UNIVERSITY OF CALIFORNIA,
IRVINE

Engineering of Yeast for the Production of Fuels and Polyketides

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemical and Biochemical Engineering

by

Jin Wook Choi

Dissertation Committee:
Professor Nancy A. Da Silva, Chair
Professor Szu-Wen Wang
Professor Suzanne B. Sandmeyer

2014

Portion of Chapter 5 © 2014 Elsevier B.V.
All other materials © 2014 Jin Wook Choi

DEDICATION

To

Olivia, Chloe, and Su

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	vi
LIST OF ABBREVIATIONS	x
ACKNOWLEDGMENTS	xiv
CURRICULUM VITAE	xv
ABSTRACT OF THE DISSERTATION	xvi
CHAPTER 1: Introduction	1
1.1. Motivation	2
1.2. Objectives	7
1.3. References	9
CHAPTER 2: Literature Review	12
2.1. Metabolic engineering of Saccharomyces cerevisiae	13
2.1.1. Significance of using Saccharomyces cerevisiae for the production of bioethanol, industrial chemicals, and polyketides	13
2.1.2. Use of versatile vector series, pXPs	14
2.2. Bioethanol production in *Saccharomyces cerevisiae*	17
2.2.1. Pentose assimilation pathway	17
2.2.1.1. Pentose transport in *Saccharomyces cerevisiae*	17
2.2.1.2. Xylose pathway	18
2.2.1.3. Arabinose pathway	21
2.2.1.4. Pentose phosphate pathway	23
2.3. Polyketide synthesis in *Saccharomyces cerevisiae*	
2.3.1. Significance of polyketides and their synthesis in *Saccharomyces Cerevisiae*	25
2.3.1.1. Polyketides as a source of drug molecules	25
2.3.1.2. Replacing carbon from fossil	26
2.3.2. Polyketide synthases	29
2.3.2.1. Domain structure of polyketide synthases	29
2.3.2.2. Product offloading	30
2.3.3. Protease effect on polyketide synthases and polyketide synthesis	31
2.3.4. 6-methylsalicylic acid (6-MSA) biosynthesis	31
2.3.5. Biosynthesis of dihydromonacolin L (DML), the lovastatin precursor	33

2.3.5.1. Lovastatin biosynthesis 34
2.3.6. Biosynthesis of acetyl-CoA and malonyl-CoA 36
2.4. References 41

CHAPTER 3: Construction and Expression of Fungal Arabinose Pathway Genes in *Saccharomyces cerevisiae* 50
3.1. Abstract 51
3.2. Introduction 52
3.3. Materials and Methods 56
3.3.1. Molecular biology techniques 56
3.3.2. Vector constructions 56
3.3.2.1. Construction of pJC vectors 56
3.3.2.2. Codon/codon-pair-optimized gene synthesis and cloning 57
3.3.2.3. *GAL2*/*HXT1* hybrid gene synthesis and plasmid construction 60
3.3.3. Strain construction 61
3.3.4. Media and Cultivations 63
3.3.5. His-tagged protein purification using Ni-NTA spin columns 66
3.3.6. Western Blot analysis 67
3.3.7. Activity assay 67
3.3.8. Xylitol assay 68
3.4. Results and Discussion 69
3.4.1. Gene synthesis and plasmid construction 70
3.4.2. Expression and activity of codon/codon-pair-optimized fungal arabinose pathway enzymes 72
3.4.3. Comparison of expression level and activity of fungal arabinose pathway enzymes in crude cell extract 74
3.4.4. Construction of *Saccharomyces cerevisiae* with three fungal arabinose pathway enzymes and comparison of xylitol production level 77
3.4.5. Construction of strains with Gal2 or Gal2/Hxt1 chimera proteins as the only sugar monomer transporter 81
3.5. Conclusions and future directions 88
3.6. References 90

CHAPTER 4: Biosynthesis of Dihydromonacolin L, a Precursor to Lovastatin in *Saccharomyces cerevisiae* 94
4.1. Abstract 95
4.2. Introduction 96
4.3. Materials and Methods 99
4.3.1. Molecular biology techniques 99
4.3.2. Vector construction 100
4.3.3. Strain construction 104
4.3.4. Media and cultivation 105
4.3.5. Ni-NTA column purification and SDS-PAGE analysis 106

4.3.6. Plasmid stability test 106
4.3.7. Quantitative Real-time PCR 107
4.3.8. DML detection 107
4.4. Results and Discussion 109
4.4.1. Comparison of LovB and TE expression 109
4.4.2. Comparison of DML level 111
4.4.3. Identification of a dedicated thioesterases, LovG in *Aspergillus terreus* 113
4.5. Conclusions 114
4.6. References 115

CHAPTER 5: Pathway Engineering for the Enhanced Synthesis of 6-MSA in *Saccharomyces cerevisiae* 117
5.1. Abstract 118
5.2. Introductions 119
5.3. Materials and Methods 122
5.3.1. Molecular biology techniques 122
5.3.2. Vector construction 123
5.3.3. Strain construction 128
5.3.4. Media and cultivation 132
5.3.5. Plasmid stability test 135
5.3.6. Expression and Purification of 6-MSAS, Acc1, and Acc1^{S1157A} 135
5.3.7. *In vitro* activity assay 137
5.3.8. Glucose, Ethanol, and 6-MSA measurements using HPLC 138
5.4. Results and Discussion 139
5.4.1. Engineering of pyruvate dehydrogenase bypass for enhanced Synthesis of 6-MSA in *Saccharomyces cerevisiae* 139
5.4.2. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase 146
5.4.2.1. Identification of Snf1 target residue on Acc1 146
5.4.2.2. *In vitro* activity assay for Acc1^{S1157A} 148
5.4.2.3. *In vivo* production of 6-MSA and fatty acids 150
5.4.3. Metabolic pathway modifications to increase 6-MSA synthesis 152
5.4.4. Combined effect of pyruvate dehydrogenase bypass gene overexpression and *PYC1* deletion 158
5.5. Conclusions 160
5.6. References 161

CHAPTER 6: Improved 6-MSA Synthesis via Enhanced 6-MSA Synthase Expression System 165
6.1. Abstract 166
6.2. Introduction 167
6.3. Background 169
6.3.1. Manipulation of UMP synthesis pathway to obtain autoselection 169

capability
6.3.2. Engineering of *URA3* marker for increased plasmid copy number 171
6.4. Materials and methods 173
6.4.1. Molecular biology techniques 173
6.4.2. Vector construction 174
6.4.3. Yeast strain construction 178
6.4.4. Media and cultivation 179
6.4.5. Analytical methods 180
6.5. Results and discussion 181
6.5.1. Copy number and media effects 181
6.5.2. Enhancing 6-MSA production by employing the N-degron *URA3* marker 182
6.5.3. Improving plasmid stability in complex medium through autoselection 184
6.6. Conclusions 187
6.7. References 188

Appendices 191
A. Codon optimization and gene assembly 192
A.1. Codon and codon pair optimized gene sequences 192
A.1.1. *cALX1* 192
A.1.2. *clxr1* 192
A.1.3. *clad1* 192
A.2. Oligos for codon/codon-pair optimized gene assembly 193
A.2.1. *clad1* 193
A.2.2. *clxr1* 195
A.2.3. *cALX1* 196
A.3. Gene assembly procedure 198
B. Molecular biology protocols 200
B.1. Gibson reaction 200
C. Primer sequences 202
References 204

LIST OF FIGURES

		Page
Figure 2.1	pXP vector series	16
Figure 2.2	Pentose pathways	19
Figure 2.3	Pentose metabolism in recombinant yeast	24
Figure 2.4	Biosynthesis of 6-MSA in *Penicillium patulum*	33
Figure 2.5	Lovastatin biosynthesis pathway	35
Figure 2.6	Malonyl-CoA biosynthesis pathway in cytoplasm of *S. cerevisiae*	37
Figure 3.1	Pentose pathways	54
Figure 3.2	Synthetic genes were assembled and confirmed on agarose gel electrophoresis	71
Figure 3.3	Western blot of LXR1, LAD1 and ALX1 from synthetic genes	73
Figure 3.4	Expression level of LAD1 and ALX1 were compared on Western blot between native version and synthetic-gene-based version	75
Figure 3.5A	Xylitol production level per cell culture volume	79
Figure 3.5B	Xylitol production level per dry cell weight	79
Figure 3.6	Amino acid sequence alignment of Gal2 with other transporter proteins to identify ubiquitinated lysine residues in Gal2	83
Figure 3.7	Strategy to construct the Gal2/Hxt1 hybrid transporter	85
Figure 3.8	Dry cell weight for y812G, y812GH1, and y812GH2 with time	86
Figure 4.1	Vector maps for pJC and pI	101
Figure 4.2	SDS-PAGE showing LovB expression in BJ5464 cell extract using CEN/ARS and 2μ-based vectors BJ-CLovB and BJ-2LovB	110
Figure 4.3	Comparison of TE and LovB expression on SDS-PAGE	111

Figure 5.1	Diagram for 6-MSAS and npgA dual gene expression plasmid (pKUTP-6MN) construction via Gibson assembly	127
Figure 5.2	6-MSA synthesis in BYPN1 strains with *ACS1* or/and *ACC1* overexpression	140
Figure 5.3	Effect of *ACS* variants and *CAB1* overexpressions in BJ5464-based strain on 6-MSA titer	142
Figure 5.4	Improved 6-MSA synthesis by *ACC1* overexpression in BJPN1	143
Figure 5.5	Media optimization to prevent flocculation in *ACC1* overexpressed strain (BJPN1C-P6M)	145
Figure 5.6	Amino acid sequence alignment between rat and *S. cerevisiae* Acc1	148
Figure 5.7	Comparison of *in vitro* activity of Acc1	150
Figure 5.8	Production of 6-MSA and fatty acids *in vivo*	152
Figure 5.9	Diagram of pathways involving genes targeted for deletion	155
Figure 5.10	6-MSA production from OptKnock predicted deletions in modified 1% SDC medium	156
Figure 5.11	Comparison of 6-MSA production from different upstream pathway deletion gene strains in 3 different media	157
Figure 5.12	6-MSA titer after combined overexpressions and knockouts	159
Figure 6.1	UMP biosynthesis pathway consists of de novo pathway and salvage pathway	170
Figure 6.2	pIU13-6MSAS features and restriction sites for linearization	175
Figure 6.3	Diagram of pKA cloning	177
Figure 6.4	Effect of copy number and media on 6-MSA production	182
Figure 6.5	Comparison of standard *URA3* marker and N-degron *URA3* marker	184
Figure 6.6	Comparison of 6-MSA producing strains between *FUR1* (open) and *fur1* (closed)	186

LIST OF TABLES

		Page
Table 2.1	Redox imbalance summary	22
Table 3.1	pXP and pJC series vectors	57
Table 3.2	List of plasmids used in this study	59
Table 3.3	List of strains used in this study	62
Table 3.4	Purified enzyme activity assay	74
Table 3.5	Enzyme activity assay in cell extract for LAD1 and ALX1: synthetic-gene-based versus native	76
Table 3.6	Strains with different combinations of *cXYL1*, *clad1*, *clxr1* or *cALX1*	78
Table 4.1	List of plasmids and strains used for DML synthesis	103
Table 4.2	DML synthesis comparison	112
Table 5.1	List of plasmids constructed for the engineering of pyruvate dehydrogenase bypass	126
Table 5.2	List of strains for the engineering of pyruvate dehydrogenase bypass	131
Table 5.3	List of various SDC media used in this chapter	134
Table 5.4	Comparison of 6-MSA yield (mmol / mol substrate) during each growth phase in different media conditions	146
Table 5.5	Glucose and ethanol levels at harvest time for Acc1 activity assay	148
Table 5.6	List of gene deletions based on the enhanced TAL synthesis	154
Table 5.7	List of gene deletions predicted by OptKnock for enhanced 6-MSA synthesis	154
Table 6.1	List of plasmids and yeast strains	176

LIST OF ABBREVIATIONS

6-MSA	6-methylsalicylic acid
6-MSAS	6-methylsalicylic acid synthase
ACAS	atrochrysone carboxylic acid synthase (*Aspergillus terreus*)
Acc1	Acetyl-CoA carboxylase, *ACC1*
ACP	Acyl carrier protein
ACS	Acetyl-CoA synthetase
Acs1	Acetyl-CoA synthetase, *ACS1*
ACS_{SE}	Acetyl-CoA synthetase (*Salmonella enterica*), ACS_{SE}
Act1	Actin, *ACT1*
ACTE	Atrochryson carboxyl ACP thioesterase (*Aspergillus terreus*)
Adh1	Alcohol dehydrogenase, *ADH1*
Adh2	Alcohol dehydrogenase, *ADH2*
AI	L-arabinose isomerase
ALD	Acetaldehyde dehydrogenase
Ald6	Acetaldehyde dehydrogenase, *ALD6*
ALX1	L-xylulose reductase (*Ambrosiozyma monospora*), *ALX1*
AMPK	AMP-activated protein kinase
AptB	metallo-β-lactamase type thioesterase (*Aspergillus nidulans*), *AptB*
AraA	L-arabinose isomerase (AI, *Bacillus subtilis*), *araA*
AraB	L-ribulokinase (RK, *Escherichia coli*), *araB*
AraD	L-ribulose-5-P 4-epimerase (R5PE, *Escherichia coli*), *araD*
AT	Acyltransferase
Cab1	Pantothenate kinase, *CAB1*
Cyc1	Cytochrome c, isoform 1, *CYC1*
DH	Dehydratase

DML	Dihydromonacolin L
Fbp1	Fructose-1,6-bisphosphatase, *FBP1*
Fur1	Uracil phosphoribosyltransferase, *URA3*
Gal2	Galactose permease, *GAL2*
Gnd1	6-phosphogluconate dehydrogenase, *GND1*
Gpd1	Glycerol-3-phosphate dehydrogenase, *GPD1*
His3	Imidazoleglycerol-phosphate dehydratase, *HIS3*
Hor2	DL-glycerol-3-phosphate phosphatase, *HOR2*
Hpm3TE	Thioesterase domain of Hpm3
HXT	Hexose transporter
Hxt1	Low-affinity glucose transporter of the major facilitator superfamily, *HXT1*
Hxt2	High-affinity glucose transporter of the major facilitator superfamily, *HXT2*
Hxt5	Hexose transporter with moderate affinity for glucose, *HXT5*
Hxt6	High-affinity glucose transporter, *HXT6*
Hxt7	High-affinity glucose transporter, *HXT7*
Inm1	Inositol monophosphatase, *INM1*
Kex2	Subtilisin-like protease (proprotein convertase), *KEX2*
LacZ	β-galactosidase, *lacZ*
LAD	L-arabitol 4-dehydrogenase, *lad1*
LAD1	L-arabitol dehydrogenase (LAD, *Trichoderma reesei*), *lad1*
Leu2	β-isopropylmalate dehydrogenase (IMDH), *LEU2*
LNKS	Lovastatin nonaketide synthase
LovB	Lovastatin nonaketide synthase (*Aspergillus terreus*), *LovB*
LovC	Enoyl reductase (*Aspergillus terreus*), *LovC*
LovD	Transesterase (*Aspergillus terreus*), *LovD*
LovF	Lovastatin diketide synthase (*Aspergillus terreus*), *LovF*
LXR	L-xylulose reductase
LXR1	L-xylulose reductase (LXR, *Trichoderma reesei*), *lxr1*

KR	Ketoreductase
KS	Ketosynthase
MAT	Malonyl acyltransferase
Met17	O-acetyl homoserine-O-acetyl serine sulfhydrylase, *MET17*
MT	Methyl transferase
NpgA	4'-phosphopantetheinyl transferase (PPT, *Aspergillus nidulans*), *npgA*
PanK	Pantothenate kinase (*Escherichia coli*), *coaA*
Pgk1	3-phosphoglycerate kinase, *PGK1*
PKS	Polyketide synthase
PKS13TE	Thioesterase domain of PKS13 (*Gibberella zeae*)
Plb1	Phospholipase B, *PLB1*
Plb2	Phospholipase B, *PLB2*
PPT	4'-phosphopantetheinyl transferase
Pyc1	Pyruvate carboxylase, *PYC1*
Pyc2	Pyruvate carboxylase, *PYC2*
Rdc1TE	Thioesterase domain of Rdc1
Rhr2	DL-glycerol-3-phosphate phosphatase, *RHR2*
R5PE	L-ribulose-5-P 40epimerase
RK	L-ribulokinase
Snf1	AMP-activated serine/threonine protein kinase, *SNF1*
TAL	Triacetic acid lactone
TE	Thioesterase
Tef1	Translational elongation factor EF-1 alpha, *TEF1*
TM	Transmembrane segment
Tpo1	Polyamine transporter, *TPO1*
Trp1	Phosphoribosylanthranilate isomerase, *TRP1*
UMP	Uridine monophosphate
Ura3	Orotidine-5'-phosphate (OMP) decarboxylase, *URA3*

Urk1	Uridine kinase, *URK1*
VrtG	Thioesterase (*Penicillium aethiopicum*), *vrtG*
XDH	Xylitol dehydrogenase
XI	Xylose isomerase
XK	Xylulokinase, *XYL3*, *XKS1*
Xks1	Xylulokinase (XK), *XKS1*
XR	Xylose reductase, *XYL1*
XYL1	Xylose reductase (XR, *Pichia stipitis*), *XYL1*
XYL2	Xylitol dehydrogenase (XDH, *Pichia stipitis*), *XYL2*
XYL3	Xylulokinase (XK, *Pichia stipitis*), *XYL3*
XylA	Xylose isomerase (XI), *XylA*
Zwf1	Glucose-6-phosphate dehydrogenase, *ZWF1*

ACKNOWLEDGMENTS

I would like to express my deep appreciation and gratitude to my advisor, Prof. Nancy A. Da Silva, for the patient mentorship she provided to me. I would like to thank Prof. Szu-Wen Wang and Prof. Suzanne B. Sandmeyer for advices and being in my committee. I would like to thank Dr. Wesley Hatfield, Dr. Kirsty Salmon, Dr. Kimberly Aeling, Elaine Ito, Dr. Fang Fang and Becky Irwin for their guidance and help during first year of my Ph.D. I would like to thank Prof. Yi Tang in University of California at Los Angeles and his group members for being excellent collaborator. I would like to thank Prof. Reuben Shaw for helpful advices in Acc1 mutant work. I would also like to thank all previous and current Da Silva group members for good discussions. This research was supported by UC Discovery, Verdezyne, National Science Foundation (Award No. EEC-0813570), and National Institute of Health (NIGMS) (Award No. 1R01GM092217). I would like to acknowledge the Elsevier B.V. for permission to use the published material in a portion of this dissertation.

My PhD was not possible without the constant and unconditional support from my parents. I thank my girls, Olivia and Chloe for their laughs and love. But the biggest thank and heart goes to my wife Su. She was the one who was always beside me when I was happy, sad, frustrated, tired, and mad. Without the support of my family, I won't be here.

CURRICULUM VITAE

Jin Wook Choi

2005-2014 Ph.D. in Chemical and Biochemical Engineering

2004-2005 M.S. in Chemical and Biochemical Engineering

1994-2001 B.S. in Chemical Engineering, Han Yang University

WORK EXPERIENCE

2002 Coreana Cosmetics Co., Ltd.

SELECTED PUBLICATIONS AND PRESENTATIONS

1. Jin W. Choi, Nancy A. Da Silva, 2014. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-coa carboxylase. Jounal of Biotechnology, 187, 56-59

2. Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., Tang, Y., 2013. LovG: the thioesterase required for dihydromonacolin l release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. Angew Chem Int Ed Engl. 52, 6472-5. Angewandte Chemi

3. Jin W. Choi, Nancy A. Da Silva, 2012. Engineering of *Saccharomyces cerevisiae* for enhanced Polyketide Production. Abstracts of Papers, UKC, Anaheim

4. Jin W. Choi, Nancy A. Da Silva Jin W. Choi, Nancy A. Da Silva, 2012. Engineering of *Saccharomyces cerevisiae* for enhanced polyketide production. Abstracts of Papers of American Chemical Society, 243

5. Jin W. Choi, Christina Tran, Nancy A. Da Silva, 2011. The optimization of the production of fungal polyketides. Abstracts of Papers, UKC, Park City

6. Jin W. Choi, Nancy A. Da Silva, 2010. Ethanol production from *Saccharomyces cerevisiae* via a heterologous fungal arabinose pathway. Abstracts of Papers of American Chemical Society, 239

7. Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X. K., Kealey, J. T., Da Silva, N. A., Vederas, J. C., Tang, Y., 2009. Complete reconstitution of a highly reducing iterative polyketide synthase. Science. 326, 589-592.

ABSTRACT OF THE DISSERTATION

Engineering of Yeast for the Production of Ethanol and Polyketides

By

Jin Wook Choi

Doctor of Philosophy in Chemical and Biochemical Engineering

University of California, Irvine, 2014

Professor Nancy A. Da Silva, Chair

Saccharomyces cerevisiae is a promising microorganism for the production of ethanol for fuel, and the synthesis of precursors to industrial chemicals and natural products. The goal of this research was to engineer *S. cerevisiae* strains for the enhanced synthesis of these products. For the economical production of bioethanol, complete use of hemicellulosic sugars is necessary. Thus, the fungal arabinose pathway was imported and evaluated. The NADH-dependent L-xylulose reductase (ALX1) from *Ambrosiozyma monospora* was introduced to balance the use of redox cofactors. Three fungal arabinose pathway genes (*XYL1*, *lad1*, and *lxr1* or *ALX1*) were codon and codon pair optimized for expression in *S. cerevisiae*. Various combinations and copy numbers of the three required genes were evaluated by measuring growth and xylitol production; use of ALX1 resulted in up to 9-fold higher xylitol titers relative to LXR1. Arabinose uptake, the rate-limiting step for arabinose utilization in *S. cerevisiae*, was also addressed by creating a chimeric

protein of Gal2 and Hxt1. Polyketides are versatile molecules that can be industrial chemical precursors as well as drug precursors. Five different thioesterases were compared for the release of dihydromonacolin L (DML), the precursor to the cholesterol lowering agent lovastatin, from LovB in *S. cerevisiae*, and AptB was identified to be the best candidate. Most polyketides use acetate as starter unit and malonate or other malonate-based molecules as extender units. 6-methylsalicylic acid synthase (6-MSAS) was chosen as the model system for the engineering of *S. cerevisiae* to increase the intracellular availability of acetate and malonate. 6-MSA titer showed improvement by 50% via deletion of *PYC1*, by 3-fold via engineering of Acc1 (9-fold in activity), and by 90% via combined use of N-degron tagged *URA3* and autoselection for the expression of 6-MSAS in complex medium. The strategies developed in this research contribute to the metabolic engineering of *S. cerevisiae* for the synthesis of ethanol, and biochemical and drug precursors.

Chapter 1

Introduction

1.1. Motivation

The baker's yeast *Saccharomyces cerevisiae*, also used for brewing and wine making, is a well-established microorganism for both research and industrial applications. Such popularity can be attributed to the many advantages of this microorganism. The great number of genetic tools and robust fermentation technologies are very attractive features. For *S. cerevisiae*, highly efficient transformation methods have been developed (Gietz and Woods, 2001) and an array of promoters with varying strength are available (Blazeck et al., 2012; Fang et al., 2010; Hadfield et al., 1993; Mumberg et al., 1994; Partow et al., 2010; Shen et al., 2012). Different vectors enable fine tuning of copy number via CEN/ARS or 2μ versions or through chromosomal integrations (Da Silva and Srikrishnan, 2012; Romanos et al., 1992), and targeted integration into chromosomes using homologous recombination gives very stable expression of heterologous genes. The availability of many antibiotic or auxotrophic selection markers is also an advantage (Gueldener et al., 2002); other useful tools include reporter genes and immunotags (Janke et al., 2004). One important advantage that should be noted is the GRAS (Generally Regarded As Safe) status of *S. cerevisiae* by the U.S. Food and Drug Administration (FDA) (Bonekamp and Oosterom, 1994). Moreover, the genome of *S. cerevisiae* was the first completely sequenced eukaryotic genome and the data is easily accessible online (Goffeau et al., 1996). Additionally, a few detailed models are available that enable the prediction of promising metabolic engineering interventions (Duarte et al., 2004; Mo et al., 2009; Zomorrodi and Maranas, 2010). Many of the above advantages also hold for other microorganisms, such as the bacterium *Escherichia coli*. However, an advantage of *S. cerevisiae* is that it is a single-celled eukaryote. Many products require eukaryotic organelles such as the endoplasmic reticulum and Golgi complex for protein

modification and secretion (Buckholz and Gleeson, 1991). Given the many positive attributes, *S. cerevisiae* is a very important host and has seen applications in the synthesis of pharmaceutical drug precursors such as artemisinic acid (Ro et al., 2006), taxadiene (Engels et al., 2008), naringenin (Jiang et al., 2005), 6-MSA (Kealey et al., 1998), and dihydromonacolin L (Ma et al., 2009; Xu et al., 2013); the production of ethanol as fuel (Cai et al., 2012; Hahn-Hagerdal et al., 2007; Ho et al., 1999; Jeffries and Jin, 2004; Matsushika et al., 2009; Naik et al., 2010); and the generation of platform industrial chemicals from biorenewable sources (Nielsen et al., 2013; Nikolau, 2010; Nikolau et al., 2008).

The biosynthesis of ethanol in *S. cerevisiae* has a promising future. *S. cerevisiae* has a very high tolerance to ethanol (Piskur et al., 2006). Due to the advantages mentioned above and also because glucose to ethanol fermentation has been most effective in this microorganism (Ho et al., 1999)*, S. cerevisiae* has become a major host strain for bioethanol production. The use of pentoses as well as hexoses in *S. cerevisiae* can help industries achieve much higher yield from sugar sources such as corn stover and cob, wheat straw, and wood (Wyman, 2003). Furthermore, the increasing demand for transportation fuel and the limited supply of petroleum has led to a growing demand for fuel ethanol from biological sources (Ragauskas et al., 2006).

Pursuing sustainability is not limited to transportation fuel. Limited oil supplies will also affect industrial chemical production. Most platform chemicals such as ethylene and propylene come from petroleum (Nikolau et al., 2008), which takes millions of years to regenerate. Replacing petroleum with biological sources and synthesizing platform chemicals from simple sugars such as hexoses and pentoses that can be obtained from agricultural products and

byproducts will be very important and also lucrative. A major emphasis of the NSF Engineering Research Center: Center for Biorenewable Chemicals (CBiRC) is to engineer microorganisms to produce an array of platform chemicals that can undergo further chemical modification to produce desired industrial chemicals (Nikolau et al., 2008). These platform chemicals include short chain carboxylic acids and pyrones, which can be synthesized by utilizing fatty acid synthases and other polyketide synthases, respectively. Introducing various polyketide synthases or domains into the fatty acid synthesis system, which is a type of polyketide synthase, will enable diversification of the library of platform chemicals (Nikolau et al., 2008). Engineering yeast for the high-level production of diverse polyketide synthases and products is essential.

Application of polyketide synthases for the production of biochemicals is a relatively recent endeavor. Natural polyketides have long been recognized as an excellent source of drug candidates for use as pharmaceuticals. More than 20 commercial drugs have come from 7000 known polyketide structures. It is estimated that more than 99% of bacteria that exist on earth cannot be cultured in the laboratory and the possible number of existing microorganisms is 3.7 x 10^{30} in the marine environment alone (Kennedy et al., 2008). Many of these might contain novel polyketide molecules, which could be potential sources for new drug candidates (Li and Vederas, 2009). Further understanding of new polyketide synthases as well as currently known ones brings opportunities for combinatorial biosynthesis, which aims to produce novel chemical entities with improved biological activities by mixing domains and assembling enzymes from well-known systems (Zhang and Tang, 2008). To maximize the opportunity from hidden or currently unreachable natural sources and make the most out of already discovered natural or natural product-derived novel biochemical pathways, securing sound microbial platforms for high-level

expression of the heterologous genes is essential. Introducing a heterologous gene brings difficulties such as adaptation of codon and codon context usage, and consideration of toxicity to the heterologous host (Li and Vederas, 2009). These problems could be solved by employing the metabolic engineering tools available in model microorganisms such as *E. coli* and *S. cerevisiae*. Building a metabolic engineering framework that encompasses combinatorial biosynthesis and heterologous gene expression for the production of natural product derived drugs would be advantageous.

Examples of commercialized polyketides include antimicrobials such as erythromycin, rifamycin, tetracycline; antifungals such as amphotericin B; immunosuppressant such as tacrolimus [TK506] and rapamycin; and anticancer agents such as doxorubicin, epothilone, and geldanamycin (Panagiotou et al., 2009). Another well-known example from natural sources is the cholesterol lowering drug, lovastatin (mevinolin, MEVACOR) (Alberts, 1988). Cholesterol biosynthesis involves reduction of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) into mevalonate by HMG-CoA reductase and this is a major rate-limiting step. An effective inhibitor of the HMG CoA reductase is the fungal metabolite lovastatin, which is produced from a filamentous fungus *Aspergillus terreus*. Moreover, the methylated analog of lovastatin, simvastatin (Zocor) has improved drug properties and is a blockbuster drug with sales of over $4.3 billion in 2006 prior to the loss of patent protection (Li and Vederas, 2009). However, current production of lovastatin or simvastatin is done in *A. terreus*. Only recently, one of lovastatin precursors was synthesized in *S. cerevisiae* (Xu et al., 2013). There are still obstacles to overcome that include *in vivo* synthesis of the final product, lovastatin toxicity to *S. cerevisiae*, and increasing lovastating titers. Solving these problems will lead to more efficient and economical

production of lovastatin or simvastatin and provide a potential platform technology for the synthesis of other polyketides.

The yeast *S. cerevisiae* holds great potential for polyketide synthesis, biorenewable chemical precursor production, and ethanol production from non-utilizable carbon sources. Furthermore, knowledge of various polyketide synthases, fatty acid synthesis, carbon utilization pathways, and efficient cultivation strategies can contribute to the development of sustainable energy, chemical, and pharmaceutical technologies. Coordinated development of yeast for ethanol and polyketide production will thus be beneficial in building a sustainable and healthy global community.

1.2. Objectives

The goal of this dissertation was to develop a systematic approach to engineering yeast metabolism for the synthesis of non-native products or enhanced production of native products under the following two product categories: ethanol (as a transportation fuel) and polyketides (as pharmaceuticals or biorenewable chemical precursors). The production of fuel ethanol has been heavily researched mainly for the utilization of hexoses and xylose as carbon sources. The first objective of this dissertation was to engineer *S. cerevisiae* for arabinose uptake and utilization. Use of all pentose sugars is essential for the cost-efficient production of ethanol. The second objective was to engineer *S. cerevisiae* for the high-level production of polyketides. Polyketides are synthesized by one or more enzymes often from multiple units of the same substrates. These enzymes are often non-native to *S. cerevisiae* but can utilize native substrates such as acetyl-CoA and malonyl-CoA. Heterologous expression of these exogenous enzymes and generating plenty of substrates are crucial for the maximum production of the final products in *S. cerevisiae*.

Objective 1. To engineer *S. cerevisiae* for arabinose uptake and utilization. Specific objectives were:

1) To construct and evaluate the initial steps of the fungal arabinose assimilation pathway in *S. cerevisiae*.
2) To engineer a more stable arabinose transporter in *S. cerevisiae*.

Objective 2. To engineer *S. cerevisiae* for the high-level production of polyketides. Specific objectives were:

1) To evaluate thioesterases for the efficient release of dihydromonacolin L (DML) from LovB.
2) To improve precursor availability for the high-level production of 6-methylsalicylic acid (6-MSA) via engineering of Acc1 and the upstream carbon metabolism pathway.
3) To engineer a *S. cerevisiae* strain for high-level expression of 6-MSA synthase (6-MSAS) and improved 6-MSA synthesis via expression system engineering.

1.3. References

Alberts, A. W., 1988. Discovery, biochemistry and biology of lovastatin. American Journal of Cardiology. 62**,** J10-J15.

Blazeck, J., Garg, R., Reed, B., Alper, H. S., 2012. Controlling promoter strength and regulation in Saccharomyces cerevisiae using synthetic hybrid promoters. Biotechnology and Bioengineering. 109**,** 2884-2895.

Bonekamp, F. J., Oosterom, J., 1994. On the safety of *Kluyveromyces lactis* - a review. Applied Microbiology and Biotechnology. 41**,** 1-3.

Buckholz, R. G., Gleeson, M. A. G., 1991. Yeast systems for the commercial production of heterologous proteins. Bio-Technology. 9**,** 1067-1072.

Cai, Z., Zhang, B., Li, Y., 2012. Engineering Saccharomyces cerevisiae for efficient anaerobic xylose fermentation: Reflections and perspectives. Biotechnology Journal. 7**,** 34-46.

Da Silva, N. A., Srikrishnan, S., 2012. Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. Fems Yeast Research. 12**,** 197-214.

Duarte, N. C., Herrgard, M. J., Palsson, B. O., 2004. Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model. Genome Research. 14**,** 1298-1309.

Engels, B., Dahm, P., Jennewein, S., 2008. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. Metabolic Engineering. 10**,** 201-206.

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2010. A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. n/a-n/a.

Gietz, R. D., Woods, R. A., 2001. Genetic transformation of yeast. Biotechniques. 30**,** 816-+.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S. G., 1996. Life with 6000 genes. Science. 274**,** 546-&.

Gueldener, U., Heinisch, J., Koehler, G. J., Voss, D., Hegemann, J. H., 2002. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Research. 30**,** 8.

Hadfield, C., Raina, K. K., Shashimenon, K., Mount, R. C., 1993. The expression and performance of cloned genes in yeasts. Mycological Research. 97**,** 897-944.

Hahn-Hagerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I., Gorwa-Grauslund, M. F., 2007. Towards industrial pentose-fermenting yeast strains. Applied Microbiology and Biotechnology. 74**,** 937-953.

Ho, N. W. Y., Chen, Z., Brainard, A. P., Sedlak, M., 1999. Successful design and development of genetically engineered Saccharomyces yeasts for effective cofermentation of glucose and xylose from cellulosic biomass to fuel ethanol. Advances in Biochemical Engineering Biotechnology; Recent progress in bioconversion of lignocellulosics. 65**,** 163-192.

Janke, C., Magiera, M. M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., Knop, M., 2004. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast. 21**,** 947-962.

Jeffries, T. W., Jin, Y. S., 2004. Metabolic engineering for improved fermentation of pentoses by yeasts. Applied Microbiology and Biotechnology. 63**,** 495-509.

Jiang, H. X., Wood, K. V., Morgan, J. A., 2005. Metabolic engineering of the phenylpropanoid pathway in Saccharomyces cerevisiae. Applied and Environmental Microbiology. 71**,** 2962-2969.

Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., Barr, P. J., 1998. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. Proceedings of the National Academy of Sciences of the United States of America. 95**,** 505-509.

Kennedy, J., Marchesi, J. R., Dobson, A. D. W., 2008. Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. Microbial Cell Factories. 7**,** 8.

Li, J. W. H., Vederas, J. C., 2009. Drug discovery and natural products: end of an era or an endless frontier? Science. 325**,** 161-165.

Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X. K., Kealey, J. T., Da Silva, N. A., Vederas, J. C., Tang, Y., 2009. Complete reconstitution of a highly reducing iterative polyketide synthase. Science. 326**,** 589-592.

Matsushika, A., Inoue, H., Kodaki, T., Sawayama, S., 2009. Ethanol production from xylose in engineered Saccharomyces cerevisiae strains: current state and perspectives. Applied Microbiology and Biotechnology. 84**,** 37-53.

Mo, M. L., Palsson, B. O., Herrgard, M. J., 2009. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. Bmc Systems Biology. 3.

Mumberg, D., Muller, R., Funk, M., 1994. Regulatable promoters of Saccharomyces cerevisiae - comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Research. 22**,** 5767-5768.

Naik, S. N., Goud, V. V., Rout, P. K., Dalai, A. K., 2010. Production of first and second generation biofuels: A comprehensive review. Renewable & Sustainable Energy Reviews. 14**,** 578-597.

Nielsen, J., Larsson, C., van Maris, A., Pronk, J., 2013. Metabolic engineering of yeast for production of fuels and chemicals. Current Opinion in Biotechnology. 24**,** 398-404.

Nikolau, B. J., 2010. An integrated strategy for generating lipid-based biorenewable chemicals: diversifying fatty acid synthesis with polyketide synthesis biocatalysts. Chemistry and Physics of Lipids. 163**,** S16-S17.

Nikolau, B. J., Perera, M., Brachova, L., Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. Plant Journal. 54**,** 536-545.

Panagiotou, G., Andersen, M. R., Grotkjaer, T., Regueira, T. B., Nielsen, J., Olsson, L., 2009. Studies of the production of fungal polyketides in Aspergillus nidulans by using systems biology tools. Applied and Environmental Microbiology. 75**,** 2212-2220.

Partow, S., Siewers, V., Bjorn, S., Nielsen, J., Maury, J., 2010. Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. Yeast. 27**,** 955-964.

Piskur, J., Rozpedowska, E., Polakova, S., Merico, A., Compagno, C., 2006. How did Saccharomyces evolve to become a good brewer? Trends in Genetics. 22**,** 183-186.

Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C. A., Frederick, W. J., Hallett, J. P., Leak, D. J., Liotta, C. L., Mielenz, J. R., Murphy, R., Templer, R., Tschaplinski, T., 2006. The path forward for biofuels and biomaterials. Science. 311**,** 484-489.

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., Keasling, J. D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature. 440**,** 940-943.

Romanos, M. A., Scorer, C. A., Clare, J. J., 1992. Foreign gene expression in yeast - a review. Yeast. 8**,** 423-488.

Shen, M. W. Y., Fang, F., Sandmeyer, S., Da Silva, N. A., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 29**,** 495-503.

Wyman, C. E., 2003. Potential synergies and challenges in refining cellulosic biomass to fuels, chemicals, and power. Biotechnology Progress. 19**,** 254-262.

Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., Tang, Y., 2013. LovG: The thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. Angew Chem Int Ed Engl. 52**,** 6472-5.

Zhang, W. J., Tang, Y., 2008. Combinatorial biosynthesis of natural products. Journal of Medicinal Chemistry. 51**,** 2629-2633.

Zomorrodi, A. R., Maranas, C. D., 2010. Improving the iMM904 S. cerevisiae metabolic model using essentiality and synthetic lethality data. Bmc Systems Biology. 4.

Chapter 2.

Literature Review

2.1. Metabolic engineering of *Saccharomyces cerevisiae*

2.1.1. Significance of using *S. cerevisiae* for the production of bioethanol, industrial chemicals, and polyketides

In addition to the advantages of using *S. cerevisiae* in both research and industrial applications reviewed in Chapter 1, *S. cerevisiae* is tolerant to acidic culture conditions unlike the popular host *E. coli*, which prefers neutral pH (Lin and Tanaka, 2006). *S. cerevisiae* can produce ethanol up to 18% (by volume) of the fermentation broth and is tolerant to the toxic components in lignocellulosic hydrolysates (Clark and Mackie, 1984; Dien et al., 2003; Dupreez, 1994; Lin and Tanaka, 2006; McMillan, 1994). Thus, *S. cerevisiae* has been used for cellulosic biofuel production and has advantages over other microorganisms as a heterologous host strain for ethanol production from hemicellulose (Hahn-Hagerdal et al., 2007). These advantages include (1) *S. cerevisiae* has been the choice for ethanol production as long as human history, and (2) many industries are already equipped with fermentation facilities for this microorganism and new developments in using pentose for ethanol fermentation in this yeast can be easily adapted into the existing facilities. The reduction in operating cost coming from the second advantage is estimated to be up to 20 % (Wooley et al., 1999). Therefore, *S. cerevisiae* is a very good host strain for bioethanol production using different sugar sources.

S. cerevisiae is also an appropriate host for heterologous expression of polyketide synthases, particularly fungal polyketide synthases. Polyketides are promising precursors to both industrial chemicals and pharmaceuticals (Cardenas and Da Silva, 2014; Pickens et al., 2011).

Synthesis of polyketides in heterologous hosts (such as *S. cerevisiae* and *E. coli*) holds many advantages due to the availability of genetic tools and the ease of cultivation relative to the native hosts (Romanos et al., 1992). However, high-level synthesis can still be difficult to achieve. Dihydromonacolin L, the lovastatin precursor, has been only synthesized in *Aspergillus terreus*, *Aspergillus nidulans*, and *S. cerevisiae* (Kennedy et al., 1999; Xu et al., 2013). One problem was the expression of LovB, the 335 kDa Lovastatin nonaketide synthase, in heterologous hosts. The successful expression of this protein in yeast was possible only after using a protease deficient strain, BJ5464 (*MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1Δ1.6R can1 GAL*) as the host (Lee, 2006; Lee et al., 2009). Even successful synthesis of these molecules does not lead to industrial production due to low quantity. 6-methylsalicylic acid (6-MSA) has been produced at 1.7g/L in *S. cerevisiae*, while *Streptomyces coelicolor* and *E. coli* produced only 60-75 mg/L and *A. nidulans* produced 300 mg/L (Fujii et al., 1996; Kealey et al., 1998; Pickens et al., 2011). Recently, triacetic acid lactone (TAL) has been produced at 2.3 g/L in *S. cerevisiae* (Cardenas and Da Silva, 2014); this compound is quite to *E. coli*. These examples show the exceptional potential of *S. cerevisiae* as a heterologous host for polyketide synthesis.

2.1.2. Use of versatile vector series, pXPs

A series of vectors (pXP series) was created by our lab in collaboration with Professor Suzanne Sandmeyer's lab (UCI) (Fang et al., 2011; Shen et al., 2012). These vectors were designed as autonomous yeast plasmids and as templates for subsequent chromosomal integration (Figure

2.1). pXP vectors have *TEF1*, *HXT7*-391, *PGK1*, *ADH2*, *GAL1*, and *CUP1* promoters; *URA3*, *TRP1*, *MET15*, *LEU2-d8*, *HIS3* and *CAN1* as auxotrophic markers; and 2μ and CEN/ARS yeast replication origins. For integration, primers can be designed with sequences homologous to the genome, and integration fragments can be generated by PCR from pXP vectors using these primers. This PCR product integrates at the specific target locus on the *S. cerevisiae* genome by homologous recombination. Efficient homologous recombination in *S. cerevisiae* allows one, two, or more pieces to be integrated simultaneously (Figure 2.1). Two-piece integration allows shorter PCR products and reduced probability of error. After integration, the selection marker should be removed from the genomic DNA so that the same markers can be reused. pXP vectors contain two *lox*P sequences flanking yeast auxotrophic markers for marker removal. These two *lox*P sequences can recombine together when CreA is expressed in the same cell (Gueldener et al., 2002; Guldener et al., 1996; Sauer, 1994). A CreA vector is transformed into the cell, expressed to remove the *lox*P-marker-*lox*P fragment, and cured from the cell by growing cells on non-selective plate. Alternatively, direct selection against certain auxotrophic markers can be used.

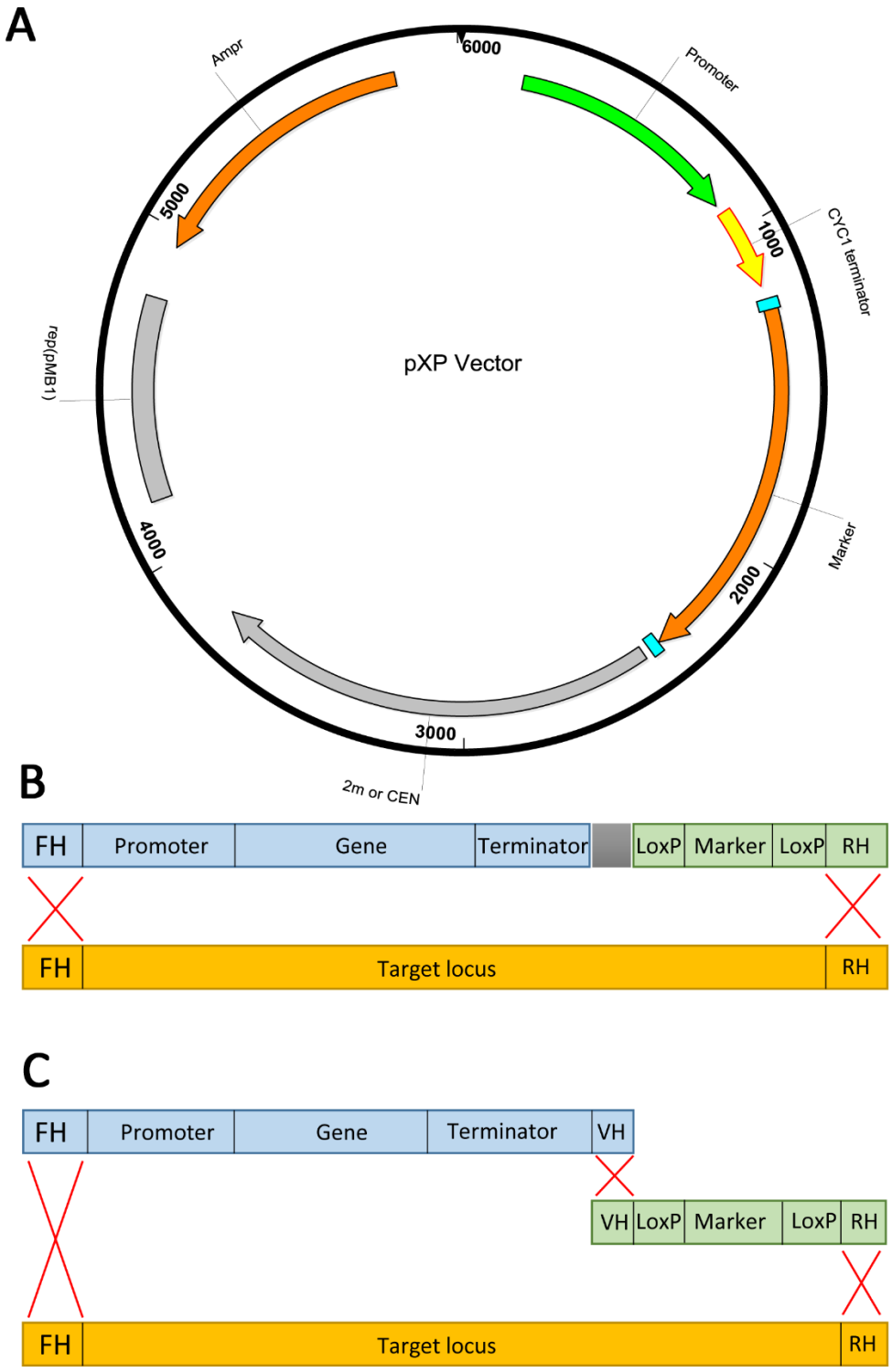


Figure 2.1 pXP vector series. Figure was adapted from Fang et al. (2011). A: Vector map for pXP vector. B: The integration of one PCR fragment via doublecrossover is shown. C: The integration of two PCR fragments via doublecrossover is shown. FH, front homology; VH, vector homology; RH, rear homology

2.2. Bioethanol production in *Saccharomyces cerevisiae*

2.2.1. Pentose assimilation pathway

2.2.1.1. Pentose transport in Saccharomyces cerevisiae

Native *S. cerevisiae* cannot utilize pentose sugars. However, pentoses in the culture medium can still be imported into *S. cerevisiae* through membrane-bound transporter proteins. Xylose is imported through most glucose transporters; however the affinity for xylose is over 200-fold lower than for glucose (Lee et al., 2002). Among hexose transporters, Hxt4, Hxt5, Hxt7 and Gal2 are the most relevant to xylose transport (Leandro et al., 2009). Unlike xylose, there is only one known transporter for L-arabinose in *S. cerevisiae*, galactose permease (Gal2) (Kou et al., 1970). When *GAL2* was introduced on a CEN/ARS plasmid (Liang and Gaber, 1996) and expressed in *S. cerevisiae*, a 32-fold increase in L-arabinose transport was observed (Becker and Boles, 2003). Overexpression of *GAL2* also slightly improved the growth rate on L-arabinose as a sole carbon source (Becker and Boles, 2003). However, Gal2 has significantly lower affinity toward L-arabinose relative to glucose and galactose. In addition, *GAL2* transcription is repressed and the Gal2 transporter is inactivated by high concentrations of glucose (Ozcan and Johnston, 1999).

Recently, Londesborough et al. expressed *LAT1* and *LAT2*, arabinose transporters from *Ambrosiozyma monospora*, in *S. cerevisiae* with no significant improvement in arabinose uptake relative to the control strain that does not have *LAT1* or *LAT2* (Londesborough et al., 2014). Interestingly, Lat1 tagged with green fluorescent protein was able to transport three-fold more

arabinose than Lat1 without GFP. The authors suspected that the C-terminus of Lat1 might be subject to inactivation that might be avoided by a C-terminal tag.

2.2.1.2. Xylose pathway

Utilization of xylose is observed in bacteria and fungi as well as in certain yeasts. Two pathways have been successfully expressed in *S. cerevisiae*. In the first pathway (from bacteria), D-xylose is converted to D-xylulose by xylose isomerase (XI, *XylA*) (Figure 2.2) (Bruinenberg et al., 1983). D-xylulose is then phosphorylated and transferred into the pentose phosphate pathway (Figure 2.2 and 2.3). Xylose isomerases have been identified from *Thermus thermophiles* (Walfridsson et al., 1996), *Clostridium phytofermentans* (Brat et al., 2009), *Bacteroides stercoris* (Ha et al., 2011) or anaerobic fungi such as *Piromyces* sp. E2 (Karhumaa et al., 2007; Kuyper et al., 2005) and *Orpinomyces* sp. (Madhavan et al., 2009). This pathway does not require redox cofactors (Kuyper et al., 2004).

In the second pathway (from fungi) (Figure 2.2A), D-xylose is reduced to xylitol in a reaction catalyzed by xylose reductase (XR) (Kim et al., 2013). Xylitol is then oxidized to D-xylulose by xylitol dehydrogenase (XDH). D-xylulose is phosphorylated to produce D-xylulose-5-phosphate by xylulokinase (XK), which enters the pentose phosphate pathway. The fungal xylose pathway has one reduction and one oxidation step. Commonly used xylose pathway enzymes in *S. cerevisiae* for xylose assimilation are XR (*XYL1*), XDH (*XYL2*), and XK (*XYL3*) from *Pichia stipitis* and *XKS1* from *S. cerevisiae* (Jin et al., 2002; Kotter et al., 1990). NADPH is preferred by XR (*P. stipitis*)

as a cofactor with 3000-fold lower Km for NADPH than NADH, while XDH (*P. stipitis*) is NAD^+-specific (Verduyn et al., 1985).

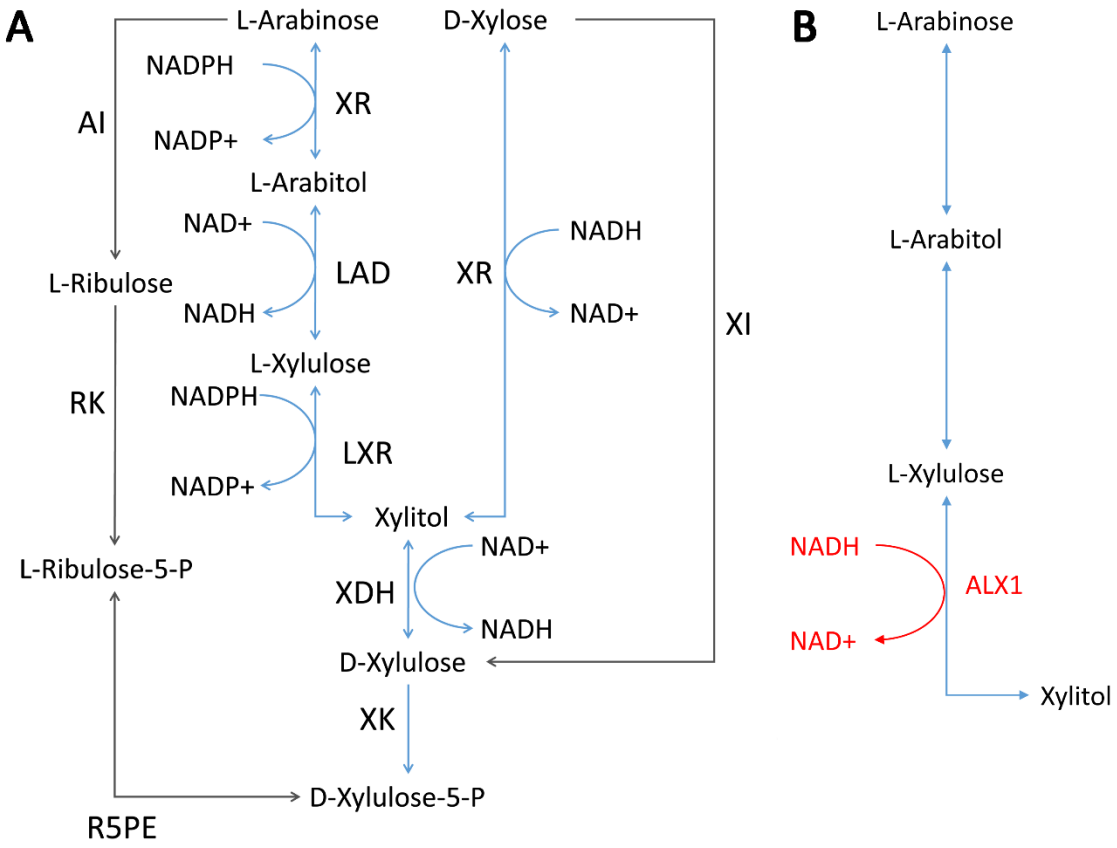


Figure 2.2 Pentose pathways. XR, xylose reductase (*XYL1*) from *P. stipitis*; LAD, L-arabitol 4-dehydrogenase (*lad1*) from *Trichoderma reesei*; LXR, L-xylulose reductase (*lxr1*) from *T. reesei*; XDH, xylitol dehydrogenase, (*XYL2*) from *P. stipitis*; XK, xylulokinase (*XYL3* from *P. stipitis* or *XKS1* from *S. cerevisiae*); XI, xylose isomerase (*XylA*); AI, L-arabinose isomerase (*araA*) from *Bacillus subtilis*; RK, ribulokinase (*araB*) from *E. coli*; R5PE, ribulose-5-phosphate-4-epimerase (*araD*) from *E. coli*. (A) Pathway contains NADPH-dependent LXR1 (*lxr1*) from *T. reesei*, (B) Pathway contains NADH-dependent ALX1 (*ALX1*) from *A. monospora*

The redox imbalance coming from XR and XDH *in S. cerevisiae* is relieved to some extent if the host is grown aerobically. However, it is hard to compensate when grown anaerobically

because of excess NADH caused by the inability to reoxidize NADH under anaerobic conditions. Efforts to avoid the redox imbalance in the fungal pathway have been published. For example, Watanabe *et al.* (2007) generated mutated *P. stipitis* XR and XDH enzymes to reverse their cofactor preferences from NADPH to NADH and NAD+ to NADP+ (Watanabe et al., 2005) respectively. It has been also shown that a copy number increase in *XKS1,* the xylulokinase from *S. cerevisiae*, also helps efficient anaerobic fermentation (Moniruzzaman et al., 1997). Xylose isomerases do not require cofactors (Amore et al., 1989; Wilhelm and Hollenberg, 1984) and the redox imbalance is avoided. However, the kinetics of XI is thermodynamically inferior to the XR/XDH system (Karhumaa et al., 2007) and most XI from heterologous sources have not been functional in *S. cerevisiae*. The XI from *Thermus thermophilus* was functional, but had very low activity and led to much lower growth on xylose relative to the XR/XDH pathway due to limited XI activity (Karhumaa et al., 2005; Walfridsson et al., 1996). Another XI from fungi, *Pyromyces* sp., was introduced into *S. cerevisiae* and led to higher activity and better growth on xylose (Harhangi et al., 2003) relative to the bacterial XIs. The stoichiometry of xylose conversion to ethanol is as follows (Jin and Jeffries, 2004).

$$\text{Xylose} + 0.5\ O_2 \rightarrow 1.5\ \text{Ethanol} + 2.0\ CO_2 \qquad \text{(by using XR and XDH)}$$

$$3\ \text{Xylose} \rightarrow 5\ \text{Ethanol} + 5\ CO_2 \qquad \text{(by using XI)}$$

The theoretical yields are 0.51 g ethanol/ g xylose and 0.46g ethanol / g xylose with the XI and XR/XDH pathway, respectively. The difference in the yield mainly arises from the employment of oxygen and production of CO_2. However, even though XI has a higher theoretical

yield and does not suffer from redox imbalance, the Pyromyces XI showed very low tolerance to fermentation inhibitors from toxic hydrolysates (Karhumaa et al., 2007). On the other hand, the XR-XDH system benefits from the furfural, a byproduct resulting from pyrolysis of biomass, because this can be used as an electron acceptor replacing oxygen under anaerobic condition (Madhavan et al., 2012; Palmqvist and Hahn-Hagerdal, 2000; Wahlbom and Hahn-Hagerdal, 2002). This will partially solve the redox imbalance problem and reduce xylitol accumulation. Therefore, the XR-XDH system has advantages over the XI system considering the industrial production of ethanol directly from toxic hydrolysates.

