGTCCGGAAGGATTAAGTGCATTTGTTGCAGGTGAAT
CCCCTGGTACAGTCATTCGCGGGGAA

Mtipk

ATGATCATTCTGAAACTGGGCGGATCTGTTATAACGAGAAAAGACTCTGAAGAACCTGCGA
TTGACAGGGATAATCTAGAAAGAATCGCCTCTGAGATTGGAAACGCATCTCCCAGCAGCCT
TATGATTGTGCATGGGGCTGGTTCTTTTGGACATCCATTTGCCGGTGAATATCGTATAGGG
AGTGAAATTGAAAATGAGGAGGATTTACGTCGCCGTCGCTTTGGATTTGCTCTCACACAGA
ATTGGGTTAAGAAGTTGAACTCCCACGTATGCGACGCACTTCTGGCAGAAGGTATTCCTGC
GGTATCCATGCAGCCTTCTGCGTTCATACGTGCGCACGCTGGCCGCATCAGCCATGCCGA
TATCTCGCTGATCCGAAGTTATTTAGAAGAAGGTATGGTACCGGTCGTGTATGGAGATGTG
GTACTGGACAGTGATAGACGACTTAAATTTTCTGTAATTTCAGGTGATCAGCTGATTAATCA
CTTCTCCTTAAGACTCATGCCGGAACGCGTTATTTTAGGCACCGACGTTGATGGCGTGTAT
ACTCGTAACCCTAAAAAACACCCTGACGCACGCTTACTCGACGTGATAGGTTCGCTGGACG
ATCTGGAATCGTTAGACGGGACTCTCAACACAGATGTAACGGGAGGGATGGTTGGAAAAA
TTCGCGAATTGCTGCTATTAGCCGAGAAGGGTGTGGAAAGCGAAATTATAAATGCCGCAGT
CCCCGGTAATATCGAACGCGCTCTGCTGGGCGAAGAAGTACGCGGCACGCGTATCACAG
GAAAACAT

Mjipk

ATGCTGACCATCCTGAAATTAGGCGGTAGTATCCTGTCAGATAAAAACGTGCCATATTCTAT
CAAGTGGGACAATTTGGAACGCATCGCGATGGAAATTAAAAACGCCTTAGACTACTACAAA
AATCAGAACAAAGAAATCAAACTGATTTTAGTACACGGGGGGGGCGCATTTGGGCACCCA
GTCGCGAAAAAGTATCTTAAGATTGAAGACGGGAAAAAAATTTTTATTAACATGGAAAAAGG
TTTTTGGGAAATCCAGCGCGCGATGCGTCGTTTTAACAACATTATTATTGATACCCTCCAAA
GTTATGATATCCCTGCAGTTTCTATCCAGCCGTCGTCGTTCGTAGTGTTTGGTGACAAACT
GATCTTTGATACCAGCGCTATCAAGGAAATGCTGAAACGTAACTTGGTGCCGGTTATTCAT
GGCGACATCGTTATCGATGATAAAAATGGGTACCGCATTATTAGCGGTGATGATATCGTGC
CATATCTGGCCAATGAGTTGAAAGCCGACCTCATTTTGTATGCTACTGATGTTGATGGTGT
GCTTATCGACAACAAACCCATCAAGCGCATCGACAAAAATAACATCTATAAAATCCTGAACT
ACCTGAGCGGCAGCAACTCGATTGACGTTACTGGAGGCATGAAATATAAAATCGATATGAT
TCGTAAAAATAAATGTCGCGGTTTCGTGTTTAACGGGAATAAGGCGAACAACATCTATAAAG
CGCTTTTAGGGGAAGTGGAAGGTACCGAGATCGACTTTAGCGAA

