Biosynthetic Production of

Aromatic Fine Chemicals

by

Shawn Pugh

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2016 by the
Graduate Supervisory Committee:

David Nielsen, Chair
Cesar Torres
Lenore Dai
Mary Laura Lind
Xuan Wang

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

This dissertation focuses on the biosynthetic production of aromatic fine chemicals in engineered *Escherichia coli* from renewable resources. The discussed metabolic pathways take advantage of key metabolites in the shikimic acid pathway, which is responsible for the production of the aromatic amino acids phenylalanine, tyrosine, and tryptophan. For the first time, the renewable production of benzaldehyde and benzyl alcohol has been achieved in recombinant *E. coli* with a maximum titer of 114 mg/L of benzyl alcohol. Further strain development to knockout endogenous alcohol dehydrogenase has reduced the in vivo degradation of benzaldehyde by 9-fold, representing an improved host for the future production of benzaldehyde as a sole product. In addition, a novel alternative pathway for the production of protocatechuate (PCA) and catechol from the endogenous metabolite chorismate is demonstrated. Titers for PCA and catechol were achieved at 454 mg/L and 630 mg/L, respectively. To explore potential routes for improved aromatic product yields, an *in silico* model using elementary mode analysis was developed. From the model, stoichiometric optimums maximizing both product-to-substrate and biomass-to-substrate yields were discovered in a co-fed model using glycerol and D-xylose as the carbon substrates for the biosynthetic production of catechol. Overall, the work presented in this dissertation highlights contributions to the field of metabolic engineering through novel pathway design for the biosynthesis of industrially relevant aromatic fine chemicals and the use of *in silico* modelling to identify novel approaches to increasing aromatic product yields.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Dr. David Nielsen for his ongoing support and guidance. I remember my first day at ASU, he gave a research presentation to the incoming graduate students and in that moment I knew that I wanted to work in his lab. To this day I consider it be one of the best decisions I have made. I have developed a passion for metabolic engineering and truly love the research I do, all of which would not have been possible without him. Also, I would like to thank my committee members Dr. Lenore Dai, Dr. Xuan Wang, Dr. Mary Laura Lind, and Dr. Cesar Torres for taking time out of their schedule to serve as a member of my committee. Finally, I would like to thank all my friends and labmates who have made everyday a pleasure to walk into the lab, including Becky McKenna, Luis Moya, Ploy Phusitkanchana, Michael Wiehn, Ibrahim Halloum, Vick Syradi, Thomas Levario, and Brian Thompson. Thank you again to everyone mentioned, I could not have done this without you!

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Abstract

This chapter provides a brief overview of metabolic engineering and the motivation for biosynthetically producing aromatic fine chemicals. The metabolic pathways to renewably synthesize benzaldehyde, benzyl alcohol, para-hydroxybenzoate, protocatechuate, and catechol via metabolites of the shikimic acid pathway are introduced. Finally, Chapter 1 concludes with the dissertation organization.

1.1 Background and motivation

Over the last several decades, metabolic engineering has proven to be a powerful means of producing a large number of specialty chemicals using readily available renewable resources as the starting material in genetically engineered microorganisms. Thanks to advancements in parallel fields, such as high throughput genomic sequencing and protein engineering, our ability to identify and characterize pathways, whose natural products are of industrial significance, has evolved. To date, numerous commodity chemicals including biofuels, such as ethanol(Ingram et al. 1998, Woodruff, Boyle, and Gill 2013, Clomburg and Gonzalez 2010), butanol(Lan and Liao 2011, Dusséaux et al. 2013, Nielsen et al. 2009), and long chain fatty acids(Zhang, Rodriguez, and Keasling 2011, Lu 2010, Radakovits et al. 2010, Work et al. 2012), fine chemicals, such as plastic precursors like styrene(McKenna and Nielsen 2011b), hydroxystyrene(Qi et al. 2007b, Verhoef et al. 2009), and isoprene(Erickson, Nelson, and Winters 2012), as well as pharmaceutical precursors, like artemisinic acid(Ro et al. 2006) and (*R*)-1,2-phenylethanediol(McKenna et al. 2013b), have been synthesized renewably in recombinant microbes.

This impressive feat has been achieved by transferring enzymes, or pathways, with desirable biocatalytic properties from one (or more) organisms, via plasmid-based recombinant DNA technology, to hosts that may be readily engineered. However, redirecting a microbial host's natural metabolism to overproduce non-natural products comes with a price, as hosts are often ill-equipped to handle the imposed metabolic burden(Jones, Kim, and Keasling 2000). Genetically modifying a microbial host inflicts a considerable degree of stress by forcing it to 1. harbor and stably maintain foreign DNA,

2. redirect key metabolic regulatory machinery involved in transcription and translation, and 3. overcome the presence of antibiotics and toxic pathway byproducts. As a result of the metabolic burden associated with the overproduction of proteins, engineered hosts most commonly suffer from reduced growth rates, lower biomass yields, and inclusion body (protein aggregate) formation. In addition, while natural systems have evolved to control pathway flux, recombinant pathways often suffer from flux imbalances which limit achievable titers and yields while imposing unnecessary burden on the host(Dueber et al. 2009). Thus, the field of metabolic engineering has been working diligently to apply advancements in synthetic biology to overcome bottlenecks, reduce burden, and achieve higher titers of biosynthetically produced commodity chemicals(Keasling 1999, Dueber et al. 2009)

1.2 Motivation for renewably synthesizing aromatic fine chemicals

While most of the attention in the field of metabolic engineering has been on the biosynthetic production of transportation fuels, in response to the growing concern of the depletion of oil reserves, what is often overlooked is the numerous fine chemicals which are also produced from petroleum. In particular, nearly all aromatic compounds are derived, in one fashion or another, from petroleum derived BTX (benzene, toluene, and the three isomers of xylene) aromatics(Radwan et al. 1997, Schobert and Song 2002). BTX aromatics are used as precursors for the synthetic production of numerous fine chemicals including phenol (used to make phenolic resins)(Schobert and Song 2002, Schmidt 2005), styrene (used to make plastics)(McKenna and Nielsen 2011b, Miller, Newhook, and Poole 1994), ethylbenzene (a precursor to styrene)(Miller, Newhook, and Poole 1994), cumene (an aromatic building-block)(Schmidt 2005), and terephthalate (a

component of PET plastics)(Sako et al. 2000), to name a few. Once petroleum resources are depleted, however, alternative and renewable routes for the production of these essential commodity chemicals will be a necessity. Metabolic engineering and the engineering of microbes via recombinant DNA technology offers one such approach to renewably synthesizing chemicals of industrial significance(Adkins et al. 2012).

To date, numerous BTX derivatives have been biosynthetically produced in engineered microbes from renewable resources, including styrene(McKenna and Nielsen 2011b), hydroxystyrene(Qi et al. 2007b), phenol(Wierckx et al. 2005b), and vanillin(Kim, Kim, and Lee 2008), to name a few. This was accomplished via the *de novo* construction of heterologous pathways in engineered microorganisms taking advantage of endogenous aromatic metabolites. To date however, neither benzene, toluene, or any isomer of xylene has been produced renewably; therefore, researchers must continue to discover new routes to biosynthetically produce their derivatives. Thus, the focus of this work will be to construct and characterize novel pathways for the renewable synthesis of aromatic fine chemicals, in particular: the flavor compounds benzaldehyde and benzyl alcohol, and the phenolic antioxidants para-hydroxybenzoate, protocatechuate and catechol.

1.3 Metabolic Pathways

To renewably synthesize aromatic fine chemicals in engineered *E. coli*, the shikimic acid pathway offers a valuable source of aromatic precursors to which engineered pathways may be linked to central metabolism. The shikimic acid pathway is a tightly regulated pathway responsible for the production of the proteinogenic aromatic amino acids phenylalanine, tyrosine and tyrptophan(Keseler et al. 2005). As illustrated in

Fig. 1.1, the pathway begins by the condensation of phosphoenolpyruvate (PEP) and D-erythrose-4-phosphate (E4P) to form 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) via the expression of DAHP synthase. *E. coli* possess three isozymes of DAHP synthase including *aroF*, *aroG*, and *aroH* which are regulated both transcriptionally and allosterically in the presence of tyrosine, phenylalanine, and tryptophan, respectively. Transcriptional regulation is mediated by the DNA-binding transcriptional repressors TyrR (tyrosine and phenylalanine) and TrpR (tryptophan), which bind to chromosomal DNA at the transcription initiation site in the presence of the respective amino acids. In addition to tight regulation of carbon entering the shikimic acid pathway, consumption of the key metabolite chorismate, at which point the pathways for the three aromatic amino branches off, is also tightly regulated. Here, chorismate may either be directed toward tryptophan via the enzyme TrpE (anthranilate synthase component I), phenylalanine via the enzyme PheA (phenylpyruvate forming bi-functional chorismate mutase/prephenate dehydratase), or tyrosine via the enzyme TyrA (4-hydroxyphenylpyruvate forming bi-functional chorismate mutase/prephenate dehydratase)(Keseler et al. 2005). Similar to the DAHP synthase isozymes, expression of TyrA and TrpE is controlled both transcriptionally (mediated by TyrR and TrpR, respectively) and allosterically; meanwhile, PheA is regulated only via allosteric enzyme inhibition(Keseler et al. 2005). To achieve high flux through the shikimic acid pathway and thus achieve high titers of biosynthetically produced aromatic fine chemicals, a suitable host in which feedback regulation has been relieved is a necessity.

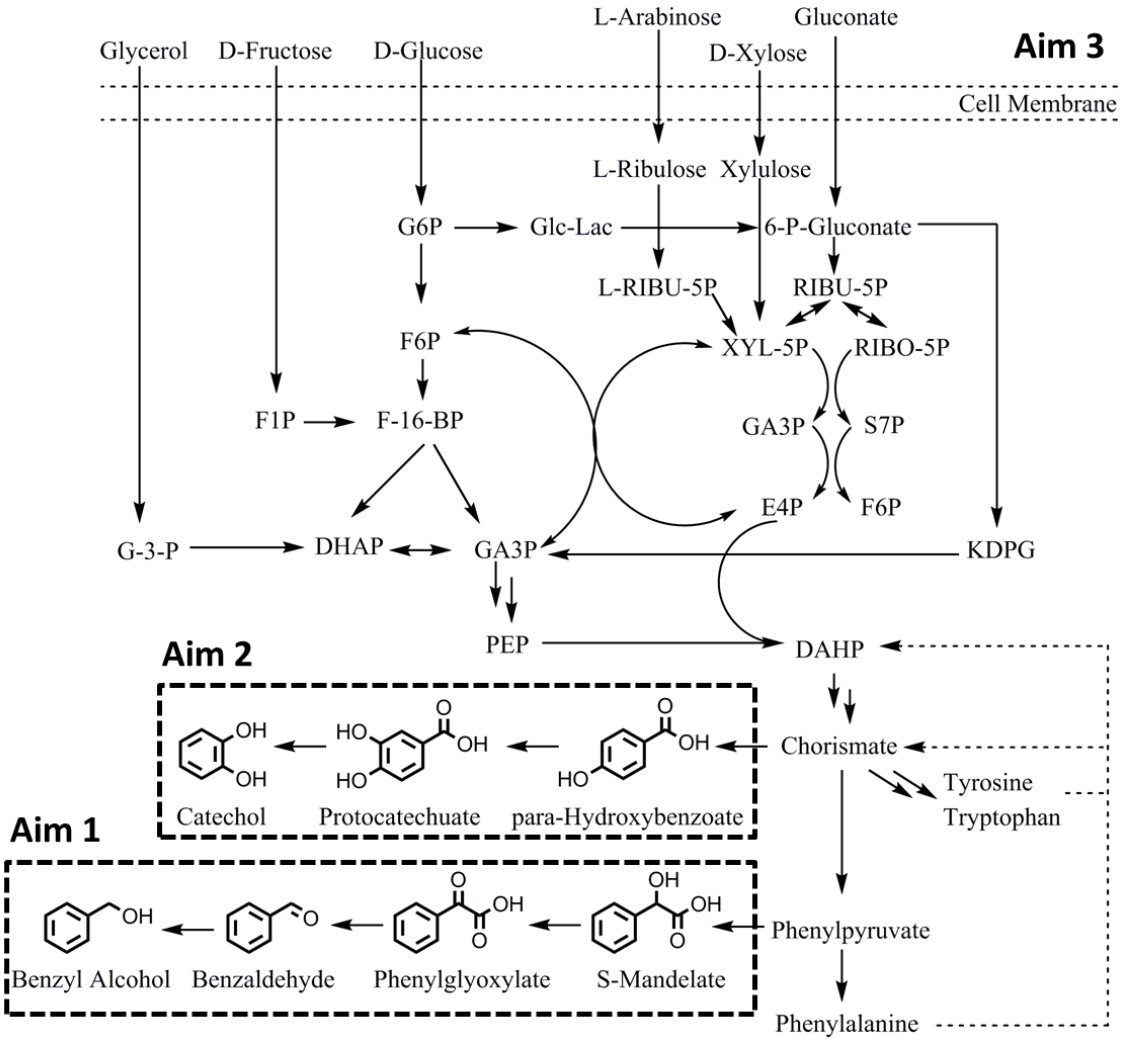

Figure 1.1 Biosynthetic routes for the production of aromatic fine chemicals. Multiple arrows indicate that several steps are occurring but not illustrated. Dashed arrows indicate that the metabolite demonstrates feedback repression via allosteric enzyme inhibition, transcriptional repression, or both.

To overcome the feedback regulation imposed on the shikimic acid pathway, a feedback resistant mutant, *E. coli* NST74 ATCC 31884, has previously been engineered to overproduce phenylalanine(Tribe 1987b). As a result of mutagenesis, mediated via the chemomutagen NTG (N-methyl-N'-nitro-N-nitrosoguanidine) and high throughput selection on phenylalanine anti-metabolite plates containing *p*-fluoro-DL-phenylalanine, NST74 possesses mutations which relieve the allosteric enzyme inhibition of AroF,

AroG, AroH, and PheA(Tribe 1987b). In addition, the phenylalanine operator PheO and the DNA-binding transcriptional repressor TyrR were also mutated to promote the production of phenylalanine. In controlled bioreactor experiments, NST74 has been shown to achieve titers as high as 1.98 g/L of phenylalanine(Tribe 1987b). With such a high flux through the shikimic acid pathway, NST74 represents an ideal host platform for further genetic engineering to biosynthetically produce aromatic fine chemicals(McKenna and Nielsen 2011b).

In addition to mutating the feedback resistant enzymes (AroF, AroG, AroH, and PheA) of the shikimic acid pathway, various other strategies have been employed to increase flux. As previously described, carbon enters the shikimic acid pathway via the condensation of PEP and E4P(Keseler et al. 2005). As a product of glycolysis and the immediate precursor of the TCA cycle, PEP is a readily available endogenous metabolite and it is believed that flux entering the pathway is limited by the bioavailability of E4P. E4P is the final product of the pentose phosphate pathway (which is responsible for purine and pyrimidine biosynthesis) and is readily incorporated into glycolysis via the reversible transketolase reaction, expressed by *tktAB* in *E. coli*(Keseler et al. 2005). Overexpression of *tktA* in *E. coli* has been shown to redirect flux from glycolysis into the pentose phosphate pathway yielding higher tiers of E4P and thus higher flux through the shikimic acid pathway(Draths et al. 1992, Curran et al. 2013b). A similar approach, while crude, involves knocking out the pentose phosphate by deletion of glucose-6-phosphate dehydrogenase (expressed in *E. coli* as *zwf*) and forcing flux to be redirected via the reversible transketolase reaction. When combined, overexpression of transketolase and deletion of the pentose phosphate pathway have been shown to work synergistically to

improve the bioavailability of E4P, as reported in the yeast *Saccharomyces cerevisiae*(Curran et al. 2013b).

As summarized by Fig. 1.1, the shikimic acid pathway possesses a plethora of endogenous precursors for the biosynthetic production of aromatic fine chemicals. To start, the metabolite chorismate, whose consumption is naturally regulated by *E. coli*, offers a means to achieve the biosynthetic production of protocatechuate (PCA) and catechol. This is achieved by first converting chorismate to the phenolic acid para-hydroxybenzoate (pHBA). pHBA may then be hydroxylated to yield PCA, which may subsequently be decarboxylated to produce the phenolic diol catechol. Previous reports have demonstrated that PCA and catechol may be biosynthetically derived from the metabolite dehydroshikimate (DHS)(Curran et al. 2013b, Draths and Frost 1994); however, we propose this alternative pathway may represent a more suitable platform. Downstream of the shikimic acid pathway, the intermediate metabolite phenylpyruvate (the immediate precursor to phenylalanine) also represents a platform for the biosynthetic production of the aromatic flavor and fragrance compounds benzaldehyde and benzyl alcohol. This is achieved by first converting phenylpyruvate to (*S*)-mandelate. Subsequently, (*S*)-mandelate is then converted to phenylglyoxylate which may then be decarboxylated to yield benzaldehyde. *E. coli*, however, has been found to possess numerous alcohol dehydrogenases with activity on benzaldehyde, and thus has the innate ability to convert benzaldehyde to benzyl alcohol. The pathways mentioned above, and illustrated in Fig. 1.1, will be described in greater detail in the proceeding chapters.

1.4 Dissertation organization

This dissertation is organized into five chapters. Chapter 1 provides a brief discussion on the motivation of the research and recent advancements in the field of metabolic engineering. This chapter also provides an outline of the metabolic pathways to achieve the biosynthetic production of the monoaromatic fine chemicals of interest from renewable resources in recombinant *E. coli*. In Chapter 2, the biosynthetic production of the fine chemicals benzaldehyde and benzyl alcohol from renewable resources in engineered *E. coli* is demonstrated, for the first time (Figure 1.1, Aim 1). In Chapter 3, a novel biosynthetic route toward the production of para-hydroxybenzoate, protocatechuate, and catechol is discussed (Figure 1.1, Aim 2). In Chapter 4, elementary mode analysis was used to screen alternative carbon substrates to increase the bioavailability of PEP and E4P and ultimately increase flux through the shikimic acid (Figure 1.1, Aim 3). Finally, in Chapter 5, conclusions and future works are discussed.

CHAPTER 2

ENGINEERING *ESCHERICHIA COLI* FOR RENEWABLE BENZYL ALCOHOL PRODUCTION

Abstract

Benzyl alcohol is an aromatic hydrocarbon used as a solvent and an intermediate chemical in the pharmaceutical, cosmetics, and flavor/fragrance industries. The *de novo* biosynthesis of benzyl alcohol directly from renewable glucose was herein explored using a non-natural pathway engineered in *Escherichia coli*. Benzaldehyde was first produced from endogenous phenylpyruvate via three heterologous steps, including hydroxymandelate synthase (encoded by *hmaS*) from *Amycolatopsis orientalis*, followed by (*S*)-mandelate dehydrogenase (encoded by *mdlB*) and phenylglyoxylate decarboxylase (encoded by *mdlC*) from *Pseudomonas putida* ATCC 12633. The subsequent rapid and efficient reduction of benzaldehyde to benzyl alcohol occurred by the combined activity and native regulation of multiple endogenous alcohol dehydrogenases and/or aldo-keto reductases. Through systematic deletion of competing aromatic amino acid biosynthesis pathways to promote endogenous phenylpyruvate availability, final benzyl alcohol titers as high as 114 ± 1 mg/L were realized, representing a yield of 7.6 ± 0.1 mg/g on glucose and a ~5-fold improvement over initial strains.

This work was published as:

Pugh, S., McKenna, R., Halloum, I., & Nielsen, D.R. Engineering Escherichia coli for renewable benzyl alcohol production. *Metab Eng Comm* 2, 39-45 (2015)

2.1 Introduction

Benzyl alcohol is a naturally occurring monoaromatic alcohol with a broad range of commercial applications and a current market price of $2000-2500 USD/ton. With both low volatility and toxicity yet strong polarity, benzyl alcohol is attractive as a safe and effective solvent, particularly for use with polymers and in applications including the production of inks, paints, glues, and hardening products (e.g., epoxy resins) (Stellman 1998, Stoye and Freitag 1998, Ash and Ash 2009). Additionally, while benzyl alcohol itself confers a floral scent, it is more commonly employed as a precursor to synthesize a variety of other ester products with numerous flavor/fragrance uses, including in the manufacture of food products (Fenaroli and Burdock 1995), as well as high value hygiene and cosmetic products. For example, prior reports have found benzyl alcohol to be used in 322 cosmetic formulations belonging to 43 product categories (Nair 2001). Meanwhile, as it renders a bacteriostatic effect at even low concentrations (Marriott 2010), benzyl alcohol is also commonly used as a topical agent and preservative in the pharmaceutical and healthcare industries (Wilson and Martin 1999, Felton 2013, Meinking et al. 2010).

Benzyl alcohol is naturally synthesized by many plants, notably accumulating in edible fruits and tea leaves, as well as in the essential oils of ylang-ylang, jasmine, and hyacinth (Budavari et al. 1989). In such cases, however, benzyl alcohol contents have rarely been found to surpass even ~30 mg/kg (COE 1992) rendering these natural sources as unsuitable for supporting a commodity scale benzyl alcohol bioproduction efforts. Accordingly, conventional production of benzyl alcohol is achieved from petroleum-derived feedstocks. Most commonly this occurs from benzyl chloride (considered a

'probable carcinogen') via alkaline hydrolysis (e.g., with sodium hydroxide) (Yadav and Mehta 1993). In addition to employing energy intensive and harsh reaction conditions, this process suffers from sustainability concerns as it involves the use of non-renewable feedstocks.

As an alternative and more sustainable approach, the *de novo* biosynthesis of benzyl alcohol directly from renewable glucose was herein explored through the systematic engineering of a non-natural biosynthetic pathway engineered in the bacterium *Escherichia coli*. The proposed pathway, which utilizes phenylpyruvate as its immediate endogenous precursor, is illustrated in Figure 2.1. First, phenylpyruvate is converted to (*S*)-mandelate via expression of hydroxymandelate synthase (*hmaS*) from *Amycolatopsis orientalis*. Though its native substrate is 4-hydroxyphenylpyruvate, HmaS has also been shown to display activity on phenylpyruvate (Sun et al. 2011). (*S*)-Mandelate is subsequently converted to benzaldehyde by co-expression of two genes derived from the mandelate degradation pathway of *Pseudomonas putida* ATCC 12633 (Tsou et al. 1990). Specifically, conversion of (*S*)-mandelate to phenylglyoxylate by (*S*)-mandelate dehydrogenase (*mdlB*) followed by decarboxylation of phenylglyoxylate to benzaldehyde by phenylglyoxylate decarboxylase (*mdlC*). The production of benzyl alcohol from benzaldehyde has been reported to occur naturally in *E. coli* as a result of the native function of multiple endogenous alcohol dehydrogenases (ADHs) and/or aldo-keto reductases (AKRs). For example, *E. coli yqhD* has been shown to display substantial activities with respect to the NADPH-dependent reduction of benzaldehyde (Sulzenbacher et al. 2004). Meanwhile, in another recent study it was demonstrated that the native regulation and activity of multiple ADHs/AKRs from *E. coli* (specifically,

yqhD, yjgB, and *yahK*) was sufficient for the rapid and efficient *in vivo* reduction of 2-phenylacetaldehyde to 2-phenylethanol – a structurally similar aromatic substrate-product pair likewise synthesized via a heterologous pathway (Koma et al. 2012b). This study outlines our recent progress towards the systematic engineering of the proposed benzyl alcohol pathway, along with preliminary efforts in host strain engineering to improve initial product titers and yields.

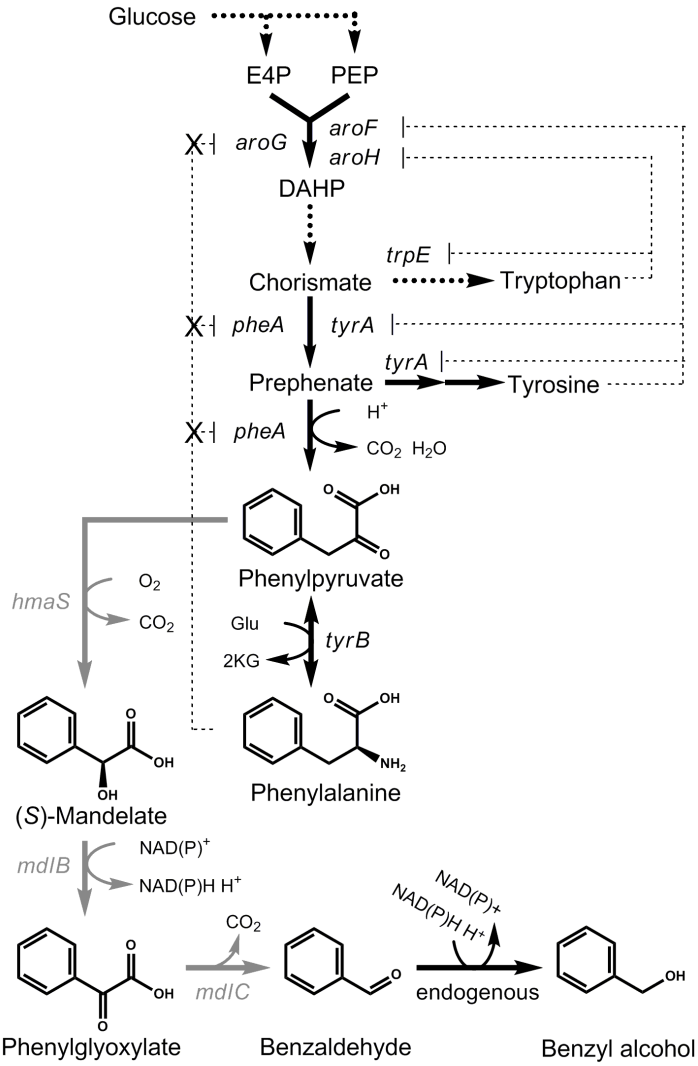

Figure 2.1 Proposed pathway for benzaldehyde and benzyl alcohol biosynthesis from glucose by engineered *E. coli*. Dashed arrows indicate multiple steps. Black and gray arrows indicate native and heterologous pathway steps, respectively. Abbreviations: phosphoenolpyruvate (PEP), D-erythrose-4-phosphate (E4P), 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP).