2.2.1.3. Arabinose pathway

In bacteria, L-arabinose is converted to l-ribulose by L-arabinose isomerase (AI, *araA*), then to l-ribulose-5-phosphate by l-ribulokinase (RK, *araB*), and eventually to D-xylulose 5-phosphate by l-ribulose-5-P 4-epimerase (R5PE, *araD*) (Figure 2.2A) (Lee et al., 1986). The downstream pathway is the same as in the xylose pathway.

In fungi, L-arabinose is reduced to l-arabitol by xylose reductase (XR) (Figure 2.2A). Then, l-arabitol is oxidized to L-xylulose by l-arabitol 4-dehydrogenase (LAD). L-xylulose is reduced further to xylitol by L-xylulose reductase (LXR). Xylitol follows the same fate as in the fungal xylose pathway. The fungal arabinose pathway has two reduction and two oxidation steps. In both reductions, xylose reductase and L-xylulose reductase reduce L-arabinose and L-xylulose, respectively, using NADPH as a cofactor. Some XR enzymes can utilize both NADH and NADPH although NADPH is preferred. In both oxidations, LAD and XDH oxidize l-arabitol and xylitol,

respectively, using NAD^+ as a cofactor. As a result, in general $NADP^+$ and NADH are generated while NADPH and NAD^+ are consumed. The redox imbalance problem is summarized in Table 2.1. There is also an L-xylulose reductase from *Ambrosiozyma monospror* (ALX1) that strictly utilizes NADH instead of NADPH (Figure 2.2B).

Table 2.1 Redox imbalance summary

***T. reesei* system**		***A. monospora* system**	
XR	NADPH → $NADP^+$	XR	NADPH → $NADP^+$
LAD1	NAD^+ → NADH	LAD1	NAD^+ → NADH
LXR1	NADPH → $NADP^+$	ALX1	NADH → NAD^+
XDH	NAD^+ → NADH	XDH	NAD^+ → NADH

This redox imbalance is avoided in the bacterial arabinose pathway. The use of the bacterial arabinose pathway with a mixed carbon source (xylose and arabinose) requires addition of the xylose pathway. When the fungal xylose pathway is included with the bacterial arabinose pathway, an inhibitor for bacterial enzymes, l-arabitol is accumulated (Karhumaa et al., 2006). Thus, the fungal xylose pathway is not desired for application with the bacterial arabinose pathway. The bacterial xylose pathway has been used with the bacterial arabinose pathway. This strategy was successful in utilizing mixed carbon sources that include both xylose and arabinose only after a careful and extensive evolutionary approach (Wisselink et al., 2009).

The effects of toxic hydrolysate on XI and the partial relief of the redox imbalance by furfural point to employing the XR-XDH system over the XI system. An ideal partner to the XR-

XDH system is the fungal arabinose pathway since these two pathways share enzymes while the bacterial arabinose pathway is inhibited by arabitol. Therefore, using the fungal xylose pathway together with fungal arabinose pathway will be advantageous.

2.2.1.4. Pentose phosphate pathway

The pentose phosphate pathway consists of oxidative and nonoxidative pathways (Figure 2.3) (Jeffries and Jin, 2004). The oxidative pathway pushes carbon flux towards xylulose-5-phosphate and the nonoxidative phase pushes xylulose-5-phosphate toward fructose-6-phosphate and glyceraldehyde-3-phosphate. In the oxidative pentose phosphate pathway, *ZWF1* is known to be responsible for the production of NADPH in *S. cerevisiae* along with *ALD6*, which is a branch of the ethanol fermentation pathway (Grabowska and Chelstowska, 2003). In the nonoxidative phase, overexpressed transketolase (*TKL1*) together with transaldolase (*TAL1*), D-ribulose-5-phosphate 3-epimerase (*RPE1*), and ribose-5-phosphate ketol-isomerase (*RKI1*) improved growth rate when host cells were cultivated on xylose as a sole carbon source (Kuyper et al., 2005). Overexpression of transaldolase (*TAL1*) was found to help the growth and ethanol production with L-arabinose as a sole carbon source (Becker and Boles, 2003). Tal1 protein level is significantly lower when *S. cerevisiae* is grown anaerobically on a limited supply of glucose compared to when it is grown aerobically (Bruckmann et al., 2009).

It has been reported that the decrease in NADPH level stimulates Zwf1 activity (Llobell et al., 1988). Thus, it might be possible that use of NADPH in the unbalanced fungal pentose utilization pathway favors the oxidative pentose phosphate pathway relative to the nonoxidative pathway,

leading to the accumulation of pentose phosphate pathway intermediates. Moreover, the oxidative phase loses one carbon atom as a CO_2 molecule by Gnd1. The overexpression of enzymes of the non-oxidative pathway (Tal1, Tkl1, Rpe1, and Rki1) might be beneficial in returning pentose phosphate pathway intermediates into glycolysis (Becker and Boles, 2003; Kuyper et al., 2005). A balanced pentose utilization pathway will not need Zwf1 to be stimulated. Therefore, a balanced pentose utilization pathway and the non-oxidative part of the pentose phosphate pathway are desired for higher ethanol production.

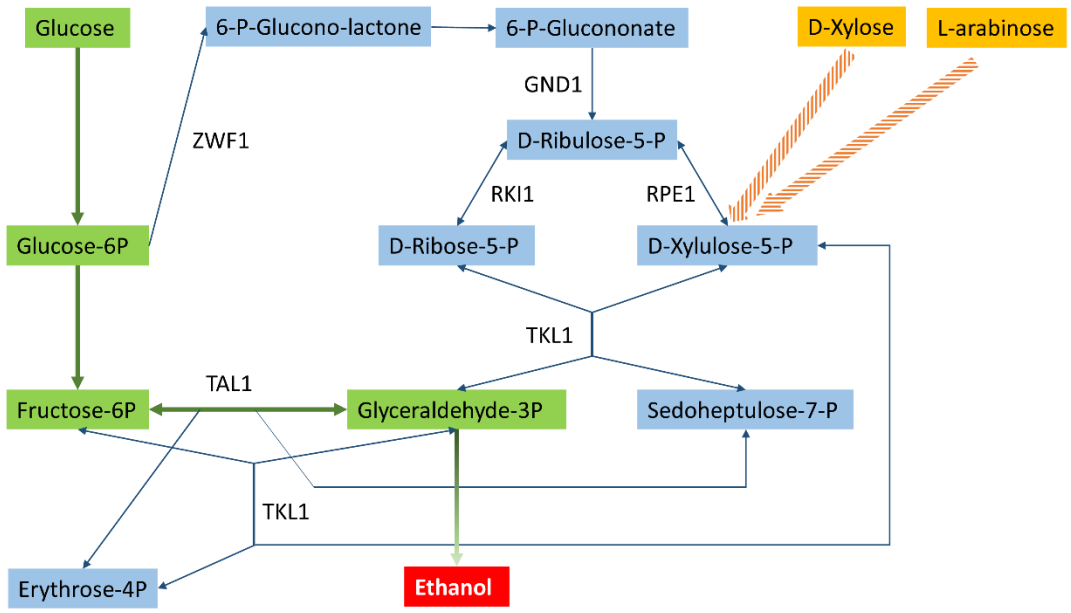


Figure 2.3 Pentose metabolism in recombinant yeast. Figure was adapted from Becker and Boles (2003). Glycolysis (Green), Pentose pathway (Orange), and *S. cerevisiae* pentose phosphate pathway (Blue) are shown.

2.3. Polyketide synthesis in *Saccharomyces cerevisiae*

2.3.1. Significance of polyketides and their synthesis in *Saccharomyces cerevisiae*

Polyketides are secondary metabolites polymerized from short-chain carboxylic acid units such as acetate, propionate, and butyrate (Chooi and Tang, 2012). The best known application of polyketides is as pharmaceutically active molecules. However, recently, the versatility of polyketide molecules attracted other industrial applications such as precursors to industrial platform chemicals, allowing a biorenewable alternative to petroleum feedstock (Nikolau et al., 2008).

2.3.1.1. Polyketides as a source of drug molecules

There are 7000 known polyketide structures and more than 20 commercial drugs including lovastatin have come from that pool (Li and Vederas, 2009). This 0.3% “hit rate” is far better than the <0.001% “hit rate” by high-throughput screening of synthetic compound libraries (Weissman and Leadlay, 2005). The structural complexity of polyketides makes chemical synthesis for large-scale production difficult (McDaniel et al., 2001). However, most native polyketide producing organisms are difficult to use for polyketide production due to very poor growth and lack of tools for genetic manipulation (Mutka et al., 2006). *S. cerevisiae* is an excellent candidate for heterologous expression of polyketide enzymes, particularly those of fungal origin. Strengths include its completely sequenced genome, a vast library of genomically specified strains, various antibiotic and auxotrophic selection markers, efficient homologous

recombination methods, immunotags, relatively fast growth rate, and GRAS (Generally Regarded As Safe) status. Examples of commercialized polyketides include antimicrobials such as erythromycin, rifamycin, and tetracycline; antifungals such as amphotericin B; immunosuppressant such as tacrolimus [TK506], rapamycin; and anticancer agents such as doxorubicin, epothilone and geldanamycin (Panagiotou et al., 2009). Moreover, 6-methylsalicylic acid (6-MSA), which is a polyketide metabolite naturally produced by *Penicillium patulum*, was produced in *S. cerevisiae* at up to 1.7 g/L quantities (Kealey et al., 1998).

Another example of a successful polyketide is the cholesterol-lowering drug lovastatin from *A. terreus*. Lovastatin is a precursor of the multi-billion dollar drug Zocor (simvastatin), which was obtained by a simple chemical modification of the side chain of lovastatin (Burr et al., 2007). Lovastatin inhibits HMG-CoA reductase at an early step of cholesterol biosynthesis (Alberts, 1988). The lovastatin variant, simvastatin, was Merck's largest selling drug and had $4 billion annual sales until 2006, when they lost U.S. patent protection (Maggon, 2005).

2.3.1.2. Replacing carbon from fossil

For more than one hundred years, mankind has been increasingly dependent on energy and chemicals from fossil-fuel-based sources (Nikolau et al., 2008). Non-renewable fossil-fuel-based carbon requires millions of years to generate. Transportation fuels account for more than 90% of the crude oil use in the USA (Marshall). Although they use less oil, the value of other products from the petrochemical industry equals the value of transportation fuels. As technology develops, various renewable alternative energy sources such as wind, solar and nuclear, have

become more and more efficient. An alternative for fossil-fuel-based chemicals is synthesis from biomass, which can be easily generated compared to fossil fuel.

The current fossil-fuel-based chemical industry has become highly efficient and integrated over the last century. For example, a number of different product chemicals are produced from propylene (Nikolau et al., 2008). These include cumene (phenol and acetone), polypropylene, acrylonitrile (acrylic fibers, adiponitrile, rubber elastomers), propylene oxide (antifreeze, polyester, polyurethane), oxo alcohols, and isopropanol. Thus, targeting replacement of fossil-carbon-based platform chemicals (e.g. propylene) with biomass-based versions is a promising strategy. Some potential platform chemicals such as glycerol and 3-hydroxypropionic acid can be produced through the fermentation of carbohydrates (Nikolau et al., 2008). However, these chemicals have a limitation in that they have a high oxygen to carbon ratio. To replace current platform chemicals with the above carbohydrate derived chemicals, it would be necessary to remove a large amount of oxygen. This process is highly energy dependent elevating the production cost (Nikolau et al., 2008). Moreover, petroleum-based platform chemicals such as ethylene, propylene and benzene have various derivatization processes to diverse chemical compounds. These chemicals are the bases for the current chemical industry. If individual chemicals produced from biomass compete in the same market, they would have to replace a whole library of products that are derivatized from the initial platform chemicals. Therefore, when considering biomass as a replacement for petroleum, it would be best to make platform chemicals that can be transformed into multiple products.

For successful production of platform biochemicals using microorganisms, CBiRC (NSF Engineering Research Center for Biorenewable Chemicals) (http://www.cbirc.iastate.edu/) was founded based on cooperation among multiple academic institutions and private company members. The participants are from various disciplines and include chemists, biologists, and engineers, and the goal is to combine biocatalysis and chemical catalysis for the synthesis of chemical precursors.

A good system for platform chemical production is the polyketide synthases. Generally, polyketide synthases can produce molecules with from 4 to over 30 carbons (Nikolau et al., 2008). Polyketide synthases utilize simple molecules as starter and extender units, and synthesize products by adding multiple extender units onto the starter unit (Nikolau et al., 2008). Examples include acetyl-CoA, propionyl-CoA and butyryl-CoA as starter units and malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA as extender units. The choice of starter unit and extender unit varies with different polyketide synthases. Also, by choosing different modification domains, specific functional groups can be added and by choosing different thieoesterases, carbon chain length can be specified. Thus, a broad array of platform chemicals can be generated.

Polyketides can be produced in their native host microorganisms. However, native microorganisms are often not optimal since these molecules are produced at very low level (Khosla and Keasling, 2003). Heterologous expression of the novel biosynthesis pathways in well known industrial host strains such as *E. coli* and *S. cerevisiae* will make the production of platform chemicals much easier due to advantages including genomic information and various cloning tools. One recent example of high-level production of a polyketide is triacetic acid lactone (TAL)

synthesis in *S. cerevisiae* (Cardenas and Da Silva, 2014). Also, new polyketide molecules can be synthesized by manipulating various functional domains from different polyketide synthases.

2.3.2. Polyketide synthases

2.3.2.1. Domain structure of polyketide synthases

Polyketide synthases (PKS) are commonly classified as 3 different types: type I PKS, type II PKS, and type III PKS (Crawford and Townsend, 2010; Shen, 2003; Smith and Tsai, 2007). Type I PKSs are modular, multifunctional enzymes and each of the modules is composed of several different catalytic domains (Shen, 2003; Smith and Tsai, 2007; Staunton and Weissman, 2001). One example of a type I PKS is 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea*, which synthesizes erythromycin A. The fungal type I PKS is similar to one bacterial type I PKS module. Type II PKSs are multi-enzyme systems with each enzyme iteratively carrying out only one catalytic reaction. An example is biosynthesis of the anthraquinone DMAC by the actinorhodin gene cluster enzymes (Carreras et al., 1996). Type III PKSs are single proteins that directly use acetyl-CoA substrates independent of acyl-carrier protein (ACP). These are found in bacteria, fungi, and plants. An example is chalcone synthase (CHS), which is responsible for the initiating step of flavonoid biosynthesis (Austin and Noel, 2003). Type III PKSs are also called CHS-like PKSs. There are also PKSs that do not fit in any of these three types. An example is the fungal iterative type I PKS, which has only one iteratively working module with multiple catalytic domains. An example of a fungal type I PKS is 6-methylsalicylic acid synthase (6-MSAS).

The active site of ACP requires a thiol moiety for substrate binding, which is obtained post-translationally through the transfer of a 4'-phosphopantetheinyl group obtained from CoA and catalyzed by a discrete enzyme, phosphopantetheinyl transferase (PPT) (Elovson and Vagelos, 1968; Lambalot et al., 1996; Lambalot and Walsh, 1995; Pfeifer and Khosla, 2001; Simoni et al., 1967; Vanaman et al., 1968). These PPTs are known to show high specificity for substrates (Pfeifer and Khosla, 2001).

2.3.2.2. Product offloading

The release of the polyketide from the synthase is typically catalyzed by a thioesterase (TE). Thioesterases belong to the α/β-hydrolase superfamily, which includes lipases, proteases and esterases, and have a size of around 28kDa (Du and Lou, 2010). Every TE has three conserved amino acids, Ser-His-Asp. Once the PK chain is complete, it is transferred from ACP to the Ser of the TE active site and released from the PKS. Some PKSs include a TE at the C-terminus of the PKS as a subdomain (Fujii et al., 2001), while other PKSs use a discrete enzyme with TE function (Awakawa et al., 2009; Szewczyk et al., 2008; Xu et al., 2013). Examples of the former includes wA PKS, a naphthopyrone synthase from *A. nidulans* (Fujii et al., 2001) and DEBS (Tsai et al., 2001). Examples of the latter includes atrochrysone carboxyl ACP thioesterase (ACTE) for atrochrysone carboxylic acid synthase (ACAS) from *A. terreus* (Awakawa et al., 2009), AptB for asperthecin biosynthesis pathway from *A. nidulans* (Szewczyk et al., 2008), and LovG for lovastatin biosynthesis pathway from *A. terreus* (Xu et al., 2013). In some cases, another catalytic domain has dual functions as can be seen with 6-MSAS from *A. terreus*. In the case of 6-MSAS, the DH

(dehydratase) domain is responsible for the release of final product after 3 rounds of decarboxylative condensation (Moriguchi et al., 2010).

2.3.3. Protease effect on polyketide synthases and polyketide synthesis

There are many proteases in various compartments of *S. cerevisiae*. Among them, two vacuolar proteases, PrA (*PEP4*, aspartic protease) and PrB (*PRB1*, serine protease) have received particular attention for their effect on the expression of heterologous proteins. PrB is generally blamed the most frequently confronting protein-related problems out of all proteases (Jones, 1991; Pringle, 1975). Activity of these proteases is higher during stationary phase than exponential phase (Jones, 1991); for example, expression of PrB when cells enter stationary phase is 100-fold higher than exponential phase. The effects of proteases were obvious with the heterologous expression of LNKS (LovB) in *S. cerevisiae*. When LovB was expressed in INVSc1 (*MATα his3Δ1 leu2 trp1-289 ura3-52*), a protease intact strain, LovB was not visible on SDS-PAGE, while it was visible when expressed in BJ5464 (*MAT*α *ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1Δ1.6R can1 GAL*) (Lee, 2006).

2.3.4. 6-methylsalicylic acid (6-MSA) biosynthesis

The first fungal PKS that was purified and the first iterative PKS to be cloned was 6-methylsalicylic acid synthase (6-MSAS) from *P. patulum* (Beck et al., 1990; Dimroth et al., 1970).

6-MSAS is a Type I polyketide synthase, which has a multi-domain structure and synthesizes 6-MSA, a precursor to the antibiotic patulin (Child et al., 1996). 6-MSAS is an iterative PKS and its domains are used repeatedly (Child et al., 1996). 6-MSAS is a tetramer of 4 identical subunits.

6-MSAS contains 5 different domains: KS (ketosynthase), AT (acyltransferase), DH (dehydratase), KR (ketoreductase), and ACP (acyl carrier protein) (Staunton and Weissman, 2001). Typically, a starter unit is added to KS and an extender unit is supplied to ACP by AT. Then, the starter unit and an extender unit go through decarboxylative Claisen condensation catalyzed by the KS domain extending the carbon chain length. This extended carbon chain repeats decarboxylative Claisen condensation with an incoming additional extender unit catalyzed by KS again (Awakawa et al., 2009; Pfeifer and Khosla, 2001).

6-MSAS catalyzes the synthesis of 6-MSA from one molecule of acetyl-CoA and three molecules of malonyl-CoA (Figure 2.4) (Dimroth et al., 1976; Dimroth et al., 1970). Two molecules of malonyl-CoA are added onto the starter unit, acetyl-CoA. Subsequently, the third malonyl-CoA is added, but this time a reducing agent NADPH is required for the reaction. Without NADPH, the intermediate product after the addition of the second malonyl-CoA is cyclized to become a triacetic acid lactone (Figure 2.4) (Spencer and Jordan, 1992). 6-MSAS from *P. patulum* has been expressed in organisms such as *Streptomyces coelicolor, E. coli, S. cerevisiae* and tobacco (Bedford et al., 1995; Kealey et al., 1998; Yalpani et al., 2001).

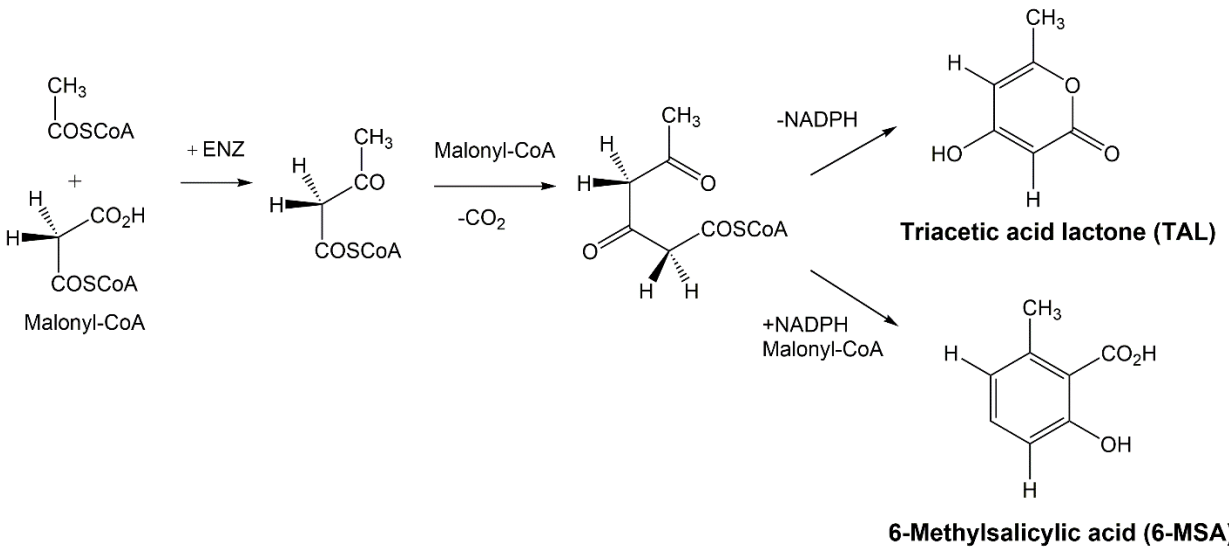


Figure 2.4. Biosynthesis of 6-MSA in *Penicillium patulum*. Figure was adapted from Spencer and Jordan (1992). Two molecules of malonyl-CoA are added repeatedly onto one molecule of acetyl-coA. The addition of the third malonyl-CoA requires NADPH as a cofactor. Without NADPH, a triacetic acid lactone is formed.

2.3.5. Biosynthesis of dihydromonacolin L (DML), the lovastatin precursor

The heterologous production of lovastatin is a subject of active research (Kennedy et al., 1999; Lee, 2006; Ma et al., 2009; Ma and Tang, 2007; Xu et al., 2013). Our laboratory previously constructed a *S. cerevisiae* strain with two lovastatin synthase enzymes (lovastatin nonaketide synthase (LovB) and a dissociative enoyl reductase (LovC)) and a phosphopantetheinyl transferase (P-pant transferase, NpgA) to produce DML (Lee, 2006). The initial strain utilized was INVSc1 (*MATα his3Δ1 leu2 trp1-289 ura3-52*) and no LovB synthesis was observed. Therefore, to avoid possible degradation of LovB by *S. cerevisiae* native proteinases, BJ5464 (*MAT*α *his3*-Δ*200 leu2*-Δ*1 trp1 ura3-52 pep4::HIS3 prb1*Δ*1.6R*) lacking vacuolar protease genes *PEP4* and *PRB1* was

evaluated. With this strain, *LovB* expression was confirmed via SDS-PAGE. The individual expression of *LovB*, *LovC* and *NpgA*, and the activity of NpgA was verified (Lee et al., 2009); however, no DML was detected (Lee, 2006). Our collaborators at UCLA (Prof. Yi Tang and his group) demonstrated that release from the synthase was the barrier. By excising a thioesterase (TE) domain from *Gibberella zeae* PKS13 and using this with our strain, a small quantity of DML was produced *in vitro* (Ma et al., 2009). After introducing a TE into our *S. cerevisiae* host, DML was also produced *in vivo*.

2.3.5.1. Lovastatin biosynthesis

Lovastatin is comprised of two polyketide chains (Figure 2.5) (Ma and Tang, 2007). One is a nonaketide and the other one is a diketide (Kennedy et al., 1999). The nonaketide is synthesized by the cooperative work of a lovastatin nonaketide synthase (LovB) and its dissociative enoyl reductase (LovC). The diketide is synthesized by lovastatin diketide synthase (LovF). The nonaketide and diketide are combined via catalysis by a distinct transesterase (LovD). An important precursor to the nonaketide chain, monacolin J, is dihydromonacolin L (DML).

DML is synthesized by the cooperative action of PPT, LovB and LovC. LovB is a highly reducing iterative type I polyketide synthase (Hendrickson et al., 1999; Kennedy et al., 1999). LovB has multiple domains, most of which have their own catalytic function. Some domains are used repeatedly in the synthesis of the nonaketide chain while one does not have any catalytic function (Figure 2.5) (Ma et al., 2009). Initially, acetyl-CoA is loaded onto ketosynthase (KS) while malonyl-CoA is loaded onto acyl carrier protein (ACP). The acetyl group is bonded to the α carbon

of the malonyl group on ACP by a decarboxylative condensation reaction (Pfeifer and Khosla, 2001). This diketide goes through reduction of the β carbon by ketoreductase (KR) and the resulting hydroxyl group is removed by dehydratase (DH) forming a double bond between the α and β carbons. The acyl group is transferred to the KS domain by malonyl acyltransferase (MAT) while a new malonyl group is added onto the ACP. Another decarboxylative condensation follows to produce a triketide. This cycle is repeated with a few modifications until the acyl chain becomes a nonaketide.

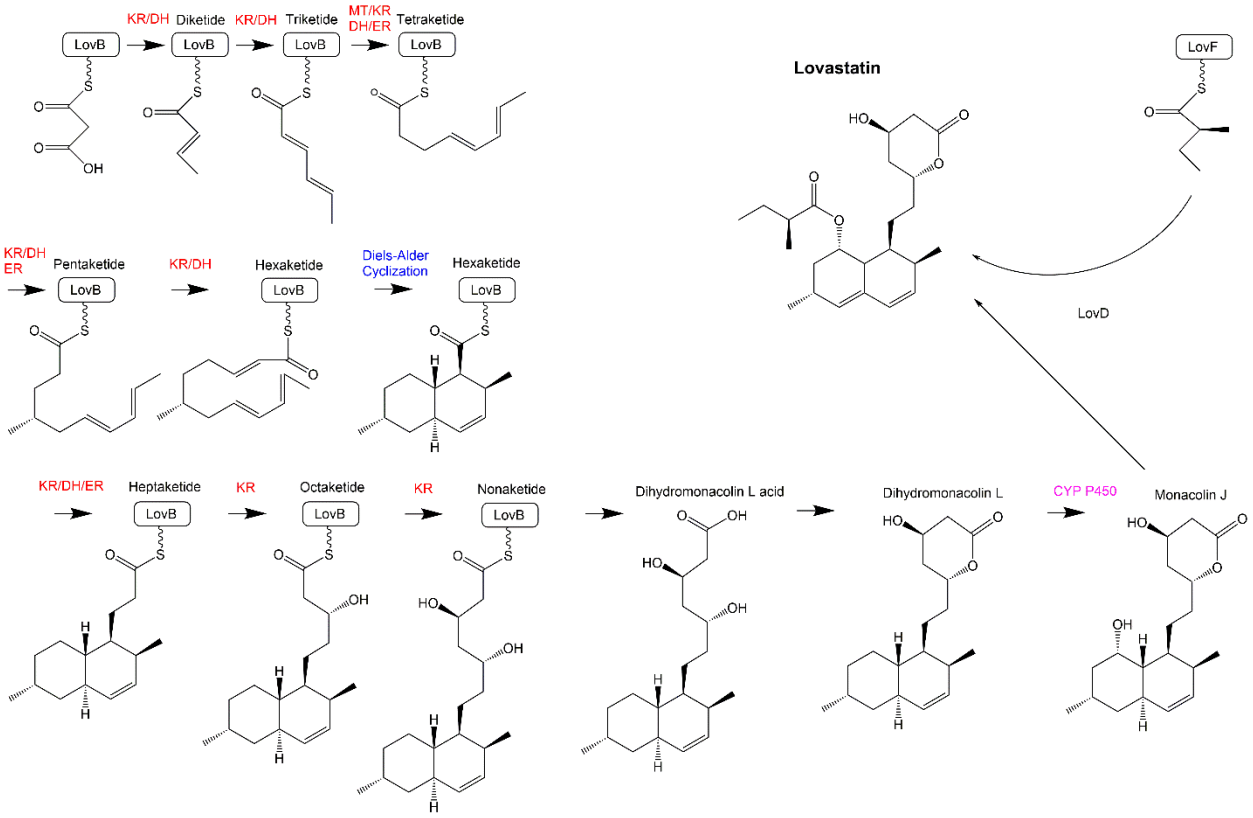


Figure 2.5. Lovastatin biosynthesis pathway. Figure was adapted from Ma and Tang (2007) and Ma et al. (Ma et al., 2009). KS, ketosynthase; DH, dehydratase; MT, methyltransferase; KR, ketoreductase; ACP, acyl carrier protein; ER, enoyl reductase (LovC); CYP P450, cytochrome P450

After dehydration of the triketide, tetraketide and hexaketide, the last double bond formed is reduced by the dissociated enoyl reductase (LovC). After the triketide is formed, a methyl group is added onto the α carbon by methyl transferase (MT). The function of the condensation (CON) domain in DML biosynthesis is unknown but seems to be necessary for normal DML synthesis (Ma et al., 2009).

The starting unit can also be malonyl-CoA instead of acetyl-CoA (Ma and Tang, 2007). The decalin portion is formed by a Diels-Alder cyclization. Diketide is synthesized by LovF and transferred to monacolin J, which is an oxidized DML, by LovD to produce lovastatin (Figure 2.5).

2.3.6. Biosynthesis of acetyl-CoA and malonyl-CoA

The biosynthesis of fatty acids and many polyketides requires acetyl-CoA as a starter unit and malonyl-CoA as an extender unit (Leibundgut et al., 2008). Malonyl-CoA availability can limit high-level production of polyketides (Wattanachaisaereekul et al., 2008). Therefore, it is a necessary step to upregulate malonyl-CoA biosynthesis.

Since one molecule of malonyl-CoA comes from one molecule of acetyl-CoA, an increase in acetyl-CoA synthesis may be necessary. Acetyl-CoA is synthesized via three pathways (Figure 2.6). The first pathway is from pyruvate by catalytic activity of pyruvate dehydrogenase complexes in the mitochondria (Pronk et al., 1996). The second pathway is via the pyruvate dehydrogenase bypass, which includes pyruvate decarboxylase (PDC), acetaldehyde dehydrogenase (ALD) and

acetyl-CoA synthetase (ACS) (Pronk et al., 1996). The third pathway is through the beta-oxidation of fatty acid in the peroxisome (Trotter, 2001).

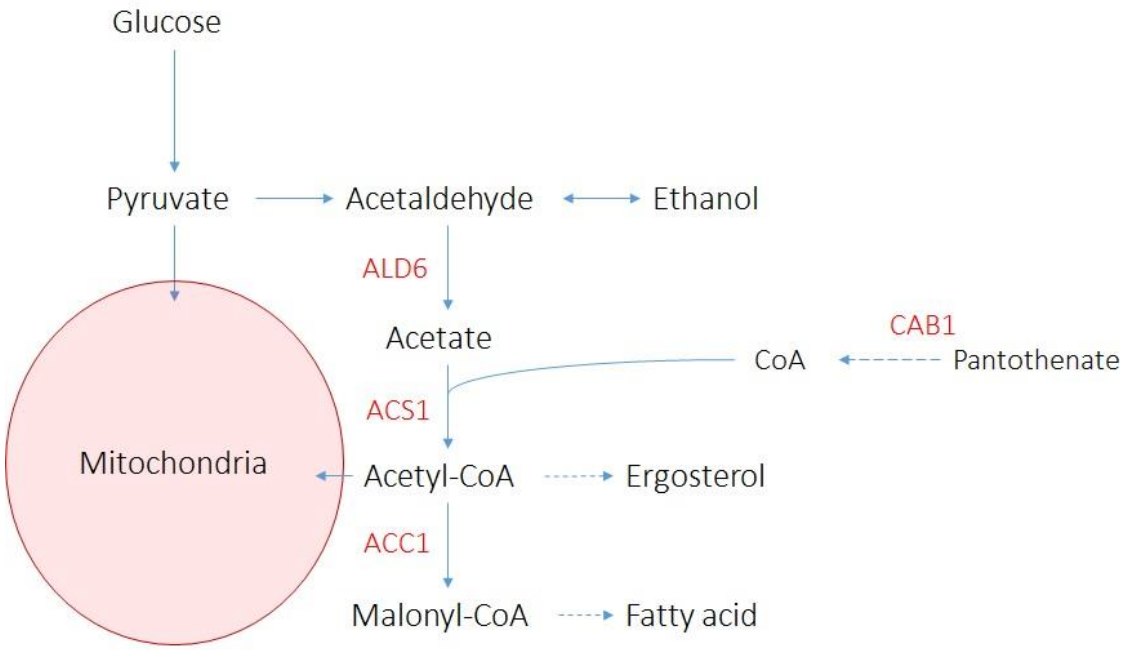


Figure 2.6. Malonyl-CoA biosynthesis pathway in cytoplasm of *S. cerevisiae*. Acetaldehyde is converted to acetate by acetaldehyde dehydrogenase (Ald6). Acetate reacts with free coenzyme A to produce acetyl-CoA catalyzed by acetyl-CoA synthetase (Acs1). Free CoA is produced through the coenzyme A biosynthesis pathway. An early step is the conversion of pantothenate to 4'-phosphopantothenate catalyzed by pantothenate kinase (Cab1). Malonyl-CoA is synthesized from acetyl-CoA, carbonate, and ATP by acetyl-CoA carboxylase (Acc1).

During glucose growth, most pyruvate is converted to acetaldehyde, which is subsequently converted to either ethanol or acetate. Acetate reacts with free CoA catalyzed by acetyl-CoA synthetases (Acs1 and Acs2). Acs1 has a 30-fold lower K_m for acetate relative to Acs2 (vandenBerg et al., 1996). However, transcription of *ACS1* is repressed and Acs1 is possibly

inactivated via degradation in the presence of glucose (deJongGubbels et al., 1997). Acs2 is neither repressed nor degraded in the presence of glucose (deJongGubbels et al., 1997).

There has been success in increasing amorphadiene production via upregulating the acetyl-CoA biosynthesis step with ACS from *Salmonella enterica* (Shiba et al., 2007). Acs1 is post-translationally controlled by the NAD^+/sirtuin dependent protein acetylation/deacetylation system (Starai et al., 2005). It is known that the L641P mutation in the *S. enterica ACS* prevents the acetylation inactivation of the protein (Starai et al., 2005). When *S. enterica* ACS^{L641P} with *ALD6* were overexpressed on a high-copy plasmid, a 1.9 fold increase in amorphadiene production was observed (Shiba et al., 2007).

Acetyl-CoA is synthesized from CoA and acetate catalyzed by Acs1 (Figure 2.6). Free CoA is synthesized through the CoA biosynthesis pathway. In the initial steps of this pathway in yeast, pantothenate kinase (Cab1) generates 4'-phosphopantothenate from pantothenate and ATP (Olzhausen et al., 2009). *CAB1* transcription is subject to glucose repression in *S. cerevisiae* (Olzhausen et al., 2009). In *E. coli* and mammalian cells, pantothenate kinase (PanK, *CoaA*), has been shown to be the rate limiting step in this pathway (Robishaw et al., 1982; Rock et al., 2000; Song and Jackowski, 1992). Vadali *et al.* overexpressed *CoaA* in *E. coli* to improve the free coenzyme A and acetyl-CoA availability (Vadali et al., 2004).

Malonyl-CoA is synthesized from acetyl-CoA, bicarbonate and ATP by the enzyme acetyl-CoA carboxylase (Acc1). Acc1 is phosphorylated and deactivated by Snf1p when glucose is not available (Woods et al., 1994). It has been shown that Acc1 obtained 3 hours after glucose

depletion had approximately 45% lower activity compared to when glucose was available (Woods et al., 1994).

Wattanachaisaereekul *et al.* observed a 60% increase in 6-MSA titer (to 250 mg/L) after changing the promoter for *ACC1* from its native promoter to a strong constitutive promoter, *TEF1* (Wattanachaisaereekul et al., 2008). They also showed that *ACC1* overexpression is more advantageous during the ethanol growth phase (diauxic growth) than the glucose growth phase. This result provides additional indirect evidence for the deactivation of Acc1 by Snf1 and also proves that *ACC1* overexpression helps to offset Acc1 deactivation.

Snf1 is a master control point of energy usage in *S. cerevisiae*. There have been many studies with AMP-activated protein kinase (AMPK), which is the mammalian counterpart of Snf1 (Davies et al., 1990; Hardie et al., 1998; Hardie and Pan, 2002; Scott et al., 2002). Depending on glucose availability, AMPK down-regulates high ATP consuming processes such as fatty acid synthesis, and turns on pathways that are required for cell survival such as fatty acid oxidation (Browne et al., 2004). One of its direct targets is Acc1, which is the initial step of fatty acid synthesis. This energy control system, including the regulation of Acc1, is well conserved in mammalian systems. Acc1 from *Rattus norvegicus* has been shown to be regulated by two protein kinases: AMPK and cAMP-activated protein kinase (PKA) (Davies et al., 1990; Munday et al., 1988). These two kinases are serine/threonine kinases and regulate Acc1 by phosphorylating S77, S79, S1200, and/or S1215. Among these serine residues, S79 and S1215 are phosphorylated by AMPK and S77 is phosphorylated by PKA. S1200 is phosphorylated by both AMPK and PKA.

S77, S79, and S1200 were found to be the critical phosphorylation sites (Ha et al., 1994). Phosphorylation at S1215 did not affect the activity of Acc1.

2.4. References

Alberts, A. W., 1988. Discovery, biochemistry and biology of lovastatin. American Journal of Cardiology. 62**,** J10-J15.

Amore, R., Wilhelm, M., Hollenberg, C. P., 1989. The fermentation of xylose - an analysis of the expression of Bacillus and Actinoplanes xylose isomerase genes in yeast. Applied Microbiology and Biotechnology. 30**,** 351-357.

Austin, M. B., Noel, A. J. P., 2003. The chalcone synthase superfamily of type III polyketide synthases. Natural Product Reports. 20**,** 79-110.

Awakawa, T., Yokota, K., Funa, N., Doi, F., Mori, N., Watanabe, H., Horinouchi, S., 2009. Physically discrete beta-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. Chemistry & Biology. 16**,** 613-623.

Beck, J., Ripka, S., Siegner, A., Schiltz, E., Schweizer, E., 1990. The multifunctional 6-methylsalicylic acid synthase gene of Penicillium-patulum - its gene structure relative to that of other polyketide synthases. European Journal of Biochemistry. 192**,** 487-498.

Becker, J., Boles, E., 2003. A modified Saccharomyces cerevisiae strain that consumes L-arabinose and produces ethanol. Applied and Environmental Microbiology. 69**,** 4144-4150.

Bedford, D. J., Schweizer, E., Hopwood, D. A., Khosla, C., 1995. Expression of a functional fungal polyketide synthase in the bacterium Streptomyces-coelicolor A3(2). Journal of Bacteriology. 177**,** 4544-4548.

Brat, D., Boles, E., Wiedemann, B., 2009. Functional Expression of a Bacterial Xylose Isomerase in Saccharomyces cerevisiae. Applied and Environmental Microbiology. 75**,** 2304-2311.

Browne, G. J., Finn, S. G., Proud, C. G., 2004. Stimulation of the AMP-activated protein kinase leads to activation of eukaryotic elongation factor 2 kinase and to its phosphorylation at a novel site, Serine 398. Journal of Biological Chemistry. 279**,** 12220-12231.

Bruckmann, A., Hensbergen, P. J., Balog, C. I. A., Deelder, A. M., Brandt, R., Snoek, I. S. I., Steensma, H. Y., van Heusden, G. P. H., 2009. Proteome analysis of aerobically and anaerobically grown Saccharomyces cerevisiae cells. Journal of Proteomics. 71**,** 662-669.

Bruinenberg, P. M., Debot, P. H. M., Vandijken, J. P., Scheffers, W. A., 1983. The role of redox balances in the anaerobic fermentation of xylose by yeasts. European Journal of Applied Microbiology and Biotechnology. 18**,** 287-292.

Burr, D. A., Chen, X. B., Vederas, J. C., 2007. Syntheses of conjugated pyrones for the enzymatic assay of lovastatin nonaketide synthase, an iterative polyketide synthase. Organic Letters. 9**,** 161-164.

Cardenas, J., Da Silva, N. A., 2014. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. Metab Eng. 25C**,** 194-203.

Carreras, C. W., Pieper, R., Khosla, C., 1996. Efficient synthesis of aromatic polyketides in vitro by the actinorhodin polyketide synthase. Journal of the American Chemical Society. 118**,** 5158-5159.

Child, C. J., Spencer, J. B., Bhogal, P., ShoolinginJordan, P. M., 1996. Structural similarities between 6-methylsalicylic acid synthase from Penicillium patulum and vertebrate type I fatty acid synthase: Evidence from thiol modification studies. Biochemistry. 35**,** 12267-12274.

Chooi, Y.-H., Tang, Y., 2012. Navigating the fungal polyketide chemical space: from genes to molecules. Journal of Organic Chemistry. 77**,** 9933-9953.

Clark, T. A., Mackie, K. L., 1984. Fermentation inhibitors in wood hydrolysates derived from the softwood Pinus-radiata. Journal of Chemical Technology and Biotechnology B-Biotechnology. 34**,** 101-110.

Crawford, J. M., Townsend, C. A., 2010. New insights into the formation of fungal aromatic polyketides. Nature Reviews Microbiology. 8**,** 879-889.

Davies, S. P., Sim, A. T. R., Hardie, D. G., 1990. Location and function of 3 sites phosphorylated on rat acetyl-CoA carboxylase by the AMP-activated protein-kinase. European Journal of Biochemistry. 187**,** 183-190.

deJongGubbels, P., vandenBerg, M. A., Steensma, H. Y., vanDijken, J. P., Pronk, J. T., 1997. The Saccharomyces cerevisiae acetyl-coenzyme A synthetase encoded by the ACS1 gene, but not the ACS2-encoded enzyme, is subject to glucose catabolite inactivation. Fems Microbiology Letters. 153**,** 75-81.

Dien, B. S., Cotta, M. A., Jeffries, T. W., 2003. Bacteria engineered for fuel ethanol production: current status. Applied Microbiology and Biotechnology. 63**,** 258-266.

Dimroth, P., Ringelmann, E., Lynen, F., 1976. 6-Methylsalicylic acid synthetase from *Penicillium-patulum* - some catalytic properties of enzyme and its relation to fatty-acid synthetase. European Journal of Biochemistry. 68**,** 591-596.

Dimroth, P., Walter, H., Lynen, F., 1970. Biosynthesis of 6-methylsalicylic acid. European Journal of Biochemistry. 13**,** 98-110.

Du, L. C., Lou, L. L., 2010. PKS and NRPS release mechanisms. Natural Product Reports. 27**,** 255-278.

Dupreez, J. C., 1994. Process parameters and environmental-factors affecting D-xylose fermentation by yeasts. Enzyme and Microbial Technology. 16**,** 944-956.

Elovson, J., Vagelos, P. R., 1968. Acyl carrier protein. X. acyl carrier protein synthetase. Journal of Biological Chemistry. 243**,** 3603-&.

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2011. A vector set for systematic metabolic engineering in *Saccharomyces cerevisiae*. Yeast. 28**,** 123-136.

Fujii, I., Ono, Y., Tada, H., Gomi, K., Ebizuka, Y., Sankawa, U., 1996. Cloning of the polyketide synthase gene atX from Aspergillus terreus and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression. Molecular & General Genetics. 253**,** 1-10.

Fujii, I., Watanabe, A., Sankawa, U., Ebizuka, Y., 2001. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of Aspergillus nidulans. Chemistry & Biology. 8**,** 189-197.

Grabowska, D., Chelstowska, A., 2003. The ALD6 gene product is indispensable for providing NADPH in yeast cells lacking glucose-6-phosphate dehydrogenase activity. Journal of Biological Chemistry. 278**,** 13984-13988.

Gueldener, U., Heinisch, J., Koehler, G. J., Voss, D., Hegemann, J. H., 2002. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Research. 30**,** 8.

Guldener, U., Heck, S., Fiedler, T., Beinhauer, J., Hegemann, J. H., 1996. A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Research. 24**,** 2519-2524.

Ha, J., Daniel, S., Broyles, S. S., Kim, K. H., 1994. Critical phosphorylation sites for acetyl-CoA carboxylase activity. Journal of Biological Chemistry. 269**,** 22162-22168.

Ha, S. J., Wei, Q. S., Kim, S. R., Galazka, J. M., Cate, J. H. D., Jin, Y. S., 2011. Cofermentation of cellobiose and galactose by an engineered Saccharomyces cerevisiae strain. Applied and Environmental Microbiology. 77**,** 7438-7438.

Hahn-Hagerdal, B., Karhumaa, K., Jeppsson, M., Gorwa-Grauslund, M. F., 2007. Metabolic engineering for pentose utilization in Saccharomyces cerevisiae. Biofuels. 108**,** 147-177.

Hardie, D. G., Carling, D., Carlson, M., 1998. The AMP-activated/SNF1 protein kinase subfamily: Metabolic sensors of the eukaryotic cell? Annual Review of Biochemistry. 67**,** 821-855.

Hardie, D. G., Pan, D. A., 2002. Regulation of fatty acid synthesis and oxidation by the AMP-activated protein kinase. Biochemical Society Transactions. 30**,** 1064-1070.

Harhangi, H. R., Akhmanova, A. S., Emmens, R., van der Drift, C., de Laat, W., van Dijken, J. P., Jetten, M. S. M., Pronk, J. T., den Camp, H., 2003. Xylose metabolism in the anaerobic fungus Piromyces sp strain E2 follows the bacterial pathway. Archives of Microbiology. 180**,** 134-141.

Hendrickson, L., Davis, C. R., Roach, C., Nguyen, D. K., Aldrich, T., McAda, P. C., Reeves, C. D., 1999. Lavastatin biosynthesis in Aspergillus terreus: characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. Chemistry & Biology. 6**,** 429-439.

Jeffries, T. W., Jin, Y. S., 2004. Metabolic engineering for improved fermentation of pentoses by yeasts. Applied Microbiology and Biotechnology. 63**,** 495-509.

Jin, Y. S., Jeffries, T. W., 2004. Stoichiometric network constraints on xylose metabolism by recombinant Saccharomyces cerevisiae. Metabolic Engineering. 6**,** 229-238.

Jin, Y. S., Jones, S., Shi, N. Q., Jeffries, T. W., 2002. Molecular cloning of XYL3 (D-xylulokinase) from Pichia stipitis and characterization of its physiological function. Applied and Environmental Microbiology. 68**,** 1232-1239.

Jones, E. W., 1991. Tackling the protease problem in *Saccharomyces cerevisiae*. Methods in Enzymology. 194**,** 428-453.

Karhumaa, K., Garcia Sanchez, R., Hahn-Hagerdal, B., Gorwa-Grauslund, M.-F., 2007. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. Microbial Cell Factories. 6.

Karhumaa, K., Hahn-Hagerdal, B., Gorwa-Grauslund, M. F., 2005. Investigation of limiting metabolic steps in the utilization of xylose by recombinant Saccharomyces cerevisiae using metabolic engineering. Yeast. 22**,** 359-368.

Karhumaa, K., Wiedemann, B., Hahn-Hagerdal, B., Boles, E., Gorwa-Grauslund, M. F., 2006. Co-utilization of L-arabinose and D-xylose by laboratory and industrial Saccharomyces cerevisiae strains. Microbial Cell Factories. 5.

Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., Barr, P. J., 1998. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. Proceedings of the National Academy of Sciences of the United States of America. 95**,** 505-509.

Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C., Hutchinson, C. R., 1999. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science. 284**,** 1368-1372.

Khosla, C., Keasling, J. D., 2003. Timeline - Metabolic engineering for drug discovery and development. Nature Reviews Drug Discovery. 2**,** 1019-1025.

Kim, S. R., Park, Y.-C., Jin, Y.-S., Seo, J.-H., 2013. Strain engineering of Saccharomyces cerevisiae for enhanced xylose metabolism. Biotechnology Advances. 31**,** 851-861.

Kotter, P., Amore, R., Hollenberg, C. P., Ciriacy, M., 1990. Isolation and characterization of the Pichia stipitis xylitol dehydrogenase gene, XYL2, and construction of a xylose-utilizing Saccharomyces cerevisiae transformant. Current Genetics. 18**,** 493-500.

Kou, S. C., Christen, Cirillo, V. P., 1970. Galactose transport in Saccharomyces cerevisiae. 2. Characteristics of galactose uptake and exchange in galacktokinaseless cells. Journal of Bacteriology. 103**,** 671-678.

Kuyper, M., Hartog, M. M. P., Toirkens, M. J., Almering, M. J. H., Winkler, A. A., van Dijken, J. P., Pronk, J. T., 2005. Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. Fems Yeast Research. 5**,** 399-409.

Kuyper, M., Winkler, A. A., van Dijken, J. P., Pronk, J. T., 2004. Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. Fems Yeast Research. 4**,** 655-664.

Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., Walsh, C. T., 1996. A new enzyme superfamily - The phosphopantetheinyl transferases. Chemistry & Biology. 3**,** 923-936.

Lambalot, R. H., Walsh, C. T., 1995. Cloning, overproduction, and characterization of the Escherichia-coli holo-acyl carrier protein synthase. Journal of Biological Chemistry. 270**,** 24658-24661.

Leandro, M. J., Fonseca, C., Goncalves, P., 2009. Hexose and pentose transport in ascomycetous yeasts: an overview. Fems Yeast Research. 9**,** 511-525.

Lee, K. K. M., Engineering of Saccharomyces cerevisiae for the biosynthesis of fungal polyketides. University of California, Irvine, 2006.

Lee, K. K. M., Da Silva, N. A., Kealey, J. T., 2009. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in Escherichia coli and Saccharomyces cerevisiae. Analytical Biochemistry. 394**,** 75-80.

Lee, N., Gielow, W., Martin, R., Hamilton, E., Fowler, A., 1986. The organization of the *araBAD* operon of *Escherichia-coli*. Gene. 47**,** 231-244.

Lee, W. J., Kim, M. D., Ryu, Y. W., Bisson, L. F., Seo, J. H., 2002. Kinetic studies on glucose and xylose transport in Saccharomyces cerevisiae. Applied Microbiology and Biotechnology. 60**,** 186-191.

Leibundgut, M., Maier, T., Jenni, S., Ban, N., 2008. The multienzyme architecture of eukaryotic fatty acid synthases. Current Opinion in Structural Biology. 18**,** 714-725.

Li, J. W. H., Vederas, J. C., 2009. Drug discovery and natural products: end of an era or an endless frontier? Science. 325**,** 161-165.

Liang, H., Gaber, R. F., 1996. A novel signal transduction pathway in Saccharomyces cerevisiae defined by Snf3-regulated expression of HXT6. Molecular Biology of the Cell. 7**,** 1953-1966.

Lin, Y., Tanaka, S., 2006. Ethanol fermentation from biomass resources: current state and prospects. Applied Microbiology and Biotechnology. 69**,** 627-642.

Llobell, A., Lopezruiz, A., Peinado, J., Lopezbarea, J., 1988. Glutathione-reductase directly mediates the stimulation of yeast glucose-6-phosphate-dehydrogenase by GSSG. Biochemical Journal. 249**,** 293-296.

Londesborough, J., Richard, P., Valkonen, M., Viljanen, K., 2014. Effect of C-terminal protein tags on pentitol and L-arabinose transport by A*mbrosiozyma monospora* Lat1 and Lat2 transporters in *Saccharomyces cerevisiae*. Applied and Environmental Microbiology. 80**,** 2737-2745.

Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X. K., Kealey, J. T., Da Silva, N. A., Vederas, J. C., Tang, Y., 2009. Complete Reconstitution of a Highly Reducing Iterative Polyketide Synthase. Science. 326**,** 589-592.

Ma, S. M., Tang, Y., 2007. Biochemical characterization of the minimal polyketide synthase domains in the lovastatin nonaketide synthase LovB. Febs Journal. 274**,** 2854-2864.

Madhavan, A., Srivastava, A., Kondo, A., Bisaria, V. S., 2012. Bioconversion of lignocellulose-derived sugars to ethanol by engineered *Saccharomyces cerevisiae*. Critical Reviews in Biotechnology. 32**,** 22-48.

Madhavan, A., Tamalampudi, S., Ushida, K., Kanai, D., Katahira, S., Srivastava, A., Fukuda, H., Bisaria, V. S., Kondo, A., 2009. Xylose isomerase from polycentric fungus Orpinomyces: gene sequencing, cloning, and expression in Saccharomyces cerevisiae for bioconversion of xylose to ethanol. Applied Microbiology and Biotechnology. 82**,** 1067-1078.

Maggon, K., 2005. Best-selling human medicines 2002-2004. Drug Discovery Today. 10**,** 739-742.

Marshall, J., Biorefineries: curing our addiction to oil. New Scientist, pp. 28-31.

McDaniel, R., Licari, P., Khosla, C., 2001. Process development and metabolic engineering for the overproduction of natural and unnatural polyketides. Advances in biochemical engineering/biotechnology. 73**,** 31-52.

McMillan, J. D., 1994. Conversion of hemicellulose hydrolyzates to ethanol. Enzymatic Conversion of Biomass for Fuels Production. 566**,** 411-437.

