UNIVERSITY OF CALIFORNIA,
IRVINE

Metabolic Engineering of *Saccharomyces cerevisiae* for the High-Level Synthesis of Polyketides

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in Chemical and Biochemical Engineering

By

Javier Cardenas

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

2015

“Science is always wrong. It never solves

a problem without creating ten more.”

-George Bernard Shaw

TABLE OF CONTENTS

Page

LIST OF FIGURES v

LIST OF TABLES ix

LIST OF ABBREVIATIONS x

ACKNOWLEDGEMENTS xii

CURRICULUM VITAE xiii

LIST OF AUTHOR CONTRIBUTIONS xv

ABSTRACT OF THE DISSERTATION xvi

CHAPTER 1: Introduction 1

1.1 Motivation 2

1.2 Objectives 10

1.3 Background 12

1.4 References 25

CHAPTER 2: Metabolic engineering of *Saccharomyces cerevisiae* for the production of triacetic acid lactone 31

2.1 Abstract 32

2.2 Introduction 33

2.3 Materials and Methods 37

2.4 Results and Discussion 45

2.5 Conclusions 63

2.6 Acknowledgements..... 64

2.7 References 65

CHAPTER 3: A combined approach for engineering a heterologous Type III PKS and enhancing polyketide biosynthesis 69

3.1 Abstract 70

3.2 Introduction 71

3.3 Materials and Methods 75

3.4 Results and Discussion..... 84

3.5 Conclusions..... 106

3.6 Acknowledgements..... 107

3.7 References 108

CHAPTER 4: Engineering cofactor and transport mechanisms in *Saccharomyces cerevisiae* for enhanced acetyl-CoA and polyketide biosynthesis 111

4.1 Abstract 112

4.2 Introduction 114

4.3 Materials and Methods 118

4.4 Results and Discussion..... 126

4.5 Conclusions..... 140

4.6 Acknowledgements..... 142

4.7 References 143

CHAPTER 5: Conclusion and Future Perspectives 150

5.1 Conclusion 151

5.2 Future Perspectives 153

5.3 References 157

APPENDIX A Functional replacement of the *Saccharomyces cerevisiae* fatty acid synthase with a bacterial type II system allows flexible product profiles 159

APPENDIX B Supplementary Material: Functional replacement of the *Saccharomyces cerevisiae* fatty acid synthase with a bacterial type II system allows flexible product profiles 175

LIST OF FIGURES

		Page
Figure 1.1	The value of a barrel of oil and its breakdown into product categories	3
Figure 1.2	Reconstructing the current petroleum industry	4
Figure 1.3	New method for construction of natural product-like compound collections	5
Figure 1.4	Chemical conversions of triacetic acid lactone	6
Figure 1.5	PKS substrates and intermediates for the Type III PKS 2-pyrone synthase	7
Figure 1.6	Pathways for deriving pyruvate from glucose and subsequent utilization	8
Figure 1.7	Comparison of fundamental processing steps in FAS and PKS	13
Figure 1.8	Mechanisms of diversification in the biosynthesis of polyketides	14
Figure 1.9	Complete pXP vector library	16
Figure 1.10	Schematic for gene integration or gene knockout by the pXP library	17
Figure 1.11	DNA Assembler method	18
Figure 1.12	Plasmid construction using 60-mer homologous sequences	18
Figure 1.13	Vector assembly by the versatile genetic assembly system (VEGAS)	19

		Page
Figure 1.14	Gibson Assembly for *in vitro* plasmid construction	20
Figure 1.15	The overall process flow for GDLS method	23
Figure 2.1	Strategy for plasmid construction	41
Figure 2.2	Comparison of promoters and varying glucose concentration	47
Figure 2.3	TAL titers using 2-PS, 3XCHS, and 1X6MSAS and protease-limited strains	48
Figure 2.4	Metabolic pathways identified for optimizing TAL precursors in *S. cerevisiae*	51
Figure 2.5	TAL produced in *S. cerevisiae* strains with single gene knockouts	53
Figure 2.6	Growth curves for various strains in this study	54
Figure 2.7	TAL levels following combination of protease and pathway gene knockouts	59
Figure 2.8	TAL Improvements on a per gram cell basis	60
Figure 2.9	TAL titers during fermentor-controlled glucose fed-batch cultivation	61
Figure 2.10	Additional characteristics of fed-batch fermentation	61
Figure 2.11	Final titer and yield summary for TAL production in *S. cerevisiae*	62
Figure 3.1	Illustration of amino acid positions conferring size in the active site	84

Page

Figure 3.2 Steady-state kinetic parameters for 2-PS and initial 2-PS variants 86

Figure 3.3 TAL titers (g/L) *in vivo* for Gen-1 variants in strain BJ5464 87

Figure 3.4 Jmol representation of 2-pyrone synthase crystal structure 89

Figure 3.5 TAL titers (g/L) *in vivo* for Gen-2 variants and comparison variants E2 and E7 90

Figure 3.6 Triacetic acid lactone titers (g/L) *in vivo* for Gen-3 variants in BJ5464 92

Figure 3.7 Triacetic acid lactone titers (g/L) *in vivo* for Gen-4 variants in BJ5464 95

Figure 3.8 Percent activity levels as a function of time in TCEP assay 96

Figure 3.9 Thiol concentrations in 2-PS cultivated in *S. cerevisiae* 98

Figure 3.10 *In vivo* TAL levels (g/L) in protease-limited strains 100

Figure 3.11 TAL titers (g/L) in batch cultivation of BJ5464 and BJ5464Δ*pyc2*Δ*nte1* 102