2.2 Materials and Methods

2.2.1 Bacterial Strains and Media

All strains constructed and used in this study are listed in Table 2.1. *E. coli* NEB10-Beta was obtained from New England Biolabs (NEB, Ipswich, MA) and was used for all cloning work and the plasmid propagation. *E. coli* NST74 (ATCC 31884) – a previously developed, feedback-deregulated phenylalanine overproducer (Tribe 1987a) – was obtained from the American Type Culture Collection (ATCC, Manassas, VA) and used as the initial host platform for aromatics production. *P. putida* ATCC 12633 was also obtained from the ATCC and served as the genetic source of *mdlB* and *mdlC*. *E. coli* JW2581-1, JW4014-2, and JW0911-1 were all obtained from the Coli Genetic Stock Center at Yale University (CGSC, New Haven, CT) and used as the source of genetic materials for the chromosomal deletion of *tyrA*, *tyrB*, and *aspC*, respectively.

E. coli and *Pseudomonas* sp. were routinely cultured in Luria-Bertani (LB) broth supplemented with ampicillin (100 mg/L), chloramphenicol (34 mg/L), and kanamycin (40 mg/L), as appropriate. For aromatics production, strains were cultured in a phosphate-limited minimal media with 15 g/L glucose (herein referred to as "MM1"), previously adapted from that of Qi *et al.* (Qi et al. 2007b) and described by McKenna and Nielsen (McKenna and Nielsen 2011b). To compensate for auxotrophies introduced in mutant strains, MM1 media was supplemented with tyrosine (0.1 g/L) and aspartate (3 g/L), as appropriate.

Table 2.1. Strains and plasmids constructed and/or used in this study.

Strains	Description	Source
E. coli NEB10-Beta	*araD139 Δ(ara,leu)7697 fhuA lacX74 galK16 galE15 mcrA f80d(lacZΔM15)recA1 relA1 endA1 nupG rpsL rph spoT1Δ(mrr-hsdRMS-mcrBC)*	NEB
E. coli NST74	*aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr), pheO352, aroG397(fbr)*	ATCC
P. putida ATCC12633	Source of *mdlB, mdlC, mdlD*	ATCC
E. coli JW2581-1	source of *tyrA*::FRT-Kan-FRT	CGSC
E. coli JW4014-2	source of *tyrB*::FRT-Kan-FRT	CGSC
E. coli JW0911-1	source of *aspC*::FRT-Kan-FRT	CGSC
E. coli NST74A	NST74 Δ*tyrA*::FRT	This study
E. coli NST74AB	NST74 Δ*tyrA*::FRT Δ*tyrB*::FRT	This study
E. coli NST74ABC	NST74 Δ*tyrA*::FRT Δ*tyrB*::FRT Δ*aspC*::FRT	This study
Plasmids	**Description**	**Source**
pTrc99A	P*trc*, pBR322 *ori*, *lacIq*, Amp^R	Prather Lab, MIT
pTrcCOLAK	P*trc*, ColA *ori*, *lacIq*, Kan^R	McKenna *et al.* 2013
pUC57-HmaS	pMB1 *ori*, Amp^R, *hmaS* (codon optimized for *E. coli*)	Genscript
pHmaS	*hmaS* of pUC57-HmaS inserted into the NcoI and EcoRI sites of pTrcCOLAK	This study
pHmaS-MdlC	*mdlC* of *P. putida* ATCC 12633 inserted into the XbaI and HindIII sites of pHmaS with second P*trc* inserted ahead of *mdlC* between the BamHI and XbaI sites	This study
pMdlB	*mdlB* of *P. putida* ATCC 12633 inserted between the NcoI and EcoRI sites of pTrc99A	This study

2.2.2 Plasmid Construction

All plasmids used in this study are also listed in Table 2.1. The plasmid pTrcCOLAK, a fusion of pTrc99A and pCOLADuet-1 (Invitrogen, Carlsbad, CA), was developed in house as previously described (McKenna et al. 2013b). The hydroxymandelate synthase encoding gene, *hmaS* from *A. orientalis*, was synthesized to

include codon optimization for high-level expression in *E. coli* by Genscript (Piscataway, NJ) and harbored in plasmid pUC57-HmaS. To construct pHmaS, codon optimized *hmaS* was PCR amplified from pUC57-HmaS and inserted between the NcoI and EcoRI sites of pTrcCOLAK. Subsequently, *mdlC* was PCR amplified from the gDNA of *P. putida* ATCC 12633 and inserted between the XbaI and HindIII sites of pHmaS. To ensure high-level expression of *mdlC,* a second P_{trc} promoter was inserted ahead of *mdlC* (between the BamHI and XbaI sites), resulting in pHmaS-MdlC. To construct pMdlB, *mdlB* was PCR amplified from *P. putida* ATCC 12633 gDNA and inserted between the NcoI and EcoRI sites of pTrc99A.

Custom DNA oligonucleotide primers were designed and synthesized by Integrated DNA Technologies (Coralville, IA). Genomic DNA (gDNA) was prepared from cultures using the ZR Fungal/Bacterial DNA MiniPrep (Zymo Research, Irvine, CA) according to vendor protocols. All genes were PCR amplified using Phusion High-Fidelity DNA Polymerase (NEB) according to standard protocols. Amplified linear DNA fragments were purified using the DNA Clean & Concentrator kit (Zymo Research) according to manufacturer protocols. Purified DNA was subsequently digested using appropriate restriction endonuclease enzymes (NEB) at 37°C for 3 h. Digested fragments were gel purified using the Zymoclean Gel DNA Recovery kit (Zymo Research) and ligated using T4 DNA ligase (NEB) at 4°C overnight. Ligation reactions were transformed into chemically competent *E. coli* NEB10-Beta (NEB) and selected by plating on LB solid agar containing appropriate antibiotics (as above). Subsequently, transformant pools were screened first by colony PCR (using the same primers as for

cloning) and then by restriction digest mapping. Final plasmid constructs were verified by sequencing.

2.2.3 Construction of E. coli deletion mutants

Chromosomal in-frame gene deletions were accomplished via a protocol adapted from the one-step inactivation method of Datsenko and Wanner (Datsenko and Wanner 2000b). Deletion cassettes for all targeted loci, each of which harbored a kanamycin resistance gene flanked by FLP recognition target sites, were PCR amplified from the gDNA of appropriate Keio collection mutants (Table 2.1) (Baba et al. 2006). In each instance, primer pairs were designed to amplify 300 bp of homology both upstream and downstream of the target gene sequence. Subsequent recombination steps were performed as previously described (Datsenko and Wanner 2000b).

2.2.4 Assaying recombinant pathway function via whole cell biotransformation studies

E. coli NST74 was co-transformed with both pHmaS-MdlC and pMdlB. Single colonies were selected from the transformant pool and seed cultures were grown in 5 mL LB broth with appropriate antibiotics at 37°C while shaking at 200 RPM overnight. Each seed culture was used to inoculate 50 mL LB supplemented with 20 g/L glucose and appropriate antibiotics in a 250 mL baffled shake flasks. Upon reaching an optical density at 600 nm (OD_{600}) of ~0.7, cultures were induced by addition of 0.25 mM IPTG. Culturing continued overnight before cells were harvested by centrifugation at 3,000 x *g*, washed twice with pH7 phosphate buffered saline (PBS) solution, and re-suspended in 50 mL pH7 PBS with 5 g/L glucose and 1 g/L phenylpyruvate. Cultures were incubated at 37°C while levels of each of phenylpyruvate, phenylalanine, (*S*)-mandelate, phenylglyoxylate, benzaldehyde, and benzyl alcohol were subsequently monitored in the

media for the next 7 h by periodic sampling for high performance liquid chromatography (HPLC) analysis. This experiment was repeated in triplicate to provide estimates of standard error.

2.2.5 Benzyl alcohol production directly from glucose by engineered E. coli strains

To test for benzyl alcohol production directly from glucose, each of *E. coli* NST74, NST74A, NST74AB, and NST74ABC were co-transformed with pHmaS-MdlC and pMdlB. Seed cultures were prepared as above and used to inoculate 50 mL MM1 media in a 250 mL baffled shake flask supplemented with appropriate antibiotics and, as needed, required amino acids. Cultures were grown at 37^{o}C while shaking at 200 RPM for 10 h prior to induction by addition of 0.25 mM IPTG. Culturing continued for an additional 96 h, during which time media samples were routinely removed and prepared for analysis by HPLC. All cultures were performed in triplicate to provide estimates of standard error.

2.2.6 Assaying benzyl alcohol toxicity

Cursory estimates of benzyl alcohol toxicity against *E. coli* NST74 were obtained by monitoring for changes in growth rate and yield that occur following its exogenous addition to growing cultures at different concentrations. Cultures were grown in 50 mL MM1 media in 250 mL baffled glass shake flasks at 37^{o}C while shaking at 200 RPM. At an OD_{600} of ~0.5, benzyl alcohol was added to the cultures at final concentrations ranging from 0 to 2 g/L. Growth was then routinely monitored by measurement of OD_{600} for an additional 6 h. All cultures were performed in triplicate to provide estimates of standard error.

2.2.7 Metabolite analysis by HPLC

Samples were prepared by centrifuging 1 mL of culture at 10,000 x *g* for 3 min to pellet and remove cells. Supernatants were transferred to an HPLC vial with a Teflon-lined cap. HPLC analysis was performed using a Hewlett Packard 1100 series HPLC system (Agilent, USA). Metabolites were separated using a reverse-phase Hypersil Gold aQ polar end capped C18 column (4.6 mm x 150 mm; Thermo Fisher, USA) maintained at 45°C and measured with a diode array detector operated at 215 nm (for phenylalanine, (*S*)-mandelate, phenylpyruvate, and benzyl alcohol) and 255 nm (for phenylglyoxylate and benzaldehyde). Samples (5 µL) were injected into a mobile phase with a constant total flow rate of 0.95 mL/min. The mobile phase consisted of 'solvent A' consisting of nanopure water, and 'solvent B' consisting of HPLC-grade methanol (99.8% pure). Beginning as a mixture (vol./vol.) of 95% solvent A and 5% solvent B, a linear gradient was then applied over 8 min until reaching 20% solvent A and 80% solvent B. This condition was then held for 2 min before a second liner gradient was applied over 4 min until reaching 95% solvent A and 5% solvent B. Under these conditions, phenylalanine, phenylglyoxylate, (*S*)-mandelate, phenylpyruvate, benzyl alcohol, and benzaldehyde were eluted at 4.3, 4.6, 5.9, 6.2, 7.1, and 7.8 min, respectively. Standard solutions were developed for each species and used as external calibrations to determine concentrations.

2.2.8 Measurement of biomass growth

Optical density measurements at 600 nm (OD_{600}), performed with a DU800 spectrophotometer (Beckman Coulter, Brea, CA), were used to determine biomass concentration. Dry cell weight (DCW) was then predicted using an established conversion factor (1 OD_{600} = 0.26 g/L) (Guo et al. 2012).

2.3 Results and discussion

2.3.1 Investigating heterologous enzyme and pathway function

Recombinant activity of all candidate enzymes as part of the composite pathway was first assayed by investigating the conversion of exogenously supplied phenylpyruvate by *E. coli* NST74 pHmaS-MdlC pMdlB resting cells. As illustrated in Figure 2.2, the entire 1 g/L of initially added phenylpyruvate was consumed within the first 7 h of the experiment. In this time, phenylglyoxylate first accumulated before then being mostly consumed within ~1.5 h. As phenylglyoxylate was consumed, both benzaldehyde and benzyl alcohol began to accumulate, doing so at similar initial rates. After 1 h, however, benzaldehyde accumulation slowed, reaching a maximum titer of 60 ± 9 mg/L at 3 h before then gradually declining through the remainder of the experiment. Benzyl alcohol finally emerged as the major pathway metabolite, approaching a maximal titer of 203 ± 5 mg/L by 7 h (note: by 24 h, (*S*)-mandelate, phenylglyoxylate, and benzaldehyde were undetected or present only at trace levels while benzyl alcohol titers reached 222 ± 3 mg/L; data not shown). In addition, however, phenylalanine also accumulated, reaching a titer of 610 ± 30 mg/L after 7 h. At this level, the competing

biosynthesis of phenylalanine was responsible for consuming 63% of supplied phenylpyruvate, with only 32% ultimately being converted to benzyl alcohol.

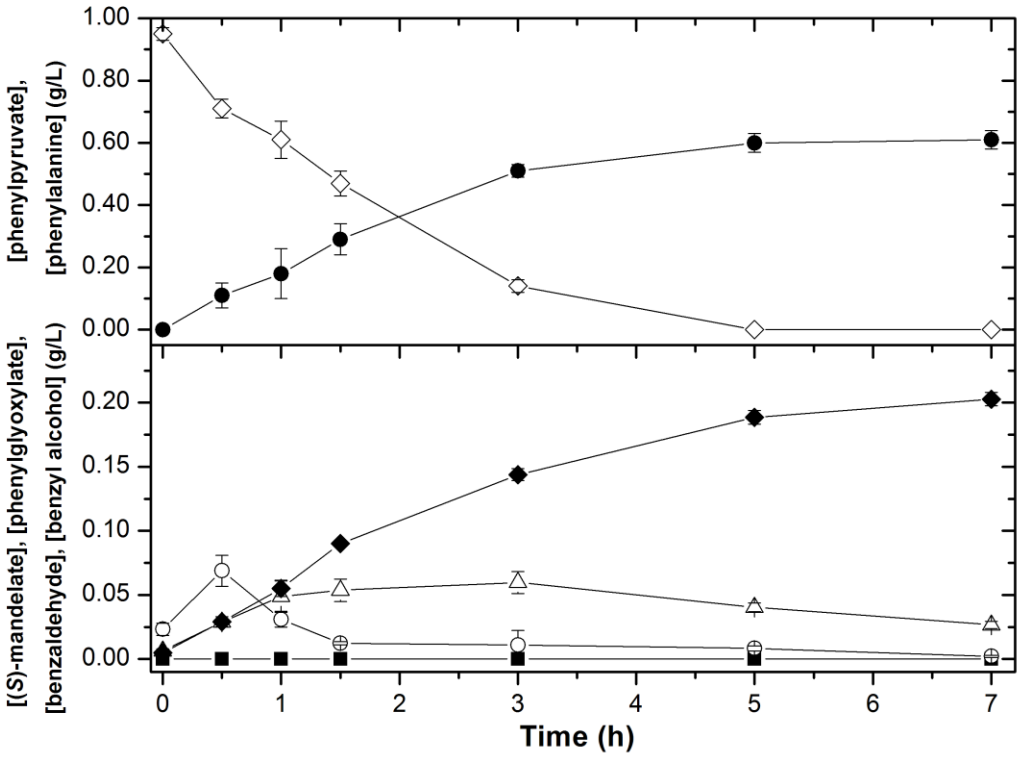

Figure 2.2 Demonstrating pathway function and monitoring metabolite flux via a whole resting cell biotransformation assay using exogenous phenylpyruvate and E. coli NST74 pHmaS-MdlC pMdlB. The conversion of 1 g/L phenylpyruvate (open diamonds) to phenylalanine (solid circles), (S)-mandelate (solid squares), phenylglyoxylate (open circles), benzaldehyde (open triangles), and benzyl alcohol (solid diamonds) was monitored over the course of 7 h. Error bars reported at one standard deviation from triplicate experiments.

Several relevant insights were gained through this initial assay. First, the candidate heterologous enzymes and proposed pathway are functionally expressed in *E. coli* under the examined conditions. Second, a flux bottleneck emerged early on at MdlC, possibly as the result of a cofactor (i.e., $NAD(P)^+$) limitation that was later balanced and eliminated as flux proceed to benzyl alcohol via the NAD(P)H-dependent reduction of

benzaldehyde. Third, despite only expressing *hmaS*, *mdlB*, and *mdlC* (which encode the three steps from phenylpyruvate to benzaldehyde; Figure 2.1), metabolite flux continued efficiently through benzaldehyde to benzyl alcohol, confirming that one or more of *E. coli*'s native AKRs and/or ADHs with activity on benzaldehyde were expressed and functional under the conditions studied. Moreover, considering that cultures were not previously adapted with or even exposed to benzaldehyde, the native regulation of said gene(s) was not part of a benzaldehyde-inducible response, but rather was likely the subject to constitutive expression. Lastly, native phenylalanine biosynthesis clearly emerged as a significant competitor for phenylpyruvate availability (Figure 2.1). This suggests that the engineered pathway, or at least its first committed step (i.e., HmaS), was poorly competitive against the native function of phenylalanine aminotransferase (i.e., TyrB).

2.3.2 Benzyl alcohol biosynthesis from glucose by using E. coli NST74 as production host

Having demonstrated the initial and promising function of the engineered pathway, benzyl alcohol biosynthesis directly from glucose was next investigated. With *E. coli* NST74 as the initial production host, however, benzyl alcohol accumulation to a maximum titer of only 23 ± 2 mg/L was possible in 96 h. Instead, phenylalanine again accumulated as the major end product, reaching up to 900 ± 160 mg/L. This observation agrees well with the outcomes of the phenylpyruvate biotransformation study (Figure 2.2), and was likely a result of the limited affinity of HmaS for phenylpyruvate as substrate. Although HmaS is known to demonstrate activity on phenylpyruvate, its natural and preferred substrate is 4-hydroxyphenylpyruvate. Accordingly, HmaS

displays a nearly 70-fold lower affinity for phenylpyruvate (K_m 0.45± 0.04 mM) versus 4-hydroxyphenylpyruvate (K_m 6.5 ± 0.8 µM) (He, Conrad, and Moran 2010). However, as no other isoenzymes specific for (*S*)-mandelate are presently known, HmaS remains the only suitable candidate at this time. Future isolation or engineering of a superior mandelate synthase will likely be required to achieve increased metabolite flux into the benzyl alcohol pathway, thereby improving achievable titers and yields. For now, the prospect of enhancing benzyl alcohol production by preserving phenylpyruvate availability was examined by eliminating competing aromatic amino acid biosynthesis pathways.

2.3.3 Host strain engineering to improve precursor availability and benzyl alcohol production

In addition to metabolite flux losses to phenylalanine (Figure 2.2), the competing biosynthesis of tyrosine has also previously been found to detract from achievable titers and yields of other aromatic chemicals similarly derived from the phenylalanine biosynthesis pathway (McKenna et al. 2013a). Accordingly, disruption of both tyrosine and phenylalanine biosynthesis was systematically examined in support of enhancing benzyl alcohol production. While deletion of *tyrA* (a bifunctional chorismate mutase/prephenate dehydratase) resulted in greater flux of prephenate through the phenylalanine branch of the pathway (leading to 53% higher phenylalanine titers; NST74A in Table 2.2), increased production of benzyl alcohol was not coincidentally observed (both titer and yield were reduced). To preserve phenylpyruvate, phenylalanine aminotransferase activity was next targeted for disruption. *E. coli* possesses three aminotransferases with reported activity on phenylpyruvate, including *tyrB*, *aspC*, and

ilvE (the latter two are nominally functional on aspartate and branched-chain amino acids, respectively) (Keseler et al. 2005). Prior studies have shown that by deleting both *aspC* and *tyrB* (which possesses nearly 1,000-fold higher activity than *aspC*) while leaving *ilvE* intact, flux of phenylpyruvate to phenylalanine can be reduced to all but the minimal level required to avoid generating a complete phenylalanine auxotroph (Sun et al. 2011, Keseler et al. 2005). Accordingly, strains NST74AB and NST74ABC were next constructed and tested as benzyl alcohol production hosts. As illustrated in Table 2.2, benzyl alcohol titers were nearly doubled to 45 ± 4 mg/L using NST74AB, and further increased to 114 ± 1 mg/L (a ~5-fold increase) using NST74ABC. In both cases, increased benzyl alcohol production was met with corresponding decreases in both phenylalanine and net biomass accumulation, with the latter likely resulting from fitness losses due to reduced amino acid biosynthesis. It should also be noted that no benzaldehyde accumulation was observed at any time for any strains, again confirming that sufficient expression of *E. coli*'s associated AKRs and/or ADHs was achieved via native regulation alone. At its maximum achievable output, the current yield of benzyl alcohol on glucose reached 7.6 ± 0.1 mg/g, or just 3.2% of its theoretical maximum value (240 mg/g; note: the theoretical yield of phenylalanine on glucose has been reported as 0.4 mol/mol (Juminaga et al. 2012)).

Table 2.2. Comparison of benzyl alcohol, phenylalanine, and biomass produced after 72 h by different *E. coli* host strains each harboring pHmaS-MdlC pMdlB.

Host Strain	Benzyl Alcohol		Phenylalanine		Biomass (DCW)	
	Titer (mg/L)	Yield (mg/g)	Titer (mg/L)	Yield (mg/g)	Titer (g/L)	Yield (g/g)
NST74	23 ± 3	1.5 ± 0.2	900 ± 160	60 ± 10	2.2 ± 0.2	0.15 ± 0.02
NST74A	7 ± 1	0.5 ± 0.1	1380 ± 20	92 ± 1	1.7 ± 0.1	0.11 ± 0.01
NST74AB	45 ± 4	3.0 ± 0.2	550 ± 20	37 ± 1	1.1 ± 0.1	0.07 ± 0.01
NST74ABC	114 ± 1	7.6 ± 0.1	410 ± 20	27 ± 1	0.9 ± 0.1	0.06 ± 0.01

2.3.4 Assaying the effect benzyl alcohol on E. coli growth

In addition to the potential fitness reducing effects caused by disrupting aromatic amino acid biosynthesis pathways (Table 2.2), it is possible that benzyl alcohol accumulation in the culture medium further contributed to the observed reduction in biomass growth. This was a particularly relevant concern in this study because, as discussed above, benzyl alcohol is known to possess and is often specifically utilized for its bacteriostatic properties (Marriott 2010). To understand the potential toxic effects associated with benzyl alcohol accumulation, as well as to estimate future limits on achievable titers, a growth challenge assay was lastly performed using exogenous benzyl alcohol. Although exogenous addition does not fully represent the expected environment when benzyl alcohol is instead synthesized intracellularly, this approach has been to provide at least useful first approximations of toxicity for other aromatic products against *E. coli* (McKenna et al. 2013a, Pugh et al. 2014). As seen in Figure 2.3, the initial growth rate was reduced in the presence of 0.25 g/L benzyl alcohol, but not growth yield. However, as the benzyl alcohol concentration was increased to 0.5 g/L and beyond, both

growth rates and yields continued to decline. In the presence of as much as 0.75 g/L benzyl alcohol growth was completely halted following exposure. From this, the toxicity limit of benzyl alcohol against *E. coli* was approximated to be ~0.75 g/L. Lucchini *et al.* reported a value of ~0.4 g/L, albeit with respect to a different strain of *E. coli* and under different culture conditions (Lucchini, Corre, and Cremieux 1990).

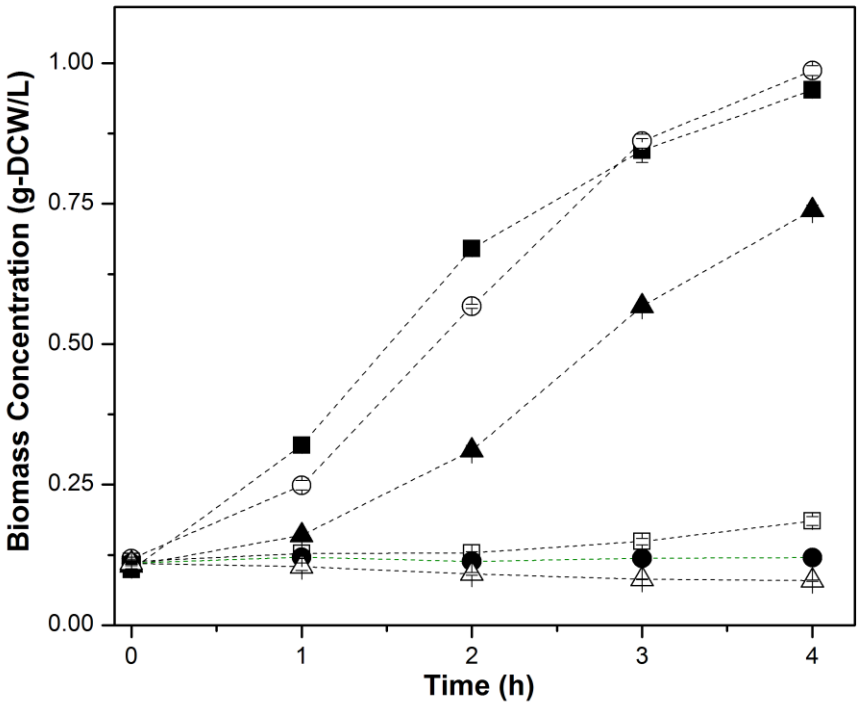

Figure 2.3 Product toxicity of benzyl alcohol. Growth response of E. coli NST74 to benzyl alcohol addition at concentrations of 0 g/L (solid squares), 0.25 g/L (open circles), 0.5 g/L (solid triangles), 0.75 g/L (open squares), 1 g/L (solid circles), and 2 g/L (open triangles). Error bars reported at one standard deviation from triplicate experiments.