Moniruzzaman, M., Dien, B. S., Skory, C. D., Chen, Z. D., Hespell, R. B., Ho, N. W. Y., Dale, B. E., Bothast, R. J., 1997. Fermentation of corn fibre sugars by an engineered xylose utilizing Saccharomyces yeast strain. World Journal of Microbiology & Biotechnology. 13**,** 341-346.

Moriguchi, T., Kezuka, Y., Nonaka, T., Ebizuka, Y., Fujii, I., 2010. Hidden function of catalytic domain in 6-methylsalicylic acid synthase for product release. Journal of Biological Chemistry. 285**,** 15637-15643.

Munday, M. R., Campbell, D. G., Carling, D., Hardie, D. G., 1988. Identification by amino-acid sequencing of 3 major regulatory phosphorylation sites on rat acetyl-CoA carboxylase. European Journal of Biochemistry. 175**,** 331-338.

Mutka, S. C., Bondi, S. M., Carney, J. R., Da Silva, N. A., Kealey, J. T., 2006. Metabolic pathway engineering for complex polyketide biosynthesis in Saccharomyces cerevisiae. Fems Yeast Research. 6**,** 40-47.

Nikolau, B. J., Perera, M., Brachova, L., Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. Plant Journal. 54**,** 536-545.

Olzhausen, J., Schubbe, S., Schuller, H. J., 2009. Genetic analysis of coenzyme A biosynthesis in the yeast Saccharomyces cerevisiae: identification of a conditional mutation in the pantothenate kinase gene CAB1. Current Genetics. 55**,** 163-173.

Ozcan, S., Johnston, M., 1999. Function and regulation of yeast hexose transporters. Microbiology and Molecular Biology Reviews. 63**,** 554-569.

Palmqvist, E., Hahn-Hagerdal, B., 2000. Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. Bioresource Technology. 74**,** 17-24.

Panagiotou, G., Andersen, M. R., Grotkjaer, T., Regueira, T. B., Nielsen, J., Olsson, L., 2009. Studies of the production of fungal polyketides in Aspergillus nidulans by using systems biology tools. Applied and Environmental Microbiology. 75**,** 2212-2220.

Pfeifer, B. A., Khosla, C., 2001. Biosynthesis of polyketides in heterologous hosts. Microbiology and Molecular Biology Reviews. 65**,** 106-+.

Pickens, L. B., Tang, Y., Chooi, Y. H., 2011. Metabolic engineering for the production of natural products. In: Prausnitz, J. M., (Ed.), Annual Review of Chemical and Biomolecular Engineering, Vol 2. vol. 2. Annual Reviews, Palo Alto, pp. 211-236.

Pringle, J. R., 1975. Methods for avoiding proteolytic artefacts in studies of enzymes and other proteins from yeasts. Methods Cell Biol. 12**,** 149-84.

Pronk, J. T., Steensma, H. Y., vanDijken, J. P., 1996. Pyruvate metabolism in Saccharomyces cerevisiae. Yeast. 12**,** 1607-1633.

Robishaw, J. D., Berkich, D., Neely, J. R., 1982. Rate-limiting step and control of coenzyme-A synthesis in cardiac-muscle. Journal of Biological Chemistry. 257**,** 967-972.

Rock, C. O., Calder, R. B., Karim, M. A., Jackowski, S., 2000. Pantothenate kinase regulation of the intracellular concentration of coenzyme A. Journal of Biological Chemistry. 275**,** 1377-1383.

Romanos, M. A., Scorer, C. A., Clare, J. J., 1992. Foreign gene expression in yeast - a review. Yeast. 8**,** 423-488.

Sauer, B., 1994. Recycling selectable markers in yeast. Biotechniques. 16**,** 1086-1088.

Scott, J. W., Norman, D. G., Hawley, S. A., Kontogiannis, L., Hardie, D. G., 2002. Protein kinase substrate recognition studied using the recombinant catalytic domain of AMP-activated protein kinase and a model substrate. Journal of Molecular Biology. 317**,** 309-323.

Shen, B., 2003. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. Current Opinion in Chemical Biology. 7**,** 285-295.

Shen, M. W. Y., Fang, F., Sandmeyer, S., Da Silva, N. A., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 29**,** 495-503.

Shiba, Y., Paradise, E. M., Kirby, J., Ro, D. K., Keasing, J. D., 2007. Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. Metabolic Engineering. 9**,** 160-168.

Simoni, R. D., Criddle, R. S., Stumpf, P. K., 1967. Fat metabolism in higher plants .31. Purification and properties of plant and bacterial acyl carrier proteins. Journal of Biological Chemistry. 242**,** 573-581.

Smith, S., Tsai, S.-C., 2007. The type I fatty acid and polyketide synthases: a tale of two megasynthases. Natural Product Reports. 24**,** 1041-1072.

Song, W. J., Jackowski, S., 1992. Cloning, sequencing, and expression of the pantothenate kinase (coaA) gene of Escherichia-coli. Journal of Bacteriology. 174**,** 6411-6417.

Spencer, J. B., Jordan, P. M., 1992. Purification and properties of 6-methylsalicylic acid synthase from *Penicillium patulum*. Biochemical Journal. 288**,** 839-846.

Starai, V. J., Gardner, J. G., Escalante-Semerena, J. C., 2005. Residue Leu-641 of acetyl-CoA synthetase is critical for the acetylation of residue Lys-609 by the protein acetyltransferase enzyme of Salmonella enterica. Journal of Biological Chemistry. 280**,** 26200-26205.

Staunton, J., Weissman, K. J., 2001. Polyketide biosynthesis: a millennium review. Natural Product Reports. 18**,** 380-416.

Szewczyk, E., Chiang, Y. M., Oakley, C. E., Davidson, A. D., Wang, C. C. C., Oakley, B. R., 2008. Identification and Characterization of the Asperthecin Gene Cluster of Aspergillus nidulans. Applied and Environmental Microbiology. 74**,** 7607-7612.

Trotter, P. J., 2001. The genetics of fatty acid metabolism in Saccharomyces cerevisiae. Annual Review of Nutrition. 21**,** 97-119.

Tsai, S. C., Miercke, L. J. W., Krucinski, J., Gokhale, R., Chen, J. C. H., Foster, P. G., Cane, D. E., Khosla, C., Stroud, R. M., 2001. Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: Versatility from a unique substrate channel. Proceedings of the National Academy of Sciences of the United States of America. 98**,** 14808-14813.

Vadali, R. V., Bennett, G. N., San, K. Y., 2004. Cofactor engineering of intracellular CoA/acetyl-CoA and its effect on metabolic flux redistribution in Escherichia coli. Metabolic Engineering. 6**,** 133-139.

Vanaman, T. C., Wakil, S. J., Hill, R. L., 1968. Complete amino acid sequence of acyl carrier protein of Escherichia coli. Journal of Biological Chemistry. 243**,** 6420-&.

vandenBerg, M. A., deJongGubbels, P., Kortland, C. J., vanDijken, J. P., Pronk, J. T., Steensma, H. Y., 1996. The two acetyl-coenzyme A synthetases of Saccharomyces cerevisiae differ with respect to kinetic properties and transcriptional regulation. Journal of Biological Chemistry. 271**,** 28953-28959.

Verduyn, C., Vankleef, R., Frank, J., Schreuder, H., Vandijken, J. P., Scheffers, W. A., 1985. Properties of the NAD(P)H-dependent xylose reductase from the xylose-fermenting yeast Pichia stipitis. Biochemical Journal. 226**,** 669-677.

Wahlbom, C. F., Hahn-Hagerdal, B., 2002. Furfural, 5-hydroxymethyl furfural, and acetoin act as external electron acceptors during anaerobic fermentation of xylose in recombinant Saccharomyces cerevisiae. Biotechnology and Bioengineering. 78**,** 172-178.

Walfridsson, M., Bao, X. M., Anderlund, M., Lilius, G., Bulow, L., HahnHagerdal, B., 1996. Ethanolic fermentation of xylose with Saccharomyces cerevisiae harboring the Thermus thermophilus xylA gene, which expresses an active xylose (glucose) isomerase. Applied and Environmental Microbiology. 62**,** 4648-4651.

Watanabe, S., Abu Saleh, A., Pack, S. P., Annaluru, N., Kodaki, T., Makino, K., 2007. Ethanol production from xylose by recombinant Saccharomyces cerevisiae expressing protein-engineered NADH-preferring xylose reductase from Pichia stipitis. Microbiology-Sgm. 153**,** 3044-3054.

Watanabe, S., Kodaki, T., Makino, K., 2005. Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. Journal of Biological Chemistry. 280**,** 10340-10349.

Wattanachaisaereekul, S., Lantz, A. E., Nielsen, M. L., Nielsen, J., 2008. Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. Metabolic Engineering. 10**,** 246-254.

Weissman, K. J., Leadlay, P. F., 2005. Combinatorial biosynthesis of reduced polyketides. Nature Reviews Microbiology. 3**,** 925-936.

Wilhelm, M., Hollenberg, C. P., 1984. Selective cloning of Bacillus-subtilis xylose isomerase and xylulokinase in Escherichia-coli genes by IS5-mediated expression. Embo Journal. 3**,** 2555-2560.

Wisselink, H. W., Toirkens, M. J., Wu, Q., Pronk, J. T., van Maris, A. J. A., 2009. Novel Evolutionary Engineering Approach for Accelerated Utilization of Glucose, Xylose, and Arabinose Mixtures by Engineered Saccharomyces cerevisiae Strains. Applied and Environmental Microbiology. 75**,** 907-914.

Woods, A., Munday, M. R., Scott, J., Yang, X. L., Carlson, M., Carling, D., 1994. Yeast Snf1 is functionally related to mammalian AMP-activated protein-kinase and regulates acetyl-CoA carboxylase *in vivo*. Journal of Biological Chemistry. 269**,** 19509-19515.

Wooley, R., Ruth, M., Sheehan, J., Ibsen, K., Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis current and futuristic scenarios. NREL/TP-580-26157, 1999.

Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., Tang, Y., 2013. LovG: The Thioesterase Required for Dihydromonacolin L Release and Lovastatin Nonaketide Synthase Turnover in Lovastatin Biosynthesis. Angew Chem Int Ed Engl. 52**,** 6472-5.

Yalpani, N., Altier, D. J., Barbour, E., Cigan, A. L., Scelonge, C. J., 2001. Production of 6-methylsalicyclic acid by expression of a fungal polyketide synthase activates disease resistance in tobacco. Plant Cell. 13**,** 1401-1409.

Chapter 3.

Construction and Expression of Fungal Arabinose Pathway Genes in *Saccharomyces cerevisiae*

3.1. Abstract

Ethanol synthesized from biological sources is considered a sustainable alternative to the use of fossil-based transportation fuel. For the cost-efficient production of ethanol, utilization of all available sugars, including L-arabinose, is imperative. In this study, we tested the fungal arabinose assimilation pathway in *S. cerevisiae*. Two critical obstacles in implementing the fungal arabinose pathway in this yeast are a cofactor imbalance arising from the use of NADPH and NADH, and poor arabinose uptake via the inefficient and unstable Gal2 transporter. To address the issue with cofactor imbalance, we introduced NADH-dependent L-xylulose reductase from a yeast strain *Ambrosiozyma monospora* (*ALX1*) along with other participating genes (*XYL1* and *lad1*) in this pathway, and tested the pathway by measuring xylitol synthesis. We compared ALX1 and the NADPH-dependent LXR1 (from *Trichoderma reesei*), and also compared *XYL1* expression from CEN/ARS-based and 2μ-based plasmids for the high level xylitol synthesis via the fungal arabinose pathway in *S. cerevisiae*. To address the inefficient arabinose uptake due to Gal2, we engineered the protein to prevent ubiquitination-triggered degradation by constructing a chimera protein including parts of Gal2 and the non-ubiquitinated Hxt1. We compared cell growth and developed a strategy for future directions in engineering Gal2.

3.2. Introduction

Ethanol has received global attention as a potential automotive fuel alternative to crude oil-based fuel (Kumar et al., 2009). Unlike traditional biomass for biofuel production, which has competed with food production and partly contributed to an increase in commodity prices, current research on biofuels focuses on the utilization of lignocelluloses, which include sugarcane bagasse, wood, corn cob, corn stover, wheat and rice straw (Sims et al., 2010). Lignocellulosic biomass includes both cellulose (20-50%) and hemicelluloses (20-35%) (Mielenz, 2001). For economical fuel ethanol production, development in the utilization of hemicelluloses, such as xylose and arabinose, as well as cellulose is crucial (Kumar et al., 2009). Xylose has received the most attention since it is the most abundant hemicellulose, while arabinose has not been as well studied (Jeffries, 2006). However, depending on the raw material, arabinose can be found in comparable amounts to xylose (Hahn-Hagerdal et al., 2007b). For example, corn stover is composed of 19% xylan and 3% arabinan, while wheat bran is composed of 19% xylan and 15% arabinan (Hahn-Hagerdal et al., 2007b). *S. cerevisiae* lacks a native fermentation capability for either xylose or arabinose. Introduction of the heterologous pentose pathways from naturally pentose fermenting microorganism, such as *Pichia stipitis* and *Ambrosiozyma monospora*, into *S. cerevisiae* is necessary for economical production of ethanol.

The construction of pentose-utilizing *S. cerevisiae* strains for ethanol production has been widely studied and frequently reviewed (Dumon et al., 2012; Hahn-Hagerdal et al., 2007a; Kim et al., 2013; Madhavan et al., 2012; Matsushika et al., 2009; Van Vleet and Jeffries, 2009). However, most studies have focused on xylose metabolism (Jin and Jeffries, 2004; Jin et al., 2000; Karhumaa

et al., 2007; Kim et al., 2013; Kotter and Ciriacy, 1993; Kuyper et al., 2005; Madhavan et al., 2009; Watanabe et al., 2007a; Watanabe et al., 2005) while heterologous expression of exogenous arabinose pathways in *S. cerevisiae* has received less attention (Becker and Boles, 2003; Bera et al., 2010; Bettiga et al., 2009; Dumon et al., 2012; Metz et al., 2013; Richard et al., 2003; Wisselink et al., 2007). Two pathways for L-arabinose utilization have been introduced into *S. cerevisiae* for ethanol production, the bacterial pathway and the fungal pathway (Becker and Boles, 2003; Richard et al., 2003), with more studies focused on the bacterial pathway (Figure 3.1) (Becker and Boles, 2003; Karhumaa et al., 2006; Wisselink et al., 2007). For xylose utilization, the fungal XR/XDH pathway is often employed in *S. cerevisiae* due to its thermodynamic advantage over the bacterial xylose isomerase pathway (Karhumaa et al., 2007). Arabitol is synthesized by XR from arabinose in fungal xylose pathway when arabinose is fed, and the bacterial arabinose pathway is inhibited by arabitol (Karhumaa et al., 2006). Therefore, there is an advantage to the fungal arabinose pathway when pairing it with the XR/XDH-based xylose pathway. There is one common difficulty between the fungal xylose pathway and the fungal arabinose pathway, the imbalance between redox cofactors (Richard et al., 2001). Reduction steps in the fungal xylose and the fungal arabinose pathway typically require NADPH, while oxidations require NAD^+ (Figure 3.1A). In the xylose pathway, there is one reduction and one oxidation. However, in the arabinose pathway, there are two reductions and two oxidations. There have been several research efforts to overcome this difficulty for these fungal pathways (Kuyper et al., 2005; Matsushika et al., 2008; Verho et al., 2004; Watanabe et al., 2007b; Watanabe et al., 2005). For the fungal arabinose pathway, a NADH dependent L-xylulose reductase was found as in figure 3.1B (Verho et al., 2004).

This enzyme was introduced from the yeast *A. monospora* to *S. cerevisiae*. One of two redox imbalances can be balanced by introducing this enzyme.

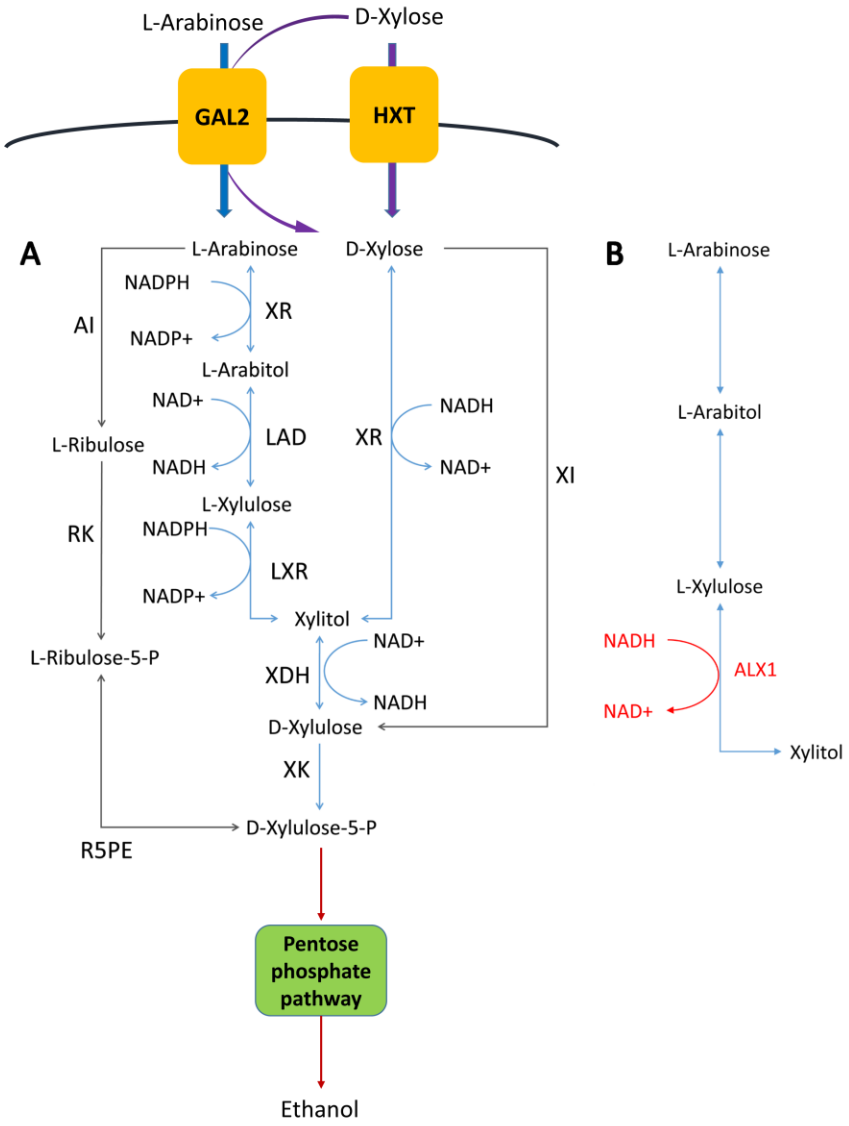


Figure 3.1. Pentose pathways. (A) Fungal and bacterial pentose utilization pathways. The fungal arabinose pathway uses an NADPH-dependent LXR1 (*lxr1*) from *T. reesei*, (B) The fungal arabinose pathway but with the NADH-dependent ALX1 (*ALX1*) from *A. monospora*. GAL2, Galactose permease (*GAL2*);HXT, hexose transporters; XR, xylose reductase (*XYL1*) from *P. stipitis*; LAD, L-arabitol 4-dehydrogenase (*lad1*) from *Trichoderma reesei*; LXR, L-xylulose reductase (*lxr1*) from *T. reesei*; XDH, xylitol dehydrogenase, (*XYL2*) from *P. stipitis*; XK, xylulokinase (*XYL3* from *P. stipitis* or *XKS1* from *S. cerevisiae*); XI, xylose isomerase (*XylA*); AI, L-arabinose isomerase (*araA*) from *Bacillus subtilis*; RK, ribulokinase (*araB*) from *E. coli*; R5PE, ribulose-5-phosphate-4-epimerase (*araD*) from *E. coli*.

Our overall goal is to create an efficient arabinose-utilizing *S. cerevisiae* strain for ethanol production. To achieve this goal, first, the exogenous fungal arabinose pathway genes including the NADH dependent L-xylulose reductase were optimized for assembly and for expression in *S. cerevisiae*. Second, we considered another rate-limiting factor in the production of ethanol using the arabinose pathway, L-arabinose uptake (Becker and Boles, 2003; Richard et al., 2003). A series of modifications of the L-arabinose transporter were made to increase L-arabinose uptake.

3.3. Materials and Methods

3.3.1. Molecular biology techniques

Oligonucleotides were synthesized by Integrated DNA technologies, Inc. PCR was carried out using KOD Hot Start DNA Polymerase (Merck KGaA, Darmstadt, Germany) or *PfuUltra* II HS DNA Polymerase (Agilent, Santa Clara, CA) for the high fidelity work or Taq DNA polymerase (New England Biolabs, Ipswich, MA). Restriction enzymes were purchased from New England Biolabs, Inc. Qiaquick Gel Extraction Kit (Qiagen) or Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Orange, CA) was used for the isolation of DNA fragments from agarose gel. T4 DNA ligase (New England Biolabs, Ipswich, MA) was used for the ligation of DNA fragments. *Escherichia coli* XL1Blue or DH5α cells were used for the storage and propagation of recombinant DNA plasmids. GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, Waltham, MA) was used for the isolation of plasmid DNA from *E. coli* cells. All primer sequences can be found in Table C.1 (Appendix C).

3.3.2. Vector constructions

3.3.2.1. Construction of pJC vectors

PmeI and RsrII were added between SpeI and XhoI of pXP218. To do that, the *CYC1* terminator of pXP218 (012209 pXP Cyct For SpeI, 012209 pXP Cyct Rev BsrG) was PCR amplified. The PCR product and pXP218 were digested with SpeI and BsrGI, and ligated to create pJC218.

By cloning the SpeI to BsrGI fragment from pJC218 to other pXP vectors, PmeI and RsrII were added to these pXP vectors creating pJC vectors as shown in Table 3.1.

Table 3.1. pXP (Fang et al., 2011; Shen et al., 2012) and pJC series vectors

Starting vector	New vectors	Promoter	Marker	Type
pXP206		*PGK1*	*TRP1*	2μ
pXP209		*PGK1*	*URA3*	2μ
pXP218	pJC218	*PGK1*	*URA3*	2μ
pXP711	pJC711	*PGK1*	*LEU2-d8*	CEN/ARS
pXP712	pJC712	*PGK1*	*URA3*	CEN/ARS
pXP741	pJC741	*ADH2*	*LEU2-d8*	CEN/ARS
pXP742	pJC742	*ADH2*	*URA3*	CEN/ARS
	pJC743	*ADH2*	*TRP1*	CEN/ARS
	pJC752	*ADH1*	*URA3*	CEN/ARS
	pJC753	*ADH1*	*TRP1*	CEN/ARS
pXP811	pJC811	*PGK1*	*LEU2-d8*	2μ
pXP812	pJC812	*PGK1*	*URA3*	2μ
pXP841	pJC841	*ADH2*	*LEU2-d8*	2μ
pXP842	pJC842	*ADH2*	*URA3*	2μ
	pJC843	*ADH2*	*TRP1*	2μ

3.3.2.2. Codon/codon-pair-optimized gene synthesis and cloning

Three codon/codon-pair optimized fungal arabinose pathway genes (L-arabinose 4-dehydrogenase (*clad1*) from *Trichoderma reesei*, L-xylulose reductase (*clxr1*) from *T. reesei*, and L-xylulose reductase (*cALX1*) from *A. monospora* were designed by Verdezyne Inc. and synthesized by our laboratory using the method developed by Larsen *et al* (2008). Each gene was PCR-assembled through multiple steps for easy PCR assembly. Initially, small fragments of 40

base pairs were synthesized at IDTDNA. Oligos contain overlapping sequences that are redundant in neighboring oligos allowed sequence based self-assembly between these oligos. 10 overlapping oligos were PCR-assembled to synthesize an intermediate fragment of about 200 base pairs. Each intermediate fragments were purified by Minelute PCR purification kit (Qiagen) and were confirmed for correct size by agarose gel electrophoresis. Five to six intermediate fragments, which also contain sequence overlap between neighboring fragments, were again PCR-assembled to synthesize the full length gene. A codon/codon-pair optimized synthetic *XYL1* (in pETH52 with CEN/ARS or pETH53 with 2μ, Table 3.2) coding for XR was obtained from the Computational Biology Research Laboratory at UCI. Sequences of oligos are included in Appendix A.2. A his-tag was added after the start codon (ATG) of each enzyme by PCR. The PCR products were inserted into pCR-Blunt II-TOPO and sequenced. His-tagged *clad1, clxr1* (*T. reesei*) and *cALX1* (*A. monospora*) were first cloned into HindIII and XhoI sites of pYES2 (Invitrogen, Inc., Carlsbad, CA) under the *GAL1* promoter to produce pOUI31, pOUI30 and pOUI34 respectively. Cloned plasmids were transformed into *E. coli* DH5α cells.

For constitutive expression via the *PGK1* promoter, synthetic genes *clad1*, *clxr1*, and *cALX1* were PCR amplified with SpeI and XhoI as flanking restriction sites. Kozak sequence (AAAAAA) was added before the start codon (Kozak, 1986), with or without the His-tag after the start codon. PCR products were inserted into pCR-Blunt II-TOPO. c*lad1* from the pTOPO vector was cloned into SpeI to XhoI of pXP209. *clxr1* and *cALX1* were cloned into pXP206 in the same way. Native *lad1*, *lxr1* and *ALX1* genes were obtained from Dr. Peter Richard (VTT Technical Research Center of Finland) and cloned into the same pXP vectors (with or without a N-terminal His-tag) as their synthetic versions.

Table 3.2 List of plasmids used in this study

Plasmid	Characteristics	Markers	Reference
pYES2	P_{GAL1}-T_{CYC1}; 2μ	*URA3*	Invitrogen
pOUI30	P_{GAL1}-Synthetic *lxr1 (clxr1)*-T_{CYC1}; 2μ	*URA3*	This study
pOUI31	P_{GAL1}-Synthetic *lad1 (clad1)*-T_{CYC1}; 2μ	*URA3*	This study
pOUI34	P_{GAL1}-Synthetic *ALX1 (cALX1)*-T_{CYC1}; 2μ	*URA3*	This study
pETH52	P_{PGK1}-Synthetic *XYL1 (cXYL1)*-T_{CYC1}; CEN/ARS	*MET17*	Unpublished
pETH53	P_{PGK1}-Synthetic *XYL1 (cXYL1)*-T_{CYC1}; 2μ	*MET17*	Unpublished
pXP206-LXRT	P_{PGK1}-*clxr1* (*T. reesei)*-T_{CYC1}; 2μ	*TRP1*	This study
pXP206-LXRT-His	P_{PGK1}-*clxr1 (T. reesei)*-T_{CYC1}; 2μ; 6X His	*TRP1*	This study
pXP206-LXRA	P_{PGK1}-*cALX1 (A. monospora)*-T_{CYC1}; 2μ	*TRP1*	This study
pXP206-LXRA-His	P_{PGK1}-*cALX1 (A. monospora)*-T_{CYC1}; 2μ; ; 6X His	*TRP1*	This study
pXP209-LAD	P_{PGK1}-*clad1 (T. reesei)*-T_{CYC1}; 2μ	*URA3*	This study
pXP209-LAD-His	P_{PGK1}-*clad1 (T. reesei)*-T_{CYC1}; 2μ; 6X His	*URA3*	This study
pXP206-1680	P_{PGK1}-*lxr1 (T. reesei)*-T_{CYC1}; 2μ	*TRP1*	This study
pXP206-1680H	P_{PGK1}-*lxr1 (T. reesei)*-T_{CYC1}; 2μ; 6X His	*TRP1*	This study
pXP206-2178	P_{PGK1}-*ALX1 (A. monospora)*-T_{CYC1}; 2μ	*TRP1*	This study
pXP206-2178H	P_{PGK1}-*ALX1 (A. monospora)*-T_{CYC1}; 2μ; 6X His	*TRP1*	This study
pXP209-2073	P_{PGK1}-*lad1 (T. reesei)*-T_{CYC1}; 2μ	*URA3*	This study
pXP209-2073H	P_{PGK1}-*lad1 (T. reesei)*-T_{CYC1}; 2μ; 6X His	*URA3*	This study
YEpADH2p	P_{ADH2}-T_{ADH2}, 2μ	*URA3*	This study
pJC752-GH1	P_{ADH1}-*GH1*(*GAL2*/*HXT1* chimera 1)-T_{CYC1}, CEN/ARS	*URA3*	This study
pJC752-GH2	P_{ADH1}- *GH1*(*GAL2*/*HXT1* chimera 2)-T_{CYC1}, CEN/ARS	*URA3*	This study
pJC752-GAL2	P_{ADH1}-*GAL2*-T_{CYC1}, CEN/ARS, *URA3*; Amp	*URA3*	This study
pJC811-GH1	P_{PGK1}-*GH1*(*GAL2*/*HXT1* chimera 1)-T_{CYC1}, 2μ	*LEU2-d8*	This study
pJC811-GH2	P_{PGK1}- *GH1*(*GAL2*/*HXT1* chimera 2)-T_{CYC1}, 2μ	*LEU2-d8*	This study
pJC811-GAL2	P_{PGK1}- *GAL2*-T_{CYC1}, 2μ	*LEU2-d8*	This study
pJC812-GAL2	P_{PGK1}- *GAL2*-T_{CYC1}, 2μ	*URA3*	This study
pJC812-GH1	P_{PGK1}-*GH1*(*GAL2*/*HXT1* chimera 1)-T_{CYC1}, 2μ	*URA3*	This study
pJC812-GH2	P_{PGK1}- *GH1*(*GAL2*/*HXT1* chimera 2)-T_{CYC1}, 2μ	*URA3*	This study

3.3.2.3. *GAL2*/*HXT1* hybrid gene synthesis and plasmid construction

GAL2 (pJC23, pJC24) was PCR amplified from BY4741 genomic DNA with RsrII and XhoI flanking the gene. The Kozak sequence (AAAAAA) was inserted in the forward primer sequence before the start codon. The PCR product was digested with RsrII and XhoI, and cloned into the same restriction sites between P_{PGK1} and T_{CYC1} of pJC811 (Table 3.1) creating pJC811-GAL2 (Table 3.2). *HXT1* gene was PCR amplified from the genome of BY4741. Two versions of *GAL2/HXT1* chimera genes were constructed. Chimera 1 had the N-terminus, the loop between transmembrane segment (TM) 6 and TM7, and the C- terminus replaced with the *HXT1* counterparts keeping all 12 *GAL2* transmembrane segments. Chimera 2 had only the N- and C-termini replaced with the *HXT1* counterparts.

To construct chimera 1, TM1 to TM6 were first PCR amplified (pJC15, pJC16) from pJC811-GAL2 adding RsrII and StuI sites at the N-terminus and AvrII and XhoI sites at the C-terminus from *GAL2.* The product was cloned into RsrII-XhoI sites of pJC811 to create pJC811-*GAL2*Fr2. The N-terminus of *HXT1* was PCR amplified (pJC13, pJC14) with RsrII at the N-terminal end. This PCR product was inserted at RsrII and StuI (Blunt end) sites of pJC811-*GAL2*Fr2 to create pJC811-*GAL2*Fr1Fr2. Next, the middle loop between TM6 and TM7 was PCR amplified (pJC17, pJC18) with AvrII upstream and PmeI and XhoI downstream. The PCR product was digested with AvrII and XhoI, and inserted at AvrII and XhoI sites of pJC811-*GAL2*Fr1Fr2 to create pJC811-*GAL2*Fr1Fr2Fr3. To clone the fragment 4, which has TM7 to TM12, first, the AvrII to XhoI segment from pJC811-*GAL2*Fr1Fr2Fr3 was inserted to the same sites of YEpADH2p to create YEpADH2p-*GAL2*Fr3. Fragment 4 was PCR amplified from the *GAL2* gene (pJC19, pJC20) with StuI and XhoI downstream.

The PCR product was digested with XhoI and inserted at the PmeI (blunt end) and XhoI sites on YEpADH2p-*GAL2*Fr3 generating YEpADH2p-*GAL2*Fr3Fr4. Then, the AvrII to XhoI fragment of YEpADH2p-*GAL2*Fr3Fr4 was cloned into the same sites of pJC811-*GAL2*Fr1Fr2 to generate pJC811-*GAL2*Fr1Fr2Fr3Fr4. The C-terminus sequence of the chimera gene was PCR amplified from *HXT1* (pJC21, pJC22) with StuI upstream and XhoI at C-terminal end. The product was digested using StuI and XhoI, and inserted to the same sites of pJC811-*GAL2*Fr1Fr2Fr3Fr4 generating the final construct pJC811-GH1 (Table 3.2).

The second version of the chimera gene was constructed by replacing GH1 internal SpeI to StuI of pJC811-GH1 with the *GAL2* counterpart from pJC811-GAL2. The resulting chimera gene and plasmid was named as *GH2* and pJC811-GH2 (Table 3.2). Both chimera genes and wild type *GAL2* were individually cloned into pJC752 and pJC812 to construct pJC752-GH1, pJC752-GH2, pJC752-GAL2, pJC812-GH1, pJC812-GH2, and pJC812-GAL2 (Table 3.2).

3.3.3. Strain construction

Three base strains were used: BY4741, yBF1587, and EBY.VW4000 (Table 3.3). yBF1587 has been obtained from Prof. Suzanne Sandmeyer's laboratory in UCI. EBY.VW4000 is a *S. cerevisiae* strain with transporters for all sugar monomers deleted and was obtained from Prof. Eckhard Boles in Goethe University Frankfurt in Germany (Wieczorke et al., 1999). Plasmid transformations into yeast strains were performed according to the Lithium acetate method (Gietz and Woods, 2001). The strains constructed are listed in Table 3.3.

Table 3.3 List of strains used in this study

Strain	Characteristics	Reference
BY4741	*MATa his3Δ1 met15Δ0 leu2Δ0 ura3Δ0*	Open Biosystems
yBF1587	*MATa his3Δ1 met15Δ0 leu2Δ0 ura3Δ0 trpΔ gre3Δ adh2Δ*	Unpublished
yOUI30	BY4741, pOUI30	This study
yOUI31	BY4741, pOUI31	This study
yOUI34	BY4741, pOUI34	This study
yLAD1H	yBF1587, pXP209-LAD-*His*	This study
yLXRTH	yBF1587, pXP206-LXRT-*His*	This study
yLXRAH	yBF1587, pXP206-LXRA-*His*	This study
y2073H	yBF1587, pXP209-2073H	This study
y1680H	yBF1587, pXP206-1680H	This study
y2178H	yBF1587, pXP206-2178H	This study
y206	yBF1587, pXP206	This study
y209	yBF1587, pXP209	This study
yXLA1	yBF1587, pETH53, pXP209-LAD, pXP206-LXRA	This study
yXLT1	yBF1587, pETH53, pXP209-LAD, pXP206-LXRT	This study
yXLA2	yBF1587, pETH52, pXP209-LAD, pXP206-LXRA	This study
yGAL	*MATa met15Δ0 leu2Δ0 ura3Δ0 trpΔ gre3Δ adh2Δ his3 Δ::*$P_{HXT7\text{-}391}$*-GAL2-*T_{CYC1}	This study
yGXLA1	yGAL, pETH53, pXP209-LAD, pXP206-LXRA	This study
EBY.VW4000	*MAT*α *Δhxt1-17 Δgal2 Δstl1Δagt1 Δmph2 Δmph3 leu2-3,112 ura3-52 trp1-289 his3-Δ1MAL2-8c SUC2*	Wieczorke *et al.* 1999
yGAL2	EBY.VW4000, *Δura3::*P_{ADH1}*-GAL2-HIS3*	This study
yGH1	EBY.VW4000, *Δura3::*P_{ADH1}*-GH1-HIS3*	This study
yGH2	EBY.VW4000, *Δura3::*P_{ADH1}*-GH2-HIS3*	This study
y812G	EBY.VW4000, pJC812-GAL2	This study
y812GH1	EBY.VW4000, pJC812-GH1	This study
y812GH2	EBY.VW4000, pJC812-GH2	This study

Gene integration into genomic DNA was carried out according to Fang et al. (2011). DNA fragments containing the genes of interest were PCR amplified from plasmids using primer pairs that include the target genome locus sequence. Amplified fragments were transformed into *S.*

cerevisiae cells using the Lithium acetate method (Gietz et al., 1995). Positive integration of the gene was confirmed by PCR with primer pairs binding at sequences flanking the integration target sequences. The size of the PCR product was confirmed using agarose gel electrophoresis.

GAL2, *GH1*, and *GH2* were integrated by the method described above. Template DNA for *GAL2*, *GH1*, and *GH2* (URA3 locus-PADH1-For, FF2288) were PCR amplified from pJC752-GAL2, pJC752-GH1, and pJC752-GH2. The *HIS3* marker (URA3-marker, FF2287) was PCR amplified from pXP220. These four linear DNA fragments were gel-purified and transformed into the host strain (EBY.VW4000) creating yGAL2, yGH1, and yGH2.

3.3.4. Media and Cultivation

Luria-Bertani (LB) medium was used for the cultivation of *E. coli* cells. Ampicilin (100 μg/ml) was used for the selection of plasmids in LB medium. *E. coli* cells were cultivated at 37°C in a 250 rpm in an agitated air shaker.

For general maintenance of *S. cerevisiae*, cells were cultivated in non-selective YPD complex medium (20 g/L dextrose, 20 g/L peptone, 10 g/L yeast extract (BD Biosciences, Sparks, MD)), or selective SDC medium (20 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate) supplemented with appropriate amino acids or nucleobases (SDC(A,U): 100 mg/L adenine hemisulfate and 100 mg/L uracil, SDC(A,T): 100 mg/L adenine hemisulfate and 100 mg/L L-tryptophan). For the selection of the *HIS3* marker, SD(-HIS) (20 g/L dextrose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base

without amino acids and ammonium sulfate, 100 mg/L adenine hemisulfate and 100 mg/L L-tryptophan, 100 mg/L uracil, 100 mg/L L-methionine, 150 mg/L L-leucine) was used. For plates, 20 g/L agar was added into the medium. Glucose was added after separately autoclaved from the other components. Yeast nitrogen base, ammonium sulfate, adenine hemisulfate, and L-tryptophan were filter-sterilized and added after autoclaving the other components.

For cultivation of yOUI30, yOUI31, and yOUI34 strains, SRC(A,T) (20g/L raffinose, 6.7g/L yeast nitrogen base without amino acids, 5g/L casamino acids, 20mg/L adenine hemisulfate salt, 20mg/L uracil, 20mg/L L-tryptophan) medium was used.

For maintenance of EBY.VW4000, YPM medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L maltose) or SMC media (20 g/L maltose, 5 g/L casamino acids, 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate with appropriate nucleobases or amino acids) were used. For integration of *GAL2*, *GH1*, and *GH2* under P_{ADH1} with the *HIS3* marker, SM(-HIS) medium (20 g/L maltose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/L L-tryptophan, 100 mg/L L-histidine) was used. For cultivation of y812-GAL2, y812-GH1, and y812-GH2 strains, modified selective SDGC(A,T) medium (10 g/L dextrose, 10 g/L galactose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/L L-tryptophan, 400mg/L L-serine, 200 mg/L L-threonine, and 20mM MES pH 5.5) or modified SGC(A,T) medium (10 g/L galactose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/L L-tryptophan, 400mg/L L-serine, 200 mg/L L-threonine, and 20mM MES pH 5.5) were used.

yOUI30, yOUI31, and yOUI34, which have L-arabinose pathway genes under the *GAL1* promoter, were cultivated in 5ml SRC(A,T) media overnight at 30°C. These cells were used to subculture 50ml SRC(A,T) and cultivated for 16 hours at 30°C. Then, D-galactose was added to a final concentration of 20g/L to induce the *GAL1* promoter and incubated for 4 hours at 30°C. Cells were harvested and pellets were washed with cold water for protein purification.

yLADH, yLXRTH, yLXRAH, y2073H, y1680H, and y2178H, which have L-arabinose pathway genes under the *PGK1* promoter, were cultivated in SDC(A,T) and SDC(A,U) media overnight at 30°C. These cells were used to subculture 50ml SDC(A,T) and SDC(A,U) media and cultivated for 12 hours at 30°C before harvesting cells. Harvested cells were washed with cold water, pelleted, and stored at -80°C.

For the xylitol assays, yBF1587 and yGAL were cultivated in SD(A,T,U,H,L,M) (20 g/L dextrose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 20 mg/L adenine hemisulfate, 20 mg/L L-tryptophan, 20 mg/L uracil, 20 mg/L 20 mg/L L-histidine, 20 mg/L L-methionine, 30 mg/L L-leucine), and yXLA1, yXLT1, yXLA2, and yGXLA1 were cultivated in SD(A,L,H) (20 g/L dextrose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 20 mg/L adenine hemisulfate, 30 mg/L L-leucine) overnight. yBF1587 and yGAL overnight cells were sub-cultured in the SDA(A,T,U,H,L,M) media with two glucose concentrations (2 g/L dextrose or 5 g/L dextrose, 20 g/L L-arabinose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 20 mg/L adenine hemisulfate, 20 mg/L L-tryptophan, 20 mg/L uracil, 20 mg/L 20 mg/L L-histidine, 20 mg/L L-methionine, and 30 mg/L L-leucine). yXLA1, yXLT1, yXLA2, and yGXLA1 overnight cells were sub-cultured in the SDA(A,L,H) media (2 g/L dextrose or 5 g/L

dextrose, 20 g/L L-arabinose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 20 mg/L adenine hemisulfate, 20 mg/L L-histidine, and 30 mg/L L-leucine). Microaerobic culture was performed in 14 ml medium in 15 ml conical tubes for 84 hours agitated at 150 rpm and 30°C. Aerobic culture was performed in 20 ml medium in 250 ml flasks for 84 hours agitated at 250 rpm and 30 °C.

yGAL, yGH1, and yGH2 were cultivated in SM(A,L,T,U) at 30°C overnight and sub-cultured into 5 ml SD(A,L,T,U) and SG(A,L,T,U). y812G, y812GH1, and y812GH2 were cultivated in SMC(A,T) medium overnight at 30°C and transferred to 5 ml SDGC(A,T) and SGC(A,T). Cells were harvested at 26 hours and 73 hours.

3.3.5. His-tagged protein purification using Ni-NTA spin columns

Collected cells were lysed using glass beads in lysis buffer (50mM sodium phosphate buffer, pH 8.0, 300mM sodium chloride, 10mM imidazole). Ni-NTA agarose (cat# 30210, Qiagen) was equilibrated with lysis buffer and the buffer was removed. Lysate was centrifuged and cleared lysate was incubated with Ni-NTA agarose at 4°C for 1 hour. Lysate Ni-NTA agarose mixture was loaded onto an empty disposable gravity column and allowed to settle. Then the soluble fraction of the sample was run through the resin bed. Ni-NTA agarose bed was washed twice using wash buffer (50 mM sodium phosphate, pH 8.0, 300 mM sodium chloride, 20 mM imidazole). Bound his-tagged proteins were eluted using elution buffer (50 mM sodium phosphate, pH 8.0, 30 mM NaCl, 250 mM imidazole). Concentrations of proteins were measured

using Bradford assay (cat# 500-0006, Bio-Rad). Glycerol (final 10%) and dithiothreitol (DTT, final 1mM) were added to purified protein samples prior to storage at -80°C until the next step.

3.3.6. Western Blot analysis

Samples were prepared in SDS sample loading buffer and run on 12% Tris-HCl gel. Separated proteins were transferred to the positively charged nylon membranes (Roche) and hybridized with monoclonal anti-polyhistidine antibody (Cat# H1029, Sigma) and rabbit anti-mouse igg-AP (Cat# 616522, Invitrogen). Hybridized antibodies were colorimetrically developed in NBT/BCIP solution (Roche).

3.3.7. Activity assay

Purified LAD1 was assayed in 100 mM Tris-HCl buffer (pH 9.0) with 0.5 mM $MgCl_2$, 4 mM NAD^+, 0.2M L-arabitol (Richard et al., 2001). Absorption at 340 nm (NADH synthesis) was measured using a Beckman Coulter DU-800 spectrophotometer. Purified ALX1 was assayed in the reverse direction using xylitol and NAD^+ as substrates (Verho et al., 2004). The assay was performed in 100 mM Tris-HCl buffer (pH 9.0) with 0.5 mM MgCl2, 2 mM NAD^+, and 0.2 M xylitol.

3.3.8. Xylitol assay

Xylitol assays were performed using a D-sorbitol/xylitol kit (Cat No. Cat.No.0670057, Roche Diagnostics GmbH). The negative control was yBF1587 for yXLA1, yXLT1, and yXLA2, and yGAL for yGXLA1. Samples were centrifuged and the supernatant was used for the xylitol detection.

3.4. **Results and Discussion**

Ethanol synthesis in *S. cerevisiae* is considered one of the sustainable alternatives to fossil fuels for transportation fuel. For economical production of fuel ethanol, utilization of all pentose sugars as well as hexose sugars is necessary. Arabinose has been much less developed compared to xylose as a carbon source for the synthesis of ethanol. For assimilation of arabinose, the fungal arabinose pathway has been less developed than the bacterial counterpart. The fungal arabinose pathway has an advantage over the bacterial pathway; it shares part of pathway with the fungal xylose pathway and accumulate less arabitol compared to the bacterial arabinose pathway. However, one bottleneck with the fungal arabinose pathway has been the redox imbalance arising from the use of phosphorylated cofactor versus non-phosphorylated cofactor (NADPH and NADH) (Figure 3.1). In this study, we tested a solution to this problem by using NADH-dependent L-xylulose reductase from the yeast *Ambrosiozyma monospora* (*ALX1*). Since this enzyme comes from a different yeast species, the gene sequence was optimized in its usage of codon and codon pairs for use in *S. cerevisiae*. Another bottleneck to ethanol synthesis via the fungal arabinose pathway in *S. cerevisiae* is the arabinose uptake from the medium to the cytosolic space. We developed a strategy to engineer Gal2, the only arabinose transporter in *S. cerevisiae*, to increase its stability in the plasma membrane.

3.4.1. Gene synthesis and plasmid construction

The first objective of this study was to enhance gene expression level for the fungal arabinose pathway. The frequently used codons and codon pairs differ greatly from organism to organism (Gutman and Hatfield, 1989). Therefore, three fungal arabinose pathway genes were optimized for enhanced expression in *S. cerevisiae* without changing the amino acid sequence: L-arabinose 4-dehydrogenase (*lad1*) from *Trichoderma reesei*, L-xylulose reductase (*lxr1*) from *T. reesei*, and L-xylulose reductase (*ALX1*) from *A. monospora*. The NADPH-dependent LXR1 was included as a control to compare to the NADH-dependent ALX1. These optimized genes were tested initially for expression and then regulated for production of the intermediate and final products. To control the expression level, the copy number of each gene in the pathway can be manipulated using the pXP vector system (Fang et al., 2011).

We constructed these genes using the Computationally Optimized DNA Assembly (CODA) method (Gutman and Hatfield, 1989; Larsen et al., 2008; Lathrop et al., 2001). The genes were designed with preferred yeast codons and codon pairs, and for easy PCR assembly. Each gene was PCR assembled through multiple steps. Initially, small fragments of 40 base pairs were synthesized by IDTDNA. These were PCR-assembled to synthesize intermediate fragments of about 200 base pairs. Intermediate fragments were again PCR-assembled to synthesize the full length genes. The genes for xylose reductase XR (encoded by *XYL1*) and xylitol dehydrogenase XDH (encoded by *XYL2*) were constructed in the same manner in the Computational Biology Research Laboratory at UCI. Detailed gene assembly procedures, sequences of oligos synthesized

by IDTDNA, and sequences of fully synthesized codon / codon-pair-optimized genes are given in Appendix A.

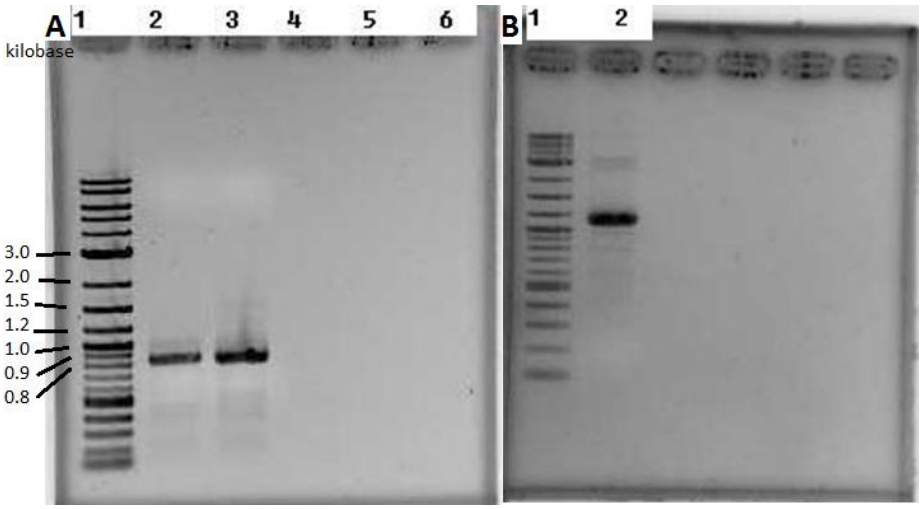


Figure 3.2 Synthetic genes were assembled and confirmed on agarose gel electrophoresis. (A) Lane 1: 2-Log DNA ladder from NEB, Lane 2: *cALX1*, Lane 3: *clxr1*, (B) Lane 1: 2-Log DNA ladder, Lane 2: *clad1* (*T. reesei*)

Three synthetic fungal arabinose pathway genes were initially cloned into pYES2 vectors under the *GAL1* promoter for initial expression tests, creating pOUI30 (*clxr1*), pOUI31 (*clad1*), and pOUI34 (*cALX1*). However, fo expression of the three synthetic genes in *S. cerevisiae*, we chose the *PGK1* promoter (a strong promoter from the glycolytic pathway) for a high expression level of each fungal arabinose pathway enzyme. The pXP vector series (Table 3.1) allows plasmid-based expression or PCR-based integration and marker excision (Fang et al., 2011). We added six adenines (and a His-tag) immediately before the start codon by PCR to aid ribosome binding

(Kozak, 1986). The PCR products were then cloned into pXP vectors and sequenced. pXP206 was used for *clxr1* and *cALX1*, and pXP209 was used for *clad1* (Table 3.2).

As controls, *lad1* (Richard et al., 2001) (pXP209-2073, pXP209-2073H for His-tagged gene) and *lxr1* (Richard et al., 2002) (pXP206-1680, pXP206-1680H for His-tagged gene) from *T. reesei* and *ALX1* (pXP206-2178, pXP206-2178H for His-tagged gene) (Verho et al., 2004) from *A. monospora* (from Dr. Peter Richard, VTT Technical Research Center of Finland) were inserted into the same expression vector system to compare expression of the optimized genes and the native fungal genes (Table 3.2).

3.4.2. Expression and activity of codon/codon-pair-optimized fungal arabinose pathway enzymes

To confirm expression and activity of the fungal arabinose pathway enzymes, assays and Western blots were performed. pOUI30 (*clxr1*), pOUI31 (*clad1*), and pOUI34 (*cALX1*), which regulate gene expression using the *GAL1* promoter, were individually transformed into *S. cerevisiae* strain BY4741 (Table 3.3). Transformants were grown in 2% synthetic raffinose liquid medium and induced for 4 hours with final 2% galactose. Cells were lysed and the his-tagged proteins were purified under native condition using Ni-NTA spin columns.

Successful expression of LAD1, LXR1 and ALX1 was confirmed on Western blots using an antibody against the his-tag (Figure 3.3). The lower band in Lane 4 may be an incomplete LAD1

or degradation product. Since the his-tag is at the N-terminus, incomplete protein products are more readily seen.

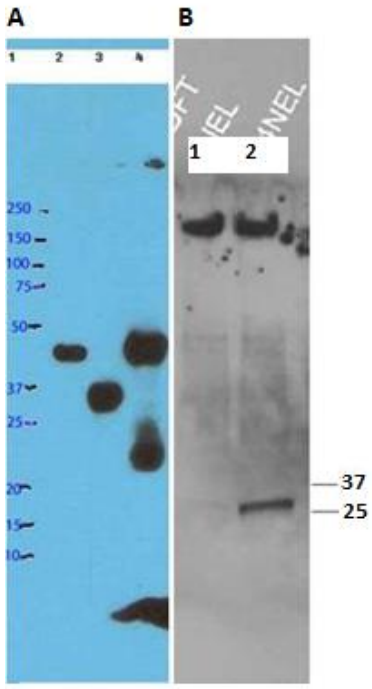


Figure 3.3 Western blot of LXR1, LAD1 and ALX1 from synthetic genes. (A) Lane 2 is a Xyl2p with 39kDa as a positive control. LXR1 is approximately 29 kDa and is in lane 3 and LAD1 is approximately 40 kDa and is the top band in lane 4. (B) Lane 1 is the negative control, which contains vector only without the *ALX1* gene. Lane 2 is ALX1 with a molecular weight of about 30.5kDa.

Activity assays with purified enzymes were performed to confirm that the fungal pathway enzymes were active (Table 3.4). The amount of protein was calculated using the Bradford assay. *In vitro* activity for purified ALX1 was obtained in a reverse reaction with xylitol and NAD as substrates (Verho et al., 2004). Absorption at 340nm was measured at 30℃ as the reaction

proceeds for 5 minutes. K_{cat} from this study is based on the purified enzyme sample. We have seen the rapid drop in activity for ALX1 after purification and avoided the loss in activity by adding glycerol in the purified enzyme and keeping samples at -80°C until the assay. The measured activity (Table 3.4) was higher than previously published activity (1115 min^{-1}) (Verho et al., 2004) but the Km for xylitol was the same. Different activity could be due to the aforementioned instability of this enzyme.

Table 3.4 Purified enzyme activity assay

Synthetic-based Enzymes		**K_m**	**K_{cat}**
LAD1	NAD	0.37 ± 0.04 mM	990 ± 110 min^{-1}
	L-arabitol	28 ± 10 mM	
ALX1 (*A. monospora*) Reverse reaction	NAD	0.22 ± 0.09 mM	2600 ± 300 min^{-1}
	Xylitol	7.9 ± 0.8 mM	1800 ± 140 min^{-1}

3.4.3. Comparison of expression level and activity of fungal arabinose pathway enzymes in crude cell extract

The use of codons as well as their pairing efficiency can be optimized according to the intracellular availability of associated tRNAs within the host microorganism (Gutman and Hatfield, 1989). Thus, the efficiency of protein translation can be enhanced. However, the amino acid sequence remains the same as native protein sequence. Therefore, codon/codon-pair

optimized genes affect only the expression level of the individual protein; specific activity of each enzyme should not change.

Strains y2073H (native *lad1*), y2178H (native *ALX1*), yLAD1H (*clad1*), and yLXRAH (*cALX1*) and negative controls, y206 (with empty pXP206) and y209 (with empty pXP209) (Table 3.3) were cultured in 2% synthetic glucose liquid medium, and harvested at the late log phase. The same dry weight of harvested cells were lysed. SDS-PAGE and Western blot were carried out with the same volume of cell extract (Figure 3.4).