Figure 3.12 Profiling of glucose batch cultivation in BJ5464Δ*pyc2*Δ*nte1* 103

Figure 3.13 Maximum observed *in vivo* TAL levels (g/L) for fed-batch fermentations 104

Figure 3.14 Biomass profiles (g/L DCW) for fed-batch fermentation 105

Figure 4.1 Vector construction of pJCT-PDH and pJCT-PDHm 123

Page

Figure 4.2 General engineering strategy applied for introducing biological driving forces 127

Figure 4.3 NADPH/$NADP^+$ ratios for BYt strains harboring PDH vectors 128

Figure 4.4 Acetyl-CoA levels for BYt strains harboring PDH vectors 129

Figure 4.5 *In vivo* TAL titers (g/L) for BYt strains harboring PDH vectors 132

Figure 4.6 Metabolic pathways related to pyruvate/acetyl-CoA transport and utilization 134

Figure 4.7 *In vivo* polyketide levels following single gene knockouts after 48 h 135

Figure 4.8 *In vivo* polyketide levels for engineered strains using PDH vectors 138

Figure 4.9 Final titer summary for polyketide production (g/L TAL) in *S. cerevisiae* 141

Figure 5.1 Overall increases in various platform strains for TAL synthesis in *S. cerevisiae* 152

Figure 5.2 Triacetic acid lactone titers for strains of the Knockout Library collection 155

Figure 5.3 Triacetic acid lactone profiles for 9 promoters using high-copy expression 156

Figure 5.4 Triacetic acid lactone profiles for strains with reduced *FAS2* expression 156

LIST OF TABLES

		Page
Table 2.1	List of plasmids and strains (constructed and used in Chapter 2)	38
Table 2.2	List of primers (constructed and used in Chapter 2)	39
Table 2.3	Growth rates and final cell densities for strain BY4741	46
Table 2.4	List of reactions and gene knockouts identified by the OptKnock algorithm	55
Table 3.1	List of plasmids and strains in *E. coli* (constructed and used in Chapter 3)	76
Table 3.2	List of plasmids and strains in *S. cerevisiae* (constructed and used in Chapter 3)	78
Table 3.3	List of primers (constructed and used in Chapter 3)	79
Table 3.4	Steady state kinetic parameters for wildtype 2-PS and 2-PS variants (Gen-1)	86
Table 3.5	Steady state kinetic parameters for wildtype 2-PS and 2-PS variants (Gen-2)	91
Table 3.6	Percent increased TAL titer (g/L) for Gen-3 variants, relative to E2 (C35S)	94
Table 4.1	List of plasmids and strains (constructed and used in Chapter 4)	119
Table 4.2	List of primers (constructed and used in Chapter 4)	120

LIST OF ABBREVIATIONS

2-PS	2-pyrone synthase
6-MSA	6-methylsalicylic acid
ATP	Adenosine triphosphate
CHS	Chalcone synthase
CoA	Coenzyme A
COBRA	Constraint based reconstruction and analysis
DCW	Dry cell weight
DO	Dissolved oxygen
DTNB	5,5'-Dithiobis(2-nitrobenzoic acid)
FAS	Fatty acid synthase
FBA	Flux balance analysis
GDLS	Genetic design through local search
GEMs	Genome scale metabolic models
GRAS	Generally regarded as safe
GSH	Glutathione
HPLC	High performance liquid chromatography
KAS	Ketoacyl synthase
LNKS	Lovastatin nonaketide synthase
MCS	Multiple cloning site
MILP	Mixed-integer linear optimization program
MINLP	Mixed-integer nonlinear optimization program
MOMA	Minimization of metabolic adjustment
MPC	Mitochondrial pyruvate carrier
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate

LIST OF ABBREVIATIONS, CONT.

PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDH	Pyruvate dehydrogenase
PKS	Polyketide synthase
PPT	Phosphopantetheinyl transferase
ROS	Reactive oxygen species
SBML	Systems Biology Markup Language
SDM	Site directed mutagenesis
TAL	Triacetic acid lactone
TCEP	Tris (2-carboxyethyl) phosphine
VEGAS	Versatile genetic assembly system
WT	Wildtype

ACKNOWLEDGEMENTS

It has been a true pleasure completing the Ph.D. at UC Irvine, and I am forever grateful to every individual who contributed to my success. I first want to thank my advisor, Dr. Nancy Da Silva, for allowing me this opportunity and for her strong support throughout the journey that is graduate school. I want to recognize all my past mentors, especially Chris Leber, Jin Wook Choi, and Corey Dodge for instilling in me valuable lab skills. Thank you to Dr. Wang and Dr. Kwon for allowing me to rotate in their labs and getting me excited for doing research.

I would like to thank all the faculty in the CBEMS department who helped me along the way, especially past committee members Prof. Hung D. Nguyen and Prof. Betty H. Olson, as well as my thesis committee members Dr. Suzanne Sandmeyer and Dr. Szu-Wen Wang. A special thanks to Prof. Martha Mecartney for her commitment in facilitating GAANN fellowship support to myself and others in the department over the years. Thanks to Steve Weinstock for his contagious good spirit and all the fun times.

I also want to thank all my undergraduate students who I TAed over the years. It was very rewarding to help them find a passion for the field of Chemical & Biochemical Engineering, and they have in turn made me a much wiser mentor.