The toxicity of most aromatic hydrocarbons against *E. coli* and other Gram-negative bacteria has commonly been suggested to be non-specific in nature, occurring as a result of species lipophilicity which leads to accumulation within the cytoplasmic

membrane, whereupon structural integrity and function become damaged (Meylan et al. 1999, Ramos et al. 2002). To this end, species with log$K_{O/W}$ values of 1-5 have been found to impose significant toxicity against bacteria (note: for benzyl alcohol, $logK_{O/W}$ = 1.10) (Ramos 2004). At the levels observed here, however, benzyl alcohol toxicity against *E. coli* was interestingly found to be poorly represented by a general toxicity model that was previously developed to relate the relative toxicity threshold of an aromatic hydrocarbon with its relative affinity for membrane accumulation (as quantified via the membrane-water partition coefficient, $K_{M/W,}$ and predicted using a previously developed model (Sikkema, de Bont, and Poolman 1994)) (McKenna et al. 2013a). Despite providing a strong linear correlation for several other monoaromatic solvents (i.e., styrene, (*S*)-styrene oxide, (*R*)-1,2-phenylethanediol, *trans*-cinnamate, *p*-hydroxystyrene), the apparent toxicity threshold of benzyl alcohol determined here was ~3-fold lower than its predicted level (~2-2.5 g/L). This implies that other, more specific stress factors may also be contributing to the overall toxicity of benzyl alcohol. Others have also found the toxicity of related aromatic alcohols to be due to both general and specific mechanisms (Lucchini, Corre, and Cremieux 1990). Although a causal relationship was not fully elucidated, exposure to the closely related aromatic alcohol 2-phenylethanol, for example, has been shown to lead to decreased rates of both DNA and RNA synthesis, on top of expected membrane stresses (Lucchini et al. 1993). Still, as current achievable benzyl alcohol titers remain well below the apparent toxicity threshold, it is unlikely that toxicity was a significant, productivity-limiting factor in this study. However, strategies to address this concern will ultimately be required as future strain and pathway engineering efforts lead to further improved benzyl alcohol

production. One attractive approach to this end would involve the expression solvent efflux pumps to actively expel inhibitory products from the cell as they are produced (Dunlop 2011, Dunlop et al. 2011). This strategy appears to be a particularly promising for aromatic products such as benzyl alcohol since several resistance-nodulation-cell division (RND) family efflux pumps from *Pseudomonas* sp. are known to display activity on aromatic species (Ramos et al. 2002, Rojas et al. 2001, Kieboom et al. 1998).

As discussed above, prior works have shown that several ADHs and/or AKRs are involved in the native ability of *E. coli* to reduce benzaldehyde to benzyl alcohol. As we began to explore this phenomenon in the context of *E. coli* benzyl alcohol production (for example, we found that initial rates of benzaldehyde reduction by NST74 *ΔyahK ΔyjgB* were nearly 5-fold lower than by NST74; data not shown), however, a concurrent study by Kunjapur *et al.* reported an insightful and comprehensive investigation to this very same end (Kunjapur, Tarasova, and Prather 2014). Ultimately, an *E. coli* strain was engineered in that study with reduced aromatic aldehyde reduction (RARE) abilities. Lacking 3 AKRs (*dkgB*, *yeaE*, *dkgA*), 3 ADHs (*yqhD*, *yahK*, *yjgB*), and the transcriptional activator *yqhC*, *E. coli* RARE converted less than 12% of exogenous benzaldehyde to benzyl alcohol after 24 h. For comparison, despite lower initial rates of benzaldehyde reduction by NST74 *ΔyahK ΔyjgB*, all added benzaldehyde was ultimately converted to benzyl alcohol within 24 h (data not shown). In related works, meanwhile, Rodriguez and Atsumi also explored the engineering of an *E. coli* strain deficient in aldehyde reductase activity (Rodriguez and Atsumi 2014). Although benzyaldehyde was not evaluated as a substrate, the related aromatic 2-phenylacetaldyde (produced endogenously from phenylpyruvate by Kivd) was investigated. In this case, a mutant

lacking 12 aldehyde reductase encoding genes (*eutE, yahK, yqhE, gldA, ybbO, yghA, yqhD, adhP, eutG, yiaY, yjgB, fucO*) was incapable of reducing 2-phenylacetaldyde to 2-phenylethanol. The similarly observed importance of deleting *yqhD*, *yahK*, and *yjgB* further underscores their likely and general role in *E. coli*'s native ability to reduce aromatic aldehydes.

This insight is important not only for understanding how the host genotype influences flux through the engineered pathway, but also, if this terminal step can be predictably controlled the proposed pathway could furthermore be leveraged to explore the *de novo* biosynthesis of benzaldehyde as an alternative end product. With a global annual production exceeding 90,000 tons (second to only vanillin), benzaldehyde, is a particularly important flavor molecule in the food and fragrance industry (Satrio and Doraiswamy 2001, Culp and Noakes 1990, Krings and Berger 1998), in addition to serving as a precursor to several fine chemicals and pharmaceutical precursors (e.g., (*L*)-phenylacetylcarbinol, or L-PAC) (Rosche et al. 2001). Such prospects will be the subject of future investigations.

2.4 Conclusion

A non-natural pathway to synthesize benzyl alcohol from glucose has been established in *E. coli*. Furthermore, through additional strain engineering to control the native reduction of benzaldehyde it is expected that the same materials could furthermore be employed to establish a biosynthetic route to benzaldehyde. As the pathway is currently limited by low activity at the first committed step, further efforts in enzyme engineering and/or bioprospecting along with the systematic optimization of expression conditions are needed to ultimately elevate key production metrics to viable levels.

2.5 Acknowledgements

This research was supported with the support of start-up funding from Arizona State University.

CHAPTER 3

RATIONAL ENGINEERING OF A NOVEL PATHWAY FOR PRODUCING THE AROMATIC COMPOUNDS P-HYDROXYBENZOATE, PROTOCATECHUATE, AND CATECHOL IN *ESCHERICHIA COLI*

Abstract

p-Hydroxybenzoate, protocatechuate, and catechol represent fine and/or commodity chemicals useful as antioxidants and building-block molecules. To date, however, these species have been largely overlooked as focal end-products. An existing route employing protocatechuate and catechol as intermediates suffers from the need for multiple auxotrophies to preserve precursor (3-dehydroshikimate) availability. A novel, modular route from endogenous p-hydroxybenzoate has been engineered in *Escherichia coli* for the individual biosynthesis of all three products from renewable glucose while minimizing auxotrophy generation. To enhance endogenous biosynthesis of p-hydroxybenzoate, native chorismate pyruvate lyase (*ubiC*) was over-expressed. p-Hydroxybenzoate was converted to protocatechuate by a hydroxylase (*pobA*) from *Pseudomonas aeruginosa*. Catechol was produced by the additional co-expression of protocatechuate decarboxylase from *Enterobacter cloacae*. Systematic expression of appropriate pathway elements in phenylalanine overproducing *E. coli* enabled initial titers of 32 ± 4, 110 ± 8, and 81 ± 15 mg/L for p-hydroxybenzoate, protocatechuate, and catechol, respectively. Disruption of chorismate mutase/prephenate dehydratase (*pheA*) to preserve endogenous chorismate then allowed maximum titers of 277 ± 2, 454 ± 11, and 451 ± 44 mg/L, respectively, at glucose yields of 5.8, 9.7, and 14.3% of their respective

theoretical maxima. Catechol titers were further improved to 630 ± 37 mg/L in a batch bioreactor study. The proposed pathway can furthermore serve as a platform for other bioproducts, including the bioplastics precursor cis,cis-muconate.

This work was published as: Pugh, S., McKenna, R., Osman, Marwan, Thompson, B. & Nielsen, D.R.. Rational engineering of a novel pathway for producing the aromatic compounds p-hydroxybenzoate, protocatechuate, and catechol in *Escherichia coli*. *Process Biochemistry* 49, 1843-1850 (2014).

3.1 Introduction

Aromatic compounds represent a broad class of fine and commodity chemicals with a diversity of industrial and consumer applications. Conventionally, however, their production occurs almost exclusively from non-renewable petroleum resources, typically using carcinogenic benzene as feedstock. As a more sustainable and 'green' alternatives, several prior studies have successfully demonstrated the engineering of microorganisms to produce a variety of aromatic chemical products directly from renewable, biomass-derived sugars. Notable recent examples have included the engineering of microbes to produce the monomer compounds *p*-hydroxystyrene and styrene (McKenna and Nielsen 2011a, Qi et al. 2007a), fragrances/flavors such as 2-phenylethanol and 2-phenylacetic acid (Koma et al. 2012a, Kang et al. 2014), as well as building-block commodity chemicals such as phenol (Kim et al. 2013, Wierckx et al. 2005a). In most cases, the synthesis of these products from renewable feedstocks has been enabled only through *i)* the functional reconstruction of naturally occurring but non-inherent biosynthetic pathways or *ii)* through the *de novo* engineering of novel and non-natural pathways (Lee et al. 2012).

p-Hydroxybenzoate (pHBA), protocatechuate (PCA), and catechol represent three additional aromatic products of commercial and industrial significance. pHBA and PCA are naturally occurring phenolic acids possessing both antioxidant and anti-inflammatory properties (Pacheco-Palencia, Mertens-Talcott, and Talcott 2008). Both species are commonly found at low levels in numerous vegetables, fruits, nuts, tea, and wine (Lin et al. 2007, Pietta et al. 1998, Tian et al. 2009), and are most notably major metabolites produced in acai fruit (*Euterpe oleracea*), where their natural accumulation can reach

levels as high as 892 ± 52 and 630 ± 36 mg/kg, respectively (Pacheco-Palencia, Mertens-Talcott, and Talcott 2008). In addition to these benefits, PCA has been shown to display chemopreventative effects in gastric carcinoma cells as a result of induced *in vitro* apoptosis (Lin et al. 2007). Meanwhile, although it too possesses antioxidant properties, catechol is most commonly used as a precursor for synthesizing fine chemicals such as artificial flavors and fragrances (notably vanillin, eugenol, and guaiacol) (Mageroy et al. 2012, Rhodia 2012), as well as in the larger scale production of many agrochemicals, pesticides, and pharmaceuticals (Rhodia 2012).

Microbial production of PCA and catechol has been demonstrated before, however, rarely as terminal end products or as part of a focused study. In prior works, for example, Draths and Frost established in *E. coli* a non-natural biosynthetic route to *cis,cis*-muconate which utilizes both PCA and catechol as pathway intermediates (Draths and Frost 1994). As illustrated in Figure 3.1, this established pathway (which was also recently reconstructed in yeast (Curran et al. 2013a, Weber et al. 2012)) utilizes as its immediate endogenous precursor 3-dehydroshikimate (DHS), an intermediate of aromatic amino acid biosynthesis used in the shikimic acid pathway (Keseler et al. 2005). DHS is first converted to PCA by DHS dehydratase before PCA decarboxylase subsequently converts PCA to catechol, with said activities most commonly imparted in *E. coli* via the co-expression of *aroZ* and *aroY* from *Klebsiella pneumoniae*, respectively. However, despite the demonstrated activity of this non-natural pathway, robust flux through this route in *E. coli* has only ultimately been achieved upon deletion of shikimate dehydrogenase (encoded by *aroE*) to conserve endogenous DHS (Niu, Draths, and Frost 2002). This strategy, however, consequently creates in *E. coli* an auxotrophy for each of

the proteinogenic aromatic amino acids phenylalanine, tyrosine, and tryptophan and the aromatic vitamins p-aminobenzoate, pHBA, and 2,3-dihydroxybenzoate (collectively used in the production of tetrahydrofolate, ubiquinone, and menaquinone). As a result, in minimal salts media, supplementation of these amino acids and vitamins is required for growth and product formation (Niu, Draths, and Frost 2002). Although nutrient supplementation is facile at the laboratory scale, this strategy is less sustainable and uneconomical as it suffers from higher media costs and poor scalability.

In view of these limitations and the potential significance of each of pHBA, PCA, and catechol as renewable bioproducts, the objective of this study is to engineer a novel and non-natural modular enzyme pathway in *E. coli* with which to systematically explore their biosynthetic potential as individual focal end products of interest. In contrast to the established route presented above, the proposed pathway, which is also illustrated in Figure 3.1, alternatively utilizes pHBA as its immediate endogenous precursor. By focusing on pHBA as the immediate endogenous pathway precursor, the biosynthesis of all aromatic amino acids and vitamins is more readily preserved, improving the overall sustainability of the overall process. To promote the endogenous production of pHBA from chorismate, chorismate pyruvate lyase can first be over-expressed. Next, pHBA can be hydroxylated to PCA via pHBA hydroxylase. Lastly, as in the established pathway, PCA can be decarboxylated to catechol by way of PCA decarboxylase. Furthermore, as seen in Figure 3.1, these heterologous chemistries render the proposed pathway as more thermodynamically favorable than the established route.

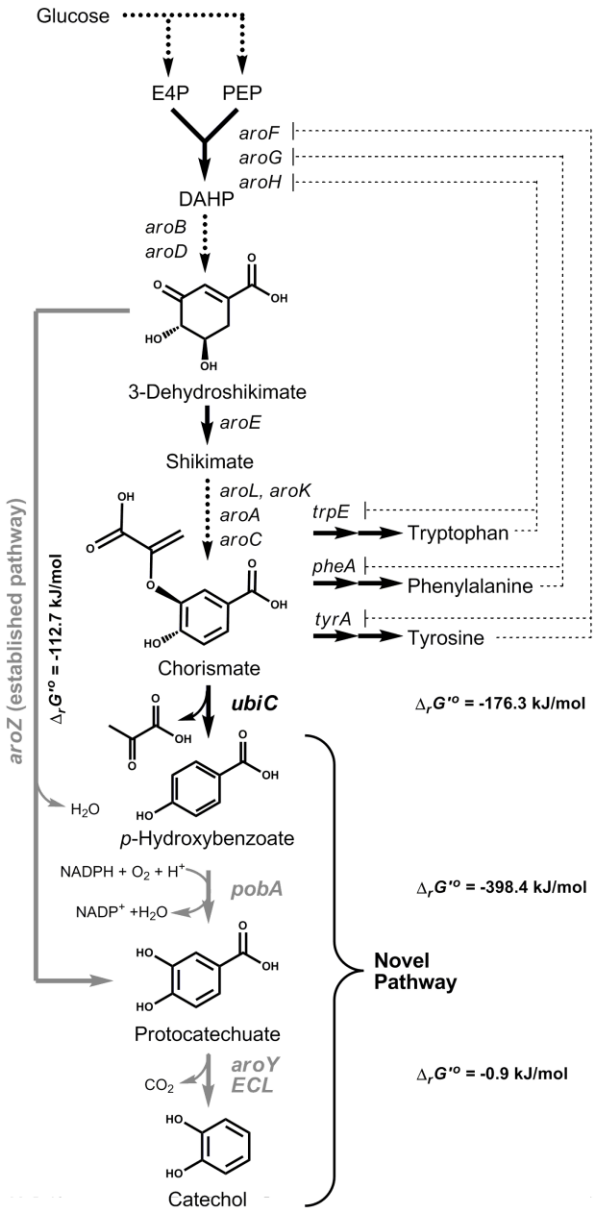

Figure 3.1. Novel and established pathways engineered for producing pHBA, PCA, and catechol from glucose. Black arrows represent enzyme steps native to E. coli whereas gray arrows are heterologous; dotted arrows represent multiple enzymatic steps; feedback regulation is shown using thin dotted lines with flat heads. Abbreviations: phosphoenolpyruvate (PEP), D-erythrose-4-phosphate (E4P), 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP). $\Delta_r G'^\circ$ is the change in Gibbs free energy due to reaction as determined using the online tool eQuilibrator (http://equilibrator.weizmann.ac.il) at a reference state of 25^{o}C, pH 7, and ionic strength of 0 M.

3.2 Materials and Methods

3.2.1 Strains and media

All strains used in this study are listed in Table 3.1. *E. coli* NEB 10-beta was obtained from New England Biolabs (Ipswich, MA) and was used for routine cloning work, as well as for the propagation and storage of all plasmids. *E. coli* NST74 (ATCC 31884; *aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr), pheO352, aroG397(fbr)*), a previously-engineered phenylalanine over-producing strain (Tribe 1987a), was obtained from the American Type Culture Collection (ATCC, Manassas, VA) and was used as the initial host platform for all aromatic chemical production. *E. coli* BW25113 was obtained from the Coli Genetic Stock Center at Yale University (CGSC, New Haven, CT) and served as the genetic source of *ubiC*, which encodes chorismate pyruvate lyase. *Pseudomonas aeruginosa* PAO1 (DSMZ 22644) was obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and served as the genetic source of *pobA*, which encodes pHBA hydroxylase. *Klebsiella pneumoniae* PZH572 (ATCC 25955) was obtained from the ATCC and served as the genetic source of *aroY*, which encodes PCA decarboxylase. *E. coli* JW2580-1 was obtained from the CGSC and served as the genetic source of the *pheA::kan*^{R} cassette used to disrupt *pheA* in *E. coli* NST74

E. coli cultures were routinely cultivated in Luria-Bertani (LB) broth. For aromatics production, a phosphate-limited minimal media (herein referred to as "MM1") was used, as adapted from McKenna *et al.* (McKenna and Nielsen 2011a). MM1 was composed of glucose (20 g/L), $MgSO_4 \cdot 7H_2O$ (0.5 g/L), $(NH_4)_2SO_4$ (4.0 g/L), MOPS (24.7 g/L), KH_2PO_4 (0.3 g/L), K_2HPO_4 (1.0 g/L) and 5 ml/L ATCC Trace Mineral

Supplement (Catalog No. MD-TMS): EDTA (0.5 g/L), $MgSO_4 \cdot 7H_2O$ (3.0 g/L), $MnSO_4 \cdot 7H_2O$ (0.5 g/L), NaCl (1.0 g/L), $FeSO_4 \cdot 7H_2O$ (0.1 g/L), $Co(NO_3)_2 \cdot 6H_2O$ (0.1 g/L), $CaCl_2$ (0.1 g/L), $ZnSO_4 \cdot 7H_2O$ (0.1 g/L), $CuSO_4 \cdot 5H_2O$ (0.01 g/L), $AlK(SO_4)_2$ (0.01 g/L), H_3BO_3 (0.01 g/L), $Na_2MoO_4 \cdot 2H_2O$ (0.01 g/L), $NaSeO_3$ (0.001 g/L), $NaWo_4 \cdot 2H_2O$ (0.1 g/L), and $NiCl_2 \cdot 6H_2O$ (0.02 g/L). As required when working with a *ΔpheA* background, ,MM1 was supplemented with phenylalanine at an initial concentration of 0.1 g/L. For plasmid maintenance, cultures were also appropriately supplemented with ampicillin (100 mg/L) and kanamycin (40 mg/L).

Table 3.1. List of strains and plasmids *engineered and/or* used in this study.

Strains	Description	Source
E. coli NEB-10 beta	*araD139 Δ(ara,leu)7697 fhuA lacX74 galK16 galE15 mcrA f80d(lacZΔM15)recA1 relA1 endA1 nupG rpsL rph spoT1Δ(mrr-hsdRMS-mcrBC)*	New England Biolabs
E. coli BW25113	source of *ubiC*	CGSC
E. coli JW2580-1	source of *pheA::kan*^{R}	CGSC
E. coli NST74	*aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr), pheO352, aroG397(fbr)*	ATCC
E. coli N74dpheA	NST74 *ΔpheA*	This study
P. aeruginosa PAO1	source of *pobA*	DSMZ
K. pneumoniae PZH572	source of *aroY*	ATCC
E. cloacae ATCC 13047	source of *ECL_01944*	ATCC

Plasmids	Description	Source
pTrc99A	*Ptrc*, pBR322 ori, *lacIq*, *Amp*^{R}	Prather Lab, MIT
pTrcCOLAK	*Ptrc*, ColA ori, *lacIq*, *Kan*^{R}	McKenna *et al.* (2013)
pKD46	*araC*, *araBp*, repA101(ts) and R101 ori, *Amp*^{R}	CGSC
pCP20	*FLP*, *ts-rep*, *[cI857](lambda)(ts)*, *Amp*^{R}	CGSC
pUbiC	*ubiC* of *E. coli* BW25113 inserted into the *NcoI* and *EcoRI* sites of pTrc99A	This study
pPobA	*pobA* of *P. aeruginosa* inserted into the *XbaI* and *HindIII* sites of pTrc99A	This study
pUbiC-PobA	*pobA* of *P. aeruginosa* inserted into the *XbaI* and *HindIII* sites of pUbiC	This study
pAroY	*aroY* of *K. pneumoniae* ATCC 25955 inserted into the *NcoI* and *EcoRI* sites of pTrcCOLAK	This study
pECL	*ECL_01944* of *E. cloacae* ATCC 13047 inserted into the *NcoI* and *EcoRI* sites of pTrcCOLAK	This study

3.2.2 Plasmid Construction.

All plasmids constructed and used in this study are listed in Table 3.1. Custom DNA oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA). Genomic DNA (gDNA) was prepared using the ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research, Irvine, CA) according to manufacturer protocols. All genes were PCR amplified using Phusion High-Fidelity DNA Polymerase (New England Biolabs) and

standard protocols. Amplified linear DNA was purified using the DNA Clean & Concentrator kit (Zymo Research, Irvine, CA) according to manufacturer protocols. Purified DNA was subsequently digested using appropriate restriction endonucleases (all from New England Biolabs) at 37^{o}C for three hours. Digested fragments were subsequently purified using the Zymoclean Gel DNA Recovery kit (Zymo Research) and ligated together using T4 DNA ligase (New England Biolabs) at 4oC overnight. Ligation reactions were transformed into chemically competent E. coli NEB10-beta, with selection performed by plating on LB solid agar with appropriate antibiotics. Transformant pools were screened by colony PCR using the same primers used in cloning initial fragments. Plasmid sequences were further confirmed by restriction digest mapping and gene sequencing.

Using the compatible expression vectors pTrc99A and pTrcCOLAK [20], all pathway genes were expressed under the control of the strong, IPTG-inducible trc promoter and lacIq repressor system. To construct pUbiC, ubiC was PCR amplified from the gDNA of E. coli BW25113 and inserted between the NcoI and EcoRI sites of pTrc99A. Subsequently, pobA was PCR amplified from the gDNA of P. aeruginosa DSMZ 22644 and inserted between the XbaI and HindIII sites of pUbiC, resulting in the construction of pUbiC-PobA. To construct pAroY, aroY was PCR amplified from the gDNA of K. pneumoniae ATCC 25955 and inserted between the NcoI and EcoRI sites of pTrcCOLAK. Lastly, to construct pECL, ECL_01944 was PCR amplified from the gDNA of E. cloacae ATCC 13047 and inserted between the NcoI and EcoRI sites of pTrcCOLAK.

3.2.3 Gene deletion

Chromosomal in-frame deletion of *pheA* in *E. coli* NST74 was accomplished via a method modified from that of Datsenko and Wanner (Datsenko and Wanner 2000a). Briefly, the deletion cassette *pheA::kan*^{R} flanked by the FLP recognition target (FRT) sites was PCR amplified from the gDNA of *E. coli* JW2580-1 using primers with 300 bp homology upstream and downstream of *pheA*. Subsequent steps were the same as previously described (Datsenko and Wanner 2000a). This resulted in the construction of *E. coli* NST74 *ΔpheA*, referred to as N74dpheA.

3.2.4 Toxicity assay of pHBA, PCA and catechol

The toxicity of pHBA, PCA, and catechol against *E. coli* NST74 was estimated by assaying for relative change in growth rate and yield following the exogenous addition of each individual species to growing cultures as a function of concentration. Cells were cultures in 50 mL MM1 media in 250 mL baffled glass shake flasks at 30^{o}C to an OD_{600} of ~0.5, at which point either pHBA, PCA, or catechol was added to the flasks at an array of final concentrations and growth was monitored via OD_{600} for an additional 6 h. All experiments were performed in triplicate to provide an estimate of standard error.

3.2.5 Assaying in vivo enzyme functionality using whole resting cells

To assess the functionality of the candidate enzymes, including PobA for pHBA hydroxylase activity as well as AroY and ECL for PCA decarboxylase activity, *in vivo* resting cell assays were performed. *E. coli* NST74 was individually transformed with each of pPobA, pAroY, and pECL before seed cultures of the resultant strains were

prepared in 5 mL LB broth with 100 mg/L ampicillin or 40 mg/L kanamycin, as appropriate, and grown overnight at 37^{o}C. Subsequently, 0.5 mL of each seed was used to inoculate 50 mL of LB broth supplemented with antibiotics and 0.5% glucose in a 250 mL baffled shake flask. Upon reaching mid-exponential phase (OD_{600} ~0.5), cultures were induced by the addition of 0.25 mM IPTG. Culturing then continued for 24 h at 37^{o}C, at which point cells were collected by centrifugation at 3,000 x *g* for 10 min. Cell pellets were washed once in pH 7 phosphate buffered saline solution (PBS) before then finally being re-suspended to an initial OD_{600} of ~4 in pH 7 PBS supplemented with 500 mg/L of pHBA to assess the hydroxylase activity of PobA or 500 mg/L of PCA to compare the decarboxylase activity of AroY and ECL. Subsequently, 10 mL of each suspension was transferred to a sealed 15 mL glass Hungate tube and cultures were incubated for 2 h at 37^{o}C with shaking at 250 rpm while aqueous samples periodically drawn for metabolite analysis by HPLC, as described below. All experiments were performed in triplicate to provide an estimate of standard error.

3.2.6 Production of pHBA, PCA, and catechol from glucose in shake flask cultures

E. coli NST74 and the phenylalanine auxotroph N74dpheA were each transformed with either pUbiC or pUbiC-PobA, for the biosynthesis of pHBA or PCA, respectively, or co-transformed with pUbiC-PobA and pAroY or pECL for the biosynthesis of catechol. Seed cultures (0.5 mL) were subsequently used to inoculate 50 mL MM1 media in 250 mL baffled shake flasks supplemented with appropriate antibiotics and, in the case of N74dpheA-derived strains, phenylalanine. The initial pH of the medium was ~7.2 and was maintained at >6 throughout the experiment by the periodic addition of 1 M monobasic potassium phosphate solution. Cultures were grown

at 32°C with shaking at 250 rpm until reaching an OD_{600} of ~0.5, at which point they were induced by the addition of 0.25 mM IPTG. Culturing then continued for an additional 72 to 96 h, during which time aqueous samples were periodically drawn for metabolite analysis by HPLC. Product yields were determined as the net moles of product synthesized per total moles of glucose consumed. Yields were represented as a percentage of their theoretical maxima using an established estimate of the theoretical maximum yield of phenylalanine from glucose of 0.40 mol/mol (Juminaga et al. 2012). All experiments were performed in triplicate to provide an estimate of standard error.

