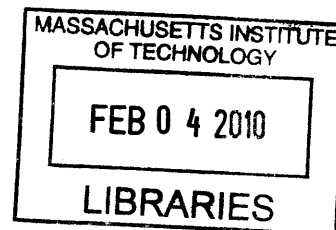


# Development of Metabolic Pathways for the Biosynthesis of Hydroxyacids and Lactones

**Collin H. Martin**

B.S., Chemical Engineering

University of Oklahoma, 2004



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

**ARCHIVES**

Massachusetts Institute of Technology

*[February 2010]*  
November 2009

© 2009 Massachusetts Institute of Technology  
All Rights Reserved

*L A A*

Signature of Author.....

Collin Martin  
Department of Chemical Engineering

Certified by.....

*[Signature]*  
Kristala Jones Prather  
Assistant Professor of Chemical Engineering  
Thesis Supervisor

Accepted by.....

William Deen  
Professor of Chemical Engineering  
Chairman, Committee for Graduate Students

# Development of Metabolic Pathways for the Biosynthesis of Hydroxyacids and Lactones

Collin H. Martin

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

## Abstract

In this thesis, metabolic routes were developed for the production of hydroxyacids and their lactones in multiple microbial systems. These compounds see widespread use in the production of pharmaceuticals, polymers, and fine chiral intermediates. First in this thesis, strategies and tools for metabolic pathway design are discussed. This is followed by the descriptions of and data for each microbial production system.

The compounds produced in this thesis and their highest titers obtained are shown below:

3-Hydroxybutyrate (3HB)	2.9 g/L
3-Hydroxyvalerate (3HV)	5.3 g/L
4-Hydroxyvalerate (4HV)	27.1 g/L
4-Valerolactone (4VL)	8.2 g/L
3,4-Dihydroxybutyrate (DHBA)	3.2 g/L
3-Hydroxybutyrolactone (3-HBL)	2.2 g/L

The production of the two hydroxyvalerates was accomplished through the reduction of the renewable substrate levulinate in *Pseudomonas putida* by endogenous host enzymes followed by the liberation of the hydroxyvalerate product through the recombinant expression of thioesterase B (*tesB*). The production of 4VL was accomplished from levulinate by adding the lactonase paraoxonase I (PON1) to the *P. putida* hydroxyvalerate production system. Because 4VL was found to exist in a pH-dependent equilibrium with 4HV, the lactonase was expressed extracytosolically in acidic media to achieve significant titers of 4VL. The addition of a second resin phase to 4VL-producing cultures with a high affinity for 4VL substantially enhanced lactone production.

3HB, 3HV, DHBA, and 3-HBL were all produced in *Escherichia coli* through the expression of an acetoacetyl-CoA thiolase (*thil*, *bktB*, or *phaA*), a 3-hydroxybutyryl-CoA reductase (*phaB* or *hbd*), and *tesB*. Supplying glucose to *E. coli* expressing these enzymes resulted in 3HB production, while supplying glucose and propionate results in 3HV production. Supplying glucose and glycolate resulted in DHBA production with

some 3-HBL, but only with the help of a fourth gene – propionyl-CoA transferase (*pct*). Removing the *tesB* gene from this four-gene system substantially increases 3-HBL titers at the expense of DHBA. This work represents the first successful production of DHBA and 3-HBL in a biological system from carbohydrate-based substrates.

In each of these systems, several broadly-applicable tools and strategies were developed to enhance product titer or discover new metabolic activities. In the *P. putida* system, cytosolic and extracytosolic biocatalysis were combined in a single metabolic pathway to realize lactone production. This catalytic strategy, termed integrated bioprocessing, is applicable to other metabolic pathways whose production suffers due to a suboptimal cytosolic enzyme. Also in the *P. putida* system, two-phase cultures were used to sequester the lactone product away from the lactonase, helping to drive lactone-hydroxyacid equilibrium towards the lactone. This methodology allows one to overcome equilibrium-based limitations of product titer. Finally in the *E. coli* work, a promiscuous pathway normally used for polyhydroxyalkanoate synthesis was exploited to give a wide range of hydroxyacid products. This substrate promiscuity was critical in achieving the production of new compounds biologically and thus substrate promiscuity was identified as a key component for metabolic pathway design and construction.

Thesis Supervisor: Kristala Jones Prather

# Table of Contents

<b>Abstract</b>	<b>2</b>
<b>List of Figures</b>	<b>7</b>
<b>List of Tables</b>	<b>8</b>
<b>List of Abbreviations</b>	<b>9</b>
<b>Chapter 1: Introduction</b>	<b>10</b>
Motivation and Objectives	10
Organization and Summary	15
<b>Chapter 2: Metabolic Pathway Design</b>	<b>18</b>
Building Novel Metabolic Routes at the Protein Level	19
<i>Theoretical Approaches</i>	21
<i>Experimental Approaches</i>	25
Building Novel Metabolic Routes at the Pathway Level	28
<i>Theoretical Approaches</i>	30
<i>Experimental Approaches</i>	37
Conclusions and Outlook for Pathway Design	43
The ReBiT Database for Pathway Design	45
<i>Methods</i>	48
<i>Implementation</i>	53
<i>Conclusions</i>	59
<b>Chapter 3: Hydroxyvalerate Production in <i>Pseudomonas putida</i></b>	<b>60</b>
Background	60
Results	64
<i>Initial Cellular System Optimization</i>	64
<i>Production of Hydroxyvalerates in LB Medium</i>	65
<i>Production of Hydroxyvalerates in M9 Minimal Medium</i>	68
<i>Mass Spectrometry Confirmation of Hydroxyvalerate Production</i>	71
<i>Investigation of Levulinate Metabolism in <i>P. putida</i></i>	74
Discussion	76
Materials and Methods	80
<i>Strains and Plasmids</i>	80
<i>Chemicals and Growth Media</i>	81
<i>Culturing of <i>P. putida</i></i>	81
<i>Analytcs</i>	82



<b>Chapter 4: Valerolactone Production in <i>Pseudomonas putida</i></b>	<b>84</b>
Background	85
Results	89
<i>Confirmation of Extracytosolic PON1 Expression</i>	89
<i>Effect of pH on 4VL Production</i>	94
<i>Production of 4VL in Shake Flasks</i>	95
<i>Bioreactor-Scale Production of 4VL</i>	97
Discussion	99
<i>Identification of G3C9 PON1 as an Extracytosolic Lactonase</i>	99
<i>pH Effect on Lactone Production</i>	101
<i>Production of 4VL using Integrated Bioprocessing</i>	101
<i>Applications and Implications of Integrated Bioprocessing</i>	103
Materials and Methods	105
<i>Chemicals</i>	105
<i>Strains and Plasmids</i>	105
<i>Culture Conditions</i>	107
<i>Lactonase Assays</i>	108
<i>Alkaline Phosphatase Assays</i>	109
<i>HPLC Analysis</i>	110
<b>Chapter 5: 3-Hydroxybutyrate Production in <i>Escherichia coli</i></b>	<b>111</b>
Introduction	111
Results	116
<i>Production of Chiral 3HB in BL21Star(DE3) E. coli</i>	116
<i>Production of Chiral 3HB in MG1655(DE3) E. coli</i>	117
<i>Confirmation of 3HB Stereochemistry</i>	119
<i>Measurement of 3HB Pathway Enzyme Activities</i>	121
<i>Measurement of Cofactor Levels in MG1655(DE3) E. coli</i>	122
Discussion	123
<i>E. coli B versus E. coli K-12 in Chiral 3HB Production</i>	124
<i>Effect of Alternative Thiolases on Chiral 3HB Production</i>	125
<i>TesB versus Ptb-Buk as a CoA Removal System</i>	126
<i>Discrepancies between Enzyme Activities for 3HB Production</i>	127
Conclusions	129
Materials and Methods	130
<i>Microorganisms</i>	130
<i>Plasmid Construction</i>	132
<i>Culture Conditions</i>	133
<i>Metabolite Analysis and Dry Cell Weight Determination</i>	134
<i>Methyl Esterification of 3HB</i>	134
<i>Chiral HPLC Analysis of Methyl-3HB</i>	135
<i>Enzyme Assays</i>	135
<i>Quantification of Intracellular Cofactor Levels</i>	136

<b>Chapter 6: Production of Higher Hydroxyacids in <i>Escherichia coli</i></b>	<b>137</b>
Introduction	137
Results	141
<i>Production of 3HV from Glucose and Propionate</i>	141
<i>Chiral Analysis of Microbially-Produced 3HV</i>	142
<i>Production of DHBA and 3-HBL from Glucose and Glycolate</i>	145
<i>Effect of tesB Expression on DHBA and 3-HBL Production</i>	147
<i>Effect of Acidic Post-Treatment on DHBA and 3-HBL Titters</i>	148
<i>Strain Effects on DHBA and 3-HBL Production</i>	149
Discussion	152
<i>Controlling Product Profiles</i>	152
<i>Controlling Product Stereochemistry</i>	155
Conclusions	156
Materials and Methods	157
<i>Strains and Chemicals</i>	157
<i>Plasmids and Primers</i>	158
<i>Culturing Conditions</i>	160
<i>Methyl Esterification of 3HV</i>	161
<i>HPLC Analyses</i>	161
<b>Chapter 7: Conclusions and Recommendations</b>	<b>164</b>
Summary of Products and Titters	164
Experimental Observations of Key Pathway Design Parameters	165
Analysis of Pathway Lactonization Strategies	168
Outlook and Recommendations	171
Concluding Thoughts	173
<b>References</b>	<b>178</b>
<b>Acknowledgements</b>	<b>211</b>

## List of Figures

<b>Figure 1.1:</b> Applications of Hydroxyacids	11
<b>Figure 1.2:</b> Synthetic Routes for the Derivatization of 3-HBL	12
<b>Figure 1.3:</b> Summary of Current Methods for 3-HBL Production	13
<b>Figure 2.1:</b> Scheme for Metabolic Pathway Creation	19
<b>Figure 2.2:</b> Flowchart for the Creation of New Enzymes	20
<b>Figure 2.3:</b> Examples of Generalized Enzyme-Catalyzed Reactions	32-34
<b>Figure 2.4:</b> Strategies for Synthetic Pathway Creation	38
<b>Figure 2.5:</b> Submission Page for a Structural Query to ReBiT	55
<b>Figure 2.6:</b> Results of an Example Query to ReBiT	57
<b>Figure 2.7:</b> Enzymatic Reactions Recommended by ReBiT for a Sample Query	58
<b>Figure 3.1:</b> Production of Hydroxyvalerates in LB Shake Flasks	67
<b>Figure 3.2:</b> Production of Hydroxyvalerates in M9 Shake Flasks	69
<b>Figure 3.3:</b> 4HV-to-3HV Production Ratios under Different Conditions	71
<b>Figure 3.4:</b> Mass Spectroscopy Analysis of 4HV and 3HV	72-73
<b>Figure 3.5:</b> Effects of Substrate Boli on Hydroxyvalerate Production	75
<b>Figure 3.6:</b> Proposed Model of Levulinate Metabolism in <i>P. putida</i>	76
<b>Figure 4.1:</b> Integrated Bioprocessing System for the Production of 4VL	87
<b>Figure 4.2:</b> Comparison of N-termini of Native PON1, G3C9 PON1, and PhoA	88
<b>Figure 4.3:</b> Lactonization of 4HV by <i>E. coli</i> Expressing PON1 or tPON1	93
<b>Figure 4.4:</b> Lactonization of 4HV by <i>E. coli</i> at Different pH Values	95
<b>Figure 4.5:</b> Shake-Flask Production of 4VL by Recombinant <i>P. putida</i>	97
<b>Figure 4.6:</b> Bioreactor-Scale Production of 4VL by Recombinant <i>P. putida</i>	98
<b>Figure 5.1:</b> 3HB Biosynthetic Pathway Schematic in Recombinant <i>E. coli</i>	115
<b>Figure 5.2:</b> Production of 3HB by <i>E. coli</i> BL21Star(DE3) and MG1655(DE3)	119
<b>Figure 5.3:</b> Chiral HPLC Spectra of 3HB Produced from Recombinant <i>E. coli</i>	120
<b>Figure 6.1:</b> Schematic of the 3-Hydroxyalkanoic Acid Pathway	139
<b>Figure 6.2:</b> 3HV Titters from the 3-Hydroxyalkanoic Acid Pathway	142
<b>Figure 6.3:</b> Chiral HPLC Spectra of 3HV Produced from Recombinant <i>E. coli</i>	144
<b>Figure 6.4:</b> Production of DHBA and 3-HBL by recombinant <i>E. coli</i>	146
<b>Figure 6.5:</b> LC/MS Analysis of DHBA produced by recombinant <i>E. coli</i>	146
<b>Figure 6.6:</b> Effect of <i>tesB</i> on DHBA and 3-HBL Product Profiles	147
<b>Figure 6.7:</b> Effect of Acid Treatment on DHBA and 3-HBL Product Profiles	149
<b>Figure 6.8:</b> Production of DHBA and 3-HBL by MG1655(DE3) <i>E. coli</i>	150
<b>Figure 6.9:</b> Highest DHBA and 3-HBL Titters obtained from <i>E. coli</i>	151
<b>Figure 6.10:</b> Pathway Scheme for DHBA and 3-HBL Production	153
<b>Figure 6.11:</b> Production of Acetate from the 3-Hydroxyalkanoic Acid Pathway	155

## List of Tables

<b>Table 3.1:</b> Hydroxyvalerate Production in Different <i>P. putida</i> Systems	64
<b>Table 4.1:</b> <i>phoA</i> Assay Results for Various PON1- <i>phoA</i> Protein Fusions	91
<b>Table 5.1:</b> Extracellular Production of Chiral 3HB by <i>E. coli</i> BL21Star(DE3)	117
<b>Table 5.2:</b> 3HB Pathway Enzyme Specific Activities of Crude <i>E. coli</i> Extracts	122
<b>Table 5.3:</b> Levels and Ratios of NAD(P) <sup>+</sup> and NAD(P)H Cofactors	123
<b>Table 5.4:</b> <i>E. coli</i> Strains, Plasmids and Oligonucleotides for 3HB Production	131
<b>Table 6.1:</b> Characteristics of 3-Hydroxyalkanoic Acid Pathway Enzymes	140
<b>Table 6.2:</b> Primers used in 3-Hydroxyalkanoic Acid Pathway Construction	160
<b>Table 7.1:</b> Summary of Products and Titers Obtained	165

## List of Abbreviations

<b>Abbreviation</b>	<b>Full Name</b>
3HB	3-Hydroxybutyrate
3HV	3-Hydroxyvalerate
3-HBL	3-Hydroxybutyrolactone
4HV	4-Hydroxyvalerate
4VL	4-Valerolactone
ADP	Adenosine Diphosphate
ATP	Adenosine Triphosphate
CoA	Coenzyme A
DHBA	3,4-Dihydroxybutyrate
<i>E. coli</i>	<i>Escherichia coli</i>
HPLC	High Pressure Liquid Chromatography
LB	Luria-Bertani (Medium)
LC/MS	Liquid Chromatography / Mass Spectroscopy
MCS	Multi-Cloning Site
NAD <sup>+</sup>	Nicotinamide Adenine Dinucleotide
NADP <sup>+</sup>	Nicotinamide Adenine Dinucleotide Phosphate
PHA	Polyhydroxyalkanoate
PHB	Poly-3-hydroxybutyric acid
<i>P. putida</i>	<i>Pseudomonas putida</i>
RBS	Ribosome Binding Site
ReBiT	Retro-Biosynthetic Tool (Database)

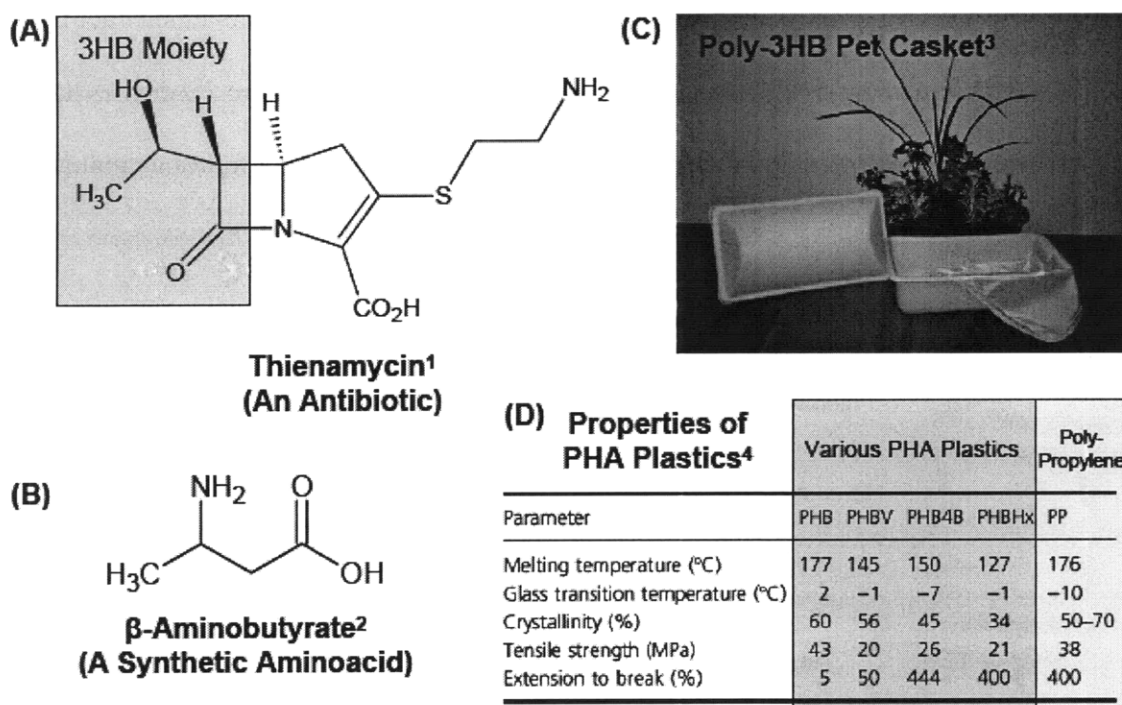
# Chapter 1: Introduction

## Motivation and Objectives

Biotechnology is rapidly becoming a major industry for the production of chemicals, fuels, and pharmaceuticals. Several modern medicines are made biologically, including the antimalarial compound artemisinin (Ro et al., 2006) and human insulin (Goeddel et al., 1979). Recently, the Energy Policy Act of 2005 (Public Law 109-58) mandated the increasing production and blending of biofuels in gasoline. The costs of non-renewable supplies of material for chemical and fuel production, namely oil, are rising as environmental and security concerns over these resources mount while the quantity of the resources dwindles. Biotechnology is being turned to not only to produce the more intricate and complex compounds for humanity, but also the everyday commodity chemicals that impact the lives of everyone daily. First and foremost, this thesis was done as a small contribution to this growing shift towards biotechnology.

More specifically, the main objective of this thesis was to develop novel, unnatural metabolic pathways for the biosynthesis of hydroxyacids and lactones. Hydroxyacids are versatile, chiral compounds that contain both a carboxyl and a hydroxyl moiety, readily allowing for the modification of this class of compounds into several useful derivatives (Lee et al., 2002; Chen and Wu, 2005). Specifically, hydroxyacids see use in the synthesis of antibiotics (Chiba and Nakai, 1985),  $\beta$ - and  $\gamma$ -aminoacids and peptides (Park et al., 2001; Seebach et al., 2001), and as chiral synthetic building blocks (Lee, 2002). Hydroxyacids can also be used directly as nutritional supplements (Tasaki et al., 1999)

and can be polymerized into biodegradable polyesters (polyhydroxyalkanoates, or PHAs) with interesting physical properties (Hazer and Steinbüchel, 2007). Figure 1.1 depicts the various uses of hydroxyacids.

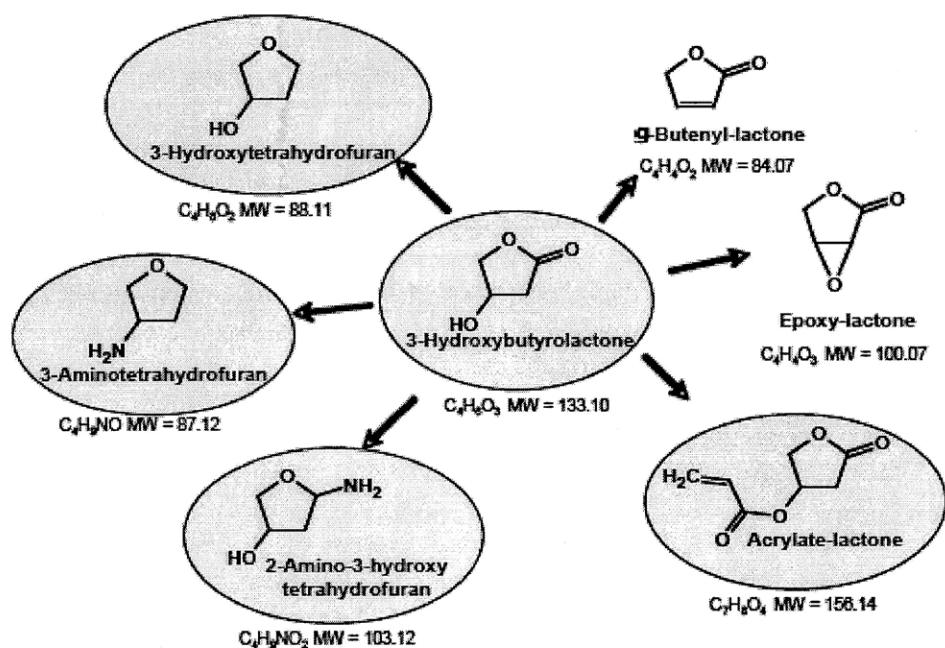


1. Chiba *et al*, Chem. Lett., 1985, 5:651-654.
2. Park *et al*, J. Chem. Res. Synop., 2001, 11:498-499.
3. Kinne Plastics Inc., Beaverton, MI, USA.
4. Verlinden *et al*, J. Appl. Microbiol, 2007, 102:1437-1449.

**Figure 1.1:** Applications of hydroxyacids for the synthesis of antibiotics (A), synthetic aminoacids (B), and plastics (C and D).

Of the lactones, 4-valerolactone (4VL), has been reported to be an ideal compound for use as a fuel and in the production of carbon-based chemicals (Horváth *et al.*, 2008). It has also seen extensive use as a component of block-copolymers for drug delivery (Attwood *et al.*, 2007; Chang and Chu, 2008), as a precursor for acrylic compounds (Manzer, 2004), and can be used as a precursor for the production of “bio-nylon”

polymers (Lange et al., 2007). Another lactone, 3-hydroxybutyrolactone (3-HBL), is widely used in the pharmaceutical industry as a building block for the class of cholesterol-reducing drugs called statins such as Crestor<sup>(R)</sup> and Lipitor<sup>(R)</sup> as well as the antibiotic Zybox<sup>(R)</sup> and the anti-hyperlipidemic medication Ezetimibe<sup>(R)</sup> (Lee et al., 2008; Lee and Park, 2009). 3-HBL has been listed as one of the top ten value-added chemicals by the U.S. Department of Energy (Werpy and Peterson, 2004). Figure 1.2 below (taken from Werpy and Peterson, 2004) illustrates a variety of derivitization routes available for 3-HBL.



**Figure 1.2:** Routes for derivitization of 3-HBL into other compounds of use proposed by Werpy and Peterson (2004). Compounds circled in yellow currently see commodity-scale commercial use.

There exists no natural, complete biochemical route for the synthesis of 3-HBL or its hydroxyacid form 3,4-dihydroxybutyrate (DHBA), nor has one yet been engineered.



Currently 3-HBL costs about \$20-50 per gram (depending on quantity and enantiopurity) and is made chemically through high-pressure hydrogenation of carbohydrates over ruthenium-based catalysts (Hollingsworth and Wang, 2000). This method produces many byproducts that are difficult to separate from 3-HBL, driving up the cost of the product. Other synthetic methods employ bacterial dehalogenation of the expensive chlorinated precursor 4-chloro-3-hydroxybutyrate (Suzuki et al., 1999; Nakagawa et al., 2006). A combination chemical and biological production scheme has been developed to reduce malic acid to 3-HBL using hydrogenation over a zinc catalyst, benzyl chloride, and immobilized lipase (Lee et al., 2008). The goal of this thesis was to create a novel completely biological route for 3-HBL synthesis. This method also suffers from the need for expensive reagents and purification. A summary of current methods for producing 3-HBL is shown in Figure 1.3 below.

R/ S	Route	Substrate	Scale	Percent ee	Yield	Remarks	Reference
S	Lipase-catalyzed, chemoenzymatic	L-malic acid	T, plants	>99.9	80	Regioselective hydrolysis	Lee et al. (2008a)
R		D-malic acid	G, lab scale	>99.9	80		
S	Chemical	Maltodextrin	G, lab scale	94	54	Chemical oxidation (cumene hydroperoxide)	Kumar et al. (2005)
S	Chemoenzymatic	Starch	Kg, plant	>99.9	57	Enzyme treatment and chemical oxidation (hydrogen peroxide)	Park et al. (2004)
R		Racemic (R,S)-CHB	G, 3-L reactor	>99	33		
S	Biological esterase- catalyzed	Racemic (R,S)-CHB	G, 5-L reactor	99	48	Recombinant <i>E. coli</i> esterase from <i>Rhizobium</i> sp. DS-S-51	Nakagawa et al. (2006, 2008)
R, S	Chemical	(R) or (S)-4-chloro-3- hydroxybutyronitrile	G, lab scale	>99.9	80	Chiral substrate (epichlorohydrin) required	Shin et al. (2005)
R	Chemical	L-arabinose	G, lab scale	>99.5	96	Chemical oxidation (hydrogen peroxide)	Hollingsworth (1999)
S		Maltodextrin	Kg, lab scale	>99.9	90		Hollingsworth (2001)

Require Purified Enzymes

Messy/Low Yield

Expensive Starting Materials

Dangerous/Expensive Reagents

**Figure 1.3:** Summary of current methods for producing 3-HBL, with their drawbacks highlighted with various colors.

Because the primary thesis goal was microbial lactone production, two different strategies for lactonization were examined. First, a system was engineered in *Pseudomonas putida* for the production of 4-valerolactone (4VL) from levulinate. In this system, the lactonization strategy was to make 4VL from its free hydroxyacid, 4-hydroxyvalerate (4HV). 4HV and 4VL exist in a pH-dependent equilibrium with one another (Teiber et al., 2003), so the *P. putida* cell system was engineered to exploit this equilibrium by performing the lactonization reaction extracytosolically in acidic medium to enhance lactone production. To further increase production, a two-phase reaction system was employed to sequester the lactone away from the cells, further driving the equilibrium towards lactone production. The second lactone production system was the 3-hydroxybutyrolactone (3-HBL) production system in *Escherichia coli*. Here, 3-HBL was produced from the activated precursor 3,4-dihydroxybutyryl-CoA. A thioesterase was employed to control production – with a thioesterase present 3,4-dihydroxybutyrate (DHBA) was produced, while without it 3-HBL was produced. Since in this system there was no lactonase, there was no appreciable equilibrium between DHBA and 3-HBL. Rather, the coenzyme A (CoA) moiety on DHBA-CoA activates the DHBA for spontaneous lactonization.

The secondary objective of this thesis was to discover pathway design or implementation strategies for enabling new metabolic routes and enhancing product titers. Given the sheer number of known enzymes (both natural and engineered) and enzyme-catalyzed reactions available, there almost certainly exists many possible metabolic pathways towards a given target compound or class of compounds (Li et al., 2004; Hatzimanikatis

et al., 2005; Prather and Martin, 2008; Martin et al., 2009). How does one sort through these possibilities and identify promising pathways? What experimental issues must be addressed to construct and implement these pathways? As a part of this thesis, a database of enzyme-catalyzed reactions called the **Retro-Biosynthetic Tool (ReBiT) was constructed to assist in metabolic pathway design by identifying enzymes that carry out a desired chemical reaction.**

### **Organization and Summary**

This thesis is organized into three distinct sections. First, metabolic pathway design and construction are discussed in Chapter 2. This chapter, which contains primarily background information on the field of metabolic pathway construction, reviews the various experimental and computational methods available for pathway design and discusses in detail the ReBiT database of enzyme-catalyzed transformations. This insight into pathway design is then put to practice experimentally in the next two sections. These sections describe two different metabolic systems constructed for the biosynthesis of two different lactones.

Chapters 3 and 4 describe the engineering of *P. putida* for the synthesis of 4VL from levulinate (Figure 3.6 and 4.1). To overcome a pH-dependent equilibrium limitation on making 4VL from its hydroxyacid in aqueous media, the lactonization reaction was catalyzed outside of the cytosol at an acidic pH. By performing the lactonization reaction outside of the cell, the pH at which the reactor occurs can be manipulated to give higher titers of lactones. Chapter 3 discusses the biosynthesis of 4HV from levulinate in *P.*

*putida*. Different culturing conditions were tested for this process, and a maximum of  $13.9 \pm 1.2 \text{ g L}^{-1}$  of 4HV could be produced from levulinate in shake flasks. This process was also found to make 3HV, and a maximum titer of  $5.3 \pm 0.1 \text{ g L}^{-1}$  of this compound could be made at the shake flask scale. Chapter 4 builds upon the 4HV-producing system from Chapter 3 by expressing the lactonase paraoxonase I (PON1) extracytosolically. The expression of PON1 outside of the cytosol was established in *E. coli* and *P. putida* (Table 4.1), and using this system  $2.1 \pm 0.1 \text{ g L}^{-1}$  and  $8.2 \text{ g L}^{-1}$  of 4VL could be produced from levulinate at the shake-flask and bioreactor scales respectively (Figures 4.5 and 4.6, respectively).

Chapters 5 and 6 describe the production of 3-hydroxyacids in *E. coli* through the 3-hydroxyalkanoic acid pathway (Figures 5.1 and 6.1). Chapter 5 details the production of the simplest 3-hydroxyacid, 3-hydroxybutyrate (3HB), as well as the CoA-manipulation tools needed to produce this molecule and the analytical methods developed to determine its chirality. Here it was found that by feeding recombinant *E. coli* glucose  $2.9 \pm 0.2 \text{ g L}^{-1}$  of (*R*)-3HB and  $2.0 \pm 0.1 \text{ g L}^{-1}$  of (*S*)-3HB could be produced at the shake-flask scale. Chapter 6 focuses on the application of the 3-hydroxyalkanoic acid pathway to the production of 3-hydroxyvalerate (3HV), DHBA and 3-HBL. Specifically, this chapter describes the successful condensation of acetyl-CoA with propionyl-CoA to form 3HV or glycolyl-CoA to form DHBA and 3-HBL. The highest titers of 3HV produced were  $537 \text{ g L}^{-1}$  of (*R*)-3HV and  $522 \text{ mg L}^{-1}$  of (*S*)-3HV. The highest titers obtained for DHBA and 3-HBL were  $3.2 \text{ g L}^{-1}$  and  $2.2 \text{ g L}^{-1}$ , respectively (Figures 6.9). This chapter also

discusses how to alter the product profiles of DHBA and 3-HBL through the expression of the *tesB* thioesterase and through an acidic post-treatment of the culture broth.

Finally, Chapter 7 first summarizes all of the pathways, products, and titers presented in Chapters 2-6 and draws broader conclusions about the case studies presented in these chapters. Generalities for pathway design are discussed, and this discussion is applied specifically to the problem of *in vivo* lactonization. This chapter concludes by providing discussion and outlook for the field of metabolic engineering – specifically the subfields of novel pathway construction and synthetic biology.

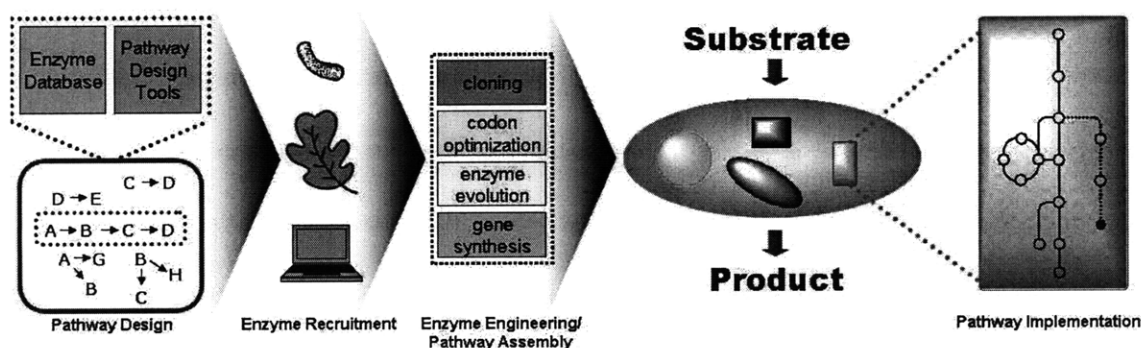
## Chapter 2: Metabolic Pathway Design

This chapter contains material published in Prather and Martin, 2008 and Martin et al., 2009.

Synthetic biology has emerged as a powerful discipline for the creation of novel biological systems (Endy, 2005; Pleiss, 2006), particularly within the subfield of metabolic pathway and product engineering (Keasling, 2008; Savage et al., 2008). Continuing efforts to characterize and understand natural enzymes and pathways have opened the door for the building of synthetic pathways towards exciting and beneficial compounds such as the anti-malarial drug precursor artemisinic acid (Ro et al., 2006) and several branched-chain alcohols for use as biofuels (Atsumi et al., 2007). The need for synthetic metabolic routes is a consequence of the fact that the array of compounds of interest for biosynthesis vastly outnumbers the availability of characterized pathways and enzymes. Several key building blocks can be made biologically (Patel et al., 2006); however, a recent report from the U.S. Department of Energy highlighted twelve biomass-derived chemical targets, only half of which have known biochemical routes (Werpy and Petersen, 2004).

With the lack of characterized natural pathways to synthesize many high-value compounds, one must learn to forge metabolic routes towards these molecular targets. Logically, it follows that for unnatural pathways, new and unnatural enzymes will be needed to compose these pathways. The parts-devices framework of synthetic biology lends itself well to this dual-sided problem of synthetic pathway creation (Endy, 2005); that is, pathways can be thought of as metabolic devices composed of individual enzyme-catalyzed reaction parts. Implicit within this framework is the idea that the challenges of

pathway creation are best approached at both the part and device levels. In this Chapter, efforts at the protein-level for broadening the array of enzyme parts that can be recruited for use in synthetic pathways are reviewed. The discussion is then expanded to pathway-level synthetic biology, where the tools available for designing metabolic pathways from enzyme-level parts and the implementation strategies for realizing these pathways experimentally are described. The overall process of pathway creation (Figure 2.1) combines experimental and theoretical components of synthetic biology at both scales. The chapter concludes with a description of the Retro-Biosynthesis Tool (ReBiT), a database of enzyme-catalyzed chemical reactions for use in the design of novel metabolic pathways.

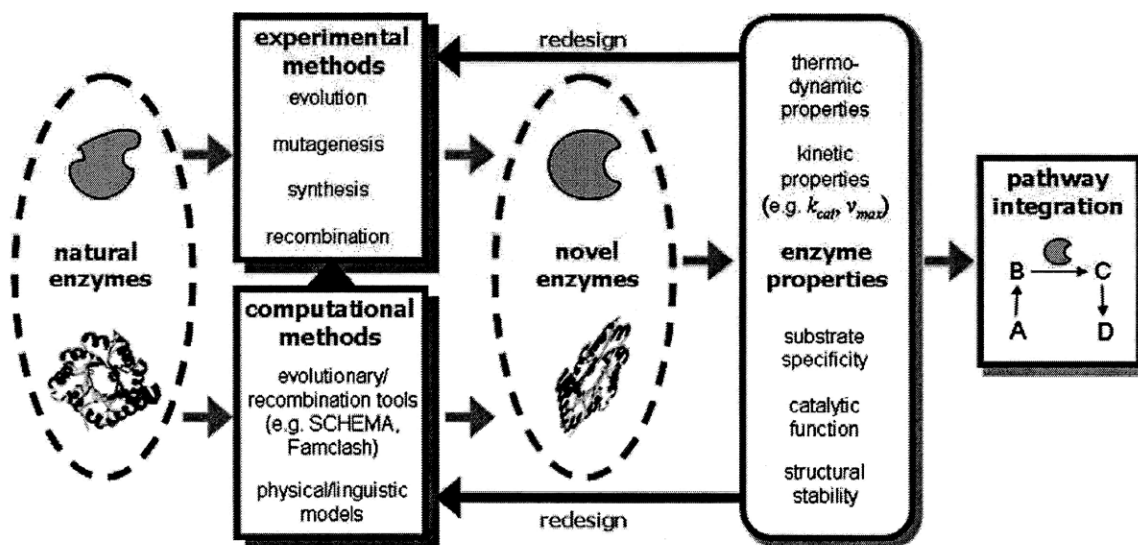


**Figure 2.1:** Overall scheme for metabolic pathway creation. The creation process includes protein-level recruitment and reengineering of enzymes and pathway-level efforts to design and assemble these enzymes into an unnatural pathway.

### **Building Novel Metabolic Routes at the Protein Level**

Through natural evolution, organisms have acquired the capacity to catalyze a multitude of diverse chemical reactions as a means to proliferate in a wide range of unique

microenvironments. Although only a small fraction of the earth's biodiversity (and an even smaller subset of its composite enzymes) has been characterized, the identification and isolation of novel proteins with unique properties or enzymatic function is a laborious procedure. One particularly promising source of new enzymes and enzymatic activities is the emerging field of metagenomics (Handelsman, 2004). Nonetheless, the physical and catalytic properties of natural enzymes often render them as incompatible or, at the very least, unoptimized for use in engineered pathways and strains. In cases where natural evolution has fallen short of industrial needs, the tools and practices of synthetic biology can be applied to aid in the creation of designer enzymes and cellular phenotypes. The challenge of building new enzymes and reengineering natural ones has been approached with the development of predictive theoretical frameworks and a range of experimental techniques (Figure 2.2).



**Figure 2.2:** Flowchart for the creation of new enzymes with experimental techniques and computational tools. New enzymes generated with these methods are examined for desired properties and either further reengineered or adapted for use in unnatural pathways.



## Theoretical Approaches

Computational tools exist to adapt the natural array of proteins for use in an increasing number of applications. For example, the effects of codon bias on expression levels (Kane, 1995; Gustafsson et al., 2004) can be resolved by design tools such as Gene Designer (Villalobos et al., 2006). Other effects such as Shine-Dalgarno sequences, promoter strength, and mRNA stability can be similarly optimized. Nonetheless, the application of these tools is still limited to the biochemical diversity found in nature. To increase the number and efficiency of biologically-catalyzed reactions, more sophisticated *in silico* techniques are needed. While full-scale protein folding and *ab initio* protein design and modeling are neither trivial nor currently practical, the use of solved protein structures, strong physical models and experimentally derived libraries allow for the design and improvement of enzymes. These theoretically designed proteins in turn have significant potential to impact pathway-level synthetic metabolism (Yoshikuni et al., 2008).

An empirical approach to synthetic protein design includes an understanding of the protein sequence/function relationship. One example is the use of a linguistic metaphor to describe a protein sequence (Searls, 1997; Searls, 2002). In language, a sentence is composed of a sequence of words whose parsed meaning is a function of not only their individual definitions but their connotations which are encoded by their type (part of speech) and their relative location to other words. Similarly, a protein ‘sentence’ is composed of residues that have not only a definitive identity but also possess chemical properties and a relative position that affect the subsequent fold and function of the

resultant protein (Przytycka et al., 2002). Building on the successes of a putative protein grammar (Przytycka et al., 2002; Naoki and Hiroshi, 1997), Loose et al. (2006) recently demonstrated its use in the design of new antimicrobial peptides. Using the TEIRESIAS algorithm (Rigoutsos and Floratos, 1998), a library with homology to known sequences restricted to below 60% was generated with approximately 50% of designs showing some antimicrobial activity. An alternative approach to modeling protein sequence/function relationships involves the use of folded protein scaffolds and quantum transition state models. Through detailed crystal structures and transition state models, Hederos et al. (2004) noted that the active site of a glutathione transferase was of the appropriate size and structure to stabilize the transition state complex of the hydrolytic degradation of a thioester. By introducing a histidine residue within the active site they were able to impart significant thioesterase activity. Finally, physics based free energy approaches have been developed to predict protein structure/function relationships in the context of antibody binding strength. While total free energy models were not a good predictor, Lippow et al. (2007) found that the electrostatic interaction contributions to total energy were well correlated with antibody binding affinity. Using this relationship, they were able to generate an improved lysozyme antibody design which demonstrated a 140-fold increase in binding. While neither of these examples fully describe protein structure/function relationships, each does offer a unique insight into the problem. Namely, they drastically reduced the sequence space of potential modifications to a manageable subset with a high probability of success. In this manner, such empirical models serve as an important tool in the design and improvement of enzymes.

Using a quantum transition state framework, great strides have been made in the *in silico* development of enzyme activities (Jiang et al., 2008; Rothlisberger et al., 2008; Kaplan and DeGrado, 2004). At the heart of these efforts is a strong understanding of the desired catalytic mechanism and its associated transition states and reaction intermediates. Once compiled, this information can be used to generate an active site of the appropriate dimensions with critical residues incorporated into appropriate locations for catalysis. At this point, the designer has two options: try to identify a suitable folded scaffold that can accommodate the active site with minimal mutations or generate a protein backbone with correctly folded active site *de novo*. Each method has its inherent advantages and challenges. While finding a host scaffold would appear to be the simpler of the two, it requires extensive searches of protein structure libraries with tools such as RosettaMatch (Zanghellini et al., 2006). Nonetheless, this approach has had some success with the catalysis of unnatural reactions such as the retro-aldol catalysis of 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone (Jiang et al., 2008) and the Kemp elimination (Rothlisberger et al., 2008). Coupled with experimental techniques, *in silico* designed enzymes can have activity levels comparable to that of evolved natural enzymes (Rothlisberger et al., 2008). In contrast, *de novo* protein scaffold development requires significant computational effort to not only consider the stability of the desired conformation of the backbone and active site but also the likelihood of destabilization. Nonetheless, Kaplan and DeGrado (2004) have successfully used such an approach to generate an O<sub>2</sub>-dependent phenol oxidase. Despite the computational overhead associated with these methods, their feasibility points to an improving and functional

understanding of protein structure/function relationships, leading to increased possibilities for the rational design of enzymes and proteins.

In the absence of rational insight, theoretical tools can assist experimental techniques in generating new and improved proteins. One common technique is protein recombination or *in vitro* shuffling which combines the best traits of two or more individual enzymes (Stemmer, 1994a; Stemmer, 1994b). However, successful recombination is contingent on shuffling at domain boundaries to ensure proper folding of each domain. The predictive algorithm SCHEMA, developed by Voigt et al. (2002), was designed to aid in the screening process of such chimeric proteins. By analyzing the nature and number of the disruptions of the intermolecular interactions, Voigt et al. were able to generate a metric correlated with the probability of active  $\beta$ -lactamase hybrids of TEM-1 and PSE-4 (2002). Subsequent studies by Meyer et al. (2003) have confirmed this correlation and used SCHEMA-guided recombination to derive functional and diverse libraries of cytochrome P-450s (Otey et al., 2004) and  $\beta$ -lactamases (Meyer et al., 2006). Another available predictive algorithm is FamClash (Saraf et al., 2004), which analyzes chimeras for the conservation of charge, volume and hydrophobicity at a given residue. Generated sequence scores have been demonstrated to be well correlated with the activities of hybrid dihydrofolate reductases. While experimental techniques are important generators of diverse protein libraries, tools such as FamClash, SCHEMA and other related sequence analysis programs enrich such chimeric libraries and vastly improve their value in the development of new and improved proteins. Currently, these tools are incapable of

predicting hits *a priori*; however, their importance in successful protein design should not be underestimated.

### **Experimental Approaches**

Rather than focusing on the prediction of protein structure and function, experimental techniques allow the improvement or modification of existing enzymes, in some instances creating entirely new enzymes and enzyme activities. These techniques include mutagenesis, enzyme engineering and evolution, and gene synthesis technology, with each boasting their own distinct advantages and inherent limitations (Bonomo et al., 2006; Alper and Stephanopoulos, 2007). Collectively, they comprise a powerful set of tools for the efficient generation of enzymes with user-specified properties. Protein recombination, for example, provides a means by which secondary structural elements, from natural or evolved proteins, can be rationally assembled in a modular fashion to integrate domains featuring desired attributes (Otey et al., 2004).

The construction of synthetic pathways typically involves the recruitment of genes from an array of sources to provide the required enzymatic function and activity (Figure 2.1). However, heterologously expressed proteins, particularly those originating from a source organism belonging to a different kingdom than that of the expression host, often suffer from poor activity as a result of dissimilarities in codon usage. In such cases, the use of synthetic genes with codon optimized sequences has been frequently employed to achieve sufficient levels of functional expression. Synthesis of a codon optimized xylanase gene from *Thermomyces lanuginosus* DSM 5826 led to a 10-fold improvement

in expression level in *E. coli* (Yin et al., 2008). Plant genes are often found to be poorly expressed in *E. coli* (Martin et al., 2001). Martin et al. (2003) synthesized a codon optimized variant of amorpha-4,11-diene synthase from the *Artemisia annua* to catalyze the conversion of farnesyl pyrophosphate to amorphadiene, a precursor used for the production of the anti-malarial drug artemisinin. As the cost associated with gene synthesis continues to decrease, imaginable applications of synthetic genes and artificial, designer proteins to include increased elements of rational design become increasingly plausible.

The versatility of directed evolution for engineering desired enzyme attributes is highlighted by a multitude of recent works employing this approach for a diverse assortment of applications, including the enhancement of thermal stability (Asako et al., 2008; Shi et al., 2008) and acid tolerance (Liu et al., 2008); promoting higher chemo-, regio-, and enantio-selectivity towards substrates (Asako et al., 2008); elimination of undesired biochemical activities (e.g., side reactions; Kelly et al., 2008); and improving heterologous expression (Mueller-Cajar et al., 2008). In the example of the stereospecific reduction of 2,5-hexanedione to (2S,5S)-hexanediol by alcohol dehydrogenase (AdhA) from the thermophilic bacteria *Pyrococcus furiosus*, laboratory evolution was used by Machielsen et al. (2008) to alter the enzyme's optimum temperature and improve its activity in recombinant *E. coli* under moderate culture conditions. Meanwhile, Aharoni et al. (2004) have achieved functional expression of mammalian paraoxonases PON1 and PON3 in *E. coli* through a directed evolution scheme that incorporated family DNA shuffling (shuffling of DNA encoding homologous genes from different genetic sources)

and random mutagenesis to achieve the first active microbial expression of recombinant PON variants. As a tool, directed evolution continues to benefit from refinements aimed at improving the efficiency at which desired mutations can be obtained from a minimal number of iterations while also reducing screening efforts (Reetz et al., 2007; Reetz et al., 2008).

In addition to improving expression and altering the thermal properties of heterologous enzymes, novel biochemical activities can be similarly engineered by the aforementioned strategies. For example, cytochrome P450 BM3 from *Bacillus megaterium* has been engineered via directed evolution using several sequential rounds of mutagenesis to alter its regioselectivity for the hydroxylation of n-alkanes from subterminal positions to that of the terminus (Meinhold et al., 2006). The approach has been employed to convert several different n-alkanes to their corresponding n-alcohols, including the hydroxylation of ethane to ethanol as a means for producing more tractable transportation fuels from petrochemical feedstocks (Meinhold et al., 2005). To promote high end-product specificity while maximizing metabolite flux, the preferential activity of an enzyme between multiple competing substrates can also be tailored. For instance, the substrate specificity of pyruvate oxidase (PoxB) from *E. coli* was altered via localized random mutagenesis to decrease its activity on pyruvate in favor of an alternative endogenous metabolite, 2-oxo-butanoate (Chang and Cronan, 2000). Synthetic pathways incorporating this PoxB mutant will accordingly display preferential synthesis of products from the four-carbon precursor. Meanwhile, Tsuge et al. (2003) utilized site directed mutagenesis to shift the substrate specificity of PhaJ, an *R*-specific enoyl-CoA

hydratase from *Aeromonas caviae* from short-chain 3-hydroxyacyl-CoA precursors towards those with longer carbon chain lengths (8 to 12). When incorporated into an engineered polyhydroxyalkanoate (PHA) synthesis pathway in *E. coli*, increased molar fractions of C<sub>8</sub> and C<sub>10</sub> 3-hydroxyacid monomer units were found to be incorporated into PHA. In this case, the capacity to distinctly manipulate the composition of PHAs makes possible the synthesis of novel bio-plastics with customizable physical properties to meet commercial requirements. The ability to finely tune the substrate specificity of an engineered enzyme is of particular importance for promoting high selectivity and product yield, as well as for reducing the ill-effects of molecular cross-talk between engineered and endogenous pathways.

At the protein level, synthetic biology aims to expand the catalog of well-characterized enzymes while also engineering novel biochemistries. Subsequent incorporation of engineered enzymes into synthetic pathways leads to the construction of devices that can be implemented to achieve a user-specified function, such as the production of biofuels or high-value pharmaceutical compounds. The design and construction of new metabolic routes from individual enzymes represents synthetic biology at the next scale, the pathway scale, and has unique challenges of its own.