Taipk

ATGATGATTCTTAAGATAGGGGGCAGCGTGATTACGGATAAATCGGCCTACCGTACCGCTC
GTACATACGCAATTCGCTCGATTGTAAAAGTGCTTAGCGGGATTGAGGATTTAGTGTGCGT
TGTTCATGGTGGTGGTTCCTTTGGTCATATCAAAGCTATGGAGTTTGGCTTGCCGGGCCCT
AAAAACCCCCGCTCAAGTATTGGTTATTCCATCGTTCATCGCGATATGGAGAACCTGGATC
TAATGGTCATAGACGCCATGATTGAGATGGGGATGCGCCCTATCAGTGTTCCAATTTCGGC
CTTACGCTACGACGGCCGTTTTGACTATACCCCTCTCATTCGGTATATCGATGCAGGCTTT
GTACCTGTTTCTTATGGTGACGTTTATATTAAAGATGAGCATAGTTACGGTATTTACAGTGG
AGATGATATTATGGCCGACATGGCTGAACTTCTCAAACCAGATGTTGCAGTATTTCTTACCG
ACGTGGATGGCATCTACAGCAAAGATCCGAAACGCAATCCGGATGCCGTATTGCTCCGGG
ATATTGACACCAACATTACATTTGATCGCGTCCAAAACGATGTCACAGGTGGTATCGGAAA
AAAATTTGAAAGTATGGTGAAAATGAAAAGCAGTGTGAAGAACGGGGTCTACCTTATTAATG
GCAACCATCCGGAACGTATAGGCGACATCGGTAAGGAGAGCTTCATCGGAACCGTCATTC
GT

Scmk

ATGTCTCTGCCATTCCTGACGTCTGCGCCAGGTAAGGTGATCATCTTCGGCGAGCACTCTG
CGGTGTACAATAAGCCGGCCGTCGCCGCCTCTGTGTCTGCGTTACGCACCTACCTGCTGA
TCAGCGAATCTTCTGCACCGGACACGATCGAGCTGGACTTTCCGGACATCAGCTTCAACCA
CAAGTGGAGCATCAACGACTTCAACGCGATCACGGAGGACCAGGTGAACAGCCAAAAGCT
GGCCAAAGCCCAGCAAGCAACCGACGGTCTGTCTCAGGAGCTGGTGTCTCTGCTGGACC
CGCTGTTAGCGCAGTTAAGCGAGAGCTTCCATTACCACGCCGCGTTCTGCTTCCTGTACAT
GTTCGTTTGCCTGTGCCCGCACGCAAAGAACATCAAGTTCAGCCTGAAGAGCACGCTGCC
GATTGGCGCAGGCTTAGGCTCTAGCGCATCTATCAGCGTGAGCCTGGCGCTGGCGATGG
CCTATCTGGGTGGCCTGATTGGCAGCAACGACCTGGAGAAACTGAGCGAAAACGACAAGC
ACATCGTGAACCAGTGGGCCTTTATCGGCGAGAAGTGCATTCATGGCACCCCGAGCGGCA
TTGACAACGCAGTTGCCACGTATGGCAACGCCCTGCTGTTCGAGAAAGACAGCCACAACG
GCACGATCAACACGAACAACTTCAAGTTCCTGGACGACTTCCCGGCGATCCCGATGATTCT
GACCTACACCCGTATCCCACGCAGCACCAAGGATTTAGTCGCCCGCGTGCGTGTTTTAGT

CACCGAAAAGTTCCCGGAGGTGATGAAGCCGATCCTGGACGCGATGGGCGAGTGCGCGC
TGCAGGGTCTGGAGATCATGACCAAGCTGAGCAAGTGCAAGGGCACCGACGATGAGGCG
GTGGAGACCAACAATGAGCTGTACGAGCAGCTGCTGGAGCTGATCCGTATCAATCACGGC
CTGCTGGTCTCTATCGGTGTGTCTCACCCGGGCCTGGAACTGATCAAAAACCTGAGCGAC
GACCTGCGCATTGGCTCTACGAAATTAACGGGTGCAGGTGGCGGTGGCTGCTCTTTAACG
CTGCTGCGCCGTGACATTACGCAGGAGCAAATCGACAGCTTCAAGAAGAAGCTGCAGGAC
GACTTCAGCTACGAGACGTTCGAGACGGACCTGGGCGGCACGGGCTGTTGCCTGCTGAG
CGCCAAAAATCTGAACAAGGACCTGAAGATCAAAAGCCTGGTGTTCCAGCTGTTCGAAAAC
AAGACGACCACGAAGCAGCAGATCGACGACCTGTTACTGCCGGGTAACACCAATCTGCCG
TGGACGTCT