3.2.7 Production of catechol from glucose in a batch bioreactor

E. coli N74dpheA was co-transformed with pUbiC-PobA and pECL and a 15 mL seed culture was prepared in LB broth supplemented with appropriate antibiotics overnight. The seed was then used to directly inoculate a 2 L BIOSTAT Aplus Sartorius bioreactor containing 1 L of MM1 supplemented with appropriate antibiotics and phenylalanine. Temperature, agitation, and aeration were held constant at 32°C, 200 rpm, and 1.5 L/min, respectively. Culture pH was maintained at 6.9 throughout by the automated addition of 1M KH_2PO_4 solution. Culturing continued for a total period of 86 h, during which time samples were taken periodically for biomass and metabolite analysis.

3.2.8 HPLC analysis

HPLC analysis of all metabolites was conducted using a Hewlett Packard 1100 series HPLC system (Santa Clara, CA) coupled with a UV/vis detector. Separation of metabolites was achieved using a reverse-phase Hypersil Gold aQ polar endcapped C18

column (4.6 mm x 150 mm; Thermo Fisher, USA). Samples (5 μL) were injected at a total constant flow rate of 1 mL/min with a mobile phase consisting of 85% 5 mM sulfuric acid (pH of 2) and 15% acetonitrile. Temperature was held constant at 45°C. The eluent was monitored via DAD detector at 215 nm for the detection of phenylalanine, PCA, and catechol, and 260 nm for pHBA. Glucose was also measured by HPLC, as previously described (McKenna et al. 2013a).

3.3 Results and Discussion

3.3.1 Assessing the toxicity of pHBA, PCA and catechol

The cytotoxic effects of pHBA, PCA, and catechol were first individually assayed to establish an estimate of how their accumulation in cultures will impact host fitness. Due to their hydrophobic nature, aromatic compounds commonly accumulate within the core of the lipid-bilayer comprising the cytoplasmic membrane, leading to increases in membrane fluidity and loss of structural integrity (Meylan et al. 1999, Ramos et al. 2002). The impact of pHBA, PCA, and catechol on *E. coli* growth was assessed following their exogenous addition to early exponential phase cultures at increasing concentrations (Figure 3.2). Although this approach does not wholly represent the culture environment expected when said species will instead be produced intracellularly, this method has been found to provide at least a reasonable first approximation of the toxicity of other aromatic products against *E. coli* (McKenna et al. 2013a). As seen in Figure 3.2, whereas *E. coli* growth rates were reduced in the presence of all three species at elevated concentrations, with critical limits (above which no growth was observed following addition to cultures) emerging at approximately 1.5, 1.5, and 3-4 g/L for pHBA, PCA, and catechol, respectively. These levels were interpreted as the maximum

toxicity thresholds for each species. For solvent-like molecules, the octanol-water partition coefficient ($\log K_{O/W}$) is correlated with intra-membrane accumulation and thus cytotoxicity. In particular, those species with $\log K_{O/W}$ values of 1-5 have been found to impose significant toxicity against bacteria (Ramos 2004). For reference, $\log K_{O/W}$ values of pHBA, PCA, and catechol are approximately 1.56, 1.06, and 0.88, respectively. Accordingly, it should be expected that the relative toxicities of the three target products would follow as: pHBA > PCA > catechol. However, the finding that pHBA and PCA imposed similar toxicity thresholds could suggest that *E. coli* possesses an improved ability to tolerate the former. This is perhaps due to the native ability of *E. coli* to assimilate pHBA into its ubiquinone (i.e., coenzyme Q) biosynthesis pathway. As an inner membrane protein, 4-hydroxybenzoate octaprenyltransferase (encoded by *ubiA*) may lessen the extent by which pHBA accumulates in the membrane via its partial conversion to 3-octaprenyl-4-hydroxybenzoate, thereby reducing its inhibitory effect. This hypothesis, however, was not further tested in a *ΔubiA* background as control due to the essential role that ubiquinone plays in *E. coli* during aerobic respiration (Kwon, Kotsakis, and Meganathan 2000). Regardless, and for the purposes of this study, these assays suggest that toxicity limitations will be most likely to arise when pHBA and PCA constitute the focal end-products of interest. Moreover, swift conversion of pHBA and PCA via pHBA hydroxylase and PCA decarboxylase, respectively, may emerge as an important consideration when the pathway is ultimately extended to catechol.

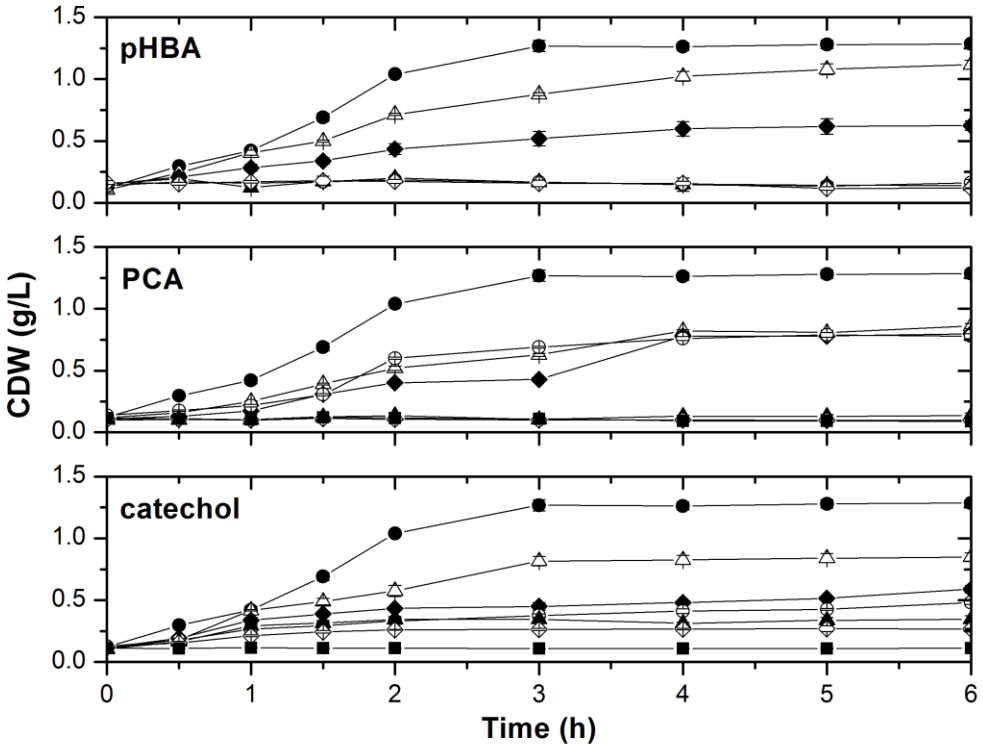

Figure 3.2. Toxicity analysis of pHBA, PCA, and catechol in E. coli. Growth response of E. coli NST74 to the exogenous addition of: pHBA (upper) at 0 g/L (solid circles), 0.5 g/L (open triangles), 1 g/L (solid diamonds), 1.5 g/L (open circles), 2 g/L (solid triangles), and 2.5 g/L (open diamonds); PCA (middle) at 0 g/L (solid circles), 0.5 g/L (open triangles), 1 g/L (solid diamonds), 1.25 g/L (open circle), 1.5 g/L (solid triangle), 2 g/L (open diamonds), and 2.5 g/L (solid squares); catechol (lower) at 0 g/L (solid circles), 1 g/L (open triangles), 2 g/L (solid diamonds), 2.5 g/L (open circles), 3 g/L (solid triangles), 3.5 g/L (open diamonds), and 4 g/L (solid squares).

3.3.2 Screening and selecting pathway enzymes

The first step of the proposed, modular biosynthetic pathway involves the conversion of endogenous chorismate to pHBA via chorismate pyruvate lyase (Figure 3.1). In *E. coli*, this inherent reaction is catalyzed by UbiC which, as alluded to above, naturally functions as the first step in the ubiquinone biosynthesis pathway (Keseler et al. 2005). However, native production of pHBA is tightly regulated such that it does not

naturally accumulate to above trace levels. In NST74 alone, pHBA production was not detected. Furthermore, in a preliminary experiment, PCA production was similarly undetected in a control strain where *pobA* was overexpressed in NST74 (data not shown). Thus, to recruit sufficient chorismate flux into the proposed pathway and enhance endogenous pHBA biosynthesis, *ubiC* was overexpressed from pUbiC in the phenylalanine overproducing strain NST74. In glucose minimal media, the resultant strain produced pHBA at titers reaching up to 32 ± 4 mg/L (or 0.23 ± 0.03 mM) after 96 h, as seen in Table 3.2. Meanwhile, in addition to pHBA, phenylalanine also accumulated as a significant co-product, reaching up to 1.28 ± 0.03 g/L (or 7.75 ± 0.18 mM). As pHBA and phenylalanine share chorismate as a common branch-point precursor, this suggested that, even when over-expressed, UbiC was outcompeted by the native activity of PheA (Figure 3.1). It should be noted that NST74 possesses a feedback-deregulated copy of PheA. Furthermore, while chorismate also serves as a precursor to tryptophan and tyrosine, as the transcription of both *trpE* and *tyrA* remain subject to tight feedback repression in NST74, neither of these species were detected in the culture media.

Table 3.2. Shake flask production of pHBA, PCA, catechol and phenylalanine (Phe), as well as biomass. Not detected: n.d., biomass dry cell weight: DCW.

Host	Plasmid	End Product	Aqueous Product Titer (mg/L) at:			Final Phe (g/L)	Yield (mg/g-glucose)	Final DCW (g/L)
			48 h	72 h	96 h			
NST74	pUbiC	pHBA	28 ± 2	29 ± 2	32 ± 4	1.28 ± 0.03	2.1 ± 0.3	2.6 ± 0.1
N74dpheA	pUbiC	pHBA	251 ± 5	269 ± 1	277 ± 2	n.d.	17.8 ± 0.2	2.0 ± 0.1
NST74	pUbiC-PobA	PCA	92 ± 7	110 ± 8	110 ± 8	1.20 ± 0.05	7.1 ± 0.6	2.6 ± 0.2
N74dpheA	pUbiC-PobA	PCA	426 ± 13	442 ± 10	454 ± 10	n.d.	33.2 ± 0.7	1.9 ± 0.1
NST74	pUbiC-PobA pECL	catechol	80 ± 18	81 ± 15	80 ± 14	1.04 ± 0.07	6.0 ± 1.0	1.3 ± 0.1
N74dpheA	pUbiC-PobA pECL	catechol	394 ± 62	440 ± 50	451 ± 44	n.d.	35.1 ± 2.6	1.1 ± 0.2

Extension of the proposed pathway to PCA would next be possible via the co-expression of pHBA hydroxylase. Several *Pseudomonas sp.* have been shown capable of performing this reaction as part of a larger pHBA degradation pathway. In this innate pathway, which ultimately links to central metabolism via the β-ketoadipate pathway (Kemp and Hegeman 1968), NADPH-dependent pHBA hydroxylase activity is encoded by *pobA*. Interestingly, the sequence of PobA and related isoenzymes is highly conserved amongst *Pseudomonas sp.*, with ~85% homology observed, for example, between *P. aeruginosa* PAO1, *P. putida* KT2440, *P. fluorescens* A506, and *P. stutzeri* A1501. PobA from *P. aeruginosa* PAO1 was selected for screening as a candidate pathway enzyme and its *in vivo* activity was evaluated in *E. coli* resting cells. When supplied with 480 ± 21 mg/L (3.5 ± 0.2 mM) of exogenous pHBA, its full and stoichiometric conversion to PCA (reaching up to 535 ± 22 mg/L or 3.5 ± 0.2 mM)

occurred in ~90 min (Figure 3.3A), confirming the functional expression of PobA in *E. coli.*

Finally, the proposed pathway could be extended to catechol by the additional co-expression of PCA decarboxylase. It has previously been shown that this can be achieved in *E. coli* by expression of *aroY* from *K. pneumoniae* (Draths and Frost 1994). In addition, the locus *ECL_01944* from *E. cloacae* (hereafter referred to as *ECL*) has also been reported to encode a PCA decarboxylase (Curran et al. 2013a). Although the heterologous activity of *ECL_01944* in *S. cerevisiae* has been confirmed, its analogous function in *E. coli* has not yet been reported. To determine which homolog possesses the greatest activity in *E. coli*, a series of comparative *in vivo* resting cell assays were performed. Here, exogenous PCA was supplied to whole resting cells prepared from NST74 pECL and NST74 pAroY, expressing *ECL_01944* and *aroY*, respectively. As illustrated in Figure 3.3B, PCA was stoichiometrically converted to catechol in just 60 min by both strains (note, whereas NST74 pECL showed slightly faster rates of PCA conversion, the difference was not statistically significant; $p > 0.5$, $n = 3$). The observed comparable activities are likely a consequence of the similarity between the predicted amino acid sequences of AroY (502 AA) and ECL (495 AA), which were found to be 89% identical (analysis not shown). Ultimately, *ECL_01944* was selected for the second step of the proposed pathway owing the novelty associated with its expression in *E. coli*.

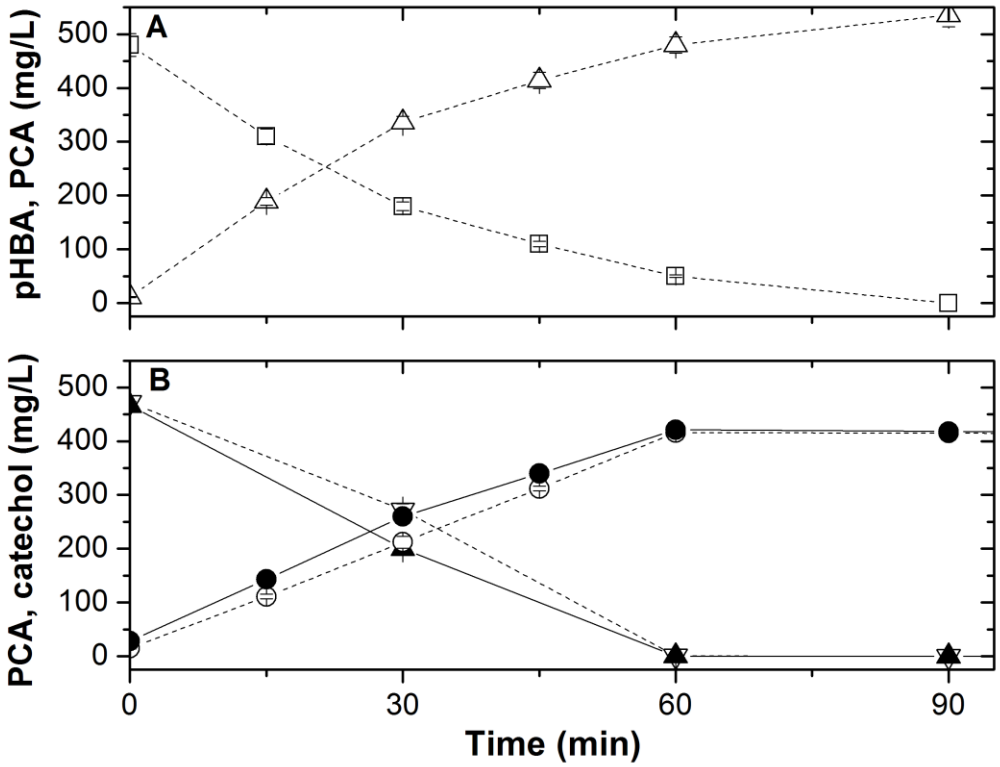

Figure 3.3. Screening of candidate pathway enzymes via in vivo resting cell assays. A) Conversion of pHBA (open squares) to PCA (open triangles) by NST74 pPobA. B) Conversion of PCA (triangles) to catechol (circles) by cultures of NST74 pECL (solid symbols) or NST74 pAroY (open symbols). Error bars reported at one standard deviation from triplicate experiments.

3.3.3 Production of pHBA, PCA, and catechol from glucose in E. coli NST74

Having confirmed the activity of all candidate pathway enzymes in *E. coli* in isolated biotransformation studies, the stepwise construction of a modular pathway for individual biosynthesis of pHBA, PCA, and catechol directly from glucose was next explored. NST74 served as the initial host background for all three products and the collective results are compared in Table 3.2 (note, the results for pHBA were also reported above). In all cases, production stopped after 48-72 h following the depletion of available

glucose. In the case of PCA, co-expression of *ubiC* and *pobA* allowed its accumulation as the major end product at up to 110 ± 8 mg/L (or 0.71 ± 0.05 mM) after 96 h. The additional co-expression of *ECL*, meanwhile, enabled catechol production as the major end product at up to 80 ± 14 mg/L (or 0.73 ± 0.13 mM) in the same time. It should also be noted that, in all cases, no intermediate accumulation was observed (i.e., no pHBA or PCA and no pHBA were detected when catechol and PCA were the respective end products of interest), suggesting that the net activity of each pathway step was sufficient to preclude the occurrence of a metabolite flux bottleneck. Meanwhile, when compared to the production of pHBA alone (i.e., when *ubiC* was solely expressed), further extension of the pathway to PCA and catechol resulted in 3.2- and 3.4-times greater net accumulation of these focal end-product, respectively. Whereas in the case of catechol it would be conceivable that this sizable increase resulted from its reduced toxicity relative to pHBA (Figure 3.2), the same argument would clearly not hold true for the equally inhibitory PCA. Rather, a more likely explanation for the increased metabolite flux through the pathway observed upon introducing *pobA* or *pobA* and *ECL* stems from the fact that UbiC suffers from severe allosteric feedback inhibition by pHBA (Siebert, Severin, and Heide 1994, Holden et al. 2002). Prior studies have shown, for example, that UbiC exhibits a 13-fold higher affinity for its product pHBA (K_p = 2.1 μM) than for its substrate chorismate (K_m = 29 μM) (Gallagher et al. 2001, Nichols and Green 1992, Holden et al. 2002). As a result, in the presence of 25 or 100 μM pHBA, reductions in reaction velocity by 51 and 83% have been observed *in vitro*, respectively (Siebert, Severin, and Heide 1994). Whereas this limitation could perhaps be overcome through protein engineering, the residues involved in pHBA binding have so far not been

specifically identified. In the present case, however, by converting pHBA to PCA via PobA (a reaction with highly favorable thermodynamics; $\Delta_r G'^{\circ}$ = -398.4 kJ/mol) soon after its endogenous biosynthesis, UbiC activity and flux through the first committed pathway step appear to be maintained at elevated levels.

Meanwhile, significant co-production and the terminal accumulation of phenylalanine was also observed for all three strains, reaching levels of at least 1.04 ± 0.07 (or 6.29 ± 0.43 mM). Again, this suggests that the engineered pathways were not able to effectively compete with native phenylalanine biosynthesis for precursor chorismate. Host engineering strategies to overcome this limitation were next systematically explored.

3.3.4 Host engineering to increase pHBA, PCA, and catechol production from glucose

To preserve chorismate availability for use in the engineered pathways, the bifunctional chorismate mutase/prephenate dehydrogenase encoded by pheA was first deleted from the background of NST74. Using the resultant strain (i.e., N74dpheA) as host, after 96 h (Table 3.2), maximal titers of pHBA, PCA, and catechol reached 277 ± 2, 454 ± 10, and 451 ± 44 mg/L, respectively (or 2.00 ± 0.01, 2.95 ± 0.06, and 4.10 ± 0.39 mM), improvements of 4.1- to 8.7-fold over NST74. At this output, product yields (which were similarly increased relative to the NST74 background; Table 3.2) from glucose represented 5.8, 9.7, and 14.3% of their respective theoretical maxima (0.40 mol/mol in each case [22], or 337, 376, 269 mg/g, respectively). One consequence to this strategy, however, was an apparent decrease in biomass production, suggesting a reduction in host fitness. In addition, further reduced biomass production in catechol producing strains was likely due to the increased metabolic burden associated with carrying as additional

plasmid and over-expression of an additional pathway enzyme. As in NST74, the greatest pathway flux was observed when PCA and catechol were the focal products of interest, again presumably due to the relieved inhibition of UbiC by pHBA.

Recently, another novel and alternative route through catechol (to cis,cis-muconate) was established in E. coli [32]. Said route stemmed instead from endogenous anthranilate, using a heterologous anthranilate 1,2-dioxygenase from P. aeruginosa to produce catechol. An improvement over the established route was also realized with this pathway in terms of the need to generate just a single (tryptophan) auxotrophy to achieve appreciable production levels. Whereas the biosynthesis of catechol as an end-product was not investigated in that study, maximal titers of cis,cis-muconate via the extended pathway reached 389 mg/L, indicating a maximal metabolite flux of 2.74 mM through the pathway. Although this is just 67% of the level demonstrated here, a direct and fair comparison is difficult to make at this point as it is unclear what effect the additional co-expression of catechol 1,2-dioxygenase might impose on either pathway or cell system.

While the generation of a single auxotrophy here is a certain improvement over the multiple auxotrophies created in support of the established route through catechol (Figure 3.1), a more robust approach might be to instead explore a gene knockdown strategy as a means to preserve chorismate availability. Previous studies have found, for example, that the use of small regulatory RNAs (sRNAs) are well suited for such applications [33, 34], and have been used to boost the production of aromatic amino acids [35]. Alternatively, replacement of the feedback deregulated copy of pheA in NST74 with its wild-type, allosterically-inhibited parent might be sufficient for restricting

chorismate flux of through this competing step. The relative merits and prospects of these approaches are now being explored.

3.3.5 Production of catechol from glucose by E. coli in a batch bioreactor

An important practical limitation experienced in shake flask cultures was the need to perform regular and manual pH adjustments, due largely to the poor buffering capacity provided by MM1 media. Moreover, since PobA is known to be an oxygen-dependent hydroxylase, it was postulated that poor rates of oxygen transfer experienced in shake flask cultures could have limited the overall activity at this step. To address these practical limitations with the goal of enhancing overall production of the terminal pathway product, catechol production was lastly investigated under controlled culture conditions in a batch bioreactor. The collective results are presented in Figure 3.4, wherein it can be seen that a total of just over 630 ± 37 mg/L catechol (or 5.72 ± 0.34 mM) was produced by *E. coli* N74dpheA pUbic-PobA pECL in 86 h. This represents a ~40% improvement over that achieved in shake flask cultures but at a nearly equivalent glucose yield (36.2 ± 1.9 mg/g; see Table 3.2). Similar to the study of Curran *et al.* (Curran et al. 2013a), the slight accumulation of PCA was observed early in the culture (i.e., within the first ~24 h), reaching as high as 137 mg/L. Although PCA was subsequently re-assimilated into the pathway and converted to catechol, its intermediary accumulation indicates the potential of a flux bottleneck existing at PCA decarboxylase. This is likely a consequence of the fact that the thermodynamics of said reaction are only slightly favorable ($\Delta_r G'^{\circ}$ = -0.9 kJ/mol). The observed lack of accumulation of pHBA through the duration of the culture, meanwhile, suggests that PobA activity was not rate limiting. This is not surprising as dissolved oxygen was abundant in the bioreactor at all

times and, again, this reaction is highly favorable thermodynamically ($\Delta_r G'° = -398.4$ kJ/mol).

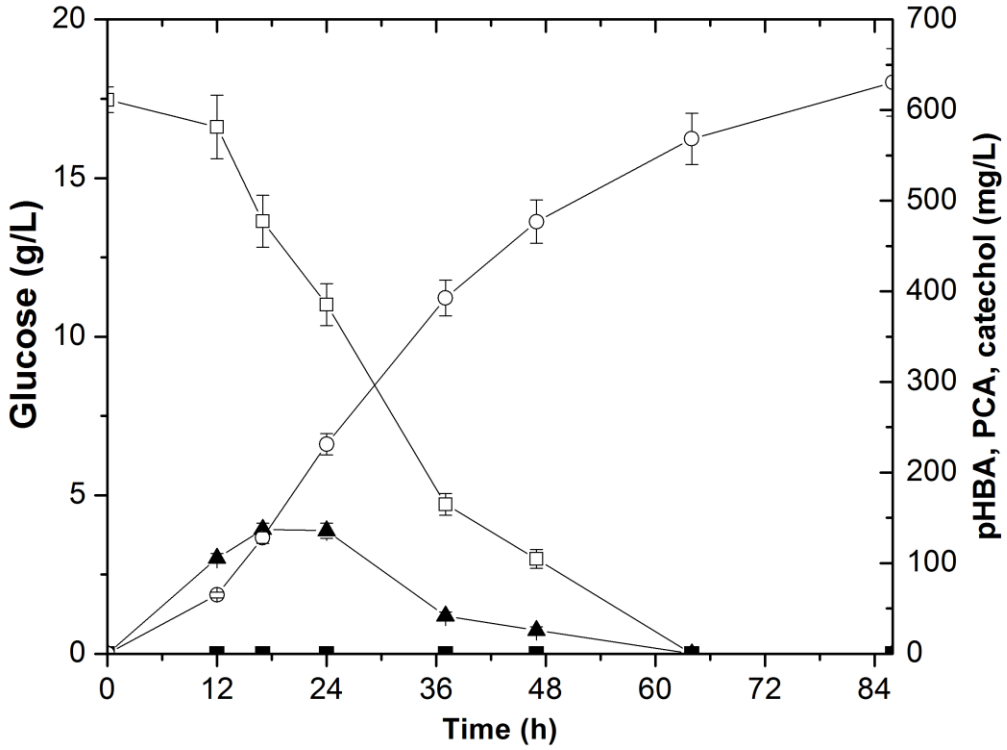

Figure 3.4. Production of catechol by *E. coli* N74dpheA pUbic-PobA pECL in a batch bioreactor. Concentrations of glucose (open squares), pHBA (solid squares), PCA (solid triangles), and catechol (open circles).