### **Building Novel Metabolic Routes at the Pathway Level**

Pathway-scale synthetic biology aims to create novel metabolic routes towards both existing metabolites and unnatural compounds. Traditionally, pathway engineering has been synonymous with metabolic engineering and its toolbox has been composed of the



same tools: gene knockouts, flux optimization, gene overexpression, and the like. The ability to manipulate natural metabolism has seen many useful applications, such as improving ethanol production in *Saccharomyces cerevistiae* (Bro et al. 2006), solventogenesis in *Clostridium acetobutylicum* (Mermelstein et al., 1993; Woods, 1995), and penicillin production in *Penicillium chrysogenum* (Casqueiro et al., 2001). A key limitation in all of these examples is the confinement of pathway engineering to the manipulation of natural metabolism. Continuing advances in characterizing, modifying, and even creating enzymes (several of them discussed in the previous section of this review) now allow us to build unnatural pathways for the biological production of compounds. Understanding synthetic biology at the protein scale affords us the opportunity to apply it at the pathway scale.

As at the protein scale, pathway-level synthetic biology has been approached from both theoretical and experimental fronts. The theoretical work centers on the concept of pathway design – assembling a logical series of enzyme-catalyzed reactions to convert an accessible substrate into a valued final compound. Theoretical pathway design probes what conversions are possible and what enzyme parts need to be assembled to create a functional metabolic device. In contrast, experimental efforts focus on the construction and application of unnatural pathways and serve as powerful real-world examples of what these pathways can accomplish. Experimental approaches enable the exploration of enzyme behaviors such as substrate promiscuity and activity, both useful properties for creating unnatural pathways that cannot readily be predicted with theoretical approaches.

## **Theoretical Approaches**

Before an unnatural metabolic pathway can be built in the laboratory, it must first be designed. The goal of pathway design is to use a series of biochemically-catalyzed reactions to connect a target product molecule to either a cellular metabolite (such as acetyl-CoA,  $\alpha$ -ketoglutarate, or L-alanine) or to a feasible feedstock (such as glucose or glycerol). This can be accomplished using either natural enzymes or engineered ones. The sheer number of known enzymes (both natural and engineered) and enzyme-catalyzed reactions available means that there will almost certainly exist many possible theoretical pathways towards a given target compound (Li et al., 2004; Hatzimanikatis et al., 2005). Identifying and ranking these different possibilities are the central challenges in pathway design.

One of the first steps in pathway design is obtaining knowledge of the enzymes and enzyme-catalyzed reactions available for use in a pathway. Comprehensive protein and metabolism databases, such as BRENDA (Schomburg et al., 2004), KEGG (Kaneshisa et al., 2006), Metacyc (Capsi et al., 2006), and Swiss-Prot (Wu et al., 2006), provide a wealth of information on the pool of natural, characterized enzymes that can be recruited. More importantly, these databases reveal chemical conversions that are achievable with enzymes. As of the preparation of this thesis, there are approximately 398,000 protein entries in Swiss-Prot (build 56.2), from which the enzymes are organized into 4757 four-digit enzyme classification (E.C.) groups in the most recent version of BRENDA (build 2007.2). Because of the large number of characterized enzymes, those performing similar reaction chemistries are typically organized into generalized enzyme-catalyzed

reactions for the purposes of pathway construction (Li et al., 2004). A generalized enzyme-catalyzed reaction is defined as the conversion of one functional group or structural pattern in a substrate into a different group or structure in its product (Figure 2.3). Structural information about the non-reacting portions of the substrate is ignored, making the identification of enzymes to carry out a desired chemical conversion a much more tractable problem. However, the logical rules for assigning enzymes to a generalized reaction can be subjective (Figure 2.3). One could for instance differentiate between reactions solely on the reacting functional groups (i.e. aldehyde to alcohol) as Li and coworkers (2004) did, or one could also include information about conserved patterns of molecular structure between similar enzyme-catalyzed reactions. Furthermore, generalized enzymatic reactions do not all fall cleanly into the existing E.C. system (Figure 2.3c).

(A)

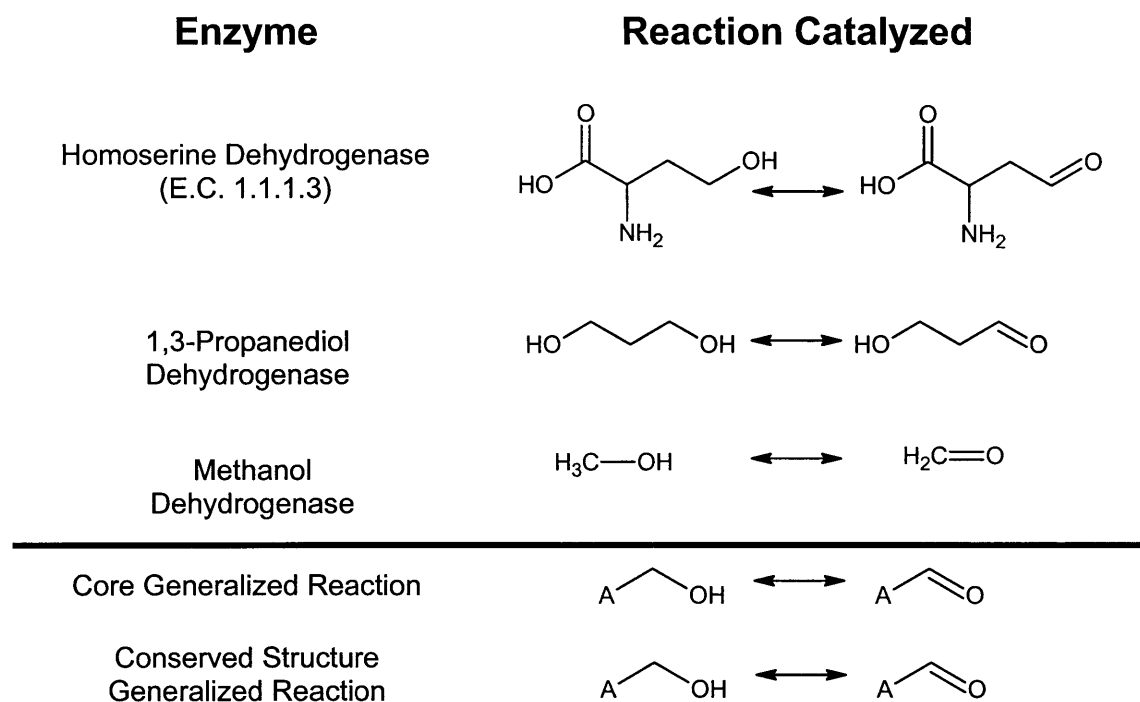


Figure 2.3a: See caption on page 34.

(B)

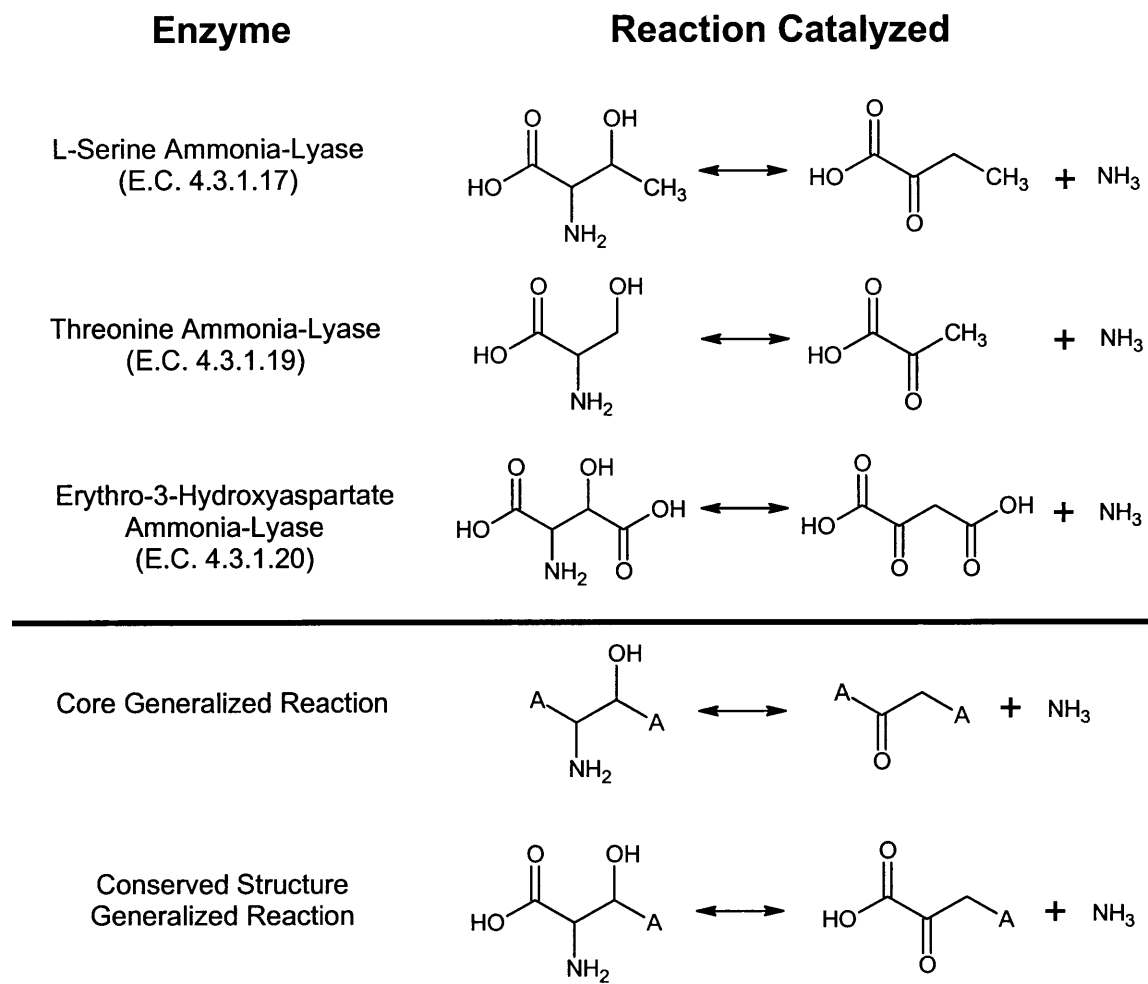
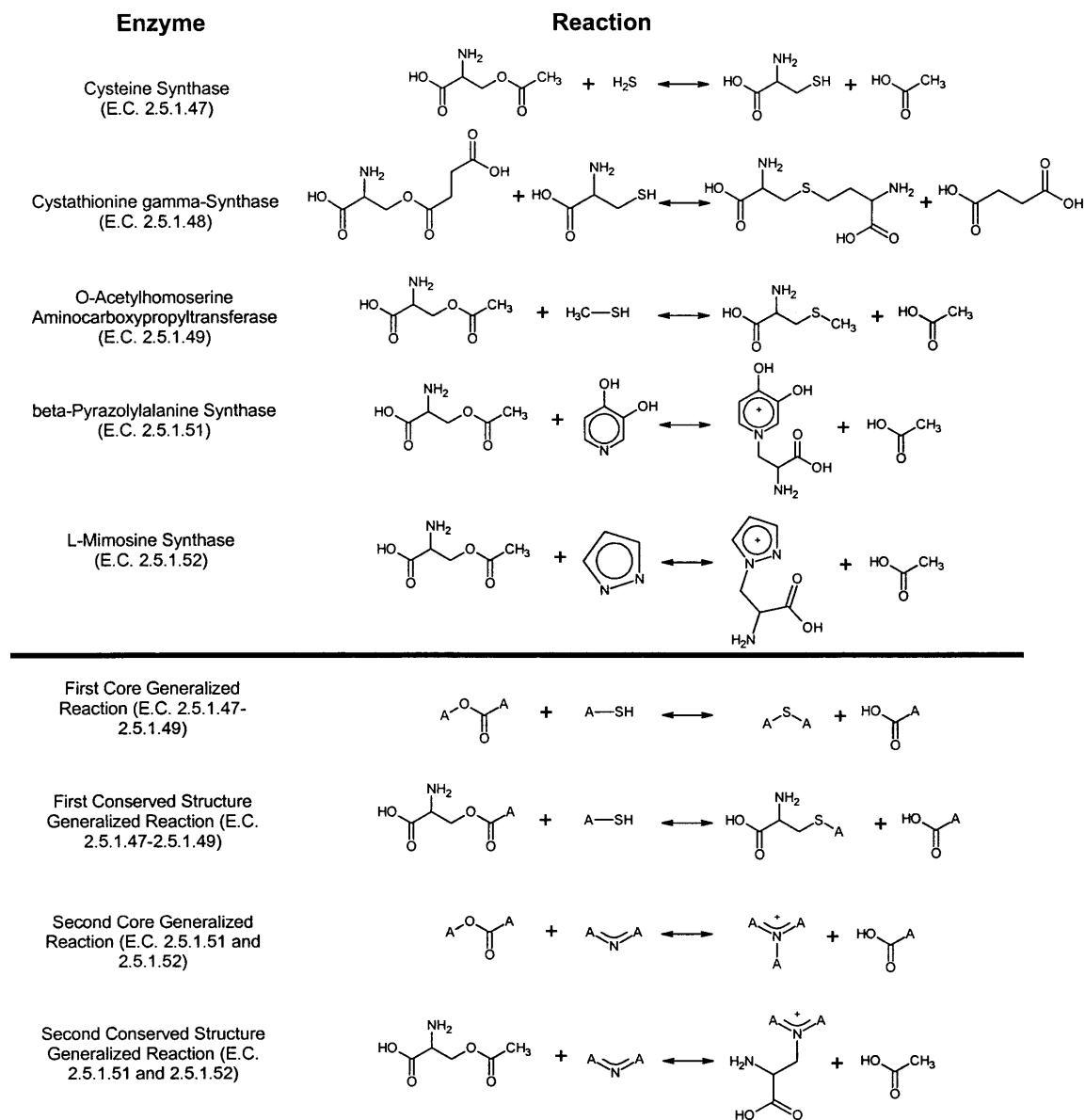


Figure 2.3b: See caption on page 34.

(C)



**Figure 2.3:** Generalized enzyme-catalyzed reactions for a subset of E.C. 1.1.1 alcohol dehydrogenases (A), E.C. 4.3.1 ammonia-lyases (B), and E.C. 2.5.1 synthases (C). The “A” atoms present in the molecular structures are wildcards. In panel A, two different methods of assigning generalized reactions, one considering only the reacting parts of the molecule (core generalized reaction) and one identifying patterns of conserved molecular structure in addition to the reacting structural elements (conserved structure generalized reaction), arrive at the same generalized reaction. In panel B, the two methods arrive at different generalized reactions, illustrating the need for a generalization standard. In panel C, a set of five enzymes within a three-digit E.C. class result in two different sets of generalized reactions, illustrating that the E.C. system does not necessarily correlate with reaction generalization.

Despite the need for a universal standard in reaction generalization, several publically-available tools utilize this approach to address the problem of pathway design. The BNICE (*B*iochemical *N*etwork *I*ntegrated *C*omputational *E*xplorer) framework allows for the discovery of numerous possible metabolic routes between two compounds (Li et al., 2004; Hatzimanikatis et al., 2005). This framework was applied to aromatic amino acid biosynthesis to find over 400,000 theoretical biochemical pathways between chorismate and phenylalanine, tyrosine, or tryptophan (Hatzimanikatis et al., 2005) and it was used to explore hundreds of thousands of novel linear polyketide structures (González-Lergier et al., 2005). As part of this thesis, the ReBiT database of over 600 conserved structure generalized enzyme-catalyzed reactions was developed and made freely available to the public at <http://www.retro-biosynthesis.com>. ReBiT accepts as input a molecular or functional group structure and returns as output all 3-digit E.C. groups capable of reacting with or producing that structure. The University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) uses a series of generalized reaction rules to propose pathways step by step, with particular emphasis on analyzing the degradation trajectories of xenobiotics (Ellis et al., 2006; Fenner et al., 2008).

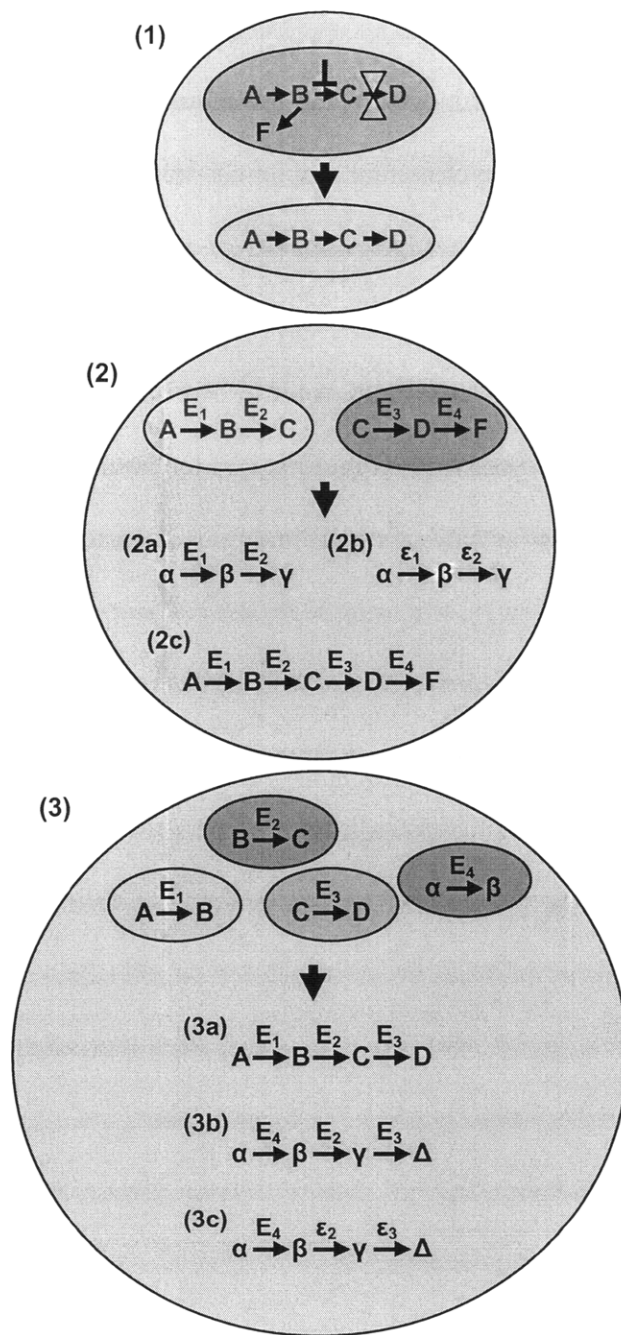
Typically multiple, and indeed in some cases, several thousand, metabolic routes can be proposed for a given compound. How does one distinguish logical, feasible pathways from frivolous, improbable ones? What metrics can be applied to judge one computationally-generated pathway as superior (i.e. more likely to be functionally constructed) to another? One way of narrowing the choice of pathways is to apply natural precedent to filter out unlikely pathway steps. In this strategy, a large set of

experimentally validated enzyme-catalyzed reactions are examined for patterns of structural change and a series of rules are developed to give preference to reaction steps containing structural changes that follow these rules. This methodology is implemented in the UM-BBD to avoid the “combinatorial explosion” that results when considering all the possible pathways that any given compound can take (Fenner et al., 2008). Another ranking strategy is to calculate the thermodynamic favorability of the steps and to penalize pathways involving steps which are energetically unfavorable. This approach is taken by the BNICE framework (Hatzimanikatis et al., 2005) using a functional group contribution method (Jankowski et al., 2008) to compute the overall change in Gibbs energy for each individual pathway step. A new pathway modeling tool, DESHARKY, quantifies and employs metabolic burden as a metric for judging unnatural pathways and, in particular, how they are connected to cellular metabolism (Rodrigo et al., 2008). DESHARKY is a Monte Carlo-based algorithm that estimates the transcriptomic and metabolic loads on cells expressing unnatural pathways and calculates the decrease in specific growth rate as a result of these additional burdens. There are still other possibilities for pathway ranking, such as the number of pathway steps taken, the known substrate specificities (or lack thereof) of the enzymes involved in each pathway, or the availability and diversity of homologous enzymes to test at each pathway step. One of the key challenges in pathway design is scoring pathways in a robust and balanced manner, and only as more non-natural pathways are designed and built will there be a better understanding as to which of these metrics are relevant and useful.



## **Experimental Approaches**

With a target compound and a proposed metabolic route to reach that compound in hand, one is now ready to begin experimental implementation of that pathway. Synthetic pathway construction occurs over several shades of novelty – from recreating natural pathways in heterologous hosts to creating synthetic pathways that parallel natural ones to building completely novel metabolic routes towards unnatural compounds from multiple, ordinarily unrelated enzymes (Figure 2.4). Here situations in which non-natural pathways prove useful and several general strategies for creating these pathways are discussed.



**Figure 2.4:** Strategies for synthetic pathway creation arranged in increasing degrees of departure from nature. A, B, C, D, F,  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\Delta$  represent metabolites, E represents an enzyme catalyzing a reaction, and  $\epsilon$  represents an engineered enzyme catalyzing a reaction. In (1), a natural pathway in its native host is transferred to a heterologous host, decoupling it from native regulation. This strategy is limited to the production of natural products using natural pathways. In (2), new pathways are made in parallel to natural ones through the use of promiscuous enzymes (2a), enzyme engineering (2b), or combinations of natural pathways (2c). Strategies 2a and 2b allow for the synthesis of new, non-natural products, while 2c allows for the creation of new metabolic routes between natural metabolites. Strategy 3 represents *de novo* pathway construction, where individual unrelated enzymes are recruited to form entirely unnatural pathways. This can be done using native enzyme activities (3a), promiscuous enzymes (3b), engineered enzymes (3c), or combinations thereof.

Through the course of evolution, nature has assembled many pathways towards several useful compounds, such as the biofuel and solvent 1-butanol in *Clostridium acetobutylicum* (Jones and Woods, 1986; Dürre et al., 2002; Lee et al., 2008), the C<sub>5</sub> terpenoid building block isopentenyl pyrophosphate (IPP) in *Saccharomyces cerevisiae* (Seker et al., 2005), and the biopolymer polyhydroxybutyrate (PHB) in *Ralstonia eutropha* (Wang and Yu, 2007). These pathways have physiological roles within their native hosts; for example, the butanol pathway from acetyl-CoA in *C. acetobutylicum* serves as an electron sink to regenerate NAD<sup>+</sup> for glycolysis while deacidifying its environment (Jones and Woods, 1986). Pathways in nature are optimized through evolution to accomplish their physiological objectives, yet in most cases of pathway engineering, it is desired to maximize the production of a target molecule in a pathway rather than to accomplish a physiological goal. Butanol production in *C. acetobutylicum*, for instance, is constricted by cellular regulation tying it to pH, redox conditions, and sporulation (Dürre et al., 2002; Lee et al., 2008). The transference of natural pathways into heterologous hosts isolates these pathways from their regulatory elements and represents a first small step towards the creation of non-natural metabolism. While heterologous pathway expression is limited to only pathways found in nature, it nonetheless has proven effective in enhancing product titers and/or deregulating compound production for a wide array of products, including the compounds in the examples above (Atsumi et al., 2007; Kang et al., 2008; Martin et al., 2003; Pitera et al., 2007).

The next level of novelty in synthetic pathway construction is creating metabolic routes that parallel natural pathways, typically by capitalizing on enzymatic promiscuity or enzyme engineering to operate natural or near-natural pathways on non-natural substrates. This pathway construction strategy allows for the biosynthesis of truly unnatural compounds. Returning to the PHB example, recombinant *R. eutrophia* have been shown to incorporate sulfur-containing short- and medium-chain length thioacids into polythioester co-polymers (Ewering et al., 2002). The synthesis of these completely unnatural polymers was made possible by taking advantage of the relatively broad substrate specificity of polyhydroxyalkanoate (PHA) synthases (Hazer and Steinbüchel, 2007), and because of that broad substrate specificity, hundreds of different monomer units of various sizes (C<sub>3</sub>-C<sub>16</sub>) and substituents have been incorporated into PHA co-polymers (Steinbüchel and Valentin, 1995). Another example of parallel pathway construction is the synthesis of triacetic acid lactone from acetyl-CoA by expressing an engineered fatty acid synthase B from *Brevibacterium ammoniagenes* (Zha et al., 2004). This multifunctional enzyme has many domains designed to catalyze the various reductions and condensations necessary for fatty acid synthesis (Meurer et al., 1991). By specifically inactivating the ketoacyl-reductase domain of this fatty acid synthase, the enzyme could no longer use NADPH to reduce its acetyl-CoA condensation products, causing them to circularize into triacetic acid lactone rather than forming linear fatty acids. Finally, natural products can be synthesized by arranging whole or partial pathways to form a mixed, synthetic metabolic route. For example, the theoretical yield of L-glutamate was improved from 1 mol glutamate per mol glucose to 1.2 mol per mol by augmenting the native *Corynebacterium glutamicum* pentose phosphate pathway with

a phosphoketolase from *Bifidobacterium lactis* (Chinen et al., 2007). This strategy allowed for the production of acetyl-CoA without the loss of carbon caused by pyruvate decarboxylation to acetyl-CoA and resulted in increased glutamate titers and productivity.

One of the most promising (and challenging) strategies for building synthetic pathways is *de novo* pathway construction: the creation of pathways using disparate enzymes to form entirely unnatural metabolic routes towards valuable compounds. This method of pathway building does not rely upon natural precedent, but rather allows one to build entirely new metabolite conduits from individual enzymatic pieces. As a result, this approach allows for the biosynthesis of the widest array of compounds. On the other hand, this strategy is the most difficult to realize given that for a completely unnatural pathway, there may not be a complete set of appropriate known enzymes in nature to build it. *De novo* pathway construction illustrates the need for a more complete set of enzymatic tools for use in building synthetic pathways, and frequently this strategy is coupled with enzyme engineering or the exploitation of enzymatic promiscuity to compensate for the absence of a natural enzyme to execute a desired conversion step.

Because of the challenge in creating functional *de novo* pathways, few examples exist. However, those that are available describe the biosynthesis of a wide range of useful compounds and illustrate the utility of the approach. For instance, a pathway for the biosynthesis of 1,2,4-butanetriol from D-xylose and L-arabinose was assembled using pentose dehydrogenases and dehydratases from *Pseudomonas fragi* and *E. coli* and

benzoylformate decarboxylase from *Pseudomonas putida* (Nui et al., 2003). In this case, multiple decarboxylases were screened to find a promiscuous decarboxylase from *P. putida* capable of acting on a 3-deoxy-glyceropentulosonic acid intermediate in the pathway. Another example of exploiting substrate promiscuity in *de novo* pathway design is in the synthesis of several higher biofuels such as 2-methyl-1-butanol, isobutanol, and 2-phenylethanol from glucose in *E. coli* (Atsumi et al., 2007). Here, several 2-keto-acid decarboxylases were screened to identify one from *Lactococcus lactis* for use in creating alcohols from 2-ketoacids (when combined with native *E. coli* alcohol dehydrogenase activity). In a third example, a synthetic pathway for the unnatural aminoacid phenylglycine from phenylpyruvate was made by combining hydroxymandelate synthase, hydroxymandelate oxidase, and D-(4-hydroxy)phenylglycine aminotransferase activities from *Amycolatopsis orientalis*, *Streptomyces coelicolor*, and *P. putida* (Müller et al., 2006). Finally, engineered enzymes can be employed to create *de novo* pathways, as in the recent case of the synthesis of 3-hydroxypropionic acid from alanine in *E. coli* (Liao et al., 2007). Here, a lysine 2,3-aminomutase from *Porphyromonas gingivalis* (Brazeau et al., 2006) was evolved to have alanine 2,3-aminomutase activity, allowing for the biosynthesis of  $\beta$ -alanine. Combining this evolved enzyme with  $\beta$ -alanine aminotransferase and endogenous alcohol dehydrogenase activities afforded the final 3-hydroxypropionic acid product. Another very recent work utilizes engineered pyruvate decarboxylase and 2-isopropylmalate synthase for the synthesis of non-natural alcohols from 2-ketoacids in *E. coli* (Zhang et al., 2008). By engineering the enzymes responsible for elongating 2-

ketoacids and carrying out their decarboxylation and reduction, the production of a broader array of longer-chain alcohols was enabled.

### **Outlook for Pathway Design**

The design and assembly of unnatural metabolic pathways represents a young and exciting field with the potential to supplement, expand upon, or even replace current industrial processes for the production of fine and commodity chemicals. Synthetic pathway engineering integrates many components and consequently is highly interdisciplinary (Figure 2.1). Key issues that need to be overcome in pathway design are (1) establishing a standard for generalized enzyme-catalyzed reactions, (2) capturing enzyme substrate preferences in these generalized reactions, and (3) determining the pathway metrics that correlate with successful pathway construction. Overcoming the first two challenges will allow for the creation of the next generation of pathway design tools that better account for enzyme behavior, while conquering the last challenge will afford us the ability to rank and choose metabolic pathways and refine the results from design tools. For experimentally implementing unnatural pathways, the central challenge is the limited number of characterized enzymes for the construction of new pathways. In particular, there is great demand for both promiscuous natural enzymes and engineered enzymes to perform specific desired reactions.

The need for new enzymes has given rise to several theoretical frameworks for relating protein sequence, structure, and function. These frameworks each address a piece of the problem – energetics, active site catalysis, and protein backbone structure, etc. – but the ability to routinely build whole enzymes is still in the distant future. In the meantime,

mimicking active sites, backbones, and protein linguistics from nature has proven fruitful in creating novel proteins. Experimental evolution and chimeragenesis of enzymes are standard ways of imparting unnatural properties, particularly in the absence of detailed information about the protein. The power of these experimental techniques is primarily limited by the size of the resulting enzyme libraries and the throughput of the screen to analyze them. Computational tools such as SCHEMA (Voigt et al., 2002) and Famclash (Saraf et al., 2004) can assist in focusing and enriching these libraries.

As biotechnology is increasingly relied upon as a means for chemical production, progress on the creation of new enzymes and unnatural pathway design and construction will flourish. These new pathways must still be expressed within a cellular context, thus improving and understanding unnatural pathway efficacy at a systems level will be important for shattering barriers in pathway expression and product titer. For example, application of flux balance analysis (Edwards et al., 2002) can guide systems-level integration of non-natural pathways with host metabolism. Furthermore, redox balancing and cofactor regeneration with respect to new pathways are critical to minimize their burden on the host cell (Endo and Koizumi, 2001). Systems-level functionality can also be coupled with unnatural pathways, for instance in the delivery of recombinant microbes to a cancerous tumor (Anderson et al., 2006). Such microbes could be engineered to simultaneously produce and deliver a drug. Established and recent advances in metabolic engineering, such as global transcription machinery engineering (Alper and Stephanopoulos, 2007), can complement synthetic biology in this regard, leading to improved performance of novel pathways.



### **The ReBiT Database for Pathway Design**

As part of this thesis, ReBiT (Retro-Biosynthetic Tool) was developed as a biochemical pathway design tool to identify enzymes catalyzing specific functional group transformations in their substrates. ReBiT allows researchers to query for enzymes that convert one specific functional group, an alcohol for instance, into another, such as an aldehyde. Combinations or clusters of multiple functional groups, like diols, ketoacids, or hydroxyamines, may also be queried using ReBiT. Rather than having to manually search each individual enzyme or enzyme class for a desired biochemical conversion, ReBiT allows such searches to be executed in a single query. Thus ReBiT allows the researcher to rapidly identify candidate enzymes for constructing specific arrangements of functional groups present in a desired final product. This functionality would be valuable to researchers interested in the construction of unnatural biochemical pathways towards high-value products.

Enzymatic synthesis of chemical compounds represents a powerful alternative to chemical synthesis, particularly in the case of chiral or highly functionalized molecules (Johannes, 1996). With the advent of recombinant DNA technology and the development of the field of metabolic engineering, it is now possible to assemble whole biochemical pathways using enzymes from multiple organisms. Furthermore the enzymes themselves are remarkable biocatalysts: specific to generally a small range of substrates yet surprisingly promiscuous when given the right reactants and conditions. This enzymatic promiscuity can be extended to unnatural substrates and amplified through genetic engineering techniques (*e.g.*, enzyme evolution) to obtain custom-made biocatalysts for

the biochemical production of a given target molecule (Farinas et al., 2001; Tao and Cornish, 2002).

Of interest here however is assembling enzymes (whether engineered or native) into *unnatural* reaction pathways. This new field of pathway development combines the breadth and catalytic power of enzymes and the potential of enzyme engineering with the field of metabolic engineering to realize the vision of efficient, versatile, and environmentally-friendly synthesis. This vision, a world where chemicals, drugs, fuels, etc. are largely produced by renewable, programmable microbes, is only just beginning to materialize. With sustained work in these fields, the barriers between robust pathway design, pathway construction, and reality will eventually fade.

Before any enzymatic pathway can be built, however, it must be designed. The goal of “retro-biosynthetic” pathway design is to connect a target small molecule product to the metabolism of the host cell through a series of enzyme-catalyzed reactions. One starts the design process at the desired target molecule and then finds enzymes that will produce that target molecule from various other substrates. These enzymes then lead the designer backwards up a hypothetical biochemical pathway to new intermediate compounds from which the process of finding enzymes and moving to new compounds is repeated, until the pathway is connected to either cellular metabolism or to a substrate to be supplied. The design process is analogous to playing a game of chess on a giant metabolic chessboard, with the goal of moving a playing piece from a starting square on the board to one or more end squares. Each possible substrate or metabolite represents a

square where the designer can place his piece and each enzyme represents a piece that can only make certain moves, constraining the designer's ability to go from one place to another. Furthermore, enzyme engineering serves to open up new moves for the designer by bridging two squares on the board that would otherwise be disconnected. Pathway design itself then is the process of figuring out how to get from one point on this board to another quickly and efficiently.

The most difficult part of the pathway design process lies in identifying enzymes that (1) allow you to make valid moves on this chessboard and (2) that move you in the right direction. The breadth and range of chemistries that enzymes catalyze is diverse enough that for many molecules, several different reactions are possible. This enzymatic diversity also means that there are almost certainly many different combinations of enzymes that can theoretically bridge two compounds with one another over the course of a pathway (Li et al., 2004). A tool is thus needed to assist the designer in sifting through this diversity, so that enzymes best connecting a target molecule to cellular metabolism can be readily identified.

Popular enzyme databases like BRENDA (Schomburg et al., 2004), MetaCyc (Capsi et al., 2006), and KEGG (Kanehisa et al., 2006) however do not readily lend themselves to pathway design, because the information contained within these databases cannot be easily queried based on the chemical or functional group transformations that the enzymes themselves catalyze. Specifically, while these databases do allow users to query their contents based on static substructures, they do not allow querying for dynamic

chemical changes that either create or convert a given substructure. This type of query would be useful when identifying enzymes that create structures and compounds of interest, i.e. in finding what moves on the metabolic chessboard are possible from a given molecule. To address this issue, ReBiT, a database of enzymatic transformations, was created to assist synthetic biologists and metabolic engineers in designing and assembling novel pathways towards valuable target compounds.

ReBiT (<http://www.retro-biosynthesis.com>) allows a user to query molecular structures against a database of enzyme-catalyzed chemical transformations. ReBiT returns hits (enzyme-catalyzed chemical transformations) in which the queried structure is produced, consumed, or changed in some way. Selecting a particular transformation lists all three-digit Enzyme Classification (E.C.) numbers that contain enzymes catalyzing that transformation as well as links to the ExPASy database (Gasteiger, 2003) for more detailed enzyme-specific information. ReBiT also displays graphical representations of all structures involved in each transformation and any cofactors or co-substrates used.

## **Methods**

ReBiT allows a user to query enzyme-catalyzed reactions by the “canonical” substrates or products involved in the reaction. For example, if a hypothetical group of enzymes catalyze the conversion of propionate into *n*-propanol and butyrate into *n*-butanol, two structures would be placed in ReBiT, one for a carboxylate (the canonical substrate) and one for a primary alcohol (the canonical product). Parts of molecules that are not canonical or that are not directly conjugated with the reacting portion of the molecule are

not included in the database. ReBiT is thus a searchable repository of dynamic molecular changes in enzyme-catalyzed reactions.

Figure 2.3 illustrates the process of mapping enzymes and E.C. numbers to the chemical structures on which they catalyze reactions. Figure 2.3a represents the relatively simple case of a subset of E.C. 1.1.1 reductases that interconvert primary alcohols and aldehydes. In generating the canonical reaction for this set of enzymes, all of the reactions catalyzed within this 3-digit E.C. class were compared and it was determined (1) what types and patterns of structural changes take place in the substrates participating in each reaction and (2) what non-reacting structural elements are conserved across each type of conversion within these reactions. For the reactions shown in Figure 2.3a, the pattern of structural change is the interconversion of an alcohol (R-OH) and an aldehyde (R-CH=O), while the conserved element across each reaction is the need for at least one carbon to be attached to the reacting group. The generalized reaction placed in ReBiT to represent these enzymes then is  $A-CH_2-OH \leftrightarrow A-CH=O$ , where A is a wildcard, and ReBiT records this reaction as the canonical structures A-CH<sub>2</sub>-OH and A-CH=O linked to each other by these E.C. 1.1.1 enzymes. Figure 2.3b depicts a somewhat more complicated example of canonical reaction deduction with a set of interconversions between 3-hydroxyaminoacids and 2-ketoacids that occur within E.C. 4.3.1. In these reactions, the pattern of structural change is the loss of the hydroxyl group at the 3-position and the conversion of the 2-amino group into a ketone. The carboxylate group in each of these reactions remains unchanged, yet this functional group is conserved across

each of the reactions. This carboxylate group is thus included in the canonical structures placed in ReBiT to represent these reactions.

Typically a given 3-digit E.C. class contains multiple types of functional group conversions. For situations like this one canonical structure-reaction mapping would be stored in ReBiT for each different pattern of conversion that was seen. Figure 2.3c illustrates this process for E.C. 2.5.1. Within this enzyme class there exists a set of enzymes that thiolate the ester bond of an aminoacid substrate (E.C. 2.5.1.47-2.5.1.49) and a set of enzymes that adds an aminoacid group to a nitrogen atom in a heteroaromatic ring (E.C. 2.5.1.51 and 2.5.1.52). Since these two groups of enzymes each produce substantially different functional groups (thiols and thioesters vs. heteroaromatic aminoacids) a canonical structure-reaction mapping was generated in ReBiT for each group. The ReBiT interface notes the 4-digit E.C. numbers of the enzymes in each group to clearly distinguish these two groups and the chemical conversions that they catalyze.

The ReBiT database itself is organized into two tables. The first table, an *Enzymes* table, matches each stored structure with one or more three-digit E.C. numbers. The structures are stored in the database as SMILES text strings (Weininger, 1988). The Enzymes table also associates each E.C. number with additional information such as cofactors and commentary for each reaction. The second table is a *Functional Groups* table that associates each SMILES structure in ReBiT with a descriptive chemical name and a list of all functional groups (alcohol, aldehyde, etc.) contained within each structure. These two tables, taken together, allow a user to search ReBiT by specifying a SMILES string,

or a structure or substructure (which is subsequently converted into a SMILES string). All functional groups can also be browsed by either structure or name.

The molecular structures in ReBiT were generated by manually examining each enzyme-catalyzed reaction in BRENDA (Schomburg et al., 2004) and MetaCyc (Capsi et al., 2006) and deducing canonical structures for both the substrates and products associated with these reactions. These structures were then gathered into a table, associated with the enzyme class or classes that generate or consume the structures, and purged of any duplicate structures or entries. A web-based interface was developed to give the central database a user-friendly form and to allow its contents to be queried. The web interface integrates the MySQL database with ChemAxon's JChem software (Csizmadia, 2000) to execute searches of ReBiT's two tables by first converting the user's structure query into a SMILES string and then comparing the converted query to the SMILES strings stored in ReBiT's database. JChem also normalizes SMILES queries since a given structure can be represented with multiple SMILES notations. After this comparison, the web interface displays any matches found to the user in an organized and easy-to-read manner.

Currently the ReBiT database contains 605 unique canonical structures derived from approximately 4,700 enzyme classes representing almost 90,000 enzymes. The 605 canonical structures are associated with each other through 637 unique enzymatic conversions in ReBiT. The definition of the E.C. nomenclature means that all enzymes with a specific 3-digit number should catalyze the same type of reaction but with different substrates (Li et al., 2004). However, there are only 246 3-digit E.C. classes as

of the preparation of this manuscript, meaning that there are on average 2.6 unique enzyme-catalyzed reactions per 3-digit E.C. class. As a specific example of this phenomenon, note that cysteine synthase (E.C. 2.5.1.47) fixes sulfur from hydrogen sulfide to form cysteine while thiamine pyridinylase (E.C. 2.5.1.2) acts as an N-transferase between aromatic heterocycles. Both enzymes catalyze markedly different chemistries yet both are classified under the same 3-digit E.C. heading (E.C. 2.5.1). Further examples of discrepancies between E.C. number and enzyme chemistry are E.C. 1.2.3.3 (pyruvate oxidase, which simultaneously decarboxylates and phosphorylates pyruvate) vs. most other enzymes in E.C. 1.2.3 (which oxidize aldehydes into carboxylic acids) and E.C. 4.2.1.22 (cystathionine beta-synthase, which forms a thioether from the alcohol group of serine and the thiol group of homocysteine) vs. E.C. 4.2.1.4, 4.2.1.62, 4.2.1.96, and others which dehydrate an alcohol into an alkene. These examples illustrate that while the E.C. system is helpful in organizing and cataloging the breadth of known enzymes, it cannot be relied upon to distinguish between enzyme groups on the basis of the functional group transformations that they catalyze. As new enzymes are created or identified, the weaknesses of this system will be amplified.

The E.C. system nonetheless serves as a common nomenclature that most researchers invoke when referring to various characterized enzymes. Because of this, the enzyme-reaction mappings encoded in ReBiT are referenced back to specific 3-digit E.C. numbers. An important consequence of this is that some of ReBiT's enzyme-reaction mappings span multiple 3-digit E.C. classes, while several 3-digit E.C. classes contain multiple enzyme-reaction mappings. The former statement highlights ReBiT's utility in



identifying similar enzyme chemistries dispersed across several E.C. classes – when queried for chemical conversions present in multiple E.C. classes ReBiT will present the user with hits from each E.C. class found to contain the queried conversion. The issue of multiple chemistries in one 3-digit E.C. class is addressed in ReBiT by the citation of specific 4-digit E.C. numbers with each mapping, so that the user is not misdirected to an enzyme within the 3-digit E.C. class that does not catalyze the queried conversion.

### **Implementation**

The ReBiT homepage presents four different ways to access the information on enzyme conversions. Structures of interest are found either by browsing the functional groups or submitting a specific query to ReBiT. Users may also browse all 605 structures, or browse a list of names associated with the functional groups. Currently the structures in ReBiT may be searched by drawing the (sub)structure in a Java applet window, or by inputting the structure as a SMILES text string.

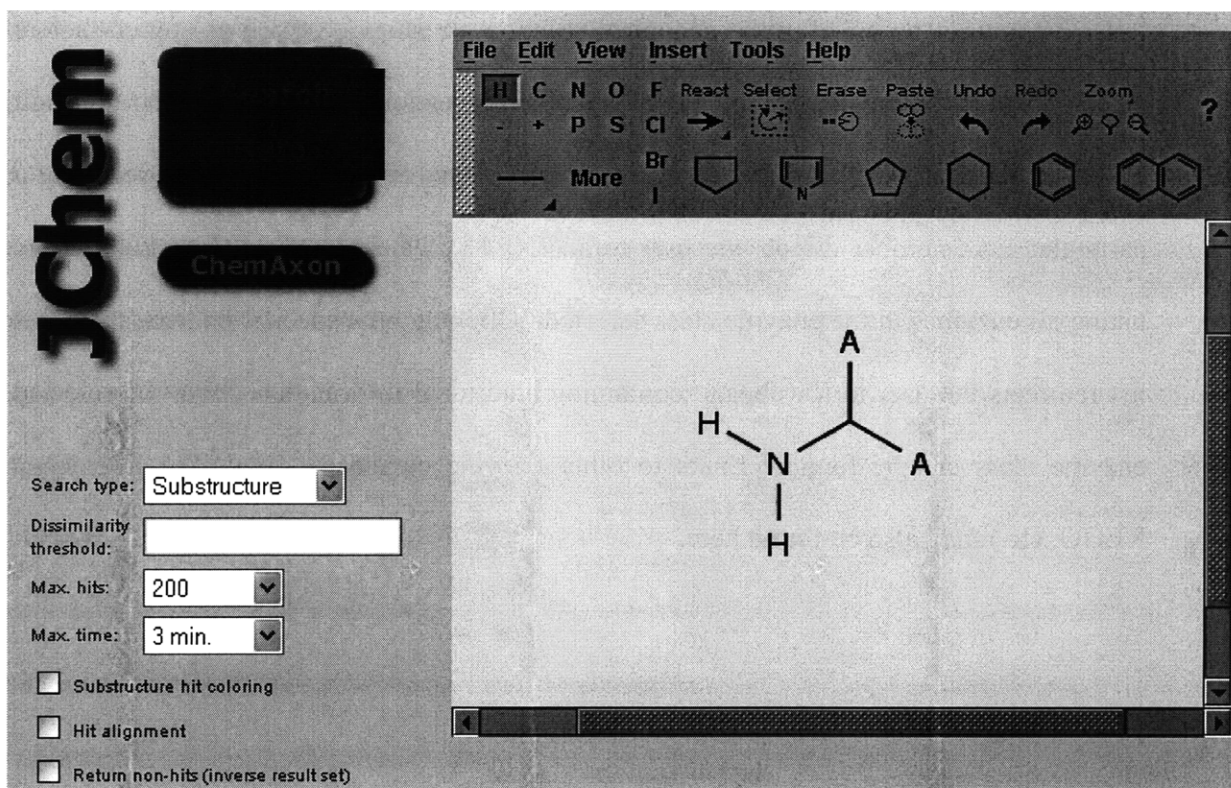
- *Browse Functional Groups* returns a table containing all 605 structures currently stored within ReBiT. Highlighting a particular functional group and choosing “Find Enzymes” will open a new window listing the enzymes, other substrates or products and co-factors involved in the reactions that consume or create the structure. This table is described in more detail below. The *Query Structure* and *Query SMILES* options can also be accessed from within the table of functional groups.

- *Browse Names* brings up an alphabetical listing of all functional groups in ReBiT. Selecting a particular name will return a secondary list with more detailed naming and

structural information. For example, selection of “Diol” produces a list of several 1,2- and 1,3-diols with structures. Selection of the text string corresponding to a particular structure returns the table of corresponding enzymes.

- *Query Structure* brings up a drawing board that allows one to directly draw a skeletal structure (Figure 2.5). After drawing the desired structure, the user selects a search strategy (substructure, superstructure, similarity, exact match, etc.) from the drop-down menu on the left-hand side of the screen. More stringent search strategies return fewer hits closer to the queried structure, while less stringent searches return more hits consisting of a broader array of structures. Selecting the search button initiates the query for the drawn structure.

- *Query SMILES* brings up a text box where SMILES code may be input. After inputting the desired SMILES string, the user can again select a search strategy and press the search button to initiate the query.



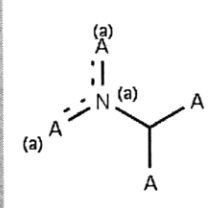
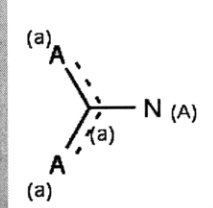
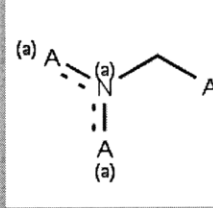
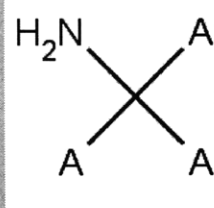
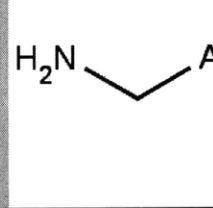
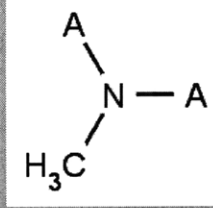
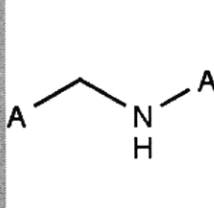
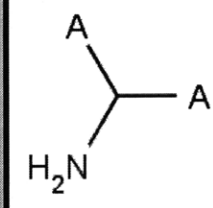
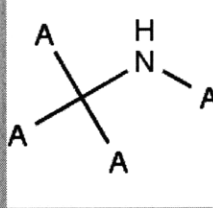
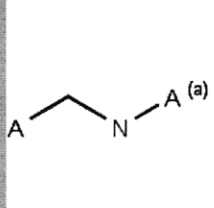
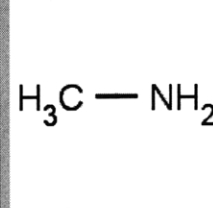
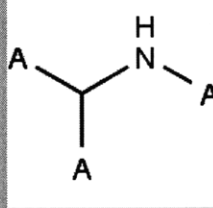
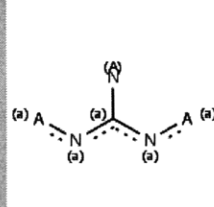
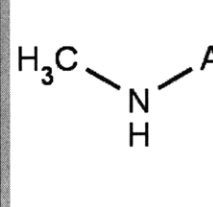
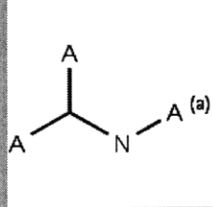
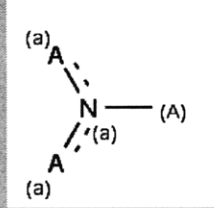
**Figure 2.5:** Submission page for a structural query to ReBiT. The box in the center of the screen is used to draw a structure to query, while the menu on the left allows the user to specify the search type (substructure, superstructure, exact match, etc.), search time, and other conditions on the hits returned by ReBiT. In the structure drawn in the palette, wildcard A atoms are used in a generic query for disubstituted primary amines.