Atipk

ATGGAACTCAATATCAGCGAAAGCCGGTCGCGCAGCATCCGGTGCATCGTGAAGCTCGGG
GGCGCGGCCATCACCTGCAAAAACGAACTCGAAAAGATCCATGACGAAAACCTCGAAGTG
GTGGCCTGCCAACTGCGGCAAGCGATGCTGGAAGGCTCCGCCCCCTCCAAAGTCATCGG
CATGGACTGGTCCAAACGGCCGGGCTCCTCCGAAATCTCCTGCGATGTGGACGACATCGG
CGACCAGAAATCGAGCGAATTCTCGAAGTTCGTGGTCGTCCACGGGGCGGGCTCGTTCG
GCCATTTCCAAGCGTCGCGGTCGGGCGTCCATAAAGGGGGCCTGGAGAAGCCCATCGTC
AAAGCCGGGTTCGTCGCGACCCGGATCTCGGTCACCAATCTCAATCTCGAGATCGTCCGC
GCCCTGGCCCGGGAAGGCATCCCGACGATCGGGATGAGCCCCTTCAGCTGCGGCTGGTC
CACCTCCAAACGGGATGTCGCGTCGGCGGATCTCGCGACGGTCGCCAAGACCATCGACT
CGGGCTTCGTGCCCGTGCTCCATGGGGACGCGGTCCTGGACAACATCCTGGGCTGCACG
ATCCTGTCCGGCGATGTCATCATCCGGCACCTGGCCGACCATCTCAAGCCGGAATACGTC
GTGTTCCTGACCGACGTGCTCGGGGTCTATGACCGGCCCCCGAGCCCGTCGGAGCCGGA
CGCCGTCCTGCTCAAGGAGATCGCCGTCGGGGAGGATGGCTCCTGGAAGGTGGTCAACC
CCCTCCTGGAACATACCGACAAAAAGGTGGATTATTCCGTCGCGGCCCACGATACCACGG
GGGGGATGGAAACCAAAATCTCGGAAGCCGCCATGATCGCCAAGCTCGGCGTGGATGTG
TACATCGTGAAAGCCGCGACCACCCATTCGCAGCGGGCGCTCAATGGCGACCTGCGCGA
CTCCGTCCCCGAAGACTGGCTGGGGACCATCATCCGCTTCAGCAAATAA

Scck

ATGGTGCAGGAGTCCCGCCCCGGCTCGGTCCGGTCGTATTCCGTGGGCTACCAGGCCCG
GTCGCGGTCGTCGTCCCAGCGCCGCCATTCGCTCACGCGGCAGCGCAGCAGCCAGCGG
CTCATCCGGACGATCTCCATCGAGAGCGATGTGAGCAATATCACGGACGATGATGATCTG
CGGGCGGTGAATGAAGGGGTGGCCGGGGTCCAGCTCGACGTCTCCGAGACGGCGAACA
AAGGGCCaCGCCGGGCCAGtGCCACCGATGTCACCGACTCGCTGGGCTCCACGTCCAGC
GAATATATCGAGATCCCCTTCGTGAAAGAGACGCTGGACGCGAGCCTCCCCTCGGATTAC
CTCAAACAAGACATCCTGAACCTGATCCAATCCCTGAAGATCTCGAAATGGTACAATAACAA
AAAGATCCAGCCCGTCGCCCAGGACATGAACCTCGTCAAAATCTCCGGCGCGATGACCAA
TGCGATCTTCAAGGTGGAGTACCCGAAACTGCCGTCCCTCCTGCTGCGGATCTATGGCCC
GAATATCGATAACATCATCGACCGCGAATATGAACTCCAGATCCTCGCGCGGCTCTCGCTG
AAAAACATCGGGCCGTCCCTGTACGGCTGCTTCGTGAATGGGCGCTTCGAGCAGTTCCTC
GAAAACTCCAAAACGCTGACCAAGGATGATATCCGGAACTGGAAAAACTCGCAACGGATC
GCCCGCCGCATGAAGGAGCTGCATGTGGGCGTGCCCCTCCTCTCGTCGGAGCGGAAGAA