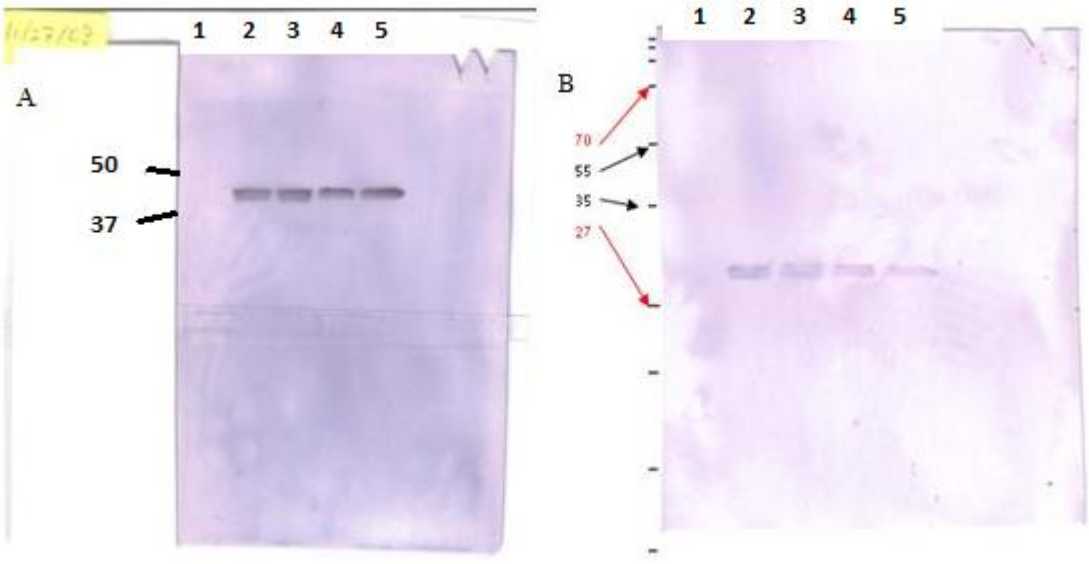


Figure 3.4 Expression level of (A) LAD1 and (B) ALX1 were compared on Western blot for the native version and synthetic-gene-based version. (A) negative control (pXP209, lane 1), native LAD1 (lane 2), synthetic-gene-based LAD1 (lane 3), 3x diluted native LAD1 (lane 4), 3x diluted synthetic-gene-based LAD1 (lane 5), (B) negative control (pXP206, lane 1), native ALX1 (lane 2), synthetic-gene-based ALX1 (lane 3), 3x diluted native ALX1 (lane 4), 3x diluted synthetic-gene-based ALX1 (lane 5).

The amount of protein was comparable between native LAD1 (lane 2) and synthetic-gene-based LAD1 (lane 3) as shown using Western blots (Figure 3.4 A). An *in vitro* activity assay (Richard et al., 2001) was performed with the synthetic-gene-based LAD1 crude cell extract and native LAD1 crude cell extract. Absorption of NADH was measured at 340nm. Synthetic-gene-based LAD1 showed similar activity compared to native LAD1 (Table 3.5).

Table 3.5 Enzyme activity assay in cell extract for LAD1 and ALX1: native versus synthetic-gene-based (n=3 for LAD1 and n=2 for ALX1)

LAD1	**Specific Activity (U/g dry cell weight)**	**ALX1**	**Specific Activity (U/g dry cell weight)**
Native	115 ± 10	**Native**	591 ± 7
Synthetic-gene-based	129 ± 10	**Synthetic-gene-based**	306 ± 7

The expression level of synthetic-gene-based ALX1 was slightly lower compared to native ALX1 as seen on the Western blot (Figure 3.4 B). Density analysis results using AlphaEaseFC (ProteinSimple, Santa Clara, CA) showed approximatley 64 % of the protein for synthetic-gene-based version relative to native. An *in vitro* activity assay was performed with the synthetic-gene-based ALX1 lysate and native ALX1 lysate (Verho et al., 2004). Absorption of NADH level was measured at 340nm. Synthetic-gene-based ALX1 showed approximately 50% of the activity compared to native ALX1 per gram cell (Table 3.5). This difference in activity between native and synthetic-gene-based ALX1 is likely due to the expression difference shown in the Western blot.

We synthesized the codon/codon-pair-optimized genes for *lad1* and *ALX1*. However, for *ALX1* and *lad1*, codon/codon-pair-optimization did not improve expression in *S. cerevisiae*

relative to the native fungal genes. We have subsequently observed little or no improvement when optimizing other fungal genes for expression in *S. cerevisiae*. For *ALX1*, the optimized gene reduced expression level on a per cell basis, and thus the measured activity on a per g cell basis.

3.4.4. Construction of *S. cerevisiae* with three fungal arabinose pathway enzymes and comparison of xylitol production level

Individual fungal arabinose pathway enzymes were evaluated in the previous set of experiments. We next evaluated the entire arabinose assimilation system by looking at the product level. While four individual genes (*XYL1*, *lad1*, *ALX1 or LXR1*, and *XYL2*) need to be expressed by the cell to produce ethanol (Figure 3.1), only three genes (*XYL1*, *lad1*, *ALX1 or LXR1)* are needed to produce the intermediate xylitol. This experiment was designed to answer the following three questions: Is *A. monospora* ALX1 better than *T. reesei* LXR1 for xylitol production? Is high *XYL1* copy number better than low *XYL1* copy number for higher xylitol production? Is overexpression of the Gal2 permease beneficial for xylitol production?

LXR1 from *T. reesei* utilizes NADPH and L-xylulose as substrates, while LAD1 requires NAD+ and L-arabitol as substrates (Figure 3.1). These create a deficit in the NADPH and NAD+ availability inside the cell. However, LXR from *A. monospora* utilizes NADH instead of NADPH, relieving this NADPH and NAD+ deficit. This should translate into higher xylitol production. The arabinose pathway introduced in the host strain does not contain XDH so xylitol is the final product that can be generated from arabinose. The xylitol produced will be excreted into the

medium (Walfridsson et al., 1995). This enzyme combination does not solve the second redox imbalance involving XR and XDH. This will need to be considered for ethanol production.

As an initial test, *cXYL1, clad1,* and *clxr1* or *cALX1* were transformed into strains yBF1587 and yGAL using 2μ or CEN/ARS plasmids (Table 3.3, Table 3.7). yGAL contains an extra copy of *GAL2* under $P_{HXT7\text{-}391}$ integrated into the genome. The strains were cultivated on a mixed carbon source medium. yBF1587 and yGAL carrying no vectors (negative controls) were cultivated in SDA(A,T,U,H,L,M), and yXLA1, yXLT1, yXLA2, and yGXLA1 were cultivated in SDC(A,L,H) with 20 g/L arabinose and two different glucose concentrations: 0.5 % glucose or 0.2 % glucose. Two glucose concentrations were evaluated because the galactose permease Gal2, that is responsible for arabinose transport, has two orders of magnitude higher affinity for glucose than arabinose (Horak and Wolf, 1997). Transport of arabinose thus starts only after glucose is depleted.

Table 3.6 Strains with different combinations of *cXYL1*, *clad*1 and *clxr1* or *cALX1*

Strain	Plasmids	Strain	Plasmids
yBF1587	None	**yGAL**	None
yXLA1	2μ *cXYL1*, 2μ *clad1*, 2μ *cALX1*	**yGXLA1**	2μ *cXYL1*, 2μ *clad1*, 2μ *cALX1*
yXLT1	2μ *cXYL1*, 2μ *clad1*, 2μ *clxr1*		
yXLA2	C/A *cXYL1*, 2μ *clad1*, 2μ *cALX1*		

Cells were cultivated under both aerobic and microaerobic conditions for 84 hours and biomass and extracellular xylitol concentrations were measured (Figure 3.5). The control strains

yBF1587 and yGAL do not contain any of the pathway genes. yBF1587 is BY4741-derived strain with *GRE3*, *ADH2*, and *TRP1* deleted.

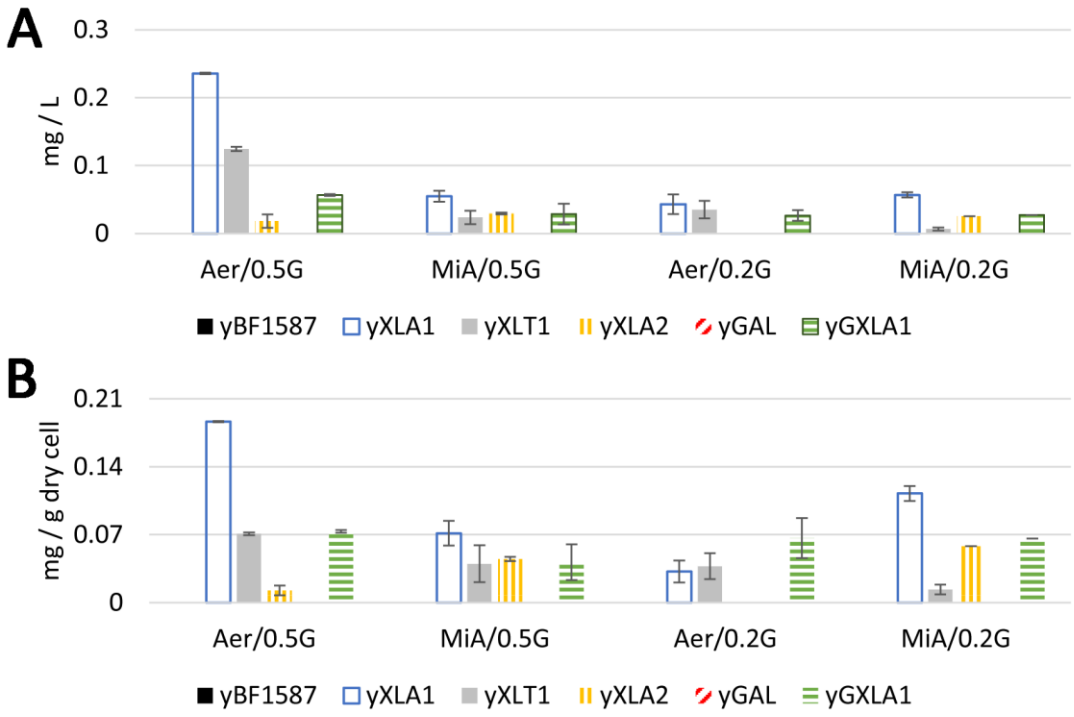


Figure 3.5 (A) Xylitol production level per cell culture volume, (B) Xylitol production level per dry cell weight. Data were averaged from two sets of experiments. * : yXLA2 for Aer/0.2G was not determined.

The comparison between the ALX1 and LXR1 can be seen by comparing yXLA1 and yXLT1 with all genes on 2μ-based plasmids. More xylitol was produced when *cALX1* was expressed with the exception of aerobic culture with 0.2% glucose. ALX1 uses NADH as cofactor generating NAD^+ and LAD1 use NAD^+ as cofactor generating NADH. Thus, NADH and NAD^+ are balanced between these two enzymes. However, XYL1 use NADPH and xylitol is the end product in current study.

Therefore NADPH is not balanced. If LXR1 is used instead of ALX1, unbalanced use of NADPH is aggravated and even NADH and NAD+ are not balanced. Thus, it is important to use ALX1 instead of LXR1 for the balanced use of cofactor and higher production of xylitol. This will likely be true for ethanol synthesis as well. However, it has to be noted that the Km of ALX1 for L-xylulose is close to half of LXR1. Moreover, specific activity for ALX1 is over 30-fold higher than LXR1 (Richard et al., 2002; Verho et al., 2004). Therefore, the higher activity and higher affinity of ALX1 might have contributed to the higher xylitol synthesis relative to the use of LXR1.

The importance of *XYL1* copy number can be witnessed via comparison of strains yXLA1 (2μ *Xyl1*) and yXLA2 (CEN/ARS *Xyl1*). The 2μ *XYL1* system produces 18-fold more xylitol than CEN/ARS *Xyl1* system under aerobic conditions with 0.5% glucose. Therefore, the higher copy number vector should be used for the complete fungal arabinose pathway strain.

From the comparison between yXLA1 and yGXLA1, *GAL2* overexpression was detrimental to both cell growth and xylitol synthesis. yBF1589 in aerobic culture with 0.5% glucose grew to OD 8.8, whereas yGAL grew only to OD 2.0. Also, yBF1589 under aerobic condition with 0.2% glucose grew to OD 4.6 and yGAL in aerobic condition with 0.2% glucose grew only to OD 1.3. It is not clear why *GAL2* overexpression slows down cell growth even without fungal arabinose pathway gene expressions. However, even specific production of xylitol in yGXLA1 is lower than in yXLA1 except for the aerobic culture condition with 0.2% glucose in the medium. It is possible that the negative effect on cell growth is due in part to the effects of Gal2 overexpression on the cell membrane.

These experiments were carried out without Xyl2p (XDH), which connects the fungal arabinose pathway to the pentose phosphate pathway. As a result, xylitol will be excreted out of

the cells and a redox imbalance still exists. Since the xylitol is the final product in our pathway, this system can be a good test system to look at the influence of two important factors, NADPH and glucose. Use of non-phosphorylated cofactor, NADH and NAD^+ was balanced by using ALX1 from *A. monospora*. However, the use of phosphorylated cofactors was left unbalanced. In need of NADPH, cells mainly depend on the oxidative pentose phosphate pathway. However, this pathway generates two NADPH at the expense of one carbon as carbon dioxide. Therefore, supply of NADPH is a crucial factor in the production of xylitol and further in the production of ethanol. Also, the presence of glucose inhibits the transport of L-arabinose. Thus, evaluation of glucose influence is also important. From the above results, we could predict that increased supply of NADPH may lead to more xylitol. Also, we were able to conclude that XR on a 2μ vector and LXR from *A. monospora* are advantageous. However, we could not draw any conclusion from the *GAL2* overexpression samples.

3.4.5. Construction of strains with Gal2 or Gal2/Hxt1 chimera proteins as the only sugar monomer transporter

The *S. cerevisiae* galactose permease Gal2 transports arabinose in addition to galactose and glucose. However, increasing arabinose transport via the native Gal2 is difficult since *GAL2* expression is repressed by glucose and Gal2 is degraded after ubiquitination and endocytosis (Horak and Wolf, 1997; Horak and Wolf, 2001; Horak and Wolf, 2005; Mylin et al., 1994). Transport of arabinose is a rate-limiting step for ethanol production from arabinose in *S.*

cerevisiae due to the absence of an efficient and stable arabinose transporter (Becker and Boles, 2003).

Glucose repression of *GAL2* expression can be avoided by employing a non-galactose-related promoter. A greater problem is the degradation of Gal2. The degradation of Gal2 starts with the tagging of the protein with a short peptide molecule, called ubiquitin. Ubiquitination requires a lysine residue in the target protein sequence. However, the responsible lysine residue is not known. Therefore, we aligned the amino acid sequence of Gal2 with other known ubiquitinated sugar transporters and a non-ubiquitinated sugar transporter (Figure 3.6) to determine potential ubiquitination sites.

Hxt5, Hxt6, and Hxt7 are all known to be ubiquitinated, while Hxt1 is not known to be ubiquitinated (Peng et al., 2003). Alignment of these 4 transporters with Gal2 showed a few consensus lysines along the sequence (Figure 3.6). In particular, three lysine residues that have been confirmed for ubiquitination of Hxt5, Hxt6, or Hxt7 are highlighted in green and are located in the N-terminus, between TM6 and TM7, and in the C-terminus (Figure 3.6, Figure 3.7). All three locations are exposed to the cytosol. Removing each lysine in these three regions of Gal2 could potentially prevent ubiquitination. However, it is also possible that the ubiquitination target lysine residues might be determined by the local structure or triggered by neighboring amino acid sequences. As a result, removing the matching lysine residue from the protein sequence may result in the ubiquitination of another lysine residue neighboring the initial lysine residue. Thus, we replaced the larger section of Gal2 protein neighboring the potential ubiquitination target lysine residue to avoid the ubiquitination of these regions.

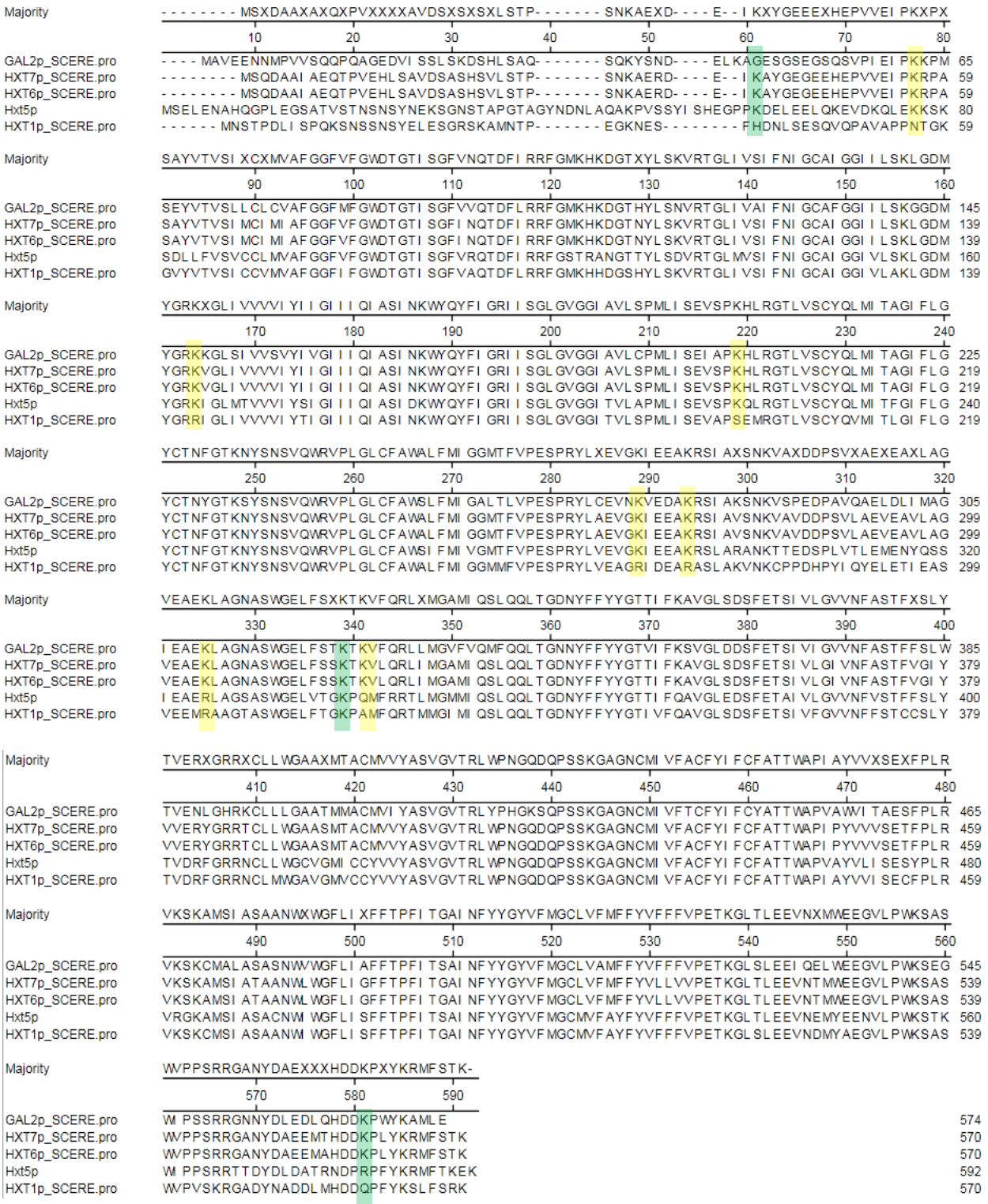


Figure 3.6 Amino acid sequence alignment of Gal2 with other transporter proteins to identify ubiquitinated lysine residues in Gal2.

Kasahara et al. proved that replacement of TM10 of Hxt2 with the counterpart from Gal2 endows galactose transport capability to Hxt2, a high affinity glucose transporter (Kasahara et al., 1996). The same group also showed that Hxt1 with its TM1, 7, 8, and 12 replaced with counterparts from Hxt2 can have identical K_m and V_{max} to Hxt2 (Kasahara and Kasahara, 2003). From these results, it appears that N- or C-terminus of the transporter protein and the loop between TM6 and TM7 are not critical in determining the sugar transport characteristics of Hxt1, Hxt2, and Gal2. Hxt1 had not been known to be subject to ubiquitination. Figure 3.6 showed that Hxt1 does not contain consensus lysine residues at the N-terminus and C-terminus. Thus, we constructed a chimera protein (GH1) by connecting the N-terminus (Hxt1), TM1 to TM6 (Gal2), Loop between TM6 and TM7 (Hxt1), TM7 to Tm12 (Gal2), and C-terminus (Hxt1) (Figure 3.7). As the second variant (GH2), we changed only both termini of Gal2 with counterparts from Hxt1. For this study, we used strain EBY.VW4000, which has all 21 hexose monomer transporter genes deleted. We constructed these hybrid transporters under the constitutive *ADH1* promoter and integrated them into the genome of the sugar monomer transporter KO strain creating yGH1 and yGH2. As a control, wild type *GAL2* under the *ADH1* promoter was integrated at the same genomic DNA locus of this strain (yGAL2).

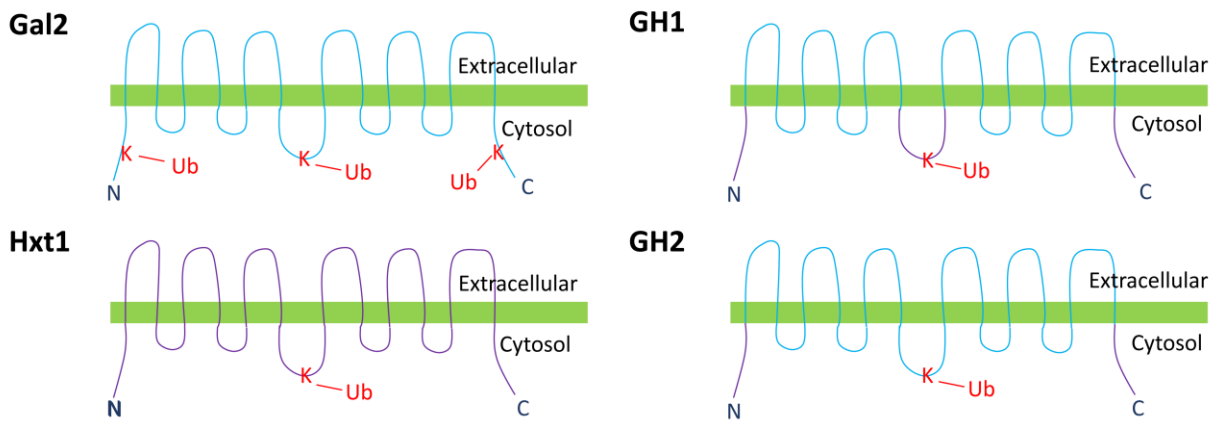


Figure 3.7 Strategy to construct the Gal2/Hxt1 hybrid transporter. This figure can also be used to understand the structure of other transporters such as Hxt2, Hxt5, Hxt6, and Hxt7 since they also have 12 transmembrane segments.

We tested these strains (yGAL2, yGH1, and yGH2) in glucose media (SD(A,L,T,U))and galactose media (SG(A,L,T,U)). However, all three strains failed to grow in either glucose or galactose media. This result might be due to the low expression level of the native and hybrid Gal2 proteins. Therefore, we expressed Gal2 and its two variants (GH1 and GH2) under the *PGK1* promoter on 2μ-based plasmids (y812G, y812GH1, and y812GH2) in the transporter KO strain, and tested growth in two media: SDGC(A,T) and SGC(A,T). The strains grew but with varying growth rates depending on the transporter (Figure 3.8). y812G grew faster than both y812GH1 and y812GH2 in both media. The final cell mass at 73 h is shown for each strain in Figure 3.8. Interestingly, y812G in SGC(A,T) media accumulated even higher mass than in SDGC(A,T) media, even though SDGC(A,T) had twice the amount of carbons relative to SGC(A,T). y812GH1 and y812GH2 started growing by 26 hour and accumulated similar cell mass to y812G (wildtype Gal2) in SDGC(A,T) medium by 73 hour: however, the cells expressing the hybrid transporters had only half the cell mass compared to y812G in SGC(A,T) medium. By 73 hour, both glucose and

galactose were consumed in both media except for y812GH1 in SDGC(A,T) medium, which had about 0.56 g/L galactose and 0.27 g/L glucose remaining. These results show that the two chimera transporters are capable of transporting both glucose and galactose although at slower rate than wild type Gal2.

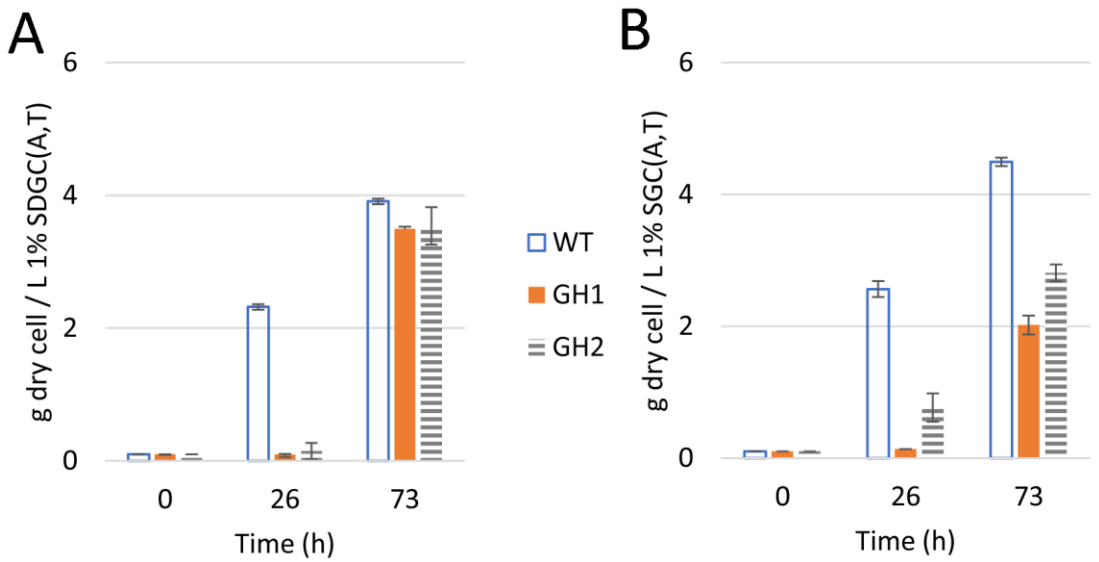


Figure 3.8. Dry cell weight for y812G, y812GH1, and y812GH2 with time in (A) SDGC(A,T) with 10 g/L glucose and 10 g/L galactose and (B) SGC(A,T) with 10 g/L galactose. Cells used up all carbon sources after 73 h except for GH1 in SDGC(A,T) medium, which had around 0.5 g/L of each glucose and galactose left.

Wild type Gal2 expressing strain had faster growth compared to Gal2/Hxt1 chimera transporter expressing strains. However, more experiments are needed to evaluate the stability of the chimera proteins and whether degradation is reduced in these hybrid transporters. The evaluation of GH1 and GH2 stability in the membrane can be done by tagging GFP at the tail of both proteins and determining their intracellular location. If the ubiquitination of these proteins

is prevented, more proteins would be displayed along the cell membrane than with the wild type Gal2 strain.

3.5. **Conclusions and Future Directions**

Thorough utilization of available sugar molecules in the biomass is a prerequisite for a cost-competitive production of fuel ethanol as well as bio-renewable chemicals in microorganisms. Although arabinose is not as abundant as xylose in nature, its proportion in several sources of hemicellulosic biomass is significant and makes utilization critical for cost-efficient production. We optimized the gene sequences of L-arabitol dehydrogenase (LAD1, *T. reesei*) and NADH-dependent L-xylulose reductase (ALX1, *A. monospora*) for expression in *S. cerevisiae*. However, the optimized *LAD1* had comparable expression to the native *LAD1* as seen with Western blot and a cell extract-based activity assay. *ALX1* had lower expression level after optimization compared to native *ALX1*. Therefore, optimization of these fungal genes was not needed for efficient expression in *S. cerevisiae*. We engineered the initial part of the arabinose assimilation pathway in *S. cerevisiae* using the recoded genes for xylose reductase (XYL1, *P. stipitis*), L-arabitol dehydrogenase (LAD1, *T. reesei*) and NADH-dependent L-xylulose reductase (ALX1, *A. monospora*). Xylitol measurements showed that the high copy *XYL1*/*ALX1* strain produces more xylitol than the high copy *XYL1*/*LXR1* strain and the low copy *XYL1*/*ALX1* strain. This is likely due in part to the redox balance between LAD1 and ALX1 in the use of NADH and NAD^+. Overexpression of the Gal2 transporter slowed down cell growth and did not help to increase xylitol production except in the aerobic culture with 0.2% glucose.

Since the arabinose is imported into the cells (by the Gal2 transporter) only after glucose is consumed and its uptake by cells is a limiting step in the ethanol synthesis from this sugar in *S. cerevisiae (Becker and Boles, 2003; Subtil and Boles, 2012)*, solving this problem will be an important step for the cost-effective use of hemicellulose in *S. cerevisiae*. Thus, we constructed

two chimera transporter using Hxt1 and Gal2 for decreasing the ubiquitination and degradation of the transporter. Both hybrids retained the ability to transport glucose and galactose, and allowed growth on both of these sugars. Continuing work with these engineered transporters will include evaluation of protein stability by localizing proteins via GFP-based fluorescence imaging followed by the evaluation of cell growth on arabinose as the single carbon source. Recently, there was a report providing evidence for Hxt1 ubiquitination and degradation (Roy et al., 2014). The target lysines were at N-terminus and might have made GH1 and GH2 more unstable during our growth study (Figure 3.8). The authors were able to prevent ubiquitination and degradation by K12A and K59A mutation in Hxt1. The same mutations on GH1 and GH2 could increase stability of two hybrid transoprters. Our efforts to prevent ubiquitination of the arabinose transporter via engineering of Gal2 aim to increase arabinose uptake. This would be beneficial for the biofuel and biochemical industries.

3.6. **References**

Becker, J., Boles, E., 2003. A modified Saccharomyces cerevisiae strain that consumes L-arabinose and produces ethanol. Applied and Environmental Microbiology. 69**,** 4144-4150.

Bera, A. K., Sedlak, M., Khan, A., Ho, N. W. Y., 2010. Establishment of l-arabinose fermentation in glucose/xylose co-fermenting recombinant Saccharomyces cerevisiae 424A(LNH-ST) by genetic engineering. Applied Microbiology and Biotechnology. 87**,** 1803-1811.

Bettiga, M., Bengtsson, O., Hahn-Hagerdal, B., Gorwa-Grauslund, M. F., 2009. Arabinose and xylose fermentation by recombinant Saccharomyces cerevisiae expressing a fungal pentose utilization pathway. Microbial Cell Factories. 8.

Dumon, C., Song, L. T., Bozonnet, S., Faure, R., O'Donohue, M. J., 2012. Progress and future prospects for pentose-specific biocatalysts in biorefining. Process Biochemistry. 47**,** 346-357.

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2011. A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 28**,** 123-136.

Gietz, R. D., Schiestl, R. H., Willems, A. R., Woods, R. A., 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast. 11**,** 355-360.

Gietz, R. D., Woods, R. A., 2001. Genetic transformation of yeast. Biotechniques. 30**,** 816-831.

Gutman, G. A., Hatfield, G. W., 1989. Nonrandom utilization of codon pairs in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America. 86**,** 3699-3703.

Hahn-Hagerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I., Gorwa-Grauslund, M. F., 2007a. Towards industrial pentose-fermenting yeast strains. Applied Microbiology and Biotechnology. 74**,** 937-953.

Hahn-Hagerdal, B., Karhumaa, K., Jeppsson, M., Gorwa-Grauslund, M. F., 2007b. Metabolic engineering for pentose utilization in Saccharomyces cerevisiae. Biofuels. 108**,** 147-177.

Horak, J., Wolf, D. H., 1997. Catabolite inactivation of the galactose transporter in the yeast Saccharomyces cerevisiae: Ubiquitination, endocytosis, and degradation in the vacuole. Journal of Bacteriology. 179**,** 1541-1549.

Horak, J., Wolf, D. H., 2001. Glucose-induced monoubiquitination of the Saccharomyces cerevisiae galactose transporter is sufficient to signal its internalization. Journal of Bacteriology. 183**,** 3083-3088.

Horak, J., Wolf, D. H., 2005. The ubiquitin ligase SCFGrr1 is required for Gal2p degradation in the yeast Saccharomyces cerevisiae. Biochemical and Biophysical Research Communications. 335**,** 1185-1190.

Jeffries, T. W., 2006. Engineering yeasts for xylose metabolism. Current Opinion in Biotechnology. 17**,** 320-326.

Jin, Y. S., Jeffries, T. W., 2004. Stoichiometric network constraints on xylose metabolism by recombinant Saccharomyces cerevisiae. Metabolic Engineering. 6**,** 229-238.

Jin, Y. S., Lee, T. H., Choi, Y. D., Ryu, Y. W., Seo, J. H., 2000. Conversion of xylose to ethanol by recombinant Saccharomyces cerevisiae containing genes for xylose reductase and xylitol dehydrogenase from Pichia stipitis. Journal of Microbiology and Biotechnology. 10**,** 564-567.

Karhumaa, K., Garcia Sanchez, R., Hahn-Hagerdal, B., Gorwa-Grauslund, M.-F., 2007. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant Saccharomyces cerevisiae. Microbial Cell Factories. 6.

Karhumaa, K., Wiedemann, B., Hahn-Hagerdal, B., Boles, E., Gorwa-Grauslund, M. F., 2006. Co-utilization of L-arabinose and D-xylose by laboratory and industrial *Saccharomyces cerevisiae* strains. Microbial Cell Factories. 5.

Kasahara, M., Shimoda, E., Maeda, M., 1996. Transmembrane segment 10 is important for substrate recognition in Ga12 and Hxt2 sugar transporters in the yeast Saccharomyces cerevisiae. Febs Letters. 389**,** 174-178.

Kasahara, T., Kasahara, M., 2003. Transmembrane segments 1, 5, 7 and 8 are required for high-affinity glucose transport by Saccharomyces cerevisiae Hxt2 transporter. Biochemical Journal. 372**,** 247-252.

Kim, S. R., Park, Y.-C., Jin, Y.-S., Seo, J.-H., 2013. Strain engineering of Saccharomyces cerevisiae for enhanced xylose metabolism. Biotechnology Advances. 31**,** 851-861.

Kotter, P., Ciriacy, M., 1993. Xylose fermentation by *Saccharomyces cerevisiae*. Applied Microbiology and Biotechnology. 38**,** 776-783.

Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell. 44**,** 283-292.

Kumar, S., Singh, S. P., Mishra, I. M., Adhikari, D. K., 2009. Recent advances in production of bioethanol from lignocellulosic biomass. Chemical Engineering & Technology. 32**,** 517-526.

Kuyper, M., Hartog, M. M. P., Toirkens, M. J., Almering, M. J. H., Winkler, A. A., van Dijken, J. P., Pronk, J. T., 2005. Metabolic engineering of a xylose-isomerase-expressing Saccharomyces cerevisiae strain for rapid anaerobic xylose fermentation. Fems Yeast Research. 5**,** 399-409.

Larsen, L. S. Z., Wassman, C. D., Hatfield, G. W., Lathrop, R. H., 2008. Computationally Optimised DNA Assembly of synthetic genes. International journal of bioinformatics research and applications. 4**,** 324-36.

Lathrop, R. H., Sazhin, A., Sun, Y., Steffin, N., Irani, S. S., 2001. A multi-queue branch-and-bound algorithm for anytime optimal search with biological applications. Genome informatics. International Conference on Genome Informatics. 12**,** 73-82.

Madhavan, A., Srivastava, A., Kondo, A., Bisaria, V. S., 2012. Bioconversion of lignocellulose-derived sugars to ethanol by engineered Saccharomyces cerevisiae. Critical Reviews in Biotechnology. 32**,** 22-48.

Madhavan, A., Tamalampudi, S., Srivastava, A., Fukuda, H., Bisaria, V. S., Kondo, A., 2009. Alcoholic fermentation of xylose and mixed sugars using recombinant Saccharomyces cerevisiae engineered for xylose utilization. Applied Microbiology and Biotechnology. 82**,** 1037-1047.

Matsushika, A., Inoue, H., Kodaki, T., Sawayama, S., 2009. Ethanol production from xylose in engineered Saccharomyces cerevisiae strains: current state and perspectives. Applied Microbiology and Biotechnology. 84**,** 37-53.

Matsushika, A., Watanabe, S., Kodaki, T., Makino, K., Inoue, H., Murakami, K., Takimura, O., Sawayama, S., 2008. Expression of protein engineered NADP plus -dependent xylitol dehydrogenase increases ethanol production from xylose in recombinant Saccharomyces cerevisiae. Applied Microbiology and Biotechnology. 81**,** 243-255.

Metz, B., Mojzita, D., Herold, S., Kubicek, C. P., Richard, P., Seiboth, B., 2013. A novel l-xylulose reductase essential for l-arabinose catabolism in *Trichoderma reesei*. Biochemistry. 52**,** 2453-2460.

Mielenz, J. R., 2001. Ethanol production from biomass: technology and commercialization status. Current Opinion in Microbiology. 4**,** 324-329.

Mylin, L. M., Bushman, V. L., Long, R. M., Yu, X., Lebo, C. M., Blank, T. E., Hopper, J. E., 1994. Sip1 is a catabolite repression-specific negative regulator of gal gene-expression. Genetics. 137**,** 689-700.

Peng, J. M., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D. M., Marsischky, G., Roelofs, J., Finley, D., Gygi, S. P., 2003. A proteomics approach to understanding protein ubiquitination. Nature Biotechnology. 21**,** 921-926.

Richard, P., Londesborough, J., Putkonen, M., Kalkkinen, N., Penttila, M., 2001. Cloning and expression of a fungal L-arabinitol 4-dehydrogenase gene. Journal of Biological Chemistry. 276**,** 40631-40637.

Richard, P., Putkonen, M., Vaananen, R., Londesborough, J., Penttila, M., 2002. The missing link in the fungal L-arabinose catabolic pathway, identification of the L-xylulose reductase gene. Biochemistry. 41**,** 6432-6437.

Richard, P., Verho, R., Putkonen, M., Londesborough, J., Penttila, M., 2003. Production of ethanol from L-arabinose by Saccharomyces cerevisiae containing a fungal L-arabinose pathway. Fems Yeast Research. 3**,** 185-189.

Roy, A., Kim, Y.-B., Cho, K. H., Kim, J.-H., 2014. Glucose starvation-induced turnover of the yeast glucose transporter Hxt1. Biochimica et biophysica acta. 1840**,** 2878-85.

Shen, M. W. Y., Fang, F., Sandmeyer, S., Da Silva, N. A., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 29**,** 495-503.

Sims, R. E. H., Mabee, W., Saddler, J. N., Taylor, M., 2010. An overview of second generation biofuel technologies. Bioresource Technology. 101**,** 1570-1580.

Subtil, T., Boles, E., 2012. Competition between pentoses and glucose during uptake and catabolism in recombinant Saccharomyces cerevisiae. Biotechnology for Biofuels. 5.

Van Vleet, J. H., Jeffries, T. W., 2009. Yeast metabolic engineering for hemicellulosic ethanol production. Current Opinion in Biotechnology. 20**,** 300-306.

Verho, R., Putkonen, M., Londesborough, J., Penttila, M., Richard, P., 2004. A novel NADH-linked L-xylulose reductase in the L-arabinose catabolic pathway of yeast. Journal of Biological Chemistry. 279**,** 14746-14751.

Walfridsson, M., Hallborn, J., Penttila, M., Keranen, S., Hahnhagerdal, B., 1995. Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the *TKL1* and *TAL1* genes encoding the pentose-phosphate pathway enzymes transketolase and transaldolase. Applied and Environmental Microbiology. 61**,** 4184-4190.

Watanabe, S., Abu Saleh, A., Pack, S. P., Annaluru, N., Kodaki, T., Makino, K., 2007a. Ethanol production from xylose by recombinant Saccharomyces cerevisiae expressing protein engineered NADP(+)-dependent xylitol dehydrogenase. Journal of Biotechnology. 130**,** 316-319.

Watanabe, S., Abu Saleh, A., Pack, S. P., Annaluru, N., Kodaki, T., Makino, K., 2007b. Ethanol production from xylose by recombinant Saccharomyces cerevisiae expressing protein-engineered NADH-preferring xylose reductase from Pichia stipitis. Microbiology-Sgm. 153**,** 3044-3054.

Watanabe, S., Kodaki, T., Makino, K., 2005. Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. Journal of Biological Chemistry. 280**,** 10340-10349.

Wieczorke, R., Krampe, S., Weierstall, T., Freidel, K., Hollenberg, C. P., Boles, E., 1999. Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in Saccharomyces cerevisiae. Febs Letters. 464**,** 123-128.

Wisselink, H. W., Toirkens, M. J., Berriel, M. D. F., Winkler, A. A., van Dijken, J. P., Pronk, J. T., van Maris, A. J. A., 2007. Engineering of Saccharomyces cerevisiae for efficient anaerobic alcoholic fermentation of L-arabinose. Applied and Environmental Microbiology. 73**,** 4881-4891.

Chapter 4.

Biosynthesis of Dihydromonacolin L, a Precursor to Lovastatin, in *Saccharomyces cerevisiae*

4.1 Abstract

The yeast *Saccharomyces cerevisiae* shows great promise for the production of fungal polyketides, including pharmaceuticals such as lovastatin. The limiting step in the synthesis of dihydromonacolin L (DML), a precursor to lovastatin, in *S. cerevisiae* is the release of DML from the lovastatin nonaketide synthase (LovB from *Aspergillus terreus*) catalyzed by a thioesterase enzyme. We compared five different thioesterases for DML synthesis levels in *S. cerevisiae* and three of the five allowed DML detection *in vivo*. The thioesterase AptB (from *Aspergillus nidulans*) produced up 3 mg/L DML, and up to 9-fold higher levels of DML relative to PKS13 TE, the only thioesterase for which release has been reported. To determine why AptB was superior to the other thioesterases for DML synthesis, we compared both transcription and translation levels of the five thioesterases. We also considered the effect of copy number of the *lovB* gene on DML levels. The strain carrying *lovB* on a multi-copy 2μ-based plasmid produced 30% more DML relative to the strain with a single integrated copy of *lovB*.

4.2 Introduction

Polyketides are secondary metabolites from microorganisms with diverse biological activities and many of them have found commercial application (Crawford and Townsend, 2010). Most native polyketide producing organisms are difficult to use for polyketide production due to difficulties in cultivation and the lack of tools for genetic manipulation (Pfeifer and Khosla, 2001). *Saccharomyces cerevisiae* is an excellent candidate for heterologous expression of polyketide enzymes, particularly those of fungal origin. Strengths include its completely sequenced genome, a vast library of genetically specified strains, various antibiotic and auxotrophic selection markers, efficient homologous recombination methods, immunotags, relatively fast growth rate, and GRAS (Generally Regarded As Safe) status (Bonekamp and Oosterom, 1994; Romanos et al., 1992). Moreover, 6-methylsalicylic acid (6-MSA), which is a polyketide metabolite naturally produced by *Penicillium patulum*, was produced in *S. cerevisiae* at 2.3 g/L quantities (Ching, 2005).

One example of a commercially successful polyketide is lovastatin, a cholesterol lowering drug natively produced by *Aspergillus terreus* (Figure 2.5) (Hendrickson et al., 1999; Kennedy et al., 1999; Maggon, 2005). Dihydromonacolin L (DML) is a precursor to lovastatin and is synthesized by a holo enzyme lovastatin nonaketide synthase (LovB), which is an iterative type I polyketide synthase, and a dissociative enoyl reductase (LovC). The activation of LovB to holo enzyme requires phosphopantetheinylation of its acyl carrier protein (ACP) domain by a discrete phosphopantetheinyl (P-pant) transferase. Our lab showed the phophopantetheinylation of LovB in *S.cerevisiae* by using two different P-pant transferases, Sfp and NpgA, from *Bacillus subtilis* and *Aspergillus nidulans,* respectively (2009). However, the expression of *lovB*, *lovC*, and *npgA* alone

did not allow synthesis of DML in *S. cerevisiae*. To troubleshoot this problem, Ma *et al.* (2009) employed a thioesterase (TE) domain from a fungal polyketide synthase PKS13 and PKS 4 (zearalenone biosynthesis pathway in *Gibberella zeae* and *Gibberella fujikuroi*) in the DML synthesis reaction *in vitro* and detected DML. This showed that the failure of product offloading from LovB automatically blocks the following synthesis of the polyketide. PKS13 and PKS4 thioesterase domains are responsible for the offloading of the product following its macrolactonization. Although not published, the same group was able to release DML from lovB using the hpm3 (hypothemycin synthesis pathway in *Hypomyces subiculosus*) (Reeves et al., 2008) thioesterase domain both *in vitro* and *in vivo.*

The objective of this study was to improve the release of DML from LovB and, thereby, improve the production of DML in *S. cerevisiae*. LovB does not have a thioesterase domain and, at that time, no discrete TE had been found in the lovastatin biosynthesis pathway. Furthermore, DML has been successfully synthesized only in *A. terreus* and *A. nidulans* (Hendrickson et al., 1999; Kennedy et al., 1999), where discrete thioesterases (e.g., AptB and ACTE in *A. nidulans* (Szewczyk et al., 2008)) might have been responsible for the offloading of DML from LovB. We thus introduced discrete thioesterases as well as TE domains for DML release. We compared five different thioesterases: PKS13 TE domain, hpm3 TE domain, Rdc1 TE domain, AptB, and VrtG. Hpm3 is a fungal iterative type I nonreducing polyketide synthase (NR-PKS) from the hypothemycin synthesis gene cluster in *Hypomyces subiculosus* (Reeves et al., 2008). Its TE domain is a macrolatonizing TE and its product, hypothemycin maintains a bicyclic ring structure just like zearalenone, in which PKS13 is involved. Rdc1 is an iterative type I nonreducing polyketide synthase (NR-PKS) from the radicicol biosynthesis pathway in *Pochonia*

chlamydosporia (Reeves et al., 2008) and the TE domain from this PKS was used. AptB is a discrete β-lactamase superfamily thioesterase from the asperthecin biosynthesis pathway of *Aspergillus nidulans* (Szewczyk et al., 2008). We also included another discrete thioesterase, VrtG from the viridicatumtoxin biosynthesis pathway in *Penicillium aethiopicum* (Chooi et al., 2010). VrtG is a β-lactamase family enzyme like AptB and ACTE. We introduced the five thioesterases into yeast and compared the DML synthesis in *S. cerevisiae* in collaboration. We also looked at the effects of copy number of the *lovB* gene by comparing strains carrying a multi-copy plasmid or one integrated copy.

4.3 Materials and Methods

4.3.1. Molecular biology techniques

Plasmids were isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific). All PCR products were purified using either the Zymoclean™ gel DNA Recover Kit (Zymo Research Corporation) or the QIAquick PCR Purification Kit (Qiagen). The Rapid DNA Ligation Kit (Thermo Scientific) was used for the ligation of DNA fragments. Restriction endonucleases and Taq DNA Polymerase were purchased from New England Biolabs Inc. KOD Hot Start DNA Polymerase (EMD Millipore) was used for high fidelity PCR while all other PCR reactions were carried out using Taq DNA Polymerase. All oligo nucleotides were synthesized by Integrated DNA Technologies, Inc. Synthesized genes were sequenced by Eton Bioscience Inc. *E. coli* competent cells were created using the calcium chloride method (Sambrook and Russell, 2001). All primer sequences are given in Table C.1 (Appendix C).

Plasmid transformation into *S. cerevisiae* was performed following the modified lithium acetate method (Gietz and Woods, 2001). Integration of linear DNA into *S. cerevisiae* cells was performed following the high-efficiency transformation method (Gietz et al., 1995).

4.3.2. Vector construction

YEpADH2p was created by replacing P_{ADH2}-*lovB*-T_{ADH2} of YEpLovB (Lee, 2006) with P_{ADH2}-T_{ADH2} from pKOS12-122c(-Nde) using NotI and KpnI. YEpADH2p was digested with NotI and blunted using T4 DNA polymerase (New England Biolabs Inc. Ipswich, MA). P_{ADH2}-T_{ADH2} was obtained after digestion of the blunted linearized YEpADH2p with XhoI. pBF3174 (Shen et al., 2012) was digested with NdeI, blunted with T4 DNA polymerase and digested with SalI (SalI has overhang that is compatible to XhoI) to remove P_{GAL1}-Rluc-T_{CYC1}. P_{ADH2}-T_{AHD2} and the backbone part of pBF3174 were ligated using Rapid DNA Ligation Kit (Thermo Scientific, Waltham, MA) to create pJC702. NpgA was digested from pKMLADH2p-npgA (Lee et al., 2009) using SalI. This fragment was ligated into the XhoI site of pJC702 to create pJC702-NpgA. pJC vectors were created by addition of PmeI and RsrII restriction sites between SpeI and XhoI (between promoter and terminator) of pXP vectors. All pJC series vectors are listed in Table 4.1 (Figure 4.1).

pIM11 was created by ligating the front and back homology sequences flanking the *met17* deletion locus of BY4741 into pJC811. The front homology forward primer (MET17FrFor) contained a NdeI site followed by a NruI site, and the front homology reverse primer (MET17FrRev) contained a NdeI site at the 5' end of the primer. All primer sequences are provided in Table C.1. (Appendix C) The front homology fragment was synthesized by PCR with BY4741 genomic DNA as the template. The PCR product was digested using NdeI and ligated into the NdeI site of pJC811 and pJC841. The rear homology fragment was synthesized via PCR (primers MET17ReFor and MET17ReRev) from BY4741 genomic DNA. MET17ReFor contained KpnI (at the 5' end of the primer) and MET17ReRev contained EcoRI, SwaI, and NotI (at the 5' end

of the primer). The PCR product was digested using KpnI and EcoRI, and ligated into the same sites in two vectors (pJC811 and pJC841) with the front homology replacing the 2μ replication origin. This created the final integrating vectors pIM11 and pIM41, respectively (Figure 4.1).

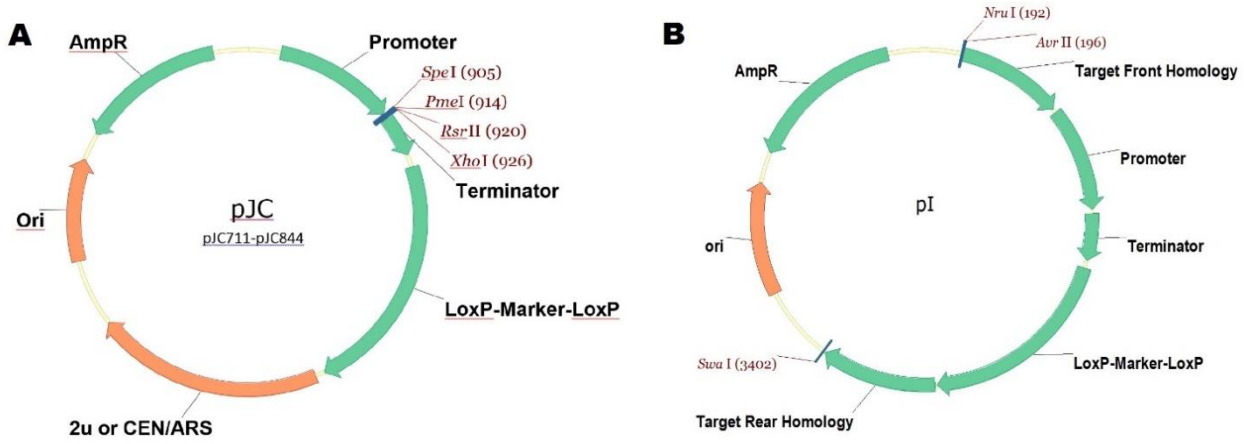


Figure 4.1 Vector maps for (A) pJC and (B) pI. (A) pJC vectors (from pJC711 to pJC844) were created (explained in chapter 3) by adding multiple cloning sites (MCS) between promoters and terminators of pXP series vectors (Shen et al., 2012). Promoter included those from *PGK1*, *ADH1*, and *ADH2*, but only the *ADH2* promoter was used in this study. The *CYC1* terminator was used for all vectors from pJC711 to pJC844. pJC702 is the only vector with the *ADH2* terminator. pXP/pJC vectors with *LEU2-d8, URA3, HIS3, TRP1* markers are available. (B) pI vector is an integration vector. pI holds two sections that are homologous to target genome sequence. These homology sequences flank the expression cassette and LoxP-marker-LoxP sequence to allow integration into the target locus through homologous recombination. The current study targeted the *MET17* and *URA3* locus for *lovB* integration.

The *lovB* sequence from the 5 prime end to +1230 (primers 012209 lovB for, 012209 lovB rev Bsu36I) was amplified by PCR with NotI and PmeI restriction enzyme sites at the 5 prime end of the ORF from YEpLovB (Lee et al., 2009). The PCR product was digested using NotI and Bsu36I after gel-purification, then inserted into YEpLovB digested with NotI and Bsu36I to create a

temporary holding vector pLovB. The *lovB* ORF was obtained from pLovB using PmeI and RsrII and ligated into the same sites in pJC842 and pIM44a to create pJC842-LovB and pIM44-LovB. *lovC* (LovC_For_SpeI, LovC_Rev_RsrII) was amplified from YEpLovC (Lee et al., 2009). The PCR product was gel-purified, digested with SpeI and RsrII, and ligated into pJC841 at the same sites creating pJC841-LovC.

Five thioesterase genes were obtained from Professor Yi Tang's group. Hpm3 TE (hpm3TE_for_AvrII_Kozak, hpm3TE_rev_noHistag_PmeI) was PCR amplified from pXK66 and gel-purified. pXK66, pPKS13TE, XW06, pYET-VrtG, and pRdc1TE were obtained from Prof. Yi Tang's laboratory at UCLA. The purified DNA was digested using AvrII and PmeI and ligated into SpeI and PmeI sites of pJC843. PKS13TE (PKS13TEFo2, PKS13TERev) was PCR amplified from pPKS13TE and gel-purified. The purified DNA was digested using SpeI and RsrII and ligated into SpeI and RsrII sites of pJC844. AptB (AptB_SpeI_F, AptB_RsrII_R) was PCR amplified from XW06 and gel-purified. The purified DNA was digested using SpeI and RsrII and ligated into SpeI and RsrII sites of pJC843. VrtG (VrtG_SpeI_F, VrtG_XhoI_R) was PCR amplified from pYET-VrtG and gel-purified. The purified DNA was digested using SpeI and XhoI and ligated into SpeI and XhoI sites of pJC843. Rdc1 TE (RadicicolTE_For, RadicicolTE_Rev) was PCR amplified from pRdc1TE and gel-purified. The purified DNA was digested using SpeI and XhoI and ligated into SpeI and XhoI sites of pJC843. Six histidine amino acid residues were added at the end of each thioesterase (Hpm3TE_rev_PmeI, PKS13TE-His_Rev, AptB-His_Rev, VrtG-His_Rev, Rdc1TE-His_Rev**)** for Ni-NTA purification and cloned into pJC843 in the same manner (Table 4.1).