This work has been supported by CBiRC, so I would like to thank the National Science Foundation as well as CBiRC's Director Brent Shanks for their strong commitment to excellence through collaborative research. A big thanks to the whole CBiRC family, but a special thanks to Tom Schwartz, Anupam Chowdhury, Kate Woods, and Chris Vickery for the guidance and support.

To all the past and present Da Silva Lab and Wang Lab members, thank you for all the help and support, and for making it fun to come in to lab. Lastly, a special thanks to my family who have been there since Day 1 and believing that this day would arrive. We made it guys.

CURRICULUM VITAE

Javier Cardenas

Education:

2010-2015	Doctor of Philosophy, Chemical and Biochemical Engineering University of California, Irvine
2008-2010	Master of Science, Chemical and Biochemical Engineering University of California, Irvine
2003-2008	Bachelor of Science, Biochemical Engineering University of California, Davis

Work Experience:

2010-2013	Teaching Assistant, CBEMS 45A: Chemical Processing & Materials Balances University of California, Irvine
2009	Validation Associate at Baxter Biosciences Glendale, CA
2008	Field Research Scientist at Valent Biosciences Fresno, CA
2006-2007	Validation Associate at Pharmaceuticals Services Corporation Vacaville, CA
2006	Process Engineering Coop at ALZA Corporation Vacaville, CA
2004-2008	Research Assistant at Plant Cell Culture Bioprocessing Laboratory (UC Davis) Davis, CA

Selected Presentations:

1. **Javier Cardenas**, Nancy A. Da Silva, Kate Woods, Marianne Bowman, Joseph Noel. Engineering *Gerbera hybrida* 2-pyrone synthase (2-PS) and host *Saccharomyces cerevisiae* for the high-level synthesis of triacetic acid lactone, Division of Biochemical Technology, 247th ACS National Meeting, Dallas, TX, March 2014.

2. **Javier Cardenas**, Nancy A. Da Silva. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone (TAL), Division of Biochemical Technology, 245th ACS National Meeting, New Orleans, LA, April 2013.

3. Thomas J. Schwartz, Robert L. Johnson, **Javier Cardenas**, Adam Okerlund, Nancy A. Da Silva, Klaus Shmidt-Rohr, James A. Dumesic. Engineering Catalyst Microenvironments for Metal-Catalyzed Hydrogenation of Biologically-Derived Platform Chemicals, Catalysis and Reaction Engineering Division, AIChE Annual Meeting, Atlanta, GA, November 2014.

CURRICULUM VITAE, CONT.

4. Thomas J. Schwartz, Robert L. Johnson, **Javier Cardenas**, Nancy A. Da Silva, Klaus Shmidt-Rohr, James A. Dumesic. Catalyst Design for the Integration of Heterogeneous Catalysis With BioCatalysis, Fuels and Petrochemicals Division, AIChE Annual Meeting, San Francisco, CA, November 2013.

5. Ruben Fernandez-Moya, **Javier Cardenas**, Chris Leber, Nancy A. Da Silva. Engineering *Saccharomyces cerevisiae* for the synthesis of short-chain carboxylic acids using a type II FAS system, Division of Biochemical Technology, 243rd ACS National Meeting, San Diego, CA, March 2012.

Selected Publications and Patents:

1. **Cardenas, J.**, Da Silva, N.A., Gordon, L., Bowman, M., Woods, K., Noel, J. Polyketide synthase variants and uses thereof. *United States Provisional Patent* (US Patent App. 61/949,082) application filed on March 6, 2014 and *PCT Patent* (PCT/US15/19058) application filed March 5, 2015.

2. **Cardenas, J.** & Da Silva, N.A. Engineering cofactor and transport mechanisms in *Saccharomyces cerevisiae* for enhanced acetyl-CoA and polyketide biosynthesis. *In preparation*

3. Vickery, C.*, **Cardenas, J.***, Da Silva, N.A. and Noel, J.P. A combined approach for engineering a heterologous Type III polyketide synthase and enhancing triacetic acid lactone biosynthesis in *Saccharomyces cerevisiae*. *In preparation* (*equal contributions)

4. Fernandez-Moya, R., Leber**,** C., **Cardenas, J.** & Da Silva, N.A. Functional replacement of the *Saccharomyces cerevisiae* fatty acid synthase with a bacterial type II system allows for flexible product profiles. Biotechnol. Bioeng. (2015)

5. Schwartz, T.J., Johnson, R.L., **Cardenas, J.**, Okerlund, A., Da Silva, N.A., Schmidt-Rohr, K. and Dumesic, J.A. Engineering catalyst microenvironments for metal-catalyzed hydrogenation of biologically derived platform chemicals. Angew. Chem. Int. Ed. (2014), 53: 12718-12722. **^**

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

^ This thesis is not based on this co-authored material and has therefore not been included.

LIST OF AUTHOR CONTRIBUTIONS

This summary highlights contributions for any co-authored materials in this thesis, as outlined above:

Paper I A combined approach for engineering a heterologous Type III polyketide synthase and enhancing triacetic acid lactone biosynthesis in *Saccharomyces cerevisiae* (This work has been included in Chapter 3):

JC and CV designed the study. JN and ND supervised the project. **JC** performed all molecular biology experiments and oxidation assays in *S. cerevisiae*, including HPLC analysis and fermentor cultivations. CV performed experiments in *Escherichia coli*, and determined kinetics and TCEP data for 2-PS and its variants.