As the inhibitory limit of catechol against *E. coli* was estimated at about 3-4 g/L, the ability to produce just 630 ± 37 mg/L of catechol after 86 h in a batch bioreactor suggests that end-product toxicity was of little concern in this study and that another factor was surely productivity limiting. As flux bottlenecks were not observed along the engineered pathway, it is likely that the pathway was limited by the availability of precursor chorismate. As the original host background (i.e., NST74) was capable of

producing up to 7.75 mM phenylalanine in shake flask cultures, if the entire chorismate supply were instead diverted to pHBA, PCA, or catechol, maximal titers of 1070, 1195, and 855 mg/L would instead be expected, respectively. However, the maximum flux of metabolites through the engineered pathways never exceeded 4.10 mM under otherwise analogous culture conditions. This could be due to the fact that, despite improving the availability of chorismate to the engineered pathways, *pheA* deletion imposed a negative impact net chorismate biosynthesis. This would be consistent with the observed fitness reduction in *ΔpheA* strains, as indicated by the 15-27% lower biomass production relative to the NST74 background.

3.3.6 Future applications

Through its stepwise extension, the modular pathway presented here was used to synthesize three unique aromatic chemicals as individual products of interest. This same approach can, however, be further applied to incorporate additional chemistries and target other products of interest. For instance, catechol has been extensively studied as an intermediate in the biosynthesis of *cis,cis*-muconate, a product of importance in bioplastics applications as it can serves as a precursor for the chemical synthesis of adipic acid (used to produce nylon-6,6 and polyurethanes) as well as terephthalic acid (an important monomer compound) (Curran et al. 2013a). The presented pathway could be readily extended from catechol to *cis,cis*-muconate via the additional co-expression of catechol-1,2-dioxygenase, as encoded, for example, by *catA* from *Acinetobacter baylyi* (Draths and Frost 1994, Curran et al. 2013b). This suggests that the presented pathway, which herein has enabled the biosynthesis of pHBA, PCA, and catechol, could

furthermore serve as a versatile platform for future studies in pathway engineering and biorefining.

3.4 Conclusion

The stepwise and individual production of pHBA, PCA, and catechol from glucose has been demonstrated through the engineering of a novel and modular biosynthetic pathway in *E. coli*. This study not only establishes a novel route to these aromatic building-block chemicals, but it also represents a sustainable platform for producing other useful products of industrial importance.

CHAPTER 4

EXPLORING STRATEGIES TO ENHANCE FLUX THROUGH THE SHIKIMIC ACID PATHWAY FOR IMPROVED PRODUCTION OF AROMATIC CHEMICALS

Abstract

Aromatic chemical products derived from precursors of the shikimic acid pathway are often produced at low titers and yields due to tight regulation of the shikimic acid pathway at several key points both through feedback regulation as well as precursor availability. Though strains of *E. coli* have been engineered to overcome feedback regulation, precursor availability, namely phospoenolpyruvate (PEP) and erythrose-4-phosphate (E4P), remains a limiting factor. Here, an *in silico* model applying elementary mode analysis (EMA) was used to identify and then examine the individual and combined effect of several novel strategies aimed at optimizing precursor stoichiometry as a means to maximize product-to-substrate (Yps) and product-to-biomass yields (Ypx). Said strategies uniquely explored the effects of both medium design and host engineering. Ultimately, a maximum theoretical product-substrate yield taking into account biomass accumulation (Yps+x) was identified in a co-fed culture of 70% D-xylose and 30% glycerol. Additional improvements were identified to increase the theoretical maximum specific productivity by 10% by further deletion of genes encoding the PEP consuming enzymes Ppc (PEP carboxylase) and PykA and PykF (pyruvate kinase).

4.1 Introduction

Biosynthesis of aromatic precursors in engineered microorganisms offers a renewable approach to producing commodity chemicals without relying on petroleum derived precursors. The shikimic acid pathway is an abundant source of aromatic precursors for the biosynthetic production of products of commercial interest, most notably styrene(McKenna and Nielsen 2011b), (*R*)- and (*S*)-styrene oxide(McKenna et al. 2013b), phenol(Thompson, Machas, and Nielsen 2016), catechol(Draths and Frost 1994, Pugh et al. 2014), benzaldehyde(Pugh et al. 2015), and benzyl alcohol(Pugh et al. 2015). However, in order for the biosynthesis of most aromatic chemicals to become economically feasible, the allocation of carbon and energy toward product formation must be improved. Currently, products derived via the shikimic acid pathway commonly suffer from low product yields, for example, catechol biosynthesis via chorismate produced in an engineered phenylalanine overproducing mutant *E. coli* NST74 achieved a maximum 14.3% of the theoretical maximum yield in a glucose fed culture(Pugh et al. 2014). Low yields of products derived from shikimic acid pathway precursors are likely a result of the numerous mechanisms of feedback regulation which limit native flux through the pathway. Formation of DHAP, the first metabolite of the shikimic acid pathway produced via the condensation of PEP and E4P, and chorismate, the key aromatic amino acid branch point, are tightly regulated through feedback regulation of phenylalanine, tyrosine, and tryptophan, as illustrated in Figure 4.1(Keseler et al. 2005).

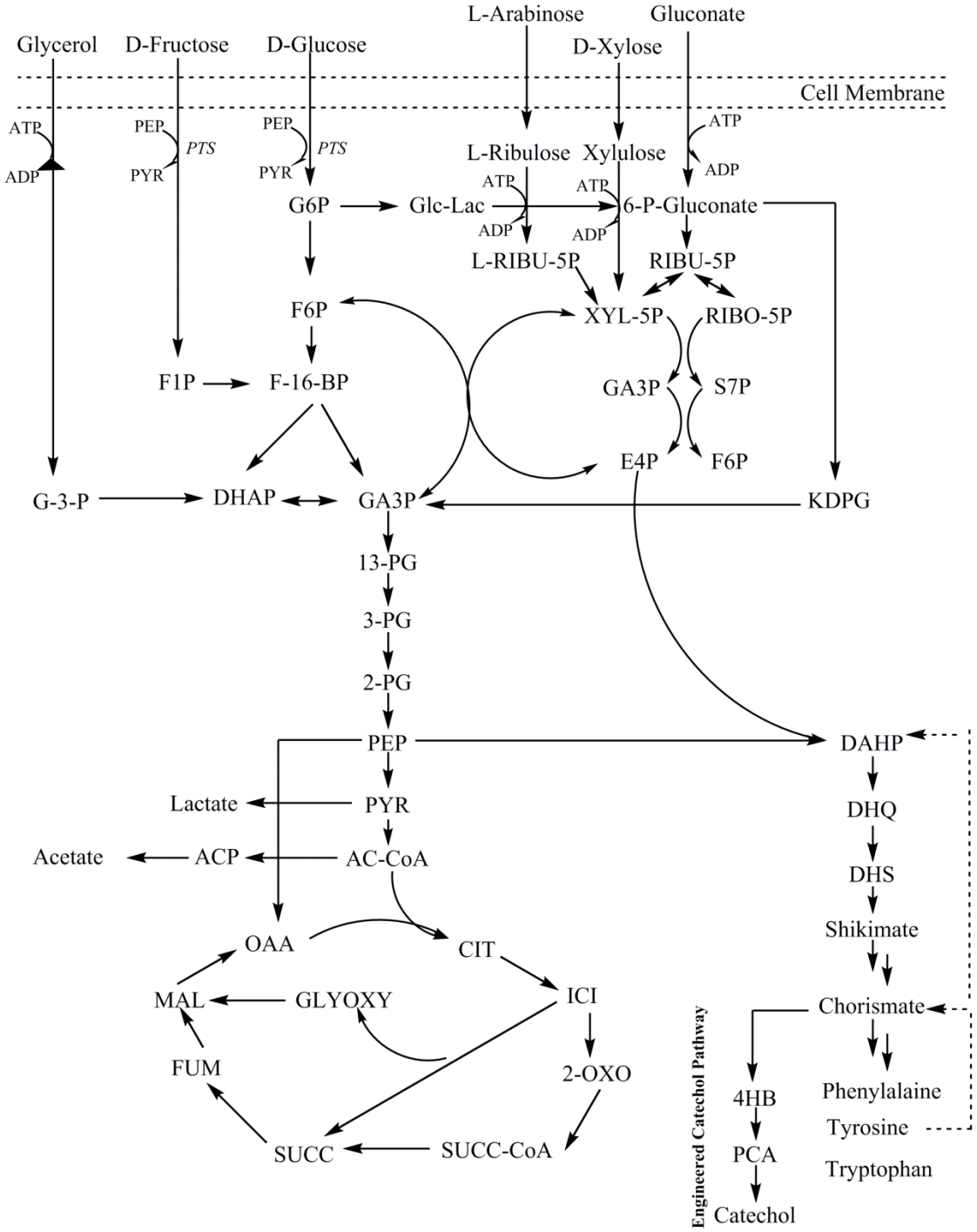

Figure 4.1. Metabolic pathways for the production of aromatics. Metabolic reactions for the consumption of glycerol, D-fructose, D-glucose, L-arabinose, D-xylose, and gluconate linked to central metabolism. The engineered pathway for the production of catechol is also illustrated. Multiple arrows indicate that several reactions are not illustrated while dashed arrows indicate feedback regulation.

While extensive work has been accomplished to reduce or remove the feedback regulation mechanisms through strain engineering, titers remain low due to poor flux through the shikimic acid pathway. Extensive research in increasing PEP and E4P bioavailability has been investigated; however, since PEP and E4P are key components of central metabolism, increasing their bioavailability remains a challenge. For example, PEP is an essential cofactor for the PEP:glucose phosphotransferase system (PTS) which is responsible for transporting and phosphorylating glucose to G6P. Researchers have constructed PTS^- mutants which, in theory, do not utilize PEP to phosphorylate glucose and thus may have increased bioavailability of PEP for the production of aromatic products; however, said PTS^- strains suffer from reduced growth rate ($\mu = 0.7\ hr^{-1}$ to $0.1\ hr^{-1}$) and are thus unsuitable as a robust host platform for aromatic chemical production(Martínez et al. 2008). In addition to the consumption of PEP via PTS, conversion of PEP to pyruvate is the last step of glycolysis prior to carbon flux entering the TCA cycle. While deletion of the responsible pyruvate kinase isozymes *pykA* and *pykF* has previously been shown to increase the product-to-biomass yield (Ypx) of the shikimic acid precursor DHAP by 3.4-fold, production of a non-natural aromatic product was not demonstrated in that study(Gosset, Yong-Xiao, and Draths 1996). Parallel routes toward increasing E4P have also been considered for increasing aromatic biosynthesis. For example, overexpression of the transketolase *tktAB*, which is directly responsible for production of E4P, has been shown to be an effective strategy for increasing flux through the shikimic acid pathway. A study by Draths *et al.* demonstrated a nearly 2-fold improvement in flux by overexpressing *tktAB*(Draths et al. 1992); however, McKenna *et al.* observed a noticeable reduction in growth rate when the enzyme was overexpressed

but also observed a similar increase in product yield(McKenna et al. 2013b). In *Saccharomyces cerevisiae*, which possesses a homologous aromatic biosynthesis pathway design to *E. coli*, Curran *et al.* overexpressed the transketolase *TKL1* and knocked out the glucose-6-phosphate dehydrogenase *ZWF1*, *zwf* in *E. coli*, to force flux from glycolysis into the pentose phosphate pathway through transketolase alone and ultimately observed a 24-fold improvement in titer achieving 141 mg/L of muconic acid(Curran et al. 2013b).

Thus far, nearly all approaches toward improving shikimic acid pathway flux have focused on overexpressing or knocking out key pathway enzymes to control flux through glycolysis, to improve PEP availability, or the pentose phosphate pathway, to improve E4P availability. Alternatively, researchers have considered carbon substrates other than D-glucose to control metabolic flux and improve product-to-substrate yields. For example, production of ethanol has been achieved using glycerol(Shams Yazdani and Gonzalez 2008, Yazdani and Gonzalez 2007, Dharmadi, Murarka, and Gonzalez 2006), D-fructose(Jain, Toran-Diaz, and Baratti 1985), L-arabinose(Becker and Boles 2003), or D-xylose(Qureshi et al. 2006) as the sole substrate. While the substrates glycerol and D-fructose enter metabolism through glycolysis, the pentose sugars L-arabinose and D-xylose as well as gluconate are metabolized via the pentose phosphate pathway. This represents a unique opportunity to potentially balance flux through glycolysis and the pentose phosphate pathway by co-feeding different carbon substrates. We hypothesize that increasing precursor availability of PEP and E4P through a co-feeding strategy of mixed carbon sources will increase achievable yields in engineered aromatic pathways. To test our hypothesis *in silico*, we utilized elementary flux mode analysis (EMA) to

calculate relevant yields in a mixed carbon feeding strategy for the biosynthesis of the aromatic product catechol; though, the methods and results discovered in this study would apply to other aromatic products derived from the shikimic acid pathway.

Previously, EMA has been used to evaluate metabolic networks and determine maximum theoretical product- and biomass-to-substrate yields(Trinh, Wlaschin, and Srienc 2009, Averesch and Krömer 2014). EMA is an efficient tool for analyzing flux through metabolic networks and the *in silico* identification/testing of novel strategies for optimizing product biosynthesis. EMA modelling utilizes the stoichiometry of all reactions composing a given metabolic network, as further confined by any the thermodynamic constraints of those reactions; reaction kinetics, however, not considered. As described by Trinh *et al.*, the conservation of mass of metabolites in a system of defined volume (i.e., a bacterial cell) can be described by Eq.1 where C is the concentration vector of the metabolite, r is the reaction rate vector, S is the stoichiometry matrix, and μ is the dilution rate of the change in volume of the system (i.e., a change in size of an individual cell)(Trinh, Wlaschin, and Srienc 2009). Within a cell, however, it may be assumed that i) the reaction rate is much greater than the dilution rate and therefore the contribution of volume changes is negligible, and ii) at steady state there is no net accumulation of the metabolite; therefore, Eq. 1 may be simplified to Eq. 2. Thermodynamically, due to the irreversibility of some reactions within the metabolic network, the reaction rate of an individual reaction (r_i) may be constrained to have a positive flux, as described by Eq. 3.

(1) $\frac{dC}{dt} = S \times r - \mu \times C$

(2) $S \times r = 0$

(3) $r_i = 0$

EMA uses Eq. 2 and 3 to calculate all possible unique solutions, also known as elementary modes (EMs).

4.2 Materials and Methods

4.2.1 Elementary flux mode analysis

Modelling was conducted using the EFMTool 4.7.1 (available at http://www.csb.ethz.ch/tools/software/efmtool.html) in MATLAB R2013a on a Hewlett Packard EliteBook at 2.60 GHz (Terzer and Stelling 2008). Each EFM represents a single, steady state flux distribution of a functional pathway within the metabolic network. Average computation time was approximately 31.7 s (N = 7).

4.2.2 Metabolic Network

The metabolic network was adapted from Averesch et al. (Averesch and Krömer 2014), and further modified to include additional metabolic reactions as compiled from literature (Keseler et al. 2005). The modified metabolic network developed in this study is thereby uniquely capable of analyzing the effects of different substrate feeding strategies, including the following carbon and energy sources as well as their mixtures: D-glucose, glycerol, L- and D-arabinose, D-xylose, gluconate, and D-fructose. The metabolic network includes glycolysis, the pentose phosphate pathway, the TCA cycle, biomass formation, the glyoxylate cycle, and the electron transport chain. Amino acid biosynthesis was not considered in the metabolic model. The biosynthetic production of catechol via chorismate(Pugh et al. 2014) was used as a model pathway for the EMA *in*

silico model as illustrated in Figure 4.1. Table 4.1 summarizes the reactions comprising the entire metabolic network developed and used in this study.

Table 4.1. Metabolic network comprising all reactions evaluated in the EMA model

Reaction	Function / Enzyme
'--> 1 Glucose + 0 Glycerol + 0 L-Arabinose + 0 D-Arabinose + 0 Xylose + 0 Gluconate + 0 Fructose'	Mixed carbon feed (subscripts modified to specify substrate)
'Product_ex -->'	Product out
'Biomass_ex -->'	Biomass_ex out
'CO2 -->'	Carbon dioxide out
'Lactate -->'	Lactate out
'Ethanol -->'	Ethanol out
'Formate -->'	Formate out
'Acetate -->'	Acetate out
'SUCC_ex -->'	Succinate out
'--> P_ex'	Phosphate in
'--> O2'	Oxygen uptake
'--> NH3'	Ammonia uptake
'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
'Catechol + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
'Glycerol + ATP <--> Glycerol-3-P + ADP'	Glycerol kinase (EC 2.7.1.30)
'Glycerol-3-P + UQ8 --> DHAP + UQ8H2'	Glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)

'L-Arabinose --> L-Ribulose'	L-arabiose isomerase (EC 5.3.1.4)
'L-Ribulose + ATP --> L-RIBU-5P + ADP'	L-ribulokinase (EC 2.7.1.16)
'L-RIBU-5P --> XYL-5P'	L-ribulose 5-phosphate 4-epimerase (EC 5.1.3.4)
'D-Arabinose --> D-Ribulose'	D-arabinose isomerase (EC 5.3.1.3)
'D-Ribulose + ATP --> RIBU-1P + ADP'	D-ribulokinase
'RIBU-1P --> DHAP + Glyald'	D-ribulose-phosphate aldolase
'Glyald + NAD --> Glycolate + NADH'	Glycoaldehyde dehydrogenase (EC 1.2.1.21)
'2 Glycolate --> 2 Glyoxylate'	Glycolate oxidase (EC 1.1.99.14)
'2 Glyoxylate --> TART'	Glyoxylate carboligase (EC 4.1.1.47)
'TART + NADH --> Glycerate + NAD'	Tartronate semialdehyde reductase
'Glycerate + ATP --> 2-PG + ADP'	Glycerate kinase (EC 2.7.1.165)
'Fructose + PEP --> F1P + PYR'	PTS
'F1P + ATP --> F-16-BP + ADP'	1-Phosphofructokinase (EC 2.7.1.56)
'Xylose --> Xylulose'	Xylose isomerase (EC 5.3.1.5)
'Xylulose + ATP <--> XYL-5P + ADP'	Xylulokinase (EC 2.7.1.17)
'Gluconate + ATP <--> 6-P-Gluconate + ADP'	Gluconate kinase (EC 2.7.1.12)
'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
'S7P + GA3P <--> E4P + F6P'	Transaldolase (EC 2.2.1.2)
'F6P + GA3P <--> E4P + XYL-5P'	Transketolase (EC 2.2.1.1)
'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
'FUM <--> MAL'	Fumarase (EC 4.2.1.2)

'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
'PEP + CO2 --> OAA + P'	Ppc: Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
'Formate --> CO2'	Formate hydrogenlyase
'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
'ATP --> ADP + P'	ATP hydrolysis
'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
'E4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)

'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
'Chorismate --> 4-Hydroxybenzoate + PYR'	Chorismate lyase (EC 4.1.3.40)
'4-Hydroxybenzoate + NADPH --> Protocatechuate + NADP'	4-hydroxybenzoate 3-monooxygenase (EC 1.14.13.2/EC 1.14.13.33)
'Protocatechuate --> Catechol + CO2'	Protocatechuate decarboxylase (EC 4.1.1.63)

4.2.3 Yield Calculations

Both product and biomass yields may be determined from the *in silico* EMA model including the product-to-substrate yield (Yps; Eq. 4) and the biomass-to-substrate yield (Yxs; Eq. 5).These yields describe the maximum achievable flux toward either product or biomass formation based on the consumption of the substrate. From a ratio of the product and biomass yields, the product-to-biomass yield (Ypx) may be determined as described by Eq. 6. The product-to-biomass yield describes the productivity of the culture on a biomass basis and can be used to understand the tradeoff between carbon flux toward product vs biomass formation. An alternative means of understanding the role of biomass formation in the model is to use the product-to-substrate plus biomass yield (Yps+x; Eq. 7). The product-to-substrate plus biomass yield is the maximum theoretical yield which takes into account the necessary flow of carbon toward biomass formation to support cell growth and reproduction. To further understand how the metabolic changes effect product formation taking into account biomass formation, the percent maximum achievable product yield may be calculated by taking a ratio of Eq. 4 and Eq. 7 (Yps^{max}; Eq. 8).

(4) $$Y_{PS} = \frac{mol\ P_{produced}}{mol\ S_{consumed}}$$

(5) $$Y_{XS} = \frac{mol\ X_{produced}}{mol\ S_{consumed}}$$

(6) $$Y_{PX} = \frac{Y_{PS}}{Y_{XS}} = \frac{mol\ P_{produced}}{mol\ X_{produced}}$$

(7) $$Y_{PS+X} = \frac{mol\ P_{produced}}{mol\ S_{consumed} + mol\ X_{produced}}$$

(8) $$Y_{PS}^{\max} = \frac{Y_{PS}}{Y_{PS+X}}$$

4.3 Results and Discussion

4.3.1 Evaluation of different carbon sources

Both PTS (including D-glucose and D-fructose) and non-PTS carbon sources (including glycerol, D-xylose, L- and D-arabinose, and gluconate) were evaluated in the *in silico* EMA model to elucidate how the different carbon sources might affect catechol (henceforth referred to as 'product') and biomass yields. As illustrated in Table 4.2, glycerol provided the highest maximum theoretical product-to-substrate and biomass-to-substrate yields (Yps = 0.125 mol P/mol C and Yxs = 0.252 mol X/mol C) when compared on a per carbon basis of the carbon sources tested. In contrast to the other compounds tested, glycerol is the most highly reduced substrates and enters central metabolism closer to the TCA cycle. Most notably, glycerol is not a PTS sugar and is metabolized by phosphorylation using ATP as a phosphate donor as opposed to consuming PEP. However, while glycerol supported the highest theoretical product-to-substrate and biomass-to-substrate yields, it also suffered from the lowest product-to-biomass yield Ypx = 0.454 mol P/mol X. The lower product-to-biomass yield resulted in

a lower maximum theoretical product-to-substrate yield plus biomass, Yps+x = 0.115 mol P/mol C achieving only 91.4% of the theoretical maximum (Yps^{max}).

D-Glucose, the preferred carbon substrate of *E. coli*, supported the highest product-to-biomass yield Ypx = 0.956 mol P/mol X and was able to achieve 96% of the theoretical product-to-substrate yield plus biomass (Yps+x = 0.113 mol P/mol C). The higher efficiency of D-glucose as a carbon substrate is likely due to the fact that, unlike glycerol, D-glucose is efficiently metabolized in a manner that results in the production of both of the required shikimic acid pathway precursors, PEP and E4P. More specifically, as illustrated in Figure 4.1, upon transport and phosphorylation of D-glucose to D-glucopyranose 6-phosphate (G6P), flux proceeds through both glycolysis, resulting in the production of PEP, as well as the pentose phosphate pathway, resulting in the production of E4P. In contrast, D-fructose, for example, is also metabolized through glycolysis, however, it enters the pathway downstream of the first committed step to the pentose phosphate pathway (Zwf; glucose 6-phosphate 1-dehydrogenase). Thus, with consequently lower flux to E4P, fructose supports a lower maximum theoretical product-to-substrate yield plus biomass (Yps+x = 0.107 mol P/mol C; Yps^{max} = 92.6%) and product-to-biomass yield (Ypx = 0.912 mol P/mol X). This observation suggests that a proper balance must be achieved between flux through both the pentose phosphate pathway (for generation of E4P) as well as glycolysis (to provide necessary essential biomass metabolites and notably PEP) to achieve high yields of shikimic acid precursors for aromatic chemical production.

Table 4.2. Carbon source evaluation of maximum theoretical yield coefficients

	Yps (mol P/ mol C)	Yps (g P/g S)	Yxs (mol X/ mol C)	Yps+x (mol P/ mol C)	Yps+x (%max)	Ypx (mol P/ mol X)
D-Glucose	0.118	0.431	0.118	0.113	96.0	0.956
D-Fructose	0.116	0.424	0.117	0.107	92.6	0.912
Glycerol	0.125	0.449	0.252	0.115	91.4	0.454
D-Xylose	0.118	0.431	0.142	0.113	95.7	0.795
L-Arabinose	0.118	0.431	0.142	0.112	95.5	0.792
D-Arabinose	0.102	0.375	0.124	0.099	96.7	0.802
Gluconate	0.108	0.363	0.108	0.097	90.3	0.900

Interestingly, the pentose sugars D-xylose and L-arabinose have comparable maximum theoretical product-to-substrate plus biomass yields to D-glucose with D-xylose having the highest, Yps+x = 0.113 mol P/mol C. The high achievable theoretical yields of the pentose sugars is possible since carbon can link directly to glycolysis and central metabolism either through the transketolase reaction (TktAB), which produces E4P as a product of the reaction, or via the Entner-Doudoroff pathway which connects 6-P-Gluconate to GA3P via KDPG (Edd; phosphogluconate dehydratase and EdA; 2-Keto-3-deoxygluconate-6-phosphate aldolase). These results demonstrate that high product yields can theoretically be achieved when carbon is fed directly to the pentose phosphate pathway as a result of enhanced E4P bioavailability as well as close linkage to glycolysis and central metabolism. Though D-xylose has a high product-to-substrate yield, it has nearly half the theoretical biomass-to-substrate yield as glycerol; thus, it was hypothesized that co-feeding glycerol and D-xylose may be able to achieve comparable, if not higher, product-to-substrate plus biomass yields than feeding D-glucose alone.