After querying ReBiT by structure or SMILES, a list of matching structures will be displayed (Figure 2.6). The user can then select the structure of interest and press the “Find Enzymes” button on the menu. A table of all enzyme-catalyzed reactions involving the selected structure will be shown, along with 3-digit E.C. numbers corresponding to each of the reactions (Figure 2.7). In parenthesis below each 3-digit E.C. group is the total number of enzymes in that 3-digit group. This number serves as an initial assessment of substrate diversity, an important metric for choosing enzymes for new pathways. Additionally for each enzyme group, graphical representations of all

ReBiT structures involved in the canonical reaction are shown as well as any co-factors or co-substrates involved. Comments (typically suggestions as to what specific 4-digit E.C. numbers catalyze each reaction) are also displayed for each reaction. Selecting a particular E.C. number directs the user to an ExPASy database (Gasteiger, 2003) page listing all enzymes in the enzyme class selected. Clicking on a specific enzyme from this list redirects the user to a webpage containing additional information about the selected enzyme class can be found. Links to other enzyme databases (BRENDA, MetaCyc, KEGG, etc.) may also be found here.

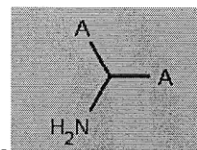
## Enzymes

Search took 0.022 seconds.

 <p>ID: 10 CN MW: 26.01744 DISS: 0.0</p>	 <p>ID: 16 CN MW: 26.01744 DISS: 0.0</p>	 <p>ID: 18 CN MW: 26.01744 DISS: 0.0</p>
 <p>ID: 38 CH2N MW: 28.03332 DISS: 0.0</p>	 <p>ID: 65 CH4N MW: 30.0492 DISS: 0.0</p>	 <p>ID: 71 CH3N MW: 29.04126 DISS: 0.0</p>
 <p>ID: 72 CH3N MW: 29.04126 DISS: 0.0</p>	 <p>ID: 107 CH3N MW: 29.04126 DISS: 0.0</p>	 <p>ID: 182 CHN MW: 27.02538 DISS: 0.0</p>
 <p>ID: 294 CN MW: 26.01744 DISS: 0.0</p>	 <p>ID: 320 CH5N MW: 31.05714 DISS: 0.0</p>	 <p>ID: 324 CH2N MW: 28.03332 DISS: 0.0</p>
 <p>ID: 445 CN3 MW: 54.03092 DISS: 0.0</p>	 <p>ID: 487 CH4N MW: 30.0492 DISS: 0.0</p>	 <p>ID: 530 CN MW: 26.01744 DISS: 0.0</p>
 <p>ID: 554 CN MW: 26.01744 DISS: 0.0</p>		

**Figure 2.6:** Results of a similarity search for the disubstituted primary amine structure drawn in the structural query palette shown in Figure 2.5. The structure exactly matching the query in Figure 2.5 is indicated with a dark rectangle. Other structures similar to the queried structure are also displayed in similarity searches like this one for the user's consideration.

## Product Functional Group: \*C(\*)N,



EC	Prim ProFG	Sec ProFG	Prim SubFG	Sec SubFG	Co-facs/subs	Comments
3.10.1 (2)	 *C(*)N	 S(=O)(=O)(O)O	 *C(*)NS(=O)(=O)O			E.C. 3.10.1.1 and 3.10.1.2
3.5.1 (92)	 *C(*)N	$H_2C=O$ C=O	 *C(*)NC(=O)			E.C. 3.5.1.68
3.5.1 (92)	 *C(*)N	$O=C=O$ O=C=O	 *C(*)NC(=O)N			E.C. 3.5.1.87, also produces ammonia
3.5.1 (92)	 *C(*)N	 *C(=O)O	 *C(=O)NC(*)*			E.C. 3.5.1.72, 3.5.1.81, 3.5.1.82, and 3.5.1.83
3.5.2 (17)	 *C(=O)O	 *C(*)N	 *C(=O)NC(*)*			Most E.C. 3.5.2 Enzymes
3.5.3 (23)	 *C(*)N	 NC(=O)N	 *C(*)NC(=N)N			E.C. 3.5.3.3
3.5.99 (7)	 *C(*)N	 ac(=O)a	 *C(*)Nc(a)a			E.C. 3.5.99.4
4.3.2 (5)	 *OC(=O)C=*	 *C(*)N	 *OC(=O)C(*)NC(*)*			E.C. 4.3.2.1, 4.3.2.2, and 4.3.2.3
6.3.4 (18)	 *C(*)N		 *C(*)O	$NH_3$ N	ATP / ADP	E.C. 6.3.4.7
6.3.5 (10)	 *C(=O)O	 *C(*)N	 *C(=O)N	 *C(*)O		E.C. 6.3.5.8

**Figure 2.7:** Enzyme-catalyzed reactions returned by ReBiT which have as their product the disubstituted primary amine structure illustrated in Figure 2.5. For each reaction, graphical representations of the structures involved are shown along with any cofactors or co-substrates. Commentary for each reaction (typically a recommendation as to one or more specific 4-digit E.C. numbers that catalyze a given reaction) is also shown. Clicking on any 3-digit E.C. number hyperlink redirects the user to the ExPASy database for detailed enzyme information for the selected reaction. The numbers in the parentheses below each 3-digit E.C. hyperlink represent the total number of enzymes in that enzyme group.

## Conclusions

ReBiT provides a new means of searching through enzyme chemistry in order to identify potential transformations that can produce or consume specific functional groups. Such a tool should be useful in selecting targets for single-step bioconversions (in whole cell or free enzyme systems), especially for those enzymes that may need to be engineered to increase substrate specificity. A database such as ReBiT is an important tool to enable metabolic engineers and synthetic biologists to begin to design new pathways for biological synthesis of small molecules. A retro-biosynthetic approach that mimics the retrosynthesis performed by organic chemists necessarily requires access to knowledge of the transformational potential of enzymes. ReBiT serves to complement other computational tools such as BNICE (Hatzimanikatis et.al. 2005) that seek to move biosynthesis towards rational design—or *re-design*—of pathways towards specific targets.

## Chapter 3: Hydroxyvalerate Production in *P. putida*

This chapter contains material published in Martin and Prather, 2009.

Hydroxyacids represent an important class of compounds that see application in the production of polyesters, biodegradable plastics and antibiotics, and that serve as useful chiral synthetic building blocks for other fine chemicals and pharmaceuticals. In this chapter, an economical, high-titer method for the production of 4-hydroxyvalerate (4HV) and 3-hydroxyvalerate (3HV) from the inexpensive and renewable carbon source levulinic acid is described. These hydroxyvalerates were produced by periodically feeding levulinate to *Pseudomonas putida* KT2440 expressing a recombinant thioesterase II (*tesB*) gene from *Escherichia coli* K12. The titer of 4HV in shake flask culture reached  $13.9 \pm 1.2 \text{ g L}^{-1}$  from *P. putida tesB*<sup>+</sup> cultured at 32°C in LB medium periodically supplemented with glucose and levulinate. The highest 3HV titer obtained was  $5.3 \pm 0.1 \text{ g L}^{-1}$  in M9 minimal medium supplemented with glucose and levulinate.

### **Background**

Hydroxyacids are versatile, chiral compounds that contain both a carboxyl and a hydroxyl moiety, readily allowing for their modification into several useful derivatives (Lee et al., 2002; Chen and Wu, 2005). Specifically, hydroxyacids are used in the synthesis of antibiotics (Chiba and Nakai, 1985),  $\beta$ - and  $\gamma$ -aminoacids and peptides (Park, et al. 2001; Seebach et al., 2001), and as chiral synthetic building blocks (Lee et al., 2002). Hydroxyacids can also be used directly as nutritional supplements (Tasaki et al., 1999) and can be polymerized into biodegradable polyesters (polyhydroxyalkanoates, or PHAs) with interesting physical properties (Hazer and Steinbüchel, 2007).



Hydroxyacids are found in nature primarily polymerized as intracellular PHAs for energy storage in numerous organisms (Lenz and Marchessault, 2005). Of all the hydroxyacids, 3-hydroxybutyrate (3HB) is the most prolific, and several papers describe different means of producing monomeric 3HB (Lee et al., 1999; Gao et al., 2002; Liu et al., 2007; Tseng et al., 2009). Longer chain length hydroxyacids, mainly 3-hydroxyvalerate (3HV), 4-hydroxyvalerate (4HV), 3-hydroxyhexanate and other medium chain length 3-hydroxyacids have been produced as constituents of various intracellular PHA copolymers (Lee et al., 1999; Gorenflo et al., 2001; Park et al., 2002; Park and Lee, 2004). However, efficient production of these longer chain hydroxyacid monomers is complicated by issues such as low yields—typically less than 10% on a g hydroxyacid per g PHA basis for *in vivo* depolymerization from PHAs (Lee et al., 1999)—or the need for complicated chemical synthesis (Jaipuri et al., 2004) or purification (De Roo et al., 2002) procedures, most of which involve the use of large quantities of organic solvents.

Currently there exist three fundamental routes to the production of monomeric hydroxyacids: chemical synthesis, *in vivo* production of PHA polymers followed by depolymerization, and biological synthesis through non-PHA pathways. Chemical routes to hydroxyacid production are hampered by the high number of chemically reactive moieties in the hydroxyacid structure and the presence of a chiral center, and very few reports on their chemical synthesis are published (Jaipuri et al., 2004). There are, however, several reports on hydroxyacid production by depolymerizing PHAs through chemical or biological means (Lee et al., 2002 and references contained within). Chemical depolymerization of PHAs (Seebach et al., 1998) typically yields derivatives of

hydroxyacids such as alkyl esters (De Roo et al., 2002). The subsequent chemical steps required to remove the chemical modifications from the hydroxyacids make this option for depolymerization unattractive. *In vivo* depolymerization can result in hydroxyacid dimer production and other products of incomplete depolymerization (Lee et al., 1999). Furthermore, both chemical and biological depolymerization methods require the initial production of a microbial PHA, which potentially complicates the process of hydroxyacid production. This additional step in the process may also result in poor product yields. Typical yields for PHA production are 0.3-0.5 g PHA per g carbon source, while typical yields for the recovery of hydroxyacids from depolymerized PHAs range from 6.7% to 87.5%, depending on the composition of the PHA and the specific depolymerization method employed (Wang and Lee, 1997; Lee et al. 1999; Gorenflo et al., 2001; Ren et al., 2007).

Direct biological production of hydroxyacid monomers has been successfully demonstrated for 3HB, and titers of 3 g L<sup>-1</sup> and 12 g L<sup>-1</sup> on the shake flask and fed-batch scales have been reported (Gao et al., 2002; Tseng et al., 2009). In these reports, 3-hydroxybutyrate is made from acetyl-CoA through the use of acetyl-CoA acetyltransferase (*phbA*), 3-hydroxybutyryl-CoA dehydrogenase (*phbB*), phosphotransbutyrylase (*ptb*), and butyrate kinase (*buk*) (Liu and Steinbüchel, 2000a; Liu and Steinbüchel, 2000b; Gao et al., 2002). The last two of these enzymes were chosen to remove the CoA moiety from 3-hydroxybutyryl-CoA to yield free 3HB and were taken from *Clostridium acetobutylicum*, where they participate in the production of butyrate from butyryl-CoA (Liu and Steinbüchel, 2000b). Recently, thioesterase II (*tesB*) from

*Escherichia coli* K12 (Naggert et al., 1991) was successfully employed to directly hydrolyze the acyl-thioester of 3HB-CoA (Liu et al., 2007). While this pathway allows for the production of 3HB from glucose, it does not efficiently make higher chain length hydroxyacids and cannot produce hydroxyacids with the hydroxyl group at different positions.

Here the production of high titers of higher chain length hydroxyacids from inexpensive and renewable carbon sources was sought. It was reported that *Pseudomonas putida* accumulates PHA copolymers containing 4HV and 3HV when fed levulinic acid (Gorenflo et al., 2001). Levulinic acid is an inexpensive ketoacid that can be readily and renewably produced by treating wheat straw (Chang et al., 2007), corn starch (Cha and Hanna, 2002), cellulose (Hayes et al., 2006) and other agricultural feedstocks with dilute acid at modestly elevated temperatures and pressures. In this chapter, a bioprocess for the production of monomeric 4HV and 3HV from levulinic acid in *P. putida* shake flask cultures is described. Two strains of *P. putida* were tested: a commercially available strain (KT2440) and a PHA synthase knockout strain (GPp104; Huisman et al., 1991). Two enzyme systems for removing CoA acyl carriers from intracellular hydroxyacids were tested: the *ptb/buk* system and *tesB*. Once a suitable strain and enzyme system was found, the process was improved at the shake flask scale in minimal and rich media for the high-titer production of both 4HV and 3HV. As of the writing of this thesis, this study represents the first time that these higher chain length hydroxyacids have been produced from a renewable feedstock in shake flasks at the g L<sup>-1</sup> scale.

## Results

### **Initial Cellular System Optimization**

Recombinant *P. putida* KT2440 or GPp104 harboring the plasmid pRK415 (empty plasmid control), pRK415-ptb/buk, or pRK415-tesB was grown in 50 mL LB shake flask cultures at 30°C. The cultures were fed with 1.25 mL of a 2.0 M levulinate stock solution at  $t = 0, 6, 22, 29,$  and 49 hours. Levulinate was supplied to the cultures in smaller doses over time because it was found that extremely high initial concentrations of levulinate were inhibitory to cell growth. The resulting titers of 4HV and 3HV reached  $4.08 \text{ g L}^{-1}$  and  $1.11 \text{ g L}^{-1}$  respectively in the KT2440 (pRK415-tesB) culture after 72 hours (Table 3.1). The molar yields of 4HV and 3HV from the levulinate consumed in this culture were 15.3% and 4.2%, respectively, at 72 hours.

**Table 3.1:** Hydroxyvalerate accumulation in shake flask culture media after 72 hours by recombinant *P. putida* KT2440 or GPp104 grown at 30°C in LB and harboring different gene combinations on the pRK415 plasmid. Levulinate was fed to these cultures from a 2.0 M stock solution in 1.25 mL increments at  $t = 0, 6, 22, 29,$  and 49 hours. Productivity is calculated on a per gram dry cell mass basis.

	KT2440	GPp104	KT2440 <i>ptb + buk</i>	GPp104 <i>ptb + buk</i>	KT2440 <i>tesB</i>	GPp104 <i>tesB</i>
Cell Density ( $\text{g L}^{-1}$ ):	4.85	1.55	2.61	2.32	2.51	1.44
Total Levulinate Consumed ( $\text{g L}^{-1}$ ):	28.8	9.5	25.3	13.6	27.1	7.2
4HV Titer ( $\text{g L}^{-1}$ ):	1.50	0.42	0.53	0.53	4.08	1.58
3HV Titer ( $\text{g L}^{-1}$ ):	0.51	0.00	0.56	0.00	1.11	0.00
4HV Productivity ( $\text{g g}^{-1}$ ):	0.31	0.34	0.20	0.29	1.63	1.37
3HV Productivity ( $\text{g g}^{-1}$ ):	0.11	0.00	0.21	0.00	0.44	0.00
4HV Molar Yield:	5.3%	4.5%	2.1%	4.0%	15.3%	22.3%
3HV Molar Yield:	1.8%	0.0%	2.2%	0.0%	4.2%	0.0%

The expression of *tesB* leads to a substantial increase in both 4HV and 3HV concentration in the culture media while the *ptb* and *buk* genes do not improve either of

the titers (Table 3.1). Furthermore, *P. putida* GPp104 in all cases had lower 4HV titers than KT2440 and did not produce detectable amounts of 3HV. GPp104 4HV productivity was comparable with KT2440, suggesting that PHA accumulation is not competing significantly with 4HV production in KT2440; the differences in titer between KT2440 and GPp104 are due to differences in stationary phase cell density rather than the presence or absence of PHA synthase activity. These results indicated that KT2440 overexpressing *tesB* was a superior strain for the production of hydroxyvalerates, and consequently all future investigations were done with KT2440 *tesB*<sup>+</sup>.

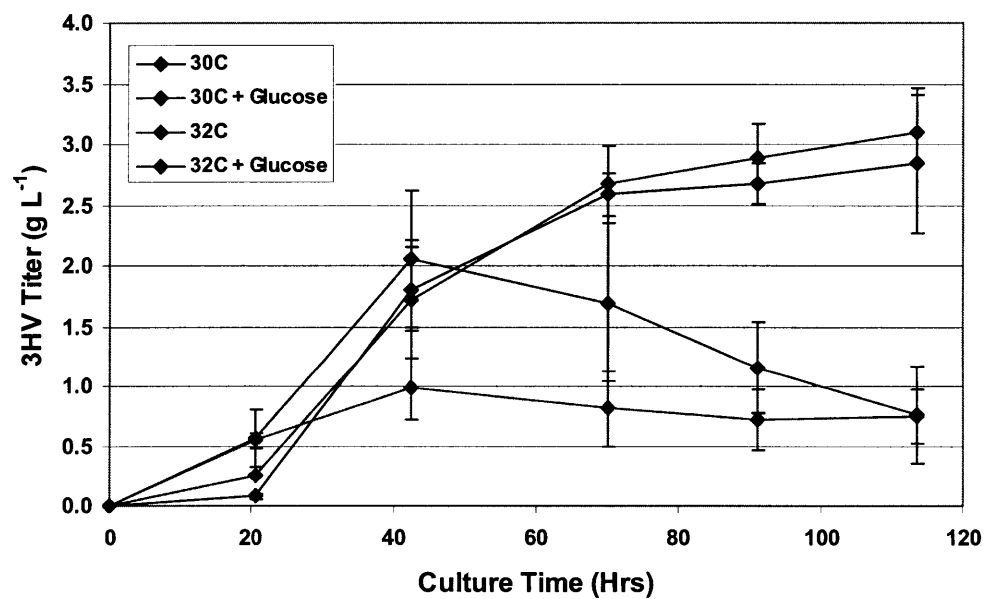
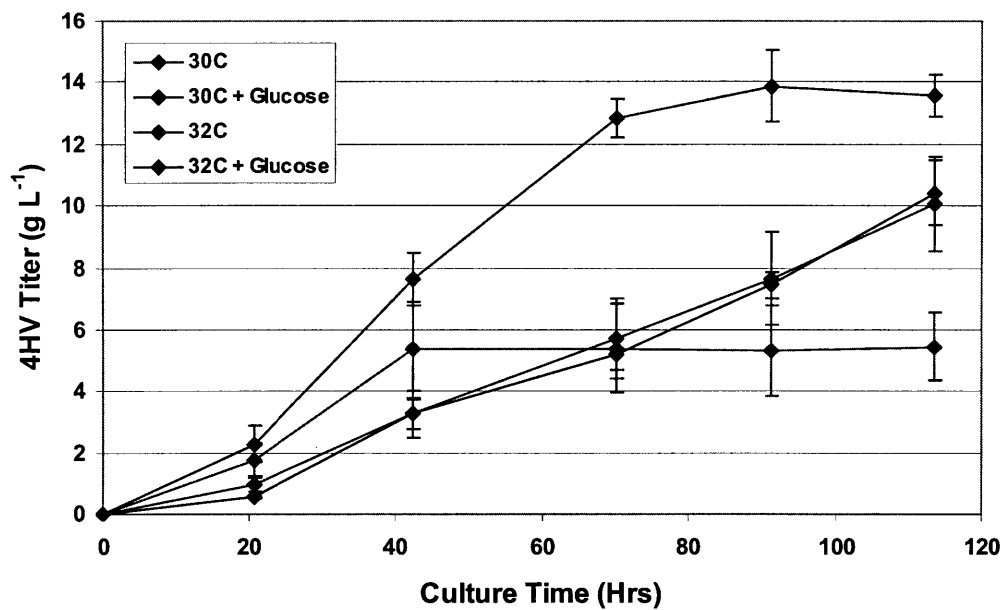
### **Production of Hydroxyvalerates in LB medium**

Recombinant *P. putida* KT2440 harboring the plasmid pRK415-*tesB* was grown in 50 mL LB shake flask cultures at 30°C or 32°C. Where indicated, cultures were initially supplemented with 0.4% glucose, followed by an additional 1 mL of 20% glucose solution at  $t = 21, 42$  and 70 hours. Glucose supplementation was done periodically because it was found that glucose, like levulinate, was inhibitory to cell growth at high concentrations. Levulinate was added to the cultures from a 2.0 M stock solution as follows: 2.5 mL at  $t = 0$  and 70 hours and 3.75 mL at  $t = 21$  and 42 hours.

The highest 4HV titer obtained ( $13.9 \pm 1.2 \text{ g L}^{-1}$ ) was from the culture incubated at 32°C and supplemented with glucose (Figure 3.1). The molar yield of 4HV from levulinate in this culture was  $25 \pm 1\%$ . The highest 3HV titer in this experiment was observed in the 30°C culture without glucose supplementation and was  $3.1 \pm 0.4 \text{ g L}^{-1}$ , corresponding to a yield of  $7 \pm 1\%$  from levulinate. Glucose supplementation did not have a significant

effect on hydroxyvalerate production in the 30°C cultures, but did have a marked effect in the 32°C cultures (Figure 3.1). The total hydroxyvalerate productivities per gram dry cell mass for the 30°C, 30°C with glucose supplementation, 32°C, and 32°C with glucose supplementation LB cultures were  $4.7 \pm 0.4$ ,  $5.0 \pm 0.4$ ,  $4.3 \pm 0.2$  and  $7.8 \pm 1.9$  g g<sup>-1</sup> respectively after 114 hours.

A probable explanation for the differences in hydroxyvalerate titers with respect to temperature is that the *tesB* enzyme is more active at 32°C. This enzyme is native to *E. coli*, an organism whose metabolism is optimum at 37°C, so *tesB* could be more active at temperatures closer to that of its native host. A more active *tesB* would cause *P. putida* to release more hydroxyvalerates, giving rise to the observed higher hydroxyvalerate titers. However, higher hydroxyvalerate titers in LB medium at 32°C were only observed when glucose was supplemented into the medium, suggesting that other factors such as metabolic burden may be affecting the titers.



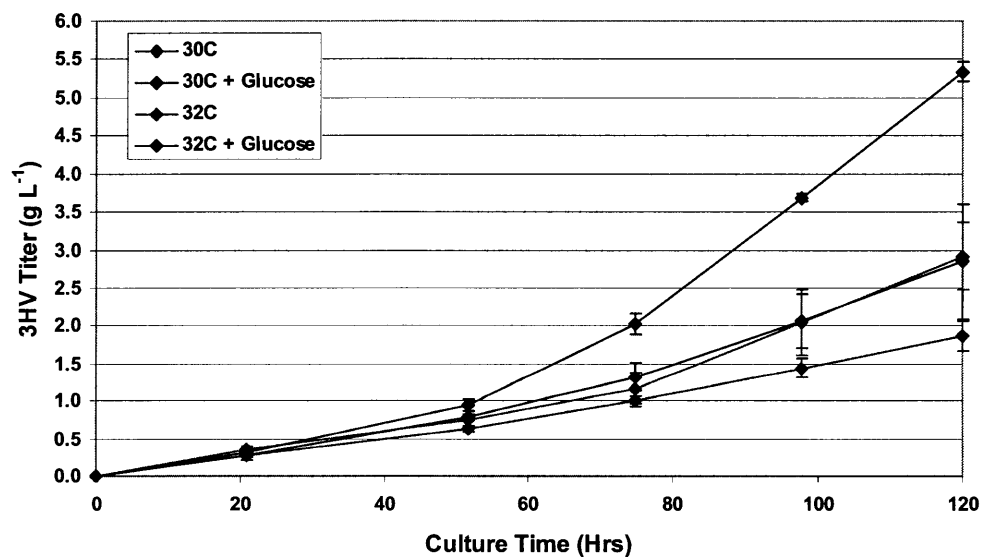
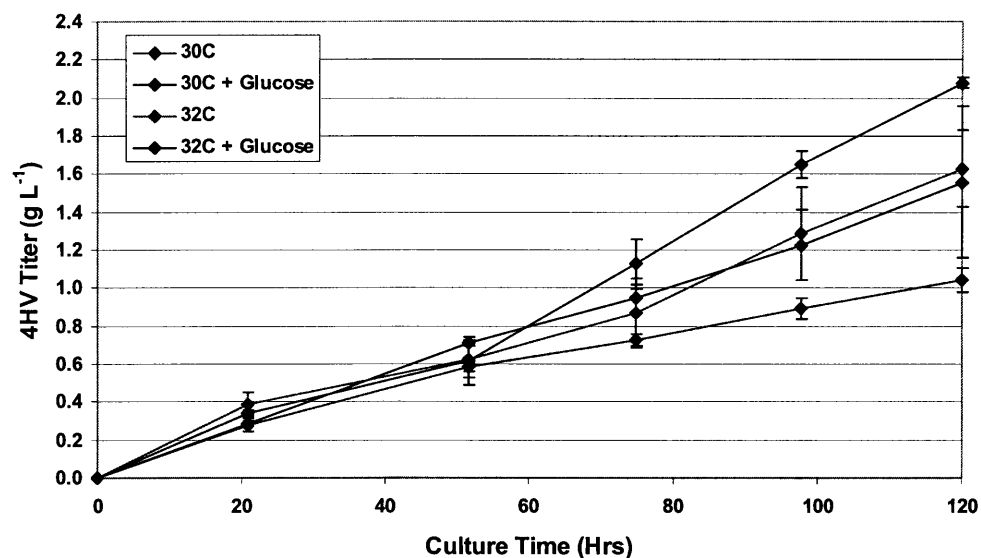
**Figure 3.1:** Time-course of 4-hydroxyvalerate (top) and 3-hydroxyvalerate (bottom) accumulation in shake flask culture media by recombinant *P. putida* KT2440 harboring pRK415-tesB grown at 30°C or 32°C in LB. Glucose was added to some cultures from a 20% stock solution where indicated, in 1 mL portions at t = 0, 21, 42, and 70 hours. Levulinate was fed to the cultures from a 2.0 M stock solution as follows: 2.5 mL at t = 0 and 70 hours and 3.75 mL at t = 21 and 42 hours. The values shown are the average of three independent cultures with the error bars representing the standard deviation.

### **Production of Hydroxyvalerates in M9 Minimal Medium**

Recombinant *P. putida* KT2440 harboring the plasmid pRK415-tesB was grown in M9 minimal medium in 50 mL shake flasks at 30°C or 32°C. The M9 medium was supplemented with 0.4% (w/v) glucose and a 1:1000 dilution of ATCC trace mineral supplement. Levulinate was added to these cultures from a 2.0 M stock solution in portions of 0.625, 0.625, 1.25, 1.25, and 1.25 mL at t = 0, 21, 52, 75, and 98 hours respectively. To some cultures where indicated, additional glucose was added from a 20% glucose stock solution in portions of 0.375 mL at t = 21, 52, 75, and 98 hours.

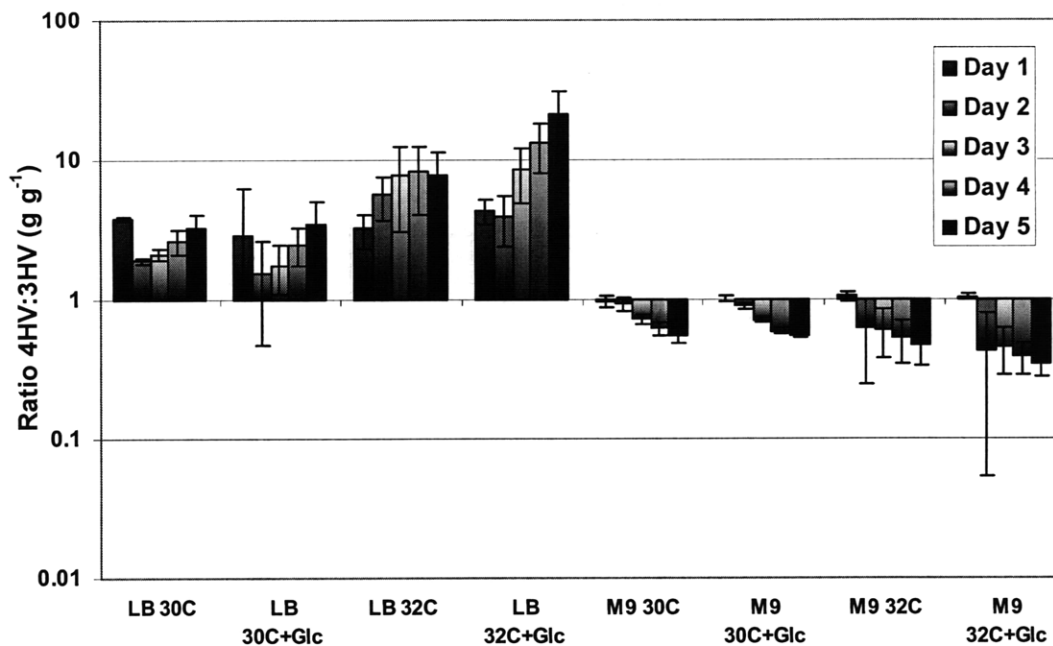
The highest titers of 4HV and 3HV in this experiment,  $2.1 \pm 0.0 \text{ g L}^{-1}$  ( $9 \pm 0\%$  molar yield of 4HV from levulinate) and  $5.3 \pm 0.1 \text{ g L}^{-1}$  ( $24 \pm 0\%$ ), were both observed in the 32°C culture with periodic glucose supplementation (Figure 3.2). The total hydroxyvalerate productivities per gram dry cell mass for the 30°C, 30°C with glucose supplementation, 32°C, and 32°C with glucose supplementation M9 cultures were  $2.1 \pm 0.1$ ,  $2.8 \pm 0.1$ ,  $3.3 \pm 0.6$  and  $5.2 \pm 0.1 \text{ g g}^{-1}$  respectively after 120 hours. In M9 medium, higher incubation temperature and periodic glucose supplementation had a cumulative positive effect on the titer of both hydroxyvalerates. The yield of 4HV from levulinate increased slightly by increasing the incubation temperature or adding additional glucose, while the 3HV yield increased substantially with these changes in culture condition.





**Figure 3.2:** Time-course of 4-hydroxyvalerate (top) and 3-hydroxyvalerate (bottom) accumulation in shake flask culture media by recombinant *P. putida* KT2440 harboring pRK415-tesB grown at 30°C or 32°C in M9 minimal media supplemented with ATCC trace mineral solution and 0.4% glucose. Additional glucose was added to some cultures from a 20% stock solution where indicated, in 0.375 mL portions at t = 21, 52, 75, and 98 hours. Levulinate was fed to the cultures from a 2.0 M stock solution as follows: 0.625 mL at t = 0 and 21 hours and 1.25 mL at t = 52, 75, and 98 hours. The values shown are the average of three independent cultures with the error bars representing the standard deviation.

Interestingly, while growth in M9 minimal medium does not yield a higher total hydroxyvalerate titer than growth in LB medium, the relative amounts of each hydroxyvalerate differ significantly between the two media. A comparison of Figures 1 and 2 indicates that more 4HV is produced in LB medium relative to 3HV, while in M9 medium more 3HV is made relative to 4HV. The ratio varied across almost two orders of magnitude, from as high as  $21.4 \pm 9.5$  in LB medium at 32°C with periodic glucose supplementation to as low as  $0.35 \pm 0.07$  in M9 medium at 32°C with periodic glucose supplementation (Figure 3.3). The choice of culture medium had the most dramatic effect on the 4HV-to-3HV ratio. In general, increasing the incubation temperature, supplementing the medium periodically with glucose, and increasing the culture time all caused the ratio to increase in LB medium and decrease in M9 medium. The underlying causes of the shifts in the 4HV-to-3HV ratio should be rooted in the expression and kinetics of the enzymes in *P. putida* that create and consume the hydroxyvalerates, and thus without further study these causes are subject only to speculation. Nonetheless, the effects of various culture conditions on the 4HV-to-3HV ratio provide a convenient means of controlling the relative amounts of each hydroxyvalerate produced.



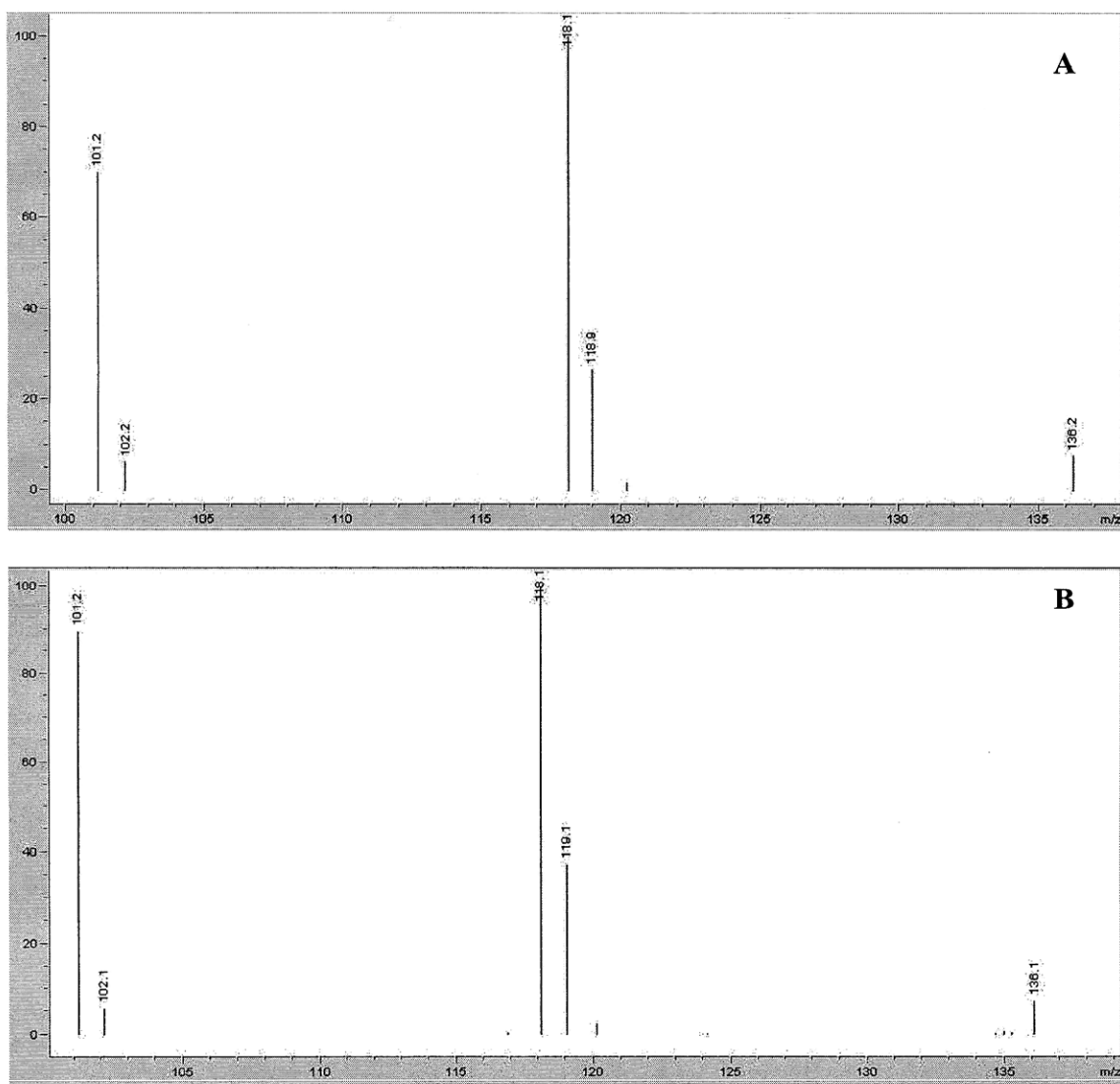
**Figure 3.3:** Ratios of 4HV-to-3HV titers produced by *P. putida tesB*<sup>+</sup> cultured under the different culture conditions from Figures 1 and 2. The conditions tested were LB or M9 media, 30°C or 32°C incubation, and periodic glucose feeding (glc) or no periodic glucose feeding.

### Mass Spectrometry Confirmation of Hydroxyvalerate Production

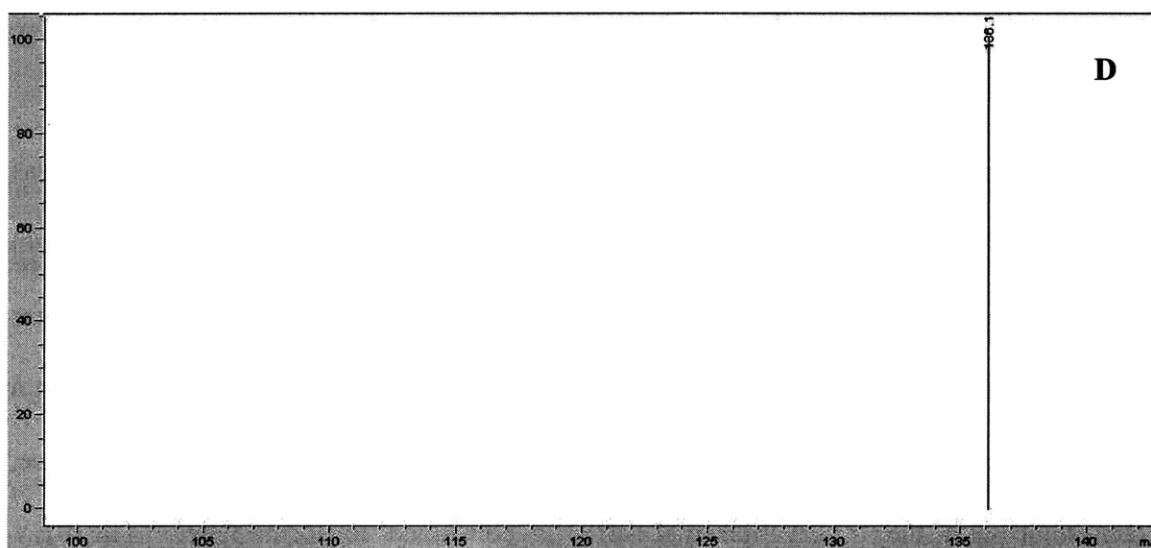
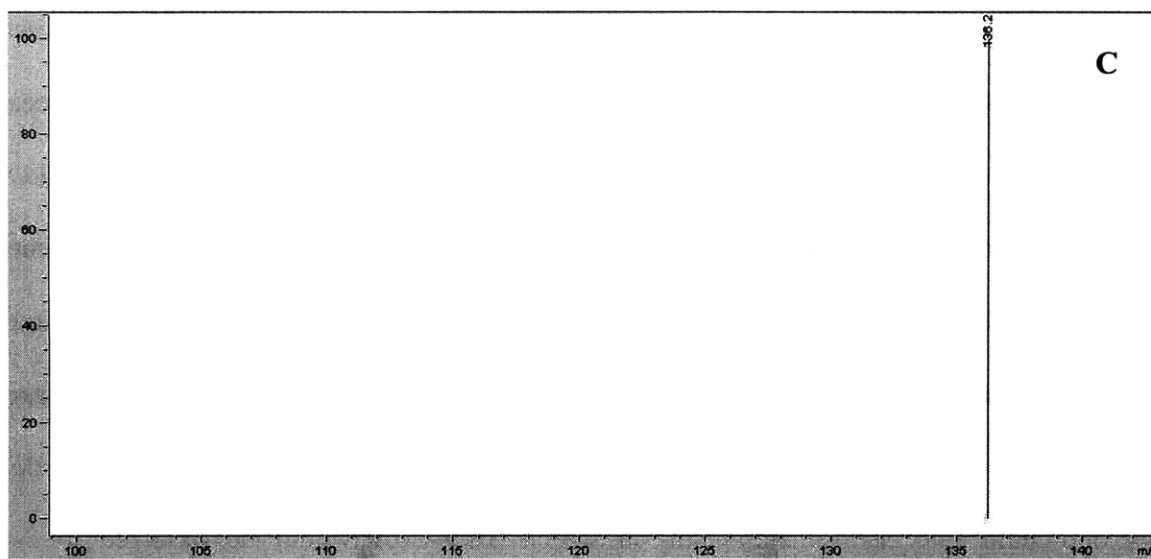
To confirm that hydroxyvalerates were being produced from levulinate, supernatants from cultures of *P. putida tesB*<sup>+</sup> supplemented with levulinate were analyzed by LC/MS and the mass spectra of each experimental hydroxyvalerate peak were compared to 4HV and 3HV standards. The mass spectra of the 4HV and 3HV peaks found in the *P. putida* cultures matches the spectra of the 4HV and 3HV standards extremely well (Figure 3.4).

In the 4HV spectra (Figures 3.4a and 3.4b), the molecular ion ( $M^+$ ) peak of 118.1 (118.134 theoretical mass) is observed as well as an  $M + H^+$  peak of 119.1. The small

peak of 136.2 is likely  $M + NH_4^+$ , given that the mobile phase for this LC/MS analysis used ammonium formate as a buffer (by the same logic, the 118.1 peak likely also represents the  $[M^+ - H_2O] + NH_4$  ion). The 101.2 peak is consistent with  $[M - H_2O] + H^+$  and is likely the result of dehydration of 4HV into its lactone form upon sample vaporization. The 3HV spectra (118.134 theoretical mass) consist of only a single peak at 136.2, representing  $M + NH_4^+$  (Figures 3.4c and 3.4d). As expected, there is no dehydration peak in the 3HV spectra, as 3HV cannot form a stable lactone.



Figures 3.4a-b: See figure caption on page 73.

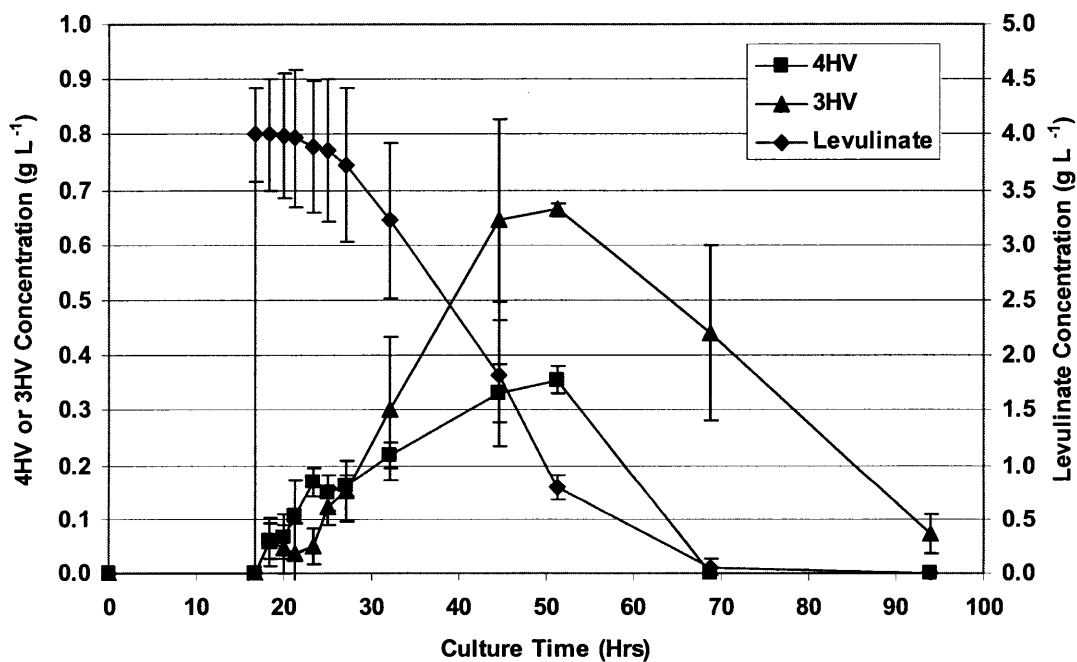
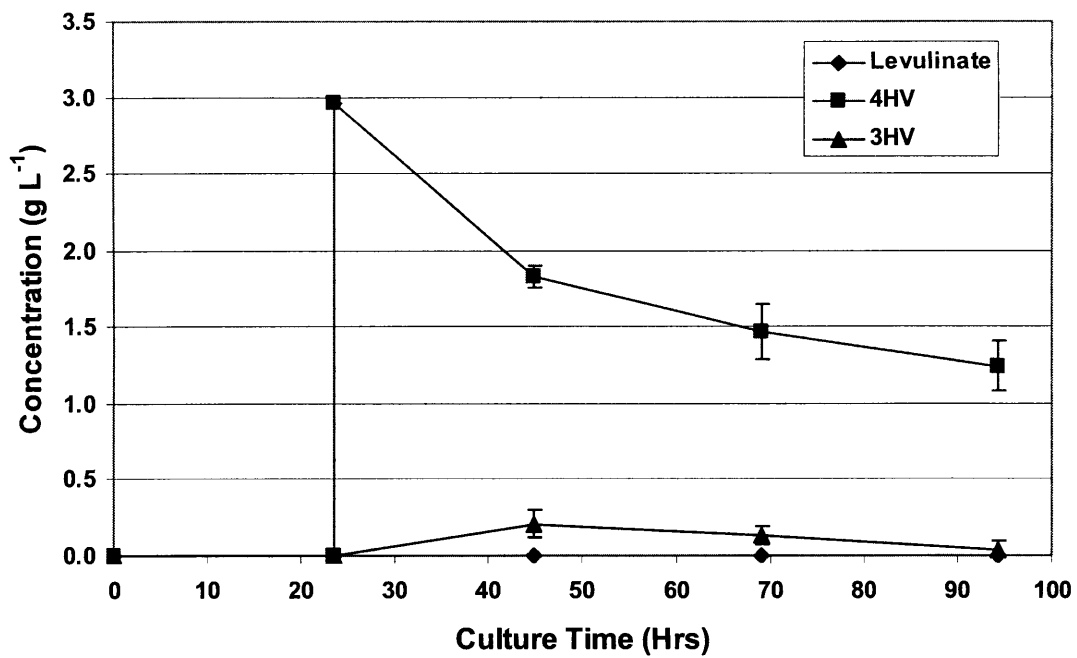


**Figure 3.4:** Positive mode electrospray mass spectra of 4HV produced by *P. putida* from levulinate (A), a 4HV standard (B), 3HV produced by *P. putida* from levulinate (C), and a 3HV standard (D). The theoretical mass of both 4HV and 3HV ( $M^+$ ) is 118.134. The additional peaks in the spectra are believed to be  $M + H^+$  (119.142),  $M + NH_4^+$  (136.173), and  $[M + H^+] - H_2O$  (101.126). The ammonium ion that gives rise to the  $M + NH_4^+$  peak at 136.1 originates from the ammonium formate buffer used in the mobile phase. The  $[M + H^+] - H_2O$  peak at 101.2 peak is likely due to lactonization induced by sample vaporization in the mass spectrometer.

### **Investigation of Levulinate Metabolism in *P. putida***

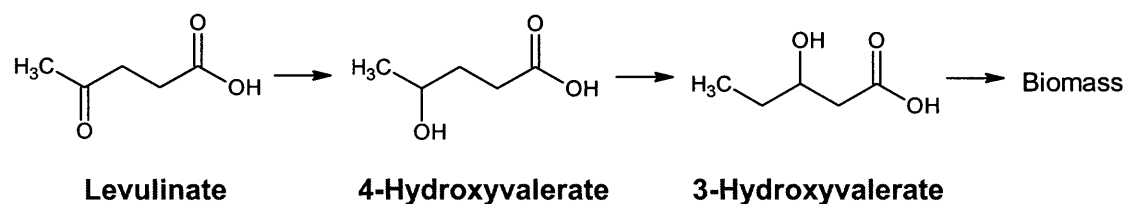
To gain insight into how hydroxyvalerates are made in *P. putida*, *P. putida tesB<sup>+</sup>* was grown in M9 minimal medium at 30°C with 0.4% (w/v) glucose used as a carbon source. This culture was then fed a bolus of either 3.0 g L<sup>-1</sup> 4HV after 24 hours or 4.0 g L<sup>-1</sup> levulinate after 17 hours. The 4HV bolus caused the production of 3HV, establishing that 3HV is produced from 4HV in *P. putida* (Figure 3.5, top). No levulinate was detected throughout the 4HV bolus experiment. The concentration of 3HV began to fall after a single day, suggesting that 3HV is metabolized by *P. putida*. One likely avenue for 3HV is odd-chain fatty acid metabolism, in which 3HV-CoA would be oxidized to 3-ketovaleryl-CoA and then split into acetyl-CoA and propionyl-CoA.

In the levulinate bolus experiment, 4HV production in the hours following the bolus was significantly more rapid than 3HV production, suggesting that 4HV is produced from levulinate before 3HV (Figure 3.5, bottom). After the concentration of 4HV reaches about 0.15 g L<sup>-1</sup>, 3HV levels begin to rapidly rise, eventually overtaking 4HV. About a day after the levulinate bolus, 4HV increases more slowly while 3HV levels rapidly rise, consistent with the observation that 4HV is consumed to form 3HV. The 4HV concentration rapidly declines once levulinate is depleted from the medium, providing further evidence that levulinate is reduced first to 4HV in *P. putida*. Finally, the decline of 3HV levels between t = 69 and 94 hours after 4HV and levulinate are both depleted further establishes that *P. putida* metabolizes 3HV.



**Figure 3.5:** Concentration profiles of levulinate, 4HV, and 3HV in a M9 shake flask culture of *P. putida* expressing *tesB* after a 3.0 g L<sup>-1</sup> bolus of 4HV at t = 24 hours (top) or a 4.0 g L<sup>-1</sup> bolus of levulinate at t = 16.7 hours (bottom). The values shown are the average of three independent cultures with the error bars representing the standard deviation.