Paper II Functional replacement of the *Saccharomyces cerevisiae* fatty acid synthase with a bacterial type II system allows for flexible product profiles (This work has been included in Appendix A):

RFM, CL, and **JC** designed the study. NAD supervised the project. RFM performed a significant amount of the experimental work, constructing the full *E. coli* FAS and evaluating the fatty acid profile *in vivo*. CL performed molecular biology experiments and constructed library of vectors. **JC** supported molecular biology experiments and performed initial characterization of the *E. coli* FAS system using SDS-PAGE, Western Blot Analysis, and NADPH assay. RFM analyzed the data and wrote the manuscript. All authors edited and approved the final manuscript.

ABSTRACT OF THE DISSERTATION

Metabolic Engineering of *Saccharomyces cerevisiae* for the High-Level Synthesis of Polyketides

By

Javier Cardenas

Ph.D. in Chemical and Biochemical Engineering

University of California, Irvine, 2015

Professor Nancy A. Da Silva, Chair

Current advances in industrial biotechnology are expanding the breadth of what is possible through the use of microbial cell factories. With global resources being exhausted at rapid rates and a heavy reliance on nonrenewable feedstocks, there is pressure to find alternatives using biorenewables. Application of these products is widespread, from reducing petroleum-based demand in the chemical industry, to synthesizing next-generation drug therapeutics at low cost. At the center of this biobased innovation is the need for a workhorse microorganism, capable of efficiently manufacturing these target molecules. In the context of biorenewable compounds, there are many cases where bacteria are not an option due to host toxicity or lack of processing ability, for example, post-translational modification. Baker's yeast, *Saccharomyces cerevisiae*, is an ideal candidate as it is well characterized and frequently used industrially. Here we report its use in producing the polyketide triacetic acid lactone (TAL), an important polyketide which can be readily converted to replace fossil-fuel derived counterparts. Due to the complexity inherent in any host cell, metabolic pathway engineering was performed

on central carbon metabolism to increase availability of acetyl-CoA and malonyl-CoA metabolites. The computational tool OptKnock in conjunction with other knockout interventions were used to create a strain (BY4741Δ*prb1*Δ*pyc2*Δ*nte1*Δ*yia6*) producing 2.2 g/L TAL, 26% of the theoretical yield on glucose, in fed-batch fermentation. In addition, enzyme engineering of the *Gerbera hybrida* 2-pyrone synthase (2-PS) catalyzing TAL formation was conducted and then evaluated in our strains. This effort screened 41 2-PS variants and elucidated key residues in the active site and the protein surface, improving catalytic efficiency and/or stability, leading to 2.5-fold increases in TAL levels in *S. cerevisiae*. The surface-modified variant 2-PS (C35S, C372S) exhibited the highest k_{cat}/K_m, over 40 times that of the native synthase. This variant was introduced into an engineered strain (BJ5464Δ*pyc2*Δ*nte1*) and cultivated in glucose or ethanol fed-batch to yield 2.7 g/L and 4.2 g/L TAL, respectively. Fermentation parameters were further improved and feed components modified for high titer (7.6 g/L) and yield (44% of theoretical) of this polyketide. Enhancement of precursor availability was also achieved by using a heterologous pathway from *Escherichia coli*. A modified *E. coli* bacterial pyruvate dehydrogenase (PDHm) was introduced in a cofactor-limited (Δ*zwf1*) strain to couple NADPH cofactor balance with the synthesis of the TAL precursor acetyl-CoA. To prevent loss of acetyl-CoA and pyruvate by transport to the mitochondria, recently elucidated pathways were disrupted (Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*). The coupling of all three approaches (Δ*zwf1*Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2* +PDHm) elevated titers to 1.6 g/L, the highest by batch cultivation (and 35% of theoretical yield). These strategies are directly applicable to the synthesis of other acetyl-CoA derived compounds, and these studies provide a framework to develop design principles for robust, industrially relevant yeast for the synthesis of polyketides.

Chapter 1:

Introduction

1.1 Motivation

For the past two centuries, mankind has extensively relied on petroleum as a feedstock for both fuels and the synthesis of many different chemicals (Nikolau et al., 2008). This flexibility in the number of compounds and also their properties has led to the success of the petrochemical industry. While utilizing just over 3% by volume of a barrel of oil, the petrochemical landscape is valued at 375 billion (USD). This is just shy of the fuels market which uses 70% of a barrel by volume (Figure 1.1) (Johnson, 2012). By downstream derivatization and synthesis, over 70,000 products can be made from the very simple starting materials found in crude oil (Vennestrøm et al., 2011). This relatively straightforward process with high selectivity and efficiency keeps production costs very low, yet the finite amount of fossil-fuel carbon in the world raises concern about the sustainability of this infrastructure. With a volatile supply and demand on petroleum (Regnier, 2007), it is clear that a renewable alternative will be required. Biobased or biorenewable chemicals are an attractive option for replacing or supplementing the chemical industry as biomass can be generated for use in biorefineries that can sustainably manufacture goods (Langeveld et al., 2010). This model has already been used in certain markets such as ethanol production capabilities in Brazil using sugarcane bagasse as a feed. Additionally, the National Science Foundation (NSF) Center for Biorenewable Chemicals (CBiRC) is dedicated to creating translational research that can serve as a framework to build a biobased industry (Figure 1.2). Current strategies in the center seek to use both fatty acid and polyketide metabolism as the platform to develop biocatalysis and chemical catalysis routes toward carboxylic acid and pyrone derivatives. While biorenewables can transform the chemical industry, sustainable biobased products can do more than just replace fossil-fuel derived chemicals.