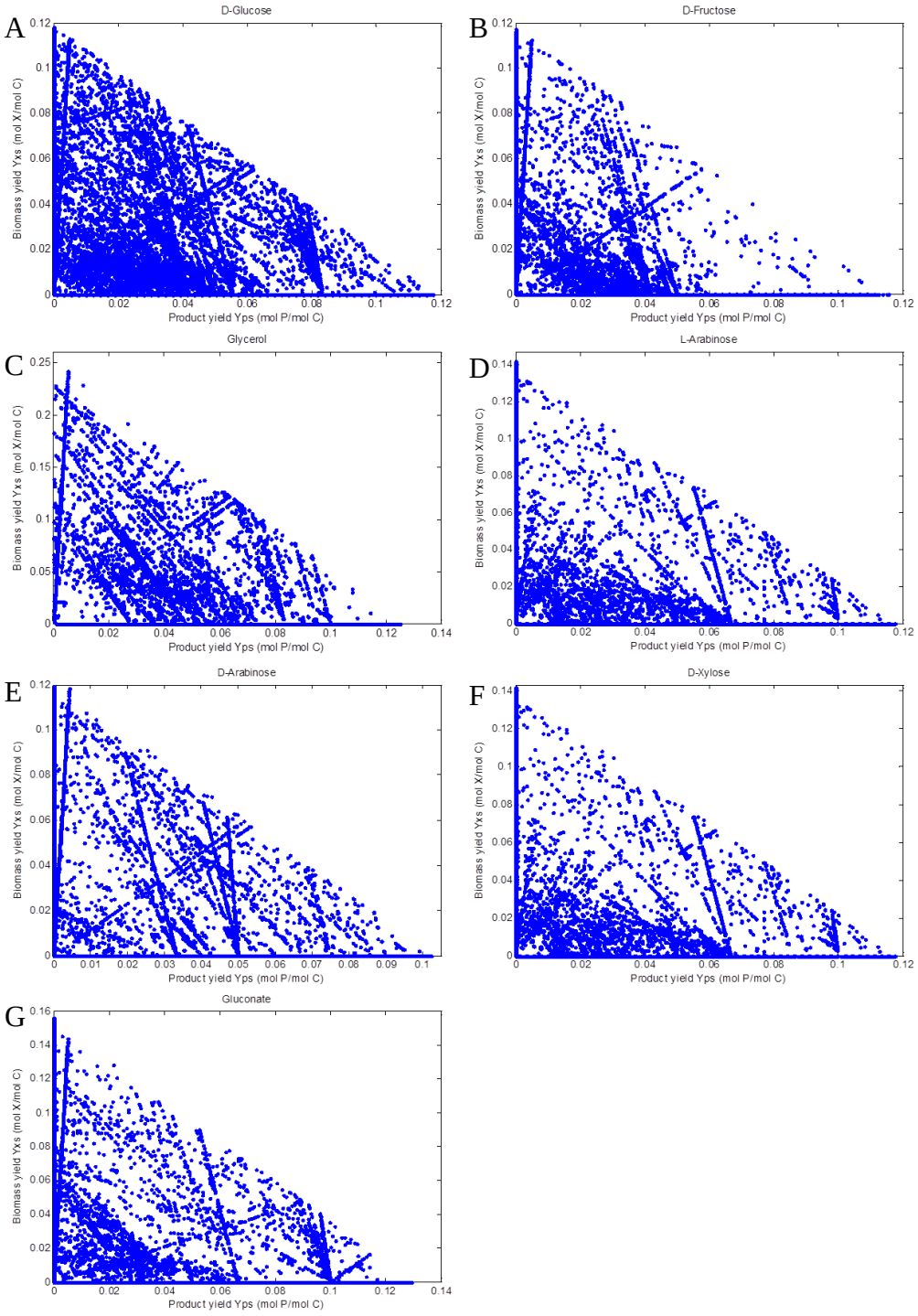

Figure 4.2. EFM distribution of biomass and product yields for the E. coli pathway networks for the production of catechol fed different carbon substrates. EFM plots include D-glucose (A), D-fructose (B), glycerol (C), L-arabinose (D), D-arabinose (E), D-xylose (F), and gluconate (G).

4.3.2 Optimization of product yield via co-feeding strategy

For a co-feeding strategy to be effective in *E. coli*, the carbon substrates should, ideally, not compete for cofactor-mediated transport and phosphorylation (i.e., PTS sugars) nor should their exist enzymatic repression mechanisms. While D-glucose and glycerol co-feeding was considered to optimize product and biomass yields, co-feeding *E. coli* D-glucose with other carbon sources has often been unsuccessful as *E. coli* will preferentially consume D-glucose before consuming other carbon sources, a mechanism commonly referred to as carbon catabolite repression (CCR) (Brückner and Titgemeyer 2002, Saier and Roseman 1976). On the other hand, no known repression or competition pathways for co-feeding glycerol and D-xylose exist; therefore, these carbon sources were considered in a co-fed *in silico* EMA model.

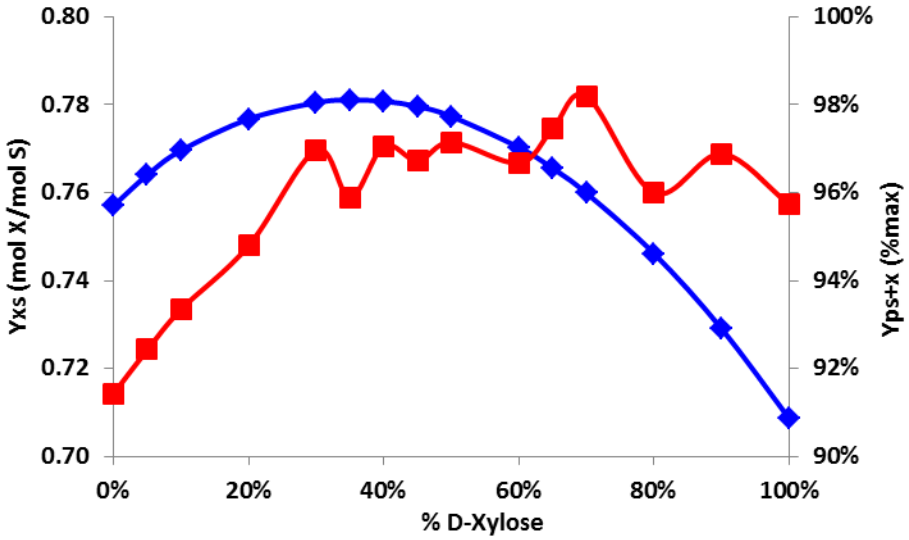

Figure 4.3. In silico EMA model yields for co-fed D-xylose and glycerol cultures. The biomass yield Yxs is plotted in blue diamonds while the maximum product-substrate plus biomass yield Yps+x is plotted in red squares.

A range of D-xylose and glycerol ratios were considered in the EMA model. As illustrated in Figure 4.3, a maximum biomass-to-substrate yield Yxs = 0.78 mol X/mol S

was achieved at 35% D-xylose and 65% glycerol. As previously discussed, glycerol fed cultures have the highest theoretical biomass yields on a per carbon basis therefore it was expected that the maximum biomass yield was achieved when fed a higher ratio of glycerol. Conversely, at a higher D-xylose ratio of 70% D-xylose and 30% glycerol, the maximum product-substrate plus biomass yield Yps+x = 0.516 mol P/mol S was achieved at 98.2% of the theoretical maximum. On a per carbon basis, co-feeding 70% D-xylose and 30% glycerol can theoretically achieve a higher product-to-substrate yield (Yps+x = 0.117 mol P/mol C) than glucose alone (Yps+x = 0.113 mol P/molC). These results can be interpreted visually in the EFM distribution plots in Figure 4.4. As illustrated, in the 100% glycerol case (A), the biomass yield is favored over product yield indicated by the higher number of EMs clustering in its favor. On the other hand in the 100% D-xylose case (B), EMs cluster more in favor of the product yield. As expected, in the co-fed cases, the EM distribution clusters either in favor of biomass yield in the case of 35% D-xylose ad 65% glycerol (C), or a more balanced distribution between biomass and product yield like the 70% D-xylose and 30% glycerol case (D).

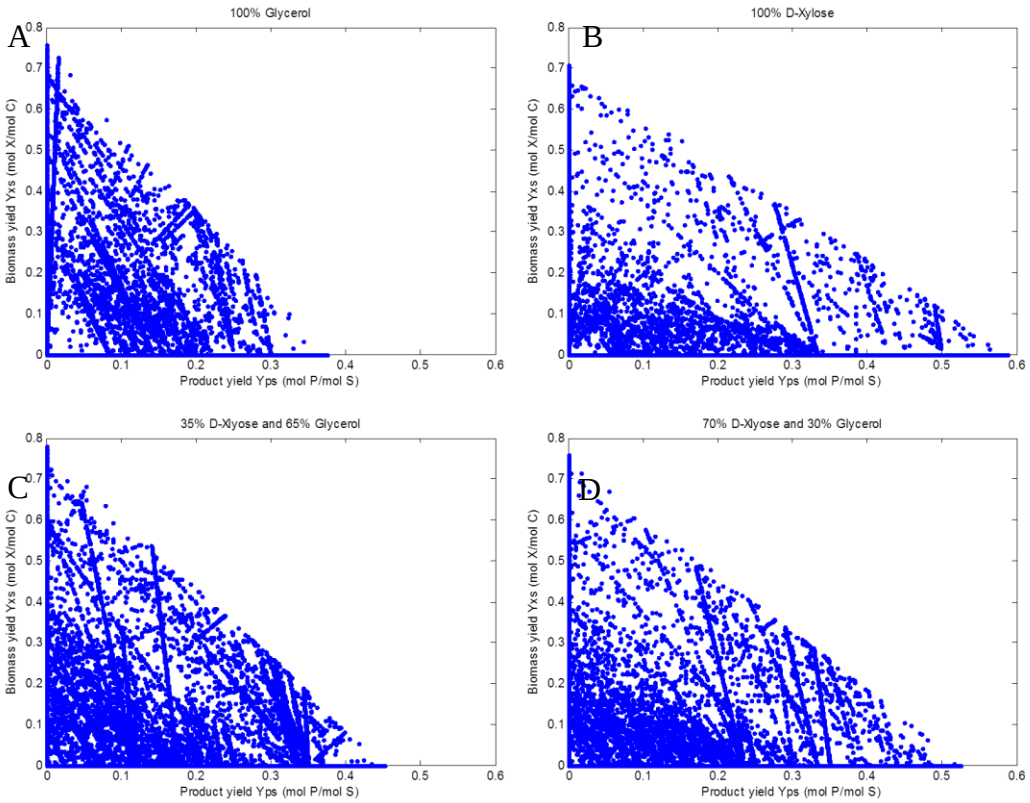

Figure 4.4. EFM distribution of biomass and product yields for the E. coli pathway networks. This includes the production of catechol co-fed D-xylose and glycerol. EFM plots include cultures fed 100% glycerol, 100% D-xylose (B), 35% D-xylose and 65% glycerol (C), and 70% D-xylose and 30% glycerol (D).

It is hypothesized that a higher theoretical product-to-substrate yield was achieved in the co-feeding strategy as a result of carbon sourced from D-xylose being directly metabolized in the pentose phosphate pathway to yield E4P while carbon sourced from glycerol is metabolized via glycolysis to yield PEP. To test this hypothesis and decouple the effects of co-feeding carbon sourced from D-xylose and glycerol, the transketolase reaction (EC 2.2.1.1; E4P + XYL-5-P = F6P + GA3P), which is the key reversible link between glycolysis and the pentose phosphate pathway and the primary route toward E4P production, was removed from the metabolic network in the EMA model to simulate a *ΔtktA* knockout mutant. In the *tktA* mutant, E4P production is still achievable via the

transaldolase reaction (EC 2.2.1.2, S7P + GA3P = E4P + F6P). As described in Table 4.3, in a glucose fed culture model, deletion of *tktA* has a significant impact to the maximum theoretical product-to-substrate yield reducing Yps = 0.118 mol P/mol C in the wild-type (WT) to Yps = 0.111 mol P/mol C in the *ΔtktA* mutant. Since the transketolase reaction is the primary mechanism for E4P biosynthesis and carbon flux to reversibly enter the pentose phosphate pathway from glycolysis, it was expected that deletion of this pathway would have a negative impact on maximum product yields. However, since carbon from glucose can enter the pentose phosphate pathway via the glucose 6-phosphate 1-dehydrogenase reaction (catalyzed by Zwf) to generate necessary biomass precursors and potentially re-enter glycolysis via the metabolites 6-P-gluconate and KDPG, the maximum theoretical biomass-to-substrate yield remained unchanged. The *tktA* knockout was then tested in the 70% D-xylose and 30% glycerol co-fed model and interestingly, no impact to the maximum theoretical product-to-substrate yield was observed compared to the wild-type. This confirms the hypothesis that the co-feeding strategy is tuned to drive flux toward product formation by feeding carbon directly to the necessary pathways to generate the shikimic acid pathway precursors E4P and PEP. However, unlike the glucose fed culture model, the co-fed model did observe a reduction in biomass-to-substrate yield changing from Yxs = 0.173 mol P/mol C in the WT to Yxs = 0.168 in the *ΔtktA* mutant. This suggests that the transketolase reaction is necessary to generate biomass forming metabolites in the co-fed model. By knocking out *tktA* in the EMA model, the pentose phosphate pathway and glycolysis were decoupled to demonstrate that the increase in product-to-substrate yield in the co-feeding strategy was a result of directly feeding carbon to the pathways necessary for E4P and PEP production.

Table 4.3: Role of transketolase on product and biomass yields

Carbon Source	Genotype	Yps (mol P/mol C)	Yxs (mol X/mol C)
Glucose	WT	0.118	0.118
Glucose	ΔtktA	0.111	0.118
70% D-Xylose + 30% Glycerol	WT	0.119	0.173
70% D-Xylose + 30% Glycerol	ΔtktA	0.119	0.168

4.3.3 Increasing product-to-biomass yield via strategic knock-out approach

Deletion of the PEP consuming pyruvate kinases PykA and PykF have previously been shown to increase the product-to-biomass yield of the shikimic acid pathway metabolite DHAP (Gosset, Yong-Xiao, and Draths 1996). Deletion of *pykAF* was tested in the EMA *in silico* model and the deletions had a marginal effect (<1% increase) on the product-to-biomass yield of a 100% D-glucose fed culture. In addition to the pyruvate kinase isozymes, deletion of phosphoenolpyruvate carboxylase (*ppc*) was also evaluated for its effect on specific productivity; however, no effect was observed in the wild-type host or the *ppc* and *pykAF* triple knock-out mutant. The developed metabolic network was accordingly modified to evaluate the individual and combined effect of *pykAF* and *ppc* knockouts in cultures co-fed with a 70% D-xylose and 30% glycerol mixture. While deletion of *ppc* alone had no effect, deletion of both pyruvate kinase isozymes (*pykAF*) provided a ~7% increase in the product-to-biomass yield, increasing from Ypx = 0.679 mol P/mol X in the wild-type background to Ypx = 0.711 mol P/mol X in the mutant. Interestingly, while *ppc* had no effect as a single knockout, the Δ*ppc* Δ*pykAF* mutant provided a further increase in product-to-biomass yield, reaching Ypx = 0.741 mol P/mol X or a net increase of 9% over the wild-type background.

Table 4.4: Effect of strategic knock-outs on maximum yield coefficients

Genotype	Carbon Feed Strategy	Yps (mol P/ mol C)	Yxs (mol X/ mol C)	Yps+x (mol P/ mol C)	Yps+x (%max)	Ypx (mol P/ mol X)
WT	100% D-Glucose	0.118	0.118	0.113	96.0	0.996
ΔpykAF	100% D-Glucose	0.118	0.117	0.113	96.0	1.003
Δppc	100% D-Glucose	0.118	0.118	0.113	96.0	0.996
ΔppcΔpykAF	100% D-Glucose	0.118	0.117	0.113	96.0	1.003
WT	70% D-Xylose + 30% Glycerol	0.119	0.173	0.117	98.2	0.679
ΔpykAF	70% D-Xylose + 30% Glycerol	0.119	0.165	0.117	98.2	0.711
Δppc	70% D-Xylose + 30% Glycerol	0.119	0.173	0.117	98.2	0.679
ΔppcΔpykAF	70% D-Xylose + 30% Glycerol	0.119	0.158	0.117	98.2	0.741
ΔppcΔpykAF	73% D-Xylose + 27% Glycerol	0.116	0.151	0.114	98.6	0.757

The effect of substrate ratio for D-xylose/glycerol mixtures was subsequently reevaluated for a host deficient in the two focal PEP consuming pathways (i.e., the Δ*ppc* Δ*pykAF* mutant). As illustrated in Figure 4.5, in this case the maximum theoretical product-substrate plus biomass yield was achieved at 73% D-xylose and 27% glycerol, Yps+x = 0.693 mol P/mol C (or 0.114 mol P/mol C) achieving 98.6% of the theoretical maximum (Yps^{max}). Overall, the maximum biomass yield shifted down from Yxs = 0.780 mol P/mol S (or 0.173 mol P/mol C) in the wild-type host to Yxs = 0.711 mol P/mol S (or 0.151 mol P/mol C) in the Δ*ppc* Δ*pykAF* mutant. Interestingly, the new biomass yield feed ratio optimum also shifted from 35% D-xylose to 40% D-xylose favoring a higher ratio of glycerol. The overall higher shift in D-xylose utilization for product and biomass yields indicates that the deletion of PEP consuming pathways increased PEP availability and E4P became the rate limiting metabolite. By increasing the D-xylose to glycerol ratio in the Δ*ppc* Δ*pykAF* mutant, the available pool of PEP and E4P was balanced leading to

new optimums for product and biomass yields. The effect on the yields can be observed in the EFM distribution illustrated in Figure 4.6.

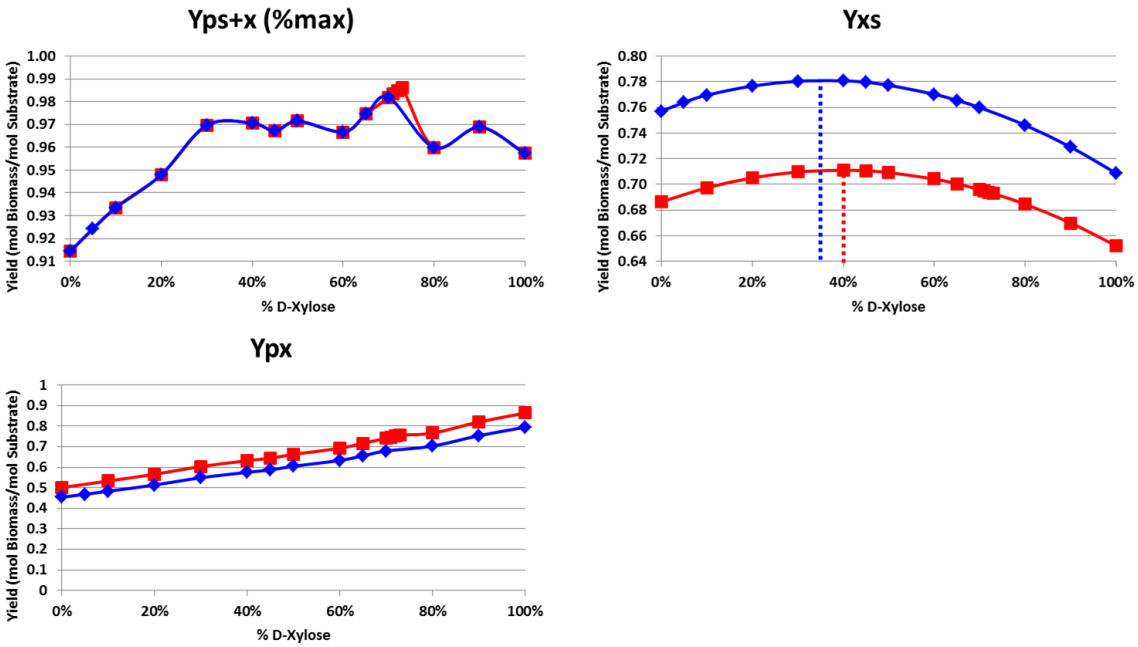

Figure 4.5. In silico EMA model yields for co-fed D-xylose and glycerol cultures in a wild-type and ΔppcΔpykAF host background. The wild-type host is plotted in blue diamonds while the ΔppcΔpykAF mutant is plotted in red squares. Plots include the maximum produce-substrate plux biomass yield Yps+x (top left), the biomass-substrate yield Yxs (top right), and the specific productivity Ypx (bottom left).

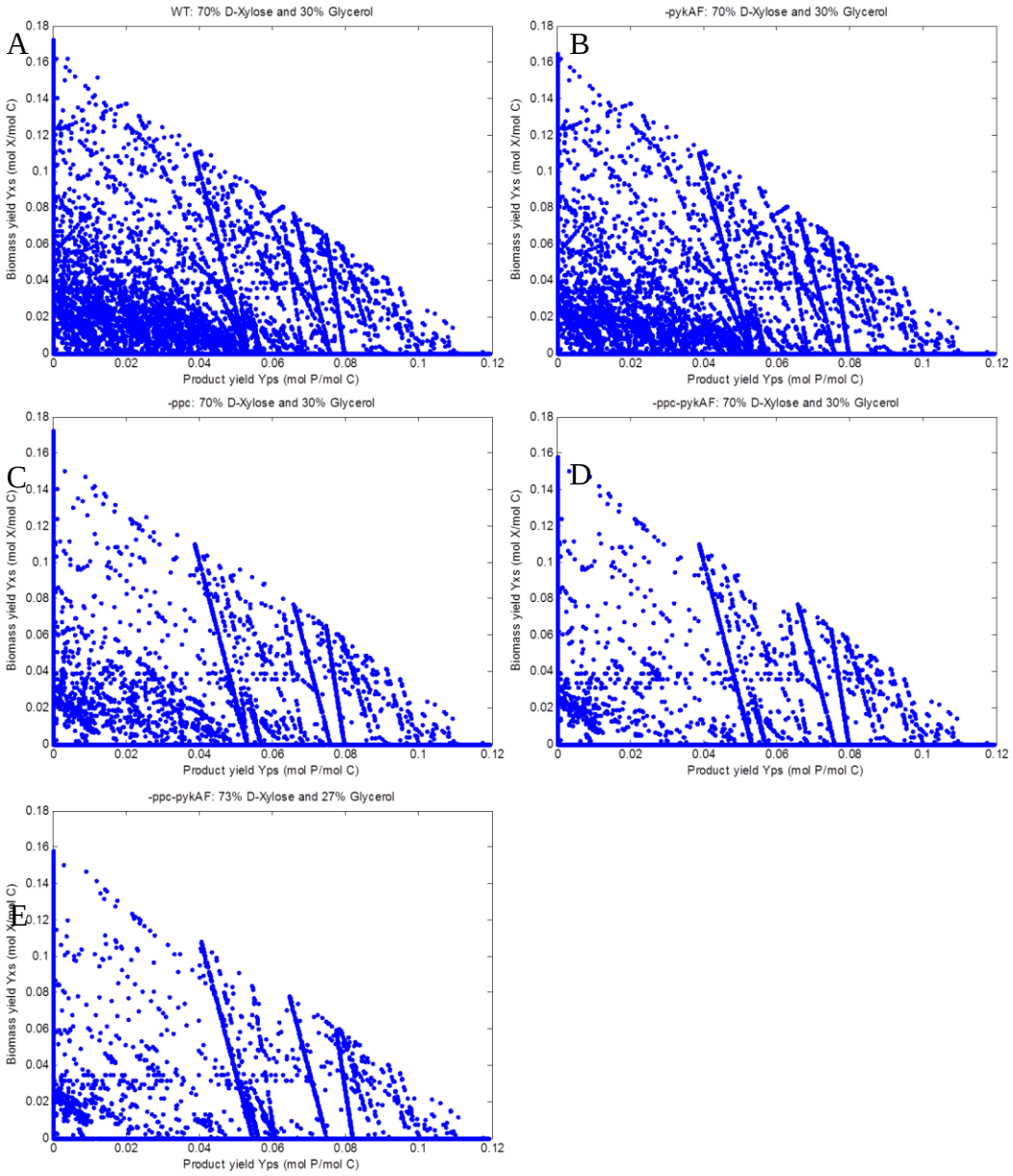

Figure 4.6. EFM distribution of biomass and product yields for the E. coli mutants to increase bioavailability of PEP. EFM plots include WT: 70% D-xylose and 30% glycerol (A), ΔpykAF: 70% D-xylose and 30% glycerol (B), Δppc: 70% D-xylose and 30% glycerol (C), ΔppcΔpykAF: 70% D-xylose and 30% glycerol (D), and ΔppcΔpykAF: 73% D-xylose and 27% glycerol (E).

4.4 Conclusion

In this study, we illustrated how EMA can be used to characterize mixed-carbon feeds and tune metabolic flux to achieve an optimized ratio of precursor metabolites. By

feeding carbon directly into the pentose phosphate pathway via D-xylose, E4P bioavailability was increased, while co-feeding glycerol promoted the formation of PEP and biomass. EMA was further utilized to identify deletions which enhanced PEP bioavailability and maximum achievable product-to-substrate plus biomass yields. Though catechol was used as a surrogate for the model, the identified approaches could be applied to any aromatic product derived from the shikimic acid pathway. Overall, we demonstrated that EMA is an efficient means of identifying stoichiometric optimums to potentially improve product yields.

CHAPTER 5

DISCUSSION AND FUTURE WORK

Abstract

The novel biosynthetic pathways presented in these works highlight new contributions to the field of metabolic engineering. However, further improvements are necessary to improve flux through the pathways to ultimately achieve higher titers and yields. Future works are discussed including the use of protein scaffolds to co-localize enzymes to overcome pathway imbalances for the biosynthesis of benzyl alcohol as well as an evaluation of alternative catechol production pathways.

5.1 Introduction

These works have demonstrated the biosynthetic production of benzyl alcohol and benzaldehyde, for the first time, as well as a novel route toward the biosynthetic production of *para*-hydroxybenzoate, protocatechuate, and catechol. Unlike the production of other aromatic products which are limited by product imposed toxicity, such as styrene for example (McKenna and Nielsen 2011b), these fine chemicals are limited by pathway flux and precursor availability. From the *in silico* EMA model, we identified that co-feeding glycerol and D-xylose in addition to targeted host knockouts may be able to enhance precursor availability, namely phosphoenolpyruvate and erythrose-4-phosphate, by feeding carbon directly to glycolysis and the pentose phosphate pathway. While increasing precursor availability may enhance flux through the engineered pathways, parallel approaches should be considered in future works in order to identify the optimal routes toward achieving high product titers and yields. For example, in the case of the benzyl alcohol pathway, the first heterologous enzymatic reaction (HmaS, hydroxymandelate synthase) has been identified to be the rate limiting step due to poor substrate affinity. Recent advances in protein scaffolds however have demonstrated that flux imbalances could be overcome by tuning the relative ratio of co-localized pathway enzymes. On the other hand, unlike the benzyl alcohol pathway which has only one identified route for synthesis, a recent renewal of interest in catechol and muconic acid has led researchers to identify new and novel routes toward the biosynthesis of catechol. In this case, a thorough evaluation of the various routes should be conducted to identify the pathway capable of generating the highest titers and yields.