TGGGAGCGCCTGCTGGCAAAAAATCAACCAATGGCTCCGCACGATCGAGAAGGTGGATCA
GTGGGTCGGGGACCCGAAGAACATCGAGAACAGCCTCCTCTGCGAAAATTGGTCCAAATT
CATGGACATCGTCGATCGGTACCACAAGTGGCTGATCAGCCAAGAACAAGGGATCGAGCA
AGTCAACAAAAATCTGATCTTCTGCCATAATGATGCCCAATACGGGAATCTCCTCTTCACCG
CGCCCGTCATGAACACCCCCTCCCTGTATACCGCGCCGAGCTCGACCTCCCTGACGTCCC
AAAGCAGCAGCCTCTTCCCCTCGTCCAGCAACGTGATCGTCGATGATATCATCAATCCCCC
GAAGCAAGAACAATCCCAAGATTCCAAACTCGTGGTCATCGATTTCGAATACGCCGGGGC
CAATCCCGCCGCGTACGATCTCGCCAATCACCTCTCGGAATGGATGTACGACTATAATAAC
GCCAAAGCCCCGCACCAGTGCCACGCCGACCGGTACCCCGACAAGGAGCAAGTGCTCAA
CTTCCTGTATTCGTATGTCAGCCATCTCCGCGGCGGGGCCAAAGAGCCCATCGATGAAGA
AGTCCAGCGCCTCTATAAATCGATCATCCAGTGGCGCCCCACGGTGCAGCTCTTCTGGTC
GCTGTGGGCGATCCTGCAAAGCGGCAAGCTGGAAAAAAAAGAAGCCAGCACCGCCATCAC
CCGCGAAGAAATCGGGCCCAATGGGAAAAAGTATATCATCAAGACGGAGCCCGAGTCGCC
CGAAGAGGACTTCGTCGAAAATGACGACGAACCCGAAGCCGGCGTGTCGATCGATACCTT
CGACTACATGGCCTACGGGCGGGACAAGATCGCGGTGTTCTGGGGGGGACCTGATCGGGC
TGGGCATCATCACGGAGGAGGAATGCAAGAACTTCTCGAGCTTCAAATTCCTCGACACCA
GCTACCTGTAA

Table A6: DOE runs to optimize taxadiene synthesis via the IUP

Table A6-i: Run 1

Run #	Temp. (oC)	IPTG conc. (mM)	OD600 at induction	mM isoprenol	IUPi promoter	Taxadiene (µg/gdcw)
1	22	0.05	0.4	10	pro5	21217
2	30	0.05	0.4	10	pro4	1622
3	22	0.2	0.4	10	pro4	20941
4	30	0.2	0.4	10	pro5	3867
5	22	0.05	0.7	10	pro4	18391
6	30	0.05	0.9	10	pro5	5861
7	22	0.2	0.7	10	pro5	17178
8	30	0.2	0.9	10	pro4	3677
9	22	0.05	0.4	30	pro4	19628
10	30	0.05	0.4	30	pro5	7855
11	22	0.2	0.4	30	pro5	20089
12	30	0.2	0.4	30	pro4	5483
13	22	0.05	0.7	30	pro5	37629
14	30	0.05	0.9	30	pro4	4910
15	22	0.2	0.7	30	pro4	21610
16	30	0.2	0.9	30	pro5	5954
17	26	0.125	0.6	20	pro4	18601
18	26	0.125	0.6	20	pro5	23366
19	26	0.125	0.6	20	pro4	14702
20	26	0.125	0.6	20	pro5	22031
21	26	0.125	0.6	20	pro4	18284
22	26	0.125	0.6	20	pro5	22675