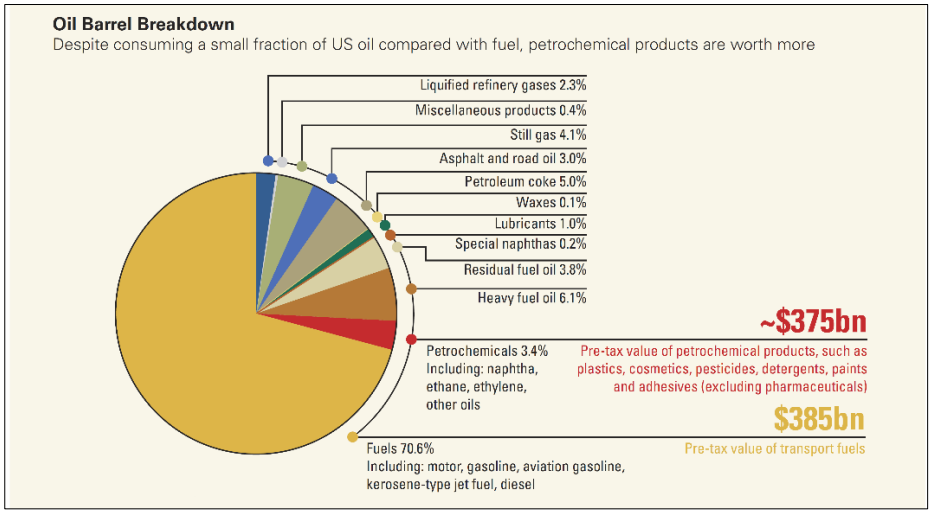


Figure 1.1 The value of a barrel of oil and its breakdown into product categories (Johnson, 2012).

The biopharmaceutical space is also invested in adopting biobased methodologies. Just as the petrochemical industry has had success leveraging chemical synthesis, biotech and biopharma companies have relied almost solely on high throughput screening techniques since it was developed as a discovery method in the early 1990s for drug candidate development, starting the era of combinatorial chemistry in industry (Rouhi, 2003). These screening methods in large part facilitated the discovery of many therapeutics over the past two decades (Herper, 2012). There are two caveats that will likely prevent similar successes using this strategy going forward. First, a significant number of FDA approvals granting patent protection from 2012-2019 creates an opportunity for biosimilars to capture a significant share of the market (~25%), also known as the patent cliff (Calo-Fernández and Martínez-Hurtado, 2012). Additionally, while the combinatorial chemistry approach generates tens of thousands of compounds, it lacks the structural and functional diversity required for next generation drug discovery

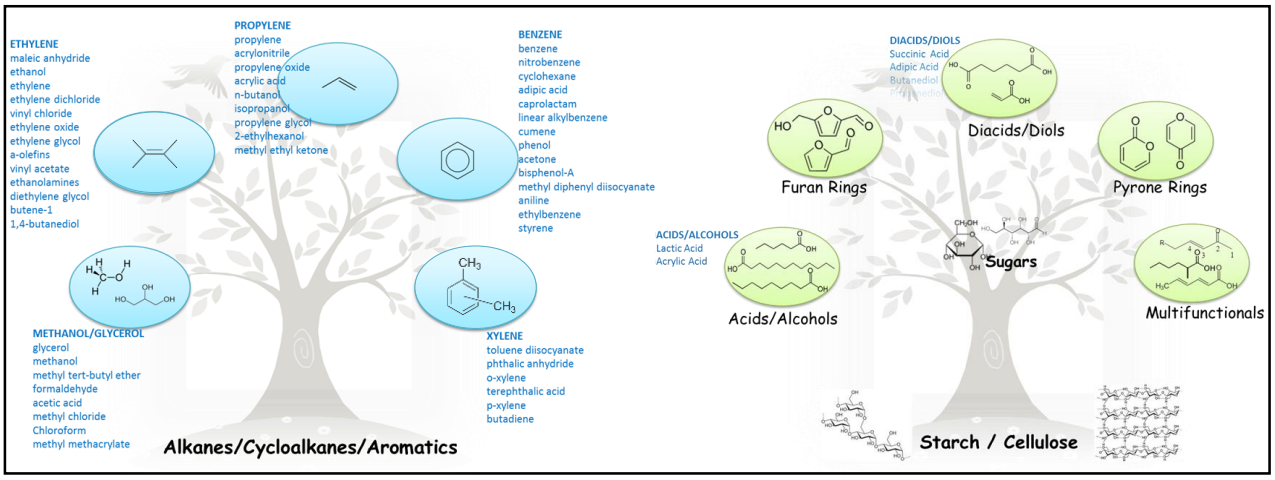


Figure 1.2 Reconstructing the current petroleum industry (left) toward a biorenewable alternative framework (right), illustrating the biological conversion of sugars toward carboxylic acid and pyrone derivatives (www.cbirc.iastate.edu).