The observations from the 4HV and levulinate bolus experiments support the levulinate metabolism pathway shown in Figure 3.6. First, levulinate is reduced by a reductase activity native to *P. putida* to 4HV. This 4HV is then isomerized into 3HV by one or more enzymes also native to *P. putida*. Finally, 3HV itself is metabolized by *P. putida*, likely by way of odd-chain fatty acid metabolism. The differences in the relative ratios of the individual hydroxyvalerates produced are likely a consequence of the effect of various culture conditions on the rates of the reaction steps in the levulinate metabolism pathway, particularly the isomerization of 4HV to 3HV.



**Figure 3.6:** Proposed model of levulinate metabolism in *P. putida* based on the experimental results shown in Figure 3.5.

## **Discussion**

In this chapter, it was demonstrated that high concentrations of 4HV and 3HV can be made from levulinate in recombinant *P. putida* KT2440 by overexpressing the *tesB* gene. Through careful modulation of the culture and feeding conditions, maximum titers of 14 g L<sup>-1</sup> 4HV and 5 g L<sup>-1</sup> 3HV at the shake flask scale were achieved.

Hydroxyvalerate production in both rich and minimal media was investigated and it was found that the choice of culture medium substantially affects the relative amounts of each



hydroxyvalerate produced. LB medium was found to be superior for 4HV production while M9 medium was better for 3HV production. The total hydroxyvalerate titer however is higher in LB medium relative to M9 medium, an expected result given the relative richness of the two media. Additionally the total hydroxyvalerate productivity on a per gram dry cell weight basis was consistently higher in LB (4.3-7.8 g g<sup>-1</sup>) than in M9 (2.1-5.2 g g<sup>-1</sup>) and was generally dependent on temperature and glucose supplementation. Despite these differences, both hydroxyvalerates were produced in g L<sup>-1</sup> quantities at the shake flask scale in both kinds of media.

While the hydroxyvalerate titers in this report are high at the shake flask scale, they are not necessarily optimal. Both the total hydroxyvalerate titer and ratio of the individual hydroxyvalerate produced were found to be strongly dependent on medium composition and temperature. The temperature effect may be due to increased *tesB* activity, since increased temperatures generally increased the concentrations of both hydroxyvalerates. Medium composition heavily influences the conversion of 4HV to 3HV, since the ratio of hydroxyvalerates produced by *P. putida* changes dramatically between rich and minimal media. Specifically, the gene or genes responsible for this conversion appear to be upregulated in defined media.

Because hydroxyvalerate levels were heavily influenced by culture conditions, analyzing a wider variety of culture conditions would almost certainly lead to even higher product titers. Furthermore, the ratio of hydroxyvalerates could be further controlled by taking advantage of the levulinate pathway structure shown in Figure 3.6. For instance, 3HV

could be made by *P. putida* without any 4HV by first allowing the cells to accumulate both hydroxyvalerates by feeding the culture levulinate and then subsequently depriving the cells of levulinate, allowing them to convert the 4HV left in the medium to 3HV. Care would need to be taken to harvest the 3HV from the culture medium immediately after 4HV is depleted as the cells do consume 3HV. The 69 hour time point from the levulinate bolus experiment in Figure 3.5 illustrates this strategy. At this time point  $0.44 \pm 0.16 \text{ g L}^{-1}$  of 3HV was present in the culture media as a result of the  $4.0 \text{ g L}^{-1}$  levulinate bolus while no 4HV was detected at this time point.

If the specific enzymes responsible for the steps in the levulinate metabolism pathway are found, the pathway could likely be engineered to produce even higher titers of 4HV or 3HV from levulinate. Additionally, discovery of the levulinate pathway enzymes would allow the reconstruction of the pathway in other organisms such as *Escherichia coli* that are easier to engineer and could allow pathway expression at higher temperatures where *tesB* may be more active. A likely candidate for both the reduction of levulinate and the isomerization of 4HV to 3HV is the enzyme complex encoded by the *P. putida fadBA* operon (Olivera et al., 2001). This multifunctional complex possesses hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA isomerase activities that could both reduce levulinate to 4HV and isomerize 4HV to 3HV. The high titers of hydroxyvalerates produced by *P. putida* from levulinate suggest that the enzymes responsible for levulinate metabolism in *P. putida* are highly active, making them attractive choices for use in biosynthetic pathways involving similar reaction steps.

Given that total hydroxyvalerate titers in excess of  $15 \text{ g L}^{-1}$  could be made in 50 mL shake flasks, one future study of immediate interest is hydroxyvalerate production in larger stirred tank reactors. Large-scale production in a better aerated reaction vessel should afford even higher titers of 4HV and 3HV with improved economic margins. Another future study of interest is the production of other hydroxyacids with *tesB* in *P. putida*. In particular, it is known that supplying *P. putida* with oleic acid results in the accumulation of a PHA containing 3-hydroxyalkanoic acids from six to fourteen carbon atoms in length (De Roo et al., 2002). Given that *tesB* is known to accept longer-chain fatty acid CoA thioesters such as decanoyl-CoA as substrates (Naggert et al., 1991), it is likely that *tesB* will be able to liberate these  $\text{C}_6\text{-C}_{14}$  3-hydroxyacids from their intracellular CoA carriers prior to their incorporation as a PHA. This would result in the accumulation of these novel, longer-chain hydroxyacids in the culture media.

In conclusion, a method was developed for producing high titers of the high-value hydroxyacids 4HV and 3HV from the renewable and inexpensive substrate levulinate in *P. putida* KT2440 at the shake flask scale by overexpressing the *tesB* gene. Hydroxyvalerate titers of several grams per liter were achieved in both rich and minimal media. Modulation of culture conditions such as temperature, feed profile, and media yielded increased hydroxyvalerate titers as well as control over the relative amounts of 4HV and 3HV produced. Chapter 4 covers the further transformation of 4HV into 4-valerolactone (4VL) using this cellular system.

## **Materials and Methods**

### **Strains and Plasmids**

*Pseudomonas putida* KT2440 (ATCC 47054; American Type Culture Collection, Manassas, VA, USA) and GPp104 (Huisman et al., 1991) were used to produce hydroxyvalerate monomers from levulinic acid. GPp104 is a polyhydroxyalkanoate (PHA) synthase deficient mutant of *P. putida* KT2442. *Escherichia coli* thioesterase II (*tesB*) was amplified from the *E. coli* K12 MG1655 (ATCC 47076) genome by PCR. The primers used were purchased from Sigma-Genosys (St. Louis, MO, USA) and were as follows (restriction sites used for cloning are underlined): EcoRI-*tesB*-FP (5'-GAATTCTACTGGAGAGTTATATGAGTCAGG-3') and SalI-*tesB*-RP (5'-GTCGACTTAATTGTGATTACGCATC-3'). HotStar HiFidelity DNA polymerase was purchased from Qiagen (Valencia, CA, USA) and used according to the manufacturer's instructions. The *tesB* gene was first cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA) to produce the plasmid pGEM-*tesB*. pGEM-*tesB* was then digested with *EcoRI* and *SalI* and cloned into a similarly digested broad-host-range expression vector pRK415 (Keen et al., 1988) to produce the plasmid pRK415-*tesB*. Molecular biology manipulations were performed using standard cloning protocols (Sambrook and Russell, 2001).

The phosphotransbutyrylase (*ptb*) and butyrate kinase (*buk*) genes were amplified as an operon by PCR from the genomic DNA of *Clostridium acetobutylicum* (ATCC 824) using the primers EcoRI-*ptb/buk*-FP (5'-GAATTCACCAGTGATTAAGAGTTTTAATG-3') and PstI-*ptb/buk*-RP (5'-

GTCGACGGTACTGGTTATATTATATTATTTATG-3'). These two genes were cloned into the pGEM-T Easy vector to yield the plasmid pGEM-ptb/buk. pGEM-ptb/buk was then digested with *EcoRI* and *PstI* and cloned into similarly digested pRK415 to produce the plasmid pRK415-ptb/buk.

### **Chemicals and Growth Media**

Dehydrated LB broth (Miller formulation) was purchased from BD Biosciences (San Jose, CA, USA) and D-glucose was purchased from Mallinckrodt Chemicals (Phillipsburg, NJ, USA). M9 minimal medium was prepared as described elsewhere (Sambrook and Russell, 2001). LB and M9 media were autoclaved prior to use, while D-glucose was prepared as a 20% (w/v) stock solution and sterile filtered. Levulinic acid was purchased from Acros Organics (Morris Plains, NJ, USA) and was neutralized to a pH of 7.0 with 10N NaOH and sterile-filtered prior to use. ATCC Trace Mineral Supplement (ATCC MD-TMS) was added to M9 cultures where indicated. 3-Hydroxyvalerate was purchased from Epsilon Chimie (Brest, FRANCE) and was used as an HPLC standard. 4-Hydroxyvalerate was made by saponification of  $\gamma$ -valerolactone purchased from Alfa Aesar (Ward Hill, MA, USA) and was used as an HPLC standard.

### **Culturing of *P. putida***

50 mL of M9 or LB media was added to a 250 mL shake flask. The medium was supplemented with 10  $\mu\text{g mL}^{-1}$  tetracycline for plasmid maintenance, 1.0 mM IPTG for induction of gene expression, and various levulinic acid concentrations. M9 cultures were given a 1000x dilution of ATCC Trace Mineral Supplement. The inocula for all

cultures was prepared from *P. putida* KT2440 or GPp104 grown overnight from frozen stock in LB at 30°C, centrifuged at 2,000 x *g* and 4°C for 5 min, and washed and resuspended in 0.9% (w/v) sodium chloride. All experimental cultures were inoculated with this resuspension to an initial optical density at 600nm (OD<sub>600</sub>) of 0.05 and were incubated at 30 or 32°C with shaking at 250 rpm. Samples were periodically withdrawn from these cultures for OD<sub>600</sub> and HPLC analysis.

### **Analytics**

Culture cell density was monitored by measuring optical density at 600 nm (OD<sub>600</sub>) on a Beckman Coulter DU800 UV/Vis spectrophotometer. Optical density readings of cell concentration were correlated to dry cell weight (DCW) per unit volume by measuring the OD<sub>600</sub> of several *P. putida* KT2440 or GPp104 cultures, filtering a known culture volume through a pre-weighed Whatman 0.45µm cellulose acetate filter, and drying the retained cells for several days in a 37°C oven. A calibration curve for both KT2440 and GPp104 was constructed to convert OD<sub>600</sub> values to g DCW-L<sup>-1</sup> and conversion factors of 0.42 g L<sup>-1</sup> OD<sub>600</sub><sup>-1</sup> and 0.53 g L<sup>-1</sup> OD<sub>600</sub><sup>-1</sup> were found for KT2440 and GPp104, respectively.

High-performance liquid chromatography (HPLC) samples were prepared by centrifuging 1 mL of culture at 14,000 x *g* for 10 min at room temperature and withdrawing the supernatant for analysis. HPLC samples were analyzed on an Agilent 1100 Series instrument equipped with an Agilent ZORBAX SB-Aq reverse phase column (4.6 x 150mm, 3.5µm). The column temperature was maintained at 65°C. Levulinic acid

and hydroxyvalerate concentrations were measured with a refractive index detector. The mobile phase was an aqueous solution of 25 mM ammonium formate, pH of 2.0. The flowrate through the column was 0.5 mL min<sup>-1</sup>. Levulinic acid, 4-hydroxyvalerate, and 3-hydroxyvalerate were used as HPLC standards and had respective retention times of 7.4, 7.0, and 8.0 minutes.

Mass spectrometry (MS) analysis was performed using an Agilent 6120 Quadrupole LC/MS unit operated in positive ion electrospray mode. The unit was operated downstream of the liquid chromatography column described above under the same chromatography conditions. Samples were vaporized at 150°C and nitrogen at 300°C, 60 psig, and 5.0 L min<sup>-1</sup> was used as the carrier gas for MS analysis.

## Chapter 4: Valerolactone Production in *P. putida*

This work is in press in Appl. Environ. Microbiol. as of the writing of this thesis.

Enzymes are powerful biocatalysts capable of performing specific chemical transformations under mild conditions. Yet as catalysts they remain subject to the laws of thermodynamics, namely that they cannot catalyze chemical reactions beyond equilibrium. In this chapter, the phenomenon and application of using extracytosolic enzymes and medium conditions such as pH to catalyze metabolic pathways beyond their intracellular catalytic limitations is described. This methodology, termed “integrated bioprocessing” because it integrates intracellular and extracytosolic catalysis, was applied to a lactonization reaction in *Pseudomonas putida* for the economical and high-titer biosynthesis of 4-valerolactone from the inexpensive and renewable source levulinic acid. Specifically, mutant paraoxonase I (PON1) was expressed in *P. putida*, shown to export from the cytosol in *Escherichia coli* and *P. putida* using an N-terminal sequence, and demonstrated to catalyze the extracytosolic and pH-dependent lactonization of 4-hydroxyvalerate to 4-valerolactone. With this production system, the titer of 4-valerolactone was enhanced substantially in reduced-pH media using extracytosolically-expressed lactonase vs. an intracellular lactonase: from  $<0.2 \text{ g L}^{-1}$  to  $2.1 \pm 0.4 \text{ g L}^{-1}$  at the shake flask scale. Based on these results, production of 4-hydroxyvalerate and 4-valerolactone was scaled up to a 2 L bioreactor and titers of  $27.1 \text{ g L}^{-1}$  and  $8.2 \text{ g L}^{-1}$  for the two respective compounds were achieved. These results illustrate the utility of integrated bioprocessing as a strategy for enabling lactone production from novel metabolic pathways and enhancing product titers.



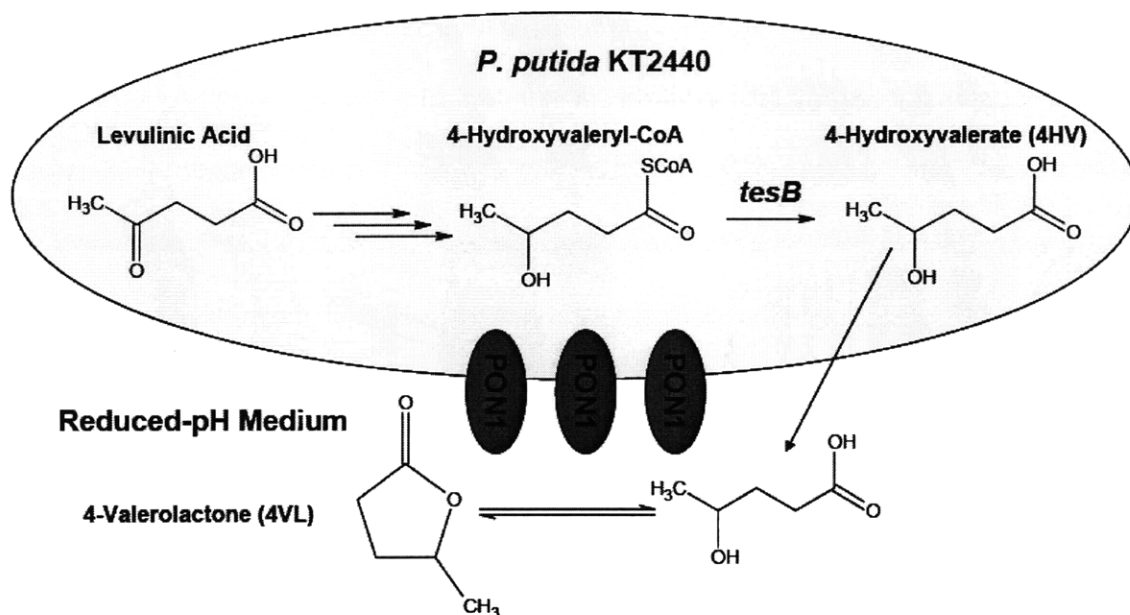
## **Background**

In nature, metabolism is not restricted to the cytosol. Metabolic activity can occur in the periplasm, on cell surfaces, or even extracellularly in the environment. When extracytosolic enzyme expression occurs, it is often to convey an advantage to the cell that is not possible with the enzyme in the cytosol. For instance, *E. coli* alkaline phosphatase is expressed in the periplasm to detoxify compounds before they can enter the cell and to allow the enzyme better access to the extracellular environment for scavenging phosphate (Derman and Beckwith, 1991). *Penicillium decumbens* secretes cellulases presumably to break down extracellular substrates into a form amenable to uptake by the cell for further metabolism (Sun et al., 2008). Each of these enzymes exists as a component of larger metabolic pathways (phosphate and cellulosic material metabolism) and each has evolved for non-cytosolic expression to facilitate the physiological goals of their respective pathways.

In microbial biotechnology, the objective is typically not physiological but commercial in nature: to increase the titer of a small-molecule metabolic product. Though the goal has changed, the lessons we can learn from natural systems remain. The cytoplasm is not always the best choice for enzyme expression because cytoplasmic conditions are not necessarily optimal for enzyme productivity. One such system is the intramolecular lactonization of hydroxyacids such as 4-hydroxyvalerate (4HV) to lactones such as 4-valerolactone (4VL). This reaction, catalyzed in this work by the G3C9 variant of paraoxonase I (PON1; Aharoni et al., 2004), is known to be highly pH-dependent. Because lactonization is acid-catalyzed and because hydroxyacids and lactones exist in

pH-dependent equilibrium with each other, control of the pH at which lactonization occurs is critical to achieving high titers of lactones (Teiber et al., 2003). The cytoplasmic pH, typically about 7.5, is too high to achieve good titers of lactones at equilibrium (Wilks and Slonczewski, 2007). This limitation in lactone titer is thermodynamic in nature, meaning that overexpressing the lactonase or other traditional metabolic engineering techniques would be ineffective at improving lactone production. However, having the lactonase perform catalysis outside the cytosol, where the pH can be lowered, would alter the equilibrium in favor of lactone production, thereby increasing product titer.

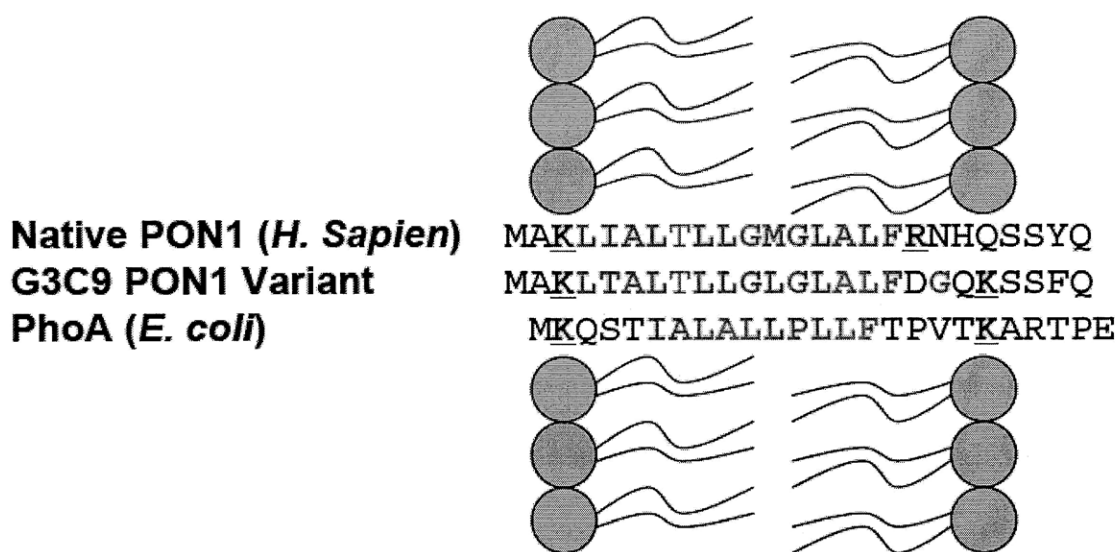
To implement such a system, we used extracytosolically-expressed PON1 to catalyze the lactonization reaction in reduced pH-media (pH ~ 6). In Chapter 3, it was found that *Pseudomonas putida* was capable of producing high titers of 4HV from the renewable carbon source levulinic acid (Martin and Prather, 2009). This process is achieved through coenzyme-A carriers and the secretion of 4HV into the extracellular medium was enhanced by the expression of an intracellular thioesterase. Thus, *P. putida* can serve as an intracellular source of 4HV. This 4HV can then be lactonized by extracytosolically-expressed PON1 in reduced-pH media to yield 4VL (Figure 4.1). This general strategy is termed “integrated bioprocessing” for its integration of cytosolic and extracytosolic biocatalysis to enhance production.



**Figure 4.1:** Integrated bioprocessing system for the production of 4VL from levulinate. Production of 4HV from levulinate occurs intracellularly, while the lactonization reaction takes place extracytosolically in reduced-pH media. The enzyme(s) responsible reaction step(s) from levulinate to 4-hydroxyvaleryl-CoA occur in *P. putida* and are currently not known.

There is conflicting evidence as to whether the G3C9 variant of PON1 expressed in microbes would be cytosolic (Aharoni et al., 2004) or whether G3C9 PON1 would be associated with the membrane as native PON1 is (Deakin et al., 2002). In particular it was noted that G3C9 PON1 had a significantly altered N-terminal primary sequence relative to native PON1 (Figure 4.2), and that the altered sequence looked similar to that of alkaline phosphatase (PhoA), a protein known to be membrane-associated in *E. coli*. Thus in this work, it was first demonstrate that the G3C9 variant of PON1 (referred to as simply PON1 throughout the rest of this report) is expressed outside of the cytosol in both *E. coli* and *P. putida* and consequently that it responds strongly to changes in the medium pH. Then, intracellular 4HV production was combined with extracytosolic

lactonization to enable the high-titer biochemical production of 4VL. This work represents the first report of 4VL synthesis in a biological system.



**Figure 4.2:** Comparison of the first 25 N-terminal residues of native human PON1, the G3C9 variant of PON1 created by Aharoni and coworkers (2004), and *E. coli* PhoA (a protein known to export into the periplasm). Aliphatic residues (which bury within the membrane) are shown in green while cationic residues (presumed to bind to the negatively-charged heads of phospholipids) are underlined in red. In this work it is hypothesized that the G3C9 variant of PON1 possesses an N-terminal signal sequence similar enough to bacterial signal sequences to allow PON1 export from the cytosol in both *E. coli* and *P. putida*.

4VL has been reported to be an ideal compound for use as a fuel and in the production of carbon-based chemicals (Horváth et al., 2008). It has also seen extensive use as a component of block-copolymers for drug delivery (Atwood et al., 2007; Chang and Chu, 2008), as a precursor for acrylic compounds (Manzer, 2004), and can be used as a precursor for the production of “bio-nylon” polymers (Lange et al., 2007). Current synthetic methods for 4VL also utilize levulinic acid as the starting material but require supercritical solvents (Manzer and Hutchenson, 2005) or carbon dioxide (Bourne et al.,

2007) with hydrogen gas under harsh conditions (~10-40 MPa and ~200°C) and ruthenium-based catalysts. In contrast, our biological method of production is done under mild conditions without the need for harsh solvents, hydrogen, rare metal catalysts, or supercritical fluids. This integrated bioprocessing system to produce 4VL from levulinate (Figure 4.1) was compared to an entirely intracellular 4VL production pathway using cytosolically-expressed PON1 to demonstrate the effectiveness of integrated bioprocessing to improve product titers at different pH values. The system was scaled up to a 2.0 L bioreactor to further increase product titers as well as take advantage of automated pH control.

## **Results**

### **Confirmation of Extracytosolic PON1 Expression**

To demonstrate the integrated bioprocessing system for the production of 4VL, it was first necessary to establish the localization of the PON1 lactonase. It was suspected that the G3C9 PON1 variant (Aharoni et al., 2004) might export from the cytosol using an N-terminal sequence (Figure 4.2). In particular, key similarities were identified between the first 23 amino acids of G3C9 PON1 and *E. coli* alkaline phosphatase, a protein known to export from the cytoplasm. First, the spacing between the two lysines in the N-termini of these two proteins was identical. These lysines are cationic at physiological pH and can associate with negatively-charged phospholipid heads in a cell membrane. Second, both N-termini have several hydrophobic residues between these two lysines, which would help anchor that part of the protein inside a cell membrane. Comparing the N-termini of native human PON1 with G3C9 PON1 shows that the spacing between the two cationic

residues differs substantially between the two proteins. It was suspected that this difference would allow G3C9 to export from the cytosol. Considering that G3C9 PON1 was evolved from human PON1 with the goal of functional expression in *E. coli* (Aharoni et al., 2004), this key difference in the N-terminal signal sequences of the two lactonases may be what allows G3C9 to be functionally expressed in *E. coli*.

To test G3C9 PON1 (hereinto referred to as PON1) for any ability to export from the cytosol, fusions of PON1 with *E. coli* alkaline phosphatase (PhoA), an enzyme only active in the periplasm (Derman and Beckwith, 1991), were constructed. PON1 fusions that successfully export from the cytoplasm enable PhoA activity. To construct the protein fusions, PON1 and a truncated version of PON1 with residues 2-23 removed, tPON1, were fused to a truncated version of phoA (tPhoA) with residues 1-23 removed. The fusions were constructed as N-PON1-tPhoA-C and N-tPON1-tPhoA-C, with an *Xba*I restriction site used as the linker between PON1 or tPON1 and tPhoA. The stop codon of PON1 and tPON1 was removed to allow translation of the entire fusion. Additionally, PhoA and tPhoA were tested as controls. Cells to be tested were streaked on LB plates supplemented with phosphate (to suppress endogenous phosphatase expression), IPTG and the indicator bromo-4-chloro-3-indolyl phosphate (XP). Colony color was detected by eye: blue or bluish colonies indicated active, periplasmic PhoA, while completely white colonies indicated that the PhoA construct was inactive. The qualitative results of this assay are shown in Table 4.1 for constructs in *E. coli* and *P. putida*, along with data from a quantitative PNP-based PhoA assay of *E. coli* construct lysates. These data show that in both organisms, PON1-tPhoA is exported from the cytosol while tPON1-tPhoA is not,

indicating that PON1 is capable of export from the cytosol and that the first 23 N-terminal residues of PON1 are essential for this process.

**Table 4.1:** PhoA assay results for PON1-PhoA protein fusions. Qualitative results of a XP plate-based assay for phoA activity for various protein constructs in *E. coli* and *P. putida* are listed along with quantitative data from a PNP-based PhoA assay of *E. coli* construct lysates. Protein fusions are written from N- to C-terminus (e.g. N-PON1-tPhoA-C). A “t” in front of a protein name indicates that the protein has had its N-terminal signal sequence removed. PNP assay data given as the averages and standard deviations of three independent experiments.

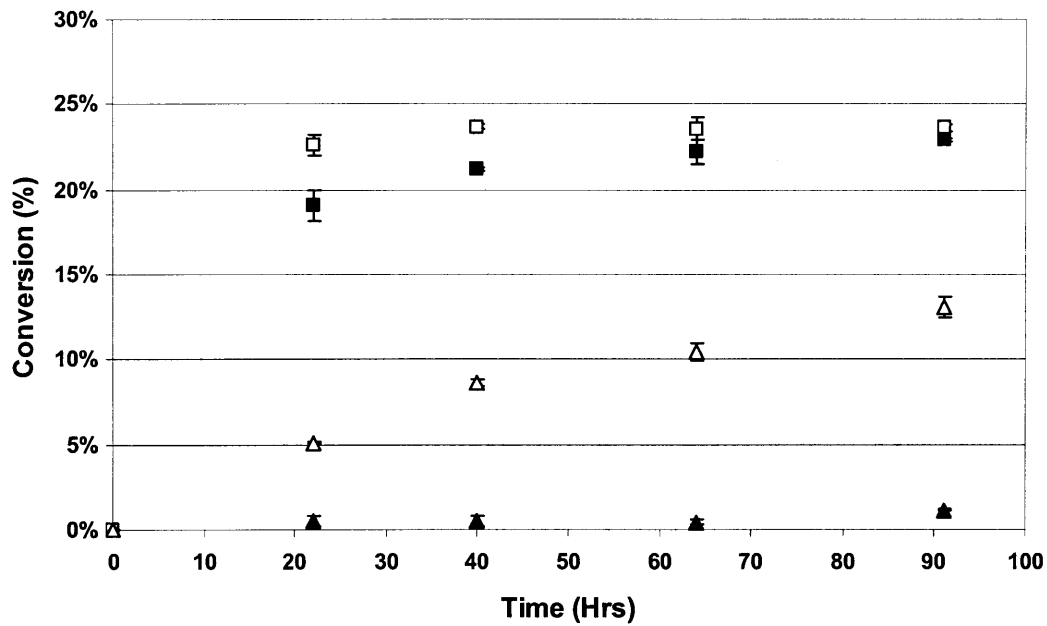
Construct	<i>E. coli</i> Colony Color on XP Plates	<i>P. putida</i> Colony Color on XP Plates	PNP PhoA Activity in <i>E. coli</i> (U/mg)
PhoA	Blue	Blue	1.79 ± 0.06
tPhoA	White	White	0.25 ± 0.05
PON1-tPhoA	Blue	Blue	3.88 ± 0.13
tPON1-tPhoA	White	White	0.26 ± 0.05

To provide additional verification of extracytosolic PON1 expression, whole-cells and lysates of *E. coli* expressing PON1 and tPON1 were assayed for lactonase activity at low pH (6.2) and high pH (7.2) (Figure 4.3). Based on the results of the PhoA fusion studies in the previous section, PON1 should be an extracytosolic lactonase while tPON1 should be a cytosolic lactonase. Thus in whole-cells, PON1 should be exposed to the extracellular pH while tPON1 should be exposed to only the intracellular pH of approximately 7.5. In lysates, however, both PON1 and tPON1 should be exposed to the medium pH, as there is no cytosolic membrane to shield them. As the degree of PON1-catalyzed lactonization is known to be pH-dependent (Teiber et al., 2003), PON1 exposed to lower pH values produces more 4VL from 4HV. Thus lactone production can be used to estimate the pH that PON1 is exposed to and consequently to identify whether PON1

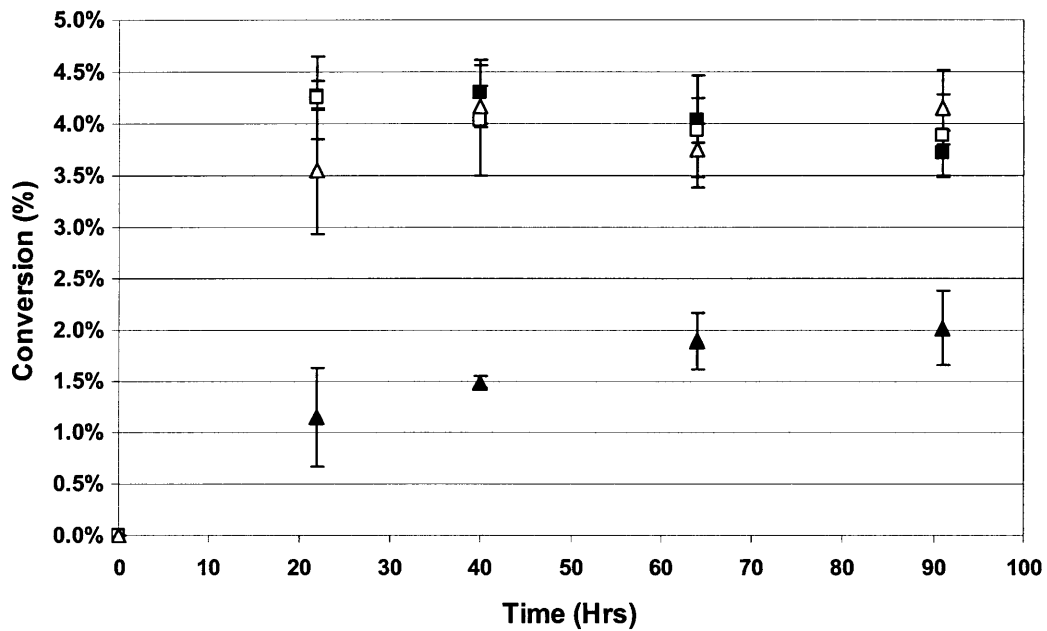
is localized intracellularly or extracytosolically. At an extracellular pH of 6.2, both whole-cells and lysates containing PON1 are highly active, while only the lysate from tPON1 cells is highly active at this pH (Figure 3a). Whole-cells expressing tPON1 are only minimally active, achieving only 1.5-2.0% conversion of 4HV into 4VL. At an extracellular pH of 7.2, the results are essentially the same (Figure 4.3b), the only difference being that the more active samples are limited to 4% conversion by the higher pH. The tPON1 sample here again achieves approximately 1.5-2.0% conversion, presumably because the tPON1 in whole-cells is exposed to the same higher intracellular pH (7.5) irrespective of the medium pH value.



(A)



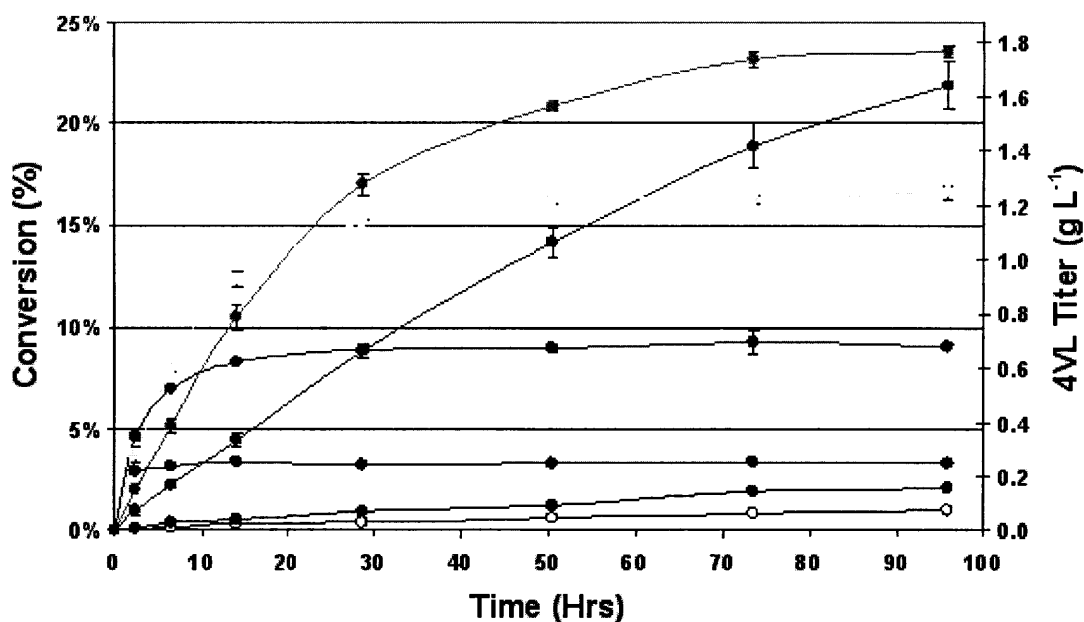
(B)



**Figure 4.3:** Conversion of 4HV to 4VL by *E. coli* whole-cells and lysates with expressed PON1 or tPON1 versus time at a pH of 6.2 (A) and 7.2 (B). Solid squares and triangles represent whole-cell data while empty squares and triangles represent lysate data. Squares are data obtained using PON1 while triangles are data obtained using tPON1. The data points shown are the averages and standard deviations of three independent experiments.

### **Effect of Medium pH on 4VL Production**

To better understand how medium pH affects lactone production, we assayed whole *E. coli* cells expressing PON1 for their ability to convert 75 mM 4HV into 4VL at pH values of 5.9, 6.2, 6.4, 6.7 and 7.2. *E. coli* cells expressing tPON1 and no PON1 (empty plasmid control) were also tested at a pH of 6.2, and the conversion of 4HV into 4VL was monitored over time. Both the rate and amount of conversion were strongly dependent on the pH, with lower pH's allowing for the highest conversions but at lower production rates (Figure 4.4). Higher pH values allowed for more rapid conversion but the conversion leveled out at a much lower level. This behavior is consistent with a pH-dependent equilibrium being established between 4HV and 4VL, a phenomenon that has previously been observed (Teiber et al., 2003). The pH effect on lactone production is significant – a decrease of a single pH unit (from 7.2 to 6.2) creates a 7-fold improvement in lactone production (Figure 4.4). The extracytosolic version of PON1 produced 11-fold more 4VL than tPON1 at a medium pH of 6.2.

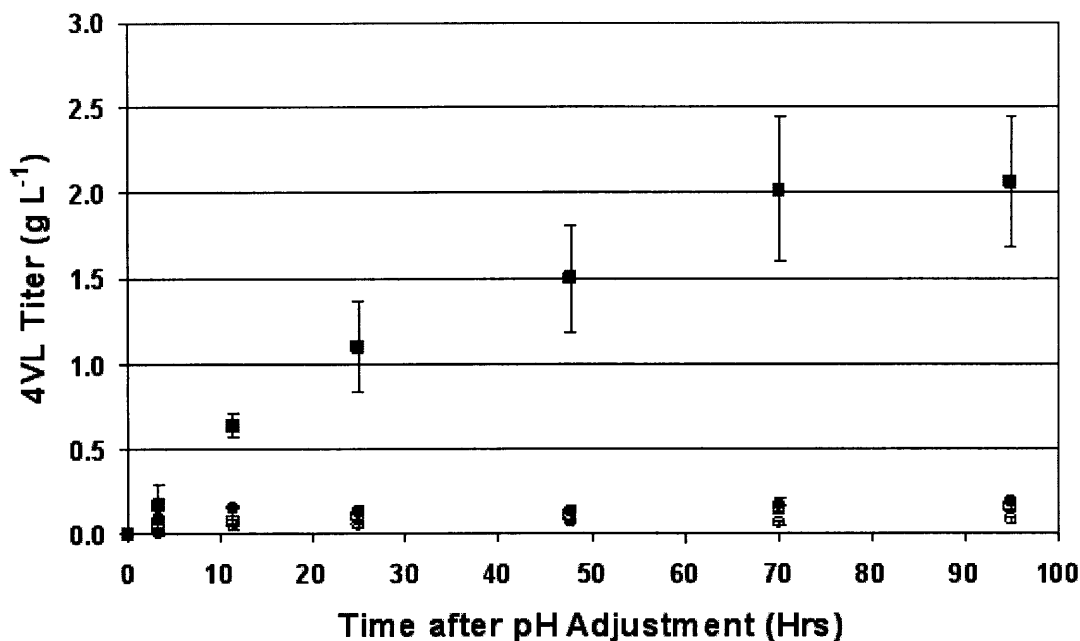


**Figure 4.4:** Conversion of 4HV to 4VL at various pH values by whole-cell *E. coli* expressing PON1 or tPON1. The solid squares, gray-filled squares, empty squares, solid triangles, and empty triangles represent samples with pH values of 5.9, 6.2, 6.4, 6.7 and 7.2 respectively for cells expressing PON1. The filled and empty circles represent lactonization by tPON1 and a no PON1 control, respectively, both at a pH value of 6.2. The data points shown are the averages and standard deviations of three independent experiments.

### Production of 4VL in Shake Flasks

The above observations of PON1 localization outside of the cytosol and the lactonization reaction being strongly pH dependent were combined to create an integrated bioprocessing approach to producing 4VL (Figure 4.1). In Chapter 3, it was found that *P. putida* is capable of producing high concentrations of 4HV from levulinate when *E. coli* thioesterase II (*tesB*) is expressed (Martin and Prather, 2009). Thus by supplying levulinate to recombinant *P. putida* expressing *tesB* and PON1, first 4HV and then 4VL is produced. Because low pH values can inhibit *P. putida* growth, 4VL production from levulinate was done in two phases. In the first phase, recombinant *P. putida* cells

expressing *tesB* and either PON1 or tPON1 were grown in LB medium supplemented with levulinate and the pH was unregulated. During this time  $10.9 \pm 1.3 \text{ g L}^{-1}$  and  $12.0 \pm 0.9 \text{ g L}^{-1}$  of 4HV were produced in the PON1 and tPON1 cultures respectively, and the pH of the cultures rose to approximately 8.0-8.5 (data not shown). No 4VL was detected during the first phase. After 96 hours, the cultures were split into two halves and the pH of the medium in each half was adjusted downward to either 6.3 or 7.3. During this second phase, lactone production was monitored for an additional 96 hours. While all cultures had similarly high concentrations of 4HV at the beginning of the second phase, the 4VL titer was significantly enhanced only by extracytosolic PON1 at pH 6.3 (Figure 4.5). The enhancement of 4VL titer was 11-fold for PON1 at pH 6.3 vs. 7.3 ( $2.1 \pm 0.4 \text{ g L}^{-1}$  4VL vs.  $0.19 \pm 0.02 \text{ g L}^{-1}$ ) and 13-fold for PON1 vs. tPON1 at pH 6.3 ( $2.1 \pm 0.4 \text{ g L}^{-1}$  4VL vs.  $0.15 \pm 0.02 \text{ g L}^{-1}$ ).

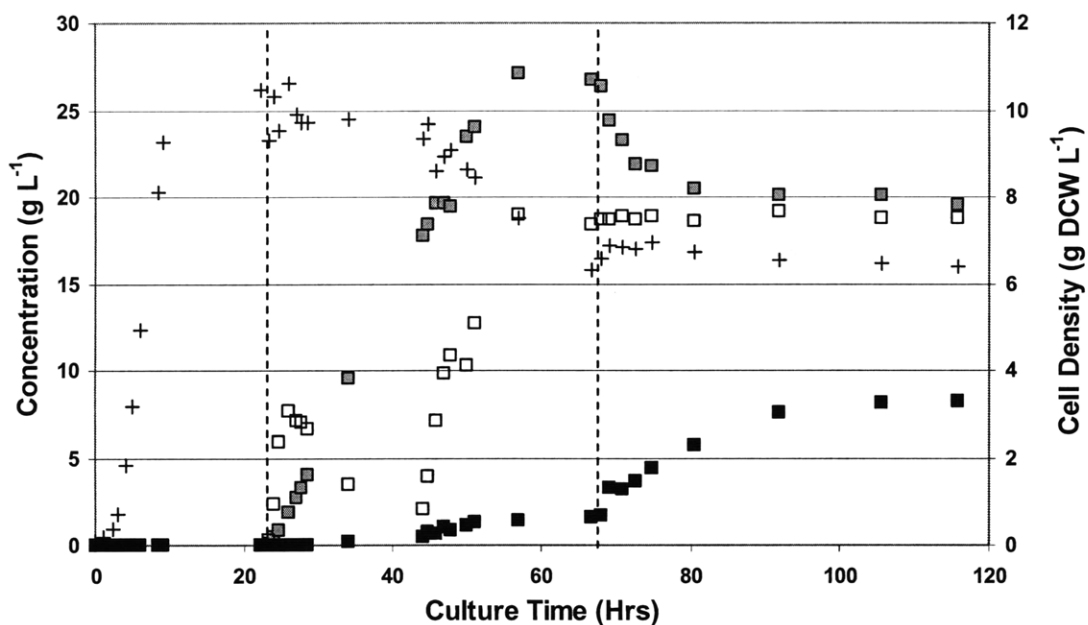


**Figure 4.5:** Production of 4VL from levulinate by recombinant *P. putida* expressing *tesB* and either PON1 (filled symbols) or tPON1 (open symbols) grown in shake flasks. Cultures were grown in LB medium supplemented with glucose and levulinate for 96 hours prior to  $t = 0$ , during which time the pH was unregulated.  $10.9 \pm 1.3 \text{ g L}^{-1}$  and  $12.0 \pm 0.9 \text{ g L}^{-1}$  of 4HV was produced in the PON1 and tPON1 cultures, respectively. At  $t = 0$  hours the pH of the medium was adjusted to either 6.3 (red squares) or 7.3 (blue circles). The data points shown are the averages and standard deviations of three independent experiments.

### Bioreactor-Scale Production of 4VL

Production of 4VL was scaled to a 2.0 L fed-batch reactor to take advantage of automated pH control as well as to further improve 4VL titer. 1.0 L of TB medium was inoculated with *P. putida* KT2440 expressing *tesB* and PON1 and the production of 4HV and 4VL was monitored over time (Figure 4.6). Levulinate feeding to the reactor did not begin until the cells were in stationary phase (22 hours after inoculation). Levulinate concentrations were maintained between 2 and 20  $\text{g L}^{-1}$  throughout the experiment to allow for 4HV production without inhibiting cellular metabolism with excessive levulinate. Having some excess levulinate was found to inhibit the metabolism of the

4HV product (data not shown). The pH was maintained at 7.0 during the 4HV accumulation phase. After 67 hours of cultivation the pH was shifted to 6.0 to allow for 4VL production and levulinate feeding was stopped.



**Figure 4.6:** Production of 4HV (gray-filled squares) and 4VL (black squares) from levulinate (empty squares) by recombinant *P. putida* expressing *tesB* and *PON1* grown in a 2.0 L bioreactor. Cell density is indicated by cross-shaped symbols. Levulinate feeding began after 23 hours (first dashed line), and at  $t = 67$  hours the pH of the medium was set to 6.0 to allow for 4VL accumulation (second dashed line). Cell density is indicated by crosses.

During the 4HV accumulation phase, the 4HV titer reached  $27.1 \text{ g L}^{-1}$  (Figure 4.6), which corresponded to a 26.2% yield from levulinate. The 4HV productivity of the reactor during this phase was  $0.81 \text{ g L}^{-1} \text{ hr}^{-1}$ . Prior to the pH shift, the titer of 4VL reached  $1.6 \text{ g L}^{-1}$ , corresponding to 6.5% lactonization of the 4HV produced. Significant 4VL accumulation did not occur until after the pH was shifted downward from 7.0 to 6.0.

After the pH shift, the titer of 4VL improved 5-fold to  $8.2 \text{ g L}^{-1}$  and the fraction of 4HV converted to 4VL increased to 33.1%.

## **Discussion**

### **Identification of G3C9 PON1 as an Extracytosolic Lactonase**

To establish the 4VL integrated bioprocessing system as a tool for product titer enhancement, it was first necessary to verify the functionality of the individual components. One of the central components of integrated bioprocessing is the utilization of extracytosolic space to carry out biochemical conversions under conditions more favorable than those in the cytosol. Thus it is critical to establish the extracytosolic expression of a component in the pathway of interest. Here, that component is the PON1 lactonase.

Because of differing evidence as to whether PON1 would export from the cytosol (Aharoni et al., 2004; Deakin et al., 2002), several experiments to test for protein export were performed. First, qualitative screening of protein fusions of PON1 with truncated PhoA (tPhoA) was performed in *E. coli* and *P. putida* by streaking these cells on plates containing the chromogenic PhoA substrate XP. Only PON1 protein fusions with an intact N-terminal signal sequence were found to have PhoA activity, and PON1 fusions without this sequence lacked PhoA activity (Table 4.1). These qualitative results were corroborated quantitatively through a *p*-nitrophenolphosphate-based assay for PhoA activity.

Additionally, lactonase assays were done with recombinant *E. coli* cells expressing PON1 and tPON1 to corroborate PON1 export. Lactone production is known to be highly pH-dependent, as lactones establish pH-dependent equilibria with their hydrolyzed forms (Figure 4.4; Teiber et al., 2003). Furthermore, it is known that the cytosolic pH remains constant for intact cells over a wide range of medium pH values (Wilks and Slonczewski, 2007). Thus by assaying lactonase activity at different pH values in cells expressing exported lactonase (PON1) versus intracellular lactonase (tPON1), one can determine if these lactonases are indeed exported or retained in the cytosol. The results of this experiment showed that PON1 exposed the medium pH (whether by protein export across the cytoplasmic membrane or in the case of lysates removal of the membrane altogether) had enhanced lactone production (Figures 4.3a and 4.3b) and that the degree of enhancement depended strongly on how much lower the medium pH is from the cytosolic pH.

These results, taken together, strongly support the hypothesis that the G3C9 PON1 variant is capable of export from the cytosol and that it uses an N-terminal signal sequence to accomplish the exportation. The primary N-terminal sequence of G3C9 PON1 contains a stretch of hydrophobic residues bracketed by appropriately-positioned lysine residues that presumably insert PON1 into the cellular membrane (Figure 4.2). This is characteristic of Sec-dependent protein transport (Manting and Drissen, 2000), though additional studies would be needed to verify this. Furthermore, it is unclear whether PON1 is periplasmic or is displayed on the outer membrane. It is clear however



that PON1 is not secreted into the medium, as we never detected lactonase activity in the supernatants of any of the cultures examined in this work (data not shown).

### **pH Effect on Lactone Production: Cytosolic versus Extracytosolic PON1**

Next, there needs to be a benefit to carrying out the extracytosolic reaction outside of the cytoplasm, namely an increase in product titer, in order to justify the use of an integrated bioprocess system. Lactone production using PON1 is highly pH-dependent (Teiber et al., 2003), however the intracellular pH is maintained at a relatively high and unfavorable level for lactone production – approximately 7.5 for *E. coli* for example (Wilks and Slonczewski, 2007). However by employing extracytosolic PON1 in media with lower pH values (relative to the cytosolic pH), the titers of lactone produced can be increased. To test this, whole-cells of recombinant *E. coli* expressing PON1 were suspended in buffered media at different pH values and supplied with 4HV to lactonize. The pH effect on lactonization was found to be quite potent: a decrease in a single pH unit increased the equilibrium amount of lactone seven-fold (23.6% conversion at pH 6.2 versus 3.3% conversion at pH 7.2; Figure 4.4). Expressing PON1 extracytosolically conveyed a full order of magnitude difference in lactone titer (23.6% conversion at a pH of 6.2 versus 2.1% for the tPON1 sample at a pH of 6.2).

### **Production of 4VL using Integrated Bioprocessing**

To complete the integrated bioprocessing lactone production system, the 4HV should be produced by the cell rather than supplied directly to the medium. This establishes that the increase in 4VL production by using an extracytosolic lactonase is due to the integrated

bioprocessing effect rather than an artifact of substrate transport across the cytosolic membrane. In Chapter 3, it was found that recombinant *P. putida* expressing thioesterase II (*tesB*) from *E. coli* was capable of producing high titers of 4HV from levulinate (Martin and Prather, 2009). Thus the full integrated bioprocessing system, one which combines both cytosolic hydroxyacid production with extracytosolic lactonization (Figure 4.1), can be tested in recombinant *P. putida* expressing both *tesB* and PON1.