Table A6-ii: Run 2

Run #	Temp. (oC)	IPTG conc. (mM)	OD600 at induction	mM isoprenol	aTC (ng/mL)	OD at pTET induction	Taxadiene (µg/gdcw)
1	22	0.05	0.421	10	10	0.421	17892
2	30	0.05	0.421	10	10	0.81	13221
3	22	0.2	0.421	10	10	0.421	15463
4	30	0.2	0.421	10	10	0.81	13536
5	22	0.05	0.82	10	10	0.421	9546
6	30	0.05	0.82	10	10	0.81	14680
7	22	0.2	0.82	10	10	0.421	16878
8	30	0.2	0.82	10	10	0.81	13647
9	22	0.05	0.421	30	10	0.421	13525
10	30	0.05	0.421	30	10	0.81	4689
11	22	0.2	0.421	30	10	0.421	6632
12	30	0.2	0.421	30	10	0.81	2077
13	22	0.05	0.82	30	10	0.421	16854
14	30	0.05	0.82	30	10	0.81	11554
15	22	0.2	0.82	30	10	0.421	15203
16	30	0.2	0.82	30	10	0.81	4747
17	22	0.05	0.421	10	20	0.421	13308
18	30	0.05	0.421	10	20	0.81	13718
19	22	0.2	0.421	10	20	0.421	17891
20	30	0.2	0.421	10	20	0.81	12454
21	22	0.05	0.82	10	20	0.421	12897
22	30	0.05	0.82	10	20	0.81	11654
23	22	0.2	0.82	10	20	0.421	11971
24	30	0.2	0.82	10	20	0.81	14430
25	22	0.05	0.421	30	20	0.421	5276
26	30	0.05	0.421	30	20	0.81	6835
27	22	0.2	0.421	30	20	0.421	18183
28	30	0.2	0.421	30	20	0.81	1220
29	22	0.05	0.82	30	20	0.421	2672
30	30	0.05	0.82	30	20	0.81	1310
31	22	0.2	0.82	30	20	0.421	11404
32	30	0.2	0.82	30	20	0.81	1474
33	26.7	0.125	0.594	20	15	0.594	18245
34	26.7	0.125	0.594	20	15	0.594	13701
35	26.7	0.125	0.594	20	15	0.594	12620