(Feher and Schmidt, 2003). To date, only a single drug, Onyx and Bayer's Nexavar (sorafenib), has been discovered using *de novo* combinatorial chemistry (Mander and Liu, 2010; Newman and Cragg, 2007). Centuries of evolution have shown that nature has done a superb job of creating high level diversity which is an opportunity for enhanced drug compound libraries. As illustrated (Figure 1.3), new approaches are taking a natural product (compound produced by a living organism) scaffold, which has stereochemistry naturally, and introducing moieties that further alter the chirality of the compound, a feature that is absent when attaching traditional elements to a core chemical scaffold. With natural products and their derivatives making up 41% of anticancer drugs and 65% of antibacterial drugs (Borman, 2013), biopharmaceutical companies will increasingly rely on these compounds for developing new therapies. Organizations that plan to capitalize on biosimilars or new natural product therapies will therefore rely on a biological

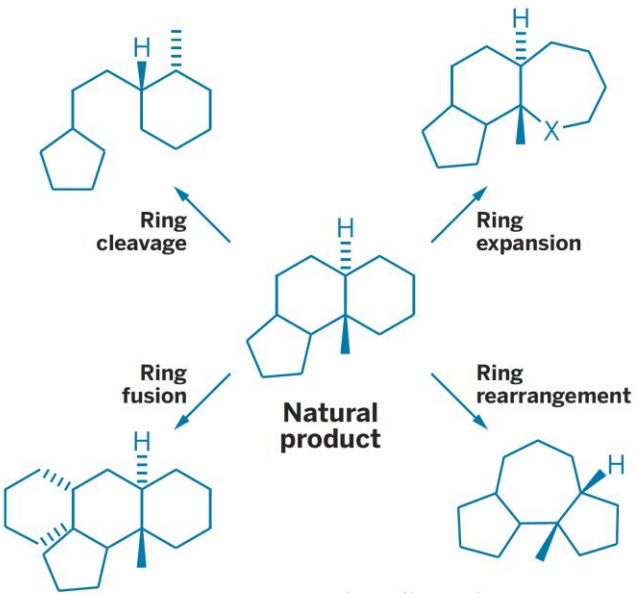


Figure 1.3 New method for construction of natural product-like compound collections exemplified for ring structures (Huigens et al., 2013).

system for drug candidate creation as well as large-scale production.

Polyketides are one class of natural products, and their biosynthesis illustrates a compound diversity that makes them an attractive platform for both biorenewables and biopharmaceuticals. This diversity has already been harnessed for cholesterol-lowering, anticancer, antifungal, antiviral, antimicrobial, and antioxidant products, stimulating the discovery of polyketide biosynthetic routes for large-scale production (Lussier et al., 2012; Pfeifer and Khosla, 2001). Recently, the polyketide triacetic acid lactone (TAL) has shown promise as a chemical intermediate in the biorenewables space (Chia et al., 2012; Saunders et al., 2015). TAL has been proposed as a platform due to the ability to readily transform it into a variety of products using well established chemical catalysis methods (Chia et al., 2012). It can be converted

to sorbic acid, currently used as a food preservative, as well as other compounds used as flavoring agents, resins, adhesives, and plastics, all of which are currently manufactured by non-renewable means (Figure 1.4).

To consider biosynthesis of polyketides in a host organism, the supply of starting materials needs to be addressed. The options for substrate pool can include acetyl-CoA, propionyl-CoA, isobutyryl-CoA, isovaleryl-CoA, malonyl-CoA, methylmalonyl-CoA, ethylmalonyl-CoA, propylmalonyl-CoA, or hydroxymalonyl-CoA, many of which are not readily available or even native to yeast (Pfeifer and Khosla, 2001). In the case of *S. cerevisiae*, the simplest available substrates of acetyl-CoA and malonyl-CoA are produced natively for many purposes like the synthesis of fatty acids (Leibundgut et al., 2008). For Type III PKSs, various enzymes require malonyl-CoA as extender unit with a respective initiator substrate (Austin and Noel, 2003). As illustrated (Figure 1.5), the 2-pyrone synthase (2-PS) derived from *Gerbera hybrida* uses one

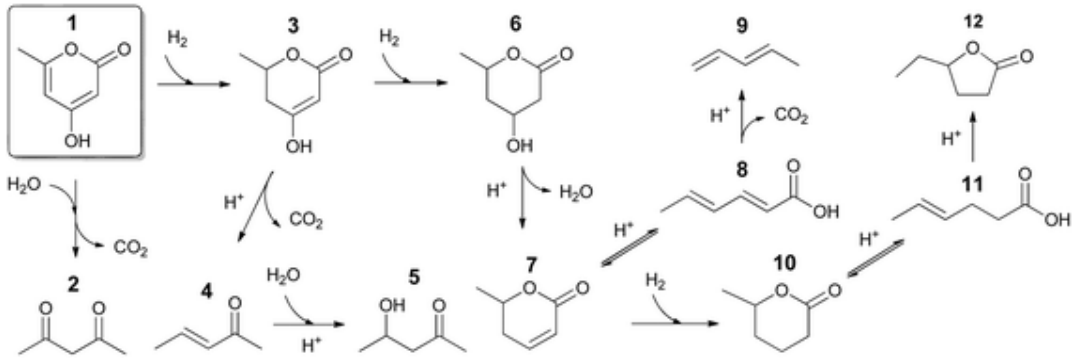


Figure 1.4 Chemical conversion of (1) triacetic acid lactone by hydrogenation, dehydration, ring opening, or decarboxylation (Chia et al., 2012). These reactions lead to the following products: (2) 2,4-pentanedione, (3) 5,6-dihydro-4-hydroxy-6-methyl-2*H*-pyran-2-one, (4) 3-penten-2-one, (5) 4-hydroxy-2-pentanone, (6) 4-hydroxy-6-methyltetra-hydro-2-pyrone, (7) 6-methyl-5,6-dihydro-2-pyrone, (8) 2,4-hexadienoic acid, (9) 1,3-pentadiene, (10) δ-hexalactone, (11) hexenoic acid, and (12) γ-caprolactone.