Integrated bioprocessing was applied to the production of 4-valerolactone from levulinate in *P. putida*, and it was found that this approach was necessary to realize significant production from this novel pathway. A 13-fold improvement in lactone titer was realized by employing an extracytosolic lactonase for the lactonization reaction versus an intracellular lactonase control (Figure 4.5). Placing the lactonase enzyme outside of the cytosol, where its environment can be controlled by manipulating medium conditions, was key in realizing improved 4VL titers, as the benefits of using an extracytosolic lactonase over an intracellular one in this system disappear when the medium pH approaches that of the cytosol ( $0.19 \pm 0.02 \text{ g L}^{-1}$  4VL with PON1 at pH 7.3 versus  $0.15 \pm 0.02 \text{ g L}^{-1}$  for tPON1). To further increase 4VL titer and allow for automated pH control, the 4VL integrated bioprocessing system was tested at the bioreactor scale, and  $8.2 \text{ g L}^{-1}$  of 4VL was produced (Figure 4.6). As expected, 4VL production in the bioreactor was highly dependent on medium pH, and a 5-fold improvement in equilibrium 4VL titer was observed when the pH of the medium was decreased to 6.0 from 7.0

## **Applications and Implications of Integrated Bioprocessing**

Integrated bioprocessing is a viable strategy for enabling and improving product production in a broad array of biological systems. In general, integrated bioprocessing is applicable to any enzyme that performs sub-optimally under cytosolic conditions. Such enzymes would exhibit poor activity and consequently may become bottlenecks in production from desired metabolic pathways. By placing these enzymes outside of the cytosol, the conditions under which the enzymes operate can be easily manipulated by altering the properties of the culture medium.

The integrated bioprocessing approach employed here can be generalized to other enzymes who better function at lower or higher pH values than those found in the cytosol. Examples of such enzymes are *E. coli* glutamate decarboxylase, which has a pH optimum of about 4.5 (Capitani et al., 2003), and *Pseudomonas pseudoalcaligenes* alkaline lipase, which has a pH optimum of 8-10 (Lin et al., 1996). By localizing such enzymes outside of the cytosol and manipulating the medium pH, the activity from these enzymes can be improved. Yet another opportunity to use integrated bioprocessing is to overcome substrate transport issues with the cell membrane. A classic example of this strategy is the use of extracellular cellulases in both natural (Sun et al., 2008) and engineered (Shin and Chen, 2008) systems to degrade cellulosic matter for cellular uptake and metabolism. This concept can be expanded, for instance, to include extracytosolic expression of oxygen-requiring enzymes (such as oxygenases) in a metabolic pathway to give them better access to molecular oxygen. By placing enzymes outside of the cytosolic membrane, where significant oxygen consumption takes place

due to oxidative phosphorylation, these enzymes would be exposed to higher concentrations of oxygen.

However, integrated bioprocessing has several limitations, chief among them is the inability to use enzymes that require expensive cofactors. Cofactors such as NAD(H), NADP(H), and ATP are cytosolic; enzymes requiring these molecules cannot function extracytosolically and therefore do not readily lend themselves to integrated bioprocessing. Also while integrated bioprocessing allows one to manipulate the reaction conditions for a given enzyme, one still cannot use conditions that would kill the cell expressing the enzyme or denature the enzyme itself. For instance in this work, we could not obtain even larger amounts of 4VL by performing integrated bioprocessing with *P. putida* cells in pH 2 medium because neither the cells nor the enzyme would tolerate a pH value that low. While integrated bioprocessing allows one to “bend” the conditions under which biocatalysis occurs, one still cannot “break” the enzyme or its host cell.

Despite these exceptions, integrated bioprocessing remains a valuable option for enhancing the activity of enzymes that underperform in the cytosol. This methodology has been successfully applied to lactonization, a reaction that is difficult to perform in aqueous and cellular systems due to the significant rate of lactone hydrolysis that occurs at neutral pH. Through the use of integrated bioprocessing in this work, lactone titer increased by over an order of magnitude. Though in this work a pH difference across the cytosolic membrane was exploited to enhance product titer, other differences across this membrane such as oxygen concentration or redox state can be exploited for integrated

bioprocessing. All one needs is a suboptimal cytosolic enzyme, an N-terminal signal sequence to export the enzyme, and a parameter (pH, dissolved oxygen, etc.) that can be manipulated in the medium to enhance the exported enzyme's activity.

## **Materials and Methods**

### **Chemicals**

All chemicals were purchased at the highest grade or purity available unless otherwise indicated. LB broth, glucose, isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), and antibiotics were purchased from Becton Dickinson and Company (Sparks, MD), Mallinckrodt (Hazelwood, MO), Teknova (Hollister, CA), and Calbiochem (San Diego, CA), respectively. The sources for other chemicals are described in the relevant methods below.

### **Strains and Plasmids**

*Escherichia coli* DH10B was used for all *E. coli* studies and molecular cloning in this work and was purchased from Invitrogen (Carlsbad, CA). *Pseudomonas putida* KT2440 was obtained from the American Type Culture Collection (ATCC #47054). pRK415 (Tet<sup>R</sup>) was used to express the thioesterase II (*tesB*) gene from *E. coli* MG1655, while a gentamycin-resistant variant of the plasmid pMMB206, called pMMB206G, was used for the expression of PON1. pRK415 was a generous gift from Prof. Keith Poole while pMMB206 was obtained from ATCC (ATCC # 37808). pMMB206G was produced from pMMB206 by the introduction of a gentamycin resistance cassette excised with *XmnI* from the plasmid pBSL141 (ATCC #87146) at an *XmnI* restriction site in

pMMB206. All molecular biology manipulations were performed using standard cloning protocols (Sambrook and Russell, 2001).

The *tesB* gene was cloned by PCR into pRK415. Primers used were purchased from Sigma-Genosys (St. Louis, MO, USA) and were as follows (restriction sites used for cloning are underlined): EcoRI-*tesB*-FP (5'-GAATTCTACTGGAGAGTTATATGAGTCAGG-3') and Sall-*tesB*-RP (5'-GTCGACTTAATTGTGATTACGCATC-3'). HotStar HiFidelity DNA polymerase was purchased from Qiagen (Valencia, CA, USA) and used according to the manufacturer's instructions. The PON1 gene was similarly cloned by PCR into pMMB206G using the primers PstI-PON1-FP (5'-GACACTGCAGATGGCTAAACTGACAGCG-3') and XbaI-PON1-RP (5'-CATATCTAGATTACAGCTCACAGTAAAGAGCTTTG-3'). Truncated PON1 (tPON1) was made by eliminating the second through the 23<sup>rd</sup> amino acids from the N-terminus of PON1. tPON1 was amplified from pMMB206G-PON1 by PCR using PstI-tPON1-FP (5'-GACACTGCAGATGTCTTCTTTCCAAACACGAT-3') and XbaI-PON1-RP and inserted into pMMB206G to make pMMB206G-tPON1.

Alkaline phosphatase (PhoA) protein fusion vectors were constructed by first cloning *phoA* by PCR from DH10B *E. coli* genomic DNA into pRK415 using the primers XbaI-*phoA*-FP (5'-GACATCTAGAATGAAACAAAGCACTATTGCAC-3') and KpnI-*phoA*-RP (5'-GACAGGTACCTTATTTTCAGCCCCAGAGC-3') to make pRK415-*phoA*. Truncated *phoA* (tPhoA) was cloned in a similar manner into pRK415 using the primers XbaI-t*phoA*-FP (5'-GACATCTAGACGGACACCAGAAATGCCT-3') and KpnI-*phoA*-

RP). pRK415-PON1-tphoA and pRK415-tPON1-tphoA vectors were made by digesting pMMB206G-PON1 and pMMB206G-tPON1 with *Pst*I and *Xba*I and ligating into a similarly digested pRK415-tphoA.

### **Culture Conditions**

For shake-flask scale 4VL production experiments, recombinant *P. putida* harboring the plasmids pRK415-tesB and pMMB206G-PON1 or pMMB206G-tPON1 was cultured in 50 mL LB in a 250 mL shake flask at 32°C. Gentamycin (20 mg L<sup>-1</sup>) and tetracycline (10 mg L<sup>-1</sup>) were added to provide selective pressure for plasmid maintenance. Cultures were inoculated to an initial optical density at 600 nm (OD<sub>600</sub>) of 0.05, and 50 µL of 1 M IPTG was added at the beginning of the culture for induction of gene expression. 200 µL of a 5 mg L<sup>-1</sup> aqueous solution of bromothymol blue was added to the cultures as an indicator of culture pH. 1 mL of 20% glucose was added at approximately t = 0, 24, 48, and 72 hours. Similarly 2.0 M levulinate was supplied to the cultures as follows: 0.875 mL at t = 0 hours, 2.5 mL at t = 24 hours, and 3.75 mL at t = 48 and 72 hours. At t = 96 hours, the pH of the cultures was adjusted to the desired value using 6 N HCl. Additional 6 N HCl was added as needed to maintain the desired pH values. Cultures were then incubated at 32°C at the desired pH for an additional 96 hours to allow for 4VL production.

Bioreactor experiments were performed in a 2.0 L Biostat B bioreactor from Sartorius AG (Goettingen, Germany) equipped with two six-blade disk impellers. The bioreactor was inoculated with 50 mL of LB containing late-exponential phase recombinant *P. putida* harboring the plasmids pRK415-tesB and pMMB206G-PON1. The working

volume of the bioreactor was 1.0 L, and it was operated at 32°C. Air (0.5-1.0 vvm) was sparged into the bioreactor and the stirrer speed was varied between 300 and 850 rpm to maintain dissolved O<sub>2</sub> levels between 10% and 40% of air saturation. The pH of the reactor was set to 7.0 and was controlled with the automatic addition of 28% ammonium hydroxide and 4.0 M phosphoric acid. The medium in the reactor consisted of terrific broth (12 g L<sup>-1</sup> tryptone, 24 g L<sup>-1</sup> yeast extract, 9.4 g L<sup>-1</sup> potassium hydrogen phosphate and 2.2 g L<sup>-1</sup> potassium dihydrogen phosphate) with 1.0 mM IPTG, 2.0% glucose, 20 mg L<sup>-1</sup> gentamycin, 10 mg L<sup>-1</sup> tetracycline, 4 mM magnesium sulfate, 0.2 mM calcium chloride, and 0.1 mg/L ferric ammonium citrate. After 22 hours of initial growth, levulinic acid was continuously fed to the reactor from a 400 g L<sup>-1</sup> feedstock to maintain the concentration of levulinate between 2 and 20 g L<sup>-1</sup>. The concentrations of levulinate, 4HV, and 4VL in the bioreactor were monitored by HPLC, while cell density was monitored by measuring optical density at OD<sub>600</sub> and converted into g DCW L<sup>-1</sup> using a conversion factor of 0.42 g DCW OD<sub>600</sub><sup>-1</sup> L<sup>-1</sup> (15). After 67 hours, the pH of the culture was adjusted to 6.0 and levulinate feeding was discontinued. The culture was continued at 32°C for an additional 50 hours to allow for 4VL production.

### **Lactonase Assays**

Whole-cell and lysate samples were tested for lactonase activity in 1X M9 salts (Sambrook and Russell, 2001) with the nitrogen source (ammonium chloride) removed and exchanged for an equal molarity of sodium chloride to prevent the growth of whole-cell samples. Unless otherwise noted, this medium was supplemented with 0.1 M CaCl<sub>2</sub> to supply a divalent ion to PON1 (Teiber et al., 2003), 50 mM 4-hydroxyvalerate and 40



mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer. The pH of this medium was adjusted with 10 N NaOH or 6 N HCl to a desired value in the range of 5.0-7.0. When testing whole-cells, cells were centrifuged at 2,500 x g for 5 minutes, their original medium removed, and the cells were resuspended in 0.9% (w/v) sterile sodium chloride to an OD<sub>600</sub> of 25. This suspension was then used to supply cells to whole-cell lactonase assay experiments to an OD<sub>600</sub> of 0.5. For the analysis of cell lysates (prepared by repeated freezing and thawing of lysozyme-treated cells), 1 mg of total protein (bovine serum albumin equivalent as assayed by the Bradford method (Bradford, 1976)) was added to the lactonase assay mixture. All samples were then incubated with shaking at 37°C and samples were periodically withdrawn for HPLC analysis to determine the amount of 4VL produced.

### **Alkaline Phosphatase Assays**

Alkaline phosphatase (PhoA) activity was qualitatively assessed on agar plates using bromo-4-chloro-3-indolyl phosphate (XP) purchased from Amresco (Solon, OH) as an indicator. Recombinant *E. coli* or *P. putida* to be tested for PhoA activity were streaked onto LB plates supplemented with 10 mg L<sup>-1</sup> tetracycline, 1 mM IPTG, 100 mg L<sup>-1</sup> XP, and 75 mM phosphate (to suppress endogenous phosphatase expression). These plates were incubated at 37°C (*E. coli*) or 30°C (*P. putida*) for 24-48 hours and their color was assessed by eye. Blue colonies indicated active, periplasmic PhoA, while white colonies indicated that the PhoA construct was inactive.

PhoA activity was quantitatively assessed in cell lysates using *p*-nitrophenyl phosphate (PNP) purchased from Amresco (Solon, OH). The assay mixture consisted of 15 mM PNP and 2.0 mM MgSO<sub>4</sub> in 1.0 M Tris-HCl, pH 8.0. To this mixture crude protein lysate (30 µg of total protein bovine serum albumin equivalent as assayed by the Bradford method (Bradford, 1976)) was added and the solution was briefly vortexed to mix. The liberation of *p*-nitrophenol was monitored by measuring the absorbance of the mixture at 405 nm at room temperature. One unit of PhoA activity is defined as 1 µmol of *p*-nitrophenol liberated per minute at room temperature.

### **HPLC Analysis**

All HPLC samples were prepared by taking 1 mL of culture broth, centrifuging for 5 minutes at 16,000 x *g* to pellet cells, and taking the supernatant for analysis. HPLC analysis was performed on an Agilent 1100 Series instrument equipped with a Zorbax SB-Aq column (0.46 cm x 15 cm, 3.5 µm) purchased from Agilent Technologies (Santa Clara, CA). The column temperature was maintained at 65°C. Levulinic acid, 4HV, and 4VL were detected on a refractive index detector and had retention times of approximately 3.43, 3.28 and 5.11 minutes respectively. The mobile phase was 25 mM ammonium formate in water (pH 2.0) and the flow rate through the column was 1.0 mL min<sup>-1</sup>. Levulinic acid and 4VL purchased from Alfa Aesar (Ward Hill, MA) were used as standards, while the 4HV standard was prepared by saponification of 4VL with 10 N sodium hydroxide at 4°C.

## Chapter 5: 3-Hydroxybutyrate Production in *E. coli*

This work was published in Tseng et al., 2009 and much of the data in this chapter was obtained in collaboration with Hsien-Chung Tseng.

In this chapter, synthetic metabolic pathways constructed for the production of enantiopure (*R*)- and (*S*)-3-hydroxybutyrate (3HB) from glucose in recombinant *Escherichia coli* are described. To promote maximal metabolic flux towards 3HB, three thiolase homologs (BktB, Thil, and PhaA) and two CoA-removal mechanisms (Ptb-Buk and TesB) were profiled. Two enantioselective 3-hydroxybutyryl-CoA dehydrogenases (PhaB, producing the (*R*)-enantiomer and Hbd, producing the (*S*)-enantiomer) were utilized to control the 3HB chirality across two *E. coli* backgrounds, BL21Star(DE3) and MG1655(DE3), representing *E. coli* B and *E. coli* K-12 derived strains, respectively. MG1655(DE3) was found to be superior for the production of each 3HB stereoisomer although the recombinant enzymes exhibited lower *in vitro* specific activities compared to BL21Star(DE3). Hbd *in vitro* activity was significantly higher than PhaB in both strains. The engineered strains achieved titers of enantiopure (*R*)-3HB and (*S*)-3HB as high as 2.92 g L<sup>-1</sup> and 2.08 g L<sup>-1</sup>, respectively, in shake flask cultures within two days. The NADPH/NADP<sup>+</sup> ratio was found to be two- to three-fold higher than the NADH/NAD<sup>+</sup> ratio under the culture conditions examined, presumably affecting *in vivo* activities of PhaB and Hbd and resulting in greater production of (*R*)-3HB than (*S*)-3HB.

### **Introduction**

The synthesis of chiral molecules is of significant interest in the pharmaceutical industry because frequently one stereoisomer of a drug has efficacy while the other has either

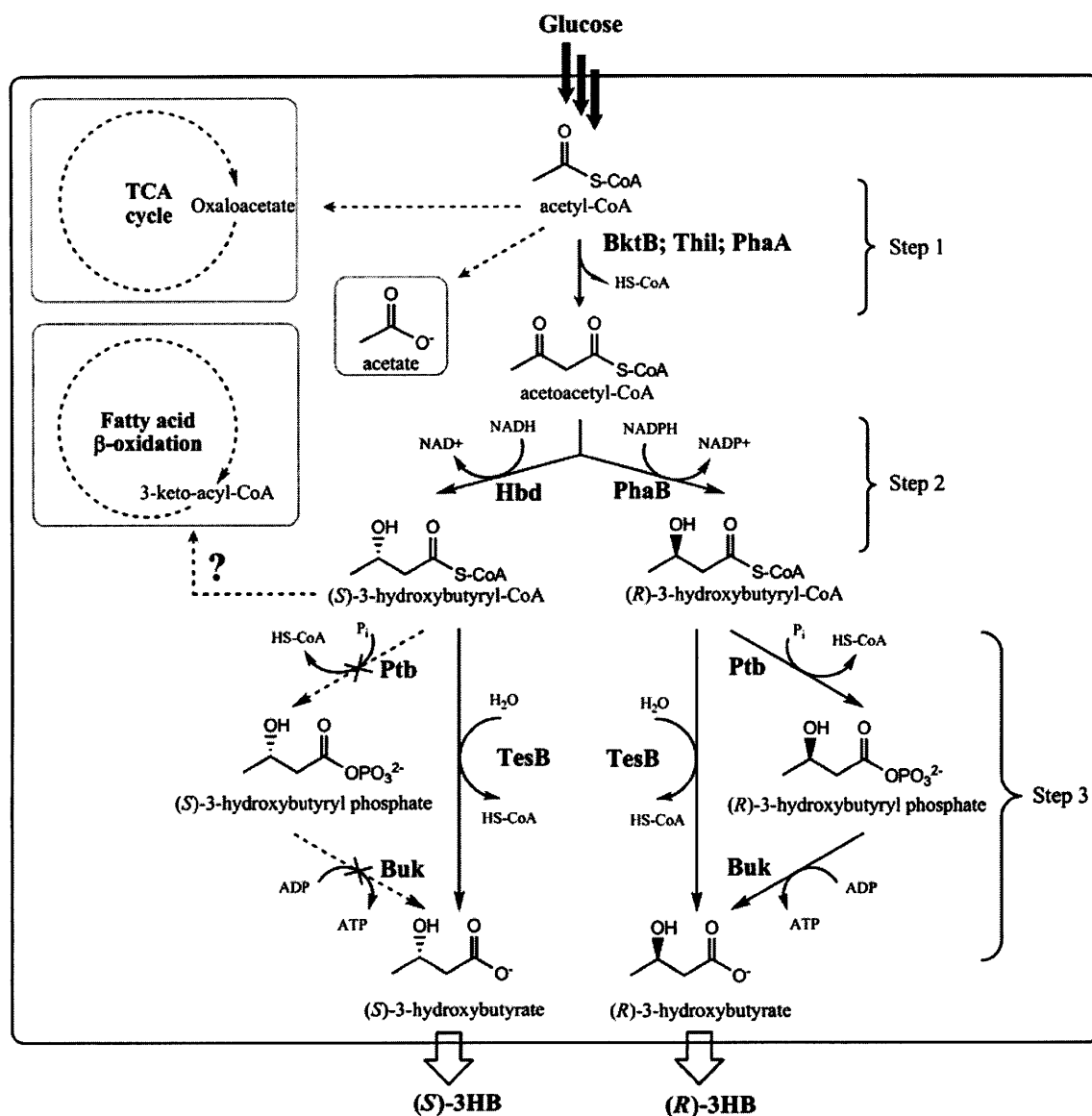
substantially reduced or no activity, or may even have adverse effects (Patel, 2006; Pollard and Woodley, 2007). Additionally, chiral molecules serve as building blocks for many pharmaceuticals and high value compounds. Thus, the ability to prepare chiral molecules with high optical purity is important. Stereoselective chemical processes generally employ expensive chiral catalysts, require harsh physical conditions and solvents, and suffer from extensive byproduct formation. In contrast, enzyme-catalyzed reactions are highly stereoselective and can be performed in aqueous solutions under mild conditions (Patel, 2006). As a result, the use of biological processes for chiral molecule production has been extensively investigated (Chen and Wu, 2005; Shiraki et al., 2006; Tokiwa and Calabia, 2008; Zhao et al., 2003). One example of such a process is the biosynthesis of 3-hydroxybutyric acid (3HB), a versatile chiral molecule containing one hydroxyl group and one carboxyl group, used as a building block for the synthesis of optically-active fine chemicals, such as vitamins, antibiotics, pheromones, and flavor compounds (Chiba and Nakai, 1985; Chiba and Nakai, 1987; Mori, 1981; Seebach et al., 1986).

The biosynthesis of 3HB has typically been achieved using two different mechanisms: depolymerization (*in vitro* or *in vivo*) of microbially synthesized poly-(*R*)-3-hydroxybutyric acid (PHB) (De Roo et al., 2002; Lee and Lee, 2003), or direct synthesis of 3HB without a PHB intermediate (Gao et al., 2002; Lee et al., 2008; Liu et al., 2007). However, due to the stereospecific constraints of PHB synthesis, in which polymers are composed exclusively of (*R*)-3HB monomer units, the synthesis of (*S*)-3HB from PHB is effectively impossible. In contrast, direct synthesis of both enantiopure (*R*)-3HB and

(*S*)-3HB is possible. Pathways facilitating (*R*)-3HB synthesis have been constructed in *Escherichia coli* by simultaneous expression of *phaA* (encoding acetoacetyl-CoA thiolase) and *phaB* (encoding (*R*)-3-hydroxybutyryl-CoA dyhydrogenase) from *Ralstonia eutropha* H16, and *ptb* (encoding phosphotransbutyrylase) and *buk* (encoding butyrate kinase) from *Clostridium acetobutylicum* ATCC 824 (Gao et al., 2002). In addition to the use of *ptb* and *buk* to catalyze the conversion of (*R*)-3HB-CoA to (*R*)-3HB, *tesB* (encoding thioesterase II from *E. coli*) has also been used for the direct hydrolysis of (*R*)-3HB-CoA to yield (*R*)-3HB (Liu et al., 2007). The production of (*S*)-3HB in *E. coli* has recently been reported using a biosynthetic pathway consisting of *phaA* from *R. eutropha* H16, *hbd* (encoding (*S*)-3-hydroxybutyryl-CoA dehydrogenase) from *C. acetobutylicum* ATCC 824, and *bch* (encoding 3-hydroxyisobutyryl-CoA hydrolase) from *Bacillus cereus* ATCC 14579 (Lee et al., 2008).

In *E. coli*, the synthesis of both enantiomers of 3HB begins with the condensation of two molecules of acetyl-CoA, catalyzed by a thiolase, to give acetoacetyl-CoA (Figure 5.1). The acetoacetyl-CoA is then reduced either to (*R*)-3-hydroxybutyryl-CoA ((*R*)-3HB-CoA) via ketone reduction mediated by an NADPH-dependent (*R*)-3-hydroxybutyryl-CoA dehydrogenase (PhaB), or to (*S*)-3-hydroxybutyryl-CoA ((*S*)-3HB-CoA) via an NADH-dependent (*S*)-3-hydroxybutyryl-CoA dehydrogenase (Hbd). (*R*)-3HB-CoA and (*S*)-3HB-CoA can each be further modified via a suitable CoA-removal reaction to form (*R*)-3HB and (*S*)-3HB, respectively. In an effort to increase chiral 3HB production, it is essential to identify a thiolase capable of efficiently catalyzing the first reaction in the 3HB biosynthetic pathways, to draw acetyl-CoA from competing endogenous pathways.

Thus, three different thiolases, BktB and PhaA from *R. eutropha* H16, and Thil from *C. acetobutylicum* ATCC 824, were examined to determine which is most proficient for 3HB synthesis. (*R*)-3HB-CoA and (*S*)-3HB-CoA synthesized via the reduction reaction catalyzed by PhaB and Hbd, respectively, must be converted to their respective free acid forms before transport or diffusion out of the cell. Two sets of CoA-removing enzyme mechanisms have been compared in this work, including the phosphotransbutyrylase (Ptb) and butyrate kinase (Buk) system encoded by the *ptb-buk* operon from *C. acetobutylicum* ATCC 824, and acyl-CoA thioesterase II (TesB) from *E. coli* MG1655. Moreover, it has long been argued whether B strains or K-12 strains of *E. coli* would serve as better hosts for the biosynthesis of small molecules. Microarrays and Northern blot analyses have suggested that several metabolic pathways, including the TCA cycle, glyoxylate shunt, glycolysis, and fatty acid degradation are different between these two strains (Phue et al., 2005; Schneider et al., 2002; Walle and Shiloach, 1998; Xia et al., 2008), implying that they may differ significantly in their ability to supply significant levels of acetyl-CoA as the precursor for 3HB synthesis. Thus, we have also compared 3HB synthesis across two representative *E. coli* strains: *E. coli* BL21Star(DE3) (B strain) and *E. coli* MG1655(DE3) (K-12 strain). 3HB chirality was examined and verified by high performance liquid chromatography (HPLC) analysis using a chiral stationary phase to provide separation.



**Figure 5.1:** Schematic representation of (S)-3HB or (R)-3HB synthesis from glucose in engineered *E. coli*. BktB, acetoacetyl-CoA thiolase from *R. eutropha* H16; Thil, acetoacetyl-CoA thiolase from *C. acetobutylicum* ATCC 824; PhaA, acetoacetyl-CoA thiolase from *R. eutropha* H16; Hbd, (S)-3-hydroxybutyryl-CoA dehydrogenase from *C. acetobutylicum* ATCC 824; PhaB, (R)-3-hydroxybutyryl-CoA dehydrogenase from *R. eutropha* H16; Ptb, phosphotransbutyrylase from *C. acetobutylicum* ATCC 824; Buk, butyrate kinase from *C. acetobutylicum* ATCC 824; TesB, acyl-CoA thioesterase II from *E. coli* MG1655.

## Results

### **Production of Chiral 3HB in BL21Star(DE3)**

The chiral 3HB pathways were first constructed in BL21Star(DE3), an *E. coli* B strain, by co-transforming one pETDuet-1 derivative (pET-H-B, pET-H-T, pET-H-P, pET-P-B, pET-P-T, or pET-P-P) and one pCDFDuet-1 derivative (pCDF-T or pCDF-PB) (Table 1). All strains were cultured at 30°C for 48 h in LB medium supplemented with 20 g L<sup>-1</sup> glucose, and accumulation of chiral 3HB in the culture medium was measured. The capabilities of two sets of CoA-removing enzymes, Ptb-Buk and TesB, were then compared for (*R*)-3HB or (*S*)-3HB production. TesB successfully removed the CoA moiety from both (*R*)-3HB-CoA and (*S*)-3HB-CoA to release the free acids (Table 5.1). In comparison, Ptb-Buk could also convert (*R*)-3HB-CoA but was essentially inactive on (*S*)-3HB-CoA. Titrers as high as 1.58 g L<sup>-1</sup> and 2.41 g L<sup>-1</sup> of (*S*)-3HB and (*R*)-3HB, respectively, could be achieved when TesB was used to remove CoA. In the case of Ptb-Buk, only trace amount of (*S*)-3HB (less than 0.10 g L<sup>-1</sup>) was produced, and 1.26 g L<sup>-1</sup> of (*R*)-3HB was observed. It was also noted that less acetate, 1.02 g L<sup>-1</sup> on average, was produced with the TesB system compared to that with the Ptb-Buk system, which was 1.55 g L<sup>-1</sup> on average.

The efficacies of three different thiolases on chiral 3HB production were then compared. Both BktB and PhaA were found to yield similar titers of (*R*)-3HB when co-expressed with PhaB and Ptb-Buk (~1.25 g L<sup>-1</sup>) or PhaB and TesB (~2.39 g L<sup>-1</sup>). However, Thil gave approximately 20% lower titers of (*R*)-3HB and increased acetate production compared to BktB and PhaA. This phenomena was not observed in (*S*)-3HB production,



where each of the thiolases resulted in similar titers of (*S*)-3HB, with an average of 1.52 g L<sup>-1</sup>. In general, TesB could outperform Ptb-Buk for CoA removal, resulting in significantly higher titers of both (*R*)-3HB and (*S*)-3HB (Table 5.1).

**Table 5.1:** Extracellular production of chiral 3HB by *E. coli* BL21Star(DE3) grown in shake flasks<sup>a, b</sup>

3HB Synthesis Pathways			Acetate (g L <sup>-1</sup> )	3HB (g L <sup>-1</sup> )	3HB specific content (g g-DCW <sup>-1</sup> )	3HB yield (g g-Glucose <sup>-1</sup> )
Step 1	Step 2	Step 3				
<i>(S)</i> -3HB	BktB		1.53 ± 0.13	<0.1	NA <sup>c</sup>	NA
	Thil	Hbd	1.79 ± 0.03	<0.1	NA	NA
	PhaA		1.55 ± 0.14	<0.1	NA	NA
<i>(R)</i> -3HB	BktB		1.33 ± 0.02	1.23 ± 0.06	1.19 ± 0.05	0.18 ± 0.01
	Thil	PhaB	1.66 ± 0.08	1.03 ± 0.02	0.69 ± 0.01	0.15 ± 0.01
	PhaA		1.41 ± 0.11	1.26 ± 0.06	1.19 ± 0.15	0.20 ± 0.03
<i>(S)</i> -3HB	BktB		0.89 ± 0.07	1.55 ± 0.05	1.50 ± 0.09	0.27 ± 0.01
	Thil	Hbd	1.09 ± 0.12	1.58 ± 0.07	1.22 ± 0.06	0.24 ± 0.01
	PhaA		1.01 ± 0.03	1.44 ± 0.04	1.43 ± 0.04	0.26 ± 0.02
<i>(R)</i> -3HB	BktB		0.88 ± 0.03	2.41 ± 0.04	2.24 ± 0.04	0.30 ± 0.00
	Thil	PhaB	1.20 ± 0.03	1.90 ± 0.06	1.52 ± 0.03	0.26 ± 0.01
	PhaA		1.05 ± 0.09	2.36 ± 0.08	2.06 ± 0.17	0.30 ± 0.02

<sup>a</sup> Cells were grown aerobically in LB media with the addition of 2% glucose at 30°C for 48 h. 1 mM IPTG was added when OD<sub>600</sub> reached 0.4~0.8.

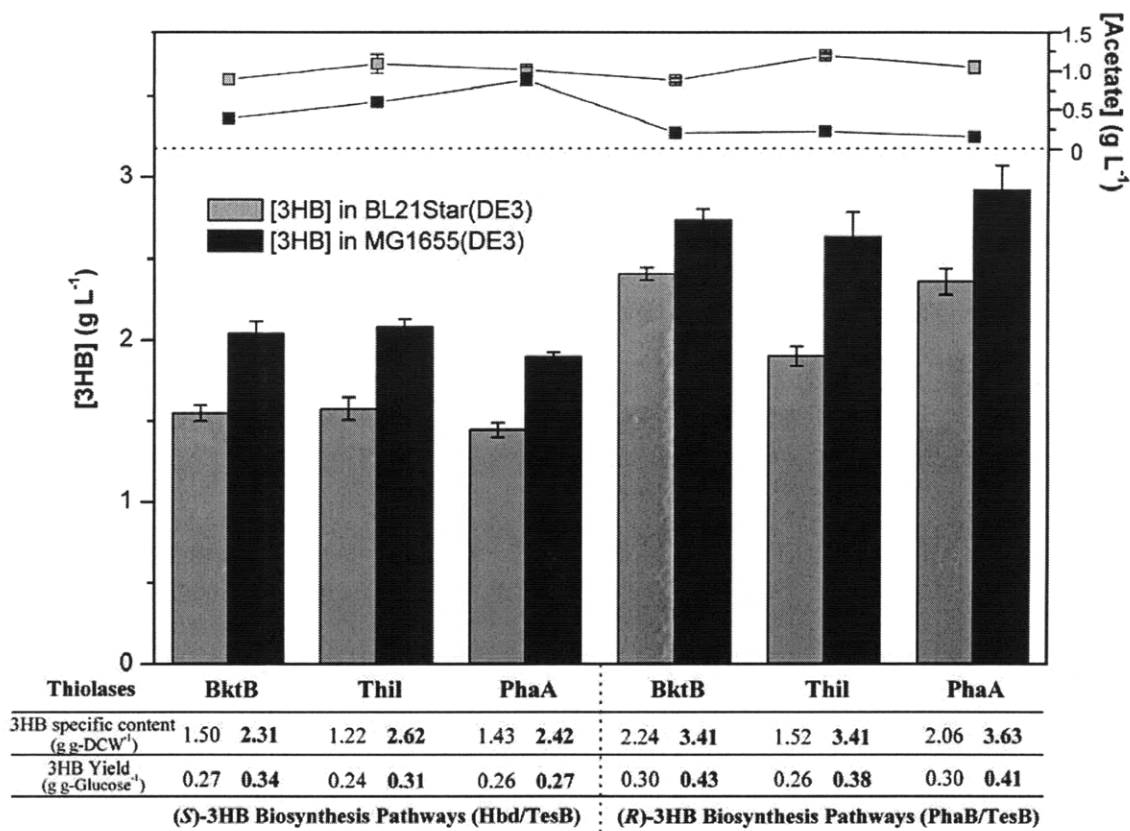
<sup>b</sup> Data are presented as the average value and standard deviation of measurements from three independent cultures.

<sup>c</sup> NA: not applicable

### Production of Chiral 3HB in MG1655(DE3)

Since TesB was identified as the most effective enzyme among those tested for CoA removal in BL21Star(DE3), further investigation of MG1655(DE3), an *E. coli* K-12 strain, exclusively employed 3HB pathways using TesB. In all cases, (*R*)-3HB or (*S*)-3HB production was substantially higher with MG1655(DE3) than with BL21Star(DE3) under the same culture conditions (Figure 5.2). MG1655(DE3) produced up to 2.08 g L<sup>-1</sup>

of (*S*)-3HB and 2.92 g L<sup>-1</sup> of (*R*)-3HB, ~30% and ~20% higher titers, respectively, than those produced by BL21Star(DE3). It is also interesting to note that generally less acetate was produced in MG1655(DE3) than in BL21Star(DE3), suggesting that more acetyl-CoA carbon flux was directed towards 3HB biosynthesis in MG1655(DE3) than towards acetate production. These two production systems were also compared in terms of 3HB specific content (g g-DCW<sup>-1</sup>) and 3HB yield (g g-Glucose<sup>-1</sup>). 3HB specific contents in MG1655(DE3) strains were ~50-120% greater than their respective BL21Star(DE3) counterparts as a result of increased 3HB production and reduced biomass accumulation (Figure 5.2). The comparison of 3HB yield on glucose shows that the efficiency of 3HB production from glucose in MG1655(DE3) was higher than in BL21Star(DE3). Generally, MG1655(DE3) was able to produce more chiral 3HB and less acetate, while also accumulating less biomass than BL21Star(DE3).

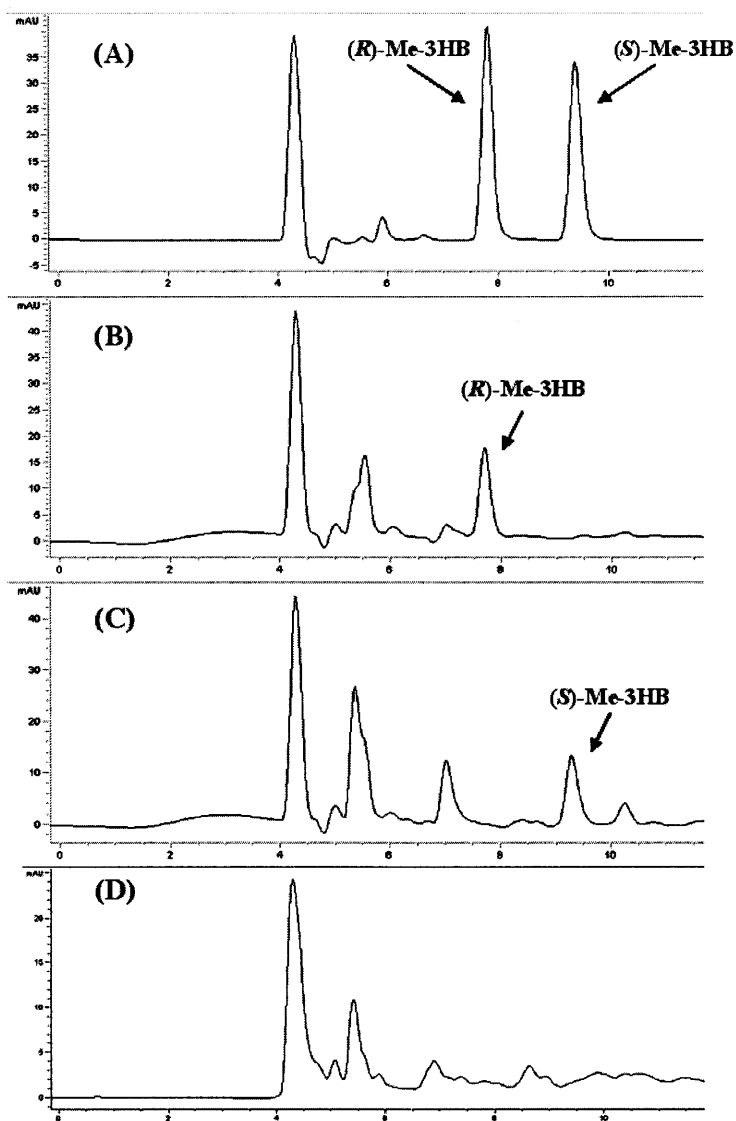


**Figure 5-2:** Extracellular production of chiral 3HB by *E. coli* BL21Star(DE3) and MG1655(DE3) grown in shake flasks. (*S*)-3HB was produced when Hbd was employed (left), while (*R*)-3HB was produced when PhaB was employed (right). In all cases, TesB was used to mediate CoA removal.

### Confirmation of 3HB Stereochemistry

*E. coli* BL21Star(DE3) harboring *bktB*, *tesB*, and either *phaB* or *hbd* was grown at 30°C in 50 mL LB supplemented with 20 g L<sup>-1</sup> glucose for 48 hours. The stereochemistry of the resulting 3HB in the media from these cultures was determined by methyl esterification of the 3HB present followed by chiral HPLC analysis as described in the Materials and Methods section at the end of this Chapter. The results confirm previous reports that cells harboring a thiolase and CoA-removal enzyme produce (*R*)-3HB if the

cells are concomitantly expressing *phaB* and synthesize (*S*)-3HB if they are expressing *hbd* (Figure 5.3; Lee et al., 2008; Liu et al., 2007)



**Figure 5.3:** HPLC spectra of (A) methyl-(*R*)-3HB and methyl-(*S*)-3HB standards, (B) culture medium from *E. coli* BL21Star(DE3) expressing *bktB*, *phaB* and *tesB* after boiling in methanol, (C) culture medium from *E. coli* BL21Star(DE3) expressing *bktB*, *hbd* and *tesB* after boiling in methanol, and (D) culture medium from *E. coli* BL21Star(DE3) after boiling in methanol as a control.

### Measurement of Specific Activities of 3HB Synthesis Enzymes Expressed in *E. coli*

To better understand the differences in 3HB titers between the different 3HB pathways in *E. coli* BL21Star(DE3) and MG1655(DE3), a correlation between *in vivo* chiral 3HB production and *in vitro* enzyme activities was sought. The activities of BktB, Thil, PhaA, Hbd and PhaB were measured in recombinant *E. coli* BL21Star(DE3) and MG1655(DE3). The background acetoacetyl-CoA thiolase activities in *E. coli* BL21Star(DE3) and MG1655(DE3) were very weak ( $<0.04$  U mg protein<sup>-1</sup>), and there were undetectable 3-hydroxybutyryl-CoA dehydrogenase activities (Table 5.2). Plasmid-based expression of BktB, Thil, PhaA, Hbd and PhaB was able to give functional enzymes in both *E. coli* strains (Table 5.2). It is interesting to note that all enzymes analyzed had greater specific activities in BL21Star(DE3) than in MG1655(DE3), showing approximately 130% to 360% higher thiolase activity and 24% to 44% higher 3-hydroxybutyryl-CoA dehydrogenase activity, although 3HB production was lower. For comparison of alternative acetoacetyl-CoA thiolases, Thil had approximately 6-fold higher specific activity than BktB and PhaA in BL21Star(DE3), and 6-fold and 3-fold higher specific activity than BktB and PhaA, respectively, in MG1655(DE3). For 3-hydroxybutyryl-CoA dehydrogenase activities of PhaB and Hbd, results show that the activities of Hbd (NADH-dependent) were approximately 140-fold and 120-fold higher than those of PhaB (NADPH-dependent) in BL21Star(DE3) and MG1655(DE3), respectively.

**Table 5.2:** Enzyme specific activities ( $\text{U mg}^{-1}$ )<sup>a</sup> of crude extracts of *E. coli* BL21Star(DE3) and MG1655(DE3)

Enzymes	<i>E. coli</i> Strains <sup>b</sup>		
	BL21Star(DE3)	MG1655(DE3)	
Acetoacetyl-CoA Thiolase	Control	0.04 ± 0.01	0.03 ± 0.00
	BktB	5.05 ± 0.69	1.10 ± 0.24
	Thil	29.66 ± 3.87	6.85 ± 0.97
	PhaA	5.90 ± 0.87	2.56 ± 0.36
3-Hydroxybutyryl-CoA Dehydrogenase	Control	ND <sup>c</sup>	ND
	Hbd	58.78 ± 7.76	40.76 ± 8.08
	PhaB	0.41 ± 0.18	0.33 ± 0.05

<sup>a</sup> One unit was defined as the conversion of 1  $\mu\text{mol}$  of substrate to product per min at 25°C.

<sup>b</sup> Data are presented as the average value and standard deviation of measurements from three independent cultures.

<sup>c</sup> ND: not detected

### Measurement of Cofactor Levels in Engineered MG1655(DE3) Strains

Although Hbd consistently showed higher *in vitro* activity than PhaB, the accumulated titers of (*S*)-3HB were lower than for (*R*)-3HB. Since these two enzymes have different pyridine nucleotide cofactor (NADPH/NADH) specificities, the intracellular levels of the reduced and oxidized cofactors were measured to gain additional insight into 3HB pathway performance (Table 5.3). At the late exponential phase (4 h post-induction), the specific contents of NADH and NADPH in all tested recombinant strains were found to be significantly lower than those of their respective oxidized forms ( $\text{NAD}^+$  and  $\text{NADP}^+$ ), with the ratios of  $\text{NADH}/\text{NAD}^+$  and  $\text{NADPH}/\text{NADP}^+$  ranging from 0.08-0.16 and 0.26-0.31, respectively. At the stationary phase (24 h post-induction), reduced cofactor increased concomitantly with the decrease of oxidized cofactor, resulting in higher ratios of  $\text{NADH}/\text{NAD}^+$  and  $\text{NADPH}/\text{NADP}^+$  ranging from 0.20-0.44 and 0.79-0.99,

respectively. Consistent with previously published reports (Brumaghim et al., 2003; Lee et al., 1996), in general NADH was found to be the predominant reducing equivalent in our *E. coli* strains while the NADPH/NADP<sup>+</sup> ratio was considerably higher than the NADH/NAD<sup>+</sup> ratio under the culture conditions examined.

**Table 5.3:** Levels and ratios of NAD<sup>+</sup>, NADH, NADP<sup>+</sup>, and NADPH cofactors in engineered MG1655(DE3) strains

Constructs	pET-H-P pCDF-T		pET-P-P pCDF-T		pETDuet pCDFDuet		
	4	24	4	24	4	24	
Post-induction time (hr)							
Cofactor levels <sup>a</sup> (nmol/mg-DCW)	NAD <sup>+</sup>	6.15 ± 0.75	3.95 ± 0.29	7.11 ± 0.08	4.94 ± 0.13	8.05 ± 1.09	4.79 ± 0.44
	NADH	0.97 ± 0.11	1.75 ± 0.02	0.59 ± 0.07	1.49 ± 0.21	0.85 ± 0.08	0.96 ± 0.06
	NADP <sup>+</sup>	1.58 ± 0.16	0.96 ± 0.08	1.89 ± 0.18	1.22 ± 0.05	2.12 ± 0.04	1.12 ± 0.12
	NADPH	0.50 ± 0.08	0.94 ± 0.09	0.49 ± 0.03	0.97 ± 0.23	0.59 ± 0.02	0.97 ± 0.09
Cofactor ratios	NADH/NAD <sup>+</sup>	0.16 ± 0.00	0.44 ± 0.03	0.08 ± 0.01	0.30 ± 0.03	0.11 ± 0.00	0.20 ± 0.01
	NADPH/NADP <sup>+</sup>	0.31 ± 0.02	0.99 ± 0.18	0.26 ± 0.01	0.79 ± 0.16	0.28 ± 0.02	0.87 ± 0.01

<sup>a</sup>Data are presented as the average and standard deviation of measurements from two independent cultures.

## Discussion

The underlying objective of this study was to explore the high level production of both (*R*)- and (*S*)-3HB in recombinant *E. coli* by investigating different host strains, thiolase homologs, and CoA-removal mechanisms. Shake flask-scale production of enantiomerically pure (*R*)-3HB and (*S*)-3HB to concentrations of up to 2.92 g L<sup>-1</sup> and 2.08 g L<sup>-1</sup>, respectively, were achieved. Biosynthesis of enantiopure (*S*)-3HB was reported at titers of 0.61 g L<sup>-1</sup> (Lee et al., 2008), which was the highest reported concentration prior to this study.

### ***E. coli* B versus *E. coli* K-12 in Chiral 3HB Production**

As previously described (Phue et al., 2005), there exist several intrinsic differences in metabolic pathways between *E. coli* B and K-12 strains, suggesting that the availability of metabolic intermediates as precursors for engineered biosynthetic pathways may also differ between these two strains. In addition, it has been generally concluded that the *E. coli* B strains are capable of producing greater amounts of proteins than the *E. coli* K-12 strains, making *E. coli* B strains better for protein production (Terpe, 2006). However, in terms of their role as microbial chemical factories, higher expression levels of recombinant proteins may not necessarily result in higher product titers, especially when substrate availability rather than enzymatic activity is rate limiting in the pathway. The results presented here support the hypothesis that host strain selection can critically influence the activity of recombinant enzymes as well as the productivity of a non-natural pathway. Although strains constructed in BL21Star(DE3) showed much higher acetoacetyl-CoA thiolase and 3-hydroxybutyryl-CoA dehydrogenase activities than those constructed in MG1655(DE3), the chiral 3HB titers from recombinant BL21Star(DE3) were roughly 20% to 30% lower than those from MG1655(DE3). Such discrepancy between product titers and enzyme activities in those strains suggests that distribution of precursors, i.e., acetyl-CoA, and not enzyme expression levels was the limiting factor for chiral 3HB production. It further implies that metabolic networks play an important role in chiral 3HB synthesis (Stelling et al., 2002). This hypothesis is supported by the observation that significantly more cell mass (data not shown) and acetate were accumulated in BL21Star(DE3) probably due to a large fraction of carbon flux drawn into the TCA cycle and acetate production pathway at the acetyl-CoA node



(Figure 5.1 and 5.2). This flux distribution resulted in less acetyl-CoA directed into engineered chiral 3HB pathways in BL21Star(DE3). In contrast, MG1655(DE3), likely with a different distribution within its metabolic network, achieved high-level production of chiral 3HB while accumulating less cell mass and acetate despite its relatively lower enzyme activities. Since excess accumulation of cell mass and byproduct formation represent an inefficient use of carbon resources that will reduce the yield of desired products, MG1655(DE3) served as the superior production strain by better balancing its growth while maintaining efficient production of chiral 3HB.

### **Effect of Alternative Acetoacetyl-CoA Thiolases on Chiral 3HB Production**

It was originally anticipated that the three different acetoacetyl-CoA thiolases (Step 1 of Figure 5.1) examined in this work would differ in terms of chiral 3HB production by directing different amounts of carbon flux (in the form of acetyl-CoA) into the engineered 3HB pathways. With the exception of Thil for (*R*)-3HB synthesis in BL21Star(DE3), the choice of thiolase had little effect on chiral 3HB titers even though the enzymes were found to display different specific activities *in vitro* in both recombinant strains (Table 5.2). Thil showed the highest *in vitro* specific activity in both strains, but recombinant BL21Star(DE3) with either Thil/PhaB/TesB or Thil/PhaB/Ptb-Buk yielded reduced (*R*)-3HB titers compared to other BL21Star(DE3) counterparts. To explain this contradiction, it should be noted that *in vitro* thiolase activity was assayed in the thiolytic direction, where two acetyl-CoA molecules were synthesized from one acetoacetyl-CoA and one CoA molecule. Thus, as a result of the combined effect of a hundred-fold lower PhaB activity compared to Hbd and a higher Thil activity compared

to BktB and PhaA in BL21Star(DE3), the acetoacetyl-CoA synthesized by Thil could accumulate in the cell. Since the thiolytic reaction is highly exergonic, thiolysis of acetoacetyl-CoA by thiolase is thermodynamically favored (Masamune et al., 1989). The accumulated acetoacetyl-CoA would then be cleaved into two acetyl-CoA molecules in the thiolytic direction, thereby supplying more acetyl-CoA to cell mass and acetate accumulation. In contrast, this did not occur in MG1655(DE3) probably due to the less active competing pathways towards cell mass and acetate accumulation and a negligible thiolytic reaction as a result of lower enzyme activities compared to BL21Star(DE3).