5.2 Using protein scaffold to overcome pathway flux imbalance to improve the biosynthetic production of benzyl alcohol

As previously described in Chapter 2, flux entering the benzaldehyde and benzyl alcohol pathway is limited by the poor activity of hydroxymandelate synthase (HmaS) which has a 70-fold lower affinity for phenylpyruvate (K_m 0.45± 0.04 mM) versus its preferred substrate 4-hydroxyphenylpyruvate (K_m 6.5 ± 0.8 μM) (He, Conrad, and Moran 2010). Since no other isoenzymes specific for (*S*)-mandelate are known, HmaS remains the only suitable candidate for the biosynthetic pathway at this time. For the biosynthetic production of these products to be economically feasible, titer and product yields should be enhanced beyond its current state. For this goal to be realized, the poor flux entering the engineered pathway must be overcome; a feat which may successfully be accomplished utilizing a recent development in the field of synthetic biology, protein scaffolds.

To address problems associated with plasmid based genetic engineering, such as metabolic burden and flux imbalances, various strategies have been developed (with mixed success), including strategies to modulate or improve transcription and translation (e.g., plasmid copy number(Jones, Kim, and Keasling 2000), promoter strength(Hammer, Mijakovic, and Jensen 2006), ribosomal binding site strength(Salis, Mirsky, and Voigt 2009), and codon utilization(Tyo, Alper, and Stephanopoulos 2007)), as well as strategies to directly control flux of a particular reaction (e.g., gene circuits like toggle switches(Gardner, Cantor, and Collins 2000, Atkinson et al. 2003), logic gates(Hasty, McMillen, and Collins 2002, Goni-Moreno and Amos 2012), or even quorum sensing(Zhu et al. 2002, Fuqua, Parsek, and Greenberg 2001)). One of the more

prominent and developing techniques in synthetic biology is protein scaffolds (Dueber et al. 2009).

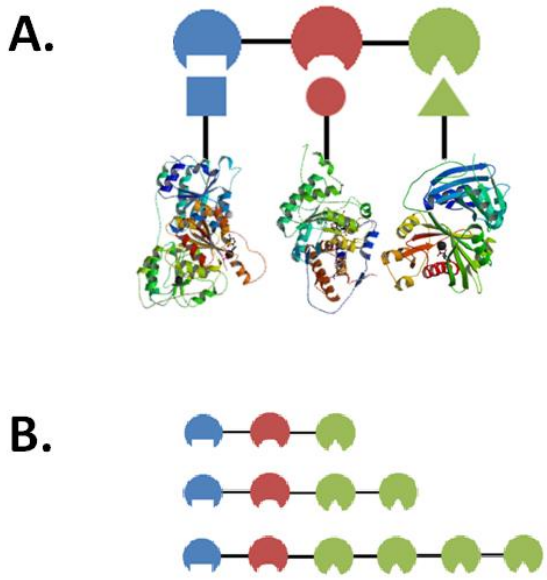

Figure 5.1 Modular control of flux via synthetic protein scaffolds. (a). Illustrates how synthetic domains (circles with cut shapes) act as a scaffold to which enzymes with ligands (square, circle, and triangle) bind and co-localize. Enzyme structures represent MdlC(Hasson et al. 1998), MdlB(Sukumar et al. 2001), and HmaS(Brownlee et al. 2008) (b). Illustrates how the domains of the scaffold can be 'tuned', three configuration in which the third domain is varied is shown.

Protein scaffolds offer a means of substrate channeling via *in vivo* co-localization of pathway enzymes. This regulatory control is accomplished by synthesizing a scaffold protein which possess 'domains' tethered via flexible poly-glycine-serine linkers. To date, the only three domains which have been characterized for this purpose include the rat derived GTPase binding domain (GBD), the mouse derived actin polymerization switch N-WASP (SH3), and the mouse derived PSD95/DIgA/Zo-1 domain from the adaptor protein syntrophin (PDZ)(Dueber et al. 2009). The domains (GBD, SH3, and PDZ) act as

a docking station, as illustrated in Fig. 5.1, which can be tuned by varying the number of a particular domain in the scaffold. Each domain has a strong affinity for a particular ligand and the strength of the binding interaction may be quantified by the dissociation constant (K_d). The domains GBD, SH3, and PDZ have an affinity for their ligands of K_d = 1 µM, 0.1 µM, and 8 µM, respectively(Dueber et al. 2009). To co-localize enzymes, via protein scaffolds, the ligands are tethered to the enzyme on either the N or C-terminal separated by a flexible poly-glycine-serine linker. Synthetic protein scaffolds have proven effective in addressing flux imbalance by circumventing unfavorable equlibria and kinetics while also reducing the metabolic burden associated with over-expression of enzymes in a heterologous host (Dueber et al. 2009). In addition, co-localization minimizes potentially toxic intermediate metabolite accumulation as well as reduces the loss of intermediates through diffusion, stability, or competing pathways. Thus far, protein scaffolds have successfully been applied to achieve a 77-fold improvement in the production of mevalonate(Dueber et al. 2009), a 3-fold improvement in glucarate(Moon et al. 2010), and a 3-fold improvement in butyrate(Baek et al. 2013).

To test the efficacy of protein scaffolds for improving pathway flux, three scaffold configurations were constructed and evaluated, described in Appendix A.1. Scaffold configurations included pTET-1:1:1, pTET-1:1:2, and pTET-1:1:4 representing GBD-SH3-PDZ, GBD-SH3-PDZ-PDZ, GBD-SH3-PDZ-PDZ-PDZ-PDZ, respectively. Said configurations were designed in a manner that holds the number of GBD and SH3 domains constant (one of domain each) while varying the number of PDZ domains (one, two, or four). Then, by creating fusions between HmaS, MdlB, and MdlC and those ligands with affinity for PDZ, SH3, and GBD, respectively, multi-protein complexes with

tunable stoichiometry of pathway enzymes could be created. More specifically, by varying the relative number of co-localized subunits of HmaS, the net activity of this rate-limiting step could be increased relative to other pathway enzymes to balance and improve pathway flux.

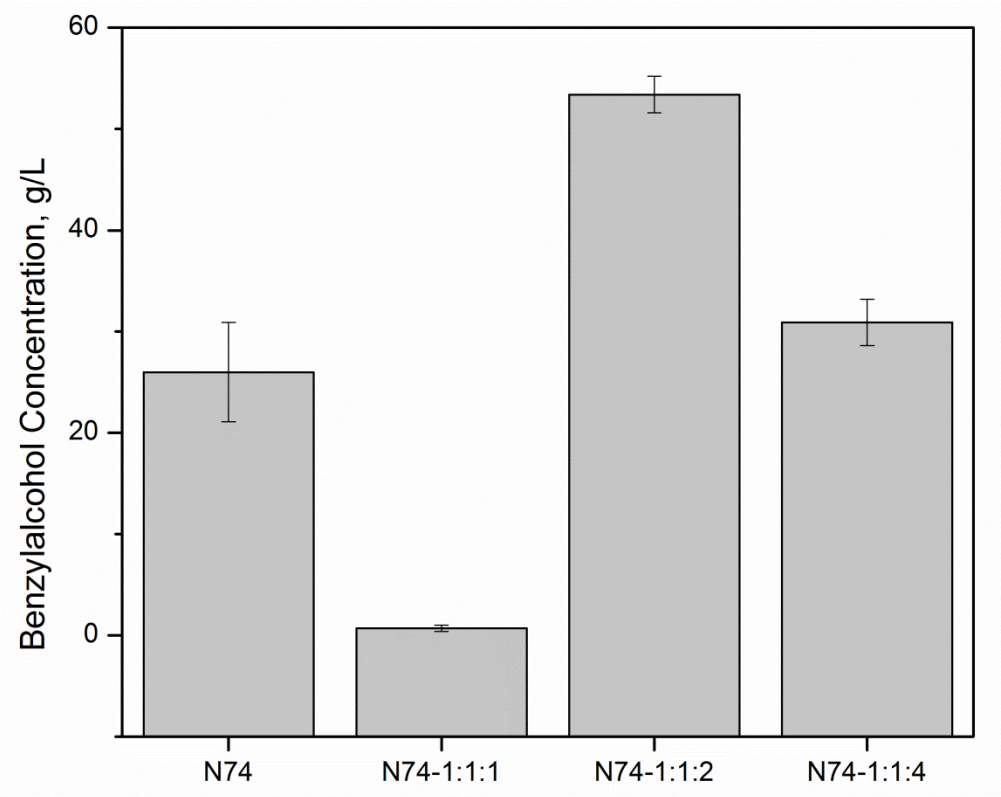

Figure 5.2 Assaying efficacy of protein scaffolds to enhance metabolic flux. Experiment compared *E. coli*NST74 (N74) harboring pHmaS^{CO}(PDZ)-trc-MdlC(GBD) and pMdlB(SH3) as a control to NST74 harboring the aforementioned plasmids and the scaffolds pTET-1:1:1, pTET-1:1:2, or pTET-1:1:4 to create strains N74-1:1:1, N74-1:1:1, and N74-1:1:4. Cultures were grown in MM1 supplemented with appropriate amino acids and induced with 0.25 mM IPTG and 0.2μMaTc. Cultures were monitored over 96 hours, at which point maximum titers of benzyl alcohol (gray columns) were achieved. Error bars reported at one standard deviation from triplicate experiments.

As illustrated in Fig 5.2, nearly a 2-fold improvement was observed in the strain N74-1:1:2 (1 MdlC: 1 MdlB: 2 HmaS) whose titers reached 53 mg/L as compared to the control N74 which reached 26 mg/L; however, it should be taken into account that strain

N74-1:1:2 achieved this feat while under the metabolic burden associated with harboring an additional plasmid, relative to the control. This result demonstrates that increasing the localized concentration of HmaS relative to MdlB and MdlC can increase flux through the metabolic pathway and confirms our hypothesis that the poor substrate affinity of HmaS for phenylpyruvate is imposing a rate-limiting flux imbalance. However, there seems to exist a local optima for co-localization of the pathway enzymes since in the cases N74-1:1:1 (1 MdlC: 1 MdlB: 1 HmaS) and N74-1:1:4 (1 MdlC: 1 MdlB: 4 HmaS) no titer improvements were observed. This preliminary experiment demonstrates that there is potential for increasing flux through the benzaldehyde and benzyl alcohol pathway; however, further experiments are required, including, i) testing alternative protein scaffold combinations which vary not only the relative number of HmaS enzymes but also MdlB and MdlC, ii) dependence of linker spacing to account for molecular space and protein folding considerations, iii) effect of domain-linker pairing to the heterologous enzymes (i.e., testing the use of the GBD or SH3 domain for HmaS as opposed to the PDZ domain), and iv) testing the effect of deleting the previously identified knockouts which enhance titer (*tyrA*, *tyrB,* and *aspC*).

5.3 Catechol: multiple routes toward the same product

A recent renewal of interest in muconic acid as a biosynthetic product has resulted in the identification of three additional pathways capable of achieving catechol as an intermediate to muconic acid. As described in Figure 5.3, in addition to the routes identified by Draths et al. (A) and by Pugh et al. (B), two additional pathways have been identified through the intermediate isochorismate, one through salicylate (C) and the other through 2,3-dihydroxybenzoate (D), as well as another pathway through the

tryptophan precursor anthranilate (E) (Draths and Frost 1994, Pugh et al. 2014, Lin et al. 2014, Sun et al. 2013).

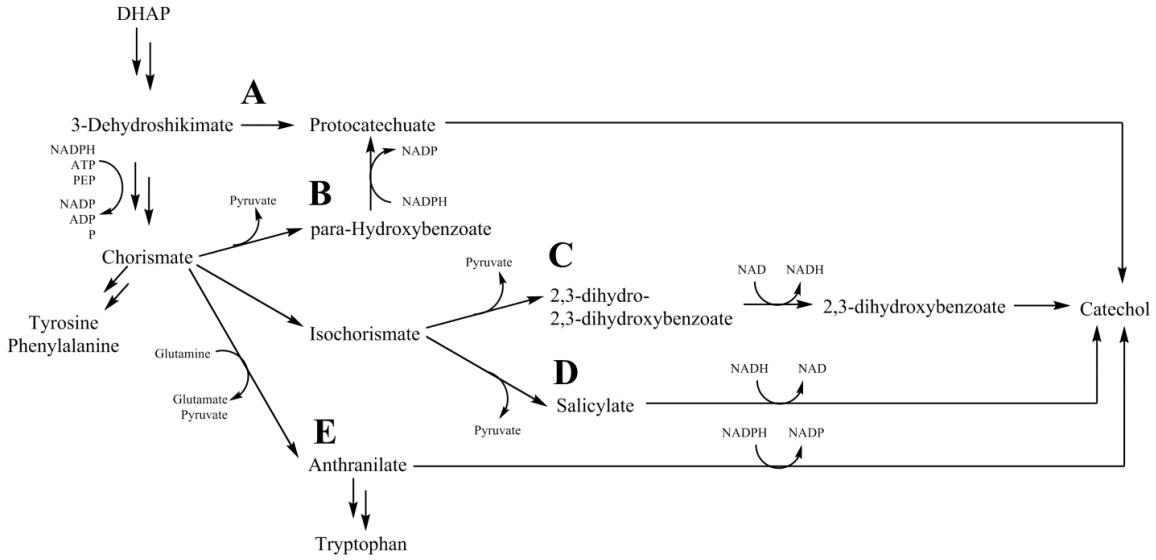

Figure 5.3 Multiple routes toward catechol. The pathways to achieve catechol from the precursors 3-dehydroshikimate (A), para-hydroxybenzoate (B), 2,3-dihydroxybenzoate (C), salicylate (D), and anthranilate (E) are illustrated. Consecutive arrows indicate that multiple steps which are not indicated.

Thus far, no fair comparison has been performed to evaluate the various pathways head-to-head; therefore, it is difficult to determine which pathway can achieve the highest titer and product-to-substrate yield. For example, pathway A was first demonstrated in an *E .coli* ΔaroE knockout strain which lacked the native shikimate dehydrogenase activity and was thus an auxotroph for phenylalanine, tyrosine, and tryptophan. While pathway A from DHS utilizes less cofactors for synthesis of catechol (1 mol NADPH, ATP, and PEP per mol of chorismate produced), synthesis from the branch point chorismate may be a more desirable alternative to prevent the auxotrophies associated with the *aroE* deletion. Since consumption of catechol is tightly regulated both through transcriptional regulation as well as allosteric enzyme inhibition by phenylalanine, tyrosine, and tryptophan, native

metabolism alone may be capable of reducing flux to these essential metabolites thereby increasing the bioavailability of chorismate for catechol synthesis. This strategy was partly demonstrated in Chapter 3 in which the phenylalanine overproducing strain *E. coli* NST74 was engineered with the chorsimate mutase/ prephenate dehydrogenase activity knocked-out (*E. coli* NST74ΔpheA^{fbr}). While the strain was an auxotroph for phenylalanine, titers for catechol achieved 630 mg/L with no measured co-production of either tyrosine or tryptophan. To further improve upon said approach and test the hypothesis for allowing native metabolism to regulate consumption of catechol thereby increasing the bioavailability of chorismate without an auxotrophy, the wild-type, feedback-regulated chorsimate mutase/prephenate dehydrogenase (PheA) could be replaced on the chromosome of NST74 to potentially create a chorismate overproducer. Alternatively, a wild-type strain of *E. coli* could be engineered to express the feedback resistant DHAP synthase isozymes (AroF^{fbr}, AroG^{fbr}, and AroH^{fbr}). In order for the biosynthesis of catechol to be economically viable, auxotrophies should be avoided to reduce the cost of necessary exogenous supplementation of aromatic amino acids.

The five catechol pathways were also compared from a thermodynamic perspective by estimating the change in the Gibbs energy of reaction under physiological conditions using eQuilibrator$^{2.0}$ (http://equilibrator.weizmann.ac.il/). As illustrated in Figure 5.4, as the pathways proceed, the respective total change in Gibbs free energy decreases indicating that all pathways are thermodynamically favorable. Though the total change in Gibbs free energy is partially magnified by the number of reactions steps, the pathways from *para*-hydroxybenzoate (B), salicylate (D), and anthranilate (E) had the largest change in Gibbs free energy which may indicate that these pathways are more

thermodynamically favorable than the pathway from 3-dehydroshikimate (A). Future investigations to identify the best route toward catechol and muconic acid should compare the described pathways in a common host and media platform as well as under similar fermentation conditions.

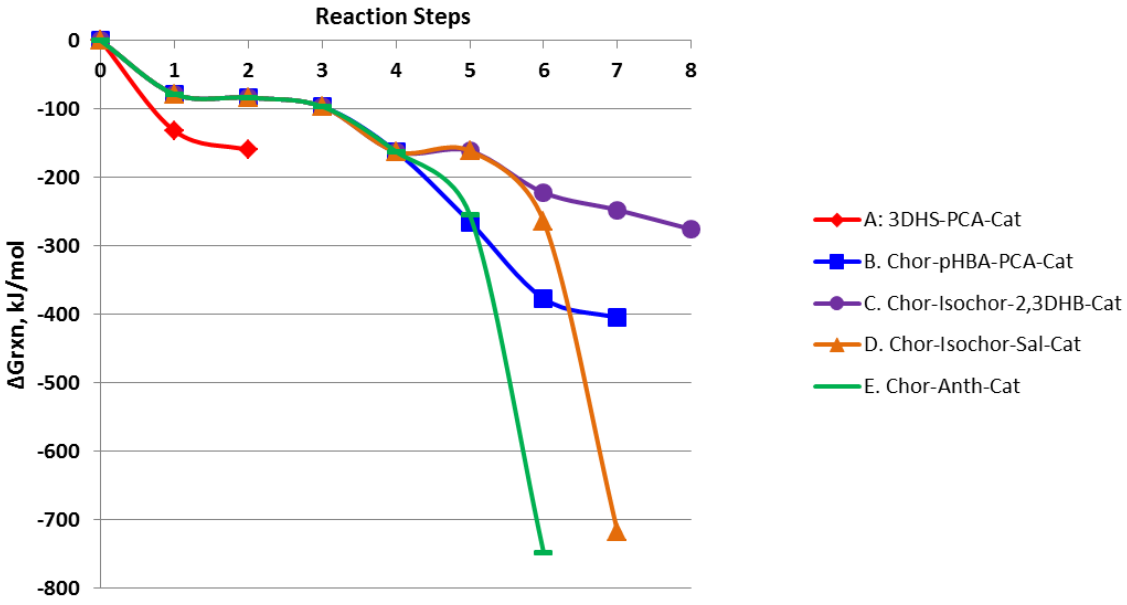

Figure 5.4 Gibbs free energy of reaction by enzyme step for the production of catechol. The change in Gibbs free energy was compared for the 5 catechol pathways illustrated in Figure 5.3.

5.4 Conclusion

These works demonstrate the feasibility of synthesizing drop-in compatible aromatic fine chemicals in engineered microorganisms. Through the development of novel pathway design, the biosynthetic production of benzyl alcohol and benzaldehyde has been realized. Though flux through the pathway is limited due to poor substrate affinity, protein scaffolds have demonstrated a feasible means of overcoming said limitations in preliminary experiments. In addition, to benzyl alcohol and benzaldehyde, an alternative pathway toward the production of para-hydroxybenzoate, protocatechuate, and catechol was demonstrated. While alternative pathways have since been discovered

to achieve catechol, future work is necessary to determine the optimal route toward the product of interest. Lastly, through the use of elementary mode analysis, an *in silico* model was developed and it was discovered that co-feeding glycerol and D-xylose may increase PEP and E4P bioavailability thereby increasing flux toward the production of aromatic fine chemicals. Ultimately, the field of metabolic engineering requires development of new and novel approaches, such as those discussed throughout this dissertation, in order to bridge the fiscal gap between petroleum derived products and their renewable drop-in compatible alternatives.

REFERENCES

Adkins, J., S. Pugh, R. McKenna, and D. R. Nielsen. 2012. "Engineering microbial chemical factories to produce renewable "biomonomers"." *Front Microbiol* no. 3:313. doi: 10.3389/fmicb.2012.00313.

Ash, M., and I. Ash. 2009. "Handbook of preservatives." In. Endicott, NY: Synapse Information Resources,.

Atkinson, M. R., M. A. Savageau, J. T. Myers, and A. J. Ninfa. 2003. "Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli." *Cell* no. 113 (5):597-607.

Averesch, Nils J. H., and Jens O. Krömer. 2014. "Tailoring strain construction strategies for muconic acid production in S. cerevisiae and E. coli." *Metabolic Engineering Communications* no. 1:19-28.

Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori. 2006. "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection." *Mol Syst Biol* no. 2:2006 0008. doi: 10.1038/msb4100050.

Baek, J. M., S. Mazumdar, S. W. Lee, M. Y. Jung, J. H. Lim, S. W. Seo, G. Y. Jung, and M. K. Oh. 2013. "Butyrate Production in Engineered Escherichia coli With Synthetic Scaffolds." *Biotechnology and Bioengineering*. doi: 10.1002/bit.24925.

Becker, Jessica, and Eckhard Boles. 2003. "A Modified Saccharomyces cerevisiae Strain That Consumes l-Arabinose and Produces Ethanol." *Applied and Environmental Microbiology* no. 69 (7):4144-4150. doi: 10.1128/aem.69.7.4144-4150.2003.

Brownlee, June, Panqing He, Graham R. Moran, and David H. T. Harrison. 2008. "Two Roads Diverged: The Structure of Hydroxymandelate Synthase from Amycolatopsis orientalis in Complex with 4-Hydroxymandelate†,‡." *Biochemistry* no. 47 (7):2002-2013. doi: 10.1021/bi701438r.

Brückner, Reinhold, and Fritz Titgemeyer. 2002. "Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization." *FEMS Microbiology Letters* no. 209 (2):141-148. doi: 10.1111/j.1574-6968.2002.tb11123.x.

Budavari, S, MJ O'neil, A Smith, and PE Heckelman. 1989. "The merck index: an encyclopedia of chemicals, drug, and biologicals."

Clomburg, JamesM, and Ramon Gonzalez. 2010. "Biofuel production in Escherichia coli: the role of metabolic engineering and synthetic biology." *Applied Microbiology and Biotechnology* no. 86 (2):419-434. doi: 10.1007/s00253-010-2446-1.

COE. 1992. Flavouring substances and natural sources of flavourings, 5th Ed., Vol. I, Council of Europe. Strasbourg.

Culp, Randolph A., and John E. Noakes. 1990. "Identification of isotopically manipulated cinnamic aldehyde and benzaldehyde." *Journal of Agricultural and Food Chemistry* no. 38 (5):1249-1255. doi: 10.1021/jf00095a021.

Curran, K. A., J. M. Leavitt, A. S. Karim, and H. S. Alper. 2013a. "Metabolic engineering of muconic acid production in Saccharomyces cerevisiae." *Metab Eng* no. 15:55-66. doi: 10.1016/j.ymben.2012.10.003.

Curran, Kathleen A., John M. Leavitt, Ashty S. Karim, and Hal S. Alper. 2013b. "Metabolic engineering of muconic acid production in Saccharomyces cerevisiae." *Metabolic Engineering* no. 15 (0):55-66.

Datsenko, K. A., and B. L. Wanner. 2000a. "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." *Proceedings of the National Academy of Sciences of the United States of America* no. 97 (12):6640-6645. doi: DOI 10.1073/pnas.120163297.

Datsenko, Kirill A., and Barry L. Wanner. 2000b. "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." *Proceedings of the National Academy of Sciences* no. 97 (12):6640-6645. doi: 10.1073/pnas.120163297.

Dharmadi, Yandi, Abhishek Murarka, and Ramon Gonzalez. 2006. "Anaerobic fermentation of glycerol by Escherichia coli: A new platform for metabolic engineering." *Biotechnology and Bioengineering* no. 94 (5):821-829. doi: 10.1002/bit.21025.

Draths, K. M., and J. W. Frost. 1994. "Environmentally Compatible Synthesis of Adipic Acid from D-Glucose." *Journal of the American Chemical Society* no. 116 (1):399-400. doi: Doi 10.1021/Ja00080a057.

Draths, K. M., D. L. Pompliano, D. L. Conley, J. W. Frost, A. Berry, G. L. Disbrow, R. J. Staversky, and J. C. Lievense. 1992. "Biocatalytic synthesis of aromatics from D-glucose: the role of transketolase." *Journal of the American Chemical Society* no. 114 (10):3956-3962. doi: 10.1021/ja00036a050.

Dueber, J. E., G. C. Wu, G. R. Malmirchegini, T. S. Moon, C. J. Petzold, A. V. Ullal, K. L. J. Prather, and J. D. Keasling. 2009. "Synthetic protein scaffolds provide modular control over metabolic flux." *Nature Biotechnology* no. 27 (8):753-U107. doi: Doi 10.1038/Nbt.1557.

Dunlop, M. J. 2011. "Engineering microbes for tolerance to next-generation biofuels." *Biotechnol Biofuels* no. 4:32.

Dunlop, M. J., Z. Y. Dossani, H. L. Szmidt, H. C. Chu, T. S. Lee, J. D. Keasling, M. Z. Hadi, and A. Mukhopadhyay. 2011. "Engineering microbial biofuel tolerance and export using efflux pumps." *Mol Syst Biol* no. 7:487.

Dusséaux, Simon, Christian Croux, Philippe Soucaille, and Isabelle Meynial-Salles. 2013. "Metabolic engineering of Clostridium acetobutylicum ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture." *Metabolic Engineering* no. 18 (0):1-8.