acetyl-CoA and two malonyl-CoA units to produce a single triketide before a single C5→C1 lactonization makes the final product. In *S. cerevisiae*, production of these metabolites proceeds through the pyruvate branch point node (Figure 1.6) following the catabolic process of glycolysis (Pronk et al., 1996). Specifically, acetate is converted to acetyl-CoA by the acetyl-CoA synthetase (encoded by *ACS1* and *ACS2*), an ATP-consuming reaction. Subsequently, acetyl-CoA acts as the substrate for malonyl-CoA formation via the acetyl-CoA carboxylase (encoded by *ACC1*). While the presence of both acetyl-CoA and malonyl-CoA greatly simplifies polyketide biosynthesis in yeast relying on these substrates, the pre-existing demand for these pools will require engineering new avenues for increasing their availability.

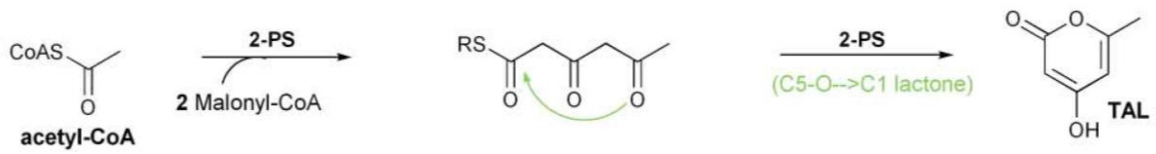


Figure 1.5 PKS substrates and intermediates for the Type III PKS 2-pyrone synthase (2-PS) and the final product triacetic acid lactone (TAL) (Austin and Noel, 2003).

Introduction of a heterologous system can also bring the challenge of identifying ideal expression conditions. Achieving optimum expression of a protein requires a large degree of control over the timing and the strength of expression. This is facilitated by the use of vectors that dictate the copy number, promoter, and terminator for the gene of interest. For *S. cerevisiae*, a variety of plasmids are readily available to introduce expression cassettes, including a large library developed by our lab and colleagues (Fang et al., 2011; Shen et al., 2012). Additional plasmid systems are available beyond this set, although it is evident that constitutive promoters

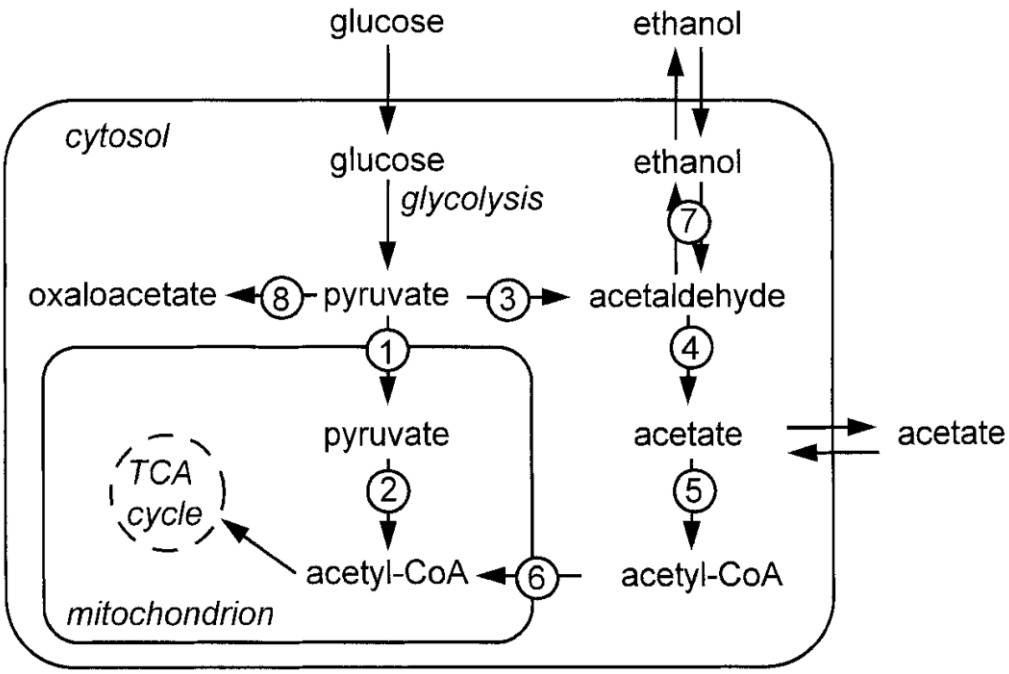


Figure 1.6 Pathways for deriving pyruvate from glucose and subsequent utilization by *S. cerevisiae* (Pronk et al., 1996).

are the most widely available (Da Silva and Srikrishnan, 2012). Heterologous expression can be difficult on the host, especially for large enzymes; therefore, having an increased flexibility in the expression platform can be useful in the case of polyketides and other target compounds. Additionally, one must take note of native processes that can target the enzyme for degradation. Proteases in the cell are the enzymes responsible for cleaving proteins to be recycled, and those found in the vacuole (e.g., Proteinase A and B) have previously been shown to affect polyketide synthase levels (Lee et al., 2009). Evaluating the expression of the enzyme in conjunction with the final product can guide appropriate strategies to increase stability of the enzyme in the microbial host system for maximum performance and productivity.