### **TesB versus Ptb-Buk as a CoA Removal System**

It has been suggested that the efficient removal of CoA from (*R*)-3HB-CoA can lead to enhanced (*R*)-3HB production (Gao et al., 2002), which could also be true for (*S*)-3HB production. To test this concept, two CoA-removal systems were assessed. The first is Ptb-Buk, encoded by an operon from *C. acetobutylicum*, which has been used for direct synthesis of polyhydroxyalkanoate (PHA) together with a PHA synthase utilizing the reverse reaction (i.e., the formation of the CoA thioester) (Liu and Steinbüchel, 2000b). The second is TesB from *E. coli*, which is reported to possess a broad substrate specificity but unknown physiological function in *E. coli* (Zheng et al., 2004). While Ptb-Buk uses a two-step CoA-removal scheme through a phosphorylated intermediate, TesB catalyzes one-step CoA-removal by direct hydrolysis (Figure 5.1). More chiral 3HB was likely produced in the TesB system than in the Ptb-Buk system due to the essentially irreversible hydrolysis by TesB. In addition, it was noted that pathways incorporating the Ptb-Buk system do not yield (*S*)-3HB (Table 5.1), which is consistent

with a previous report (Lee et al., 2008). The low level production ( $<0.10 \text{ g L}^{-1}$ ) of (*S*)-3HB produced by strains containing the Ptb-Buk system may have been due to the endogenous TesB activity in *E. coli*. In fact, in recombinant strains of *E. coli* BL21Star(DE3) in which *phaA* and *hbd* were solely expressed (i.e., no over-expression of *tesB* or *ptb-buk*), similarly low levels of (*S*)-3HB were also produced (data not shown).

### **Discrepancies between Enzyme Activities of Hbd and PhaB and 3HB Titers**

Although Hbd demonstrated much higher *in vitro* specific activities than PhaB, significantly lower titers of (*S*)-3HB than (*R*)-3HB were achieved in all strains expressing Hbd (Tables 5.1 and 5.2). This contradictory behavior may have been influenced by the following three factors *in vivo*: (1) the cofactor balance between NADH and NADPH and their respective oxidized counterparts; (2) the substrate preference of TesB for (*R*)-3HB-CoA over (*S*)-3HB-CoA; and (3) the competing pathway of fatty acid  $\beta$ -oxidation where (*S*)-3HB-CoA is an intermediate.

Given that Hbd and PhaB are NADH- and NADPH-dependent dehydrogenases, respectively, the physiological levels of NADH and NADPH and their redox ratios in *E. coli* most likely influences the *in vivo* catalytic activities of Hbd and PhaB, affecting chiral 3HB titers accordingly. Since NADH is the predominant reducing equivalent found in *E. coli* under normal conditions (Brumaghim et al., 2003) as well as conditions examined here (Table 5.3), these results indicate that Hbd should theoretically show higher *in vivo* activity than PhaB resulting in greater (*S*)-3HB than (*R*)-3HB production. The opposite, however, was observed in this study, suggesting that the physiological

NADH/NADPH ratio alone can not resolve the contradiction. It may instead be explained by differences in the physiological ratios of NADH/NAD<sup>+</sup> and NADPH/NADP<sup>+</sup>. It was shown in this study that the ratio of NADPH/NADP<sup>+</sup> is substantially higher than that of NADH/NAD<sup>+</sup> in *E. coli* MG1655(DE3) under the culture conditions used in this work (Table 5.3). These results, together with previous findings that the PHB synthesis is likely affected by the intracellular NADPH/NADP<sup>+</sup> ratio (Lee et al., 1996), suggest that in the case of chiral 3HB synthesis, a higher NADPH/NADP<sup>+</sup> ratio may result in more favorable reduction by PhaB while reduction by Hbd may be limited by the lower NADH/NAD<sup>+</sup> ratio. Overall, these observations are consistent with correspondingly higher yields of (*R*)-3HB than (*S*)-3HB in the present study.

In a similar manner to the substrate preference of Ptb-Buk for (*R*)-3HB-CoA, TesB might cleave (*R*)-3HB-CoA more efficiently than (*S*)-3HB-CoA. To verify this possibility, it will be informative to perform *in vitro* enzyme assays using (*S*)-3HB-CoA and (*R*)-3HB-CoA as substrates. Unfortunately, these chemicals are not commercially available, preventing the pursuit of this experiment.

The third hypothesis was tested using an *E. coli* mutant with an impaired fatty acid  $\beta$ -oxidation pathway. The *E. coli* MG1655(DE3) mutant was created using the method of Datsenko and Wanner (2000) by insertional inactivation of *fadB*, encoding 3-hydroxyacyl-CoA dehydrogenase. This *fadB* mutation should presumably attenuate or block the fatty acid  $\beta$ -oxidation cycle, thereby reducing the degradation of (*S*)-3HB-CoA. However, this mutant failed to achieve higher titers of (*S*)-3HB compared to its native

counterpart (data not shown). It is possible that other FadB homologs were involved in the degradation of (*S*)-3HB-CoA. For example, *fadJ* (previously called *yfcX*), encodes a subunit of enoyl-CoA hydratase that has been shown to possess the same catalytic function as FadB (Si Jae Park, 2004; Snell et al., 2002).

## Conclusions

The findings of this study suggest that the distribution of acetyl-CoA is likely the key factor affecting the production of chiral 3HB between *E. coli* BL21Star(DE3) and MG1655(DE3). Thus, in order to alter the distribution, further research should focus on blocking competing pathways for acetyl-CoA, for example, by deletion of acetate synthesis pathways comprised of acetate kinase and phosphotransacetylase (encoded by *ackA-pta*) or pyruvate oxidase (encoded by *poxB*) (Lin et al., 2005). Also, we can not entirely rule out the possibility that the overall higher titers of chiral 3HB in recombinant MG1655(DE3) might have been due to their higher TesB specific activities compared to BL21Star(DE3) counterparts in the case where CoA-removal is rate-limiting. Therefore, *in vitro* enzyme assays of both TesB and Ptb-Buk should be able to further elucidate the cause of superior titers of chiral 3HB in MG1655(DE3) than in BL21Star(DE3). Overall, production of  $\sim 3 \text{ g L}^{-1}$  (*R*)-3HB and  $\sim 2 \text{ g L}^{-1}$  (*S*)-3HB were achieved in shake flask cultures within two days. Further strain engineering should lead to more economical production of chiral 3HB. The high productivity of this 3HB pathway and the availability of several homologous pathway enzymes makes this pathway an attractive candidate for the biosynthesis of a broader range of 3-hydroxyalkanoic acids, including

3,4-dihydroxybutyric acid (DHBA). The substrate range of the 3HB pathway is explored in great detail in Chapter 6.

## **Materials and Methods**

### **Microorganisms**

The bacterial strains used in this work are listed in Table 5.4. *C. acetobutylicum* ATCC 824 was purchased from the American Type Culture Collection (ATCC, Manassas, VA). *R. eutropha* H16 was provided by Professor Anthony Sinskey of the Department of Biology at the Massachusetts Institute of Technology, USA. *E. coli* DH10B (Invitrogen, Carlsbad, CA) and ElectroTen-Blue (Stratagene, La Jolla, CA) were used for transformation of cloning reactions and propagation of all plasmids. BL21Star(DE3) (Invitrogen, Carlsbad, CA) and MG1655(DE3) were used as host strains for the biosynthesis of chiral 3HB, where MG1655(DE3) was constructed using a  $\lambda$ DE3 Lysogenization Kit (Novagen, Darmstadt, Germany) for site-specific integration of  $\lambda$ DE3 prophage into *E. coli* MG1655 (kindly donated by Professor Gregory Stephanopoulos of the Department of Chemical Engineering at the Massachusetts Institute of Technology, USA).

**Table 5.4:** *E. coli* strains, plasmids and oligonucleotides used

Name	Relevant Genotype	Reference
<b>Strains</b>		
DH10B	F <sup>-</sup> <i>mcrA</i> Δ( <i>mrr-hsdRMS-mcrBC</i> ) φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74</i> <i>recA1 endA1 araD139</i> Δ( <i>ara, leu</i> )7697 <i>galU galK</i> λ <sup>-</sup> <i>rpsL nupG</i>	Invitrogen
ElectroTen-Blue	Δ( <i>mcrA</i> )183 Δ( <i>mcrCB-hsdSMR-mrr</i> )173 <i>endA1 supE44 thi-1</i> <i>recA1 gyrA96 relA1 lac</i> Kan <sup>r</sup> [F' <i>proAB lac</i> <sup>F</sup> ΔM15 Tn10 (Tet <sup>r</sup> )]	Stratagene
MG1655	F <sup>-</sup> λ <sup>-</sup> <i>ilvG- rfb-50 rph-1</i>	ATCC 700926
BL21Star(DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) <i>gal dcm rne131</i> (DE3)	Invitrogen
MG1655(DE3)	F <sup>-</sup> λ <sup>-</sup> <i>ilvG- rfb-50 rph-1</i> (DE3)	This study
<b>Plasmids</b>		
pGEM-T easy	PCR cloning vector; F1 <i>ori</i> , Amp <sup>r</sup>	Promega
pETDuet-1	ColE1(pBR322) <i>ori, lacI, T7lac</i> , Amp <sup>r</sup>	Novagen
pCDFDuet-1	CloDF13 <i>ori, lacI, T7lac</i> , Str <sup>r</sup>	Novagen
pET-H	pETDuet-1 harboring <i>hbd</i> from <i>C. acetobutylicum</i> ATCC 824	This study
pET-H-B <sup>a</sup>	pETDuet-1 harboring <i>bktB</i> from <i>R. eutropha</i> H16, and <i>hbd</i> from <i>C. acetobutylicum</i> ATCC 824	This study
pET-H-T <sup>a</sup>	pETDuet-1 harboring <i>thil</i> and <i>hbd</i> from <i>C. acetobutylicum</i> ATCC824	This study
pET-H-P <sup>a</sup>	pETDuet-1 harboring <i>phaA</i> from <i>R. eutropha</i> H16, and <i>hbd</i> from <i>C. acetobutylicum</i> ATCC 824	This study
pET-P	pETDuet-1 harboring <i>phaB</i> from <i>R. eutropha</i> H16	This study
pET-P-B <sup>a</sup>	pETDuet-1 harboring <i>bktB</i> and <i>phaB</i> from <i>R. eutropha</i> H16	This study
pET-P-T <sup>a</sup>	pETDuet-1 harboring <i>thil</i> from <i>C. acetobutylicum</i> ATCC 824, and <i>phaB</i> from <i>R. eutropha</i> H16	This study
pET-P-P <sup>a</sup>	pETDuet-1 harboring <i>phaA</i> and <i>phaB</i> from <i>R. eutropha</i> H16	This study
pCDF-T	pCDFDuet-1 harboring <i>tesB</i> from <i>E. coli</i> MG1655	This study
pCDF-PB	pCDFDuet-1 harboring <i>ptb-buk</i> operon from <i>C. acetobutylicum</i> ATCC 824	This study
<b>Primers<sup>b</sup></b>		
	<b>Sequence 5'→3'<sup>c</sup></b>	
bktB_US_EcoRI	<u>GAATTC</u> ATGACGCGTGAAGTGGTAGTG	Sigma-Genosys
bktB_DS_XhoI	<u>CTCGAG</u> CGCAAGGCTAACCTCAGAT	Sigma-Genosys
thil_US_EcoRI	ATA <u>GAATTC</u> CATGAGAGATGTAGTAATAGTAAGTG	Sigma-Genosys
thil_DS_XhoI	TATTGAACCTC <u>CTCGAG</u> AACTTAGTTATAT	Sigma-Genosys
phaA_US_EcoRI	<u>GAATTC</u> GACTACACAATGACTGACGTTGTC	Sigma-Genosys
phaA_DS_XhoI	<u>CTCGAG</u> AAAACCCCTTCCTTATTTGC	Sigma-Genosys
hbd_US_EcoRI	<u>GAATTC</u> GGGAGGTCTGTTTAATGAAAA	Sigma-Genosys
hbd_DS_NotI	<u>GCGGCCG</u> CTGTAACCTTATTTTG	Sigma-Genosys
phaB_US_EcoRI	<u>GAATTC</u> AACGAAGCCAATCAAGGAG	Sigma-Genosys
phaB_DS_NotI	<u>GCGGCCG</u> CGCAGGTCAGCCCATATG	Sigma-Genosys
tesB_US_EcoRI	<u>GAATTC</u> TACTGGAGAGTTATATGAGTCAGG	Sigma-Genosys
tesB_DS_SalI	<u>GTCGACT</u> TAATTGTGATTACGCATC	Sigma-Genosys

<sup>a</sup> Each gene is under the control of the *T7lac* promoter with a ribosome binding site.

<sup>b</sup> Primers were synthesized at Sigma-Genosys, St. Louis, MO.

<sup>c</sup> Restriction enzyme sites used in the cloning are shown in underlined italics.

## Plasmid Construction

Genes derived from *C. acetobutylicum* ATCC 824 (*thil*, *hbd*, and *ptb-buk* operon), *R. eutropha* H16 (*phaA*, *bktB*, and *phaB*), and *E. coli* K-12 (*tesB*) were obtained by polymerase chain reaction (PCR) using genomic DNA (gDNA) templates. All gDNAs were prepared using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI). Custom oligonucleotides (primers) were purchased for all PCR amplifications (Sigma-Genosys, St. Louis, MO) as shown in Table 5.4. In all cases, HotStar HiFidelity Polymerase (Qiagen, Valencia, CA) was used for DNA amplification. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA). Recombination DNA techniques were performed according to standard procedures (Sambrook and Russell, 2001).

Two co-replicable vectors, pETDuet-1 and pCDFDuet-1 (Novagen, Darmstadt, Germany), were used for construction of chiral 3HB biosynthetic pathways (Tolia and Joshua-Tor, 2006). Both Duet vectors contain two multiple cloning sites (MCS), each of which is preceded by a T7*lac* promoter and a ribosome binding site (RBS), affording high-level expression of each individual gene.

For cloning genes, PCR products incorporated with desired restriction sites within the 5' and 3' primers were first cloned into pGEM-T Easy vector (Promega, Madison, WI) using TA cloning. The resulting plasmids were then digested and the genes were subcloned into pETDuet-1 or pCDFDuet-1. The *hbd* gene was inserted in between the *EcoRI* and *NotI* sites (MCS I) of pETDuet-1 to create pET-H. The *phaB* gene was inserted between



the *NcoI* and *NotI* sites (MCS I) of pETDuet-1 to create pET-P. Plasmids of pET-H-B, pET-H-T, and pET-H-P were constructed by inserting *bktB*, *thil*, and *phaA*, respectively, between the *MfeI* and *XhoI* sites (MCS II) of pET-H. In a similar manner, plasmids pET-P-B, pET-P-T, and pET-P-P were constructed by inserting *bktB*, *thil*, and *phaA*, respectively, between the *MfeI* and *XhoI* sites (MCS II) of pET-P. Plasmid pCDF-PB was created by inserting the *ptb-buk* operon between the *EcoRI* and *SaII* sites of pCDFDuet-1. Cloning the *tesB* gene between the *EcoRI* and *SaII* sites of pCDFDuet-1 resulted in plasmid pCDF-T. All constructs were confirmed to be correct by restriction enzyme digestion and nucleotide sequencing. Plasmids constructed in the present work are listed in Table 5.4.

### **Culture Conditions**

Seed cultures of the recombinant strains were grown in LB medium at 30°C overnight on a rotary shaker at 250 rpm. For the biosynthesis of chiral 3HB, the seed cultures were used to inoculate 50 mL LB medium in 250 mL flasks supplemented with 20 g L<sup>-1</sup> glucose at an inoculation volume of 5%. Cultures were incubated at 30°C on a rotary shaker until OD<sub>600</sub> reached 0.4~0.8. At this point, 1 mM IPTG was added to the cultures to induce recombinant protein expression. In all cases, LB medium was supplemented with 50 mg L<sup>-1</sup> ampicillin and 25 mg L<sup>-1</sup> streptomycin, and cultures were incubated aerobically. Culture broth was sampled at 4 h post-induction for enzyme assays, and at 48 h post-induction for HPLC analysis.

### **Metabolite Analysis and Dry Cell Weight Determination**

Samples were centrifuged to pellet cells while the aqueous supernatant was collected for HPLC analysis. Concentrations of 3HB, glucose, and acetate were analyzed via HPLC using an Agilent 1100 series instrument equipped with a refractive index detector (RID). Analyte separation was achieved using an Aminex® HPX-87H anion exchange column (Bio-Rad Laboratories, Hercules, CA) according to the method of Buday et al. (1990) using 5 mM H<sub>2</sub>SO<sub>4</sub> as the mobile phase.

The dry cell weight (DCW) was determined in triplicates by filtering culture broth (25 mL) through a pre-weighed filter paper (0.45- $\mu$ m cellulose acetate filters, Whatman, Maidstone, UK), followed by washing with deionized water (25 mL). The filter paper was then dried for two days in a 70°C oven and weighed for DCW determination. One OD<sub>600</sub> unit was equivalent to 0.42 g-DCW L<sup>-1</sup> for BL21Star(DE3) and 0.39 g-DCW L<sup>-1</sup> for MG1655(DE3).

### **Methyl Esterification of 3HB**

50 mL of 3HB-containing LB microbial culture was centrifuged at 6000 x *g* for 10 min and the supernatant was transferred to an evaporation dish for evaporating overnight in a 50°C oven, resulting in a brown residue. To this residue, 2 mL of acidic methanol (4:1 methanol:concentrated HCl) was added and the dish was shaken for two minutes to remove 3HB from the residue. The methanol solution was transferred into a test tube, sealed, and heated at 100°C for 3 hours. After heating, the solution was allowed to cool to room temperature. After cooling, 200  $\mu$ L of the methanol solution was mixed with 1

mL of pure isopropanol in a 1.7 mL tube, vortexed, and allowed to stand for 5 minutes. Precipitate in this sample was removed by centrifugation at  $13,200 \times g$  for 10 minutes. The supernatant was taken for chiral HPLC analysis.

### **Chiral HPLC Analysis of Methyl-3HB**

Chiral HPLC analysis was performed on an Agilent 1100 Series instrument equipped with a Chiralcel OD-H column (0.46 cm  $\phi$  x 25 cm) purchased from Daicel Chemical Industries (West Chester, PA). The column temperature was maintained at 40°C. Methyl-3HB was detected on a diode array detector at 210 nm. The mobile phase was 9:1 *n*-hexane:isopropanol and the flow rate through the column was 0.7 mL min<sup>-1</sup>. Methyl-(*R*)-3HB and methyl-(*S*)-3HB purchased from Sigma-Aldrich (St. Louis, MO) were used as standards and had retention times of 7.78 and 9.37 min, respectively.

### **Enzyme Assays**

Samples of 5 mL culture broth were centrifuged at  $5,000 \times g$  and 4°C for 10 min. The pellets were resuspended in 200  $\mu$ L of lysing buffer (1 Roche protease inhibitor cocktail tablet in 10.5 mL of 10 mM Tris-HCl supplemented with 1 mg mL<sup>-1</sup> lysozyme; pH = 8.0). Crude cell lysates were prepared by freeze-thaw lysis, carried out by subjecting cells to five cycles of freezing in liquid nitrogen followed by thawing at 37°C. After centrifugation at  $14,000 \times g$  and 4°C for 20 min, the supernatant was taken for enzyme assays. The total protein concentration of this supernatant was determined by the method of Bradford (1976) using bovine serum albumin as a standard.

Activities of thiolases, including BktB, Thil, and PhaA, were assayed using acetoacetyl-CoA and CoA as substrates, and the decrease in acetoacetyl-CoA concentration was followed at 303 nm (Inui et al., 2008). Activities of PhaB and Hbd were assayed using acetoacetyl-CoA as substrate and NADPH or NADH as cofactor, respectively. The decrease of NADPH or NADH concentration resulting from 3-hydroxybutyryl-CoA formation from acetoacetyl-CoA was measured at 340 nm (Inui et al., 2008; Schubert et al., 1988).

### **Quantification of Intracellular Cofactor Levels**

The NAD<sup>+</sup>, NADH, NADP<sup>+</sup>, and NADPH were each determined using Enzychrom™ NAD<sup>+</sup>/NADH and NADP<sup>+</sup>/NADPH assay kits (Bioassay Systems, Hayward, CA). Manufacturer's protocols were followed except that for sample preparation cells were resuspended in 100 µl extraction buffer and sonicated for 10 seconds on ice before heat extraction. The assays utilized alcohol dehydrogenase and glucose dehydrogenase for NAD(H) and NADP(H) quantifications, respectively, and colorimetric changes in the samples were measured at 565 nm.

## Chapter 6: Production of Higher Hydroxyacids in *E. coli*

In this chapter, the 3HB pathway developed in Chapter 5 is expanded to produce more complex 3-hydroxyacids – specifically 3HV and DHBA. Also, the lactonization of DHBA to 3-HBL is investigated in detail. These other products are made by supplying the non-natural substrates propionate (for 3HV production) or glycolate (for DHBA and 3-HBL production) to the 3HB pathway. These non-natural substrates are thiolated by propionyl-CoA transferase (*pct*) to form CoA thioesters capable of condensing with acetyl-CoA to eventually form the novel hydroxyacid products. For 3HV production, titers of 537 mg L<sup>-1</sup> of (*R*)-3HB and 522 mg L<sup>-1</sup> of (*S*)-3HV were achieved. For DHBA production, 3.2 g L<sup>-1</sup> of DHBA along with 230 mg L<sup>-1</sup> of 3-HBL were produced. The DHBA/3-HBL product profile can be shifted towards 3-HBL in one of two ways – by either removing the *tesB* gene and allowing for the auto-lactonization of DHBA-CoA or by treating DHBA-rich culture supernatants with acid to chemically lactonize the DHBA into 3-HBL. Removing the *tesB* gene resulted in an average improvement in 3-HBL production of 52%. The latter method of chemically lactonizing the culture supernatant with acid yielded 420 mg L<sup>-1</sup> of DHBA and 2.17 g L<sup>-1</sup> of 3-HBL, a 9-fold improvement in 3-HBL titer.

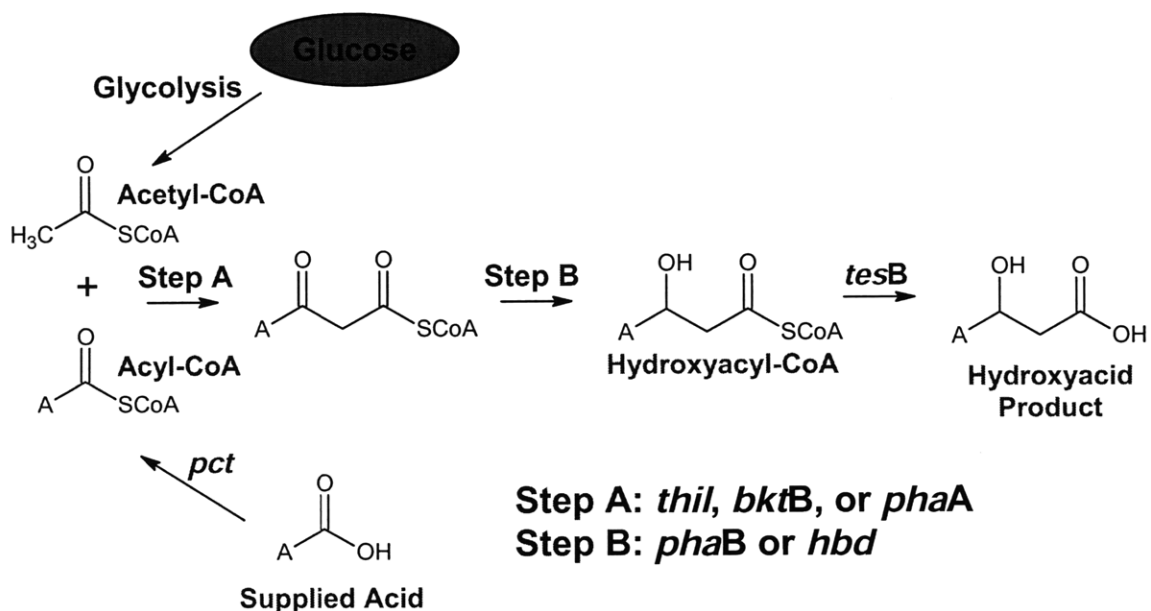
### **Introduction**

As discussed in Chapter 5, hydroxyacids are versatile, chiral compounds that contain both a carboxyl and a hydroxyl moiety, readily allowing for their modification into several useful derivatives (Lee et al., 2002; Chen and Wu, 2005) and making them suitable for

applications in the synthesis of antibiotics (Chiba and Nakai, 1985),  $\beta$ - and  $\gamma$ -aminoacids and peptides (Park, et al. 2001; Seebach et al., 2001), and as chiral synthetic building blocks (Lee et al., 2002). Direct biological production of hydroxyacid monomers has been successfully demonstrated for 3HB, and titers of 3 g L<sup>-1</sup> and 12 g L<sup>-1</sup> on the shake flask and fed-batch scales have been reported (Gao et al., 2002). In these reports, 3HB is made from acetyl-CoA through the use of acetyl-CoA acetyltransferase (*phbA*), 3-hydroxybutyryl-CoA dehydrogenase (*phbB*), phosphotransbutyrylase (*ptb*), and butyrate kinase (*buk*) (Liu and Steinbüchel, 2000a; Liu and Steinbüchel, 2000b; Gao et al., 2002; Figure 5.1). The last two of these enzymes were chosen to remove the CoA moiety from 3-hydroxybutyryl-CoA to yield free 3HB and were taken from *Clostridium acetobutylicum*, where they participate in the production of butyrate from butyryl-CoA (Liu and Steinbüchel, 2000b). Recently, thioesterase II (*tesB*) from *Escherichia coli* K12 (Naggert et al., 1991) was successfully employed to directly hydrolyze the acyl-thioester of 3HB-CoA (Liu et al., 2007; Tseng et al., 2009; Figure 5.1).

The success of direct production of 3HB from the condensation of two acetyl-CoA moieties opens up the possibility of producing more structurally diverse hydroxyacids through generalizing this pathway to the condensation of an acyl-CoA molecule with acetyl-CoA (Figure 6.1). In the generalized 3HB pathway, referred to as the 3-hydroxyalkanoic acid pathway, one of the three acetoacetyl-CoA thiolases from Chapter 5 (*thil*, *bktB*, or *phaA*) is employed to perform the condensation reaction to yield a 3-ketoacyl-CoA intermediate, which is subsequently reduced by either *phaB* or *hbd* to yield a 3-hydroxyacyl-CoA compound. Finally, *tesB* is used to hydrolyze the CoA moiety

from the 3-hydroxyacyl-CoA, liberating the free 3-hydroxyacid. These steps are exactly identical to those used in 3HB biosynthesis (Figure 5.1). One additional complication in the 3-hydroxyalkanoic acid pathway is that unlike acetyl-CoA, short chain acyl-CoA compounds that would be candidate substrates for this pathway are not readily available metabolites in *E. coli*. To circumvent this, propionyl-CoA transferase (*pct*), a broad substrate-specificity enzyme from *Megasphaera elsdenii* (Schweiger and Buckel, 1984; Taguchi et al., 2008) that exchanges CoA moieties between short-chain organic acids, is employed. In the 3-hydroxyalkanoic acid pathway, *pct* is used to transfer CoA from acetyl-CoA to an acid supplied to the cell (such as propionate), forming the acyl-CoA required by the pathway. Table 6.1 summarizes the characteristics and sources for each enzyme employed in the 3-hydroxyalkanoic acid pathway shown in Figure 6.1.



**Figure 6.1:** Schematic of the 3-hydroxyalkanoic acid pathway, a generalized version of the 3HB pathway shown in Figure 5.1. One key difference between these two pathways is the addition of *pct* to generate acyl-CoA molecules used in the condensation reaction with acetyl-CoA.

**Table 6.1:** General characteristics and sources of enzymes used in the 3-hydroxyalkanoic acid pathway shown in Figure 6.1.

<b>Enzyme</b>	<b>Organism of Origin</b>	<b>Properties</b>	<b>Reference</b>
<i>bktB</i>	<i>R. eutropha</i> H16	Acetyltransferase, broad substrate range (C <sub>4</sub> -C <sub>6</sub> )	Slater et al., 1998
<i>hbd</i>	<i>C. acetobutylicum</i> 824	Dehydrogenase, forms S stereoisomer product	Boynton et al., 1996
<i>pct</i>	<i>M. elsdenii</i>	CoA-transferase, wide substrate range (C <sub>2</sub> -C <sub>4</sub> )	Schweiger and Buckel, 1984
<i>phaA</i>	<i>R. eutropha</i> H16	Acetyltransferase, high activity and somewhat wide substrate range	Schubert et al., 1988
<i>phaB</i>	<i>R. eutropha</i> H16	Dehydrogenase, forms R stereoisomer product	Schubert et al., 1988
<i>tesB</i>	<i>E. coli</i> K12	Thioesterase, very broad substrate range (C <sub>4</sub> -C <sub>16</sub> )	Huisman et al., 1991
<i>thil</i>	<i>C. acetobutylicum</i> 824	Acetyltransferase, high activity	Stim-Herndon et al., 1995

The ultimate goal of this work was to use the 3-hydroxyalkanoic acid pathway to synthesize 3-hydroxybutyrolactone (3-HBL) and/or its hydrolyzed form 3,4-dihydroxybutyric acid (DHBA) from sugars and sugar-derived substrates. 3-HBL is widely used in the pharmaceutical industry as a building block for the class of cholesterol-reducing drugs called statins such as Crestor<sup>(R)</sup> and Lipitor<sup>(R)</sup> as well as the antibiotic Zybox<sup>(R)</sup> and the anti-hyperlipidemic medication Ezetimibe<sup>(R)</sup> (Lee et al., 2008; Lee and Park, 2009). Other pharmaceuticals derived from 3-HBL include HIV inhibitors (Kim et al., 1995) and the nutritional supplement carnitine (Wang and Hollingsworth, 1999b). 3-HBL has been listed as one of the top ten value-added chemicals by the U.S. Department of Energy (Werpy and Peterson, 2004). 3-HBL can readily be transformed into a variety of three-carbon building blocks (Wang and Hollingsworth, 1999a), and Figure 1.2 (taken from Werpy and Peterson, 2004) illustrates a variety of additional

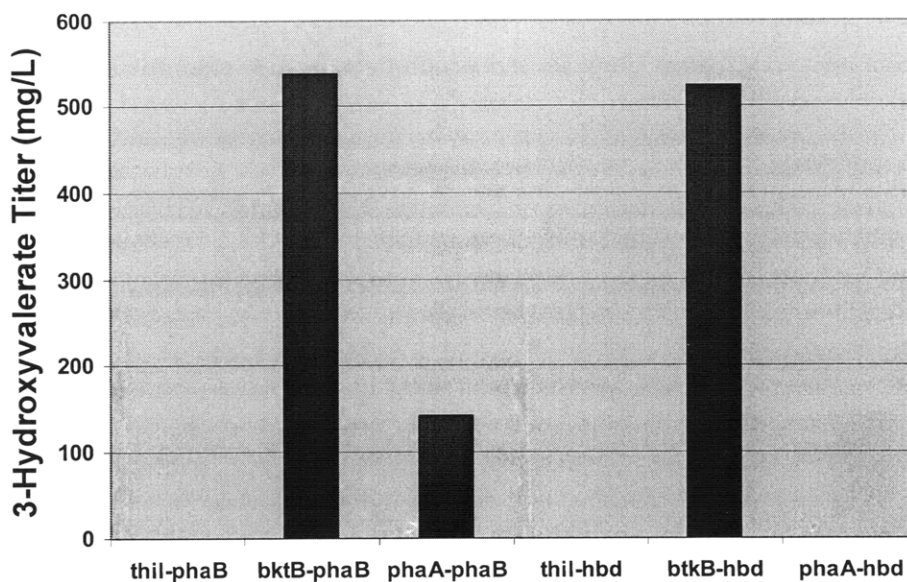


derivitization routes available for 3-HBL. In this work, the production of 3-hydroxyvalerate (3HV) from glucose and propionate is first examined to test the ability of the hydroxyalkanoic acid pathway to uptake free acids and convert them to CoA thioesters. The work then progresses onto DHBA and 3-HBL biosynthesis from glucose and glycolate and examines several factors controlling the titers of and ratio between these two products.

## **Results**

### **Production of 3HV from Glucose and Propionate**

The 3-hydroxyalkanoic acid pathway was first tested using propionate as a substrate. Neutralized propionate was supplied along with glucose to recombinant *E. coli* cells expressing one of three acetoacetyl-CoA thiolases (*thil*, *bktB*, or *phaA*), one of two 3-hydroxybutyryl-CoA reductases (*phaB* or *hbd*), *tesB*, and *pct*. These cultures were incubated for 48 hours at 30°C, and the resulting 3HV titers were measured by HPLC (Figure 6.2). 3HV was detected in cultures expressing *bktB-phaB* (537 mg L<sup>-1</sup>), *phaA-phaB* (142 mg L<sup>-1</sup>), and *bktB-hbd* (522 mg L<sup>-1</sup>). No 3HV was detected in the other three pathway gene combinations. The result that pathways with *bktB* yield the highest 3HV titer is consistent with reports that BktB has high activity towards the condensation of a C<sub>3</sub> substrate with a C<sub>2</sub> to form a C<sub>5</sub> product (Slater et al., 1998). The choice of reductase in the 3-hydroxyalkanoic acid pathway did not significantly influence 3HV production, except in the case of *phaA*, where 3HV was only detected in the *phaA-phaB* gene combination.

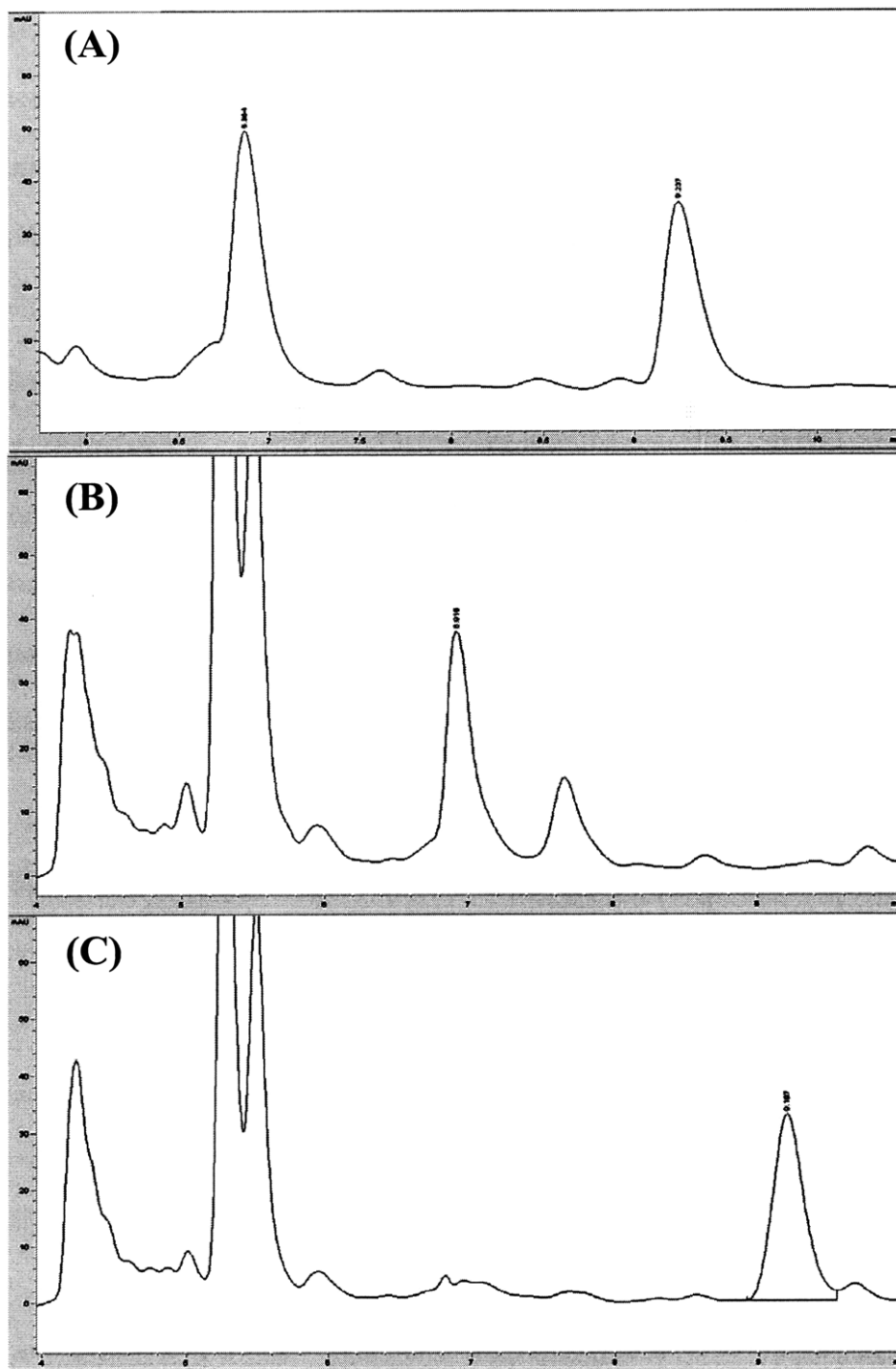


**Figure 6.2:** Production of 3HV by recombinant *E. coli* expressing *tesB*, *pct*, one of three acetoacetyl-CoA thiolases (*thil*, *bktB*, or *phaA*), and one of two 3-hydroxybutyryl-CoA reductases (*phaB* or *hbd*).

### Chiral Analysis of Microbially-Produced 3HV

Recombinant *E. coli* expressing the 3-hydroxyalkanoic acid pathway expressing *bktB* and either *phaB* or *hbd* as the 3-hydroxybutyryl-CoA reductase was grown at 30°C in 50 mL LB supplemented with 20 g L<sup>-1</sup> glucose and 20 mM propionate for 48 hours. The stereochemistry of the resulting 3HV in the media from these cultures was determined by methyl esterification of the 3HV present followed by chiral HPLC analysis as described in the Materials and Methods section at the end of this Chapter. These spectra were compared to a racemic 3HV standard. The results confirm that 3-hydroxyalkanoic acid pathways expressing *phaB* make a different stereoisomer of 3HV than pathways employing *hbd* (Figure 6.3). Unfortunately the lack of an enantiopure 3HV standard prevents the assignment of absolute stereochemistry (e.g. *R* or *S*) to each sample, though

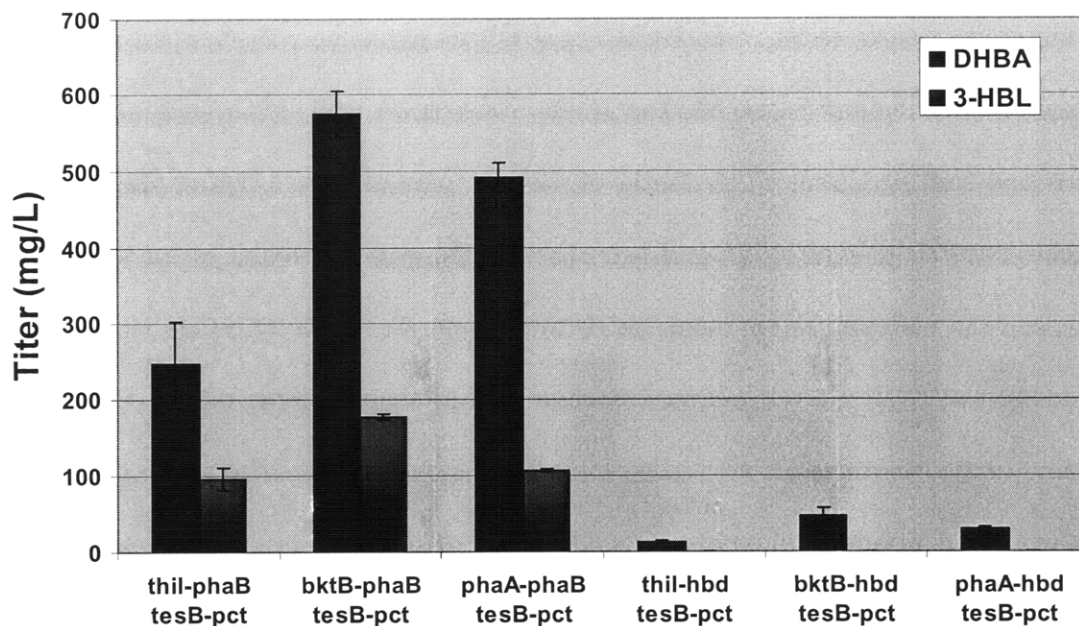
it is clear that the *phaB* and *hbd* cultures produce different stereoisomers. However, based on previous reports regarding the product stereochemistry of *phaB* and *hbd* (Lee et al., 2008; Liu et al., 2007) and the observation that (*R*)-3HB has a faster retention time relative to (*S*)-3HB (Figure 5.3), it is highly likely that here that the *phaB* pathway sample is making (*R*)-3HV while the *hbd* pathway sample produces (*S*)-3HV.



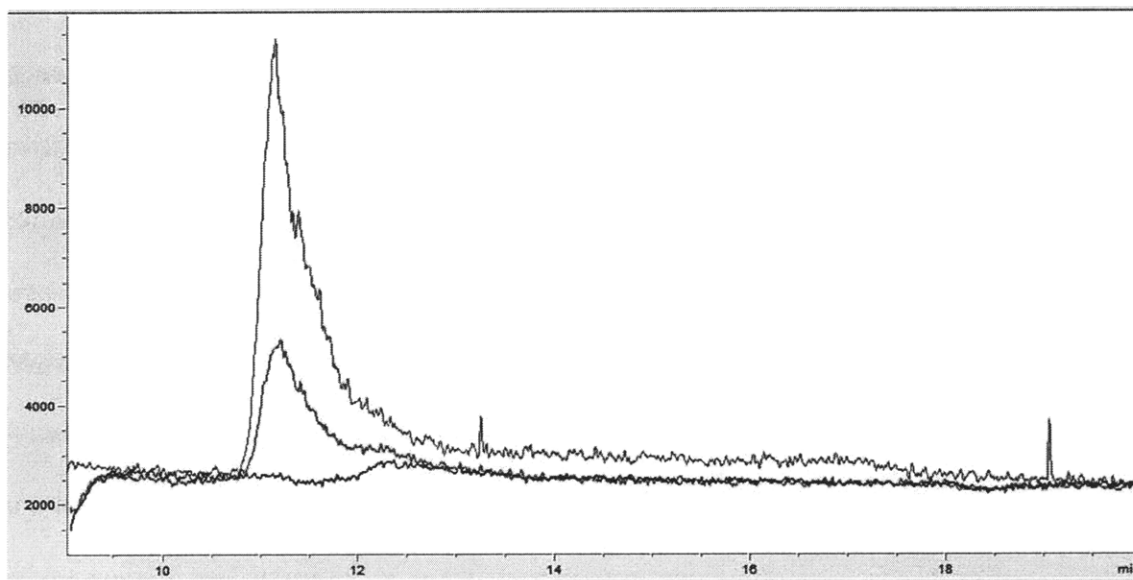
**Figure 6.3:** HPLC spectra of methyl esters of 3HV. (A) represents a racemic 3HV standard with the two stereoisomers eluting at 6.9 and 9.2 min, (B) is culture medium from *E. coli* expressing *ptb-buk*, *phaB*, and *bktB* with a peak at 6.9 min, and (C) is culture medium from *E. coli* expressing *ptb-buk*, *hbd*, *bktB*, and *tesB* with a peak at 9.2 min. The *ptb-buk* operon encodes a phosphotransbutyrase (*ptb*) and a butyryl-CoA synthetase (*buk*) from *Clostridium acetobutylicum* that functions to produce high levels of propionyl-CoA from propionate, similar to *pct*. All samples were boiled in acidic methanol to form methyl-3HV esters prior to HPLC analysis. Chiral 3HB experiments using this same chiral HPLC analysis protocol (Figure 5.3) strongly suggests that the 6.9 and 9.2 min peaks represent (*R*)-3HV and (*S*)-3HV respectively.

### **Production of DHBA and 3-HBL from Glucose and Glycolate**

For the production of DHBA and 3-HBL, neutralized glycolate was supplied along with glucose to recombinant *E. coli* MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> cells expressing one of three acetoacetyl-CoA thiolases (*thil*, *bktB*, or *phaA*), one of two 3-hydroxybutyryl-CoA reductases (*phaB* or *hbd*), *tesB*, and *pct*. These cultures were incubated for 96 hours at 30°C, and the resulting DHBA and 3HBL titers were measured by HPLC (Figure 6.4). The identity of DHBA was confirmed through LC/MS analysis (Figure 6.5). DHBA was detected in all 3-hydroxyalkanoic acid pathway gene combinations, however 3-HBL was only detected in pathways expressing *phaB*. The *bktB-phaB* combination yielded the highest DHBA and 3-HBL titers of  $573 \pm 30 \text{ mg L}^{-1}$  and  $178 \pm 4 \text{ mg L}^{-1}$  respectively. The *phaA-phaB* combination performed almost as well, yielding DHBA and 3-HBL titers of  $492 \pm 19 \text{ mg L}^{-1}$  and  $107 \pm 3 \text{ mg L}^{-1}$  respectively. The *thil-phaB* combination did not perform as well as the other *phaB* combinations, yielding only  $247 \pm 57 \text{ mg L}^{-1}$  of DHBA and  $96 \pm 14 \text{ mg L}^{-1}$  of 3-HBL. Pathway combinations using *hbd* made only DHBA, and the maximum DHBA titer of  $47 \pm 10 \text{ mg L}^{-1}$  was from *bktB-hbd*.



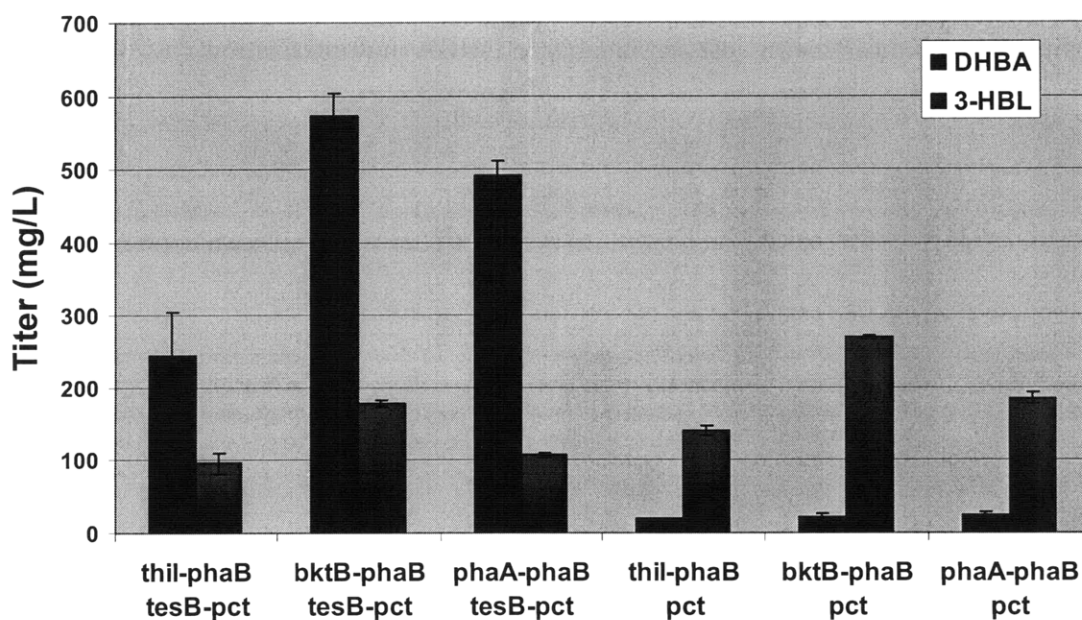
**Figure 6.4:** Production of DHBA (blue bars) and 3-HBL (red bars) by recombinant MG1655(DE3) *endA*<sup>-</sup> *recA*<sup>-</sup> *E. coli* expressing different 3-hydroxyalkanoic acid pathway genes. The specific genes used in each pathway are shown on the x-axis.



**Figure 6.5:** LC/MS analysis of DHBA produced by recombinant *E. coli*. The green line represents a DHBA standard, while the red line is DHBA produced by recombinant *E. coli* expressing *phaA*, *phaB*, *pct*, and *tesB*. The blue line is a negative control. The MS signal is an ion trace at  $m/z = 138$ , which represents the ammonium adduct of protonated DHBA.

### Effect of *tesB* Expression on DHBA and 3-HBL Production

The production of DHBA and 3-HBL was tested with and without *tesB* in 3-hydroxyalkanoic acid pathways co-expressing *phaB*, one of three acetoacetyl-CoA thiolases (*thil*, *bktB*, or *phaA*), and *pct*. This experiment was conducted similar to the one above, but now with the goal of understanding how 3-HBL is being produced by the 3-hydroxyalkanoic acid pathway. The resulting DHBA and 3-HBL titers are shown in Figure 6.6. As is seen in this figure, pathways lacking *tesB* produced on average 57% more 3-HBL, but 95% less DHBA. The *bktB-phaB-pct* combination yielded the highest 3-HBL titer of  $270 \pm 3 \text{ mg L}^{-1}$ . The performance of pathways with different acetoacetyl-CoA thiolases followed that of Figure 6.4, with *bktB* reaching the highest product titers and *thil* having the lowest titers.