Erickson, B., Nelson, and P. Winters. 2012. "Perspective on opportunities in industrial biotechnology in renewable chemicals." *Biotechnol J* no. 7 (2):176-85. doi: 10.1002/biot.201100069.

Felton, L. 2013. *Remington Essentials of Pharmaceutics*: Pharmaceutical Press.

Fenaroli, Giovanni, and George A. Burdock. 1995. *Fenaroli's handbook of flavor ingredients*. 3rd ed. 2 vols. Boca Raton, Fla.: CRC Press.

Fuqua, C., M. R. Parsek, and E. P. Greenberg. 2001. "Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing." *Annu Rev Genet* no. 35:439-68. doi: 10.1146/annurev.genet.35.102401.090913.

Gallagher, D. T., M. Mayhew, M. J. Holden, A. Howard, K. J. Kim, and V. L. Vilker. 2001. "The crystal structure of chorismate lyase shows a new fold and a tightly retained product." *Proteins-Structure Function and Genetics* no. 44 (3):304-311. doi: Doi 10.1002/Prot.1095.

Gardner, T. S., C. R. Cantor, and J. J. Collins. 2000. "Construction of a genetic toggle switch in Escherichia coli." *Nature* no. 403 (6767):339-42. doi: 10.1038/35002131.

Goni-Moreno, A., and M. Amos. 2012. "Continuous computation in engineered gene circuits." *Biosystems* no. 109 (1):52-6. doi: 10.1016/j.biosystems.2012.02.001.

Gosset, G., J. Yong-Xiao, and K. M. Draths. 1996. "A direct comparison of approaches for increasing carbon flow to aromatic biosynthesis in Escherichia coli." *J Ind Microbiol* no. 17. doi: 10.1007/bf01570148.

Guo, T., B. J. Sun, M. Jiang, H. Wu, T. F. Du, Y. Tang, P. Wei, and P. K. Ouyang. 2012. "Enhancement of butanol production and reducing power using a two-stage controlled-pH strategy in batch culture of Clostridium acetobutylicum XY16." *World Journal of Microbiology & Biotechnology* no. 28 (7):2551-2558. doi: Doi 10.1007/S11274-012-1063-9.

Hammer, Karin, Ivan Mijakovic, and Peter Ruhdal Jensen. 2006. "Synthetic promoter libraries – tuning of gene expression." *Trends in Biotechnology* no. 24 (2):53-55.

Hasson, M. S., I. Schlichting, J. Moulai, K. Taylor, W. Barrett, G. L. Kenyon, P. C. Babbitt, J. A. Gerlt, G. A. Petsko, and D. Ringe. 1998. "Evolution of an enzyme active site: the structure of a new crystal form of muconate lactonizing enzyme compared with mandelate racemase and enolase." *Proc Natl Acad Sci U S A* no. 95 (18):10396-401.

Hasty, J., D. McMillen, and J. J. Collins. 2002. "Engineered gene circuits." *Nature* no. 420 (6912):224-30. doi: 10.1038/nature01257.

He, Panqing, John A Conrad, and Graham R Moran. 2010. "The rate-limiting catalytic steps of hydroxymandelate synthase from Amycolatopsis orientalis." *Biochemistry* no. 49 (9):1998-2007.

Holden, M. J., M. P. Mayhew, D. T. Gallagher, and V. L. Vilker. 2002. "Chorismate lyase: kinetics and engineering for stability." *Biochim Biophys Acta* no. 1594 (1):160-7.

Ingram, L. O., P. F. Gomez, X. Lai, M. Moniruzzaman, B. E. Wood, L. P. Yomano, and S. W. York. 1998. "Metabolic engineering of bacteria for ethanol production." *Biotechnology and Bioengineering* no. 58 (2-3):204-214. doi: 10.1002/(sici)1097-0290(19980420)58:2/3<204::aid-bit13>3.0.co;2-c.

Jain, W. K., I. Toran-Diaz, and J. Baratti. 1985. "Continuous production of ethanol from fructose by immobilized growing cells of Zymomonas mobilis." *Biotechnology and Bioengineering* no. 27 (5):613-620. doi: 10.1002/bit.260270510.

Jones, K. L., S. W. Kim, and J. D. Keasling. 2000. "Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria." *Metabolic Engineering* no. 2 (4):328-338.

Juminaga, D., E. E. K. Baidoo, A. M. Redding-Johanson, T. S. Batth, H. Burd, A. Mukhopadhyay, C. J. Petzold, and J. D. Keasling. 2012. "Modular Engineering of L-Tyrosine Production in Escherichia coli." *Applied and Environmental Microbiology* no. 78 (1):89-98.

Kang, Z., C. Zhang, G. Du, and J. Chen. 2014. "Metabolic engineering of Escherichia coli for production of 2-phenylethanol from renewable glucose." *Appl Biochem Biotechnol* no. 172 (4):2012-21. doi: 10.1007/s12010-013-0659-3.

Keasling, Jay D. 1999. "Gene-expression tools for the metabolic engineering of bacteria." *Trends in Biotechnology* no. 17 (11):452-460.

Kemp, M. B., and G. D. Hegeman. 1968. "Genetic control of the beta-ketoadipate pathway in Pseudomonas aeruginosa." *Journal of Bacteriology* no. 96 (5):1488-99.

Keseler, Ingrid M., Julio Collado-Vides, Socorro Gama-Castro, John Ingraham, Suzanne Paley, Ian T. Paulsen, Martín Peralta-Gil, and Peter D. Karp. 2005. "EcoCyc: a comprehensive database resource for Escherichia coli." *Nucleic Acids Research* no. 33 (suppl 1):D334-D337. doi: 10.1093/nar/gki108.

Kieboom, J., J. J. Dennis, G. J. Zylstra, and J. A. de Bont. 1998. "Active efflux of organic solvents by Pseudomonas putida S12 is induced by solvents." *J Bacteriol* no. 180 (24):6769-72.

Kim, B., H. Park, D. Na, and S. Y. Lee. 2013. "Metabolic engineering of Escherichia coli for the production of phenol from glucose." *Biotechnol J*. doi: 10.1002/biot.201300263.

Kim, H. U., T. Y. Kim, and S. Y. Lee. 2008. "Metabolic flux analysis and metabolic engineering of microorganisms." *Mol Biosyst* no. 4 (2):113-20. doi: 10.1039/b712395g.

Koma, D., H. Yamanaka, K. Moriyoshi, T. Ohmoto, and K. Sakai. 2012a. "Production of aromatic compounds by metabolically engineered Escherichia coli with an expanded shikimate pathway." *Appl Environ Microbiol* no. 78 (17):6203-16. doi: 10.1128/AEM.01148-12.

Koma, Daisuke, Hayato Yamanaka, Kunihiko Moriyoshi, Takashi Ohmoto, and Kiyofumi Sakai. 2012b. "Production of aromatic compounds by metabolically engineered Escherichia coli with shikimate pathway expansion." *Applied and Environmental Microbiology*. doi: 10.1128/aem.01148-12.

Krings, U., and R. G. Berger. 1998. "Biotechnological production of flavours and fragrances." *Applied Microbiology and Biotechnology* no. 49 (1):1-8. doi: 10.1007/s002530051129.

Kunjapur, Aditya M., Yekaterina Tarasova, and Kristala L. J. Prather. 2014. "Synthesis and Accumulation of Aromatic Aldehydes in an Engineered Strain of Escherichia coli." *Journal of the American Chemical Society* no. 136 (33):11644-11654. doi: 10.1021/ja506664a.

Kwon, O., A. Kotsakis, and R. Meganathan. 2000. "Ubiquinone (coenzyme Q) biosynthesis in Escherichia coli: identification of the ubiF gene." *FEMS Microbiol Lett* no. 186 (2):157-61.

Lan, Ethan I., and James C. Liao. 2011. "Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide." *Metabolic Engineering* no. 13 (4):353-363.

Lee, J. W., D. Na, J. M. Park, J. Lee, S. Choi, and S. Y. Lee. 2012. "Systems metabolic engineering of microorganisms for natural and non-natural chemicals." *Nat Chem Biol* no. 8 (6):536-46. doi: 10.1038/nchembio.970.

Lin, H. H., J. H. Chen, C. C. Huang, and C. J. Wang. 2007. "Apoptotic effect of 3,4-dihydroxybenzoic acid on human gastric carcinoma cells involving JNK/p38 MAPK signaling activation." *Int J Cancer* no. 120 (11):2306-16. doi: 10.1002/ijc.22571.

Lin, Yuheng, Xinxiao Sun, Qipeng Yuan, and Yajun Yan. 2014. "Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in Escherichia coli." *Metabolic Engineering* no. 23:62-69.

Lu, Xuefeng. 2010. "A perspective: Photosynthetic production of fatty acid-based biofuels in genetically engineered cyanobacteria." *Biotechnology Advances* no. 28 (6):742-746.

Lucchini, J. J., N. Bonnaveiro, A. Cremieux, and F. LeGoffic. 1993. "Mechanism of bactericidal action of phenethyl alcohol in *Escherichia coli*." *Current Microbiology* no. 27:295-300.

Lucchini, J. J., J. Corre, and A. Cremieux. 1990. "Antibacterial activity of phenolic compounds and aromatic alcohols." *Res Microbiol* no. 141 (4):499-510.

Mageroy, Melissa H., Denise M. Tieman, Abbye Floystad, Mark G. Taylor, and Harry J. Klee. 2012. "A Solanum lycopersicum catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol." *The Plant Journal* no. 69 (6):1043-1051. doi: 10.1111/j.1365-313X.2011.04854.x.

Marriott, John F. 2010. *Pharmaceutical compounding and dispensing*. 2nd ed. London: Pharmaceutical Press.

Martínez, Karla, Ramón de Anda, Georgina Hernández, Adelfo Escalante, Guillermo Gosset, Octavio T. Ramírez, and Francisco G. Bolívar. 2008. "Coutilization of glucose and glycerol enhances the production of aromatic compounds in an Escherichia coli strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system." *Microbial Cell Factories* no. 7 (1):1-12. doi: 10.1186/1475-2859-7-1.

McKenna, R., and D. R. Nielsen. 2011a. "Styrene biosynthesis from glucose by engineered E. coli." *Metab Eng* no. 13 (5):544-54.

McKenna, R., S. Pugh, B. Thompson, and D. R. Nielsen. 2013a. "Microbial Production of the Aromatic Building-Blocks (S)-Styrene Oxide and (R)-1,2-Phenylethanediol from Renewable Resources." *Biotechnol. J.* no. 8 (12):1465-1475.

McKenna, Rebekah, and David R. Nielsen. 2011b. "Styrene biosynthesis from glucose by engineered E. coli." *Metabolic Engineering* no. 13 (5):544-554.

McKenna, Rebekah, Shawn Pugh, Brian Thompson, and David R. Nielsen. 2013b. "Microbial production of the aromatic building-blocks (S)-styrene oxide and (R)-1,2-phenylethanediol from renewable resources." *Biotechnology Journal* no. 8 (12):1465-1475. doi: 10.1002/biot.201300035.

Meinking, T. L., M. E. Villar, M. Vicaria, D. H. Eyerdam, D. Paquet, K. Mertz-Rivera, H. F. Rivera, J. Hiriart, and S. Reyna. 2010. "The clinical trials supporting benzyl alcohol lotion 5% (Ulesfia): a safe and effective topical treatment for head lice (pediculosis humanus capitis)." *Pediatr Dermatol* no. 27 (1):19-24. doi: 10.1111/j.1525-1470.2009.01059.x.

Meylan, W. M., P. H. Howard, R. S. Boethling, D. Aronson, H. Printup, and S. Gouchie. 1999. "Improved method for estimating bioconcentration/bioaccumulation factor from octanol/water partition coefficient." *Environmental Toxicology and Chemistry* no. 18 (4):664-672. doi: Doi 10.1897/1551-5028(1999)018<0664:Imfebb>2.3.Co;2.

Miller, R. R., Ronald Newhook, and Alan Poole. 1994. "Styrene Production, Use, and Human Exposure." *Critical Reviews in Toxicology* no. 24 (s1):S1-S10. doi: doi:10.3109/10408449409020137.

Moon, Tae Seok, John E. Dueber, Eric Shiue, and Kristala L. Jones Prather. 2010. "Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli." *Metabolic Engineering* no. 12 (3):298-305.

Nair, B. 2001. "Final report on the safety assessment of Benzyl Alcohol, Benzoic Acid, and Sodium Benzoate." *Int J Toxicol* no. 20 Suppl 3:23-50.

Nichols, B P, and J M Green. 1992. "Cloning and sequencing of Escherichia coli ubiC and purification of chorismate lyase." *Journal of Bacteriology* no. 174 (16):5309-5316.

Nielsen, David R., Effendi Leonard, Sang-Hwal Yoon, Hsien-Chung Tseng, Clara Yuan, and Kristala L. Jones Prather. 2009. "Engineering alternative butanol production platforms in heterologous bacteria." *Metabolic Engineering* no. 11 (4–5):262-273.

Niu, W., K. M. Draths, and J. W. Frost. 2002. "Benzene-free synthesis of adipic acid." *Biotechnol Prog* no. 18 (2):201-11. doi: 10.1021/bp010179x.

Pacheco-Palencia, Lisbeth A., Susanne Mertens-Talcott, and Stephen T. Talcott. 2008. "Chemical Composition, Antioxidant Properties, and Thermal Stability of a Phytochemical Enriched Oil from Açai (Euterpe oleracea Mart.)." *Journal of Agricultural and Food Chemistry* no. 56 (12):4631-4636. doi: 10.1021/jf800161u.

Pietta, P. G., P. Simonetti, C. Gardana, A. Brusamolino, P. Morazzoni, and E. Bombardelli. 1998. "Catechin metabolites after intake of green tea infusions." *Biofactors* no. 8 (1-2):111-8.

Pugh, S., R. McKenna, M. Osman, and D. R. Nielsen. 2014. "Rational engineering ot a novel pathway for producing the aromatic compounds p-hydroxybenzoate, protocatechuate, and catechol in Escherichia coli." *Process Biochemistry* no. 49 (11):1843–1850.

Pugh, Shawn, Rebekah McKenna, Ibrahim Halloum, and David R. Nielsen. 2015. "Engineering Escherichia coli for renewable benzyl alcohol production." *Metabolic Engineering Communications* no. 2:39-45.

Qi, W. W., T. Vannelli, S. Breinig, A. Ben-Bassat, A. A. Gatenby, S. L. Haynie, and F. S. Sariaslani. 2007a. "Functional expression of prokaryotic and eukaryotic genes in Escherichia coli for conversion of glucose to p-hydroxystyrene." *Metab Eng* no. 9 (3):268-76.

Qi, Wei Wei, Todd Vannelli, Sabine Breinig, Arie Ben-Bassat, Anthony A. Gatenby, Sharon L. Haynie, and F. Sima Sariaslani. 2007b. "Functional expression of prokaryotic and eukaryotic genes in Escherichia coli for conversion of glucose to -hydroxystyrene." *Metabolic Engineering* no. 9 (3):268-276.

Qureshi, N., B. S. Dien, N. N. Nichols, B. C. Saha, and M. A. Cotta. 2006. "Genetically Engineered Escherichia Coli for Ethanol Production from Xylose: Substrate and Product Inhibition and Kinetic Parameters." *Food and Bioproducts Processing* no. 84 (2):114-122.

Radakovits, Randor, Robert E. Jinkerson, Al Darzins, and Matthew C. Posewitz. 2010. "Genetic Engineering of Algae for Enhanced Biofuel Production." *Eukaryotic Cell* no. 9 (4):486-501. doi: 10.1128/ec.00364-09.

Radwan, Gamil M., Shaheen A. Al-Muhtaseb, Ali M. Dowaidar, and Mohamed A. Fahim. 1997. "Extraction of Aromatics from Petroleum Naphtha Reformate by a 1-Cyclohexyl-2-pyrrolidone/Ethylene Carbonate Mixed Solvent." *Industrial & Engineering Chemistry Research* no. 36 (2):414-418. doi: 10.1021/ie960395o.

Ramos, J. L., E. Duque, M. T. Gallegos, P. Godoy, M. I. Ramos-Gonzalez, A. Rojas, W. Teran, and A. Segura. 2002. "Mechanisms of solvent tolerance in gram-negative bacteria." *Annual Review of Microbiology* no. 56:743-768. doi: DOI 10.1146/annurev.micro.56.012302.161038.

Ramos, Juan-Luis. 2004. "Pseudomonas. Volume 3, Biosynthesis of macromolecules and molecular metabolism." In. Dordrecht ; London: Springer,.

Rhodia. 2012. GPS Safety Summary: Catechol. Rhodia Member of the Solvay group.

Ro, D. K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Y. Chang, S. T. Withers, Y. Shiba, R. Sarpong, and J. D. Keasling. 2006. "Production of the antimalarial drug precursor artemisinic acid in engineered yeast." *Nature* no. 440 (7086):940-943. doi: Doi 10.1038/Nature04640.

Rodriguez, G. M., and S. Atsumi. 2014. "Toward aldehyde and alkane production by removing aldehyde reductase activity in Escherichia coli." *Metab Eng* no. 25:227-37. doi: 10.1016/j.ymben.2014.07.012.

Rojas, A., E. Duque, G. Mosqueda, G. Golden, A. Hurtado, J. L. Ramos, and A. Segura. 2001. "Three efflux pumps are required to provide efficient tolerance to toluene in Pseudomonas putida DOT-T1E." *Journal of Bacteriology* no. 183 (13):3967-3973.

Rosche, B., V. Sandford, M. Breuer, B. Hauer, and P. Rogers. 2001. "Biotransformation of benzaldehyde into (R)-phenylacetylcarbinol by filamentous fungi or their extracts." *Applied Microbiology and Biotechnology* no. 57 (3):309-315. doi: 10.1007/s002530100781.

Saier, M H, and S Roseman. 1976. "Sugar transport. 2nducer exclusion and regulation of the melibiose, maltose, glycerol, and lactose transport systems by the phosphoenolpyruvate:sugar phosphotransferase system." *Journal of Biological Chemistry* no. 251 (21):6606-6615.

Sako, T., I. Okajima, T. Sugeta, K. Otake, S. Yoda, Y. Takebayashi, and C. Kamizawa. 2000. "Recovery of constituent monomers from polyethylene terephthalate with supercritical methanol." *Polymer Journal* no. 32 (2):178-181. doi: DOI 10.1295/polymj.32.178.

Salis, H. M., E. A. Mirsky, and C. A. Voigt. 2009. "Automated design of synthetic ribosome binding sites to control protein expression." *Nature Biotechnology* no. 27 (10):946-50. doi: 10.1038/nbt.1568.

Satrio, Justinus A. B., and L. K. Doraiswamy. 2001. "Production of benzaldehyde: a case study in a possible industrial application of phase-transfer catalysis." *Chemical Engineering Journal* no. 82 (1–3):43-56.

Schmidt, Robert J. 2005. "Industrial catalytic processes—phenol production." *Applied Catalysis A: General* no. 280 (1):89-103.

Schobert, H. H., and C. Song. 2002. "Chemicals and materials from coal in the 21st century." *Fuel* no. 81 (1):15-32.

Shams Yazdani, Syed, and Ramon Gonzalez. 2008. "Engineering Escherichia coli for the efficient conversion of glycerol to ethanol and co-products." *Metabolic Engineering* no. 10 (6):340-351.

Siebert, M., K. Severin, and L. Heide. 1994. "Formation of 4-hydroxybenzoate in Escherichia coli: characterization of the ubiC gene and its encoded enzyme chorismate pyruvate-lyase." *Microbiology* no. 140 (Pt 4):897-904.

Sikkema, J., J. A. de Bont, and B. Poolman. 1994. "Interactions of cyclic hydrocarbons with biological membranes." *J Biol Chem* no. 269 (11):8022-8.

Stellman, J.M. 1998. *Encyclopaedia of occupational health and safety*. 4th ed. Geneva: International Labour Office.

Stoye, Dieter, and Werner Freitag. 1998. *Paints, coatings, and solvents*. 2nd, completely rev. ed. Weinheim ; New York: Wiley-VCH.

Sukumar, N., Y. Xu, D. L. Gatti, B. Mitra, and F. S. Mathews. 2001. "Structure of an active soluble mutant of the membrane-associated (S)-mandelate dehydrogenase." *Biochemistry* no. 40 (33):9870-8.

Sulzenbacher, G., K. Alvarez, R. H. Van Den Heuvel, C. Versluis, S. Spinelli, V. Campanacci, C. Valencia, C. Cambillau, H. Eklund, and M. Tegoni. 2004. "Crystal structure of E.coli alcohol dehydrogenase YqhD: evidence of a covalently modified NADP coenzyme." *J Mol Biol* no. 342 (2):489-502. doi: 10.1016/j.jmb.2004.07.034.

Sun, X., Y. Lin, Q. Huang, Q. Yuan, and Y. Yan. 2013. "A novel muconic acid biosynthesis approach by shunting tryptophan biosynthesis via anthranilate." *Appl Environ Microbiol* no. 79 (13):4024-30. doi: 10.1128/AEM.00859-13.

Sun, Z. T., Y. Y. Ning, L. X. Liu, Y. M. Liu, B. B. Sun, W. H. Jiang, C. Yang, and S. Yang. 2011. "Metabolic engineering of the L-phenylalanine pathway in Escherichia coli for the production of S- or R-mandelic acid." *Microbial Cell Factories* no. 10.

Terzer, M., and J. Stelling. 2008. "Large-scale computation of elementary flux modes with bit pattern trees." *Bioinformatics* no. 24 (19):2229-2235. doi: 10.1093/bioinformatics/btn401.

Thompson, Brian, Michael Machas, and David R Nielsen. 2016. "Engineering and comparison of non-natural pathways for microbial phenol production." *Biotechnology and bioengineering*.

Tian, R. R., Q. H. Pan, J. C. Zhan, J. M. Li, S. B. Wan, Q. H. Zhang, and W. D. Huang. 2009. "Comparison of phenolic acids and flavan-3-ols during wine fermentation of grapes with different harvest times." *Molecules* no. 14 (2):827-38. doi: 10.3390/molecules14020827.

Tribe, D.E. 1987a. "Novel microorganism and method." *US Patent 4,681,852*.

Tribe, David E. 1987b. Novel Microorganisms and Method. Austgen-Biojet International Pty.

Trinh, C. T., A. Wlaschin, and F. Srienc. 2009. "Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism." *Applied Microbiology and Biotechnology* no. 81 (5):813-826. doi: 10.1007/s00253-008-1770-1.

Tsou, Amy Y, Stephen C Ransom, John A Gerlt, Douglas D Buechter, Patricia C Babbitt, and George L Kenyon. 1990. "Mandelate pathway of Pseudomonas putida: sequence relationships involving mandelate racemase,(S)-mandelate dehydrogenase, and benzoylformate decarboxylase and expression of benzoylformate decarboxylase in Escherichia coli." *Biochemistry* no. 29 (42):9856-9862.

Tyo, K. E., H. S. Alper, and G. N. Stephanopoulos. 2007. "Expanding the metabolic engineering toolbox: more options to engineer cells." *Trends in Biotechnology* no. 25 (3):132-7. doi: 10.1016/j.tibtech.2007.01.003.

Verhoef, Suzanne, Nick Wierckx, R. G. Maaike Westerhof, Johannes H. de Winde, and Harald J. Ruijssenaars. 2009. "Bioproduction of p-Hydroxystyrene from Glucose by the Solvent-Tolerant Bacterium Pseudomonas putida S12 in a Two-Phase Water-Decanol Fermentation." *Applied and Environmental Microbiology* no. 75 (4):931-936. doi: 10.1128/aem.02186-08.

Weber, C., C. Bruckner, S. Weinreb, C. Lehr, C. Essl, and E. Boles. 2012. "Biosynthesis of cis,cis-muconic acid and its aromatic precursors, catechol and protocatechuic acid, from renewable feedstocks by Saccharomyces cerevisiae." *Appl Environ Microbiol* no. 78 (23):8421-30. doi: 10.1128/AEM.01983-12.

Wierckx, N. J., H. Ballerstedt, J. A. de Bont, and J. Wery. 2005a. "Engineering of solvent-tolerant Pseudomonas putida S12 for bioproduction of phenol from glucose." *Appl Environ Microbiol* no. 71 (12):8221-7.

Wierckx, Nick J. P., Hendrik Ballerstedt, Jan A. M. de Bont, and Jan Wery. 2005b. "Engineering of Solvent-Tolerant Pseudomonas putida S12 for Bioproduction of Phenol from Glucose." *Applied and Environmental Microbiology* no. 71 (12):8221-8227.

Wilson, Lance, and Steven Martin. 1999. "Benzyl Alcohol as an Alternative Local Anesthetic." *Annals of Emergency Medicine* no. 33 (5):495-499.

Woodruff, Lauren B. A., Nanette R. Boyle, and Ryan T. Gill. 2013. "Engineering improved ethanol production in Escherichia coli with a genome-wide approach." *Metabolic Engineering* no. 17 (0):1-11.

Work, Victoria H., Sarah D'Adamo, Randor Radakovits, Robert E. Jinkerson, and Matthew C. Posewitz. 2012. "Improving photosynthesis and metabolic networks for the competitive production of phototroph-derived biofuels." *Current Opinion in Biotechnology* no. 23 (3):290-297.

Yadav, GD, and PH Mehta. 1993. "Theoretical and experimental analysis of capsule membrane phase transfer catalysis: selective alkaline hydrolysis of benzyl chloride to benzyl alcohol." *Catalysis letters* no. 21 (3-4):391-403.

Yazdani, Syed Shams, and Ramon Gonzalez. 2007. "Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry." *Current Opinion in Biotechnology* no. 18 (3):213-219.

Zhang, Fuzhong, Sarah Rodriguez, and Jay D. Keasling. 2011. "Metabolic engineering of microbial pathways for advanced biofuels production." *Current Opinion in Biotechnology* no. 22 (6):775-783.