Table 4.1. List of plasmids and strains used for DML synthesis

Plasmids	**Characteristics**	**Reference**
pBF3060	P_{GAL1}-CreA-T_{CYC1}, 2μ, *URA3* marker	Fang et al., 2011
pBF3174	P_{GAL1}-Rluc-T_{CYC1}, CEN/ARS, *LEU2* marker	Shen et al., 2012
pJC702-NpgA	P_{ADH2}-NpgA-T_{ADH2}, CEN/ARS, *LEU2* marker	This study
pJC841-LovC	P_{ADH2}-lovC-T_{CYC1}, 2μ, *LEU2* marker	This study
YEplovB	P_{ADH2}-*lovB*-T_{ADH2}, 2μ, *URA3* marker	Lee et al, 2009
pJC742-LovB	P_{ADH2}-*lovB*-T_{CYC1}, CEN/ARS, *URA3* marker	This study
pJC842-LovB	P_{ADH2}-*lovB*-T_{CYC1}, 2μ, *URA3* marker	This study
pIM44-LovB	P_{ADH2}-*lovB*-T_{CYC1}, Integration vector targeting *MET17* locus, *HIS3*	This study
pPKS13TE	P_{ADH2}-PKS13TE-T_{ADH2}	Unpublished
pXK66	P_{ADH2}-hpm3TE-T_{ADH2}, 2μ, *LEU2*	Unpublished
pRdc1TE	P_{ADH2}-Rdc1TE-T_{ADH2}	Unpublished
pYET-VrtG	P_{ADH2}-VrtG-T_{ADH2}, 2μ, *TRP1*	Unpublished
XW06	P_{ADH2}-AptB-T_{ADH2}, 2μ, *TRP1*	Unpublished
pJC843-PKS13TE	P_{ADH2}-PKS13TE-T_{CYC1}, 2μ, *TRP1*	This study
pJC843-hpm3TE	P_{ADH2}-hpm3TE-T_{CYC1}, 2μ, *TRP1*	This study
pJC843-Rdc1TE	P_{ADH2}-Rdc1TE-T_{CYC1}, 2μ, *TRP1*	This study
pJC843-AptB	P_{ADH2}-AptB-T_{CYC1}, 2μ, *TRP1*	This study
pJC843-VrtG	P_{ADH2}-VrtG-T_{CYC1}, 2μ, *TRP1*	This study
pJC843-PKS13TE-His	pJC843-PKS13TE, with His-tag	This study
pJC843-hpm3TE-His	pJC843-hpm3TE, with His-tag	This study
pJC843-Rdc1TE-His	pJC843-Rdc1TE, with His-tag	This study
pJC843-AptB-His	pJC843-AptB, with His-tag	This study
pJC843-VrtG-His	pJC843-VrtG, with His-tag	This study
Yeast strains	**Characteristics**	Reference
BJ5464	*MAT* α *ura3-52 his3-* Δ *200 leu2-* Δ *1 trp1 pep4::HIS3 prb1* Δ *1.6R can1 GAL*	Jones, 1990
BJN	BJ5464, YDRWTy1-5::P_{ADH2}-NpgA-T_{ADH2}	This study
BJNC	BJN, *Δpep4*::*Δhis3*:: P_{ADH2}-LovC-T_{CYC1}-*LEU2*	This study
BJNCA	*Δleu2*::P_{ADH2}-AptB-T_{CYC1}-*LEU2*	This study
BJNCBA	BJNCB, *Δleu2*::P_{ADH2}-*AptB*-T_{CYC1}-*LEU2*	This study
BJNCB	BJNC, *Δmet17*::P_{ADH2}-*lovB*-T_{CYC1}-*HIS3*	This study
BJ-CLovB	BJ5464, pJC742-LovB	This study
BJ-2LovB	BJ5464, pJC842-LovB	This study
BJ-P	BJ5464, pJC843-PKS13TE-His	This study
BJ-H	BJ5464, pJC843-hpm3TE-His	This study
BJ-R	BJ5464, pJC843-Rdc1TE-His	This study
BJ-A	BJ5464, pJC843-AptB-His	This study
BJ-V	BJ5464, pJC843-VrtG-His	This study
BJNC-PB	BJNC, pJC843-PKS13TE, pJC842-LovB	This study
BJNC-HB	BJNC, pJC843-hpm3TE, pJC842-LovB	This study
BJNC-RB	BJNC, pJC843-PKS13TE, pJC842-LovB	This study
BJNC-AB	BJNC, pJC843-AptB, pJC842-LovB	This study
BJNC-VB	BJNC, pJC843-VrtG, pJC842-LovB	This study
BJNCB-P	BJNCB, pJC843-PKS13TE	This study
BJNCB-H	BJNCB, pJC843-hpm3TE	This study
BJNCB-R	BJNCB, pJC843-Rdc1TE	This study
BJNCB-A	BJNCB, pJC844-AptB	This study
BJNCB-V	BJNCB, pJC844-VrtG	This study
BJNCA-B	BJNCA, pJC842-LovB	This study

4.3.3. Strain construction

BJ5464 were used as the base *S. cerevisiae* strain. We followed previously reported gene integration methods (Fang et al., 2011) using double crossover homologous recombination for the construction of the strains used in this study. Strains constructed in this chapter are listed in Table 4.1.

S. cerevisiae strain BJN was constructed by integrating a copy of *npgA* under the *ADH2* promoter into the genome of BJ5464. P_{ADH2}-*npgA*-T_{ADH2} and LoxP-*LEU2*-LoxP were PCR amplified (YDRWTy1-5-ADH2PF, FF2325) from pJC702-NpgA and gel purified. The linear DNA fragment was integrated into the YDRWTy1-5 locus of BJ5464 via homologous recombination (Gietz et al., 1995). pBF3060 containing CreA recombinase under *GAL1* promoter was transformed into BJN. *LEU2* was removed by expressing CreA recombinase as reported by Fang et al. (Fang et al., 2011) pBF3060 was lost using 5-fluoroorotic acid (5-FOA) selection creating BJN.

BJNC was created by integrating a copy of LovC under the *ADH2* promoter into the genome of BJN. P_{ADH2}-LovC-T_{CYC1} and LoxP-*LEU2*-LoxP were PCR amplified (HIS3-ADH2-For and HIS3-LoxP-Rev) from pJC841-LovC and gel purified. The linear DNA fragment was integrated into the *Δpep4*::*HIS3* locus creating BJNC. A linear DNA fragment carrying P_{ADH2}-*lovB*-T_{CYC1} and a *HIS3* marker was excised from pIM44-LovB using AvrII and SwaI, and transformed into BJNC creating BJNCB. The *LEU2* marker was removed from BJNC and BJNCB by expressing CreA recombinase as explained above. P_{ADH2}-AptB-T_{CYC1} (LEU2-ADH2-For and FF2288) and a *LEU2* marker (LEU2-LoxP-Rev and FF2287) were PCR amplified from pJC843-AptB and pJC811, respectively. Two linear DNA fragments were gel purified and transformed into BJNC and BJNCB creating BJNCA and BJNCBA.

4.3.4. Media and cultivation

Luria-Bertani (LB) medium was used for the cultivation of *E. coli* cells. Ampicilin (100 μg/ml) was used for the selection of plasmids in LB medium. *E. coli* cells were cultivated at 37°C and 250 rpm in an air shaker.

S. cerevisiae was cultivated in non-selective YPD complex medium (10 or 20 g/L dextrose, 20 g/L peptone, 10 g/L yeast extract (BD Biosciences, Sparks, MD)), or selective SDC medium (20 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate) supplemented with appropriate amino acids or nucleotides (SDC(A): 100 mg/L adenine hemisulfate, SDC(A,U): 100 mg/L adenine hemisulfate and 100 mg/L uracil, SDC(A,T): 100 mg/L adenine hemisulfate and 100 mg/L l-tryptophan). For the selection using the *LEU2* marker, SD(-LEU) (20 g/L dextrose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine hemisulfate and 100 mg/L l-tryptophan, 100 mg/L uracil, 100 mg/L l-histidine) was used. For the selection using the *HIS3* marker, SD(-HIS) (20 g/L dextrose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine hemisulfate and 100 mg/L l-tryptophan, 100 mg/L uracil, 150 mg/L l-leucine) was used. For plates, 20 g/L agar was added. Glucose was autoclaved separately and then added to the other components. Yeast nitrogen base, ammonium sulfate, adenine hemisulfate, l-histidine, and l-tryptophan were filter-sterilized and added after autoclaving the other components. For DML production, YPD with 10 g/L glucose was used.

For DML production, cells were initially cultivated in selective medium overnight and transferred into the YPD. *S. cerevisiae* cells were incubated at 30°C and 250 rpm in an air shaker.

Culture volume was 5 ml in test tubes or 200 ml in 1 L flasks. Optical density was measured at 600 nm using a UV-2450 UV/Vis spectrophotometer (Shimadzu, Kyoto, Japan).

4.3.5. Ni-NTA column purification and SDS-PAGE analysis

Harvested cells were pelleted and washed with cold water twice. Cells were pelleted again and resuspended in lysis buffer (50 mM sodium phosphate, 1 mM DTT, 1% protease inhibitor cocktail (Sigma, Cat No. P8215), 300 mM sodium chloride, 10 mM imidazole). Cells were broken using glass beads and cell debris was removed by centrifugation. His-tagged thioesterases were purified using His-Pur™ Ni-NTA spin column (Thermo Scientific, Cat No. 88227). His-tagged proteins were eluted in elution buffer (50 mM sodium phosphate, 1 mM DTT, 300 mM sodium chloride, 250 mM imidazole). Proteins were quantified using Bradford assay (Bio-Rad, Cat No. 500-0006). The same amount of total proteins or purified proteins were separated in a 7% SDS-PAGE gel for LovB and in a 12% SDS-PAGE gel for the thioesterases.

4.3.6. Plasmid stability test

Cells were collected after 72 hours of expression in 1% YPD and plated onto YPD plates (200 cells/plate). After colonies appeared, they were replica-plated onto selective plates and YPD plates (to confirm viability). The TE genes were on *TRP1*-marked plasmids and the *lovB* gene was on a *URA3*-marked plasmid. Thus, SDC(A,U) was used for TE and SDC(A,T) was used for *lovB* as

the selective plates. These new plates were incubated at 30°C for 2 days and the colonies on each plate were counted. The percentage of plasmid containing cells was calculated as the number of colonies on the selective plates divided by the total viable colonies transferred.

4.3.7. Quantitative Real-time PCR

Quantitative Real-time PCR was performed to measure transcription level of the thioesterases. *ACT1* was used as the reference gene. Thioesterases were expressed for 48 hours in 5 ml 1% YPD. 1 ml of cells was collected and total RNA was extracted using Yeast StarTM RNA Kit (Zymo Research Corporation). Subsequently, DNA was removed using Turbo DNA-*free*™ Kit (Life Technologies Corporation). From the RNA, cDNA was generated using High Capacity Reverse Transcription Kit (Life Technologies Corporation). For RT-PCR, primers were obtained from IDTDNA, Inc. RT-PCR was done in triplicate for all reactions including reference cDNA. Results were normalized by the reference cDNA.

4.3.8. DML detection

DML was extracted and detected by Dr. Wei Xu in Prof. Yi Tang's group as reported (Xu et al., 2013). Cells were harvested in 50 ml 1% YPD at 30°C with agitation. At specified times, 0.5 ml samples were collected from the flasks. Cells were pelleted and supernatant was collected. The sample was extracted twice using ethyl acetate (1% TFA). After the organic phase was dried, the

sample was redissolved in methanol. DML (m/z = 307) was measured using a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer (Shimadzu, Kyoto, Japan).

4.4 Results and Discussion

DML, the precursor to lovastatin, has been successfully synthesized *in vitro* when LovC, LovB and a selected thioesterase were included in the reaction mixture along with NADPH, MatB to supply malonyl-CoA, and SAM (S-adenosyl methionine) (Ma et al., 2009). The limiting step for DML synthesis was the release of DML from LovB catalyzed by the thioesterase. In this study, a set of thioesterases from a variety of microorganisms was chosen and compared for *in vivo* DML synthesis. Two expression systems were evaluated, one with both lovB and the thioesterases on 2μ-based plasmids, and one with the thioesterases on 2μ plasmids and *lovB* integrated into the genome.

4.4.1. Comparison of LovB and TE expression

One effective way of increasing enzyme expression is to increase DNA copy number. The 2μ replication origin enabled a *S. cerevisiae* cell to maintain up to nearly 60 copies per cell in BY4743 strain and 35 copies per cell in BY4741 strain (Karim et al., 2013), although values of approximately ten are also common. However, maintaining large numbers of genes can be deleterious due to the burden of high level expression (Parekh et al., 1995). DML synthesis requires the expression of *lovB*, *npgA* (a 4'-phosphopantetheinyl transferase), *lovC*, and a thioesterase. Maintaining more than one multi-copy plasmid in a cell is often not sustainable. To avoid carrying two or three multi-copy plasmids, we compared the enzyme expression level between CEN/ARS and 2μ-based *lovB* plasmids using SDS-PAGE. This was to determine the effect

of copy numbers on expression level of LovB, and help us determine whether *lovB* should be maintained in the genome or on a multi-copy plasmid.

LovB enzyme expression level was compared in strain BJ5464 carrying BJ-CLovB (CEN/ARS) and BJ-2LovB (2μ) (Figure 4.2). While the 2μ-based plasmid resulted in greater LovB levels relative to the CEN/ARS system, the difference was less than 2-fold. Therefore, we decided to integrate the *lovB* gene into the genome for the DML synthesis system.



Figure 4.2. SDS-PAGE showing LovB expression in BJ5464 cell extract using CEN/ARS and 2μ-based vectors BJ-CLovB and BJ-2LovB.

We also checked thioesterase expression level using SDS-PAGE. Strains with the five different thioesterases (strains BJ-P, BJ-H, BJ-R, BJ-A, BJ-V) were expressed in YPD (10 g/L dextrose). Thioesterases were his-tagged at the C-terminus and were purified using Ni-NTA columns. Purified thioesterases were compared on SDS-PAGE (Figure 4.3 A). hpm3 TE showed strong expression while Radicicol TE (Rdc1 TE) and AptB were comparable. PKS 13 TE showed comparably low expression level and the VrtG band was barely visible. These comparisons assume that all TEs were similarly purified via the his-tag, which may not be fully accurate. LovB expression level from the integrated copy was also compared among different strains (BJNCB with TE plasmids, Table 4.1) and shown to be comparable among all strains (figure 4.3 B).

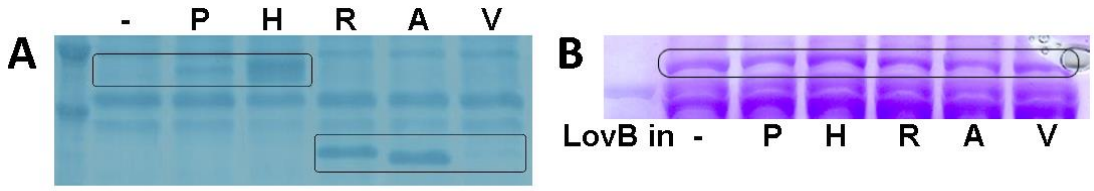


Figure 4.3. Comparison of (A) TE and (B) LovB expression on SDS-PAGE. (A) 5 His-tagged TEs are shown on SDS-PAGE after Ni-NTA purification. (B) LovB expressed from BJNCB strain carrying thioesterase plasmids. '–' negative control strain without TE; 'P' PKS13 TE domain; 'H' hpm3 TE domain; 'R' Rdc1 TE domain; 'A' AptB; 'V' VrtG

4.4.2. Comparison of DML level

For comparison of DML levels, we constructed two sets of strains (Table 4.1). One set contained an integrated copy of *lovB* and a 2μ-based thioesterase plasmid. For comparison, the second set contained both *lovB* and thioesterase on separate 2μ-based plasmid. Both sets contain one copy of NpgA and one copy of LovC integrated into the genome. Using these strains, we compared the five different thioesterases in terms of DML production. We then compared the DML synthesis from the strain with a genomic copy of *lovB* to the strain with a 2μ-based *lovB*.

Ten strains with the five different thioesterases were cultivated in YPD (1% glucose) for 48 h or 72 h at 30°C. DML was extracted and measured by Dr. Wei Xu in our collaborator's (Prof. Yi Tang) laboratory at UCLA. In the single plasmid system, AptB showed the highest DML titer with 4.5-fold and 9-fold higher than hpm3 TE and PKS13-TE, respectively after 72 hours of cultivation. However, Rdc1 TE and VrtG showed no detectable levels of DML. We then compared DML levels for the single and dual plasmids system for the AptB TE. Interestingly, similar to the LovB SDS-PAGE results in Figure 4.1, the 2μ based *lovB* strain with AptB had only 36% higher DML

titers than the integrated *lovB* strain after 48 hours of cultivation (Table 4.2). With the integrated *lovB*, DML levels increased from 48 to 72 h, while they decreased with the plasmid-based *lovB*. After 72 hours of cultivation, DML titers were lower for this dual plasmid system.

Table 4.2 DML synthesis comparison

	Single Plasmid System			Dual Plasmid System		
Strains	TE mRNA	DML (ug/L) (48h/72h)	TE Plasmid Stability (72h)	DML (ug/L) (48h/72h)	TE Plasmid Stability (72h)	*lovB* Plasmid Stability (72h)
PKS13 TE	217%	120/250	96%	NT	87%	74%
hpm3 TE	168%	380/480	87%	NT	94%	67%
Rdc1 TE	168%	0	90%	NT	NT	NT
AptB	153%	1900/2200	87%	3000/1300	87%	63%
VrtG	100%	0	95%	NT	87%	89%

NT: Not tested.

We also checked plasmid stability of both the TE and *lovB* plasmids in all strains (BJNCB-P, BJNCB-H, BJNCB-R, BJNCB-A, BJNCB-V, BJNC-PB, BJNC-HB, BJNC-RB, BJNC-AB, and BJNC-VB). All five TE plasmids had comparable stability (87 % to 96 %) with both the single plasmid system and the dual plasmid system. *lovB* stability was lower than TEs, although maintained above 60%. All 6 genes including *lovB* and the five thioesterases were expressed under the late-phase *ADH2* promoter, which is turned on after glucose is depleted. This is the reason that most plasmids were relatively stable. LovB is 335 kDa protein and its gene, *lovB* is a 9.1 kilo-base gene; this may have contributed to the generally lower *lovB* plasmid stability relative to the thioesterase plasmids.

To see if the expression level of the thioesterases affected DML synthesis, we looked at transcription level. With the single plasmid strains, mRNA levels were determined via RT-PCR after extracting RNA from each strain. The data for the individual TEs was normalized by the data for the actin that was obtained from the same RNA sample. The RT-PCR results showed relatively comparable levels with PKS13 TE showing the highest (217%) and VrtG the lowest (set at 100%).

Based on the above results, the large differences in DML synthesis among the 6 strains is likely not due to expression level differences. They may be due to the efficiency of each thioesterase at releasing DML from LovB. To prove this, each thioesterase could be compared for *in vitro* DML synthesis.

4.4.3. Identification of a dedicated thioesterases, LovG in *Aspergillus terreus*

In the course of these studies, our collaborator, Prof. Yi Tang's group (UCLA) identified a dedicated thioesterase, LovG for DML release from LovB (Xu et al., 2013). Previously, it was annotated as an oxidoreductase or hypothetical protein. However, the mutation study of this gene in *A. terreus* drastically reduced the DML level. *lovB*, *lovC*, *npgA*, and *lovG* were expressed together in BJ5464, and up to 35 mg/L DML was produced *in vivo*. LovG was more efficient than the heterologous TEs and at least 10-fold higher levels of DML were produced. Therefore, further studies with the heterologous TEs were suspended.

4.5 Conclusions

The main goal of this study was to improve the synthesis of DML, the lovastatin precursor, in *S. cerevisiae*. DML is synthesized using a PPT, LovB, LovC and a discrete TE enzyme. The current bottleneck for DML synthesis in *S. cerevisiae* appears to be the release of nonaketide chain from LovB. According to previous published reports, DML release can be carried out by PKS13 TE from *G. zeae in vitro* (Ma et al., 2009). In this study, we looked at PKS13 TE and hpm3 TE and additional thioesterases (Rdc1 TE, AptB and VrtG) for the release of DML from LovB *in vivo*. Among these five thioesterases, AptB showed the highest DML synthesis while Rdc TE and VrtG did not show any detectable level of DML. To find out what factor led to the different DML production levels, we considered plasmid stability of the TE and *lovB* plasmids, transcription levels of the TEs, and protein levels for the TEs. Based on our results, we hypothesize that the difference in DML production levels is due to the difference in DML releasing efficiency of each TE. For more direct evidence, an *in vitro* assay is necessary. After these studies were done, a native thioesterase responsible for the DML release in *A. terreus* was identified to be LovG by our collaborators (Xu et al., 2013). It will be interesting to include LovG for *in vitro* activity measurements with PKS13 TE, hpm3 TE, and AptB. For the synthesis of the final product lovastatin, another set of enzymes has to be introduced: LovF and LovD. In addition, lovastatin toxicity to *S. cerevisiae* also has to be circumvented for higher-level production. Although, this work will require effort, lovastatin synthesis in *S. cerevisiae* will contribute to the industrial production of lovastatin and other polyketides.

4.6 References

Bonekamp, F. J., Oosterom, J., 1994. On the safety of *Kluyveromyces lactis* - a review. Applied Microbiology and Biotechnology. 41**,** 1-3.

Ching, C., Sequential cloned gene integration: enhancements in *Saccharomyces cerevisiae*, extension to polyploid yeast strains, and application to polyketide production., Chemical and Biochemical Engineering, Vol. PhD. University of California, Irvine, Irvine, 2005.

Chooi, Y. H., Cacho, R., Tang, Y., 2010. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. Chemistry & Biology. 17**,** 483-494.

Crawford, J. M., Townsend, C. A., 2010. New insights into the formation of fungal aromatic polyketides. Nature Reviews Microbiology. 8**,** 879-889.

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2011. A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 28**,** 123-136.

Gietz, R. D., Schiestl, R. H., Willems, A. R., Woods, R. A., 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast. 11**,** 355-360.

Gietz, R. D., Woods, R. A., 2001. Genetic transformation of yeast. Biotechniques. 30**,** 816-831.

Hendrickson, L., Davis, C. R., Roach, C., Nguyen, D. K., Aldrich, T., McAda, P. C., Reeves, C. D., 1999. Lavastatin biosynthesis in *Aspergillus terreus*: characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. Chemistry & Biology. 6**,** 429-439.

Karim, A. S., Curran, K. A., Alper, H. S., 2013. Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. Fems Yeast Research. 13**,** 107-116.

Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C., Hutchinson, C. R., 1999. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science. 284**,** 1368-1372.

Lee, K. K. M., Engineering of *Saccharomyces cerevisiae* for the biosynthesis of fungal polyketides. University of California, Irvine, 2006.

Lee, K. K. M., Da Silva, N. A., Kealey, J. T., 2009. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. Analytical Biochemistry. 394**,** 75-80.

Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X. K., Kealey, J. T., Da Silva, N. A., Vederas, J. C., Tang, Y., 2009. Complete reconstitution of a highly reducing iterative polyketide synthase. Science. 326**,** 589-592.

Maggon, K., 2005. Best-selling human medicines 2002-2004. Drug Discovery Today. 10**,** 739-742.

Parekh, R., Forrester, K., Wittrup, D., 1995. Multicopy overexpression of bovine pancreatic trypsin-inhibitor saturates the protein-folding and secretory capacity of *Saccharomyces cerevisiae*. Protein Expression and Purification. 6**,** 537-545.

Pfeifer, B. A., Khosla, C., 2001. Biosynthesis of polyketides in heterologous hosts. Microbiology and Molecular Biology Reviews. 65**,** 106-+.

Reeves, C. D., Hu, Z., Reid, R., Kealey, J. T., 2008. Genes for the biosynthesis of the fungal polyketides hypothemycin from *Hypomyces subiculosus* and radicicol from *Pochonia chlamydosporia*. Applied and Environmental Microbiology. 74**,** 5121-5129.

Romanos, M. A., Scorer, C. A., Clare, J. J., 1992. Foreign gene expression in yeast - a review. Yeast. 8**,** 423-488.

Sambrook, J., Russell, D. W., 2001. Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Shen, M. W. Y., Fang, F., Sandmeyer, S., Da Silva, N. A., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in *Saccharomyces cerevisiae*. Yeast. 29**,** 495-503.

Szewczyk, E., Chiang, Y. M., Oakley, C. E., Davidson, A. D., Wang, C. C. C., Oakley, B. R., 2008. Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. Applied and Environmental Microbiology. 74**,** 7607-7612.

Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., Tang, Y., 2013. LovG: The thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. Angew Chem Int Ed Engl. 52**,** 6472-5.

Chapter 5.

Pathway Engineering for the Enhanced Synthesis of 6-MSA in *Saccharomyces cerevisiae*

A portion of this chapter has been slightly modified from the publication: Jin Wook Choi, Nancy A. Da Silva, 2014. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase. Journal of Biotechnology. 187, 56-59

5.1. Abstract

Polyketides have immense potential as candidates for pharmaceuticals and industrial chemicals. The high-level production of polyketides require a sufficient supply of precursors. Key precursors to polyketides are acetyl-CoA and malonyl-CoA. To increase the intracellular acetyl-CoA and malonyl-CoA levels, we overexpressed *ACS1*, *ACS* from *Salmonela enterica*, and its L641P mutant, *CAB1,* and *ACC1* in *Saccharomyces cerevisiae*. Overexpression of these precursor pathway genes led to modest improvements in the synthesis of the polyketide, 6-methylsalicylic acid (6-MSA). The Acc1 enzyme is also deactivated by AMP-activated serine/threonine protein kinase (Snf1) when glucose is depleted. We created a S1157A mutation in the *S. cerevisiae* native Acc1 to prevent deactivation following glucose depletion. This resulted in a 9-fold increase in *in vitro* activity and a 3-fold increase in *in vivo* synthesis of 6-MSA and native fatty acids. We also tested upstream pathway gene deletions to redirect the carbon flow toward malonyl-CoA synthesis; deletion of *PYC1* led to a 50 % increase in 6-MSA level in *S. cerevisiae*. OptKnock was also used to predict deletion target genes for optimized 6-MSA synthesis, and these gene knockouts were tested. We then evaluated overexpression of different combinations of ACS_{SE}^{L641P}, $ACC1^{S1157A}$ and *ADH2*, leading to slight improvement. We overexpressed $ACC1^{S1157A}$ in the strains with knockouts of *PYC1* or *PLB1*. This led to another step increase relative to before $ACC1^{S1157A}$ overexpression. Utilization of these various strategies will benefit the production of other polyketides as well.

5.2. Introduction

Polyketides are secondary metabolites polymerized from short-chain carboxylic acid units such as acetate, malonate, propionate, and butyrate (Chooi and Tang, 2012). Polyketides have been widely used as pharmaceutically active molecules; commercially successful examples include erythromycin, tetracycline, doxorubicin, and lovastatin (Crawford and Townsend, 2010). In addition, polyketides (e.g., triacetic acid lactone, TAL) are useful as precursors for the synthesis of industrial chemicals (Chia et al., 2012). Fatty acids are also synthesized from acetyl-CoA and malonyl-CoA, and are promising as chemical precursors (Leber and Da Silva, 2014; Nikolau, 2010) and for use as biofuels (Fortman et al., 2008). Two important production hosts for polyketides and fatty acids are *Escherichia coli* and *Saccharomyces cerevisiae*. The latter has shown promise for both simple and complex fungal polyketides such as 6-methylsalisylic acid (6-MSA) (Kealey et al., 1998) and dihydromonacolin L (DML), a precursor to lovastatin (Ma et al., 2009; Xu et al., 2013).

Fatty acids and polyketides use multiple units of the same building block (e.g., malonyl-CoA) repeatedly. 6-MSA is synthesized from one molecule of acetyl-CoA and three molecules of malonyl-CoA (Dimroth et al., 1970), DML requires one molecule of acetyl-CoA and eight molecules of malonyl-CoA (Kennedy et al., 1999), and TAL uses one acetyl-CoA molecule and two malonyl-CoA molecules (Eckermann et al., 1998). Native yeast strains also require high levels of malonyl-CoA for the synthesis of fatty acids, and the synthesis of heterologous polyketides in *S. cerevisiae* competes with native fatty acid synthesis for the available malonyl-CoA. Therefore, high-level production of these polyketides can be quickly limited by malonyl-CoA availability. The

cytosolic acetyl-CoA is synthesized from acetate and coenzyme A by acetyl-CoA synthetase 1 (Acs1) and the cytosolic malonyl-CoA is synthesized from acetyl-CoA by acetyl-CoA carboxylase (Acc1). Therefore, to avoid malonyl-CoA limitation, the yeast metabolic pathways have been engineered to increase flux to acetyl-CoA (Chen et al., 2014; Kozak et al., 2014; Shiba et al., 2007), and the *ACS1* and *ACC1* gene has been overexpressed (Chen et al., 2014; Shiba et al., 2007; Wattanachaisaereekul et al., 2008).

Acc1 is negatively regulated by AMP-activated protein kinase (AMPK) in mammalian cells and by Snf1 in *S. cerevisiae* (Davies et al., 1990; Munday et al., 1988; Scott et al., 2002; Woods et al., 1994). In rat liver cells, AMPK is activated by an increased AMP to ATP ratio and phosphorylates acetyl-CoA carboxylase at S79, S1200, and S1215 (Davies et al., 1990; Ha et al., 1994), with S79 considered the critical residue. In *S. cerevisiae*, activation of Snf1 is triggered by glucose depletion leading to partial deactivation of cytosolic acetyl-CoA carboxylase (Acc1) via phosphorylation at one or more serine residues (Woods et al., 1994). Consequently, the deactivation of Acc1 will lower the cytosolic malonyl-CoA supply.

In this study, our objectives were to increase intracellular availability of malonyl-CoA. To achieve that, we overexpressed various acetyl-CoA synthetase genes and pantothenate kinase as well as acetyl-CoA carboxylase. We also took an alternate approach to increasing cytosolic malonyl-CoA levels by preventing the deactivation of the *S. cerevisiae* Acc1 enzyme. A critical serine responsible for deactivation via phosphorylation was identified and mutated to an alanine. *In vitro* Acc1 assays confirmed activity during the ethanol phase following glucose depletion, and expression of the modified Acc1 resulted in higher polyketide product levels *in vivo*. Finally, we

applied a computational approach to predict target genes to delete for improved 6-MSA synthesis, and tested these strains for improved 6-MSA production.

5.3. Materials and Methods

5.3.1. Molecular biology techniques

Plasmids were isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific). All PCR products were purified using either the Zymoclean™ gel DNA Recover Kit (Zymo Research Corporation) or the QIAquick PCR Purification Kit (Qiagen). The Rapid DNA Ligation Kit (Thermo Scientific) was used for the ligation of DNA fragments. Restriction endonucleases, Taq DNA Polymerase, and NEB 10-beta competent cells were purchased from New England Biolabs Inc. KOD Hot Start DNA Polymerase (EMD Millipore) was used for high fidelity PCR while all other PCR reactions were carried out using Taq DNA Polymerase. For site-directed mutagenesis, PfuUltraII Fusion HS polymerase (Agilent Technologies, Santa Clara, CA) was used. All oligo nucleotides were synthesized by Integrated DNA Technologies, Inc. Synthesized genes were sequenced by Eton Bioscience Inc. *E. coli* competent cells were created using the Calcium chloride method (Sambrook and Russell, 2001). All primer sequences are given in Table C.1 (Appendix C).

Plasmid transformation into *S. cerevisiae* was performed following the modified lithium acetate method (Gietz and Woods, 2001). Integration of linear DNA into *S. cerevisiae* cells was performed following the high-efficiency transformation method (Gietz et al., 1995).

5.3.2. Vector construction

Escherichia coli strain XL1-Blue (Stratagene) or DH5α (Invitrogen) was used as the primary strain for maintenance of plasmids and general cloning procedures. The pXP series of vectors were used for yeast expression (Table 3.1) (Fang et al., 2011; Shen et al., 2012). The vectors constructed are listed in Table 5.1. Primer sequences are listed in Table C.1 (Appendix C).

npgA was digested from pKMLADH2p-npgA (Lee et al., 2009) using SalI. This fragment was ligated into the XhoI site of pJC118 and pJC742 to create pJC118-NpgA and pJC742-NpgA. pJC812-6MSAS and pJC842-6MSAS were constructed by digesting 6-MSAS from YEp6MSAS (Lee et al., 2009) using SpeI and RsrII and cloned into the same sites of pJC812 and pJC842 to create pJC812-6MSAS and pJC842-6MSAS, respectively.

ACS1 (ScAcs1_for_AvrII_Kozak, ScAcs1_rev_PmeI) and *CAB1* (Cab1_for_AvrII_Kozak, Cab1_rev_PmeI_v2) were amplified from BY4741 genomic DNA. ACS_{SE} and ACS_{SE}^{L641P} (SEAcs_for_SpeI_Kozak, SEAcs_rev_XhoI) were amplified from pESC-ALD6-SEacs and pESC-ALD6-$SEacs^{L641P}$ (Shiba et al., 2007). *ACS1* and *CAB1* were digested with AvrII and PmeI, then cloned into pJC811 between SpeI and PmeI. ACS_{SE} and ACS_{SE}^{L641P} were cloned into SpeI and XhoI of pJC811.

An integration vector pIU43 was constructed as following. Front homology (Ura3_FH_For2, Ura3_FH_Rev) and rear homology (Ura3_RH_For, Ura3_RH_Rev2) sequences were PCR amplified. Front homology was digested with NdeI and inserted into the same site of pJC743. Then, the rear homology fragment was digested and inserted between KpnI and EcoRI of this vector creating pIU43. *PGK1* promoter was digested with NdeI and SpeI from pJC812 and inserted into the same sites of pIU43 replacing P_{ADH2} and created pIU13.

The *HXT7* promoter (PHXT7F2, PHXTR2) was amplified from BJ5464 flanked with NdeI and SpeI upstream and downstream, respectively. The PCR product was gel-purified, digested with NdeI and SpeI, cloned into the same sites of pJC752. *ADH2* (ADH2F, ADH2R) was amplified from BJ5464 flanked with SpeI and XhoI. PCR product was gel-purified, digested with SpeI and XhoI, and cloned into the same sites of above *HXT7* promoter vector creating a temporary vector, pHXT7-ADH2. P_{HXT7}-*ADH2* was digested with NdeI and BamHI and cloned into the same sites of pIU13. However, pIU13 contains two NdeI sites flanking the *URA3* locus front homology sequence. Thus, NdeI site was partially digested and the correct size fragment was gel-purified followed by the cloning of P_{HXT7}-*ADH2* between NdeI and BamHI sites creating pIUTHA.

ACC1 (pBF3054) was received from Professor Suzanne Sandmeyer at the University of California, Irvine. pBF3054 carries the *S. cerevisiae ACC1* between SpeI (upstream) and XhoI (downstream) sites. A 1kb fragment from the start codon to the internal SpeI (+1063) was PCR amplified with RsrII at the upstream of the start codon down to the internal SpeI site. This PCR fragment was digested with RsrII and SpeI. pJC118 was used as the holding vector for *ACC1* and digested with RsrII and XhoI. pBF3054 was digested with SpeI, XhoI, and NdeI. Then, the three DNA fragments: the PCR product (RsrII to SpeI), the rear part of *ACC1* (SpeI to XhoI), and pJC118 (XhoI to RsrII), were ligated to create pJC118-ACC1. pIM11-ACC1 was created by digesting with RsrII and XhoI from pJC118-ACC1 and ligating *ACC1* into pIM11. The S1157A mutant of *ACC1* was created by site-directed mutagenesis from pIM11-ACC1 with primers (ACC1_S1157A_SDM_F and ACC1_S1157A_SDM_R. pIM11-ACC1 was digested with KasI to NaeI (1.2kb) which flank S1157 and replaced with the counterpart from $ACC1^{S1157A}$ creating pIM11-ACC1m. The amplified 1.2kb section was sequenced to ensure the correct final sequence.

A 6X Histidine tag was added at the C-terminus of *ACC1* and *ACC1*^{S1157A}. The internal BsrGI to XhoI segment of *ACC1*^{S1157A} was PCR amplified (ACC14351For and ACC1HisRev). This fragment was digested using BsrGI and XhoI, and ligated with the RsrII to BsrGI fragment of *ACC1* or *ACC1*^{S1157A} and the XhoI to RsrII fragment of pJC811 to create pJC811-ACC1H and pJC811-ACC1mH, respectively.

A dual gene plasmid, pKUTP-6MN, containing P_{TEF1}-6MSAS-T_{ADH2} and P_{PGK1}-NpgA-T_{CYC1} was constructed with 4 fragments by Gibson assembly (Figure 5.1). The 4 fragments were the *TEF1* promoter, 6-MSAS and *ADH2* terminator, *PGK1* promoter and *npgA*, and linearized pKA (XhoI to NotI) in order. Fragment 1 was obtained by PCR of the *TEF1* promoter from pXP418 (TEF1F and TEF1R). Fragment 2 was obtained by restriction digestion of YEp6MSAS at SpeI and KpnI. The 3rd fragment was obtained by PCR of the *PGK1* promoter and *npgA* from pJC118-NpgA (PGK1F and NpgAR). The 4th fragment was obtained by restriction digestion of pKA using NotI and XhoI. Gibson reaction was done as reported (Appendix B) (Gibson, 2009) and the mixture was transformed into NEB 10-beta competent cells. The correct sequence was confirmed by Eton Bioscience Inc.

Table 5.1. List of plasmids constructed for the engineering of pyruvate dehydrogenase bypass.

Plasmids	Characteristics	Reference
pESC-ALD6-SEacs	P_{GAL1}-*ALD6*, P_{GAL10}-ACS_{SE}	Shiba et al., 2007
pESC-ALD6-SEacs^{L641P}	P_{GAL1}-*ALD6*, P_{GAL10}-ACS_{SE}^{L641P}	Shiba et al., 2007
YEp6MSAS	2μ, P_{ADH2}-T_{ADH2}, *URA3*	Lee et al., 2006
pXP118	CEN/ARS, P_{PGK1}-T_{CYC1}, *URA3*	Fang et al., 2011
pXP418	2μ, P_{TEF1}-T_{CYC1}, *URA3*	Fang et al., 2011
pXP742	CEN/ARS, P_{ADH2}-T_{CYC1}, *URA3*	Shen et al., 2012
pXP811	2μ, P_{PGK1}-T_{CYC1}, *LEU2*	Shen et al., 2012
pXP812	2μ, P_{PGK1}-T_{CYC1}, *URA3*	Shen et al., 2012
pXP842	2μ, P_{ADH2}-T_{CYC1}, *URA3*	Shen et al., 2012
pXP843	2μ, P_{ADH2}-T_{CYC1}, *TRP1*	Unpublished
pBF3060	2μ, P_{GAL1}-*CreA*-T_{CYC1}, *URA3*	Fang et al., 2011
YEpADH2p	2μ, P_{ADH2}-T_{ADH2}, *URA3*	Chapter 4
pJC702	CEN/ARS, P_{ADH2}-T_{ADH2}, *URA3*	Chapter 4
pJC118	pXP118, SpeI-PmeI-RsrII-XhoI	Chapter 3
pJC742	pXP742, SpeI-PmeI-RsrII-XhoI	Chapter 3
pJC743	CEN/ARS, P_{ADH2}-T_{ADH2}, *TRP1*	Chapter 3
pJC811	pXP811, SpeI-PmeI-RsrII-XhoI	Chapter 3
pJC812	pXP812, SpeI-PmeI-RsrII-XhoI	Chapter 3
pJC842	pXP842, SpeI-PmeI-RsrII-XhoI	Chapter 3
pJC118-NpgA	pJC118, *npgA*	This study
pJC702-NpgA	pJC702, *npgA*	Chapter 4
pJC742-NpgA	pJC742, *npgA*	This study
pJC811-ACS1	pJC811, *ACS1*	This study
pJC811-ACS_{SE}	pJC811, ACS_{SE}	This study
pJC811-ACS_{SE}^{L641P}	pJC811, ACS_{SE}^{L641P} (ACS_{SE}*m*)	This study
pJC811-CAB1	pJC811, *CAB1*	This study
pJC811-ACC1H	pJC811, *ACC1-His*	This study
pJC812-ACC1mH	pJC811, $ACC1^{S1157A}$-*His*	This study
pJC812-6MSAS	pJC812, 6-MSAS	This study
pJC842-6MSAS	pJC842, 6-MSAS	This study
pIM11	Integrating vector targeting *MET17* locus, P_{PGK1}-T_{CYC1}, *LEU2*	Chapter 4
pIU43	Integrating vector targeting *URA3* locus, P_{ADH2}-T_{CYC1}, *TRP1*	This study
pIU13	Integrating vector targeting *URA3* locus, P_{PGK1}-T_{CYC1}, *TRP1*	This study
pIM11-ACC1	pIM11, *ACC1*	This study
pIM11-ACC1m	pIM11, $ACC1^{S1157A}$	This study
pIUTHA	Integrating vector targeting *URA3* locus, P_{HXT72}-T_{CYC1}, *TRP1*, P_{HXT7}-*ADH2*	This study
pKA	2μ, P_{ADH2}-T_{CYC1}, P_{KEX2}-*Ubi-R-URA3* *	This study
pKA-6MSAS	pJC842-6MSAS, P_{KEX2}-*Ubi-R-URA3* *	This study
pKP-6MSAS	pJC812-6MSAS, P_{KEX2}-*Ubi-R-URA3* *	This study
pKUTP-6MN	2μ, P_{TEF1}-6MSAS-T_{ADH2}, P_{PGK1}-*npgA*-T_{CYC1}, P_{KEX2}-*Ubi-R-URA3* *	This study

* Construction of P_{KEX2}-*Ubi-R-URA3* is explained in Chapter 6 (section 6.3.2).

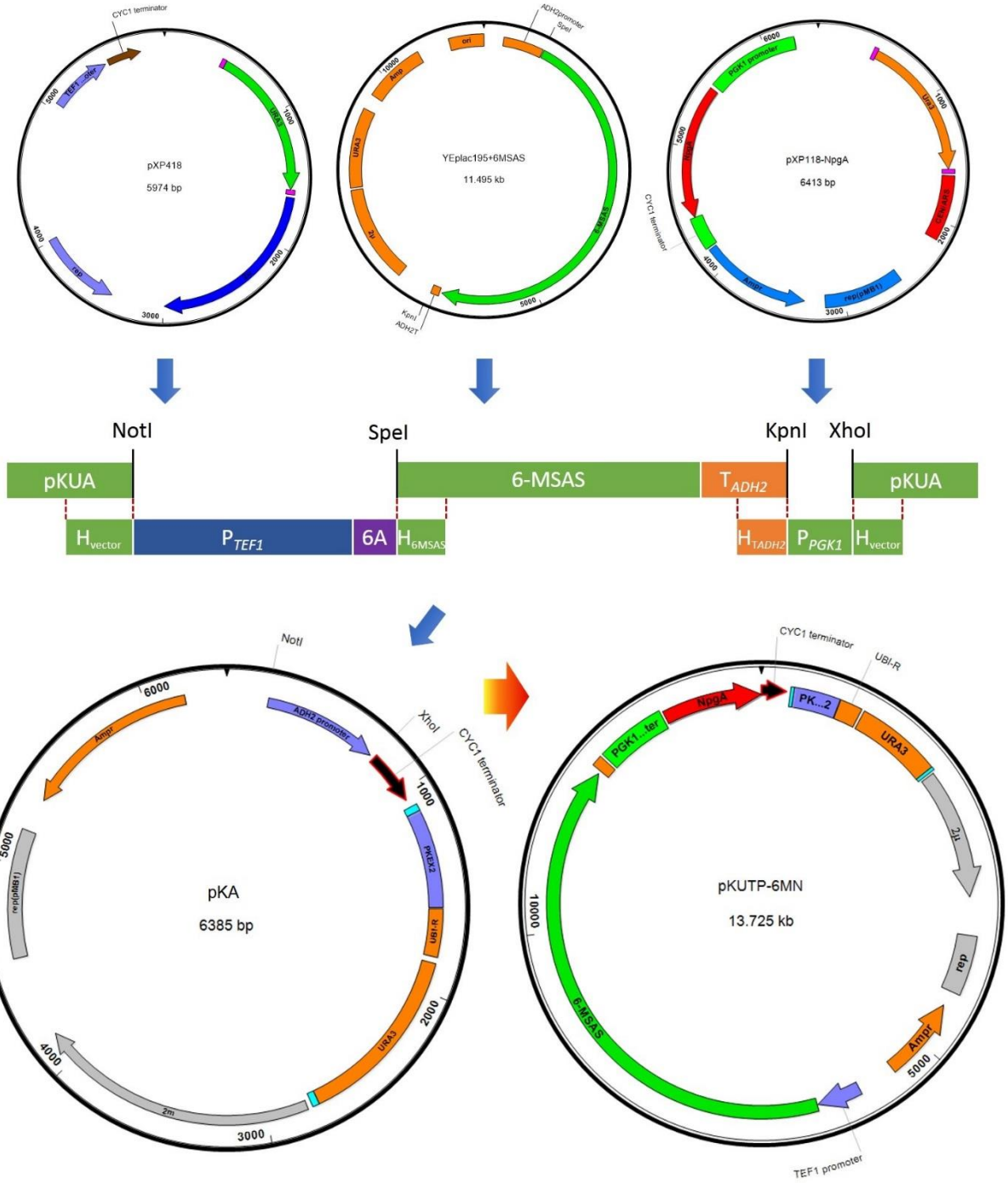


Figure 5.1. Diagram for 6-MSAS and *npgA* dual gene expression plasmid (pKUTP-6MN) construction via Gibson assembly

5.3.3. Strain construction

BJ5464 (*MATα his3-Δ200 leu2-Δ1 trp1 ura3-52 pep4::HIS3 prb1Δ1.6R can1 GAL*) and BY4741 (*MATa his3Δ1 met150 leu2Δ0 ura3Δ0*) were used as the base *S. cerevisiae* strains. We followed previously reported gene integration methods (Fang et al., 2011) using double crossover homologous recombination for the construction of the strains used in this study (Table 5.2).

Strain BYPN1 was constructed by integrating one copy of P_{PGK1}-*npgA*-T_{CYC1} into the YDRWTy1-5 locus of BY4741. P_{PGK1}-*npgA*-T_{CYC1} (YDRWTy1-5-PGKF, FF2288) and LoxP-*HIS3*-LoxP was PCR amplified from pXP220 (FF2325 and FF2287). Both PCR fragments were gel purified and transformed into BY4741 creating BYPN1.

To remove the *HIS3* marker gene from BYPN1, pBF3060 was transformed into this strain. *HIS3* was removed from BYPN1 by expressing CreA recombinase (pBF3060) as reported by Fang et al. (Fang et al., 2011). P_{PGK1}-*ACS1*-T_{CYC1} was PCR amplified from pJC811-ACS1 (HIS3-PGK1-For, FF2288) and LoxP-HIS3-LoxP was PCR amplified from pXP220 (HIS3-LoxP-Rev and FF2287). Both PCR fragments were gel purified and transformed into BY4741 creating BYPN1S. pIM11-ACC1 was digested using NruI and NotI. DNA fragments containing the *MET17* locus front homology, P_{PGK1}-*ACC1*-T_{CYC1}, *LEU2*, and *MET17* locus rear homology were separately transformed into BYPN1, creating BYPN1C. pKP-6MSAS was transformed into BYPN1, BYPN1S, BYPN1C, BYPN1SC creating BYPN1-KP6M, BYPN1S-KP6M, BYPN1C-KP6M, BYPN1SC-KP6M.

S. cerevisiae strain BJN2 was constructed by integrating an additional copy of *npgA* under the *ADH2* promoter into the genome of BJN (Chapter 4). *LEU2* was removed from BJN by expressing CreA recombinase as explained above. P_{ADH2}-*npgA*-T_{CYC1} was PCR amplified from

pJC742-NpgA (YDRWTy1-5-ADH2PF and FF2288) and LoxP-*TRP1*-LoxP was PCR amplified from pXP843 (FF2325 and FF2287). Both PCR fragments were gel purified and transformed into BJN creating BJN2.

Strain BJPN1 was constructed by integrating one copy of P_{PGK1}-*npgA*-T_{CYC1} into the *Δleu2* locus of BJ5464. P_{PGK1}-*npgA*-T_{CYC1} (LEU2-PGK1-For and FF2288) was PCR amplified from pJC118-NpgA. LoxP-*TRP1*-LoxP (LEU2-LoxP-Rev and FF2287) was PCR amplified from pXP843. Both PCR fragments were gel purified and transformed together into BJ5464 creating BJPN1b. An empty pIM11 was digested using NruI and NotI. The linear DNA fragment with *MET17* locus front homology, P_{PGK1}-T_{CYC1}, *LEU2*, and *MET17* locus rear homology was transformed into this strain creating BJPN1dm.

BJPN1S was created by integrating one copy of P_{PGK1}-*ACS1*-T_{CYC1} into the *Δhis3* locus of BJPN1. P_{PGK1}-*ACS1*-T_{CYC1} and *LEU2* marker (HIS3-PGK1-For and HIS3-LoxP-Rev) were PCR amplified from pJC811-ACS1. This linear DNA was gel purified and transformed into BJPN1 creating BJPN1S. BJPN1S_{SE}, BJPN1S_{SE}m, and BJPN1B were created in the same way as BJPN1S using pJC811-ACS_{SE}, pJC811-ACS_{SE}^{L641P}, and pJC711-CAB1 with the same set of primers, respectively. The same single linear fragment containing P_{PGK1}-*CAB1*-T_{CYC1} and *LEU2* marker amplified and gel purified from pJC711-CAB1 was integrated into BJPN1S_{SE}m after removing the *LEU2* and *TRP1* markers, creating BJPN1S_{SE}mB. pXP812-6MSAS was transformed into BJPN1S, BJPN1S_{SE}, BJPN1S_{SE}m, BJPN1B, and BJPN1S_{SE}mB creating BJPN1S-P6M, BJPN1S_{SE}-P6M, BJPN1S_{SE}m-P6M, BJPN1B-P6M, and BJPN1S_{SE}mB-P6M.

BJPN1C and BJPN1Cm were created as following. pIM11-ACC1 and pIM11-ACC1m were digested using NruI and NotI. DNA fragments containing the *MET17* locus front homology, P_{PGK1}-

ACC1-T_{CYC1} or P_{PGK1}-*ACC1*^{S1157A}-T_{CYC1}, *LEU2*, and *MET17* locus rear homology were separately transformed into BJPN1b, creating BJPN1C and BJPN1Cm, respectively.

pKA-6MSAS was transformed into BJN2 creating BJN2-KA6M. pJC811-ACC1H and pJC811-ACC1mH were transformed into BJ5464 creating BJ-C and BJ-Cm. pJC812-6MSAS was transformed into BJPN1dm, BJPN1C, and BJPN1Cm creating BJPN1dm-P6M, BJPN1C-P6M, and BJPN1Cm-P6M, respectively. pJC812 was transformed into BJPN1dm to create a negative control strain, BJPN1dm-P.

For evaluation of OptKnock guided predictions (Mo et al., 2009; Schellenberger et al., 2011), we transformed pKUTP-6MN (Figure 5.1) into selected strains from the Yeast Knockout Collection (YSC1053) (Table 5.2).

To remove existing auxotrophic marker genes, pBF3060 was transformed into BJPN1 (*TRP1*), BJPN1S_{SE}m (*LEU2*), and BJPN1Cm (*LEU2*). *TRP1* and *LEU2* were removed from these strains by expressing CreA recombinase (pBF3060) as reported by Fang et al. (Fang et al., 2011). P_{PGK1}-*ACC1m* was integrated into the MET17 locus of the marker-removed BJPN1S_{SE}m creating BJPN1S_{SE}mCm in the same method as for pIM11-ACC1. pIUTHA was digested with NruI and NotI and transformed into marker-removed BJPN1, BJPN1S_{SE}m, BJPN1Cm, and BJPN1S_{SE}mCm. The P_{HXT7}-*ADH2* cassette was integrated into BJPN1, BJPN1S_{SE}m, and BJPN1Cm creating BJPN1A, BJPN1S_{SE}mA, BJPN1CmA, and BJPN1S_{SE}mCmA. pKP-6MSAS was transformed into these strains creating BJPN1A-KP6M, BJPN1S_{SE}mA-KP6M, BJPN1CmA-KP6M, and BJPN1S_{SE}mCmA-KP6M.