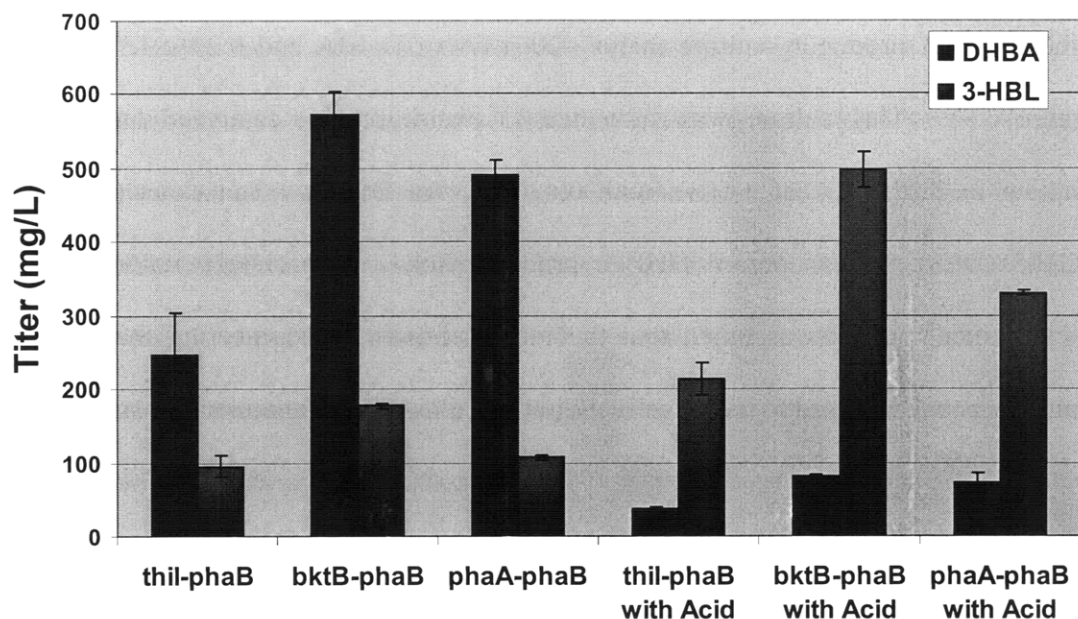


**Figure 6.6:** Production of DHBA (blue bars) and 3-HBL (red bars) by recombinant *E. coli* MG1655(DE3) *endA*<sup>-</sup> *recA*<sup>-</sup> expressing different 3-hydroxyalkanoic acid pathway genes in the presence or absence of *tesB*. The specific genes used in each pathway are shown on the x-axis.

### **Effect of Acidic Post-Treatment on DHBA and 3-HBL Titrers**

In organic chemistry, it is well documented that lactonization is acid-catalyzed (Carey, 2000). Since the *phaB* cultures shown in Figure 6.4 produced much DHBA but relatively little 3-HBL, the 96-hour supernatants from these cultures were acidified with 50 mM of hydrochloric acid, reducing the pH below 1.0. These culture supernatants were then incubated at 37°C overnight to allow for acid-catalyzed lactonization to occur. After this overnight incubation, these samples were subjected to HPLC analysis, and the resulting DHBA and 3-HBL titers are shown in Figure 6.7. The acid post-treatment of the culture supernatants improved 3-HBL titers by 171% on average, and the highest 3-HBL titer obtained was  $497 \pm 25 \text{ mg L}^{-1}$  for the *bktB-phaB* pathway combination. DHBA titers dropped by 85% on average during the process due to the lactonization of DHBA to 3-HBL. Approximately 80-85 mol% of the DHBA was converted to 3-HBL, with the remaining 15-20 mol% being lost presumably to DHBA multimerization.



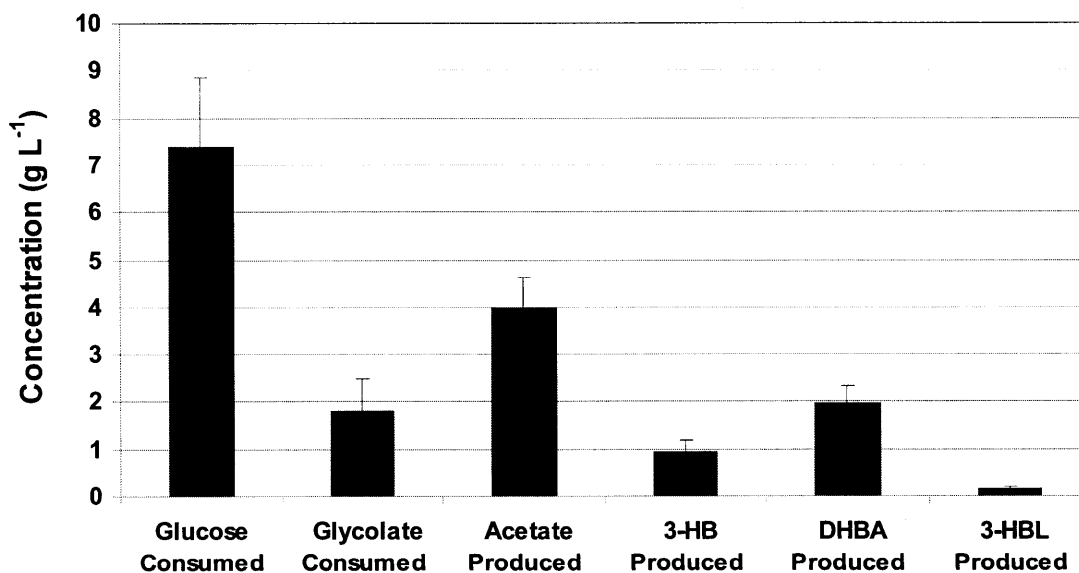


**Figure 6.7:** Production of DHBA (blue bars) and 3-HBL (red bars) by recombinant *E. coli* MG1655(DE3) *endA*<sup>-</sup> *recA*<sup>-</sup> expressing different 3-hydroxyalkanoic acid pathway genes along with *tesB* and *pct*. The right three samples represent product titers that result when the left three samples are treated with acid. The specific acetoacetyl-CoA thiolase and 3-hydroxybutyryl-CoA reductase used in each pathway are shown on the x-axis.

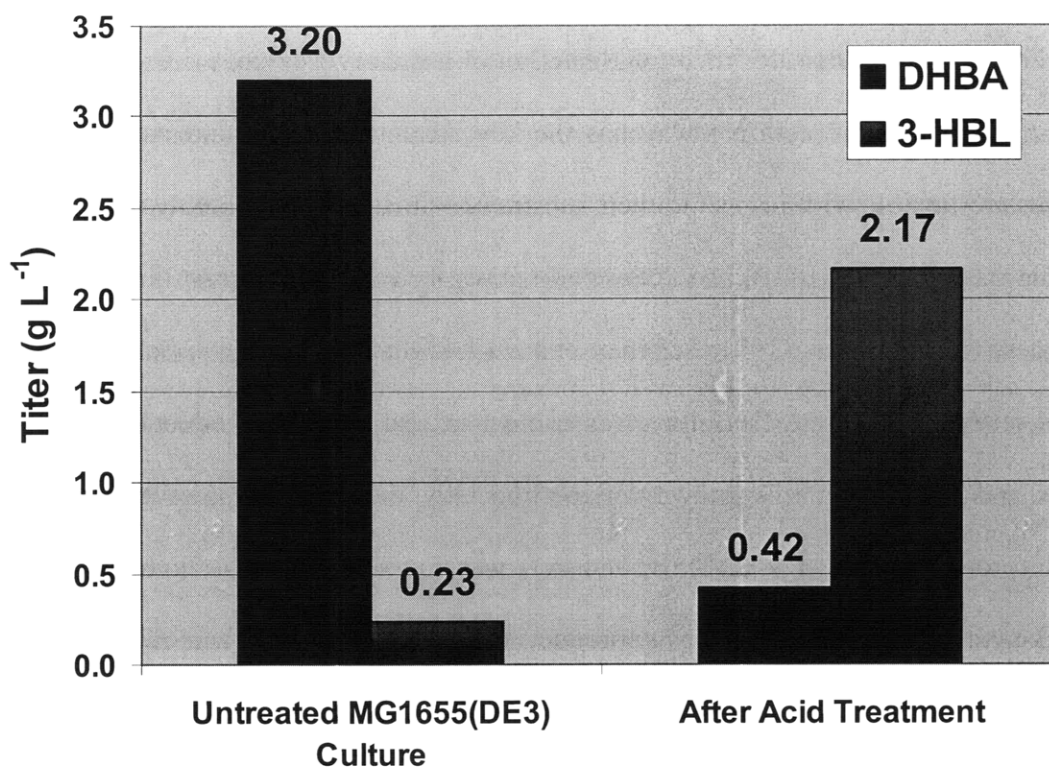
### Strain Effects on DHBA and 3-HBL Production

The 3-hydroxyalkanoic acid pathway was further tested in *E. coli* MG1655(DE3) with intact *recA* and *endA* genes to see if the titers of DHBA and 3-HBL could be further improved in a more wild-type *E. coli* strain. The results of this experiment are shown in Figures 6.8 and 6.9. In Figure 6.8, the total consumption of glucose and glycolate and the production of acetate, 3HB, DHBA, and 3-HBL for 11 independent cultures is shown. The relatively large sample size helps to counteract the natural variability in gene expression that occurs in *recA*<sup>+</sup> strains. In this experiment,  $1.99 \pm 0.36$  g L<sup>-1</sup> of DHBA was produced along with  $0.16 \pm 0.04$  g L<sup>-1</sup> of 3-HBL. This represents more than a 3-fold

improvement in DHBA production over MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> (Figure 6.4). The highest DHBA producing culture made 3.20 g L<sup>-1</sup> of DHBA and 0.23 g L<sup>-1</sup> of 3-HBL (Figure 6.9). This culture was subjected to overnight acid-catalyzed lactonization, resulting in 2.17 g L<sup>-1</sup> of 3-HBL with 0.42 g L<sup>-1</sup> of DHBA. The exact reason why MG1655(DE3) is a superior DHBA producer to MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> is undetermined, but it is assumed that the increased genetic engineering that the latter strain has been subjected to has negatively impacted its growth and metabolism.



**Figure 6.8:** Production profile of the 3-hydroxyalkanoic acid pathway in recombinant *E. coli* MG1655(DE3) expressing *bktB*, *phaB*, *pct*, and *tesB* after 96 hours. The average levels of glucose and glycolate consumed, along with the average levels of acetate, 3HB, DHBA, and 3-HBL produced are shown for 11 independent cultures, with the error bars representing the standard deviation.



**Figure 6.9:** Production of DHBA and 3-HBL in the highest-producing culture of recombinant *E. coli* MG1655(DE3) expressing *bktB*, *phaB*, *pct*, and *tesB* after 96 hours. This culture was subjected to an overnight acid treatment to catalyze lactonization, and the resulting DHBA and 3-HBL levels are shown on the right side of this figure.

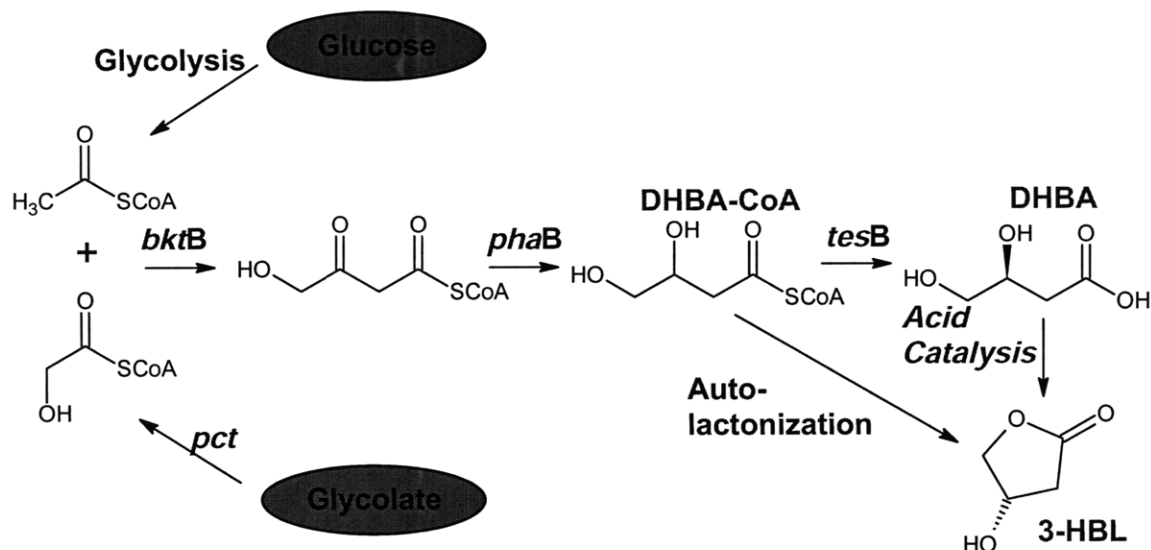
## Discussion

### **Controlling Product Profiles through Selection of Pathway Enzymes**

The 3-hydroxyalkanoic acid pathway has the key advantage that it and its constituent enzymes are flexible with respect to their substrates – the pathway is highly versatile. In particular, TesB, PhaB, BktB, and Pct are all known to work with at least three different substrates (Huisman et al., 1991; Schubert et al., 1988; Slater et al., 1998; Schweiger and Buckel, 1984). Substrate flexibility was critical in realizing the production of 3HV, DHBA, and 3-HBL from this pathway. Experimentally, the combination of the four most flexible proteins (TesB, PhaB, BktB, and Pct) was found to yield the highest titers of novel 3-hydroxyalkanoic acid pathway products (Figures 6.3, 6.4, 6.6, and 6.7).

Furthermore, the products obtained through the 3-hydroxyalkanoic acid pathway can be precisely controlled through the substrates and genes supplied to the pathway. For instance, 3HV is only produced in the presence of propionate, while DHBA and 3-HBL are only made when glycolate is added. The pathway can also be modulated to produce more DHBA or 3-HBL through the presence or absence of the thioesterase TesB. With recombinant TesB, significantly more DHBA is made ( $573 \text{ mg L}^{-1}$  for the *bktB-phaB* pathway combination) than 3-HBL ( $178 \text{ mg L}^{-1}$ ). Without TesB, more 3-HBL is made ( $270 \text{ mg L}^{-1}$  for *bktB-phaB*) than DHBA ( $21 \text{ mg L}^{-1}$ ). The small amount of DHBA formed in the absence of recombinant *tesB* may be due to the expression of genomic *tesB*, as this gene is native to *E. coli* (Huisman et al., 1991). Nonetheless, this result suggests that without a thioesterase to hydrolyze the CoA thioester of DHBA (DHBA-CoA), DHBA-CoA can self-lactonize into 3-HBL *in vivo* (Figure 6.10). While it is not

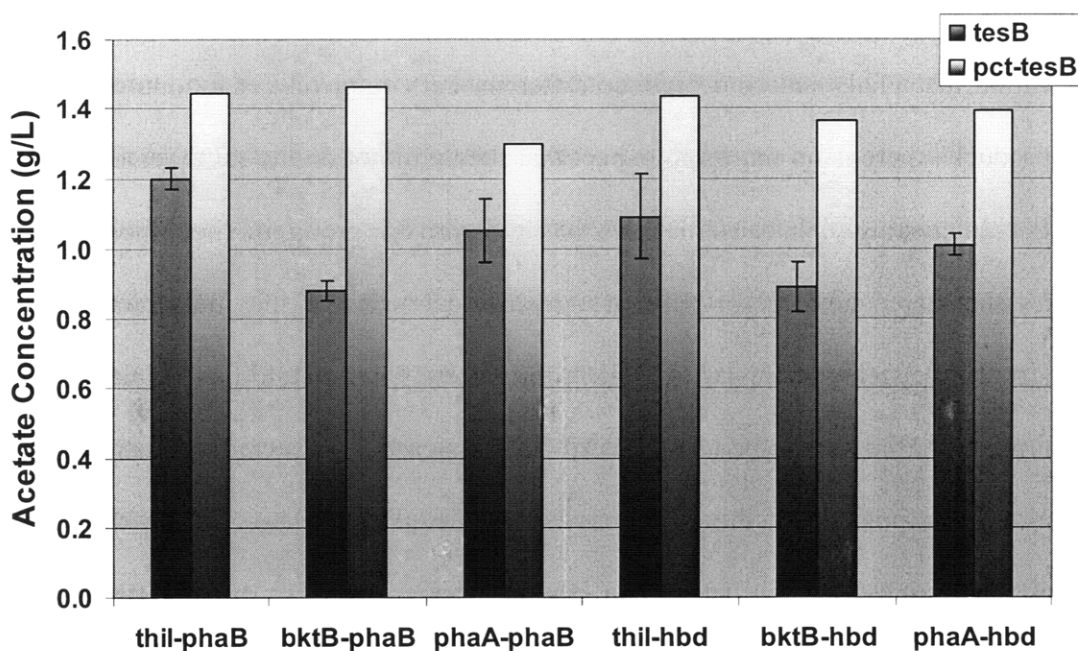
known whether there is an enzyme in *E. coli* that catalyzes this reaction or whether this reaction occurs spontaneously *in vivo*, it is clear that 3-HBL is produced from glucose and glycolate and that *tesB* plays an important role in determining the fate of the DHBA-CoA intermediate (Figure 6.6).



**Figure 6.10:** 3-Hydroxyalkanoic acid pathway scheme for DHBA and 3-HBL production from glucose and glycolate, highlighting the key branch point at DHBA-CoA. With the expression of recombinant *tesB*, more DHBA is produced by the pathway than 3-HBL (Figure 6.6), however in the absence of *tesB*, more 3-HBL is produced while little DHBA is made, suggesting that DHBA-CoA can self-lactonize into 3-HBL *in vivo*. Here the pathway is shown with *bktB* and *phaB* as the acetoacetyl-CoA thiolase and 3-hydroxybutyryl-CoA reductase, as this pair of genes was shown to yield the highest levels of DHBA and 3-HBL (Figure 6.4).

Another interesting point is that in all of the 3HV, DHBA, and 3-HBL production experiments done in this work in MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup>, very little 3HB (< 100 mg L<sup>-1</sup>) was detected (data not shown). Even in MG1655(DE3), there was twice as much DHBA produced relative to 3HB (Figure 6.9), even though 3HB is the natural substrate for the enzymes employed in the 3-hydroxyalkanoic acid pathway. 3HB is made from

the condensation of two acetyl-CoA moieties (Figure 5.1) and should theoretically be a major side product in the production of the other 3-hydroxyacids in this work since 3HB is the natural product of the action of acetoacetyl-CoA thiolases with 3-hydroxybutyryl-CoA reductases and 3HB has been produced at 2-3 g L<sup>-1</sup> with these enzymes (Chapter 5 of this thesis; Tseng et al., 2009). The only difference between the 3HB and 3-hydroxyalkanoic acid pathways is the presence of *pct* in the latter pathway. Pct is a CoA-transferase that exchanges CoA moieties between short-chain organic acids, including acetate and acetyl-CoA (Schweiger and Buckel, 1984). That very little 3HB is made in the presence of recombinant *pct* suggests that the Pct protein may be removing CoA from acetyl-CoA in the cell and transferring it onto other acids, including propionate for 3HV production and glycolate for DHBA and 3-HBL production. This CoA transfer could significantly reduce intracellular acetyl-CoA concentrations, making the second-order condensation of two acetyl-CoA molecules much less likely than the condensation of acetyl-CoA with another acyl-CoA (a first-order reaction with respect to acetyl-CoA). A byproduct of removing CoA from acetyl-CoA would be increased acetate production, and this effect was confirmed by HPLC in comparing the 3-hydroxyalkanoic acid pathway with and without *pct* (Figure 6.11). The significant increase in acetate production corroborates the hypothesis that Pct is harvesting CoA from acetyl-CoA for use in thiolating other short-chain acids within the cell. This action likely reduces acetyl-CoA levels to the point that 3HB biosynthesis is significantly inhibited.



**Figure 6.11:** Acetate production after 48 hours from the 3-hydroxyalkanoic acid pathway in *E. coli* MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> with and without *pct*. Experiments without the *pct* were performed in triplicate, while those with *pct* represent single trials.

### Controlling Stereochemistry via Choice of 3-Hydroxybutyryl-CoA Reductase

The 3-hydroxyalkanoic acid pathway's flexibility also extends to the stereochemistry of its final product - by using *phaB* or *hbd* as the 3-hydroxybutyryl-CoA reductase in the pathway, different stereochemical outcomes can be realized. It is well documented in the literature that PhaB results in (*R*)-3-hydroxyacids while Hbd results in the (*S*) stereoisomer (Lee et al., 2008; Liu et al., 2007). This phenomenon was experimentally verified both in pathways producing 3HB (Figure 5.3; Tseng et al., 2009) and 3HV (Figure 6.3). Unfortunately, the methyl esterification chiral analysis method used to analyze the stereochemistry of 3HB and 3HV does not work on DHBA because of its ability to lactonize and because of the inability of the 3-HBL lactone stereoisomers to

resolve from one another during the analysis (data not shown). However, based on the observation that PhaB and Hbd produce different and completely enantiopure 3HB and 3HV products, there is no reason to expect that they will not do the same for DHBA and 3-HBL. Interestingly, because DHBA has a hydroxyl group in the  $\delta$ -position that changes the stereochemical priority of the different parts of the molecule about its stereocenter, the stereochemistry of DHBA formed by PhaB should be (*S*)-DHBA while that formed by Hbd should be (*R*)-DHBA. This assignment of stereochemistry is based on the absolute position of the  $\gamma$ -hydroxyl group in DHBA relative to the carboxylate group. Since PhaB is much better than Hbd at producing DHBA and 3-HBL in the 3-hydroxyalkanoic acid pathway (Figure 6.4), the pathway is better suited for the production of (*S*)-DHBA and (*S*)-3-HBL than the (*R*) stereoisomers. Fortunately, the (*S*) stereoisomer is predominately used in the production of pharmaceuticals and high-value compounds (Lee and Park, 2009). The screening of *hbd* homologs could identify a 3-hydroxybutyryl-CoA reductase with identical stereochemical preference but with an increased substrate range for the production of (*R*)-DHBA and (*R*)-3-HBL.

### **Conclusions**

In this work, 3HV (0.54 g L<sup>-1</sup>), DHBA (3.2 g L<sup>-1</sup>), and 3-HBL (2.2 g L<sup>-1</sup>) were produced using the 3-hydroxyalkanoic acid pathway (Figure 6.1) in *E. coli*. These compounds have applications ranging from chiral building blocks to polymeric materials to high-value pharmaceuticals. This work represents one of the first reports on 3HV biosynthesis in *E. coli*, and is the first report on the complete biological production of DHBA and 3-HBL from inexpensive sugar and sugar-derived substrates. Further screening of



additional 3-hydroxyalkanoic acid pathway homologs should provide a rich resource of pathway combinations for the production of other high value 3-hydroxyacids such as 3-hydroxypropionate from formyl-CoA and acetyl-CoA. Molecular analogs of DHBA and 3-HBL could be made using the 3-hydroxyalkanoic acid pathway as well, such as 3,4-dihydroxyvalerate and 4-methyl-3-HBL from the condensation of lactyl-CoA and acetyl-CoA, potentially leading to new intermediates for more effective pharmaceuticals.

Product titers from the 3-hydroxyalkanoic acid pathway could be further enhanced by additional optimization of culture conditions and host strain. The screening of homologous pathway enzymes could improve titers in addition to broadening the substrate range of the pathway. One promising route for increasing productivity is to engineer *E. coli* to produce pathway substrates “in-house.” For instance, glycolate could be made intracellularly from glyoxylate through the action of the *ycaW* glyoxylate reductase gene (Nuñez et al., 2001). Intracellular production of pathway substrates would eliminate the need for substrate transport across the cellular envelope, accelerating production. Intracellular production of substrates would also allow the feeding of just glucose to *E. coli* for product formation, simplifying the bioprocess while lowering material costs.

## **Materials and Methods**

### **Strains and Chemicals**

*E. coli* DH10B (Invitrogen, Carlsbad, CA) and ElectroTen-Blue (Stratagene, La Jolla, CA) were used for transformation of cloning reactions and propagation of all plasmids.

*E. coli* MG1655(DE3) was constructed from *E. coli* MG1655 using a  $\lambda$ DE3 Lysogenization Kit (Novagen, Darmstadt, Germany). The *recA* and *endA* genes in *E. coli* MG1655(DE3) were then knocked out by the method of Datsenko and Wanner (2000), to yield *E. coli* MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> (Tseng et al., 2009). Both *E. coli* MG1655(DE3) and *E. coli* MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> were used for the production of 3HV, DHBA, and 3-HBL. Luria-Bertani (LB) medium, D-glucose, propionic acid, and glycolic acid were purchased from BD Biosciences (Sparks, MD), Mallinckrodt Chemicals (Phillipsburg, NJ), Sigma-Aldrich (St. Louis, MO), and MP Biomedicals (Solon, OH), respectively. Unless otherwise noted, all chemicals were purchased at the highest grade available.

### **Plasmids and Primers**

Genes derived from *C. acetobutylicum* ATCC 824 (*thil* and *hbd*), *R. eutropha* H16 (*phaA*, *bktB*, and *phaB*), *E. coli* K-12 (*tesB*), and *M. elsdenii* (*pct*) were obtained by polymerase chain reaction (PCR) using genomic DNA (gDNA) templates. All gDNAs were prepared using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI). Custom oligonucleotides (primers) were purchased for all PCR amplifications (Sigma-Genosys, St. Louis, MO) and are shown in Table 6.2. In all cases, HotStar HiFidelity Polymerase (Qiagen, Valencia, CA) was used for DNA amplification. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA). DNA digestions, ligations, and transformations were performed according to standard procedures (Sambrook and Russell, 2001).

Two co-replicable vectors, pETDuet-1 and pCDFDuet-1 (Novagen, Darmstadt, Germany), were used for construction of the 3-hydroxyalkanoic acid pathways. Both Duet vectors contain two multiple cloning sites (MCS), each of which is preceded by a *T7lac* promoter and a ribosome binding site (RBS), affording high-level expression of each individual gene. For cloning the pathway genes into these vectors, PCR products were made using 5' and 3' primers with desired restriction sites incorporated within them. These sites, used for cloning the genes, are underlined in Table 6.2. These PCR products were digested with restriction enzymes corresponding to the restriction site incorporated into them by their respective primers and ligated directly into similarly digested Duet vector. Ligation reactions using pETDuet-1 as the vector were transformed into DH10B *E. coli*, while ligations using pCDFDuet-1 were transformed into ElectroTen-Blue *E. coli*. One of three acetoacetyl-CoA thiolases (*thil*, *bktB*, and *phaA*) and one of two 3-hydroxybutyryl-CoA reductases (*phaB* and *hbd*) were cloned into pETDuet-1, resulting in six pETDuet-based plasmids. The pCDFDuet-based plasmids contained either *pct* and *tesB* or just *pct* alone. Once all plasmids were constructed, one pETDuet-based plasmid and one pCDFDuet-based plasmid were co-transformed into *E.coli* MG1655(DE3) *recA*<sup>-</sup> *endA*<sup>-</sup> or *E.coli* MG1655(DE3) to create a complete 3-hydroxyalkanoic acid pathway-expressing production strain.

**Table 6.2:** List of DNA oligonucleotide primers used in the cloning of genes for the 3-hydroxyalkanoic acid project. Restriction sites used for cloning are underlined. Primer names correspond to the name of the gene that the primer amplifies, whether the primer is the forward primer (FP) or reverse primer (RP) of that gene, and the restriction site incorporated into the primer sequence for cloning.

Primer	Sequence 5'→3'	Source
bktB-FP-NcoI	GAC <u>ACCATGGGC</u> ATGACGCGTGAAGTGGTAG	Sigma-Genosys
bktB-RP-EcoRI	GACAGAATTCTCAGATACGCTCGAAGATGG	Sigma-Genosys
hbd-FP-NdeI	TATCCATATGAAAAAGGTATGTGTTATAGGTGC	Sigma-Genosys
hbd-RP-XhoI	GACACTCGAGTTATTTTGAATAATCGTAGAAACCTTTTC	Sigma-Genosys
phaA-FP-NcoI	GAC <u>ACCATGGGC</u> ATGACTGACGTTGTCATCGTATC	Sigma-Genosys
phaA-RP-EcoRI	GACAGAATTCTTATTTGCGCTCGACTGC	Sigma-Genosys
phaB-FP-NdeI	GACACATATGACTCAGCGCATTGC	Sigma-Genosys
phaB-RP-XhoI	GACACTCGAGTCAGCCCATGTGCAGG	Sigma-Genosys
pct-FP-EcoRI	GACAGAATTCATGAGAAAAGTAGAAATCATTACAGCTG	Sigma-Genosys
pct-RP-PstI	GACACTGCAGTTATTTTTTTCAGTCCCATGGG	Sigma-Genosys
tesB-FP-NdeI	GACACATATGAGTCAGGCGCTAAAAAAT	Sigma-Genosys
tesB-RP-XhoI	GACACTCGAGTTAATTGTGAATTACGCATCACC	Sigma-Genosys
thil-FP-NcoI	GAC <u>ACCATGGGC</u> ATGAGAGATGTAGTAATAGTAAGTGC TGTAAGA	Sigma-Genosys
thil-RP-EcoRI	GACAGAATTCTATTTAGTCTCTTTCAACTACGAGAGC	Sigma-Genosys

### Culturing Conditions

Recombinant *E. coli* harboring one of the acetoacetyl-CoA thiolases (*thil*, *bktB*, or *phaA*) and one of two 3-hydroxybutyryl-CoA reductases (*phaB* or *hbd*) on the pETDuet-1 vector and *tesB* and/or *pct* on the pCDFDuet-1 vector were cultured in 50 mL LB supplemented with 1% glucose in a 250 mL shake flask at 30°C unless otherwise indicated in the text. Ampicillin (50 mg L<sup>-1</sup>) and streptomycin (50 mg L<sup>-1</sup>) were added to provide selective pressure for plasmid maintenance. Additional supplemental compounds, such as neutralized propionate or glycolate, were added to cultures where indicated in the text from 2.0M stock solutions. Cultures were inoculated to an initial optical density at 600 nm (OD<sub>600</sub>) of 0.05. Once the cells reached mid-exponential phase (OD<sub>600</sub> ~ 0.5), 50 µL of 1.0 M IPTG and 1 mL of 2.0 M neutralized propionate or glycolate was added to the cultures for induction of gene expression and to provide the

substrates needed for the 3-hydroxyalkanoic acid pathway. 1 mL of culture was withdrawn daily for HPLC analysis.

### **Methyl Esterification of 3HV**

50 mL of 3HV-containing LB microbial culture was centrifuged at 6000 x *g* for 10 min and the supernatant was transferred to an evaporation dish for evaporating overnight in a 50°C oven, resulting in a brown residue. To this residue, 2 mL of acidic methanol (4:1 methanol:concentrated HCl) was added and the dish was shaken for two minutes to remove 3HV from the residue. The methanol solution was transferred into a test tube, sealed, and heated at 100°C for 3 hours. After heating, the solution was allowed to cool to room temperature. After cooling, 200 µL of the methanol solution was mixed with 1 mL of pure isopropanol in a 1.7 mL tube, vortexed, and allowed to stand for 5 minutes. Precipitate in this sample was removed by centrifugation at 13,200 x *g* for 10 minutes. The supernatant was taken for chiral HPLC analysis.

### **HPLC Analyses**

Achiral HPLC analysis was performed on an Agilent 1100 Series instrument equipped with an Aminex HPX-87H column (0.7 cm x 30 cm) purchased from Bio-Rad Laboratories (Hercules, CA). For all achiral analyses, the mobile phase was 5 mM sulfuric acid, 5 µL of culture supernatant was injected for the analyses, and products were quantified using a refractive index detector. For the achiral analysis of 3HV, the column was operated at a temperature of 35.0 °C. The flowrate started at 0.55 mL min<sup>-1</sup> and was linearly ramped up to 0.8 mL min<sup>-1</sup> for the first 12 minutes and maintained at 0.8 mL

min<sup>-1</sup> for an additional 8 minutes. 3HV purchased from Epsilon Chimie (Brest, France) was used as a standard and had a retention time of approximately 14.5 min. For the analysis of DHBA and 3-HBL, the column was operated at a temperature of 40.0 °C and a constant flowrate of 0.75 mL min<sup>-1</sup>. 3-HBL purchased from Sigma-Aldrich (St. Louis, MO) was used as a standard, and DHBA was prepared from 3-HBL by saponification with 10N sodium hydroxide and used as a standard. The retention times for DHBA and 3-HBL were approximately 8.9 min and 13.7 min respectively.

Chiral HPLC analysis was performed on an Agilent 1100 Series instrument equipped with a Chiralcel OD-H column (0.46 cm  $\phi$  x 25 cm) purchased from Daicel Chemical Industries (West Chester, PA). The column temperature was maintained at 40°C. Methyl-3HV was detected on a diode array detector at 210 nm. The mobile phase was 9:1 *n*-hexane:isopropanol and the flow rate through the column was 0.7 mL min<sup>-1</sup>. Racemic 3HV purchased from Epsilon Chimie (Brest, France) was boiled in acidic methanol as described above to form racemic methyl-3HV, which was used as a standard. The retention times for the two enantiomers were 6.86 and 9.24 min. Prior chiral HPLC analysis with 3HB suggests that the earlier retention time is (*R*)-3HV, while the latter is (*S*)-3HV.

Mass spectrometry (MS) analysis was performed using an Agilent 6120 Quadrupole LC/MS unit operated in positive ion electrospray mode. The LC mobile phase was 25 mM ammonium formate, pH 2, and the flowrate was 0.6 mL min<sup>-1</sup>. The unit was operated downstream of the liquid chromatography column described above under the

same chromatography conditions. Samples were vaporized at 150°C and nitrogen at 300°C, 60 psig, and 5.0 L min<sup>-1</sup> was used as the carrier gas for MS analysis. The ion detector was calibrated to scan for cations at an *m/z* of 138, representing the ammonium adduct of DHBA.

## Chapter 7: Conclusions and Recommendations

### Summary of Products and Titters

In this thesis, metabolic routes were developed for the production of hydroxyacids and their lactones in two different microbial systems. The first system described in Chapters 3 and 4 was for the production of 3HV, 4HV, and 4VL from levulinate in *P. putida*. In this work, the renewable carbon source levulinate was reduced by enzymes native to *P. putida* to 4HV and 3HV, and the inclusion of a non-cytosolic lactonase enzyme allowed for the production of 4VL under acidic conditions. This system allowed for the high-titer production of five-carbon hydroxyacids and their lactones from feasible, renewable substrates. The second system discussed in Chapters 5 and 6 depicted the production of various  $\gamma$ -hydroxyacids and lactones through the 3-hydroxyalkanoic acid pathway. This pathway had the advantage of being flexible with regard to substrate and stereochemical specificity, allowing for the production of a wide range of products of different and controlled chiralities. 3HB, 3HV, DHBA, and 3-HBL were all synthesized using this platform. The different products from each of these systems and their respective maximum titers obtained are shown in Table 7.1 below.



**Table 7.1:** Summary of products and titers of hydroxyacids and lactones produced in this thesis. Stereochemistry is indicated where it is known.

<b><i>P. putida</i> Hydroxyvalerate System (Chapters 3-4)</b>	
<b>Product</b>	<b>Maximum Titer</b>
3HV	5.3 g L <sup>-1</sup>
4HV	27.1 g L <sup>-1</sup>
4VL	8.2 g L <sup>-1</sup>
<b><i>E. coli</i> 3-Hydroxyalkanoic Acid Pathway System (Chapters 5-6)</b>	
<b>Product</b>	<b>Maximum Titer</b>
( <i>R</i> )-3HB	2.9 g L <sup>-1</sup>
( <i>S</i> )-3HB	2.1 g L <sup>-1</sup>
( <i>R</i> )-3HV	537 mg L <sup>-1</sup>
( <i>S</i> )-3HV	522 mg L <sup>-1</sup>
DHBA	3.2 g L <sup>-1</sup>
3-HBL	2.2 g L <sup>-1</sup>

### **Experimental Observations of Key Pathway Design Parameters**

In Chapter 2 of this thesis, there was an extensive discussion on pathway design, with emphasis on theoretical or database tools available to aid in this endeavor. Chapters 3-6 however provide several important experimental design parameters or “rules-of-thumb” for pathway design as well. Chief among these is the strategic choice of enzymes and enzyme homologs for metabolic pathways. The 3-hydroxyalkanoic acid pathway illustrates this point extensively. By examining multiple homologs of acetoacetyl-CoA thiolases and 3-hydroxybutyryl-CoA reductases, the array of products from the pathway was broadened. For instance, Figure 6.2 shows that only half of the combinations of acetoacetyl-CoA thiolases and 3-hydroxybutyryl-CoA reductases were able to produce 3HV. Had not multiple enzymes been screened, 3HV might not have been detected, and had not different 3-hydroxybutyryl-CoA reductases been screened, the production of different product enantiomers would not have been realized (Figures 5.3 and 6.3). This

observation indicates that the availability of homologous enzymes is a key parameter in the design and construction of novel metabolic pathways.

Another key design “rule-of-thumb” was found to be using enzymes with broad substrate specificities in building unnatural pathways. Completely non-natural biosynthetic products like DHBA and 3-HBL were only possible when several enzymes with wide substrate ranges were assembled together. Figure 6.4 illustrates this point well. Here, PhaB was shown to give much higher DHBA and 3-HBL titers than Hbd. PhaB was cloned from the polyhydroxyalkanoate synthesis pathway in *R. eutropha*, a pathway that is designed to store a wide range of hydroxyacids as a carbon-storage polymer (Schubert et al., 1988). Because this polymer is used for energy, the cell is not particularly picky about the nature of the carbon stored (otherwise it would not store as much carbon as it could, resulting in an evolutionary disadvantage), and the PhaB enzyme consequently has very loose substrate specificity. Hbd on the other hand evolved as part of a very specific butyrate production pathway designed as a sink for electrons generated during the anaerobic *C. acetobutylicum* fermentations on glucose (Boynton et al., 1996). Because there is no evolutionary advantage for Hbd to have broad substrate specificity, there has been no pressure for it to perform catalysis on anything other than 3-ketobutyryl-CoA. Consequently Hbd does not tolerate the  $\delta$ -hydroxyl group in 3-keto-4-hydroxybutyryl-CoA (a precursor to DHBA) as well as PhaB and DHBA titers are much lower with Hbd than PhaB. This substrate specificity effect on product titers is also seen amongst the acetoacetyl-CoA reductases, with BktB and PhaA having much higher titers than Thil (Figure 6.4), again due to the same evolutionary pressure argument on BktB and PhaA

(also used for carbon storage in *R. eutropha*; Schubert et al., 1988; Slater et al., 1998) versus the lack of that pressure on Thil (which is also used in the butyrate pathway in *C. acetobutylicum*; Stim-Herndon et al., 1995). Having broad substrate range “auxiliary” enzymes for the 3-hydroxyalkanoic acid pathway – Pct and TesB – also made 3HV, DHBA, and 3-HBL production possible (Schweiger and Buckel, 1984; Huisman et al., 1991). By examining the known substrate tolerances of candidate enzymes for a pathway and understanding those substrate tolerances in the evolutionary context of the enzymes’ native functions, one can identify those enzymes that are most likely to accept non-natural substrates in *de novo* biosynthetic pathways for the production of novel products.

Another interesting observation is that sometimes simply expressing an enzyme that catalyzes a known reaction is not enough to achieve product. For instance, it was known that 4HV was a substrate for PON1 (Draganov et al., 2005), yet simply expressing PON1 does not result in substantial 4VL accumulation (Figures 4.3 - 4.5). Enzymes must be expressed in the proper context to realize their true catalytic potential. Intracellularly, PON1 could not produce significant quantities of 4VL due to the unfavorable equilibrium between 4HV and 4VL. At an acidic pH, however, PON1 can better catalyze 4HV lactonization, because at reduced pH values the equilibrium between 4HV and 4VL shifted more towards 4VL production. Because the cytosol pH is slightly basic (Wilks and Slonczewski, 2007), PON1 had to be localized outside of the cell to take advantage of a more acidic medium pH for catalysis. Just as not all enzymes in nature are cytosolic, all recombinantly-expressed enzymes need not be cytosolic as well. In this particular case, relocating the PON1 lactonase outside of the cytosol where the conditions under

which it performed catalysis could be controlled allowed for an order of magnitude improvement in lactone titer (Figure 4.5). This strategy is generalizable to bottleneck enzymes in other metabolic pathways that do not perform well under cytosolic conditions.

### **Analysis of Lactonization Strategies**

Lactonization in aqueous biological systems is challenging for two reasons: (1) water hydrolyzes lactone rings and (2) lactonization generally does not occur without an acid catalyst (Carey, 2000). Aqueous culture media has water concentrations in excess of 50 molar, strongly encouraging the hydrolysis of lactones while suppressing the formation of lactone rings. Ironically, functional group contribution theory calculations show that lactonization should have a standard free energy change of  $-2.7 \text{ kcal mol}^{-1}$ , indicating that lactonization under standard thermodynamic conditions is spontaneous (Mavrovouniotis, 1991). Unfortunately, aqueous solution is anything but “standard thermodynamic conditions”, and the high concentration of water drives lactone equilibrium towards the hydrolyzed state. Since microbial cells must be cultured in aqueous solution, removing the water during the culturing phase is not an option. Removing the water after the culture is done would result in lactonization, but the energy cost for evaporating the water would be very high. Enzymes like PON1 exist that can catalyze the lactonization reaction (Aharoni et al., 2004), however as enzymes are catalysts, they only speed the reaction along to equilibrium and cannot produce more lactone than would otherwise be present at equilibrium. In this case, that equilibrium lies far in favor of lactone

hydrolysis, and in the case of the *P. putida* 4VL lactone production system described in Chapter 4, the equilibrium was approximately 4 mol% lactone (Figures 4.3 and 4.4).

One strategy for overcoming this thermodynamic reality is to drive the lactonization reaction forward with acid. Acid accelerates the lactonization reaction by counteracting the net alkaline generation that occurs when a deprotonated carboxylate group condenses with a hydroxyl group. Unfortunately inside a microbial cell where biocatalysis normally occurs, the pH is maintained at a relatively high value of approximately 7.5 (Wilks and Slonczewski, 2007). In this thesis (Chapter 4), this problem was overcome by localizing a lactonase outside of the cytosol, where the enzyme is exposed to the medium pH. This allowed for control of the pH of lactonization by manipulation of the medium pH, and consequently by lowering the medium pH, lactonization was substantially enhanced (Figure 4.4). However, the medium pH cannot be reduced below the threshold at which the cells would die due to acid toxicity. In the *P. putida* system where this technique was employed, that threshold was approximately pH 6.0, where the equilibrium amount of lactone (4VL) was approximately 30-35 mol% (Figure 4.6). While this strategy is a solid improvement over the original 4 mol% lactone, this approach still leaves 65-70 mol% of the lactone in the hydrolyzed state.

A much better approach was to circumvent the lactone-hydroxyacid equilibrium entirely by starting with a higher-energy substrate. In the case of the 3-hydroxyalkanoic acid pathway biosynthesis of 3-HBL (Chapter 6), 3-HBL was made spontaneously from DHBA-CoA through a type of auto-lactonization reaction (Figure 6.10). DHBA-CoA

has substantially more energy than DHBA due to the activation of its carboxylate group by a CoA thioester (Mavrovouniotis, 1991). This additional energy helps drive lactonization to completion by using the CoA moiety as a leaving group. When *tesB* was removed from the 3-hydroxyalkanoic acid pathway to allow auto-lactonization of DHBA-CoA to occur, 89-94 mol% of the product obtained was in the lactone (3-HBL) form (Figure 6.6). Furthermore these numbers are a conservative estimate as genomic *tesB* (which hydrolyzes DHBA-CoA into DHBA) was still active in this experiment. This near-complete degree of lactonization indicates that using high-energy substrates for lactonization is a highly effective option, though it may be difficult to find systems where such a strategy is possible. For instance, 4HV-CoA in theory can lactonize, however this spontaneous lactonization was never observed experimentally.

A more general strategy for lactonization is the post-treatment of culture broths containing large amount of hydroxyacid precursor with copious amounts of acid, driving the lactonization reaction forward chemically. When this method was employed for 3-HBL production from DHBA, lactone fractions were improved from 20-31 mol% to 84-88 mol%, though admittedly with some loss of DHBA to multimerization (Figures 6.7 and 6.9). While this method essentially terminates the microbial cell culture and requires the purchase and addition of acid, it is a quick and simple culture post-treatment that is broadly applicable to any microbial production system so long as that system can produce the hydroxyacid precursor.

## **Outlook and Recommendations**

The *P. putida* 4VL production system (Chapters 3 and 4) has already shown tremendous promise in producing 4VL, a potentially valuable next generation biofuel and bioprocess plant chemical (Horváth et al., 2007), from levulinate, a renewable biomass-derived feedstock listed as one of the top ten value-added chemicals derived from biomass (Werpy and Peterson, 2004). This process, the first of its kind, allows for the utilization of levulinate made from the processing of agricultural waste. The hydroxyvalerates produced as an intermediate in 4VL production can be polymerized into “green” biopolyesters with favorable physical properties (Gorenflo et al., 2001). Shake-flask and bioreactor scale experiments done in this thesis have shown that high titers of hydroxyvalerates (Figures 3.1 and 3.2) and 4VL (Figures 4.5 and 4.6) are possible even without any strain or enzyme engineering. Future efforts to improve upon the titers of 3HV, 4HV, and 4VL reported in this work should focus first on identifying the enzyme(s) in *P. putida* responsible for levulinate metabolism so that the pathway can be reconstituted in a more accessible microorganism such as *E. coli*. This would allow for metabolic and enzyme engineering of these enzymes in order to improve pathway productivity. Alternatively, *P. putida* could be engineered directly if the levulinate metabolism pathway was known.

The integrated bioprocessing effect for 4VL production in *P. putida* in acidic media, while making 4VL production possible, only converts 30-35 mol% of 4HV into 4VL. This means that 4VL titers could be improved by up to three-fold simply by finding a way to achieve complete lactonization. One quick and easy way of doing this is by

acidifying 4HV-rich culture media to  $\text{pH} < 1$ , however the addition of so much acid could render the process economically unattractive. A viable alternative to this would be to operate this bioprocess in two-phases. If the 4VL produced in the aqueous phase could be sequestered away from the PON1 lactonase, the equilibrium between 4HV and 4VL would shift even further towards 4VL production through a Le Chatelier-type effect. As of the writing of this thesis this experiment is being tested using two solid phase polymeric resins: Dowex™ Optipore™ L493 and Dowex™ Optipore™ SD-2 (Dow Chemical Company (Midland, MI)). Preliminary results indicate that the equilibrium level of lactonization at  $\text{pH} 6$  in the presence of these two resins increases from approximately 30 mol% to 55 mol% (data not shown). This result shows that implementation of a two-phase system has excellent potential for driving lactonization forward while recovering and purifying the 4VL product.

The 3-hydroxyalkanoic acid pathway also shows tremendous promise as a general platform for the biosynthesis of a wide range of 3-hydroxyacids and their derivatives. While this thesis focused on the production of 3HB, 3HV, DHBA, and 3-HBL, it is certainly possible to produce other products through the condensation of acetyl-CoA with different acyl-CoA molecules. The main challenges in realizing broader production of 3-hydroxyacids with this pathway are thiolating the necessary organic acids so that they may be substrates for condensation and finding additional acetoacetyl-CoA thiolases and 3-hydroxybutyryl-CoA reductases capable of working with different and ever more divergent substrates. In this work, only one *pct* homolog was cloned and screened, and fortunately it was tolerant of a wide range of substrates. Examining other *pct* homologs,



such as one from *Clostridium propionicum* (Selmer et al., 2002) may lead to new substrates and new discoveries with the 3-hydroxyalkanoic acid pathway. Similarly, the screening of new acetoacetyl-CoA thiolase and 3-hydroxybutyryl-CoA reductase homologs should enable the condensation and reduction of an even more diverse array of pathway intermediates. Finally as with the *P. putida* 4VL work, enzyme and strain engineering is always a viable option for further improving pathway productivity and achieving higher titers. Scale-up of the 3-hydroxyalkanoic acid pathway in a bioreactor is also another option, as this was not explored in this work. Given the value of 3-HBL (currently \$20-50 per gram depending on quantity and enantiopurity) and its large market as a substrate in the production of many high-value compounds and pharmaceuticals (Lee and Park, 2009), efforts should be focused in bringing 3-HBL titers to commercially viable levels.

### **Concluding Thoughts**

Fundamentally this thesis was an exercise in metabolic exploration. As stated in Chapter 2, there are many enzymes out there – both known and unknown – and there are many potential and productive combinations of these enzymes that result in the creation of new and/or useful chemical compounds. Currently this metabolic terrain remains largely unmapped, and opportunities for revolutionizing how chemicals are made are out in that terrain somewhere. The work done here represents only one homestead in this vast metabolic wilderness. Nonetheless, the metabolic wilderness, the author hopes, is one step closer to being tamed.