Table A6-iii: Run 3

Run #	C-source (g/L)	mM isoprenol	C-source	N-source	Taxadiene (µg/gdcw)
1	7.5	5.2	Glucose	CSM 1%	3725
2	17.5	5.2	Glucose	CSM 1%	7608
3	7.5	20	Glucose	CSM 1%	5122
4	17.5	20	Glucose	CSM 1%	3061
5	5.428932	12.6	Glucose	CSM 1%	4550
6	19.57107	12.6	Glucose	CSM 1%	4145
7	12.5	2.13482	Glucose	CSM 1%	3704
8	12.5	23.06518	Glucose	CSM 1%	4611
9	12.5	12.6	Glucose	CSM 1%	4433
10	12.5	12.6	Glucose	CSM 1%	3743
11	12.5	12.6	Glucose	CSM 1%	1669
12	12.5	12.6	Glucose	CSM 1%	3345
13	12.5	12.6	Glucose	CSM 1%	3771
14	7.5	5.2	Glycerol	CSM 1%	2936
15	17.5	5.2	Glycerol	CSM 1%	3036
16	7.5	20	Glycerol	CSM 1%	4536
17	17.5	20	Glycerol	CSM 1%	3894
18	5.428932	12.6	Glycerol	CSM 1%	2389
19	19.57107	12.6	Glycerol	CSM 1%	1808
20	12.5	2.13482	Glycerol	CSM 1%	3016
21	12.5	23.06518	Glycerol	CSM 1%	2100
22	12.5	12.6	Glycerol	CSM 1%	3777
23	12.5	12.6	Glycerol	CSM 1%	2928
24	12.5	12.6	Glycerol	CSM 1%	4350
25	12.5	12.6	Glycerol	CSM 1%	5511
26	12.5	12.6	Glycerol	CSM 1%	5026
27	7.5	5.2	Glucose	YE 1%	3446
28	17.5	5.2	Glucose	YE 1%	2966
29	7.5	20	Glucose	YE 1%	1588
30	17.5	20	Glucose	YE 1%	2262
31	5.428932	12.6	Glucose	YE 1%	2425
32	19.57107	12.6	Glucose	YE 1%	1984
33	12.5	2.13482	Glucose	YE 1%	1727
34	12.5	23.06518	Glucose	YE 1%	2958
35	12.5	12.6	Glucose	YE 1%	3131
36	12.5	12.6	Glucose	YE 1%	3307
37	12.5	12.6	Glucose	YE 1%	3097
38	12.5	12.6	Glucose	YE 1%	4100
39	12.5	12.6	Glucose	YE 1%	1855
40	7.5	5.2	Glycerol	YE 1%	2182
41	17.5	5.2	Glycerol	YE 1%	2827
42	7.5	20	Glycerol	YE 1%	3114
43	17.5	20	Glycerol	YE 1%	4839
44	5.428932	12.6	Glycerol	YE 1%	3294
45	19.57107	12.6	Glycerol	YE 1%	2615
46	12.5	2.13482	Glycerol	YE 1%	3129
47	12.5	23.06518	Glycerol	YE 1%	4047
48	12.5	12.6	Glycerol	YE 1%	4387
49	12.5	12.6	Glycerol	YE 1%	3340
50	12.5	12.6	Glycerol	YE 1%	1532
51	12.5	12.6	Glycerol	YE 1%	5028
52	12.5	12.6	Glycerol	YE 1%	1052

Table A6-iv: N-source supplementation optimization

Run #	C-source (g/L)	mM isoprenol	C-source	N-source	Taxadiene (µg/gdcw)
1	20	10	Glucose	TRP 1%	6789
2	20	10	Glucose	TRP 1%	10063
3	20	10	Glucose	TRP 1%	8861
4	20	10	Glucose	CSM 1%	6207
5	20	10	Glucose	CSM 1%	1844
6	20	10	Glucose	CSM 1%	4721
7	20	10	Glucose	YE 1%	2408
8	20	10	Glucose	YE 1%	3594
9	20	10	Glucose	YE 1%	3867

A.R. References

1. Ferrières L, et al. (2010) Silent mischief: Bacteriophage Mu insertions contaminate products of Escherichia coli random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J Bacteriol* 192(24):6418–6427.
2. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* 330(2010):70–74.
3. Baba T, et al. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* 2. doi:10.1038/msb4100050.
4. Hirota Y, Suzuki H, Nishimura Y, Yasuda S (1977) On the process of cellular division in Escherichia coli: a mutant of E. coli lacking a murein-lipoprotein. *Proc Natl Acad Sci U S A* 74(4):1417–1420.
5. Sun TP, Webster RE (1986) fii, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant tolA and tolB. *J Bacteriol* 165(1):107–115.
6. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol*

21(7):796–802.

7. Alonso-Gutierrez J, et al. (2013) Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production. *Metab Eng* 19:33–41.

8. Cunningham FX, Lee H, Gantt E (2007) Carotenoid biosynthesis in the primitive red alga Cyanidioschyzon merolae. *Eukaryot Cell* 6(3):533–545.

9. Silva-Rocha R, et al. (2013) The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* 41(D1):666–675.

10. Lee TS, et al. (2011) BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J Biol Eng* 5(1):12.

11. Jiang Y, et al. (2015) Multigene editing in the Escherichia coli genome using the CRISPR-Cas9 system. *Appl Environ Microbiol* 81(7):2506–2514.

12. Zhou K, Qiao K, Edgar S, Stephanopoulos G (2015) Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat Biotechnol* 33(4):377–383.

13. Santos CNS, Koffas M, Stephanopoulos G (2011) Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metab Eng* 13(4):392–400.