Table 5.2. List of strains for the engineering of pyruvate dehydrogenase bypass

Strains	Characteristics	Reference
BJ5464	*MAT*α *his3-Δ200 leu2-Δ1 trp1 ura3-52 pep4::HIS3 prb1Δ1.6R can1 GAL*	Jones, 1991
BY4741	*MAT*a *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Open Biosystems
BYPN1	BY4741, YDRWTy1-5::P_{PGK1}-*npgA*-T_{CYC1}-*HIS3*	This study
BYPN1C	BYPN1, *Δmet17*:: P_{PGK1}-*ACC1*-T_{CYC1}-*LEU2*	This study
BYPN1S	BYPN1 (*HIS3* removed), *Δhis3*:: P_{PGK1}-*ACC1*-T_{CYC1}-*HIS3*	This study
BYPN1SC	BYPN1S, *Δmet17*:: P_{PGK1}-*ACC1*-T_{CYC1}-*LEU2*	This study
BYPN1-KP6M	BYPN1, pKP-6MSAS	This study
BYPN1C-KP6M	BYPN1C, pKP-6MSAS	This study
BYPN1S-KP6M	BYPN1S, pKP-6MSAS	This study
BYPN1SC-KP6M	BYPN1SC, pKP-6MSAS	This study
BJN	BJ5464, YDRWTy1-5::P_{ADH2}-*npgA*-T_{ADH2}	This study
BJN2	BJN , *ura3*::P_{ADH2}-*npgA*-T_{ADH2}-*TRP1*	This study
BJN2-KA6M	BJN2, pKA-6MSAS	This study
BJ-C	BJ5464, pJC811-ACC1H	This study
BJ-Cm	BJ5464, pJC811-ACC1mH	This study
BJPN1	BJ5464, *leu2*::P_{PGK1}-*npgA*-*TRP1*	This study
BJPN1dm	BJPN1, *Δmet17*::*LEU2*	This study
BJPN1S	BJPN1, (*pep4*::*Δhis3*):: P_{PGK1}-*ACS1*-T_{CYC1}-*LEU2*	This study
BJPN1S$_{SE}$	BJPN1, (*pep4*::*Δhis3*):: P_{PGK1}-ACS_{SE}-T_{CYC1}-*LEU2*	This study
BJPN1S$_{SE}$m	BJPN1, (*pep4*::*Δhis3*):: P_{PGK1}-ACS_{SE}^{L641P}-T_{CYC1}-*LEU2*	This study
BJPN1B	BJPN1, (*pep4*::*Δhis3*):: P_{PGK1}-*CAB1*-T_{CYC1}-*LEU2*	This study
BJPN1S$_{SE}$mB	BJ5464, *Δleu2*::P_{PGK1}-*npgA*, (*Δpep4*::*Δhis3*):: P_{PGK1}-ACS_{SE}^{L641P}-T_{CYC1}, *Δtrp1*:: P_{PGK1}-*CAB1*-T_{CYC1}-*LEU2*	This study
BJPN1C	BJPN1, *Δmet17*::P_{PGK1}-*ACC1-LEU2*	This study
BJPN1Cm	BJPN1, *Δmet17*::P_{PGK1}-$ACC1^{S1157A}$-*LEU2*	This study
BJPN1dm-P	BJPN1dm, pJC812	This study
BJPN1dm-P6M	BJPN1dm, pJC812-6MSAS	This study
BJPN1C-P6M	BJPN1C, pJC812-6MSAS	This study
BJPN1Cm-P6M	BJPN1Cm, pJC812-6MSAS	This study
BJPN1S-P6M	BJPN1S, pJC812-6MSAS	This study
BJPN1S$_{SE}$-P6M	BJPN1S_{SE}, pJC812-6MSAS	This study
BJPN1S$_{SE}$m-P6M	BJPN1S_{SE}m, pJC812-6MSAS	This study
BJPN1B-P6M	BJPN1B, pJC812-6MSAS	This study
BJPN1S$_{SE}$mB-P6M	BJPN1BS_{SE}m, pJC812-6MSAS	This study
BJPN1A	BJ5464, *leu2*::P_{PGK1}-*npgA, ura3*::P_{HXT7}-*ADH2*	This study
BJPN1S$_{SE}$mA	BJPN1, (*pep4*::*Δhis3*):: P_{PGK1}-ACS_{SE}^{L641P}-T_{CYC1}, *ura3*::P_{HXT7}-*ADH2*	This study
BJPN1CmA	BJ5464, *leu2*::P_{PGK1}-*npgA*, *Δmet17*::P_{PGK1}-$ACC1^{S1157A}$, *ura3*::P_{HXT7}-*ADH2*	This study
BJPN1S$_{SE}$mCmA	BJ5464, *leu2*::P_{PGK1}-*npgA*, (*pep4*::*Δhis3*):: P_{PGK1}-ACS_{SE}^{L641P}-T_{CYC1}, *Δmet17*::P_{PGK1}-*ACC1-LEU2, ura3*::P_{HXT7}-*ADH2*	This study
BJPN1A-KP6M	BJPN1A, pKP-6MSAS	This study
BJPN1S$_{SE}$mA-KP6M	BJPN1S_{SE}mA, pKP-6MSAS	This study
BJPN1CmA-KP6M	BJPN1CmA, pKP-6MSAS	This study
BJPN1S$_{SE}$mCmA-KP6M	BJPN1S_{SE}mCmA, pKP-6MSAS	This study
BY4741Δzwf1	BY4741, *zwf1*::*KanMX*	Open Biosystems
BY4741Δfbp1	BY4741, *fbp1*::*KanMX*	Open Biosystems

Table 5.2 (Continued)

Strains	Characteristics	Reference
BY4741Δpyc1	BY4741, *pyc1*::*KanMX*	Open Biosystems
BY4741Δpyc2	BY4741, *pyc2*::*KanMX*	Open Biosystems
BY4741Δgpd1	BY4741, *gpd1*::*KanMX*	Open Biosystems
BY4741Δplb1	BY4741, *plb1*::*KanMX*	Open Biosystems
BY4741Δplb2	BY4741, *plb2*::*KanMX*	Open Biosystems
BY4741Δrhr2	BY4741, *rhr2*::*KanMX*	Open Biosystems
BY4741Δhor2	BY4741, *hor2*::*KanMX*	Open Biosystems
BY4741Δtpo1	BY4741, *tpo1*::*KanMX*	Open Biosystems
BY4741Δinm1	BY4741, *inm1*::*KanMX*	Open Biosystems
BYΔzwf1-6MN	BY4741Δzwf1, pKUTP-6MN	This study
BYΔfbp1-6MN	BY4741Δfbp1, pKUTP-6MN	This study
BYΔpyc1-6MN	BY4741Δpyc1, pKUTP-6MN	This study
BYΔpyc2-6MN	BY4741Δpyc2, pKUTP-6MN	This study
BYΔgpd1-6MN	BY4741Δgpd1, pKUTP-6MN	This study
BYΔplb1-6MN	BY4741Δplb1, pKUTP-6MN	This study
BYΔplb2-6MN	BY4741Δplb2, pKUTP-6MN	This study
BYΔrhr2-6MN	BY4741Δrhr2, pKUTP-6MN	This study
BYΔhor2-6MN	BY4741Δhor2, pKUTP-6MN	This study
BYΔtpo1-6MN	BY4741Δtpo1, pKUTP-6MN	This study
BYΔinm1-6MN	BY4741Δinm1, pKUTP-6MN	This study
BYΔpyc1Cm	BYΔpyc1, *Δmet17*::P_{PGK1}-*ACC1*^{S1157A}-*LEU2*	This study
BYΔplb1Cm	BYΔplb1, *Δmet17*::P_{PGK1}-*ACC1*^{S1157A}-*LEU2*	This study
BYΔpyc1Cm-6MN	BYΔpyc1Cm, pKUTP-6MN	This study
BYΔplb1Cm-6MN	BYΔplb1Cm, pKUTP-6MN	This study
BYΔzwf1-6MN	BYΔpyc1Cm, pKUTP-6MN	This study

5.3.4. Media and cultivation

Luria-Bertani (LB) medium was used for the cultivation of *E. coli* cells (Sambrook and Russell, 2001). Ampicilin (100 μg/ml) was used for the selection of plasmids in LB medium. *E. coli* cells were cultivated at 37°C in a 250 rpm in an air shaker.

S. cerevisiae was cultivated in non-selective YPD complex medium (20 g/L dextrose, 20 g/L peptone, 10 g/L yeast extract (BD Biosciences, Sparks, MD)), modified selective SDC(A)

medium (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 400 mg/L L-serine, 200 mg/L L-threonine, and 20mM MES pH 5.5), or modified selective SD(-LEU) medium (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100ug/L Biotin, 100 mg/L adenine sulfate, 100 mg/l uracil, 100 mg/l L-tryptophan, 100 mg/l L-histidine, 400 mg/l L-serine, 200 mg/l L-threonine, 1.4 g/L Yeast Synthetic Drop-out Medium Supplements-without leucine, histidine, tryptophan, uracil (Y2001, Sigma-Aldrich, St. Louis, MO), and 20mM MES at pH 5.5) according to the marker requirement (Hanscho et al., 2012; Sherman et al., 1986). For plates with leucine or histidine deficiency, SD(-LEU) plates (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/l uracil, 100 mg/l L-tryptophan, 100 mg/l L-histidine, 100 mg/L L-methionine, 20 g/L agar) or SD(-HIS) plates (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/l uracil, 100 mg/l L-tryptophan, 150 mg/l L-leucine, 100 mg/L L-methionine, 20 g/L agar) were used. The detailed recipes of the various SDC media used in this chapter are given in Table 5.3. *S. cerevisiae* cells were incubated at 30°C and 250 rpm in an air shaker. Culture volume was 5 ml in test tubes or 200ml in 1L flasks. Optical density was measured at 600 nm using a UV-2450 UV/Vis spectrophotometer (Shimadzu, Kyoto, Japan).

Table 5.3. List of various SDC media used in this chapter.

Medium	Recipe (1L medium)
SDC(A,T) medium (2g1y1a1c)	20 g dextrose, 1.7 g yeast nitrogen base, 5 g ammonium sulfate, 5 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
0.5g1y1a1c	5 g dextrose, 1.7 g yeast nitrogen base, 5 g ammonium sulfate, 5 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1g1y1a1c	10 g dextrose, 1.7 g yeast nitrogen base, 5 g ammonium sulfate, 5 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1g2y2a1c	10 g dextrose, 3.4 g yeast nitrogen base, 10 g ammonium sulfate, 5 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1g1y1a2c	10 g dextrose, 1.7 g yeast nitrogen base, 5 g ammonium sulfate, 10 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
0.5g2y2a2c	5 g dextrose, 3.4 g yeast nitrogen base, 10 g ammonium sulfate, 10 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1g2y2a2c	10 g dextrose, 3.4 g yeast nitrogen base, 10 g ammonium sulfate, 10 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
2g2y2a2c	20 g dextrose, 3.4 g yeast nitrogen base, 10 g ammonium sulfate, 10 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1g3y3a3c	10 g dextrose, 5.1 g yeast nitrogen base, 15 g ammonium sulfate, 15 g casamino acids, 100 mg adenine hemisulfate, 100 mg L-tryptophan
Modified 2 % SDC medium (2 % mSDC)	20g dextrose, 1.7 g yeast nitrogen base, 5 g ammonium sulfate, 0.2 g threonine, 0.4 g serine, 20 mM MES pH 5.5.
Modified 1 % SDC medium (1 % mSDC)	10g dextrose, 1.7 g yeast nitrogen base, 5 g ammonium sulfate, 0.2 g threonine, 0.4 g serine, 20 mM MES pH 5.5.
Enhanced 1 % SDC medium (1 % eSDC)	Modified 1 % SDC medium but with 5.1g yeast nitrogen base and 15g ammonium sulfate.
2 % mSDC(A,T)	2 % mSDC, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1 % mSDC(A,T)	1 % mSDC, 100 mg adenine hemisulfate, 100 mg L-tryptophan
1 % eSDC(A,T)	1 % eSDC, 100 mg adenine hemisulfate, 100 mg L-tryptophan

5.3.5. Plasmid stability test

Stability of plasmids in non-selective YPD medium was tested as following. 200 cells were plated onto YPD plates. The plates were incubated for 2 days and replica-plated onto selective plates and incubated at 30°C for 2 days. The ratio of colonies that grew on selective plates relative to the number of total colonies on YPD plates was calculated.

5.3.6. Expression and Purification of 6-MSAS, Acc1, and Acc1^{S1157A}

BJN2 and BJN2-KA6M were cultivated in modified SDC(A) medium overnight and used to inoculate 200 mL 1 % YPD medium in 1L flasks. 6-MSAS was expressed for 36 hours. Cells were harvested and washed using cold water, and cell pellets were stored at -80°C until cell lysis. Cells were lysed in lysis buffer (200 mM potassium phosphate, pH 7.6, 5 mM β–mercaptoethanol, 1 mM EDTA, 15 % Glycerol, 1 % protease inhibitor cocktail (Cat# P8215, Sigma), 50 mM sodium fluoride, 5 mM sodium pyrophosphate) using a French Press Cell Disrupter (Thermo Fisher) at 16,000 psi. Cell extracts were obtained by centrifugation of lysate. 6-MSAS was purified following published method with modifications (Spencer and Jordan, 1992; Vogel and Lynen, 1975). Decreasing fractions of proteins were collected with ammonium sulfate precipitation (22 % - 45 % fraction collected), PEG6000 precipitation (6 % - 18 % fraction collected), and hydroxyapatite chromatography (Bio-Gel HTP, cat# 139-0420, Bio-Rad). Collected proteins contained native Acc1 according to an Acc1 activity assay. Therefore, biotinylated proteins,

which include Acc1, were removed by incubating proteins with avidin-agarose resin (cat# A9207-1ML, Sigma).

BJ5464, BJ-C, and BJ-Cm were cultivated overnight in 5 ml media. Modified SD(-LEU) medium was used for BJ-C and BJ-Cm, and L-leucine was added to 150 mg/L into SD(-LEU) for BJ5464. The same media were used for protein expression but in 200 mL volume. Glucose level was monitored to harvest cells at the mid-exponential growth phase. Ethanol phase samples were harvested during the late ethanol growth phase. Glucose and ethanol concentrations in the harvested samples were measured using HPLC. Harvested cells were washed in cold water and stored at -80°C until cell lysis.

Purification of Acc1 and Acc1^{S1157A} was performed according to previously reported methods after modification (Matsuhashi, 1969; Shirra et al., 2001). Cells were broken using the French press in lysis buffer (100mM potassium phosphate, pH 7.2, 5mM β-mercaptoethanol, 1 % protease inhibitor cocktail (Cat# P8215, Sigma), 1mM EDTA, 50mM sodium fluoride, 5mM sodium pyrophosphate, 0.25M sucrose). Cell lysates were centrifuged and biotinylated proteins were purified from supernatants using avidin affinity chromatography (Pierce Monomeric Avidin Agarose, cat# 20228, Thermo Scientific). Protein concentration was measured using Bradford assay and adjusted after comparing bands on SDS-PAGE. SDS-PAGE band density was analyzed by AlphaEaseFC (Alpha Innotech Corporation, San Leandro, CA).

5.3.7. *In vitro* activity assay

6-MSAS activity was assayed according to a previous report (Vogel and Lynen, 1975). The reaction was carried out at room temperature with 50 mM potassium phosphate at pH 7.6, acetyl-CoA (10 μM), NADPH (40 μM), BSA (0.25 mg/ml), and 6-MSAS. The reaction was started by adding malonyl-CoA (5 μM) to the reaction mixture. 6-MSA increase was measured by detecting the native fluorescence of 6-MSA using 310 nm excitation and 390 nm emission wavelengths in an Aminco-Bowman Series 2 luminescence spectrometer (SLM Aminco). The unit activity of 6-MSAS is defined as the amount of 6-MSAS that produces 1 μmol of 6-MSA per minute.

The activity of Acc1 was assayed in combination with 6-MSAS and other previously reported reaction components (Matsuhashi, 1969). The reaction was carried out at room temperature with ATP (2.5 mM), potassium bicarbonate (12.5 mM), L-cysteine (5 mM), magnesium chloride (10 mM), BSA (0.25 mg/ml), EDTA dipotassium magnesium salt (10 mM), potassium phosphate pH 7.6 (50 mM), 6-MSAS, and Acc1. The reaction was initiated by adding acetyl-CoA into the reaction mixture, and the increase in 6-MSA was measured by detecting native fluorescence of 6-MSA on a SpectraMAX M2 microplate reader (Molecular Devices, LLC) (310 nm excitation and 390 nm emission wavelengths). The unit activity of Acc1 is defined as the amount of Acc1 that catalyzes 1 μmol of acetyl-CoA per minute.

5.3.8. Glucose, Ethanol, and 6-MSA measurements using HPLC

An HPLC (LC-10ATvp pumps, Shimadzu, Kyoto, Japan) was used for the detection and measurement of glucose, ethanol and 6-MSA. Glucose and ethanol were separated in an Aminex HPX-87H column (Bio-Rad, Hercules, CA) and detected using a RID-10A refractive index detector. 5mM sulfuric acid was used as mobile phase and run at 0.5 ml/min, 65°C. 6-MSA was separated at 0.2 ml/min and room temperature using a Zorbax SB-C18 column (Agilent Technologies, Santa Clara, CA) and detected at 306 nm using a SPD-10Avp UV detector. Acetonitrile with 1 % acetic acid and water with 1 % acetic acid were used in gradient mode. The method was started with 80 % aqueous phase decreasing to 40 % by 20 minutes. The aqueous phase was further decreased to 0 % by 24 minutes and kept constant at 0 % until 35 minutes. Then the aqueous phase was increased to 80 % by 42 minutes. The column was equilibrated at 80 % aqueous phase for 18 minutes finishing the run by 60 minutes.

5.4. Results and Discussion

5.4.1. Engineering of pyruvate dehydrogenase bypass for enhanced synthesis of 6-MSA in *Saccharomyces cerevisiae*.

Acs1 and Acc1 are responsible for the synthesis of acetyl-CoA and malonyl-CoA, respectively (Figure 2.6). *ACS1* expression is repressed and its protein is deactivated during the glucose growth phase (Shiba et al., 2007). Acc1 is deactivated after glucose is depleted (Woods et al., 1994). Overexpression of *ACS1* and *ACC1* genes can be beneficial for the production of 6-MSA, which is synthesized from one acetyl-CoA and three malonyl-CoA molecules.

The *ACS1* and *ACC1* genes were overexpressed in BYPN1 by integrating a single copy of each under the strong *PGK1* promoter into the genome individually or together (Figure 5.2). 6-MSA production was compared among four different strains after cultivation in 2 % mSDC(A,T) for 72 hours. *ACC1* overexpression (BYPN1C-KP6M) led to a 35 % improvement in 6-MSA titer compared to WT (BYPN-KP6M). However, *ACS1* overexpression (BYPN1S-KP6M) led to a significant decrease in 6-MSA titer with standard deviation as large as the average titer. The simultaneous overexpression of *ACC1* and *ACS1* in BYPN1 led to only a 22 % increase in 6-MSA titer, which is lower than the *ACC1*-only overexpression strain. Clearly, *ACS1* overexpression was detrimental to 6-MSA synthesis in BYPN1.

We evaluated the same set of strains with two other media, 1 % mSDC(A,T) and 1% eSDC(A,T) to see the effect of media components on 6-MSA level, especially for the *ACS1* overexpression strain (Figure 5.2). 1 % mSDC(A,T) and 1 % eSDC(A,T) contain 1 % glucose while

2 % mSDC(A,T) contains 2 % glucose (Table 5.3). The 6-MSA level from the *ACS1* overexpressing strain (BYPN1S-KP6M) in 1 % mSDC(A,T) and 1 % eSDC(A,T) dropped in the new media and were comparable with about 7 mg/L and 6 mg/L, respectively. However, 6-MSA levels for the control and *ACC1*(only)-overexpressing strain showed noticeable improvement in the new media, particularly 1% eSDC(A,T) (2.4-fold for WT, 2.6-fold for *ACC1* overexpressed strain in 1% eSDC(A,T) compared to 2% mSDC(A,T)).

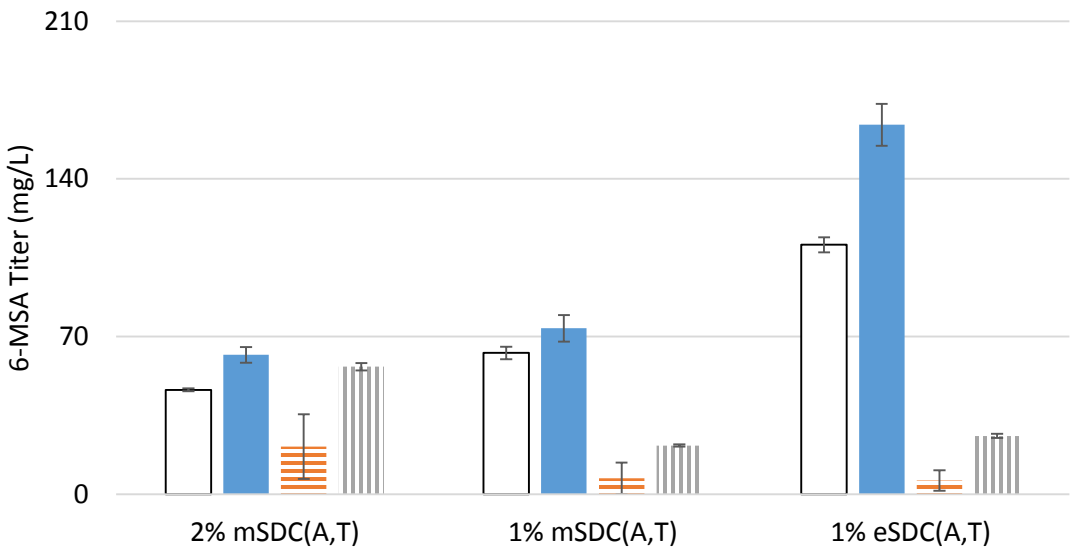


Figure 5.2. 6-MSA synthesis in BYPN1 strains with *ACS1* or/and *ACC1* overexpression. Overexpression strains had a single copy of *ACS1* (BYPN1S-KP6M, horizontally striped bar, orange) or *ACC1* (BYPN1C-KP6M, Closed bar, blue) or both genes (BYPN1SC-KP6M, vertically stripped bar, gray) integrated in the genome of BYPN1. 6-MSAS was expressed under the *PGK1* promoter on a 2μ-based plasmid (pKP-6MSAS) and *ACS1*/*ACC1* overexpression strains were compared to control strain (BYPN1-KP6M, empty bar).

We also constructed strains with *npgA* and *ACS1* integrated into the genome under the *ADH2* promoter, and with the 6-MSAS gene also under the *ADH2* promoter on a 2μ-based plasmid. Interestingly, no 6-MSA was detected. We suspected that the *ADH2* promoter might be significantly attenuated. To confirm this, we transformed the 2μ-based plasmid with 6-MSAS under the *ADH2* promoter into our original strain used for Figure 5.2 (BYPN1S, *npgA* and *ACS1* integrated under the *PGK1* promoter). Again, 6-MSA was not detectable. This result supports the hypothesis of an attenuated *ADH2* promoter. If the *ADH2* promoter is attenuated, it will slow down ethanol consumption since the native Adh2 is the only alcohol dehydrogenase that converts ethanol to acetaldehyde. To confirm that ethanol consumption slows down in *ACS1* overexpressed strains, we measured glucose and ethanol levels in both strains. For both strains, the ethanol concentration remained constant in the medium even after glucose was depleted (data not shown). This result shows that *ADH2* promoter is not active in these *ACS1* overexpressed strains. This experiment also explains why BYPN1S-P6M, the *ACS1* overexpression strain in Figure 5.2 (with 6-MSAS under the *PGK1* promoter) showed significantly lowered 6-MSA relative to the WT strain. The small amount of 6-MSA made in this strain must have been made during the glucose phase.

To see if this observation is strain specific, we also tested overexpression of *ACS1* and *ACC1* in our alternate expression host BJPN1 strain, which has *PEP4* and *PRB1* protease gene knockouts (Table 5.2). 6-MSAS was again expressed under the *PGK1* promoter on a 2μ-based plasmid. We also tested overexpression of the ACS_{SE} and the mutant ACS_{SE}^{L641P} from *Salmonella enterica*, which was reported to improve the synthesis of mevalonate (Shiba et al., 2007).

Overexpression of *ACS1*, ACS_{SE}, and ACS_{SE}^{L641P} resulted in a comparable levels of 6-MSA as the control (WT) (Figure 5.3).

We also overexpressed *CAB1*, a gene for pantothenate kinase, which is responsible for one of early steps in Coenzyme A biosynthesis. In *Escherichia coli*, pantothenate kinase (PanK) was found to be the rate-limiting step in the CoA synthesis pathway (Robishaw et al., 1982; Rock et al., 2000; Song and Jackowski, 1992). *CAB1* overexpression in *S. cerevisiae* did not increase 6-MSA level, which was comparable to the control (Figure 5.3). The simultaneous overexpression of ACS_{SE}^{L641P} and *CAB1* was only slightly better than WT. The specific 6-MSA level decreased compared to WT in all samples.

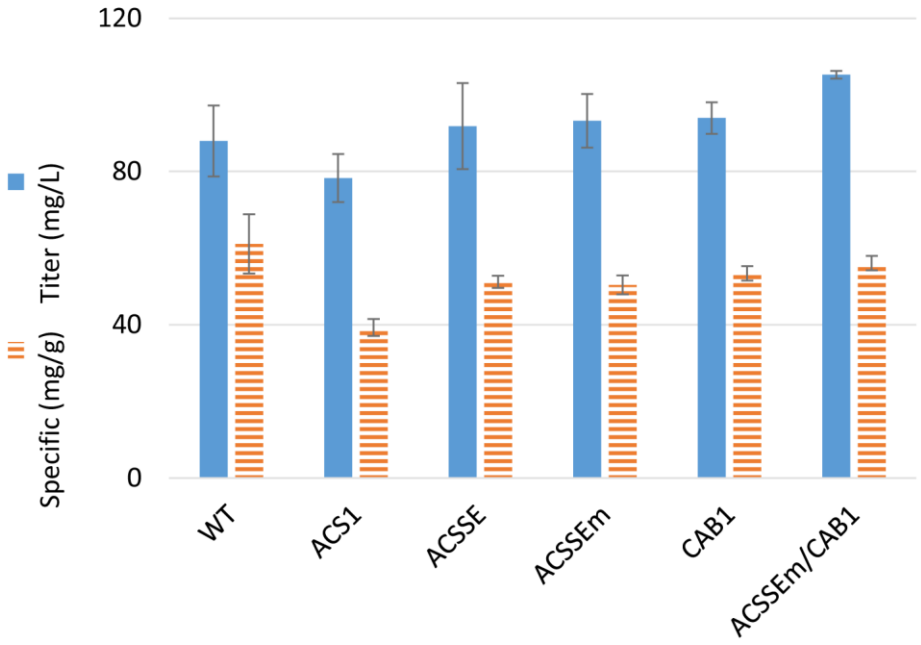


Figure 5.3 Effect of *ACS* variants and *CAB1* overexpressions in BJ5464-based strain on 6-MSA titer

We overexpressed *ACC1* in BJPN1 (the protease knockout strain) using a single copy integration under the *PGK1* promoter in the genome. After 72 hours of cultivation, the strain overexpressing *ACC1* strain in standard SDC(A,T) medium showed about 45 % improvement in terms of titer (Figure 5.4). However, this strain also showed significant flocculation in this medium. In another experiment, we were able to disperse the aggregated cells after resuspension in 2mM EDTA resulting in increased optical density (data not shown). This shows that the optical density measurements obtained for samples in Figure 5.4 were underestimated due to the flocculation resulting in overestimated specific 6-MSA levels.

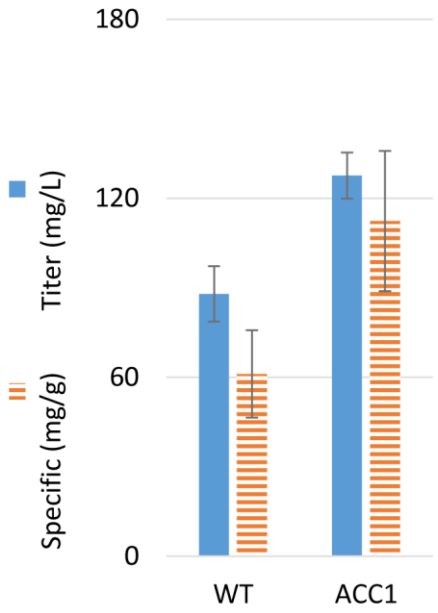


Figure 5.4. Improved 6-MSA synthesis by *ACC1* overexpression in BJPN1

To avoid flocculation during 6-MSA synthesis, we added EDTA to the SDC(A,T) medium. Flocculation was no longer observed in this strain but biomass accumulation suffered significantly leading to drastically lowered 6-MSA titer (data not shown). The flocculation was observed only in BJ5464-based strains not in BY4741-based strains. Slight flocculation was even observed in the BJ5464-based control (WT) (BJPN1-P6MSAS).

It is not clear why BJ5464-based strains are prone to flocculation that becomes severe after *ACC1* overexpression. However, flocculation was observed only with 6-MSA producing strains. We used YPD medium for the untransformed strain (no 6-MSAS plasmid) while we used SDC(A,T) medium for the 6-MSAS containing strains. Thus, we enriched the selective medium by increasing casamino acids, ammonium sulfate, and yeast nitrogen base to see if we can avoid flocculation by producing 6-MSA in rich medium. The modification of these media components (0.5g2y2a2c, 1g2y2a2c, 1g3y2a3c) prevented flocculation and increased 6-MSA production levels (Figure 5.5).

When casamino acids, ammonium sulfate, yeast nitrogen base, or a combination of these were increased 2-fold with 20 g/L glucose (2g1y1a1c, 2g2y2a1c, 2g1y1a2c, 2g2y2a2c), 6-MSA titer showed slight improvement. When glucose level was lowered again from 20 g/L to 5 g/L or 10 g/L, the titer improved significantly (2g2y2a2c, 0.5g2y2a2c, 1g2y2a2c, 1g3y2a3c) with the largest improvement achieved with 1g3y2a3c after 72 hours of cultivation. We measured 6-MSA, glucose, and ethanol levels at 24 h and 60 h (or 72h) for the *ACC1*-overexpression strain (BJPN1C-P6M) and calculated yield. In 2g1y1a1c, 6-MSA yield (per mol glucose) was much higher in the glucose growth phase than in the ethanol growth phase (Table 5.4). However, yield during the ethanol phase was higher in 0.5% glucose medium, and 6-MSA yield in 1g3y2a3c was almost 4-

fold higher in the ethanol growth phase than in the glucose growth phase. Glucose in the 0.5g or 1g media is depleted by 15 to 18 h; therefore, actual yield in the ethanol growth phase will be higher than our calculation since 20 g/L glucose will take more time to be fully consumed than 5 g/L or 10 g/L. As the concentration of glucose relative to other nutrients is decreased, more nutrients are available for 6-MSA synthesis during the ethanol growth phase and 6-MSA level increases. This media study shows more 6-MSA is synthesized during the ethanol growth phase.

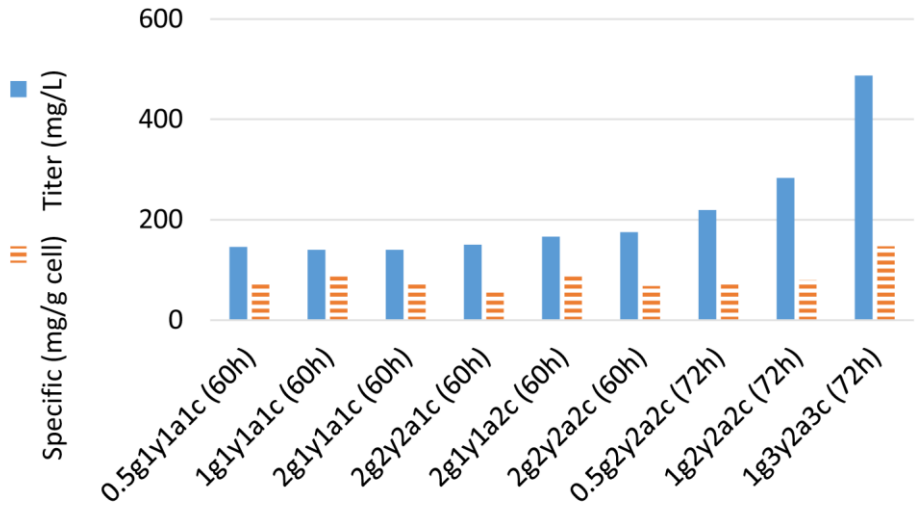


Figure 5.5. Media optimization to prevent flocculation in *ACC1* overexpressed strain (BJPN1C-P6M).

Table 5.4. Comparison of 6-MSA yield (mmol 6-MSA / mol substrate) during each growth phase in different media conditions. $Y_{Glu,6MSA}$ is based on period of 0 h to 24 h and $Y_{EtOH,6MSA}$ is based on period of 24 h to 60 h (72 h for 1g3y2a3c).

	$Y_{Glu,6MSA}$	$Y_{EtOH,6MSA}$
0.5G1y1a1c	12.8	13.4
1g1y1a1c	10.4	6.6
2g1y1a1c	7.3	1.9
1g3y2a3c	8.3	36.0

Results from both BY4741-based and BJ5464-based strains clearly show that *ACC1* overexpression benefits 6-MSA synthesis in *S. cerevisiae*. However, the results vary with regard to *ACS1* overexpressions. BY4741-based strains didn't derepress the *ADH2* promoter resulting in almost no 6-MSA synthesis during the ethanol growth phase, while BJ5464-based strains showed 6-MSA levels close to the control (WT) after overexpression of *ACS1*, *ACSSE*, and ACS_{SE}^{L641P}. However, in both strains, significantly more 6-MSA was produced during the ethanol growth phase unless other nutrients were limiting. Thus, the correct strategy to produce 6-MSA, and potentially other polyketides as well, is to produce 6-MSA in enriched medium after glucose depletion.

5.4.2. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase

5.4.2.1. *Identification of Snf1 target residue on Acc1*

In the previous sections, we confirmed that *ACC1* overexpression looked promising. Thus, we decided to also look at an alternate promising strategy to increase Acc1 levels: preventing

deactivation of this enzymbe. Acc1 is negatively regulated by AMP-activated protein kinase (AMPK) in mammalian cells and by Snf1 in *S. cerevisiae* (Davies et al., 1990; Munday et al., 1988; Scott et al., 2002; Woods et al., 1994). In rat liver cells, AMPK is activated by an increased AMP to ATP ratio and phosphorylates acetyl-CoA carboxylase at S79, S1200, andS1215 (Davies et al., 1990; Ha et al., 1994), with S79 considered the critical residue. In *S. cerevisiae*, activation of Snf1 is triggered by glucose depletion leading to partial deactivation of cytosolic acetyl-CoA carboxylase (Acc1) via phosphorylation at one or more serine residues (Woods et al., 1994). Consequently, the deactivation of Acc1 will lower the cytosolic malonyl-CoA supply. In this study, we took an alternate approach to increasing cytosolic malonyl-CoA levels by preventing the deactivation of the *S. cerevisiae* Acc1 enzyme. A critical serine responsible for deactivation via phosphorylation was identified and mutated to an alanine. *In vitro* Acc1 assays confirmed activity during the ethanol phase following glucose depletion, and expression of the modified Acc1 resulted in higher polyketide product levels *in vivo*.

To identify the potential phosphorylation target residues, we aligned the amino acid sequence of the *S. cerevisiae* Acc1 to the rat liver Acc1 sequence (Figure 5.6). The alignment result showed only one matching serine residue (S1157) in *S. cerevisiae* corresponding to S1215 of the rat Acc1. S1157 fits the AMPK phosphorylation target motif, which is a hydrophobic residue (M,L,F,I or V) for P-5 and P + 4, and basic residues (R,K or H) for P-3 or P-4 (Dale et al., 1995; Scott et al., 2002; Weekes et al., 1993). Therefore, we created a Ser1157 to Ala mutant using site-directed mutagenesis of *ACC1*.

```
Rat  GMTHVASVSDV-LLDNAFT   1226
     ||....||||: .:.|:.:
S.c. GMNRAVSVSDLSYVANSQS   1169
```

Figure 5.6. Amino acid sequence alignment between rat and *S. cerevisiae* Acc1. Amino acid sequences of the Acc1 enzyme from rat liver (top) and *S. cerevisiae* (bottom) were aligned using EMBOSS Stretcher. S1215 in the rat Acc1 corresponds to S1157 in *S. cerevisiae*.

5.4.2.2. *In vitro activity assay for Acc1^{S1157A}*

To evaluate the activity of the Acc1^{S1157A} enzyme, malonyl-CoA consumption was indirectly measured by a novel *in vitro* method coupling the reaction with a sensitive 6-methylsalicylic acid synthase (6-MSAS) fluorescence assay (Spencer and Jordan, 1992; Vogel and Lynen, 1975). Both wild type *ACC1* and the S1157A mutant were overexpressed under the glycolytic *PGK1* promoter on 2μ-based vectors (Table 5.1) in the protease knock-out yeast strain BJ5464 (MATα *his3-Δ200 leu2-Δ1 trp1 ura3-52 pep4*::*HIS3 prb1Δ1.6R can1 GAL*) (Jones, 1991). Cells were harvested during mid-exponential growth and during the ethanol growth phase. Glucose and ethanol levels at the harvest time points were comparable for all three strains (Table 5.5).

Table 5.5. Glucose and ethanol levels at harvest time for Acc1 activity assay.

Strains	Glucose (g/L)[a]	Ethanol (g/L)[b]
Native	5.98	0.76
ACC1 WT	7.83	0.81
S1157A	7.76	0.99

[a] Glucose concentrations for glucose phase samples (Initial glucose level: 10 g/L)
[b] Ethanol concentrations for ethanol phase samples

The Acc1 enzymes were purified from the cell extracts using avidin-biotin chromatography (Shirra et al., 2001). To obtain the 6-MSAS for the assay, this synthase was expressed under the *S. cerevisiae ADH2* promoter on 2μ-based plasmid in strain BJN2 (BJ5464 with two integrated copies of the *Aspergillus nidulans* gene coding for NpgA, a 4'-phosphopantetheinyl transferase for activation of 6-MSAS), and cells were harvested at 36 h. To purify the 6-MSAS from cell extracts, we adapted the method described for *Penicillium patulum* (Spencer and Jordan, 1992; Vogel and Lynen, 1975). Purified 6-MSAS showed a specific activity of 3.025U/mg. A negative control strain (BJN2), which is identical except that it lacks the 6-MSAS plasmid, did not show 6-MSAS activity. To determine the activity of the wild type Acc1 and $Acc1^{S1157A}$, the purified enzyme and 6-MSAS were added in the same reaction with required substrates except for malonyl-CoA, and the reaction was initiated by the addition of acetyl-CoA. The synthesis of 6-MSA over time was measured using 310 nm excitation and 390 nm emission wavelengths. 1 mole of 6-MSA synthesis is equivalent to 3 moles of malonyl-CoA synthesis by Acc1. To calculate the specific activity of the Acc1 variants, Acc1 concentration was determined using a combined Bradford assay and band density analysis on SDS-PAGE (Fig. 5.6 A) using AlphaEase FC (ProteinSimple, Santa Clara, CA).

The *in vitro* activity results (Figure 5.7 B) clearly show the impact of the S1157A mutation in Acc1. Specific activity of the wild type Acc1 drops over 4-fold in the ethanol growth phase relative to the glucose growth phase. In contrast, there is no decrease in the specific activity of $Acc1^{S1157A}$ following glucose depletion. Interestingly, S1157A showed a 2-fold higher specific activity compared to the wild type during the glucose growth phase. In the ethanol phase, a nearly 9-fold higher specific activity was observed for $Acc1^{S1157A}$ relative to wild type.

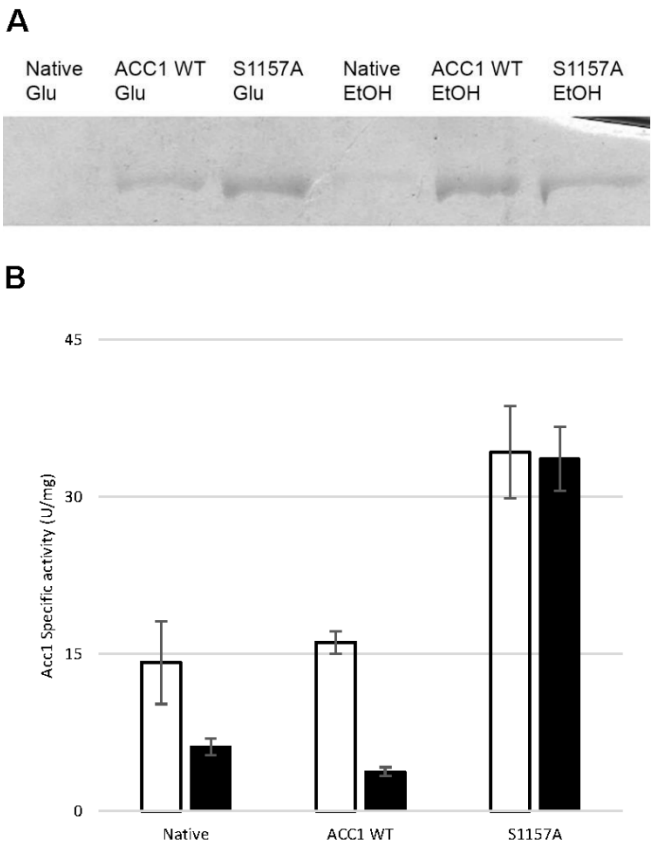


Figure 5.7. Comparison of *in vitro* activity of Acc1. (A) Acc1 bands on SDS-PAGE. Glu: Glucose phase; EtOH: Ethanol phase. Native samples show weak bands as the strain contains only the single native copy of *ACC1*. (B) *In vitro* Acc1 specific activity for the three strains during the glucose phase (open bars) and the ethanol phase (closed bars). Protein concentration was determined based on a Bradford assay and SDS-PAGE (A) band density analysis. Native: from native control strain; ACC1 WT: native + plasmid-based wild type *ACC1* gene; S1157A: native + plasmid-based $ACC1^{S1157A}$ gene. Error bars represent mean ± standard deviation (n = 2 or 3).

5.4.2.3. *In vivo production of 6-MSA and fatty acids*

The Acc1 activity assay confirmed that the S1157A mutation prevents deactivation of the enzyme during the shift in carbon source from glucose to ethanol. To determine the effect on the production of malonyl-CoA derived products, we compared the wild type and mutant Acc1 on

the synthesis of 6-MSA. A single copy of *ACC1* (BJPN1C) or *ACC1*^{S1157A} (BJPN1Cm) was integrated into the BJ5464 genome under the control of the *PGK1* promoter (Table 5.2). A copy of *npgA* was also integrated and 6-MSAS was expressed on a 2μ plasmid (both under the *PGK1* promoter). The strains were cultivated for 48 h late into the ethanol phase, and samples were assayed for 6-MSA via HPLC. The titers of 6-MSA were compared for the negative control strain (BJPN1dm-P, *npgA* integrated, empty plasmid, Table 5.2), the native strain (BJPN1dm-P6M, *npgA* integrated, 6-MSAS plasmid), the wild type Acc1 strain (BJPN1C-P6M, *npgA* integrated, wild type *ACC1* integrated, 6-MSAS plasmid), and the S1157A strain (BJPN1Cm-P6M, *npgA* integrated, *ACC1*^{S1157A} integrated, 6-MSAS plasmid). All strains still contain the native *ACC1* gene under the *ACC1* promoter. After 48 h of batch culture, the S1157A strain had produced substantially more 6-MSA with titers 2.8-fold and 3.7-fold higher than the wild type Acc1 strain and the native strain, respectively (Figure 5.8 A). On a per cell basis, the S1157A strain had a 2.4-fold and2.8-fold increase relative to the wild type Acc1 and native strain. Final values were 343 ± 64.0 mg/L, 168 ± 3.00 mg/g dry cell weight, and a yield of 34.3 ± 6.40 mg/g glucose for the S1157A strain. Fatty acids are also built from malonyl-CoA. To evaluate the effect of Acc1^{S1157A} on fatty acid biosynthesis in yeast, the same samples were assayed for intracellular fatty acid levels. Similar to the results for 6-MSA, the S1157A strain had a 3-fold increase in total fatty acid titer (Figure 5.8 B) with C18 and C18:1 showing 5.9-fold and 7.3-fold increases, respectively. Although low, both C24 and C26 levels also increased substantially in the strain with the modified Acc1 (Figure 5.8 B inset). This result clearly shows the significant impact of improved Acc1 activity on both fatty acid and 6-MSA synthesis. The Acc1^{S1157A} mutant will benefit the synthesis of other polyketides as well in *S. cerevisiae*.

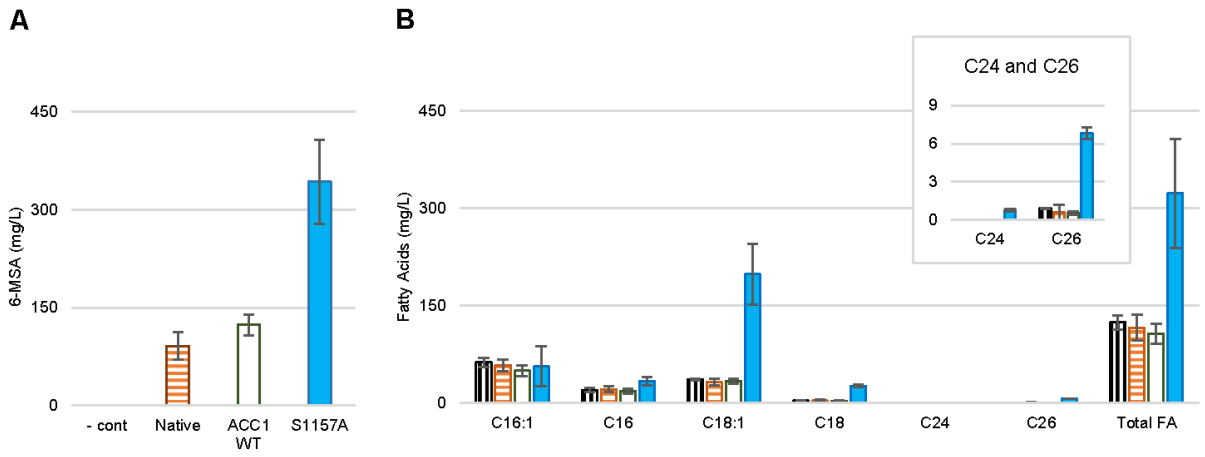


Figure 5.8. Production of 6-MSA and fatty acids *in vivo*. The strains were cultivated in 1 % mSDC(A) medium for 48 h. The samples were centrifuged and HPLC analysis of the supernatant was used to quantify 6-MSA (A) and GC-MS was used to quantify intracellular fatty acids (B) in the cell pellets. C24 and C26 levels are displayed separately in the inset of panel B. –cont: negative control (vertically-striped bars); Native: native *ACC1* (horizontally-striped bars); ACC1 WT: native + integrated wild type *ACC1* gene (open bars); S1157A: native + integrated *ACC1*^{S1157A} gene (closed bars). All strains have one integrated copy of *npgA* and carry a plasmid for expression of 6-MSAS except the negative control (-cont) which carries an empty vector. Error bars represent mean ± standard deviation (n = 3).

5.4.3. Metabolic pathway modifications to increase 6-MSA synthesis

In the previous sections, we focused on the engineering of the pyruvate dehydrogenase bypass, which immediately precedes 6-MSA synthesis, to increase the production of 6-MSA in *S. cerevisiae*. Our next step was to increase the flux through the 6-MSA synthesis pathway by looking at a broad array of competing pathways for elimination. We used OptKnock from the COBRA Toolbox (Schellenberger et al., 2011) for the prediction of host gene deletion targets for

the increased synthesis of 6-MSA or its precursors, and then tested these knockout strains for 6-MSA production.

In other work in our lab, strains have been engineered for increased synthesis of TAL (triacetic acid lactone) (Cardenas and Da Silva, 2014). TAL synthesis requires 1 acetyl-CoA and 2 malonyl-CoA, while 6-MSA synthesis requires 1 acetyl-CoA, 3 malonyl-CoA and 1 NADPH. We have seen improvements in TAL synthesis by deleting a few upstream pathway genes. Since both TAL and 6-MSA use malonyl-CoA as a major building block, selected gene deletions from the TAL work were tested for 6-MSA synthesis (Table 5.6, Figure 5.9). Among these deletion candidate genes, *ZWF1* and *FBP1* was selected to redirect carbon flux toward malonyl-CoA synthesis from pentose sugars and glycogen, respectively. *PYC1* and *PYC2* were selected for optimized synthesis of pyruvate, while *GPD1* was selected for optimized synthesis of acetyl-CoA. We, also, performed an OptKnock *in silico* analysis specifically for increased synthesis of 6-MSA.

The COBRA toolbox, libSBML, the SBMLToolbox, and gurobi 5.6 (Gurobi Optimization www.gurobi.com) were employed (Mo et al., 2009; Schellenberger et al., 2011) using MATLAB. A modified iMM904 model was loaded (Zomorrodi and Maranas, 2010) and the 6-MSA synthesis reaction was added to the model. The model objective was set to optimize for 6-MSA synthesis. OptKnock was run to obtain gene deletion predictions to achieve the maximum level of 6-MSA synthesis. Predicted pathway deletions (Table 5.7, Figure 5.9) were tested for increased 6-MSA synthesis.

6-MSA synthesis requires expression of both 6-MSAS and a PPT (4'-phosphopantetheinyl transferase). Thus, a 2μ plasmid (pKUTP-6MN) carrying the genes for both 6-MSAS and NpgA, a PPT from *Aspergillus nidulans*, was constructed. To avoid loss via recombination between

repeated sequences, two different promoters and two different terminators were selected for 6-MSAS and *npgA*. The *TEF1* promoter has been shown to result in stronger gene expression than the *PGK1* promoter (Partow et al., 2010). Previously, it has been shown that two integrated copy of *npgA* produced higher 6-MSA titer compared to 2μ-based *npgA* (Lee, 2006). Therefore, the *TEF1* promoter was used for 6-MSAS, while the *PGK1* promoter was used for NpgA. The *ADH2* terminator was used for 6-MSAS and the *CYC1* terminator was used for NpgA. This plasmid was transformed into a selection of gene deletion strains.

Table 5.6. List of gene deletions based on the enhanced TAL synthesis (Cardenas and Da Silva, 2014)

Deleted genes	**Enzyme and Pathway**
ZWF1	Glucose-6-phosphate dehydrogenase, initial step of pentose phosphate pathway
FBP1	Fructose-1,6-bisphosphatase, gluconeogenesis pathway
GPD1	Glycerol-3-phosphate dehydrogenase, glycerol synthesis pathway
PYC1	Pyruvate carboxylase, gluconeogenesis pathway
PYC2	Pyruvate carboxylase, gluconeogenesis pathway, *PYC1* isoform

Table 5.7. List of gene deletions predicted by OptKnock for enhanced 6-MSA synthesis

Deleted genes	**Enzyme and Pathway**
PLB1	Phospholipase B, glycerophospholipid metabolic process, phosphatidylcholine acyl-chain remodeling
PLB2	Phospholipase B, glycerophospholipid metabolic process
RHR2	DL-glycerol-3-phosphate phosphatase, glycerol biosynthesis
HOR2	DL-glycerol-3-phosphate phosphatase, glycerol biosynthesis, *RHR2* paralog
TPO1	Polyamine transporter
INM1	Inositol monophosphatase, myo-inositol biosynthesis

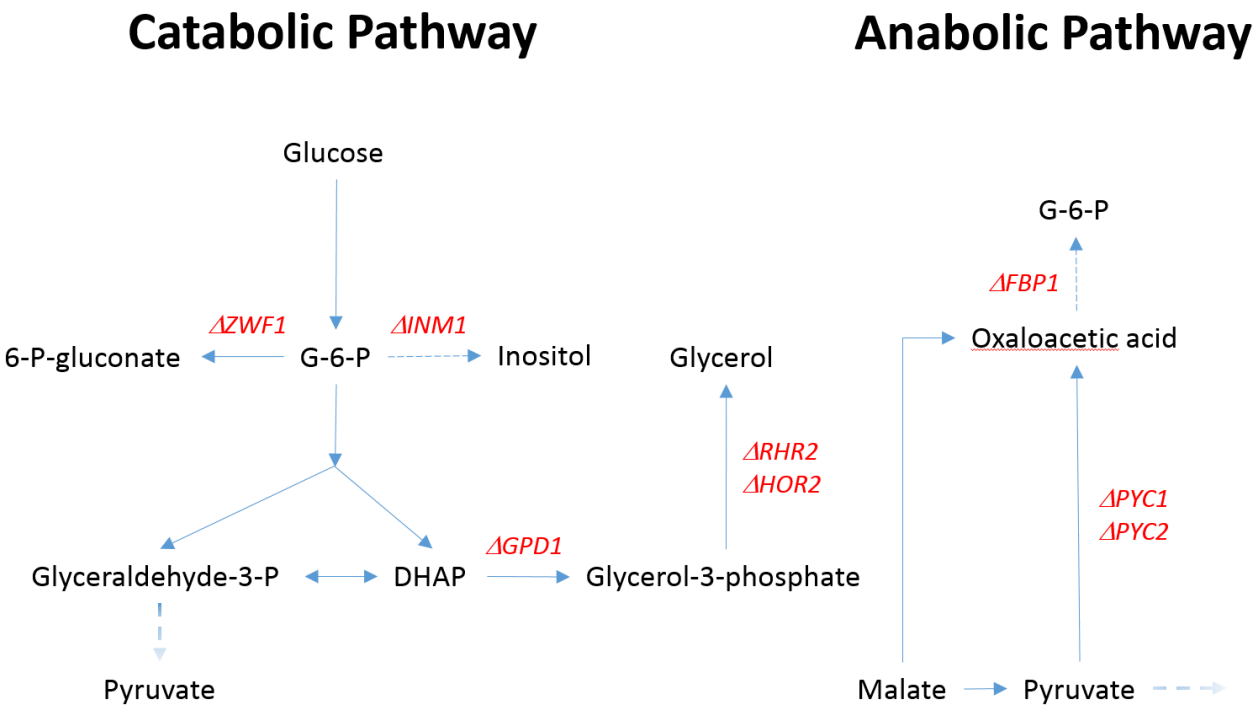


Figure 5.9. Diagram of pathways involving genes targeted for deletion.

The deletion strains containing the dual-gene plasmid (pKUTP-6MN) were cultivated in modified 1 % SDC medium (M2) (Table 5.3) for 48 hours and the supernatant was analyzed for 6-MSA. The HPLC results (Figure 5.10) identified a promising gene deletion (*PYC1*) for improved 6-MSA synthesis. The deletion of *PYC1* resulted in 50 % improvement for both 6-MSA titer and specific level. *PYC1* is responsible for the early part of gluconeogenesis and the deletion of this gene should redirect more carbon toward 6-MSA synthesis. *Δpyc2* and *Δplb1* showed modest improvements of 23 % and 16 %, respectively.

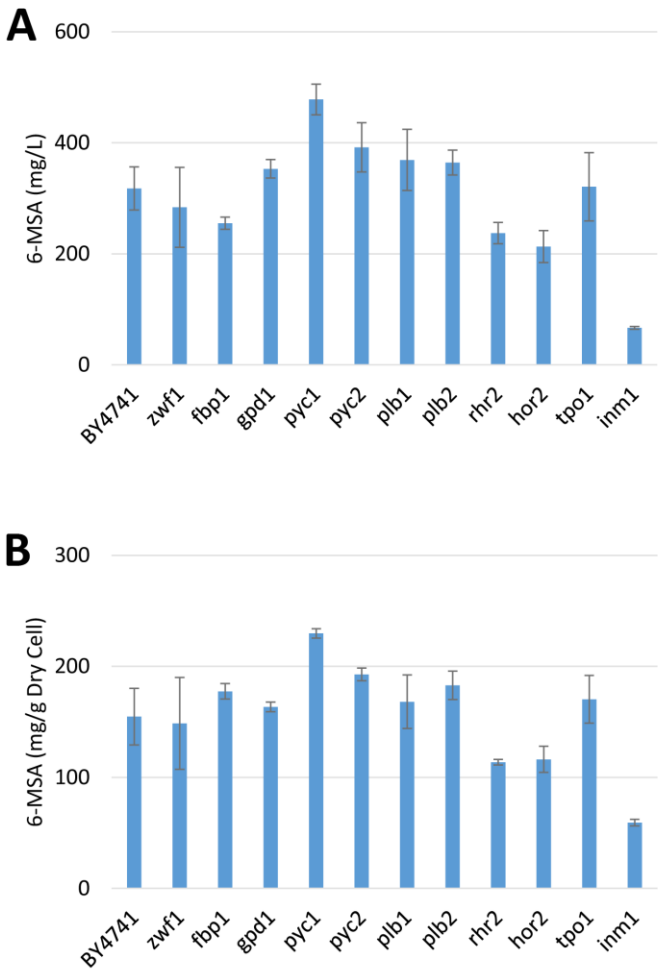


Figure 5.10. 6-MSA production from OptKnock predicted deletions in modified 1 % SDC medium. 6-MSA was produced in a selection of strains with gene deletions predicted by OptKnock for maximum 6-MSA synthesis. (A) 6-MSA titer. (B) 6-MSA production per gram

The same set of deletion strains was tested in two other media: 2 % mSDC (M1) and enhanced 1 % eSDC media (M3) (Table 5.3, Figure 5.11). While the trend from M2 is repeated in M1 and M3, *Δfbp1* improved specific 6-MSA production by 35 % (g/g dry cell) relative to BY4741 in the enriched M3 medium, although titer decreased (Figure 5.11). 6-MSA levels in M1 medium did not show significant improvement, with only *Δpyc1* showing 28% improvement over control (BY4741). *Δpyc1* showed identical levels of specific 6-MSA production between M2 and M3.