Even in the much more mature field of chemistry, new reactions and processes are being discovered daily. Every month or so one can read through the many chemistry journals that are published – each one full of new discoveries and new knowledge. Today the chemistry landscape is much better colonized than the biochemical one. With chemistry, whole broad classes of reactions have been worked out to obtain desired products. For instance, reducing agents such as lithium aluminum hydride have been found to react with and reduce many different compounds on demand. This kind of reaction is well worked out in chemistry, so much so that the chemist can be confident that the reaction has an excellent chance of working on completely untested substrates. In the biochemical production of compounds, figuring out how to develop and construct a pathway for the production of a desired chemical is a daunting and laborious process. Enzymes are all too often “black-box catalysts” that, while powerful, are poorly understood. Biotechnology is not yet developed to the point where enzyme behavior can be predicted when enzymes are challenged by non-natural substrates in novel metabolic pathways.

The development of the *P. putida* hydroxyvalerate and the *E. coli* 3-hydroxyalkanoic acid pathways has this author thinking that the future and promise of metabolic pathway engineering lies not in the construction of individual pathways but in the creation of whole biocatalytic platforms. The building of individual pathways, while helping to map the metabolic wilderness, only sheds light on the smallest of areas. Just as in chemistry, there is a great need to develop biocatalysts and pathways for the production of whole classes of compounds rather than single molecular targets. How nice it would be to pull a tube of reductase enzyme out of a freezer and to be able to use that enzyme on a vast

array of substrates with the confidence that this enzyme will catalyze their reduction. The 3-hydroxyalkanoic acid pathway (Chapters 5 and 6) was a small example of one broad biochemical system. Still, this pathway has its limits – the further away the pathway substrate is from acetyl-CoA the lower the product titer becomes. However, the 3-hydroxyalkanoic acid pathway is a seed – a starting point from which a more generalized, broader reaching 3-hydroxyalkanoic acid pathway can be engineered. Through the examination of enzyme homologs and the power of protein engineering, this seed can indeed sprout into a plant with many leaves and branches – with many different pathways substrates and valuable products. The ReBiT database discussed in Chapter 2 depicts the construction of pathways based on changes in functional group structure, but it disregards enzyme substrate specificity to simplify the problem of computational pathway design. Yet ReBiT highlights the need for enzymes that are not picky with their substrates. Just as lithium aluminium hydride is not picky with what it reacts with, so too is there a need for enzymes that can be chosen for a reaction only on the chemical reactions they catalyze and not the substrates they prefer.

Beyond building generalized metabolic pathways, there is a need for constructs, systems, or otherwise “bags of tricks” for those difficult systems that do not lend themselves to biocatalysis. The integrated bioprocessing work (Chapter 4) was one such broadly applicable system. Other engineered systems, such as the recent metabolic pathway scaffolding work (Dueber et al., 2009) have been developed to solve specific challenges in metabolic pathway construction. The 4VL production work made clear that simply co-expressing enzymes in a pathway is often not enough to truly realize the potential of that

pathway. Enzymes must be expressed in context – in the right location, in the right amount, and at the right time – in order to maximize their effect and impact on product titers. The key message in integrated bioprocessing is to integrate the pathway into the context and environment of its host. This is not simply host strain engineering, but true integration of the pathway through the use of tools that weld a recombinant pathway to the cell in a manner that maximizes pathway productivity. This is not to say that host engineering is not important – indeed host engineering is one very critical component of the weld. But to meet the challenge of generalized biocatalysis, not only must pathways be composed correctly and the host engineered for optimum productivity, but strategies and tools must be developed and applied to mesh a recombinant pathway with its host.

Biotechnology is the clear choice for clean, renewable chemical production. From medicines to fuels to materials, the human world is made of chemicals, and chemicals dramatically enhance the human quality of life. In today's world, oil is the source of most chemicals. Through oil wealth is made, national agendas are set, and humanity lives in relative comfort. Oil however is not eternal, and this fact is the primary driver of innovation within biotechnology. The metabolic wilderness is still thick, and most biochemical pathways cannot currently compete with their oil-derived counterparts, but ultimately the leap to renewable resources must be made. Ultimately, the metabolic wilderness will be colonized and tamed, the only question is when.

If there was a mantra for this thesis, it would be “Where there's an enzyme, there's a way.” The author is a firm believer that any compound can be made through

biocatalysis, though finding the right way remains quite the challenge. Enzymes are truly remarkable catalysts, yet so little is known about how they do what they do. Today, science struggles to understand native enzymatic function – predicting unnatural enzyme function is an educated guess at best. If no way exists to make a desired product, this only means that efforts should be focused on making the enzyme. While the wilderness is vast, it is not truly infinite. Yet it will not be tamed through the engineering of individual pathways, but through the creation of broader biocatalytic platforms. Rather than making one pathway at a time, efforts should be concentrated on developing the tools needed to make more pathways within a given period of time. Today's jungles are cut down not with axes, but with bullbozers.

The author truly hopes that this thesis inspires and empowers others to claim their stake, their homestead, on the frontier that is metabolic pathway engineering. Though the work is hard, the field is young, and the potential is great. Today the engineering of successful pathways is luck, but tomorrow it will be skill. A teacher very dear to the author once said that fortune favors the prepared mind. Today is the day to work hard and prepare, for biotechnology's best days have yet to come.

## References

Aharoni, A., Gaidukov, L., Tagur, S., Toker, L., Silman, I., and Tawfik, D.S. (2004). Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization. *Proc. Natl. Acad. Sci. USA*, 101:482-487.

Alper, H. and Stephanopoulos, G. (2007). Global transcription machinery engineering: A new approach for improving cellular phenotype. *Metab. Eng.*, 9:258-267.

Anderson, J.C., Clarke, E.J., Arkin, A.P., and Voigt, C.A. (2006). Environmentally Controlled Invasion of Cancer Cells by Engineered Bacteria. *J. Mol. Biol.*, 355:619-627.

Asako, H., Shimizu, M., and Itoh, N. (2008). Engineering of NADPH-dependent aldo-keto reductase from *Penicillium citrinum* by directed evolution to improve thermostability and enantioselectivity. *Appl. Microbiol. Biotechnol.*, 80:805-812.

Atsumi, S., Cann, A.F., Connor, M.R., Shen, C.R., Smith, K.M., Brynildsen, M.P., Chou, K.J.Y., Hanai, T., and Liao, J.C. (2008). Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab. Eng.*. Doi:10.1016/j.ymben.2007.08.003.

Atsumi, S., Hanai, T., and Liao, J.C. (2007). Non-fermentive pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, 451:86-90.

Atwood, D., Booth, C., Yeates, S.G., Chaibundit, C., and Ricardo, N.M.P.S. (2007). Block copolymers for drug solubilization: Relative hydrophobicities of polyether and polyester micelle-core-forming blocks. *Int. J. Pharm.*, 345:35-41.

Bonomo, J., Warnecke, T., Hume, P., Marizcurrena, A., and Gill, R.T. (2006). A comparative study of metabolic engineering anti-metabolite tolerance in *Escherichia coli*. *Metab. Eng.*, 8:227-239.

Bourne, R.A., Stevens, J.G., Ke, J., and Poliakoff, M. (2007). Maximising opportunities in supercritical chemistry: the continuous conversion of levulinic acid to  $\gamma$ -valerolactone in CO<sub>2</sub>. *Chem. Comm.*, 4632-4634.

Boynton, Z.L., Bennett, G.N., and Rudolph, F.B. (1996). Cloning, sequencing, and expression of genes encoding phosphotransacetylase and acetate kinase from *Clostridium acetobutylicum* ATCC 824. *Appl. Environ. Microbiol.*, 62:2758-2766.

Bradford, M.M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.*, 72:248-254.

Brazeau, B.J., Gort, S.J., Jessen, H.J., Andrew, A.J., and Liao, H.H. (2006). Enzymatic Activation of Lysine 2,3-Aminomutase from *Porphyromonas gingivalis*. *Appl. Environ. Microbiol.*, 72:6402-6404.

Bro C., Regenberg B., Förster J., and Nielsen, J (2006). In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab. Eng.*, 8:102-111.

Brumaghim, J.L., Li, Y., Henle, E., and Linn, S. (2003). Effects of hydrogen peroxide upon nicotinamide nucleotide metabolism in *Escherichia coli*: changes in enzyme levels and nicotinamide nucleotide pools and studies of the oxidation of NAD(P)H by Fe(III). *J. Biol. Chem.*, 278:42495-42504.

Buday, Z., Linden, J., and Karim, M. (1990). Improved acetone-butanol fermentation analysis using subambient HPLC column temperature. *Enzyme Microb. Technol.*, 12:24-27.

Capitani, G., De Biase, D., Aurizi, C., Gut, H., Bossa, F., and Grütter, M.G. (2003). Crystal structure and functional analysis of *Escherichia coli* glutamate decarboxylase. *EMBO J.*, 22:4027-4037.

Capsi, R., Foerster, H., Fulcher, C.A., Hopkinson, R., Ingraham, J., Kaipa, P., Krummenacker, M., Paley, S., Pick, J., Rhee, S.Y., Tissier, C., Zhang, P., and Karp, P.D. (2006). MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nuc. Acids Res.*, 34:D511-D516.

Carey, F.A. (2000). *Organic Chemistry*. 4<sup>rd</sup> ed., McGraw Hill, New York, NY.



Casqueiro, J., Bañuelos, O., Gutiérrez, S., and Martín, J.F. (2001). Metabolic Engineering of the Lysine Pathway for  $\beta$ -Lactam Overproduction in *Penicillium chrysogenum*. In Focus on Biotechnology, Van Broekhoven, A., Shapiro, F., and Anne, J. eds. (Dordrecht, Netherlands: Kluwer Academic Publishers), pp. 147-159.

Cha, J.Y. and Hanna, M.A. (2002). Levulinic acid production based on extrusion and pressurized batch reaction. *Ind. Crop. Prod.*, 2002, 16:109-118.

Chang, C., Cen, P. and Ma, X. (2007). Levulinic acid production from wheat straw. *Bioresource Technol.*, 2007, 98:1448-1453.

Chang, Y.-C. and Chu, I.-M. (2008). Methoxy poly(ethylene glycol)-*b*-poly(valerolactone) diblock polymeric micelles for enhanced encapsulation and protection of camptothecin. *Eur. Polymer J.*, 44:3922-3930.

Chang, Y.Y. and Cronan, J.E., Jr. (2000). Conversion of *Escherichia coli* pyruvate oxidase to an 'alpha-ketbutyrate oxidase'. *Biochem. J.*, 352 Pt. 3, 717-724.

Chen, G.Q. and Wu, Q. (2005). Microbial production and applications of chiral hydroxyalkanoates. *Appl. Microbiol. Biotechnol.*, 67:592-599.

Chiba, T. and Nakai, T.A. (1985). Synthetic approach to (1)-thienamycin from methyl (*R*)-(2)-3-hydroxybutanoate. A new entry to (3*R*,4*R*)-3-[(*R*)-1-hydroxyethyl-4-acetoxy-2-azetidinone. *Chem. Lett.*, 5:651-654.

Chiba, T., and Nakai, T. (1987). A new synthetic approach to the carbapenem antibiotic PS-5 from ethyl(*S*)-3-hydroxybutanoate. *Chem. Lett.*, 11:2187-2188.

Chinen, A., Kozlov, Y.I., Hara, Y., Izui, H., and Yasueda, H. (2007). Innovative Metabolic Pathway Design for Efficient L-Glutamate Production by Suppressing CO<sub>2</sub> Emission. *J. Biosci. Bioeng.*, 103:262-269.

Csizmadia, F. (2000). JChem: Java applets and modules supporting chemical database handling from web browsers. *J. Chem. Inf. Comput. Sci.*, 40:323-324.

Datsenko, K.A. and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA*, 97:6640-6645.

De Roo, G., Kellerhals, M.B., Ren, Q., Witholt, B. and Kessler, B. (2002). Production of chiral *R*-3-hydroxyalkanoic acids and *R*-3-hydroxyalkanoic acid methylesters via hydrolytic degradation of polyhydroxyalkanoate synthesized by *Pseudomonads*. *Biotechnol. Bioeng.*, 77:717-722.

Deakin, S., Leviev, I., Gomaraschi, M., Calabresi, L., Franceschini, G., and James, R.W. (2002). Enzymatically Active Paraoxonase-1 is Located at the External Membrane of Producing Cells and Released by a High Affinity, Saturable, Desorption Mechanism. *J. Biol. Chem.*, 277:4301-4308.

Derman, A.I. and Beckwith, J. (1991). *Escherichia coli* Alkaline Phosphatase Fails to Acquire Disulfide Bonds when Retained in the Cytoplasm. *J. Bacteriol.*, 173:7719-7722.

Draganov, D.I., Teiber, J.F., Speelman, A., Osawa, Y., Sunahara, R., and La Du, B.N. (2005). Human paraoxonases (PON1, PON2, and PON3) are lactonases with overlapping and distinct substrate specificities. *J. Lipid Res.*, 46:1239-1247.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.J., and Keasling, J.D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.*, 27:753-759.

Dürre, P., Böhringer, M., Nakotte, S., Schaffer, S., Thormann, K., and Zickner, B. (2002). Transcriptional Regulation of Solventogenesis in *Clostridium acetobutylicum*. *J. Mol. Microbiol. Biotechnol.*, 4:295-300.

Edwards, J.S., Covert, M., and Palsson, B. (2002). Metabolic modeling of microbes: the flux-balance approach. *Environ. Microbiol.*, 4:133-140.

Ellis, L.B.M., Roe, D., and Wackett, L.P. (2006). The University of Minnesota Biocatalysis/Biodegradation Database: the first decade. *Nuc. Acids Res.*, 34:D517-D521.

Endo, T. and Koizumi, S. (2001). Microbial Conversion with Cofactor Regeneration using Genetically Engineered Bacteria. *Adv. Synth. Catal.*, 343:521-528.

Endy, D. (2005). Foundations for engineering biology. *Nature*, 438:449-453.

Ewering, C., Lütke-Eversloh, T., Luftmann, H., and Steinbüchel, A. (2002). Identification of novel sulfur-containing bacterial polyesters: biosynthesis of poly(3-hydroxy-S-propyl- $\omega$ -thioalkanoates containing thioether linkages in the side chains. *Microbiology*, 148:1397-1406.

Fenner, K., Gao, J., Kramer, S., Ellis, L., and Wackett, L. (2008). Data-driven extraction of relative reasoning rules to limit combinatorial explosion in biodegradation pathway prediction. *Bioinformatics*, 24:2079-2085.

Farinas, E.T., Bulter, T., and Arnold, F.H. (2001). Directed enzyme evolution. *Curr. Opin. Biotechnol.*, 12:545-551.

Gao, H.J., Wu, Q. and Chen, G.Q. (2002). Enhanced production of D-(-)-3-hydroxybutyric acid by recombinant *Escherichia coli*. *FEMS Microbiol. Lett.*, 2002, 213:59-65.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.*, 31:3784-3788.

Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., and Crea, Roberto (1979). Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc. Natl. Acad. Sci.*, 76:106-110.

González-Lergier, J., Broadbelt, L.J., and Hatzimanikatis, V. (2005). Theoretical Considerations and Computational Analysis of the Complexity in Polyketide Synthesis Pathways. *J. Am. Chem. Soc.*, 127:9930-9938.

Gorenflo, V., Schmack, G., Vogel, R. and Steinbüchel, A. (2001). Development of a Process for the Biotechnological Large-Scale Production of 4-Hydroxyvalerate-Containing Polyesters and Characterization of Their Physical and Mechanical Properties. *Biomacromolecules* 2:45-57.

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends Biotechnol.*, 22:346-353.

Handelsman, J. (2004). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev.*, 68:669-685.

Hatzimanikatis, V., Li, C., Ionita, J.A., Henry, C.S., Jankowski, M.D., and Broadbelt, L.J. (2005). Exploring the diversity of complex metabolic networks. *Bioinformatics*, 21:1603-1609.

Hayes, D.J., Fitzpatrick, S., Hayes, M.H.B. and Ross, J.R.H. (2006). The Biofine process – production of levulinic acid, furfural, and formic acid from lignocellulosic feedstocks. In: Kamm B., Gruber P.R., Kamm M. (Eds.), *Biorefineries—Industrial Processes and Products*. 2006, Wiley-VCH, Weinheim, pp. 139-164.

Hazer, B. and Steinbüchel, A. (2007). Increase diversification of polyhydroxyalkanoates by modification reactions for industrial and medical applications. *Appl. Microbiol. Biotechnol.*, 74:1-12.

Hederos, S., Broo, K.S., Jakobsson, E., Kleywegt, G.J., Mannervik, B., and Baltzer, L. (2004). Incorporation of a single His residue by rational design enables thiol-ester hydrolysis by human glutathione transferase A1-1. *Proc. Natl. Acad. Sci.USA*, 101:13163-13167.

Horváth, I.T., Mehdi, H., Fábos, V., Boda, L., and Mika, L.T. (2008).  $\gamma$ -Valerolactone – a sustainable liquid for energy and carbon-based chemicals. *Green Chem.*, 10:238-242.

Huisman, G.W., Wonink, E., Meima, R., Kazemier, B., Terpstra, P., and Witholt, B. (1991). Metabolism of Poly(3-hydroxyalkanoates) (PHAs) by *Pseudomonas oleovorans*. *J. Biol. Chem.* 266:2191-2198.

Hollingsworth, R.I. (1999). Taming carbohydrate complexity: a facile, high-yield route to chiral 2,3-dihydroxybutanoic acids and 4-hydroxytetrahydrofuran-2-ones with very high optical purity from pentose sugars. *J. Org. Chem.*, 64:7633-7634.

Hollingsworth, R.I. and Wang, G. (2001). Process for the preparation of 3,4-dihydroxybutanoic acid and salts and lactones derived therefrom. US patent no. 6,239,311.

Hollingsworth, R.I. and Wang, G. (2000). Toward a carbohydrate-based chemistry: progress in the development of general-purpose chiral synthons from carbohydrates. *Chem. Rev.*, 100:4267-4282.

Horváth, I.T., Mehdi, H., Fábos, V., Boda, L., and Mika, L.T. (2008).  $\gamma$ -Valerolactone – a sustainable liquid for energy and carbon-based chemicals. *Green Chem.*, 10:238-242.

Inui, M., Suda, M., Kimura, S., Yasuda, K., Suzuki, H., Toda, H., Yamamoto, S., Okino, S., Suzuki, N., and Yukawa, H. (2008). Expression of *Clostridium acetobutylicum* butanol synthetic genes in *Escherichia coli*. *Appl. Microbiol. Biotechnol.*, 77:1305-1316.

Jaipuri, F.A., Jofre, M.F., Schwarz, K.A. and Pohl, N.L. (2004). Microwave-assisted cleavage of Weinreb amide for carboxylate protection in the synthesis of a (*R*)-3-hydroxyalkanoic acid. *Tetrahed. Lett.* 45:4149-4152.

Jankowski, M.D., Henry, C.S., Broadbelt, L.J., and Hatzimanikatis, V. (2008). Group Contribution Method for Thermodynamic Analysis of Complex Metabolic Networks. *Biophys. J.*, 95:1487-1499.

Jiang, L., Althoff, E.A., Clemente, F.R., Doyle, L., Rothlisberger, D., Zanghellini, A., Gallaher, J.L., Betker, J.L., Tanaka, F., Barbas, C.F., III, Hilvert, D., Houk, K.N., Stoddard, B.L., and Baker, D. (2008). *De novo* Computational Design of Retro-Aldol Enzymes. *Science* 319:1387-1391.

Jones, D.T. and Woods, D.R. (1986). Acetone-Butanol Fermentation Revisited. *Microbiol. Rev.*, 50:484-524.

Kane, J.F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.*, 6:494-500.

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K.F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., and Hiraoka, M. (2006). From genomics to chemical genomics: new developments in KEGG. *Nuc. Acids Res.*, 34:D354-D357.



Kang, Z., Wang, Q., Zhang, H., and Qi, Q. (2008). Construction of a stress-induced system in *Escherichia coli* for efficient polyhydroxyalkanoates production. *Appl. Microbiol. Biotechnol.*, 79:203-208.

Kaplan, J., and DeGrado, W.F. (2004). *De novo* design of catalytic proteins. *Proc. Natl. Acad. Sci. USA*, 101:11566-11570.

Keasling, J.D. (2008). Synthetic Biology for Synthetic Chemistry. *ACS Chem. Biol.*, 2008, 3:64-76.

Keen, N.T., Tamaki, S., Kobayashi, D. and Trollinger, D. (1988). Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. *Gene* 70:191-197.

Kelly, R.M., Leemhuis, H., Rozeboom, H.J., van Oosterwijk, N., Dijkstra, B.W., and Dijkhuizen, L. (2008). Elimination of competing hydrolysis and coupling side reactions of a cyclodextrin glucanotransferase by directed evolution. *Biochem. J.*, 413:517-525.

Kim, E.E., Baker, C.T., Dwyer, M.D., Murcko, M.A., Rao, B.G., Tung, R.D., and Navia, M.A. (1995). Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc.*, 117:1181-1182.

Kumar, P., Deshmukh, A.N., Upadhyay, R.K., and Gurjar, M.K. (2005). A simple and practical approach to enantiomerically pure (*S*)-3-hydroxy- $\gamma$ -butyrolactone: synthesis of (*R*)-4-cyano-3-hydroxybutyric acid ethyl ester. *Tetrahedron Asymmetry*, 16:2717-2727.

Lange, J.-P., Vestering, J.Z., and Haan, R.J. (2007). Towards 'bio-based- Nylon: conversion of  $\gamma$ -valerolactone to methyl pentenoate under catalytic distillation conditions. *Chem. Comm.*, 3488-3490.

Lee, I.Y., Kim, M.K., Park, Y.H., and Lee, S.Y. (1996). Regulatory effects of cellular nicotinamide nucleotides and enzyme activities on poly(3-hydroxybutyrate) synthesis in recombinant *Escherichia coli*. *Biotechnol. Bioeng.*, 52:707-712.

Lee, S.H. and Park, O.-J. (2009). Uses and production of chiral 3-hydroxy- $\gamma$ -butyrolactones and structurally related chemicals. *Appl. Microbiol. Biotechnol.*, 84:817-828.

Lee, S.H., Park, O.-J., and Uh, H.-S. (2008). A chemoenzymatic approach to the synthesis of enantiomerically pure (*S*)-3-hydroxy- $\gamma$ -butyrolactone. *Appl. Microbiol. Biotechnol.*, 79:355-362.

Lee, S.Y. and Lee, Y. (2003). Metabolic engineering of *Escherichia coli* for the production of enantiomerically pure (*R*)-(-)-hydroxycarboxylic acids. *Appl. Environ. Microbiol.*, 69:3421-3426.

Lee, S.Y., Lee, Y. and Wang, F. (1999). Chiral Compounds from Bacterial Polyesters: Sugars to Plastics to Fine Chemicals. *Biotechnol. Bioeng.* 65:363-368.

Lee, S.Y., Park, J.H., Jang, S.H., Nielsen, L.K., Kim, J., and Jung, K.S. (2008). Fermentive Butanol Production by Clostridia. *Biotechnol. Bioeng.*, 101:209-228.

Lee, S.Y., Park, S.H., Lee, Y., and Lee, S.H. (2002). Production of chiral and other valuable compounds from microbial polyesters. In: Dio, Y., Steinbüchel, A. (eds.) *Biopolymers, polyesters III*. Wiley-VCH, Weinheim, pp. 375-387.

Lenz, R.W. and Marchessault, R.H. (2005). Bacterial polyesters: biosynthesis, biodegradable plastics and biotechnology. *Biomacromolecules*, 6:1-8.

Li, C., Henry, C.S., Jankowski, M.D., Ionita, J.A., Hatzimanikatis, V., and Broadbelt, L.J. (2004). Computational discovery of biochemical routes to specialty chemicals. *Chem. Eng. Sci.*, 59:5051-5060.

Liao, H.H., Gokarn, R.R., Gort, S.J., Jensen, H.J., and Selifonova, O (2007). Alanine 2,3-aminomutase. U.S. Patent 7,309,597.

Lin, H., Bennett, G.N., and San, K.Y. (2005). Metabolic engineering of aerobic succinate production systems in *Escherichia coli* to improve process productivity and achieve the maximum theoretical succinate yield. *Metab. Eng.*, 7:116-127.

Lin, S.-F., Chiou, C.-M., Yeh, C.-M. and Tsai, Y.-C. (1996). Purification and Partial Characterization of an Alkaline Lipase from *Pseudomonas pseudoalcaligenes* F-111. *Appl. Environ. Microbiol.*, 62:1093-1095.

Lippow, S.M., Wittrup, K.D., and Tidor, B. (2007). Computational design of antibody-affinity improvement beyond *in vivo* maturation. *Nat. Biotechnol.*, 25:1171-1176.

Liu, Q., Ouyang, S.P., Chung, A., Wu, Q. and Chen, G.Q. (2007). Microbial production of *R*-3-hydroxybutyric acid by recombinant *E. coli* harboring genes of *phbA*, *phbB*, and *tesB*. *Appl. Microbiol. Biotechnol.* 76:811-818.

Liu, S.J. and Steinbüchel, A. (2000a). A novel genetically engineered pathway for synthesis of poly(hydroxyalkanoic acids) in *Escherichia coli*. *Appl. Environ. Microbiol.* 66:739-743.

Liu, S.J. and Steinbüchel, A. (2000b). Exploitation of butyrate kinase and phosphotransbutyrylase from *Clostridium acetobutylicum* for the *in vitro* biosynthesis of poly(hydroxyalkanoic acid). *Appl. Microbiol. Biotechnol.* 53:545-552.

Liu, Y.H., Lu, F.P., Li, Y., Wang, J.L., and Gao, C. (2008). Acid stabilization of *Bacillus licheniformis* alpha amylase through introduction of mutations. *Appl. Microbiol. Biotechnol.*, 80:795-803.

Loose, C., Jensen, K., Rigoutsos, I., and Stephanopoulos, G. (2006). A linguistic model for the rational design of antimicrobial peptides. *Nature*, 443:867-869.

Machielsen, R., Leferink, N.G., Hendriks, A., Brouns, S.J., Hennemann, H.G., Daubetammann, T., and van der Oost, J. (2008). Laboratory evolution of *Pyrococcus furiosus* alcohol dehydrogenase to improve the production of (2S,5S)-hexanediol at moderate temperatures. *Extremophiles*, 12:587-594.

Manting, E.H. and Drissen, A.J.M. (2000). *Escherichia coli* translocase: the unravelling of a molecular machine. *Mol. Microbiol.*, 37:226-238.

Manzer, L.E. (2004). Catalytic synthesis of  $\alpha$ -methylene- $\gamma$ -valerolactone: a biomass-derived acrylic monomer. *Appl. Catal. A*, 272:249-256.

Manzer, L.E. and Hutchenson, K.W. (2005). Production of 5-methyl-dihydro-furan-2-one from levulinic acid in supercritical media. Sept. 2005, U.S. patent 2004/0254384.

Martin, C.H., Nielsen, D.R., Solomon, K.V., and Prather, K.J. (2009). Synthetic Metabolism: Engineering Biology at the Protein and Pathway Scales. *Chem. Biol.*, 16:277-286.

Martin, C. and Prather, K.L.J. (2009). High-titer production of monomeric hydroxyvalerates from levulinic acid in *Pseudomonas putida*. *J. Biotechnol.*, 139:61-67.

Martin, V.J., Yoshikuni, Y., and Keasling, J.D. (2001). The *in vivo* synthesis of plant sesquiterpenes by *Escherichia coli*. *Biotechnol. Bioeng.*, 75:497-503.

Martin, V.J.J., Pitera, D.J., Withers, S.T., Newman, J.D., and Keasling, J.D. (2003). Engineering a mevalonate pathway in *Escherichia coli* for the production of terpenoids. *Nat. Biotechnol.*, 21:796-802.

Masamune, S., Walsh, C.T., Sinskey, A.J., and Peoples, O.P. (1989). Poly-(*R*)-3-hydroxybutyrate (PHB) biosynthesis: mechanistic studies on the biological Claisen condensation catalyzed by  $\beta$ -ketoacyl thiolase. *Pure Appl. Chem.*, 61:303-312.

Mavrovouniotis, M.L. (1991). Estimation of Standard Gibbs Energy Changes of Biotransformations. *J. Biol. Chem.*, 266:14440-14445.

Meinhold, P., Peters, M.W., Chen, M.M., Takahashi, K., and Arnold, F.H. (2005). Direct conversion of ethane to ethanol by engineered cytochrome P450 BM3. *Chembiochem*, 6:1765-1768.

Meinhold, P., Peters, M.W., Hartwick, A., Hernandez, A.R., and Arnold, F.H. (2006). Engineering Cytochrome P450 BM3 for Terminal Alkane Hydroxylation. *Adv. Synth. Catal.*, 348:763-772.

Mermelstein L.D., Papoutsakis E.T., Petersen D.J., and Bennett G.N. (1993). Metabolic Engineering of *Clostridium acetobutylicum* ATCC 824 for Increased Solvent Production by Enhancement of Acetone Fromation Enzyme Activities Using a Synthetic Acetone Operon. *Biotechnol. Bioeng.*, 42:1053-1060.

Meurer, G., Biermann, G., Schütz, A., Harth, S., and Schweizer, E. (1991). Molecular structure of the mutlifunctional fatty acid synthetase gene of *Brevibacterium ammoniagenes*: its sequence of catalytic domains is formally consistent with a head-to-tail fusion of the two yeast genes *FAS1* and *FAS2*. *Mol. Gen. Genet.*, 232:106-116.

Meyer, M.M., Hochrein, L., and Arnold, F.H. (2006). Structure-guided SCHEMA recombination of distantly related  $\beta$ -lactamases. *Protein Eng., Design and Selection*, 19:563-570.

Meyer, M.M., Silberg, J.J., Voigt, C.A., Endelman, J.B., Mayo, S.L., Wang, Z.-G., and Arnold, F.H. (2003). Library analysis of SCHEMA-guided protein recombination. *Protein Sci.*, 12:1686-1693.

Mori, K. (1981). A simple synthesis of (*S*)-(+)-sulcatol, the pheromone of *Gnathotrichus tetusus* employing baker's yeast for asymmetric reduction. *Tetrahedron*, 37:1341-1342.

Mueller-Cajar, O. and Whitney, S.M. (2008). Evolving improved *Synechococcus rubisco* functional expression in *Escherichia coli*. *Biochem. J.*, 414:205-214.

Müller, U., van Assema, F., Gunsior, M., Orf, S., Kremer, S., Schipper, D., Wagemans, A., Townsend, C.A., Sonke, T., Bovenberg, R., and Wubbolts, M. (2006). Metabolic engineering of the *E. coli* phenylalanine pathways for the production of D-phenylglycine (D-Phg). *Metab. Eng.*, 8:196-208.

Naggert, J., Narasimhan, M.L., DeVeaux, L., Cho, H., Randhawa, Z.I., Cronan, J.E., Green, B.N. and Smith, S. (1991). Cloning, sequencing, and characterization of *Escherichia coli* thioesterase II. *J. Biol. Chem.* 266:11044-11050.

Nakagawa, A., Idogaki, H., Kato, K., Shinmyo, A., and Suzuki, T. (2006). Improvement on the production of (R)-4-chloro-3-hydroxybutyrate and (S)-3-hydroxy- $\gamma$ -butyrolactone with recombinant *Escherichia coli* cells. *J. Biosci. Bioeng.*, 101:97-103.

Naoki, A., and Hiroshi, M. (1997). Predicting Protein Secondary Structure Using Stochastic Tree Grammars. *Mach. Learn.*, 29:275-301.

Nui, W., Molefe, M.N., and Frost, J.W. (2003). Microbial Synthesis of the Energetic Material Precursor 1,2,4-Butanetriol. *J. Am. Chem. Soc.*, 125:12998-12999.

Nuñez, M.F., Pellicer, M.T., Badia, J., Aguilar, J., and Baldoma, L. (2001). Biochemical characterization of the 2-ketoacid reductases encoded by *ycdW* and *yiaE* genes in *Escherichia coli*. *Biochem. J.*, 354:707-715.



Olivera, E.R., Carnicero, D., García, B., Miñambres, B., Moreno, M.A., Cañedo, L., DiRusso, C.C., Naharro, G. and Luengo, J.M. (2001). Two different pathways are involved in the  $\beta$ -oxidation of *n*-alkanoic and *n*-phenylalkanoic acids in *Pseudomonas putida* U: genetic studies and biotechnological applications. *Mol. Microbiol.* 39:863-874.

Otey, C.R., Silberg, J.J., Voigt, C.A., Endelman, J.B., Bandara, G., and Arnold, F.H. (2004). Functional evolution and structural conservation in chimeric cytochromes p450: calibrating a structure-guided approach. *Chem. Biol.*, 11:309-318.

Park, S.J. and Lee, S.Y. (2004). Biosynthesis of poly(3-hydroxybutyrate-co-3-hydroxyalkanoates) by metabolically engineered *Escherichia coli* strains. *Appl. Biochem. Biotechnol.* 114:335-346.

Park, S.H., Lee, S.H., and Lee, S.Y. (2001). Preparation of optically active  $\beta$ -aminoacids from microbial polyester polyhydroxyalkanoates. *J. Chem. Res. Synop.*, 11:498-499.

Park, S.J., Park, J.P., and Lee, S.Y. (2002). Metabolic engineering of *Escherichia coli* for the production of medium-chain-length polyhydroxyalkanoates rich in specific monomers. *FEMS Microbiol. Lett.* 214:217-222.

Park, Y.M., Chun, J.P., Rho, K.R., Yu, H.S., and Hwang, I. (2004). Process for preparing optically pure (*S*)-3-hydroxy- $\gamma$ -butyrolactone. US patent no. 6,713,290.

Patel, M., Dornburg, V., Hermann, B., Roes, L., Hüsing, B., Overbeek, L., Terragni, F., and Recchia, E. (2006). Medium and Long-term Opportunities and Risks of the Biotechnological Production of Bulk Chemicals from Renewable Resources – The Potential of White Biotechnology. Utrecht, Netherlands: Utrecht University, September 2006.

Patel, R.N. (2000). Stereoselective Biocatalysis. CRC Press, Boca Raton, FL.

Patel, R.N. (2006). Biocatalysis: synthesis of chiral intermediates for drugs. *Curr. Opin. Drug Discov. Devel.*, 9:741-764.

Phue, J.N., Noronha, S.B., Hattacharyya, R., Wolfe, A.J., and Shiloach, J. (2005). Glucose metabolism at high density growth of *E. coli* B and *E. coli* K: Differences in metabolic pathways are responsible for efficient glucose utilization in *E. coli* B as determined by microarrays and Northern blot analyses. *Biotechnol. Bioeng.*, 90:805-820.

Pitera, D.J., Paddon, C.J., Newman, J.D., and Keasling, J.D. (2007). Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab. Eng.*, 9:193-207.

Pleiss, J. (2006). The promise of synthetic biology. *Appl. Microbiol. Biotechnol.*, 73: 735-739.

Pollard, D.J. and Woodley, J.M. (2007). Biocatalysis for pharmaceutical intermediates: the future is now. *Trends Biotechnol.*, 25:66-73.

Prather, K.L.J. and Martin, C.H. (2008). *De novo* biosynthetic pathways: rational design of microbial chemical factories. *Curr. Opin. Biotechnol.*, 19:468-474.

Przytycka, T., Srinivasan, R., and Rose, G.D. (2002). Recursive domains in proteins. *Protein Sci.* 11:409-417.

Reetz, M.T. and Carballeira, J.D. (2007). Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.*, 2:891-903.

Reetz, M.T., Kahakeaw, D., and Lohmer, R. (2008). Addressing the numbers problem in directed evolution. *Chembiochem*, 9:1797-1804.

Ren, Q., Ruth, K., Thöny-Meyer, L. and Zinn, M. (2007). Process Engineering for Production of Chiral Hydroxycarboxylic Acids from Bacterial Polyhydroxyalkanoates. *Macromol. Rapid Commun.* 28:2131-2136.

Rigoutsos, I. and Floratos, A. (1998). Combinatorial pattern discovery in biological sequences: The TEIRESIAS algorithm [published erratum appears in *Bioinformatics* 1998;14(2):229]. *Bioinformatics*, 14:55-67.

Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y., Withers, S.T., Shiba, Y., Sarpong, R., and Keasling, J.D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 440:940-943.

Rodrigo, G., Carrera, J., Prather, K.J., and Jaramillo, A. (2008). DESHARKY: Automatic design of metabolic pathways for optimal cell growth. *Bioinformatics*, 24: 2554-2556.

Rothlisberger, D., Khersonsky, O., Wollacott, A.M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J.L., Althoff, E.A., Zanghellini, A., Dym, O., Albeck, S., Houk, K.N., Tawfik, D.S., and Baker, D. (2008). Kemp elimination catalysts by computational enzyme design. *Nature*, 453:190-195.

Sambrook, J. and Russell, D.W. (2001). *Molecular Cloning: a Laboratory Manual*. 3<sup>rd</sup> Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Saraf, M.C., Horswill, A.R., Benkovic, S.J., and Maranas, C.D. (2004). FamClash: A method for ranking the activity of engineered enzymes. *Proc. Natl. Acad. Sci. USA*, 101:4142-4147.

Savage, D.F., Way, J., and Silver, P.A. (2008). Defossilizing Fuel: How Synthetic Biology Can Transform Biofuel Production. *ACS Chem. Biol.*, 3:13-16.

Schneider, D., Duperchy, E., Depeyrot, J., Coursange, E., Lenski, R., and Blot, M. (2002). Genomic comparisons among *Escherichia coli* strains B, K-12, and O157:H7 using IS elements as molecular markers. *BMC Microbiol.*, 2:18-25.

Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G., and Schomburg, D. (2004). BRENDA, the enzyme database: updates and major new developments. *Nuc. Acids Res.*, 32:D431-D433.

Schubert, P., Steinbüchel, A., and Schlegel, H.G. (1988). Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly-beta-hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. *J. Bacteriol.*, 170:5837-5847.

Schweiger, G. and Buckel, W. (1984). On the dehydration of (*R*)-lactate in the fermentation of alanine to propionate by *Clostridium propionicum*. *FEBS Lett.*, 171:79-84.

Searls, D.B. (1997). Linguistic approaches to biological sequences. *Comput. Appl. Biosci.*, 13:333-344.

Searls, D.B. (2002). The language of genes. *Nature* 420:211-217.

Seebach, D., Albert, M., Arvidsson, P.I., Rueping, M., and Schreiber, J.V. (2001). From the biopolymer PHB to biological investigations of unnatural  $\beta$ - and  $\gamma$ -peptides. *Chimia*, 55:345-353.

Seebach, D., Beck, A.K., Breitschuh, R. and Job, K. (1998). Direct degradation of the biopolymer [(*R*)-3-hydroxybutyric acid] to (*R*)-3-hydroxybutanoic acid and its methyl ester. *Org. Synth.* 71:39-43.

Seebach, D., Chow, H.F., Jackson, R.F.W., Sutter, M.A., Thaisrivongs, S., and Zimmermann, J. (1986). (+)-11,11'-di-O-methylelaiophylidene preparation from Elaiophyllin and total synthesis from (*R*)-3-hydroxybutyrate and (*S*)-malate. *Liebigs Ann. Chem.*, 1986:1281-1308.

Seker, T., Møller, K., and Nielsen, J. (2005). Analysis of acyl CoA ester intermediates of the mevalonate pathway in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.*, 67:119-124.

Selmer, T., Willanzheimer, A., and Hetzel, M. (2002). Propionate CoA-transferase from *Clostridium propionicum*: cloning of the gene and identification of glutamate 324 at the active site. *Eur. J. Biochem.*, 269:372-380.

Shi, C., Lu, X., Ma, C., Ma, Y., Fu, X., and Yu, W. (2008). Enhancing the thermostability of a novel beta-agarase AgaB through directed evolution. *Appl. Biochem. Biotechnol.*, 151:51-59.

Shin, H.-D. and Chen, R.R. (2008). Extracellular recombinant protein production from an *Escherichia coli* lpp deletion mutant. *Biotechnol. Bioeng.*, 101:1288-1296.

Shin, H.I., Chang, J.H., Woo, Y.M., and Kim, Y.S. (2005). A process for the synthesis of 3-hydroxy-gamma-butyrolactone. WO patent 2005/030747.

Shiraki, M., Endo, T., and Saito, T. (2006). Fermentative production of (*R*)-(-)-3-hydroxybutyrate using 3-hydroxybutyrate dehydrogenase null mutant of *Ralstonia eutropha* and recombinant *Escherichia coli*. *J. Biosci. Bioeng.*, 102:529-534.

Si Jae Park, S.Y.L. (2004). New *fadB* homologous enzymes and their use in enhanced biosynthesis of medium-chain-length polyhydroxyalkanoates in *fadB* mutant *Escherichia coli*. *Biotechnol. Bioeng.*, 86:681-686.

Slater, S., Houmiel, K.L., Tran, M., Mitsky, T.A., Taylor, N.B., Padgett, S.R., and Gruys, K.J. (1998). Multiple  $\beta$ -ketothiolases mediate poly( $\beta$ -hydroxyalkanoate) copolymer synthesis in *Ralstonia eutropha*. *J. Bacteriol.*, 180:1979-1987.

Snell, K.D., Feng, F., Zhong, L., Martin, D., and Madison, L.L. (2002). YfcX enables medium-chain-length poly(3-hydroxyalkanoate) formation from fatty acids in recombinant *Escherichia coli fadB* strains. *J. Bacteriol.*, 184:5696-5705.

Steinbüchel, A. and Valentin, H.E. (1995). Diversity of microbial polyhydroxyalkanoic acids. *FEMS Microbiol. Lett.*, 128:219-228.

Stelling, J., Klamt, S., Betterbrock, K., Schuster, S., and Gilles, E.D. (2002). Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420:190-193.

Stemmer, W.P. (1994a). DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA*, 91:10747-10751.

Stemmer, W.P.C. (1994b). Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, 370:389-391.

Stim-Herndon, K.P., Petersen, D.J., and Bennett, G.N. (1995). Molecular characterization of the acetyl coenzyme A acetyltransferase (thiolase) from *Clostridium acetyobutylicum* ATCC 824. *Gene*, 154:81-85.



Sun, X., Liu, Z., Qu, Y., and Li, X. (2008). The Effects of Wheat Bran Composition on the Production of Biomass-Hydrolyzing Enzymes by *Penicillium decumbens*. *Appl. Biochem. Biotechnol.*, 146:119-128.

Suzuki, T., Idogaki, H., and Kasai, N. (1999). Dual production of highly pure methyl (R)-4-chloro-3-hydroxybutyrate and (S)-3-hydroxy-g-butyrolactone with *Enterobacter* sp. *Enzyme Microb. Technol.*, 24:13-20.

Taguchi, S., Yamada, M., Matsumoto, K., Tajima, K., Satoh, Y., Masanobu, M., Ohno, K., Kohda, K., Shimamura, T., Kambe, H., and Shusei, O. (2008). A microbial factory for lactate-based polyesters using a lactate-polymerizing enzyme. *Proc. Natl. Acad. Sci.*, 105:17323-17327.

Tao, H. and Cornish, V.W. (2002). Milestones in directed enzyme evolution. *Curr. Opin. Chem. Biol.*, 6:858-864.

Tasaki, O., Hiraide, A., Shiozaki, T., Yamamura, H., Ninomiya, N., and Sugimoto, H. (1999). The dimer and trimer of 3-hydroxybutyrate oligomer as a precursor of ketone bodies for nutritional care. *J. Parenter. Enteral. Nutr.*, 23:321-325.

Teiber, J.F., Draganov, D.I., and La Du, B.N. (2003). Lactonase and lactonizing activities of human serum paraoxonase (PON1) and rabbit serum PON3. *Biochem. Pharmacol.*, 66:887-896.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.*, 72:211-222.

Tokiwa, Y. and Calabia, B.P. (2008). Biological production of functional chemicals from renewable resources. *Can. J. Chem.*, 86:548-555.

Tola, N.H. and Joshua-Tor, L. (2006). Strategies for protein coexpression in *Escherichia coli*. *Nat. Methods*, 3:55-64.

Tseng, H.-C., Martin, C., Nielsen, D., and Prather, K.L.J. (2009). Metabolic Engineering of *Escherichia coli* for Enhanced Production of (*R*)- and (*S*)-3-Hydroxybutyrate. *Appl. Environ. Microbiol.*, 75:3137-3145.

Tsuge, T., Hisano, T., Taguchi, S., and Doi, Y. (2003). Alteration of chain length substrate specificity of *Aeromonas caviae* R-enantiomer-specific enoyl-coenzyme A hydratase through site-directed mutagenesis. *Appl. Environ. Microbiol.*, 69:4830-4836.

Verlinden, R.A.J., Hill, D.J., Kenward, M.A., Williams, C.D., and Radecka, I. (2007). Bacterial synthesis of biodegradable polyhydroxyalkanoates. *J. Appl. Microbiol.*, 102:1437-1449.

Villalobos, A., Ness, J.E., Gustafsson, C., Minshill, J., and Govindarajan, S. (2006). Gene Designer: a synthetic biology tool for construction artificial DNA segments. *BMC Bioinformatics*, 7:285.

Voigt, C.A., Martinez, C., Wang, Z.-G., Mayo, S.L., and Arnold, F.H. (2002). Protein building blocks preserved by recombination. *Nat. Struct. Mol. Biol.*, 9:553-558.

Walle, M.v.d. and Shiloach, J. (1998). Proposed mechanism of acetate accumulation in two recombinant *Escherichia coli* strains during high density fermentation. *Biotechnol. Bioeng.*, 57:71-78.

Wang, F. and Lee, S.Y. (1997). Poly(3-hydroxybutyrate) production with high productivity and high polymer content by a fed-batch culture of *Alcaligenes latus* under nitrogen limitation. *Appl. Environ. Microbiol.* 63:3703-3706.

Wang, G. and Hollingsworth, R.I. (1999a). Direct conversion of (*S*)-3-hydroxy- $\gamma$ -butyrolactone to chiral three carbon building blocks. *J. Org. Chem.*, 64:1036-1038.

Wang, G. and Hollingsworth, R.I. (1999b). Synthetic routes to L-carnitine and L-gamma-amino-beta hydroxybutyric acid from (*S*)-3-hydroxy- $\gamma$ -butyrolactone by functional group priority switching. *Tetrahedron Asymmetry*, 10:1895-1901.

Wang, J. and Yu, H.-Q. (2007). Biosynthesis of polyhydroxybutyrate (PHB) and extracellular polymeric substances (EPS) by *Ralstonia eutropha* ATCC 17699 in batch cultures. *Appl. Microbiol. Biotechnol.*, 75:871-878.

Werpy, T. and Petersen, G.: Top value added chemicals from biomass, Vol 1: results of screening for potential candidates from sugars and synthesis gas. Oak Ridge, TN: U.S. Department of Energy, August 2004.

Wilks, J.C. and Slonczewski, J.L. (2007). pH of the Cytoplasm and Periplasm of *Escherichia coli*: Rapid Measurement by Green Fluorescent Protein Fluorimetry. *J. Bacteriol.*, 189:5601-5607.

Woods, D.R. (1995). The genetic engineering of microbial solvent production. *Trends biotechnol.*, 13:259-264.

Wu, C.H., Apweiler, R., Bairoch, A., Natale, D.A., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., et al. (2006). The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nuc. Acids Res.*, 34:D187-D191.

Xia, X.X., Han, M.J., Lee, S.Y., and Yoo, J.S. (2008). Comparison of the extracellular proteomes of *Escherichia coli* B and K-12 strains during high cell density cultivation. *Proteomics*, 8:2089-2103.

Yin, E., Le, Y.L., Pei, J.J., Shao, W.L., and Yang, Q.Y. (2008). High-level expression of the xylanase from *Thermomyces lanuginosus* in *Escherichia coli*. *World. J. Microbiol. Biotechnol.*, 24:275-280.

Yoshikuni, Y., Dietrich, J.A., Nowroozi, F.F., Babbitt, P.C., and Keasling, J.D. (2008). Redesigning Enzymes Based on Adaptive Evolution for Optimal Function in Synthetic Metabolic Pathways. *Chem. Biol.*, 15:607-618.

Zanghellini, A., Jiang, L., Wollacott, A.M., Cheng, G., Meiler, J., Althoff, E.A., Rothlisberger, D., and Baker, D. (2006). New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci*, 15:2785-2794.

Zha, W., Shao, Z., Frost, J.W., and Zhao, H. (2004). Rational Pathway Engineering of Type I Fatty Acid Synthase Allows the Biosynthesis of Triacetic Acid Lactone from D-Glucose in Vivo. *J. Am. Chem. Soc.*, 2004:4534-4535.

Zhang, K., Sawaya, M.R., Eisenberg, D.S., and Liao, J.C. (2008). Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc. Natl. Acad. Sci.*, 105:20653-20658.

Zhao, K., Tian, G., Zheng, Z., Chen, J.C., and Chen, G.Q. (2003). Production of *D*-(-)-3-hydroxyalkanoic acid by recombinant *Escherichia coli*. *FEMS Microbiol. Lett.*, 218:59-64.

Zheng, Z., Gong, Q., Liu, T., Deng, Y., Chen, J.C., and Chen, G.Q. (2004). Thioesterase II of *Escherichia coli* plays an important role in 3-hydroxydecanoic acid production. *Appl. Environ. Microbiol.*, 70:3807-3813.

## **Acknowledgements**

The author is very grateful for the guidance and support of Prof. Kristala Jones Prather, for the love of Leah Octavio, for the tremendous help of Hsien-Chung Tseng with the 3HB work, for the assistance of Tyler DeWitt, Himanshu Dhamankar, Kayla Menard, Anudha Mittal, Jeffrey Mo, Danyi Wu, and Neil Zimmerman, for the funding of the NSF, SynBERC, Shell, and the MIT Presentential Scholarship program, for the friendship of the Prather lab, and for the strength to have come so far.