Δfbp1 blocks the only reversible step in the glycolysis pathway (Fbp1 is responsible for the reaction toward anabolic pathway) with Pfk1 responsible for reaction toward catabolic pathway (Lin et al., 2001). Pyc1 is responsible for the conversion of pyruvate to oxaloacetic acid and is the first committed step in gluconeogenesis. *Δfbp1* lead to the significant drop in biomass accumulation while *Δpyc1* did not have a significant effect.

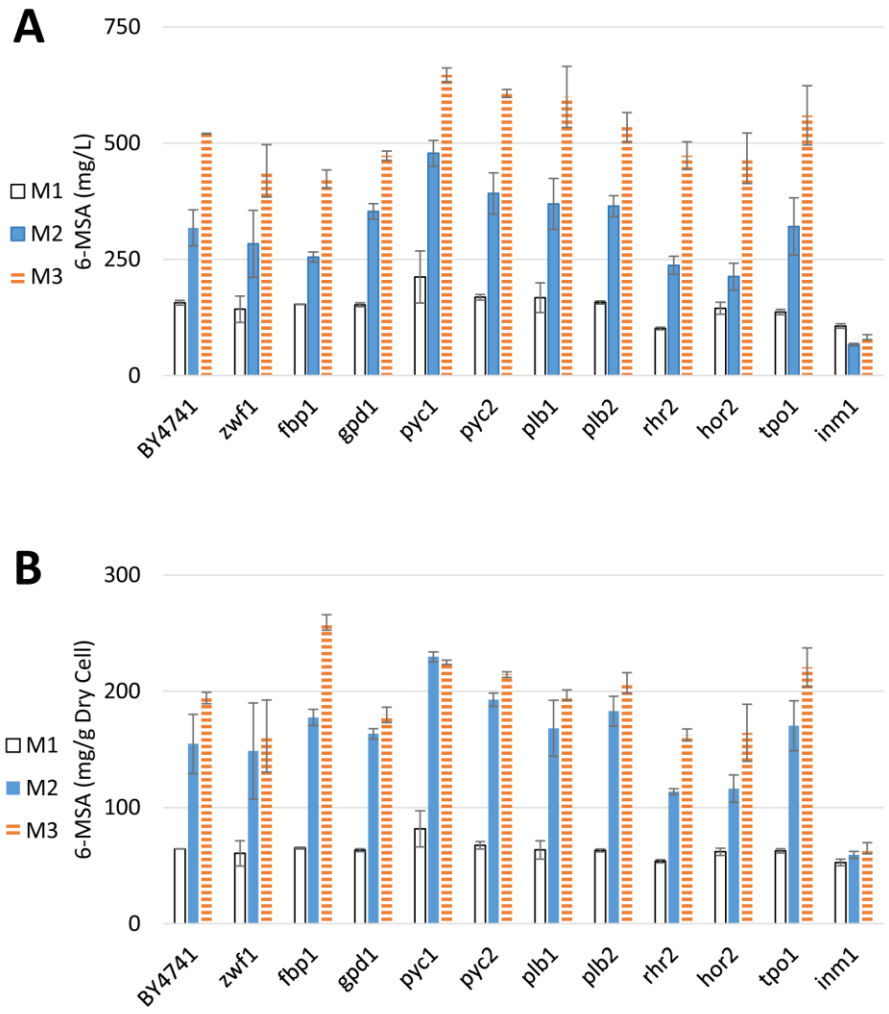


Figure 5.11. Comparison of 6-MSA production using strains with different upstream pathway gene deletions in 3 different media. 6-MSA was produced in 3 different media for 48 hours. M1, M2, and M3 are described in Table 6.2. (A) 6-MSA Titer. (B) Specific 6-MSA level per gram cell.

5.4.4. Combined effect of pyruvate dehydrogenase bypass gene overexpression and *PYC1* deletion

In the previous sections, we presented results on the overexpression of genes in the pyruvate dehydrogenase bypass including *S. cerevisiae* native *ACS1*, and *S. enterica ACS* and its mutant, native *CAB1*, and native *ACC1* and its mutant. We also tested knockouts of upstream pathway genes. While the overexpression of *ACS1*, ACS_{SE}, ACS_{SE}^{L641P} ($ACS_{SE}m$), and *CAB1* did not improve 6-MSA level, overexpression of the *ACC1* mutant and deletion of *PYC1* resulted in significant improvement in 6-MSA levels.

We thus tested if combined overexpression of ACS_{SE}^{L641P} and $ACC1^{S1157A}$ can further improve 6-MSA production. However, the new strain lowered the 6-MSA level relative to the $ACC1^{S1157A}$ overexpression strain (BJPN1Cm-P6M) with a large error bar (data not shown). Recently, Chen et al. reported about 40% improvement in 3-hydroxypropionic acid (3-HP) when they combined *ACC1* overexpression with *ADH2*, *ALD6*, and ACS_{SE}^{L641P}. Thus, we overexpressed *ADH2* (under the *HXT7* promoter) in BJPN1, BJPN1S_{SE}m, BJPN1Cm, and BJPN1S_{SE}mCm anticipating improved carbon flow from ethanol toward malonyl-CoA. These cells were cultivated in 5 ml of 1% mSDC(A,T) medium for 48 h. Overexpression of *ADH2*, ACS_{SE}^{L641P} and $ACC1^{S1157A}$ simultaneously resulted in about 20% improvement over the *ADH2* and $ACC1^{S1157A}$ overexpressed strain (Figure 5.12 A). However, *ADH2* overexpression under HXT7 promoter reduced 6-MSA titer about 10% for all samples. Thus, the final improvement with *ADH2*, ACS_{SE}^{L641P}, and $ACC1^{S1157A}$ simultaneous overexpression was only about 10% relative to $ACC1^{S1157A}$ overexpressed strain

(data not shown). *ALD6* overexpression was also tried together with *ADH2* and *ACC1*^{S1157A} overexpression but was detrimental to 6-MSA synthesis (data not shown).

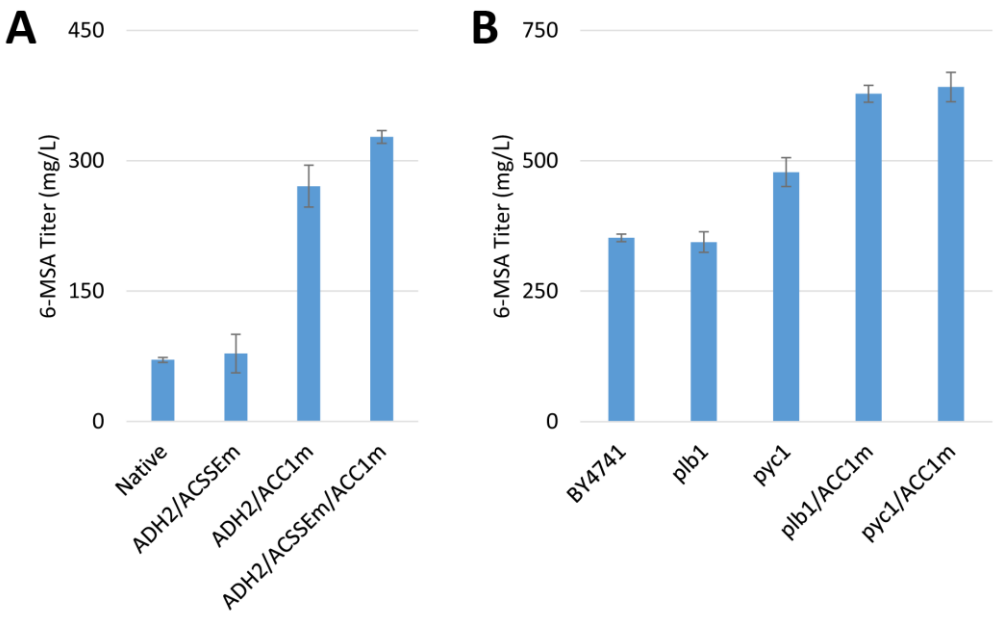


Figure 5.12. 6-MSA titer after combining gene overexpression and knockouts. (A) Native: *ADH2* overexpressed in BJPN1; *ADH2*/*ACS*$_{SE}$*m*: *ADH2* and *ACS*$_{SE}$*m* (*ACS*$_{SE}^{L641P}$) overexpressed; ADH2/ACSSEm/ACC1m: *ADH2*, *ACS*$_{SE}^{L641P}$, and *ACC1*^{S1157A} (ACC1m) overexpressed. (B) *ACC1*^{S1157A} (ACC1m) was overexpressed with *PLB1* or *PYC1* knockouts.

We also overexpressed *ACC1*^{S1157A} in the *PYC1* and *PLB1* knockout strains in 5 ml 1% mSDC(A,T) medium for 72 h. This led to approximately an 80% improvement for the *PLB1* knockout strain and a 34% improvement for the *PYC1* knockout strain relative to *plb1* or *pyc1* only strains, respectively (Figure 5.12 B).

In combined gene overexpression and deletion test, only *ACC1*^{S1157A} overexpression and *PYC1* and *PLB1* deletions proved to be effective while more deletions need to be tested.

5.5. Conclusions

Acetyl-CoA and malonyl-CoA are key building blocks for polyketide synthesis. Our objective in this study was to increase the intracellular availability of malonyl-CoA. To do that, we overexpressed *ACS1* and ACS_{SE}, ACS_{SE}^{L641P}, *CAB1*, and *ACC1* alone or in combination. 6-MSA level was not increased by overexpression of ACS_{SE}^{L641P} or *CAB1*. The combination of ACS_{SE}^{L641P} and *CAB1* led to only slight improvement (19 %). The overexpression of *ACC1* led to mild improvement (35 % for BY4741 strain and 45 % for BJ5464 strain) in 6-MSA titer. In contrast, incorporating the S1157A mutation in Acc1 to prevent deactivation of Acc1 following glucose depletion had a significant effect on both 6-MSA and fatty acid levels. Activity of the enzyme during the ethanol phase increased 9-fold (*in vitro* assay), and 6-MSA titers increased 3-fold (*in vivo*). The simultaneous 3-fold improvement in total fatty acid levels further demonstrated the promise of this enzyme. Additional pathway engineering focused on an array of competing pathways and considered gene deletions on 6-MSA production. *Δpyc1* improved 6-MSA titer by 50 %, while *Δfbp*1 showed 35 % improved specific 6-MSA level. These two knockouts combined with the modified Acc1 showed an additional 80 % improvement for *plb1* and 34 % improvement for *pyc1* strain. The strategy reported in this study will be beneficial for the production of a variety of polyketide or fatty acid products for pharmaceutical, biorenewable chemical, or biofuels applications.

5.6. References

Cardenas, J., Da Silva, N. A., 2014. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. Metab Eng. 25C**,** 194-203.

Chen, Y., Bao, J., Kim, I.-K., Siewers, V., Nielsen, J., 2014. Coupled incremental precursor and co-factor supply improves 3-hydroxypropionic acid production in Saccharomyces cerevisiae. Metabolic Engineering. 22**,** 104-109.

Chia, M., Schwartz, T. J., Shanks, B. H., Dumesic, J. A., 2012. Triacetic acid lactone as a potential biorenewable platform chemical. Green Chemistry. 14**,** 1850-1853.

Chooi, Y.-H., Tang, Y., 2012. Navigating the Fungal Polyketide Chemical Space: From Genes to Molecules. Journal of Organic Chemistry. 77**,** 9933-9953.

Crawford, J. M., Townsend, C. A., 2010. New insights into the formation of fungal aromatic polyketides. Nature Reviews Microbiology. 8**,** 879-889.

Dale, S., Wilson, W. A., Edelman, A. M., Hardie, D. G., 1995. Similar substrate recognition motifs for mammalian AMP-activated protein-kinase, higher-plant HMG-CoA reductase kinase-A, yeast SNF1, and mammalian calmodulin-dependent protein kinase I. Febs Letters. 361**,** 191-195.

Davies, S. P., Sim, A. T. R., Hardie, D. G., 1990. Location and function of 3 sites phosphorylated on rat acetyl-CoA carboxylase by the AMP-activated protein-kinase. European Journal of Biochemistry. 187**,** 183-190.

Dimroth, P., Walter, H., Lynen, F., 1970. Biosynthesis of 6-methylsalicylic acid. European Journal of Biochemistry. 13**,** 98-110.

Eckermann, S., Schroder, G., Schmidt, J., Strack, D., Edrada, R. A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpelainen, I., Proksch, P., Teeri, T. H., Schroder, J., 1998. New pathway to polyketides in plants. Nature. 396**,** 387-390.

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2011. A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 28**,** 123-136.

Fortman, J. L., Chhabra, S., Mukhopadhyay, A., Chou, H., Lee, T. S., Steen, E., Keasling, J. D., 2008. Biofuel alternatives to ethanol: pumping the microbial well. Trends in Biotechnology. 26**,** 375-381.

Gibson, D., 2009. One-step enzymatic assembly of DNA molecules up to several hundred kilobases in size.

Gietz, R. D., Schiestl, R. H., Willems, A. R., Woods, R. A., 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast. 11**,** 355-360.

Gietz, R. D., Woods, R. A., 2001. Genetic transformation of yeast. Biotechniques. 30**,** 816-831.

Ha, J., Daniel, S., Broyles, S. S., Kim, K. H., 1994. Critical phosphorylation sites for acetyl-CoA carboxylase activity. Journal of Biological Chemistry. 269**,** 22162-22168.

Jones, E. W., 1991. Tackling the protease problem in *Saccharomyces cerevisiae*. Methods in Enzymology. 194**,** 428-453.

Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., Barr, P. J., 1998. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. Proceedings of the National Academy of Sciences of the United States of America. 95**,** 505-509.

Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C., Hutchinson, C. R., 1999. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science. 284**,** 1368-1372.

Kozak, B. U., van Rossum, H. M., Benjamin, K. R., Wu, L., Daran, J.-M. G., Pronk, J. T., van Maris, A. I. J. A., 2014. Replacement of the Saccharomyces cerevisiae acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. Metabolic Engineering. 21**,** 46-59.

Leber, C., Da Silva, N. A., 2014. Engineering of *Saccharomyces cerevisiae* for the synthesis of short chain fatty acids. Biotechnology and Bioengineering. 111**,** 347-358.

Lee, K. K. M., Engineering of *Saccharomyces cerevisiae* for the biosynthesis of fungal polyketides. University of California, Irvine, 2006.

Lee, K. K. M., Da Silva, N. A., Kealey, J. T., 2009. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in Escherichia coli and Saccharomyces cerevisiae. Analytical Biochemistry. 394**,** 75-80.

Lin, S. S., Manchester, J. K., Gordon, J. I., 2001. Enhanced gluconeogenesis and increased energy storage as hallmarks of aging in Saccharomyces cerevisiae. J Biol Chem. 276**,** 36000-7.

Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X. K., Kealey, J. T., Da Silva, N. A., Vederas, J. C., Tang, Y., 2009. Complete reconstitution of a highly reducing iterative polyketide synthase. Science. 326**,** 589-592.

Matsuhashi, M., 1969. Acetyl-CoA Carboxylase from Yeast EC 6.4.1.2 Acetyl-CoA: carbon-dioxide ligase (ADP). Methods in enzymology. 95.

Mo, M. L., Palsson, B. O., Herrgard, M. J., 2009. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. Bmc Systems Biology. 3.

Munday, M. R., Campbell, D. G., Carling, D., Hardie, D. G., 1988. Identification by amino-acid sequencing of 3 major regulatory phosphorylation sites on rat acetyl-CoA carboxylase. European Journal of Biochemistry. 175**,** 331-338.

Nikolau, B. J., 2010. An integrated strategy for generating lipid-based biorenewable chemicals: diversifying fatty acid synthesis with polyketide synthesis biocatalysts. Chemistry and Physics of Lipids. 163**,** S16-S17.

Partow, S., Siewers, V., Bjorn, S., Nielsen, J., Maury, J., 2010. Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. Yeast. 27**,** 955-964.

Robishaw, J. D., Berkich, D., Neely, J. R., 1982. Rate-limiting step and control of coenzyme-A synthesis in cardiac-muscle. Journal of Biological Chemistry. 257**,** 967-972.

Rock, C. O., Calder, R. B., Karim, M. A., Jackowski, S., 2000. Pantothenate kinase regulation of the intracellular concentration of coenzyme A. Journal of Biological Chemistry. 275**,** 1377-1383.

Sambrook, J., Russell, D. W., 2001. Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., Kang, J., Hyduke, D. R., Palsson, B. O., 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nature Protocols. 6**,** 1290-1307.

Scott, J. W., Norman, D. G., Hawley, S. A., Kontogiannis, L., Hardie, D. G., 2002. Protein kinase substrate recognition studied using the recombinant catalytic domain of AMP-activated protein kinase and a model substrate. Journal of Molecular Biology. 317**,** 309-323.

Shen, M. W. Y., Fang, F., Sandmeyer, S., Da Silva, N. A., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 29**,** 495-503.

Shiba, Y., Paradise, E. M., Kirby, J., Ro, D. K., Keasing, J. D., 2007. Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. Metabolic Engineering. 9**,** 160-168.

Shirra, M. K., Patton-Vogt, J., Ulrich, A., Liuta-Tehlivets, O., Kohlwein, S. D., Henry, S. A., Arndt, K. M., 2001. Inhibition of acetyl coenzyme a carboxylase activity restores expression of the INO1 gene in a snf1 mutant strain of Saccharomyces cerevisiae. Molecular and Cellular Biology. 21**,** 5710-5722.

Song, W. J., Jackowski, S., 1992. Cloning, sequencing, and expression of the pantothenate kinase (coaA) gene of Escherichia-coli. Journal of Bacteriology. 174**,** 6411-6417.

Spencer, J. B., Jordan, P. M., 1992. Purification and properties of 6-methylsalicylic acid synthase from *Penicillium patulum*. Biochemical Journal. 288**,** 839-846.

Vogel, G., Lynen, F., 1975. 6-Methylsalicylic acid synthetase. Methods in enzymology. 43**,** 520-30.

Wattanachaisaereekul, S., Lantz, A. E., Nielsen, M. L., Nielsen, J., 2008. Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. Metabolic Engineering. 10**,** 246-254.

Weekes, J., Ball, K. L., Caudwell, F. B., Hardie, D. G., 1993. Specificity determinants for the AMP-activated protein-kinase and its plant homolog analyzed using synthetic peptides. Febs Letters. 334**,** 335-339.

Woods, A., Munday, M. R., Scott, J., Yang, X. L., Carlson, M., Carling, D., 1994. Yeast Snf1 is functionally related to mammalian AMP-activated protein-kinase and regulates acetyl-CoA carboxylase *in vivo*. Journal of Biological Chemistry. 269**,** 19509-19515.

Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., Tang, Y., 2013. LovG: The Thioesterase Required for Dihydromonacolin L Release and Lovastatin Nonaketide Synthase Turnover in Lovastatin Biosynthesis. Angew Chem Int Ed Engl. 52**,** 6472-5.

Zomorrodi, A. R., Maranas, C. D., 2010. Improving the iMM904 S. cerevisiae metabolic model using essentiality and synthetic lethality data. Bmc Systems Biology. 4.

Chapter 6.

Improved 6-MSA Synthesis via Enhanced 6-MSA Synthase Expression System

6.1. Abstract

Saccharomyces cerevisiae is an excellent host for the heterologous gene expression and product synthesis. To improve 6-MSAS expression and 6-MSA production in this yeast, we focused on evaluating an engineered auxotrophic marker, an autoselection system, and complex media. We made two initial comparisons by expressing an integrated copy of 6-MSAS under the strong glycolytic *PGK1* promoter. Use of complex media resulted in a 6-fold improvement in 6-MSA levels over standard non-selective SDC(A,T) medium. In the latter medium, 6-MSAS expression from a 2μ-based plasmid led to 24-fold higher 6-MSA titers compared to the 6-MSAS integrated strain. These two comparisons showed the impact of using rich media and the 2μ-based multi-copy plasmid for expression of polyketide synthases. We then introduced an engineered *URA3* marker (N-degron) to further increase the 6-MSAS expression level using the 2μ-based plasmid and combined it with an autoselection system via deletion of *FUR1* to ensure stability in rich complex medium. Use of the N-degron *URA3* marker led to more than 30% higher 6-MSA titers and employing the autoselection system via *Δfur1* led to a greater than 90% improvement in 6-MSA titer.

6.2. Introduction

Polyketides provide an immense pool of candidate molecules for application as pharmaceuticals and industrial chemicals. Use of well-developed and widely-used microorganisms such as *Escherichia coli* and *Saccharomyces cerevisiae* makes the development of polyketide production systems more efficient. The overall goal of Chapters 5 and 6 was to improve polyketide synthesis in *S. cerevisiae*. To achieve this goal, Chapter 5 explored pathway engineering with the overexpression of ACS_{SE}^{L641P}, *ACC1*, and $ACC1^{S1157A}$, and upstream gene deletions such as *PYC1 and PLB1*. In this study, we have considered additional factors that affect product synthesis to further improve 6-MSA production.

An important factor affecting 6-MSA levels is the copy number of the 6-MSAS plasmid. The recent development of the N-degron *URA3* marker by Chen et al. (2012) led to an increase in the plasmid copy number; this can be a promising tool for improved 6-MSAS expression. another critical factor is the selection of appropriate medium for the synthesis of 6-MSA. The seminal work on 6-MSA synthesis in *S. cerevisiae* (Kealey et al., 1998) was carried out using the *ADH2* promoter and rich but non-selective YPD medium. The level of 6-MSA synthesis in this media was 1.7 g/L. However, the use of earlier phase promoters such as the *PGK1* promoter requires the use selective media to avoid plasmid loss. Even with the late phase *ADH2* promoter some plasmid loss occurs (Shen et al., 2012). It would be beneficial to use complex media for the synthesis of polyketides without suffering significant plasmid loss. Thus, we looked at the possibility of combining the Acc1 mutant from Chapter 5, the N-degron URA3 marker, and the

use of rich media to improve 6-MSA synthesis in *S. cerevisiae*. To improve the plasmid stability in the complex media, we also introduced the use of an autoselection system (Loison *et al*. (1986).

6.3. Background

6.3.1. Manipulation of UMP synthesis pathway to obtain autoselection capability

Pyrimidine ribonucleotides are synthesized via two different pathways: the *de novo* biosynthesis pathway and the salvage pathway (Figure 6.1), that both lead to UMP (Uridine monophosphate). The *de novo* pathway is comprised of six enzymes, Ura1, 2, 3, 4, 5, and 10. Ura2, carbamoyl phosphate synthase, initiates the pathway using L-glutamate as its substrate (Denis-Duphil, 1989; Souciet et al., 1982). The UMP synthesis step is catalyzed by Ura3, orotidine-5'-phosphate decarboxylase, which is encoded by the widely used auxotrophic selection gene *URA3* (Umezu et al., 1971). *URA3* null mutants require supplementation of uracil or uridine in the growth media (Jones, 1992). Salvage pathway synthesizes UMP directly from uridine or uracil by Urk1 and Fur1, respectively (Grenson, 1969). Uridine can be either imported via Fui1 or synthesized from cytidine by Cdd1 (Kurtz et al., 1999). Uracil can be either imported through Fur4 or synthesized from cytidine by Urh1 and Fcy1 (Erbs et al., 1997; Jund et al., 1988). Uracil can also be synthesized from uridine by Urh1 (Mitterbauer et al., 2002). A *Δurk1 Δurh1* mutant cannot survive with uridine as the only pyrimidine source. Both cytidine and cytosine are transported into cells via Fcy2 (Schmidt et al., 1984). Mitterbauer et al. (2002) was not able to utilize cytidine as the only source of pyrimidine even though Fcy2 and Cdd1 were intact. According to the authors, this might be because Fcy2 has very low affinity toward cytidine and, consequently, cytidine uptake is limited.

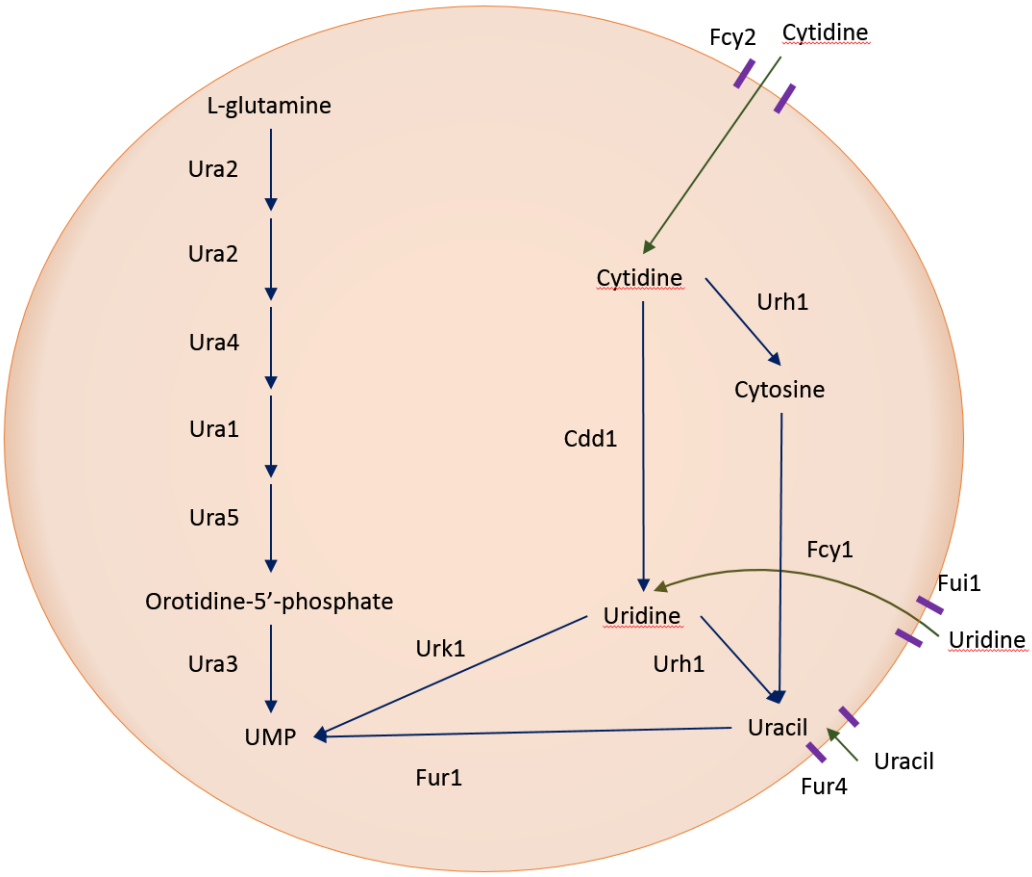


Figure 6.1 UMP biosynthesis pathway consists of *de novo* pathway and salvage pathway.

URA3 is extensively used as an auxotrophic marker for the selection of plasmids in *S. cerevisiae*. However, for this selection system to be effective, the media need to be devoid of uracil or uridine. Thus, complex media is not able to confer selection pressure on the *URA3* containing plasmid. However, *Δfur1 Δura3* double mutants will have to depend on uridine as the only way to synthesize UMP resulting in UMP limited growth. A *Δfur1* strain (Matα *fur1*) has been reported to have a three-fold longer doubling time than a *FUR1* strain (Matα *FUR1*) in YPD (Karpova et al., 1998). This result indicates the significance of the Fur1 contribution to UMP supply for cell survival. The deletion of both *FUR1* and *URA3* (de novo pathway) and the

introduction of *URA3* on a 2μ-based plasmid has been explored, and resulted in plasmid maintenance and an increase in foreign gene expression in the previously non-selective complex media (Loison et al., 1986; Napp and Da Silva, 1993).

6.3.2. Engineering of *URA3* marker for increased plasmid copy number

Ubiquitination of proteins signals proteasomal degradation or nonproteolytic activity such as in DNA repair, chromatin dynamics, mRNA export, and membrane protein trafficking (Finley et al., 2012). Hundreds of proteins in *S. cerevisiae* are subject to this post-translational modification process. Ubiquitin is usually attached to lysine residues on a target protein. Ubiquitin itself has multiple lysine residues that can be attached to another ubiquitin. This leads to polyubiquitin chain formation in diverse ways. Polyubiquitination signals the next modification of the target protein and it can be degradation signal. For these proteins to be recognized by the ubiquitination machinery, the target protein has to have specific amino acid sequences. One of these sequences is N-degron (Bachmair et al., 1986; Varshavsky, 1996). N-degron is formed by binding of a destabilizing amino acid residue to the N-terminus of the target protein usually preceded by cleavage of its initial N-terminus. This process follows diverse rules and methods and has been reviewed many times (Varshavsky, 1996; Varshavsky, 2011). Arginine has been shown to lead to the shortest protein half-life of beta-galactosidase (Bachmair et al., 1986). N-degron causes the ubiquitination at one or more specific lysine residue(s) of the protein and leads to proteolysis by proteasome 26S.

Ura3 is an interesting enzyme that is subject to ubiquitination (Peng et al., 2003). Chen *et al.* (2012) tested 4 different amino acids (Arg, Met, Glu, Gln) as N-degron by placing the tag between single ubiquitin (Ubi4) and the N-terminus of *URA3* with the native *URA3* promoter. These 4 N-degron marker variants were tested with LacZ on a 2μ plasmid to compare activity and plasmid copy number. The Arg-tethered marker increased both the copy number and activity by 70-80%. Subsequently, the *KEX2* promoter, which is a weak constitutive promoter, was used for expression of *URA3*. This led to combined improvement of LacZ activity and copy number by 3-fold compared to plasmids with the native *URA3* marker under its native promoter. Therefore, the combination of *Δura3 Δfur1* strain and a 2μ-plasmid with the N-degron *URA3* selection marker should be beneficial for 6-MSA synthesis.

6.4. Materials and methods

6.4.1. Molecular biology techniques

Plasmids were isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific). All PCR products were purified using either the Zymoclean™ gel DNA Recovery Kit (Zymo Research Corporation) or the QIAquick PCR Purification Kit (Qiagen). The Rapid DNA Ligation Kit (Thermo Scientific) was used for the ligation of DNA fragments. Restriction endonucleases, Taq DNA Polymerase, and NEB 10-beta competent cells were purchased from New England Biolabs Inc. KOD Hot Start DNA Polymerase (EMD Millipore) was used for high fidelity PCR while all other PCR reactions were carried out using Taq DNA Polymerase. All oligo nucleotides were synthesized by Integrated DNA Technologies, Inc. Synthesized genes were sequenced by Eton Bioscience Inc. *E. coli* competent cells were created using the calcium chloride method (Sambrook and Russell, 2001). All primer sequences are given in Table C.1 (Appendix C).

Plasmid transformation into *S. cerevisiae* was performed following the modified lithium acetate method (Gietz and Woods, 2001). Integration of linear DNA into *S. cerevisiae* cells was performed following the high-efficiency transformation method (Gietz et al., 1995).

6.4.2. Vector construction

Escherichia coli strains XL1-Blue (Stratagene) and DH5α (Invitrogen) were used as the primary strains for maintenance of plasmids and general cloning procedures. The pXP series of vectors (Table 3.1) were used for yeast expression (Fang et al., 2011; Shen et al., 2012).

Based on pIU13, an integration plasmid targeting the *URA3* locus, pIU13-6MSAS, was constructed for insertion of a single copy of the P_{PGK1}-6MSAS cassette into the genome. 6-MSAS, which is flanked by SpeI and RsrII, was digested and extracted from pJC842-6MSAS (Chapter 5, Table 5.1). The backbone vector was further digested by KpnI to differentiate the 6-MSAS fragment since the two initial fragments are similar in size. 6MSAS was ligated into pIU13 to create pIU13-6MSAS (Table 6.1, Figure 6.2).

The N-degron *URA3* marker, P_{KEX2}-Ubi-R-*URA3* was synthesized based on published primer sequences (Chen et al., 2012) using multiple PCR reactions. The *KEX2* promoter (PPKex2F and PPKex2R) and ubiquitin with an arginine residue (UBI-RF and UBI-RR) was amplified from BJ5464 genomic DNA. There is a 10 amino acid linker sequence between Arg and Ura3, and both UBI-RR and URA3F contained part of the linker sequence. *URA3* (URA3F and URA3R) was amplified from pJC812. The three PCR fragments were fused using a final round of PCR (PPKex2F and URA3R). The final fragment was digested using BamHI and KpnI, and ligated into the same sites on pJC842 to create pKA (Figure 6.3). pKA was digested with BsrGI (in the *CYC1* terminator) and KpnI, and the fragment containing the N-degron *URA3* marker was ligated into the same sites of pJC812-6MSAS and pJC842-6MSAS to create pKP-6MSAS and pKA-6MSAS, respectively (Table 6.1).

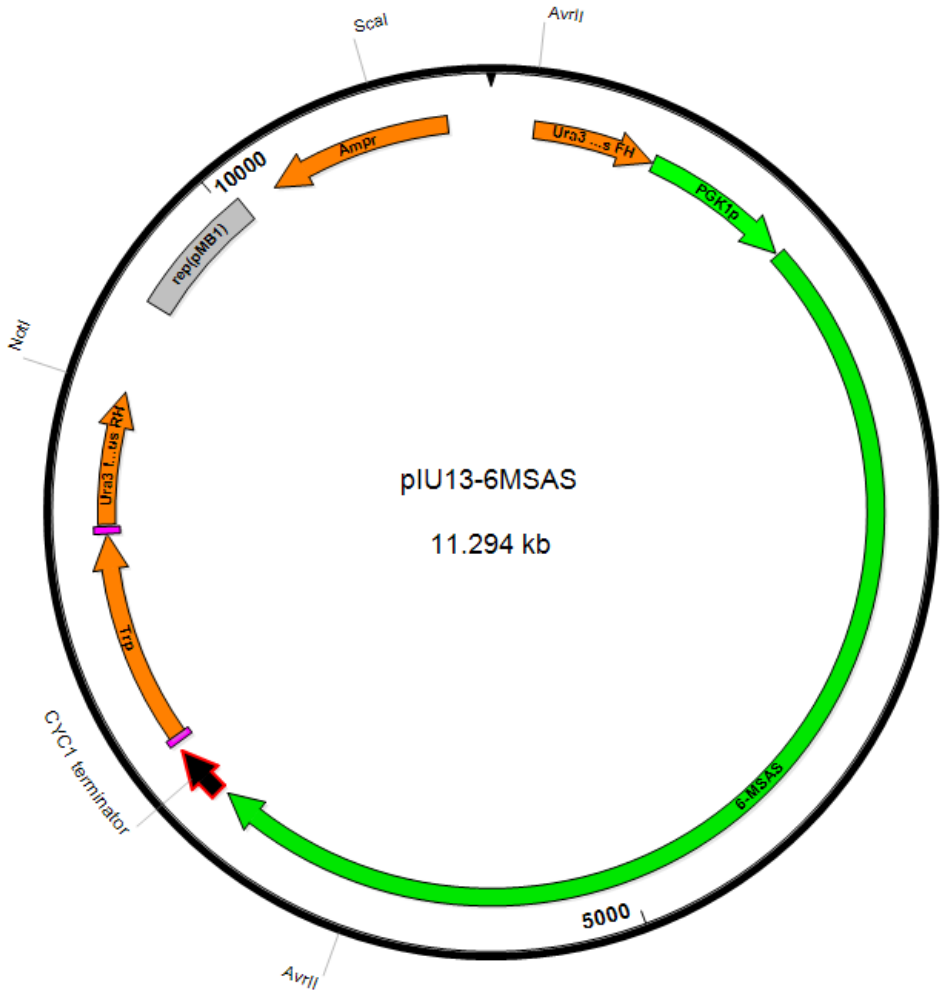


Figure 6.2 pIU13-6MSAS features and restriction sites for linearization.

Table 6.1. List of plasmids and yeast strains

Plasmids	Characteristics	Reference
pJC812	2μ, P_{PGK1}-T_{CYC1}, *URA3*	Chapter 3
pJC842	2μ, P_{ADH2}-T_{CYC1}, *URA3*	Chapter 3
pJC118-NpgA	CA, P_{PGK1}-*npgA*-T_{CYC1}, *URA3*	Chapter 5
pJC812-6MSAS	pJC812, 6-MSAS	Chapter 5
pJC842-6MSAS	pJC842, 6-MSAS	Chapter 5
YEplac195+6MSAS (YEp6MSAS)	2μ, P_{ADH2}-6MSAS-T_{ADH2}, *URA3*	Lee et al., 2009
pIU13	Integrating vector targeting *Δura3* locus, P_{PGK1}-T_{CYC1}, *TRP1*	Chapter 5
pIU13-6MSAS	pIU13, 6-MSAS	This study
pIM11-ACC1	pIM11, *ACC1*	Chapter 5
pIM11-ACC1m	pIM11, $ACC1^{S1157A}$	Chapter 5
pKA	pJC842, P_{KEX2}-*Ubi-R-URA3*	Chapter 5
pKA-6MSAS	pJC842-6MSAS, P_{KEX2}-*Ubi-R-URA3*	Chapter 5
pKP-6MSAS	pJC812-6MSAS, P_{KEX2}-*Ubi-R-URA3*	This study
pBF3060	2μ, P_{GAL1}-*CreA*-T_{CYC1}, *URA3* (constructed on pYES2 of Invitrogen Corporation)	Fang et al., 2011

Strains	Characteristics	Reference
BJ5464	*MATα his3-Δ200 leu2-Δ1 trp1 ura3-52 pep4::HIS3 prb1Δ1.6R*	(Jones, 1991)
BJPN1b	BJ5464, *Δleu2*::P_{PGK1}-*npgA*	This study
BJPN1Cmb	BJPN1, *Δmet17*::P_{PGK1}-$ACC1^{S1157A}$	This study
BJPN1CmP6M	BJPN1Cmb, *Δura3*::P_{PGK1}-6MSAS-*TRP1*	This study
BJPN1Cmb-P6M	BJPN1Cmb, pJC812-6MSAS	This study
BJPN1Cmb-A6M	BJPN1Cmb, pJC842-6MSAS	This study
BJPN1Cmb-KP6M	BJPN1Cmb, pKP-6MSAS	This study
BJPN1Cmb-KA6M	BJPN1Cmb, pKA-6MSAS	This study
BJPN1CmdF-KP6M	BJPN1Cmb, pKP-6MSAS, *Δfur1*::*TRP1*	This study
BJPN1CmdF-KA6M	BJPN1Cmb, pKA-6MSAS, *Δfur1*::*TRP1*	This study

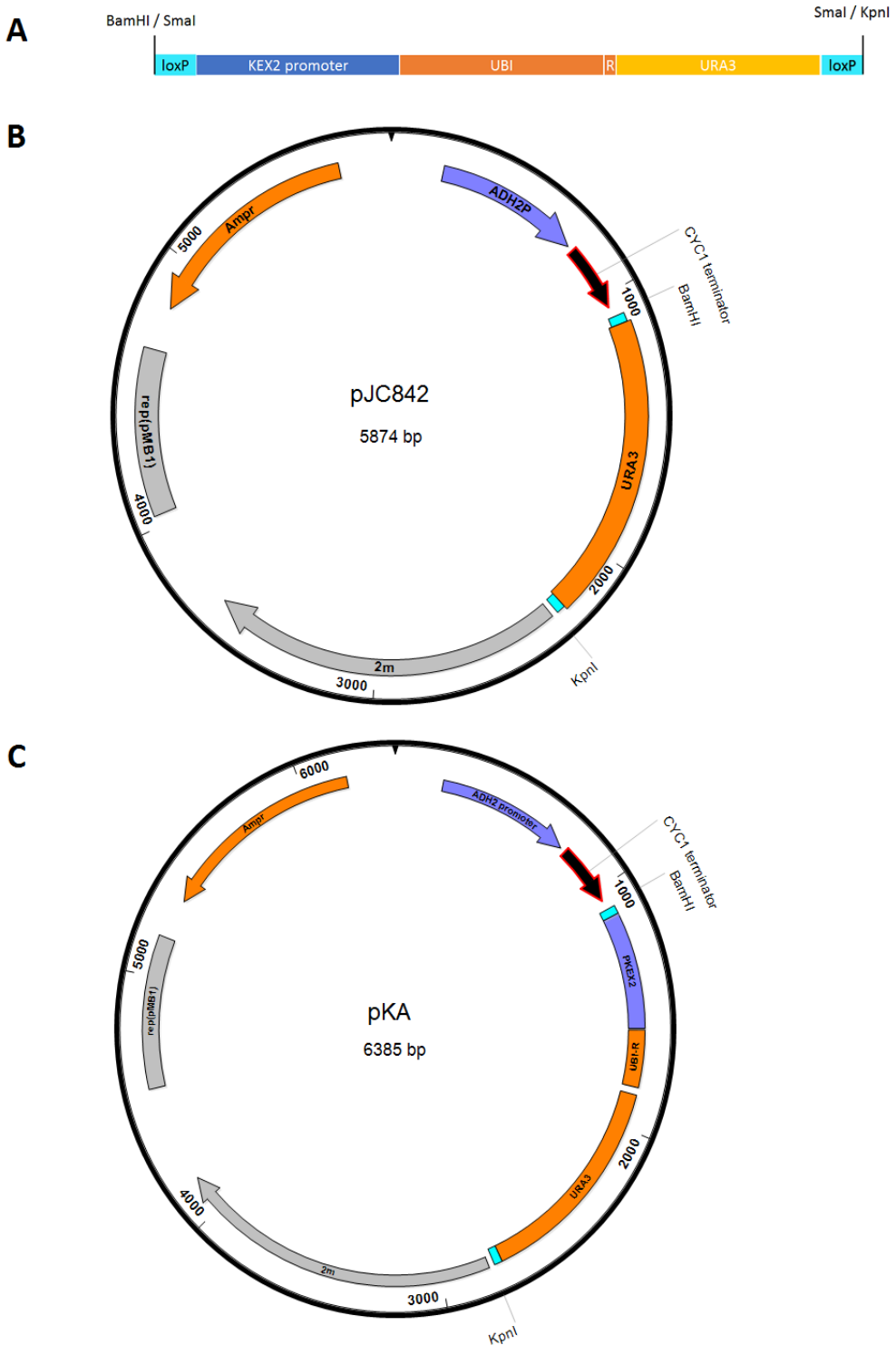


Figure 6.3 Diagram of pKA cloning. (A) N-degron *URA3* construct; (B) starting vector pJC842; (C) final construct, pKA.

6.4.3. Yeast strain construction

BJ5464 (*MAT*α *his3*-Δ*200 leu2*-Δ*1 trp1 ura3-52 pep4::HIS3 prb1*Δ*1.6R*) was used as base *S. cerevisiae* strains. CreA plasmid (pBF3060, Table 6.1) was transformed into BJPN1 and BJPN1Cm. The *TRP1* and *LEU2* markers were removed from BJPN1 and BJPN1Cm (Table 5.1), respectively, by expressing CreA. pBF3060 was removed by plating strains on 5-FOA (1 mg/ml) creating BJPN1b and BJPN1Cmb (Table 6.1). Correct auxotrophy was confirmed by replica-plating on selective plates with appropriate amino acid or nucleobase deficiency. Strain BJPN1CmP6M was constructed by integrating a single copy of the P_{PGK1}-6MSAS cassette into the genome of BJPN1Cmb. The *URA3* locus targeting integration vector, pIU13-6MSAS, was linearized by digestion with NotI and ScaI. The larger fragment was extracted followed by AvrII digestion. The linearized 8.8 kb fragment was then gel-purified and transformed into BJPN1Cmb (Table 6.1) to create BJPN1CmP6M.

N-degron evaluation strains were constructed by transforming two N-degron plasmids (pKP-6MSAS and pKA-6MSAS) into BJPN1Cmb creating BJPN1Cmb-KP6M and BJPN1Cmb-KA6M, respectively. The *FUR1* gene was deleted from strains BJPN1Cmb-KP6M and BJPN1Cmb-KA6M using double crossover homologous recombination to create BJPN1CmdF-KP6M and BJPN1CmdF-KA6M. For the deletions, the *TRP1* marker was PCR amplified (primers Pfur1delF and Pfur1delR) and transformed into the strains after gel-purification. Correct deletion of *FUR1* was confirmed by PCR (Pfur1delchkF and Pfur1delchkR).

6.4.4. Media and cultivation

Luria-Bertani (LB) medium was used for the cultivation of *E. coli* cells. Ampicilin (100 μg/ml) was used for the selection of plasmids in LB medium. *E. coli* cells were cultivated at 37°C and 250 rpm in an agitated air shaker.

S. cerevisiae was cultivated in non-selective YPD complex medium (2% YPD) (20 g/L dextrose, 20 g/L peptone, 10 g/L yeast extract (BD Biosciences, Sparks, MD)), standard SDC(A,T) medium (20 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/L L-tryptophan), modified 1% SDC(A,T) medium (1% mSDC(A,T)) (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/L L-tryptophan, 400mg/L L-serine, 200 mg/L L-threonine, and 20mM MES pH 5.5), modified SDC(A,U) medium (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100 mg/L adenine sulfate, 100 mg/L uracil, 400mg/L L-serine, 200 mg/L L-threonine, and 20mM MES pH 5.5), or modified selective SD(-LEU) medium (10 g/L dextrose, 5 g/L casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate, 100ug/L Biotin, 100 mg/L adenine sulfate, 100mg/L uracil, 100mg/L L-tryptophan, 100mg/L L-histidine, 400mg/L L-serine, 200mg/L L-threonine, 1.4 g/L Yeast Synthetic Drop-out Medium Supplements-without leucine, histidine, tryptophan, uracil (Y2001, Sigma-Aldrich, St. Louis, MO), and 20mM MES at pH 5.5) according to the marker requirement. 20 g/L agar was added into appropriate medium to prepare plates. Glucose was added as specified. *S. cerevisiae*

cells were incubated at 30°C and 250 rpm in an agitated air shaker. Cells were cultivated in 5 ml of medium in test tubes unless otherwise stated. Optical density was measured at 600 nm with UV/Vis spectrophotometer (Shimadzu, Kyoto, Japan).

6.4.5. Analytical methods

Cells were harvested via centrifugation and supernatants were analyzed by HPLC (LC-10ATvp pumps, Shimadzu, Kyoto, Japan) to measure the 6-MSA level. 6-MSA was separated at 0.2 ml/min and room temperature using a Zorbax SB-C18 column (Agilent Technologies, Santa Clara, CA) and detected at 306 nm using a SPD-10Avp UV detector. Acetonitrile with 1% acetic acid and water with 1% acetic acid were used in gradient mode. The method was started with 80% aqueous phase decreasing to 40% by 20 minutes. The aqueous phase was further decreased to 0% by 24 minutes and kept constant at 0% until 35 minutes. Then the aqueous phase was increased to 80% by 42 minutes. The column was equilibrated at 80% aqueous phase for 18 minutes finishing the run by 60 minutes. Retention time of 6-MSA was approximately 13 minutes.

6.5. Results and discussion

6.5.1. Copy number and media effects

Maintaining multiple copies of exogenous genes in *S. cerevisiae* is not always favored in host cells (Parekh et al., 1995). Sometimes, one integrated copy of the gene results in similar expression levels as a gene on a multi-copy plasmid. This could be due to the poor stability of multi-copy plasmids, and the burden to the host cells from both protein expression and enzymatic reaction products. In Chapter 4, one integrated copy of lovastatin nonaketide synthase (LovB) produced approximately 70% of the protein obtained on a 2μ-based multi-copy plasmid as seen in the SDS-PAGE (Figure 4.2). An advantage of using integrated copies is that the strain can be cultivated in a complex medium leading to higher biomass and product accumulation. Another advantage is that the copy number of the gene can be precisely controlled. Thus, we constructed strains carrying a single integrated copy of 6-MSAS (BJPN1CmP6M) or a multi-copy plasmid with 6-MSAS (BJPN1Cm-P6M) and compared 6-MSA production levels. 6-MSAS was expressed for 72 hours under the *PGK1* promoter. However, unlike the results observed for LovB, 6-MSA was produced at 24-fold higher level on the 2μ-based plasmid than with the single genomic copy (Figure 6.4). The 6-MSA integrated strain showed a 6-fold higher titer in complex medium than in selective medium. However, it was still 3-fold lower compared to the strain with the multi-copy 6-MSAS plasmid in selective medium. Therefore, the plasmid-based system was chosen for our studies.

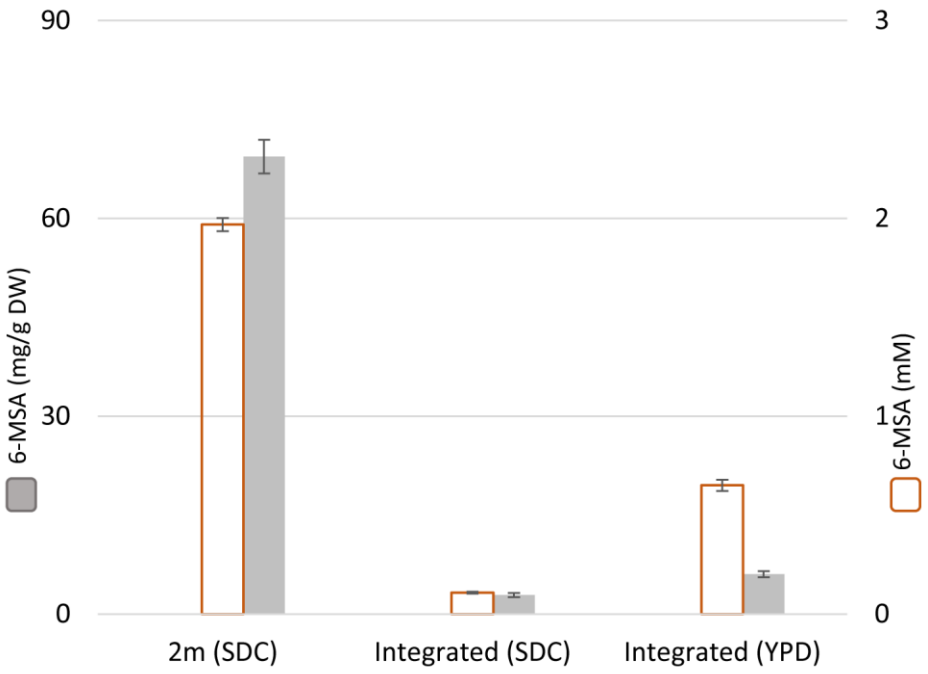


Figure 6.4. Effect of copy number and media on 6-MSA production. BJPN1Cmb was the host strain for both the 2μ-based plasmid system and the single-copy integrated system. Each vertical axis label is color-coded corresponding to the color of each bar. SDC: standard SDC(A,T) medium for 2μ, standard SDC(A,U) medium for Integrated.

6.5.2. Enhancing 6-MSA production by employing the N-degron *URA3* marker

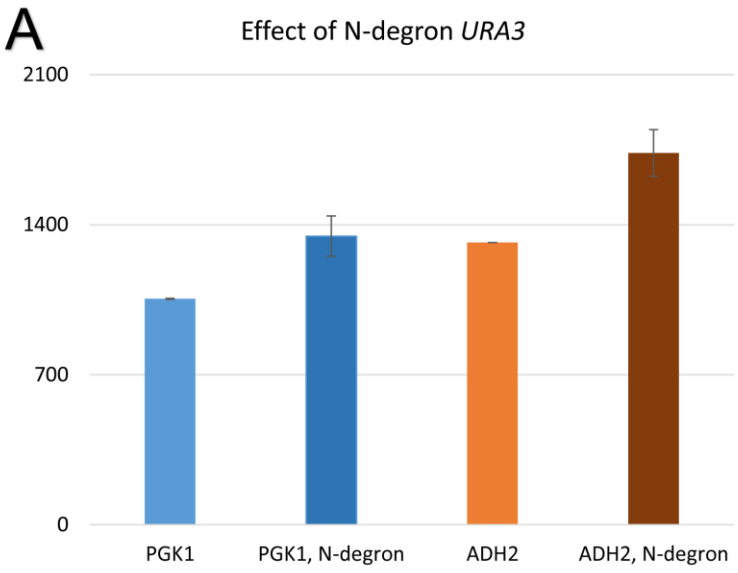
A recent report of an engineered *URA3* marker resulted in an increase in plasmid-based gene expression (Chen et al., 2012). The *URA3* promoter was replaced by the *KEX2* promoter and a single Ubiquitin sequence followed by an arginine codon was connected via a linker sequence to the N-terminus of *URA3*. The authors reported a 3-fold increase in both plasmid copy number and expressed enzyme activity. Deliberate addition of a single ubiquitin at the N-terminus of the *URA3* open reading frame (ORF) combined with the weaker constitutive *KEX2* promoter leads to

the reduced expression of *URA3* per copy of plasmid and shortened half-life of the Ura3 protein. Cells compensate for the reduced Ura3 levels by increasing the copy number of *URA3*-containing plasmids. We evaluated this marker to see the effects on 6-MSAS expression and 6-MSA synthesis.

The engineered N-degron *URA3* marker was tested with multi-copy plasmids carrying 6-MSAS under the *PGK1* promoter or the *ADH2* promoter. With the *PGK1* promoter, a 28% improvement in 6-MSA titer was observed for the N-degron *URA3* marker (BJPN1Cmb-KP6M) relative to the standard *URA3* marker (BJPN1Cmb-P6M) (Figure 6.5). With the *ADH2* promoter, a 32% improvement was observed (BJPN1Cmb-KA6M with N-degron *URA3*, BJPN1Cmb-A6M with standard *URA3*). The plasmid stability was not 100% for all strains since non-selective YPD medium was used. Use of the *PGK1* promoter system resulted in lower plasmid stability than the *ADH2* promoter because 6-MSAS expression starts while cells are growing exponentially. In contrast, the *ADH2* promoter initiates 6-MSAS transcription only after glucose is depleted (Price et al., 1990). Thus, the *PGK1* promoter system will give a greater metabolic burden to the cells than the *ADH2* promoter system.

The improvement in 6-MSA titer is substantially below the 3-fold increase in LacZ activity achieved by Chen *et al*. (Chen et al., 2012). The results indicate that the limitation in 6-MSA synthesis is only partially due to 6-MSAS copy number. This is not surprising given the high level of 6-MSA synthesis observed with the plasmid versus integrated systems (6.4.1). Other factors limiting 6-MSA production may be substrate availability for 6-MSA synthesis (e.g., acetyl-CoA, malonyl-CoA and NADPH). However, the improvement was still significant and we chose the N-degron *URA3* marker for our subsequent work.

A



B

Samples	Plasmid Stability
PGK1 (2% YPD)	58%
PGK1, N-degron (2% YPD)	61%
ADH2 (1% YPD)	75%
ADH2, N-degron (1% YPD)	75%

Figure 6.5. Comparison of standard *URA3* marker and N-degron *URA3* marker. (A) 6-MSA titer (mg/L); (B) Plasmid stability of each strain. Base strain is BJPN1Cmb.

6.5.3. Improving plasmid stability in complex medium through autoselection

Deletion of *FUR1* (coding for uracil phosphoribosyltransferase) disables uracil utilization from the culture medium and either uridine or synthesis via the *de novo* pathway is required for cell survival even in complex medium (Loison et al., 1986). This allows the use of plasmid-based

gene expression system using the *URA3* marker for plasmid selection in complex medium (Loison et al., 1986; Napp and Da Silva, 1993).

Previously, we tested 6-MSA production under both the *PGK1* and *ADH2* promoters in YPD medium (Figure 6.5). The plasmid stability varied from 58% to below 75% range. To increase plasmid stability in this non-selective complex medium, we deleted the *FUR1* gene from the strains expressing 6-MSAS from a 2μ-based plasmid with the N-degron marker (BJPN1Cmb-KP6M for *PGK1* promoter system, BJPN1Cmb-KA6M for *ADH2* promoter system) to create strains BJPN1CmdF-KP6M (*PGK1* promoter) and BJPN1CmdF-KA6M (*ADH2* promoter). Both strains were cultivated in modified 1% mSDC(A,T) and YPD medium for 48 hours. The *FUR1* deletion conferred no improvement in the modified 1% mSDC(A,T) medium (Figure 6.6). However, in YPD medium,

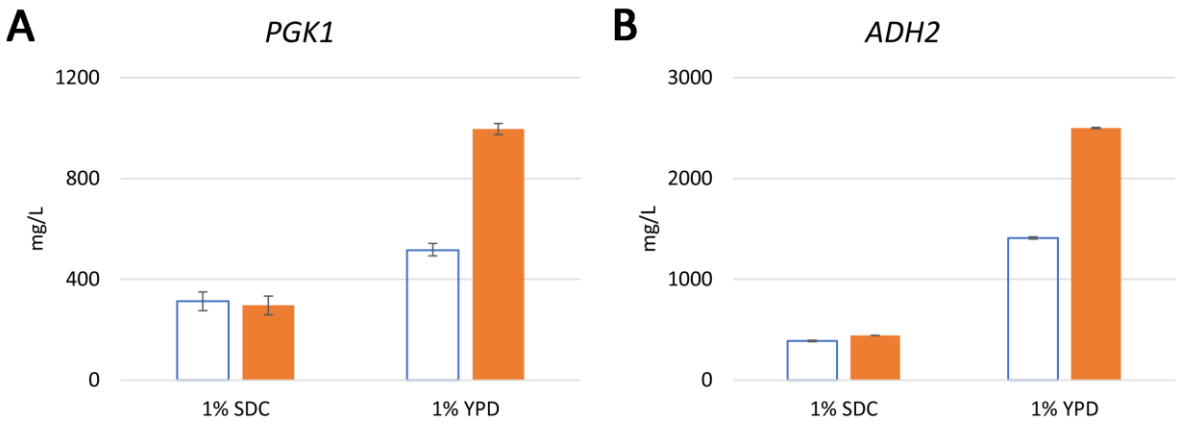


Figure 6.6. Comparison of 6-MSA producing strains between *FUR1* (open) and *fur1* (closed). All strains were cultivated for 48 hours with N-degron *URA3* marker in both modified 1% mSDC(A,T) media and 1% YPD media. Strains expressing 6-MSAS under the *PGK1* (A) promtoter and *ADH2* (B) promoter

the *FUR1* deletion improved 6-MSA levels by 93% and 77% with the *PGK1* and *ADH2* promoter system, respectively, relative to *FUR1* intact strains. The improved 6-MSA production in YPD clearly shows the substantial impact on plasmid stability when complex medium is now selective.

6.6. Conclusions

We utilized the engineering of expression vector (via promoter and auxotrophic marker choice), strain variations (gene knockouts), and comparison of different media to improve 6-MSAS expression and 6-MSA synthesis in *S. cerevisiae*. Using the *ACC1*^{S1157A} strain (Ch. 5), we compared different media conditions and 6-MSAS copy number, and found that both are limiting 6-MSA synthesis. We improved the 6-MSAS expression system by introducing an engineered *URA3* marker for higher copy number, and relying on autoselection via the *FUR1* deletion to improve plasmid stability in complex medium. The two improvements *Δfur1*/N-degron *URA3* led to a 92% increase in 6-MSA titer in 1% YPD for the *ADH2* promoter strain, and a 6.4-fold increase when compared to the *FUR1*/N-degron *URA3* strain in 1% mSDC(A,T) medium. Combining the overexpression of *ACC1*^{S1157A}, N-degron *URA3* marker, *Δfur1* and *Δura3*, upstream deletions in genes such as *PYC1* or *PLB1* or *FBP1*, and use of complex medium may allow further improvements in 6-MSA production in *S. cerevisiae*. The results are useful not only for 6-MSA synthesis, but also for the many other polyketides that use acetyl-CoA or malonyl-CoA as building blocks.

6.7. References

Bachmair, A., Finley, D., Varshavsky, A., 1986. In vivo half-life of a protein is a function of its amino-terminal residue. Science. 234**,** 179-186.

Chen, Y., Partow, S., Scalcinati, G., Siewers, V., Nielsen, J., 2012. Enhancing the copy number of episomal plasmids in Saccharomyces cerevisiae for improved protein production. Fems Yeast Research. 12**,** 598-607.

Denis-Duphil, M., 1989. Pyrimidine biosynthesis in Saccharomyces cerevisiae : the URA2 cluster gene, its multifunctional enzyme product, and other structural or regulatory genes involved in de novo UMP synthesis. Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire. 67**,** 612-631.

Erbs, P., Exinger, F., Jund, R., 1997. Characterization of the Saccharomyces cerevisiae FCY1 gene encoding cytosine deaminase and its homologue FCA1 of Candida albicans. Current Genetics. 31**,** 1-6.

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2011. A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 28**,** 123-136.

Finley, D., Ulrich, H. D., Sommer, T., Kaiser, P., 2012. The Ubiquitin-Proteasome System of Saccharomyces cerevisiae. Genetics. 192**,** 319-360.

Gietz, R. D., Schiestl, R. H., Willems, A. R., Woods, R. A., 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast. 11**,** 355-360.

Gietz, R. D., Woods, R. A., 2001. Genetic transformation of yeast. Biotechniques. 30**,** 816-831.

Grenson, M., 1969. The utilization of exogenous pyrimidines and recycling of uridine-5'-phosphate derivatives in saccharomyces cerevisiae, as studied by means of mutants affected in pyrimidine uptake and metabolism. European Journal of Biochemistry. 11**,** 249-60.

Jones, E. W., 1991. Tackling the protease problem in *Saccharomyces cerevisiae*. Methods in Enzymology. 194**,** 428-453.

Jones, M. E., 1992. Orotidylate decarboxylase of yeast and man. Current Topics in Cellular Regulation. 33**,** 331-342.

Jund, R., Weber, E., Chevallier, M. R., 1988. Primary structure of the uracil transport protein of saccharomyces cerevisiae. European Journal of Biochemistry. 171**,** 417-424.

Karpova, T. S., Moltz, S. L., Riles, L. E., Guldener, U., Hegemann, J. H., Veronneau, S., Bussey, H., Cooper, J. A., 1998. Depolarization of the actin cytoskeleton is a specific phenotype in Saccharomyces cerevisiae. Journal of Cell Science. 111**,** 2689-2696.

Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., Barr, P. J., 1998. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. Proceedings of the National Academy of Sciences of the United States of America. 95**,** 505-509.

Kurtz, J. E., Exinger, F., Erbs, P., Jund, R., 1999. New insights into the pyrimidine salvage pathway of Saccharomyces cerevisiae: requirement of six genes for cytidine metabolism. Current Genetics. 36**,** 130-136.

Lee, K. K. M., Da Silva, N. A., Kealey, J. T., 2009. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. Analytical Biochemistry. 394**,** 75-80.

Loison, G., Nguyenjuilleret, M., Alouani, S., Marquet, M., 1986. Plasmid-transformed ura3 fur1 double-mutants of Saccharomyces cerevisiae : an autoselection system applicable to the production of foreign proteins. Nature Biotechnology. 4**,** 433-437.

Mitterbauer, R., Karl, T., Adam, G., 2002. Saccharomyces cerevisiae URH1 (encoding uridine-cytidine N-ribohydrolase): Functional complementation by a nucleoside hydrolase from a protozoan parasite and by a mammalian uridine phosphorylase. Applied and Environmental Microbiology. 68**,** 1336-1343.

Napp, S. J., Da Silva, N. A., 1993. Enhancement of cloned gene product synthesis via autoselection in recombinant Saccharomyces cerevisiae. Biotechnology and Bioengineering. 41**,** 801-810.

Parekh, R., Forrester, K., Wittrup, D., 1995. Multicopy overexpression of bovine pancreatic trypsin-inhibitor saturates the protein-folding and secretory capacity of Saccharomyces cerevisiae. Protein Expression and Purification. 6**,** 537-545.

Peng, J. M., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D. M., Marsischky, G., Roelofs, J., Finley, D., Gygi, S. P., 2003. A proteomics approach to understanding protein ubiquitination. Nature Biotechnology. 21**,** 921-926.

Price, V. L., Taylor, W. E., Clevenger, W., Worthington, M., Young, E. T., 1990. Expression of heterologous proteins in *Saccharomyces cerevisiae* using the *ADH2* promoter. Methods in Enzymology. 185**,** 308-318.

Sambrook, J., Russell, D. W., 2001. Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Schmidt, R., Manolson, M. F., Chevallier, M. R., 1984. Photoaffinity labeling and characterization of the cloned purine cytosine transport system in Saccharomyces cerevisiae. Proceedings

of the National Academy of Sciences of the United States of America-Biological Sciences. 81**,** 6276-6280.

Shen, M. W. Y., Fang, F., Sandmeyer, S., Da Silva, N. A., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast. 29**,** 495-503.

Souciet, J. L., Hubert, J. C., Lacroute, F., 1982. Cloning and restriction mapping of the yeast URA2 gene coding for the carbamyl phosphate synthetase aspartate-transcarbamylase complex. Molecular and General Genetics. 186**,** 385-390.

Umezu, K., Amaya, T., Yoshimot.A, Tomita, K., 1971. Purification and properties of orotidine-5'-phosphate pyrophosphorylase and orotidine-5'-phosphate decarboxylase from baker's yeast. Journal of Biochemistry. 70**,** 249-262.

Varshavsky, A., 1996. The N-end rule: Functions, mysteries, uses. Proceedings of the National Academy of Sciences of the United States of America. 93**,** 12142-12149.

Varshavsky, A., 2011. The N-end rule pathway and regulation by proteolysis. Protein Science. 20**,** 1298-1345.

Appendices

A. Codon optimization and gene assembly

A.1. Codon and codon pair optimized gene sequences

A.1.1. *cALX1*

atgactgattatattccaacttttagatttgatggtcatttgactattgttacaggtgcatgtggaggtttggctgaagctctaattaagggtt
tgttggcttatggttctgatatagctctattggatatagatcaagaaaagactgctgctaaacaagctgaatatcataaatatgctacaga
agaattaaaattaaaagaagttccaaaaatgggttcttatgcttgtgacatttctgattctgatactgttcataaggtttttgctcaagttgc
taaagattttggaaaattgccattgcatttggttaataccgctggttattgtgaaaatttcccatgtgaagattatccagctaaaaatgcag
agaaaatggttaaggttaatttgttgggttctttgtatgtttcacaagcatttgctaaaccattgattaaagaaggaattaaaggtgcttca
gttgtacttattggttctatgtctggagctattgttaatgatccacaaaatcaagttgtttataatatgtctaaagctggtgtaattcatttgg
caaaaactcttgcttgtgaatgggctaaatataacattagagttaattctttaaatccaggttatatttatggtccattgactaaaaatgta
attaatggtaatgaagaattgtataatagatggatttctggaataccacaacaaagaatgtctgaacctaaagaatatattggtgctgttt
tgtatttgttgtctgagtctgctgcttcttatactacaggtgcatcattgttggttgatggtggttttacttcttgg

A.1.2. *clxr1*

atgccacaacctgttccaactgctaatagattgttagatttgttttctttaaaaggtaaagtagttgttgttactggtgcttctggtcctagag
gtatgggaattgaagctgctagaggatgtgctgaaatgggagcagatttagctattacttattcttctagaaaggaaggagctgagaaaa
atgctgaggaattgactaaagaatatggtgttaaagtaaaagtatataaggttaatcaatctgattataatgatgttgaaagatttgtaaa
tcaagttgtttcagatttcggtaaaattgatgcttttattgctaacgctggagctaccgctaattctggtgttgttgatggttctgcttcagatt
gggatcatgttatacaagttgatttgtctggaactgcttattgtgcaaaggctgtaggtgctcattttaaaaagcaaggtcatggttctttgg
ttataacagcttctatgtctggtcatgttgctaattatccacaagaacaaacttcatataatgttgcaaaagctggttgtattcatttggcta
gatctttggctaatgaatggagagattttgctagagttaattctatttcaccaggttatatagataccggtttgtctgattttattgatgaaaa
gactcaagaattgtggagatctatgattccaatgggtagaaatggtgatgctaaagaattaaaaggtgcttatgtttatttggtttcagatg
cttcttcttatactactggtgctgatatagtaattgatggtggttatactactaga

A.1.3. *clad1*

Atgtctccgtctgctgttgacgacgcgccaaaggctaccggtgctgcaatctctgttaaaccgaacatcggcgttttcaccaacccgaaac
acgacctgtggatctctgaagctgagccatctgcggacgctgttaaatctggtgcggacctgaaaccgggtgaagttaccatcgcggttc
gttctaccggtatttgcggttctgatgttcacttttggcacgcgggttgtatcggtccgatgatcgttgaaggcgaccacatcctgggtcacg
aatctgctggtgaagttatcgctgttcacccgaccgttagctctctccagatcggtgaccgtgttgcgatcgaaccgaacatcatctgcaac
gcttgcgaaccatgcctgaccggtcgttataacggctgcgaaaaagttgaatttctgtctaccccacctgttccgggcctcctgcgtcgtta
cgttaatcacccggcggtttggtgccacaaaatcggtaacatgtcttgggaaaatggtgctctgctcgaaccgctgtctgttgctctggctg
gtatgcaacgtgcgaaagttcagctgggcgacccggttctggtttgcggtgctggcccaatcggtctggtttcgatgctgtgtgctgctgct
gcgggtgcgtgtccactggttatcactgacatctctgaatctcgtctcgcgtttgcgaaagaaatctgtccacgtgtgactacccaccgtat
cgaaatcggtaaatctgctgaggaaaccgcgaaatctatcgttagctctttcggtggcgttgaaccagctgttaccctcgaatgtaccggt
gttgaatcttctatcgcggctgctatttgggcgtctaaattcggtggtaaagttttcgttatcggtgttggtaaaaacgaaatctctatcccgt
tcatgcgtgcgtctgttcgtgaagttgacatccagctccagtaccgttactctaacacttggccgcgtgctatccgtctgatcgaatctggtg
ttattgacctgtctaaattcgttacccaccgtttcccactcgaagacgcggttaaggctttcgaaacctctgcggacccgaaatctggcgct
atcaaagttatgatccagtctctggac

A.2. Oligos for codon/codon-pair optimized gene assembly

A.2.1. *clad1*

Seq Name	Sequence
2454 On1 fwd (Pf1-1)	ATGTCTCCGTCTGCTGTTGACGACGCGCCAAAGGCTACCGGTG
2454 On 2 rev (Pf1-2)	TTCGGTTTAACAGAGATTGCAGCACCGGTAGCCTTTGGCGCGTCG
2454 On 3 fwd (Pf1-3)	CTGCAATCTCTGTTAAACCGAACATCGGCGTTTTCACCAACCCGA
2454 On 4 rev (Pf1-4)	TCAGAGATCCACAGGTCGTGTTTCGGGTTGGTGAAAACGCCGATG
2454 On 5 fwd (Pf1-5)	AACACGACCTGTGGATCTCTGAAGCTGAGCCATCTGCGGACGCTG
2454 On 6 rev (Pf1-6)	TCAGGTCCGCACCAGATTTAACAGCGTCCGCAGATGGCTCAGCT
2454 On 7 fwd (Pf1-7)	TTAAATCTGGTGCGGACCTGAAACCGGGTGAAGTTACCATCGCG
2454 On 8 rev (Pf1-8)	CGCAAATACCGGTAGAACGAACCGCGATGGTAACTTCACCCGGTT
2454 On 9 fwd (Pf1-9 Pf2-1)	GTTCGTTCTACCGGTATTTGCGGTTCTGATGTTCACTTTTGGCAC
2454 On 10 rev (Pf1-10 Pf2-2)	CATCGGACCGATACAACCCGCGTGCCAAAAGTGAACATCAGAAC
2454 On 9 fwd (Pf1-9 Pf2-1)	GTTCGTTCTACCGGTATTTGCGGTTCTGATGTTCACTTTTGGCAC
2454 On 10 rev (Pf1-10 Pf2-2)	CATCGGACCGATACAACCCGCGTGCCAAAAGTGAACATCAGAAC
2454 On 11 fwd (Pf2-3)	GCGGGTTGTATCGGTCCGATGATCGTTGAAGGCGACCACATCCT
2454 On 12 rev (Pf2-4)	TTCACCAGCAGATTCGTGACCCAGGATGTGGTCGCCTTCAACGAT
2454 On 13 fwd (Pf2-5)	GGGTCACGAATCTGCTGGTGAAGTTATCGCTGTTCACCCGACC
2454 On 14 rev (Pf2-6)	GTCACCGATCTGGAGAGAGCTAACGGTCGGGTGAACAGCGATAAC
2454 On 15 fwd (Pf2-7)	GTTAGCTCTCTCCAGATCGGTGACCGTGTTGCGATCGAACCGAAC
2454 On 16 rev (Pf2-8)	TGGTTCGCAAGCGTTGCAGATGATGTTCGGTTCGATCGCAACACG
2454 On 17 fwd (Pf2-9 Pf3-1)	ATCATCTGCAACGCTTGCGAACCATGCCTGACCGGTCGTTATAA
2454 On 18 rev (Pf2-10 Pf3-2)	CAGAAATTCAACTTTTTCGCAGCCGTTATAACGACCGGTCAGGCA
2454 On 17 fwd (Pf2-9 Pf3-1)	ATCATCTGCAACGCTTGCGAACCATGCCTGACCGGTCGTTATAA
2454 On 18 rev (Pf2-10 Pf3-2)	CAGAAATTCAACTTTTTCGCAGCCGTTATAACGACCGGTCAGGCA
2454 On 19 fwd (Pf3-3)	CGGCTGCGAAAAAGTTGAATTTCTGTCTACCCCACCTGTTCCGGG
2454 On 20 rev (Pf3-4)	TTAACGTAACGACGCAGGAGGCCCGGAACAGGTGGGGTAGA
2454 On 21 fwd (Pf3-5)	CCTCCTGCGTCGTTACGTTAATCACCCGGCGGTTTGGTGC
2454 On 22 rev (Pf3-6)	TCCCAAGACATGTTACCGATTTTGTGGCACCAAACCGCCGGGTGA
2454 On 23 fwd (Pf3-7)	CACAAAATCGGTAACATGTCTTGGGAAAATGGTGCTCTGCTCGAA
2454 On 24 rev (Pf3-8)	TACCAGCCAGAGCAACAGACAGCGGTTCGAGCAGAGCACCATTT
2454 On 25 fwd (Pf3-9 Pf4-1)	CCGCTGTCTGTTGCTCTGGCTGGTATGCAACGTGCGAAAGTTCAG
2454 On 26 rev (Pf3-10 Pf4-2)	GCAAACCAGAACCGGGTCGCCCAGCTGAACTTTCGCACGTTGCA
2454 On 25 fwd (Pf3-9Pf4-1)	CCGCTGTCTGTTGCTCTGGCTGGTATGCAACGTGCGAAAGTTCAG
2454 On 26 rev (Pf3-10 Pf4-2)	GCAAACCAGAACCGGGTCGCCCAGCTGAACTTTCGCACGTTGCA
2454 On 27 fwd (Pf4-3)	CTGGGCGACCCGGTTCTGGTTTGCGGTGCTGGCCCAATCGGTCTG
2454 On 28 rev (Pf4-4)	AGCAGCAGCACACAGCATCGAAACCAGACCGATTGGGCCAGCACC
2454 On 29 fwd (Pf4-5)	GTTTCGATGCTGTGTGCTGCTGCTGCGGGTGCGTGTCCACTGGTT
2454 On 30 rev (Pf4-6)	ACGAGATTCAGAGATGTCAGTGATAACCAGTGGACACGCACCCGC

2454 On31 fwd (Pf4-7)	ATCACTGACATCTCTGAATCTCGTCTCGCGTTTGCGAAAGAAATC
2454 On32 rev (Pf4-8)	GTGGGTAGTCACACGTGGACAGATTTCTTTCGCAAACGCGAG
2454 On33 fwd (Pf4-9 Pf5-1)	TGTCCACGTGTGACTACCCACCGTATCGAAATCGGTAAATCTGC
2454 On34 rev (Pf4-10 Pf5-2)	GATAGATTTCGCGGTTTCCTCAGCAGATTTACCGATTTCGATACG
2454 On33 fwd (Pf4-9 Pf5-1)	TGTCCACGTGTGACTACCCACCGTATCGAAATCGGTAAATCTGC
2454 On34 rev (Pf4-10 Pf5-2)	GATAGATTTCGCGGTTTCCTCAGCAGATTTACCGATTTCGATACG
2454 On35 fwd (Pf5-3)	TGAGGAAACCGCGAAATCTATCGTTAGCTCTTTCGGTGGCGTT
2454 On36 rev (Pf5-4)	ATTCGAGGGTAACAGCTGGTTCAACGCCACCGAAAGAGCTAAC
2454 On37 fwd (Pf5-5)	GAACCAGCTGTTACCCTCGAATGTACCGGTGTTGAATCTTCTATC
2454 On38 rev (Pf5-6)	TAGACGCCCAAATAGCAGCCGCGATAGAAGATTCAACACCGGTAC
2454 On39 fwd (Pf5-7)	GCGGCTGCTATTTGGGCGTCTAAATTCGGTGGTAAAGTTTTCGTT
2454 On40 rev (Pf5-8)	TTTCGTTTTTACCAACACCGATAACGAAAACTTTACCACCGAATT
2454 On41 fwd (Pf5-9 Pf6-1)	ATCGGTGTTGGTAAAAACGAAATCTCTATCCCGTTCATGCGTGC
2454 On42 rev (Pf5-10 Pf6-2)	GATGTCAACTTCACGAACAGACGCACGCATGAACGGGATAGAGA
2454 On41 fwd (Pf5-9 Pf6-1)	ATCGGTGTTGGTAAAAACGAAATCTCTATCCCGTTCATGCGTGC
2454 On42 rev (Pf5-10 Pf6-2)	GATGTCAACTTCACGAACAGACGCACGCATGAACGGGATAGAGA
2454 On43 fwd (Pf6-3)	GTCTGTTCGTGAAGTTGACATCCAGCTCCAGTACCGTTACTCTAA
2454 On44 rev (Pf6-4)	ACGGATAGCACGCGGCCAAGTGTTAGAGTAACGGTACTGGAGCTG
2454 On45 fwd (Pf6-5)	CACTTGGCCGCGTGCTATCCGTCTGATCGAATCTGGTGTTATTGA
2454 On46 rev (Pf6-6)	GTGGGTAACGAATTTAGACAGGTCAATAACACCAGATTCGATCAG
2454 On47 fwd (Pf6-7)	CCTGTCTAAATTCGTTACCCACCGTTTCCCACTCGAAGACGCG
2454 On48 rev (Pf6-8)	CGCAGAGGTTTCGAAAGCCTTAACCGCGTCTTCGAGTGGGAAACG
2454 On49 fwd (Pf6-9)	GTTAAGGCTTTCGAAACCTCTGCGGACCCGAAATCTGGCGCTAT
2454 On50 rev (Pf6-10)	GTCCAGAGACTGGATCATAACTTTGATAGCGCCAGATTTCGGGTC
2454 On1 fwd (Pf1-1)	ATGTCTCCGTCTGCTGTTGACGACGCGCCAAAGGCTACCGGTG
2454 On10 rev (Pf1-10 Pf2-2)	CATCGGACCGATACAACCCGCGTGCCAAAAGTGAACATCAGAAC
2454 On9 fwd (Pf1-9 Pf2-1)	GTTCGTTCTACCGGTATTTGCGGTTCTGATGTTCACTTTTGGCAC
2454 On18 rev (Pf2-10 Pf3-2)	CAGAAATTCAACTTTTTCGCAGCCGTTATAACGACCGGTCAGGCA
2454 On17 fwd (Pf2-9 Pf3-1)	ATCATCTGCAACGCTTGCGAACCATGCCTGACCGGTCGTTATAA
2454 On26 rev (Pf3-10 Pf4-2)	GCAAACCAGAACCGGGTCGCCCAGCTGAACTTTCGCACGTTGCA
2454 On25 fwd (Pf3-9 Pf4-1)	CCGCTGTCTGTTGCTCTGGCTGGTATGCAACGTGCGAAAGTTCAG
2454 On34 rev (Pf4-10 Pf5-2)	GATAGATTTCGCGGTTTCCTCAGCAGATTTACCGATTTCGATACG
2454 On33 fwd (Pf4-9 Pf5-1)	TGTCCACGTGTGACTACCCACCGTATCGAAATCGGTAAATCTGC
2454 On42 rev (Pf5-10 Pf6-2)	GATGTCAACTTCACGAACAGACGCACGCATGAACGGGATAGAGA
2454 On41 fwd (Pf5-9 Pf6-1)	ATCGGTGTTGGTAAAAACGAAATCTCTATCCCGTTCATGCGTGC

2454 On50 rev (Pf6-10)	GTCCAGAGACTGGATCATAACTTTGATAGCGCCAGATTTCGGGTC
2454 On52 fwd (Tf1-5P2)	GAATACTCCATATGATGTCTCCGTCTGCTGTTGACGACGCGC
2454 On56 rev (Tf1-3P2)	AGGCATCGACTAGTGTCCAGAGACTGGATCATAACTTTGATA

A.2.2. *clxr1*

Seq Name	Sequence
2473 Onl fwd (Pf1-1)	ATGCCACAACCTGTTCCAACTGCTAATAGATTGTTAGA
2473 On2 rev (Pf1-2)	ACCTTTTAAAGAAAACAAATCTAACAATCTATTAGCAG
2473 On3 fwd (Pf1-3)	TTTGTTTTCTTTAAAAGGTAAAGTAGTTGTTGTTACTG
2473 On4 rev (Pf1-4)	CCTCTAGGACCAGAAGCACCAGTAACAACAACTACTTT
2473 On5 fwd (Pf1-5)	GTGCTTCTGGTCCTAGAGGTATGGGAATTGAAGCTGCT
2473 On6 rev (Pf1-6)	CCCATTTCAGCACATCCTCTAGCAGCTTCAATTCCCATA
2473 On7 fwd (Pf1-7 Pf2-1)	AGAGGATGTGCTGAAATGGGAGCAGATTTAGCTATTACTT
2473 On8 rev (Pf1-8 Pf2-2)	TCCTTCCTTTCTAGAAGAATAAGTAATAGCTAAATCTGCT
2473 On7 fwd (Pf1-7 Pf2-1)	AGAGGATGTGCTGAAATGGGAGCAGATTTAGCTATTACTT
2473 On8 rev (Pf1-8 Pf2-2)	TCCTTCCTTTCTAGAAGAATAAGTAATAGCTAAATCTGCT
2473 On9 fwd (Pf2-3)	ATTCTTCTAGAAAGGAAGGAGCTGAGAAAAATGCTGAGG
2473 On10 rev (Pf2-4)	CCATATTCTTTAGTCAATTCCTCAGCATTTTTCTCAGC
2473 On11 fwd (Pf2-5)	AATTGACTAAAGAATATGGTGTTAAAGTAAAAGTATATAAGGT
2473 On12 rev (Pf2-6)	TCATTATAATCAGATTGATTAACCTTATATACTTTTACTTTAACA
2473 On13 fwd (Pf2-7)	TAATCAATCTGATTATAATGATGTTGAAAGATTTGTAAAT
2473 On14 rev (Pf2-8)	CCGAAATCTGAAACAACTTGATTTACAAATCTTTCAACA
2473 On15 fwd (Pf2-9 Pf3-1)	CAAGTTGTTTCAGATTTCGGTAAAATTGATGCTTTTATTG
2473 On16 rev (Pf2-10 Pf3-2)	AGCGGTAGCTCCAGCGTTAGCAATAAAAGCATCAATTTTA
2473 On15 fwd (Pf2-9 Pf3-1)	CAAGTTGTTTCAGATTTCGGTAAAATTGATGCTTTTATTG
2473 On16 rev (Pf2-10 Pf3-2)	AGCGGTAGCTCCAGCGTTAGCAATAAAAGCATCAATTTTA
2473 On17 fwd (Pf3-3)	CTAACGCTGGAGCTACCGCTAATTCTGGTGTTGTTGATG
2473 On18 rev (Pf3-4)	TCCCAATCTGAAGCAGAACCATCAACAACACCAGAATT
2473 On19 fwd (Pf3-5)	GTTCTGCTTCAGATMGGATCATGTTATACAAGTTGAT
2473 On20 rev (Pf3-6)	AATAAGCAGTTCCAGACAAATCAACTTGTATAACATGA
2473 On21 fwd (Pf3-7)	TTGTCTGGAACTGCTTATTGTGCAAAGGCTGTAGGTGCT
2473 On22 rev (Pf3-8)	GACCTTGCTTTTTAAAATGAGCACCTACAGCCTTTGCAC
2473 On23 fwd (Pf3-9 Pf4-1)	CATTTTAAAAAGCAAGGTCATGGTTCTTTGGTTATAAC
2473 On24 rev (Pf3-10 Pf4-2)	ATGACCAGACATAGAAGCTGTTATAACCAAAGAACCAT
2473 On23 fwd (Pf3-9 Pf4-1)	CATTTTAAAAAGCAAGGTCATGGTTCTTTGGTTATAAC
2473 On24 rev (Pf3-10 Pf4-2)	ATGACCAGACATAGAAGCTGTTATAACCAAAGAACCAT
2473 On25 fwd (Pf4-3)	AGCTTCTATGTCTGGTCATGTTGCTAATTATCCACAAG
2473 On26 rev (Pf4-4)	AACATTATATGAAGTTTGTTCTTGTGGATAATTAGCAAC
2473 On27 fwd (Pf4-5)	AACAAACTTCATATAATGTTGCAAAAGCTGGTTGTATTC
2473 On28 rev (Pf4-6)	GCCAAAGATCTAGCCAAATGAATACAACCAGCTTTTGC

2473 On29 fwd (Pf4-7)	ATTTGGCTAGATCTTTGGCTAATGAATGGAGAGATTTT
2473 On30 rev (Pf4-8)	GAAATAGAATTAACTCTAGCAAAATCTCTCCATTCATTA
2473 On31 fwd (Pf4-9 Pf5-1)	GCTAGAGTTAATTCTATTTCACCAGGTTATATAGATACCG
2473 On32 rev (Pf4-10 Pf5-2)	ATCAATAAAATCAGACAAACCGGTATCTATATAACCTGGT
2473 On31 fwd (Pf4-9 Pf5-1)	GCTAGAGTTAATTCTATTTCACCAGGTTATATAGATACCG
2473 On32 rev (Pf4-10 Pf5-2)	ATCAATAAAATCAGACAAACCGGTATCTATATAACCTGGT
2473 On33 fwd (Pf5-3)	GTTTGTCTGATTTTATTGATGAAAAGACTCAAGAATTGT
2473 On34 rev (Pf5-4)	ATTGGAATCATAGATCTCCACAATTCTTGAGTCTTTTC
2473 On35 fwd (Pf5-5)	GGAGATCTATGATTCCAATGGGTAGAAATGGTGATGCT
2473 On36 rev (Pf5-6)	AAGCACCTTTTAATTCTTTAGCATCACCATTTCTACCC
2473 On37 fwd (Pf5-7)	AAAGAATTAAAAGGTGCTTATGTTTATTIGGTTTCAGA
2473 On38 rev (Pf5-8)	AGTAGTATAAGAAGAAGCATCTGAAACCAAATAAACAT
2473 On39 fwd (Pf5-9)	TGCTTCTTCTTATACTACTGGTGCTGATATAGTAATTGA
2473 On40 rev (Pf5-10)	TCTAGTAGTATAACCACCATCAATTACTATATCAGCACC
2473 On1 fwd (Pf1-1)	ATGCCACAACCTGTTCCAACTGCTAATAGATTGTTAGA
2473 On8 rev (Pf1-8 Pf2-2)	TCCTTCCTTTCTAGAAGAATAAGTAATAGCTAAATCTGCT
2473 On7 fwd (Pf1-7 Pf2-1)	AGAGGATGTGCTGAAATGGGAGCAGATTTAGCTATTACTT
2473 On16 rev (Pf2-10 Pf3-2)	AGCGGTAGCTCCAGCGTTAGCAATAAAAGCATCAATTTTA
2473 On15 fwd (Pf2-9 Pf3-1)	CAAGTTGTTTCAGATTTCGGTAAAATTGATGCTTTTATTG
2473 On24 rev (Pf3-10 Pf4-2)	ATGACCAGACATAGAAGCTGTTATAACCAAAGAACCAT
2473 On23 fwd (Pf3-9 Pf4-1)	CATTTTAAAAAGCAAGGTCATGGTTCTTTGGTTATAAC
2473 On32 rev (Pf4-10 Pf5-2)	ATCAATAAAATCAGACAAACCGGTATCTATATAACCTGGT
2473 On31 fwd (Pf4-9 Pf5-1)	GCTAGAGTTAATTCTATTTCACCAGGTTATATAGATACCG
2473 On40 rev (Pf5-10)	TCTAGTAGTATAACCACCATCAATTACTATATCAGCACC
2473 On42 fwd (Tf1-5P2)	GAATACTCCATATGATGCCACAACCTGTTCCAACTGCTAATA
2473 On45 rev (Tf1-3P2)	AGGCATCGACTAGTTCTAGTAGTATAACCACCATCAATTACT

A.2.3. *cALX1*

Seq Name	Sequence
2470 On1 fwd (Pf1-1)	ATGACTGATTATATTCCAACTTTTAGATTTGATGGTCA
2470 On2 rev (Pf1-2)	ACCTGTAACAATAGTCAAATGACCATCAAATCTAAAAG
2470 On3 fwd (Pf1-3)	TTTGACTATTGTTACAGGTGCATGTGGAGGTTTGGCTG
2470 On4 rev (Pf1-4)	AAACCCTTAATTAGAGCTTCAGCCAAACCTCCACATGC
2470 On5 fwd (Pf1-5)	AAGCTCTAATTAAGGGTTTGTTGGCTTATGGTTCTGAT
2470 On6 rev (Pf1-6)	CTATATCCAATAGAGCTATATCAGAACCATAAGCCAAC
2470 On7 fwd (Pf1-7 Pf2-1)	ATAGCTCTATTGGATATAGATCAAGAAAAGACTGCTGC
2470 On8 rev (Pf1-8 Pf2-2)	ATGATATTCAGCTTGTTTAGCAGCAGTCTTTTCTTGAT

2470 On7 fwd (Pf1-7 Pf2-1)	ATAGCTCTATTGGATATAGATCAAGAAAAGACTGCTGC
2470 On8 rev (Pf1-8 Pf2-2)	ATGATATTCAGCTTGTTTAGCAGCAGTCTTTTCTTGAT
2470 On9 fwd (Pf2-3)	TAAACAAGCTGAATATCATAAATATGCTACAGAAGAAT
2470 On10 rev (Pf2-4)	GGAACTTCTTTTAATTTTAATTCTTCTGTAGCATATTT
2470 On11 fwd (Pf2-5)	TAAAATTAAAAGAAGTTCCAAAAATGGGTTCTTATGCT
2470 On12 rev (Pf2-6)	CAGAATCAGAAATGTCACAAGCATAAGAACCCATTTTT
2470 On13 fwd (Pf2-7)	TGTGACATTTCTGATTCTGATACTGTTCATAAGGTTTTT
2470 On14 rev (Pf2-8)	AAATCTTTAGCAACTTGAGCAAAAACCTTATGAACAGTAT
2470 On15 fwd (Pf2-9 Pf3-1)	GCTCAAGTTGCTAAAGATTTTGGAAAATTGCCATTGCATT
2470 On16 rev (Pf2-10 Pf3-2)	ATAACCAGCGGTATTAACCAAATGCAATGGCAATTTTCCA
2470 On15 fwd (Pf2-9 Pf3-1)	GCTCAAGTTGCTAAAGATTTTGGAAAATTGCCATTGCATT
2470 On16 rev (Pf2-10 Pf3-2)	ATAACCAGCGGTATTAACCAAATGCAATGGCAATTTTCCA
2470 On17 fwd (Pf3-3)	TGGTTAATACCGCTGGTTATTGTGAAAATTTCCCATGTG
2470 On18 rev (Pf3-4)	TTTTTAGCTGGATAATCTTCACATGGGAAATTTTCACA
2470 On19 fwd (Pf3-5)	AAGATTATCCAGCTAAAAATGCAGAGAAAATGGTTAAG
2470 On20 rev (Pf3-6)	AAGAACCCAACAAATTAACCTTAACCATTTTCTCTGCA
2470 On21 fwd (Pf3-7)	GTTAATTTGTTGGGTTCTTTGTATGTTTCACAAGCATTT
2470 On22 rev (Pf3-8)	TCTTTAATCAATGGTTTAGCAAATGCTTGTGAAACATACA
2470 On23 fwd (Pf3-9)	GCTAAACCATTGATTAAAGAAGGAATTAAAGGTGCTTCAG
2470 On24 rev (Pf3-10)	CATAGAACCAATAAGTACAACTGAAGCACCTTTAATTCCT
2470 On23 fwd (Pf3-9 Pf4-1)	GCTAAACCATTGATTAAAGAAGGAATTAAAGGTGCTTCAG
2470 On24 rev (Pf3-10 Pf4-2)	CATAGAACCAATAAGTACAACTGAAGCACCTTTAATTCCT
2470 On25 fwd (Pf4-3)	TTGTACTTATTGGTTCTATGTCTGGAGCTATTGTTAATGATC
2470 On26 rev (Pf4-4)	TTATAAACAACTTGATTTTGTGGATCATTAACAATAGCTCCAGA
2470 On27 fwd (Pf4-5)	CACAAAATCAAGTTGTTTATAATATGTCTAAAGCTGGTGTAATT
2470 On28 rev (Pf4-6)	AGCAAGAGTTTTTGCCAAATGAATTACACCAGCTTTAGACATA
2470 On29 fwd (Pf4-7)	CATTTGGCAAAAACTCTTGCTTGTGAATGGGCTAAATATAACATT
2470 On30 rev (Pf4-8)	TGGATTTAAAGAATTAACTCTAATGTTATATTTAGCCCATTCACA
2470 On31 fwd (Pf4-9 Pf5-1)	AGAGTTAATTCTTTAAATCCAGGTTATATTTATGGTCCATT
2470 On32 rev (Pf4-10 Pf5-2)	ACCATTAATTACATTTTTAGTCAATGGACCATAAATATAACC
2470 On31 fwd (Pf4-9 Pf5-1)	AGAGTTAATTCTTTAAATCCAGGTTATATTTATGGTCCATT
2470 On32 rev (Pf4-10 Pf5-2)	ACCATTAATTACATTTTTAGTCAATGGACCATAAATATAACC
2470 On33 fwd (Pf5-3)	GACTAAAAATGTAATTAATGGTAATGAAGAATTGTATAATAGATG
2470 On34 rev (Pf5-4)	TTGTGGTATTCCAGAAATCCATCTATTATACAATTCTTCATT
2470 On35 fwd (Pf5-5)	GATTTCTGGAATACCACAACAAAGAATGTCTGAACCTA
2470 On36 rev (Pf5-6)	ACAGCACCAATATATTCTTTAGGTTCAGACATTCTTTG
2470 On37 fwd (Pf5-7)	AAGAATATATTGGTGCTGTTTTGTATTTGTTGTCTGAG
2470 On38 rev (Pf5-8)	GTAGTATAAGAAGCAGCAGACTCAGACAACAAATACAAA
2470 On39 fwd (Pf5-9)	TCTGCTGCTTCTTATACTACAGGTGCATCATTGTTGGTTG
2470 On40 rev (Pf5-10)	CCAAGAAGTAAAACCACCATCAACCAACAATGATGCACCT
2470 On1 fwd (Pf 1-1)	ATGACTGATTATATTCCAACTTTTAGATTTGATGGTCA
2470 On8 rev (Pf1-8 Pf2-2)	ATGATATTCAGCTTGTTTAGCAGCAGTCTTTTCTTGAT
2470 On7 fwd (Pf 1-7 Pf2-1)	ATAGCTCTATTGGATATAGATCAAGAAAAGACTGCTGC

2470 On16 rev (Pf2-10 Pf3-2)	ATAACCAGCGGTATTAACCAAATGCAATGGCAATTTTCCA
2470 On15 fwd (Pf2-9 Pf 3-1)	GCTCAAGTTGCTAAAGATTTTGGAAAATTGCCATTGCATT
2470 On24 rev (Pf3-10 Pf4-2)	CATAGAACCAATAAGTACAACTGAAGCACCTTTAATTCCT
2470 On23 fwd (Pf3-9 Pf 4-1)	GCTAAACCATTGATTAAAGAAGGAATTAAAGGTGCTTCAG
2470 On32 rev (Pf4-10 Pf5-2)	ACCATTAATTACATTTTTAGTCAATGGACCATAAATATAACC
2470 On31 fwd (Pf4-9 Pf5-1)	AGAGTTAATTCTTTAAATCCAGGTTATATTTATGGTCCATT
2470 On40 rev (Pf5-10)	CCAAGAAGTAAAACCACCATCAACCAACAATGATGCACCT
2470 On42 fwd (Tf1-5P2)	GAATACTCCATATGATGACTGATTATATTCCAACTTTTAGAT
2470 On45 rev (Tf1-3P2)	AGGCATCGACTAGTCCAAGAAGTAAAACCACCATCAACCAAC

A.3. Gene assembly procedure

1. Oligo mix preparation and PCR assembly

i. Resuspend each Pf (Primary fragment) oligos in 100 ul water.

ii. Mix 5 ul Pf oligos with water to prepare total 100 ul for each Pf.

iii. Prepare 5’3’ oligo mix (the first and last oligo in each Pf group. E.g. clad1, Pf1: Pf1-1 and Pf1-10). Mix 5 ul each oligo in total 50 ul water.

iv. PCR assembly

Oligo mix	4 ul
5’3’ mix	1 ul
dNTPs (10mM)	1 ul
PfuUltra II HS DNA Polymerase	1 ul
10x Pfu Buffer	5 ul
H2O	38 ul

Step 1: 95 °C / 10 min
Step 2: 95 °C / 20 sec
Step 3: 62 °C / 30 sec
Step 4: 72 °C / 15 sec
Step 5: GOTO Step 2, 30X
Step 6: 72 °C / 5 min
Step 7: 4 °C / Hold

v. PCR purification

2. Full length DNA Fragment assembly

i.Prepare full length 5’3’ oligo mix (Tf in the oligo table). Mix 5 ul oligos in 50 ul water.

ii. PCR assembly

Previously constructed Pf mix (final 2 nM)	~ ul
5’3’ Tf oligo mix	1 ul
dNTPs (10mM)	1 ul
PfuUltra II HS DNA Polymerase	1 ul
10x Pfu Buffer	5 ul
H2O	up to total 50 ul

Step 1: 95 °C / 10 min
Step 2: 95 °C / 20 sec
Step 3: 62 °C / 30 sec
Step 4: 72 °C / 15 sec
Step 5: GOTO Step 2, 30X
Step 6: 72 °C / 5 min
Step 7: 4 °C / Hold

iii. PCR purification

3. Purified PCR product is cloned (NdeI/SpeI) into pCODA-OUT vector.

4. Sequence the product.

B. Molecular biology protocols

B.1. Gibson Reaction (Gibson, 2009)

Reagents:

1. 5X isothermal (ISO) reaction buffer (25% PEG-8000, 500 mM Tris-HCl pH 7.5, 50 mM MgCl2, 50 mM DTT, 1 mM each of the 4 dNTPs, and 5 mM NAD). This is prepared as described below.
2. T5 exonuclease (Epicentre)
3. Phusion DNA polymerase (New England Biolabs)
4. Taq DNA ligase (New England Biolabs)

Procedure:

1. Prepare 5X ISO buffer. Six ml of this buffer can be prepared by combining the following:

   3 ml of 1 M Tris-HCl pH 7.5
   150 µl of 2 M $MgCl_2$
   60 µl of 100 mM dGTP
   60 µl of 100 mM dATP
   60 µl of 100 mM dTTP
   60 µl of 100 mM dCTP
   300 µl of 1 M DTT
   1.5 g PEG-8000
   300 µl of 100 mM NAD
   Add water to 6 ml
   Aliquot 100 µl and store at -20 °C

2. Prepare an assembly master mixture as following.

   320 µl 5X ISO buffer
   0.64 µl of 10 U/ µl T5 exo
   20 µl of 2 U/µl Phusion pol
   160 µl of 40 U/µl Taq lig
   Add water to 1.2 ml
   Aliquot 15 µl and store at -20 °C. This assembly mixture can be stored at -20 °C for at least one year. The enzymes remain active following at least 10 freeze-thaw cycles. This is ideal for the assembly of DNA molecules with 20-150 bp overlaps.
   For DNA molecules overlapping by larger than 150 bp, prepare the assembly mixture by using 3.2 µl of 10 U/ µl T5 exo.

3. Thaw a 15 µl assembly mixture aliquot and keep on ice until ready to be used.

4. Add 5 µl of DNA to be assembled to the master mixture. The DNA should be in equimolar amounts. Use 10-100 ng of each ~6 kb DNA fragment. For larger DNA segments, increasingly proportionate amounts of DNA should be added (e.g. 250 ng of each 150 kb DNA segment).

5. Incubate at 50 °C for 15 to 60 min (60 min is optimal).

6. Transform into competent *E. coli* cells.

Appendix C. Primer sequences

Table C.1 List of primers

Primer	sequence (5’ – 3’)
Chapter 3	
pJC13	acatcggaccgaaaaaaatgaattcaactcccgatc
pJC14	acctttttccggtgtttggagggg
pJC15	acatcggaccgaggcctaagaagcccatgtctgaatatg
pJC16	aattctcgagcctaggggattcaggaactaacgtc
pJC17	acaacctaggtatttggttgaagctggcag
pJC18	gagactcgaggtttaaacatggctggtttaccag
pJC19	caacgtttgttgatgggtgtgtttgttc
pJC20	attgctcgagaggcctttagtttctggaac
pJC21	ggcccaggcctatcattagaagaagttaatg
pJC22	cgggctcgagttatttcctgctaaacaaac
pJC23	atatcggaccgaaaaaaatggcagttgaggagaacaatatg
pJC24	ggttctcgagttattctagcatggccttgtac
012209_pXP_Cyct_For_SpeI	atcagactagtgtttaaaccggaccgctcgagtcatgtaattagt
012209_pXP_Cyct_Rev_BsrG	gttacatgcgtacacgcgtctg
Chapter 4	
012209_lovB_for	acaatgcggccgcgtttaaacaaaaaaatggctcaatctatgtatcc
012209_lovB_rev_Bsu36I	gggtagcttctgtcggaatcctcagg
LovC_For_SpeI	cggcactagtaaaaaaatgggcgaccagccattc
LovC_Rev_RsrII	atatcggtccgttacggcccctcgagcc
hpm3TE_for_AvrII_Kozak	gccacctaggaaaaaaatggtcgtagattatccc
hpm3TE_rev_noHistag_PmeI	cccggtttaaacttaatttgcctcgctg
Hpm3TE_rev_PmeI	atctgtttaaacttagtggtggtggtgg
PKS13TEFo2	cccgcactagtaaaaaaatgatcatagactatccggc
PKS13TERev	taaacggtccgttaccccgcctcgttaaag
PKS13TE-His_Rev	taaacggtccgttaatgatgatgatgatgccccgcctcgttaaag
AptB_SpeI_F	gggcactagtaaaaaaatggccttcagaataccatttgcccag
AptB_RsrII_R	ttaaatcggtccgttaattcggccggcgcgac
AptB-His_Rev	ttaaatcggtccgttaatgatgatgatgatgattcggccggcgcgac
VrtG_SpeI_F	cgcgactagtaaaaaaatggccacacgaattccttttc
VrtG_XhoI_R	actactcgagctactgctgcgttac
VrtG-His_Rev	taatctcgagctaatgatgatgatgatgctgctgcgttaccagtcc
RadicicolTE_For	cgccactagtaaaaaaatgcagcaacccaggtctattg
RadicicolTE_Rev	caccctcgagtcatcgtccaaagtgctcaaac
Rdc1TE-His_Rev	atatctcgagtcaatgatgatgatgatgatgtcgtccaaagtgctcaaacg
Chapter 5	
ScAcs1_for_AvrII_Kozak	atacctaggaaaaaaatgtcgccctctgccgtac
ScAcs1_rev_PmeI	gcggcgtttaaacttacaacttgaccgaatc
Cab1_for_AvrII_Kozak	ttaacctaggaaaaaaatgccgcgaattactcaa
Cab1_rev_PmeI_v2	cggcggtttaaactacgtacttgttttcttag
SEAcs_for_SpeI_Kozak	ggccgggcactagtaaaaaaatgagccaaacacataaac
SEAcs_rev_XhoI	aactctcgagttatgacggcatcgcgatg
ACC1-For-KOZAK-RsrII	agagcggaccgaaaaaaatgagcgaagaaagcttattc
ACC1-Rev-SpeI	accgactagtttccccagtctgac

MET17FrFor	ctccatatgtcgcgatttttctcttgaggtcacatgatcgc
MET17FrRev	aatcatatgcctgaggtgccaaccaccacagttccccaatatc
MET17ReFor	cacggtacccagatatagtcggattgccctttaagc
MET17ReRev	cccgaattcatttaaatgcggccgccaagttaacatcttatag
MET17FrFor2	ctctcgcgacctaggtttttctcttgaggtcacatgatcgc
Ura3_FH_For2	gcaccatatgtcgcgacctaggctcatcaataaaatcgaaattcc
Ura3_FH_Rev	gaggcatatgtatggaccctgaaaccacagccac
Ura3_RH_For	cctaggtaccccgggaatctcggtcgtaatg
Ura3_RH_Rev2	aattgaattcatttaaatgcggccgcttccagcccatatccaacttc
pBF55	ctagggaagacaagcaacg
PHXT7F2	acatcatatgccgtggaaatgaggggtatgcaggaatttgtgc
PHXT7R2	gtgcgaccgaccctcggactagtttttgattaaaattaaaaaaactttttgtttttgtg
ADH2F	cgccactagtatgtctattccagaaactca
ADH2R	ggccctcgagttatttagaagtgtcaacaacg
ACC1_S1157A_SDM_F	gggtatgaacagggctgttgctgtttcagatttgtc
ACC1_S1157A_SDM_R	gacaaatctgaaacagcaacagccctgttcataccc
ACC14361For	cgccgcagaaatgtacaccgaagtcaagaacgc
ACC1HisRev	gcgcctcgagttagtggtggtggtggtggtttcaaagtcttcaacaatttttc
PGK1F	tttttgctcccagcgcgtttgctcccctcgagggtaccattaggcatttgcaagaattac
NpgAR	ggggagggcgtgaatgtaagcgtgacataactaattacatgactcgagtcgacttaggataggcaattac

Chapter 6

6MSAS_Int	ctgctgtgtcactcccagac
PPKex2F	ctagaggatccccgggataacttcgtatagcatacattatacgaagttatgtagatacacgtatctcgacatg
PPKex2R	ctgataatgggttagtagtttataattatgtgacgaggccaaaaaacaacggg
UBI-RF	ctactaacccattatcagaaaacaatgcagattttcgtcaagactttgaccgg
UBI-RR	gtcgaccaagcttccctctaccacctcttagccttagcacaagatg
URA3F	agagggaagcttggtcgacaaccggtcgccaccatgtcgaaagctacatataag
URA3R	ggtgggaattcgagctcggtacccgggtaataactg
TEF1F	atcagagcagattgtactgagagtgcaccatatgatttagcggccgcaccgcgaatccttacatcac
TEF1R	ggggatgttttcccagaggggtatgtagaagttgcagcggaatgcatttttttactttgtaattaaaacttag
Pfur1delF	ccccgccacaaactatttttttgaagacatgctttctcatgactgcctaataaccgactctagaggatccccggg
Pfur1delR	ttgatatgcggcttcagccgttctgaaccttcaagatggtgttcgggtgtggaattcgagctcggtaccc

Primer sequences for targeted integration into genomic DNA (Fang et al., 2011)

URA3-marker	ccaatttttttttttttcgtcattatagaaatcattacgaccgagattcccgggaattcgagctcggtacccggg
URA3 locus-PADH1-For	gttttgaccatcaaagaaggttaatgtggctgtggtttcagggtccataaggggatcgaagaaatgatg
YDRWTy1-5-PGK_F	cacagagttgtatttgcgcttctgagcgatgcttccgagattgttgaagcaaggcatttgcaagaattactcgtg
YDRWTy1-5-marker	gattattgaagagggatgcgtttggtacaataaaaaacataggttcccaaaccgaattcgagctcggtacccggg
YDRWTy1-5-ADH2F	cacagagttgtatttgcgcttctgagcgatgcttccgagattgttgaagcagcaaaacgtaggggcaaacaaac
LEU2-ADH2-For	ccatgtataatcttcattattacagccctcttgacctctaatcatgaatgttaaaacgtaggggcaaacaaac
LEU2-PGK1-For	ccatgtataatcttcattattacagccctcttgacctctaatcatgaatgttaggcatttgcaagaattactcgtg
LEU2-LoxP-Rev	gcgtatatagtttcgtctaccctatgaacatattccattttgtaatttcgtgtcgaattcgagctcggtacccggg
HIS3-ADH2-For	gctttgctgtgggaaaaacttatcgaaagatgacgacttttttcttaattctaaaacgtaggggcaaacaaac
HIS3-PGK1-For	gctttgctgtgggaaaaacttatcgaaagatgacgacttttttcttaattctaggcatttgcaagaattactcgtg
HIS3-LoxP-Rev	caccacaactaacttttttcccgttcctccatctcttttatatttttttttcgaattcgagctcggtacccggg
FF2287	cgctcgaaggctttaatttgcggccggtcgactctagaggatccccggg
FF2288	cccggggatcctctagagtcgaccggccgcaaattaaagccttcgagcg

References

Fang, F., Salmon, K., Shen, M. W. Y., Aeling, K. A., Ito, E., Irwin, B., Tran, U. P. C., Hatfield, G. W., Da Silva, N. A., Sandmeyer, S., 2011. A vector set for systematic metabolic engineering in *Saccharomyces cerevisiae*. Yeast. 28**,** 123-136.

Gibson, D., 2009. One-step enzymatic assembly of DNA molecules up to several hundred kilobases in size.