The evidence presented above suggests that microbial production of polyketides can be extremely valuable toward the biorenewables and biopharma/biotech fields. In the case of host selection, plants, bacteria, and fungi all are naturally capable of synthesizing polyketides. Nonetheless, controlling the biosynthetic machinery and forcing a host to make large product quantities can be difficult and deleterious. For example, strains often are unable to tolerate exogenous materials that nature has designed to be toxic to cells (Pfeifer and Khosla, 2001). Therefore, having a well-characterized organism in which pathways are known and engineering tools are available would aid in achieving this goal of producing polyketides. This makes *S. cerevisiae* a suitable microbial host candidate as it has been generally regarded as safe (GRAS) by the FDA ("Generally Recognized as Safe (GRAS)," 1986), was the first eukaryote to be fully sequenced in 1996 (Goffeau et al., 1996), and to date 78% of the nearly 7,000 known open reading frames have been verified ("Saccharomyces cerevisiae Genome Overview," 2015). This expansive wealth of information has spurred metabolic engineering tools that facilitate the study and engineering of this yeast species. This thesis focuses on engineering *S. cerevisiae* for high-level production of polyketides, specifically acetyl-CoA derived compounds.

1.2 Objectives

The primary goal of the research was to optimize *S. cerevisiae* biosynthetic pathways for producing polyketides, and to determine critical factors for increasing the carbon flow to the precursor acetyl-CoA. In addition, factors influencing the activity and stability of the polyketide synthase both *in vitro* and *in vivo* were determined. The knowledge gathered by our studies will facilitate further proliferation of microbial production platforms and create highly efficient processes for the synthesis of target molecules. Specifically, this thesis aimed to enhance *in vivo* synthesis of the polyketide triacetic acid lactone (TAL) in *S. cerevisiae* by expressing engineered variants of the Type III PKS 2-pyrone synthase (2-PS) from *Gerbera hybrida*, and engineering diverse yeast metabolic pathways.

Objective 1:

The first aim was to identify gene disruptions in *S. cerevisiae* pathways for improved production of triacetic acid lactone, and specific aims included:

- Selection of a 2-PS expression system for improved TAL titers *in vivo*
- Disruption of native yeast pathways using known features of central carbon metabolism and the *in silico* prediction algorithm OptKnock
- Construction of an engineered strain incorporating simultaneous disruption of multiple pathways
- Fed-batch fermentation of the engineered strain for increased production of TAL

Objective 2:

The second aim was the design and evaluation of engineered 2-PS variants for improving performance *in vivo* as demonstrated through increased TAL biosynthesis. Specific aims included:

- Determination of TAL-synthesizing potential of 1st generation 2-PS variants designed for altered active site conformation
- Iteration of new enzyme variants (2nd / 3rd generation) incorporating design rules from prior generation constructs
- Evaluation of 2-PS variants with reduced surface cysteine incorporation (4th generation) for isolating the mechanism of 2-PS degradation
- Combination of engineered strain (from Objective 1) and optimal 2-PS variant, and fed-batch glucose and ethanol cultivation for enhanced TAL production

Objective 3:

The third aim was to create both a resupply and a biological driving force for the synthesis of acetyl-CoA using an engineered *S. cerevisiae* strain and heterologous pyruvate dehydrogenase pathway from *E. coli*. The specific aims included:

- Validation of an actively expressed heterologous pyruvate dehydrogenase mutant (PDHm) by determining the NADPH/NADP$^+$ cofactor ratio
- Evaluation of a yeast strain lacking its major NADPH supply to improve flux through the heterologous NADPH-generating pyruvate dehydrogenase and increase acetyl-CoA pools
- Disruption of major mitochondrial transport mechanisms of pyruvate and acetyl-CoA to resupply metabolite pools and increase TAL synthesis

- Construction of an engineered strain harboring both transport and NADPH-generating gene disruptions, and the overexpression of the modified $NADP^+$-utilizing PDHm, for increased TAL synthesis

1.3 Background

1.3.1 Polyketide biosynthesis

Polyketides are secondary metabolites, and their synthase systems are differentiated into three classes (Type I, Type II, and Type III), similar to fatty acid synthases (FASs). Due to this similarity (Figure 1.7), it is believed that PKS and FAS systems may have had an evolutionary divergence at an early stage (Gokhale et al., 2007; Hertweck, 2009). The biotechnology industry has taken advantage of polyketides; an example is statin biosynthesis for the development of cholesterol-lowering drugs like lovastatin, creating a market valued in the billions (Comfort, 2013). Polyketide biosynthesis, similar to fatty acid biosynthesis, can be very complicated and require many reactions to make the final product. These required activities include (Figure 1.7): (1) acyl transferase (AT) activity for shuttling substrates, (2) ketoacyl synthase (KAS) activity for condensing of initiator and elongation chains, (3) dehydrogenase (DH) / reductase (ER / KR) activity for recycling chains in the elongation cycle, and (4) thioesterase (TE) activity for cleavage of the carrier molecule (Austin and Noel, 2003; White et al., 2005). This complex biological framework exposes the divergence of FAS versus PKS systems. First, fatty acids are confined to specific chain lengths and mandatory ketoreductase processing. This system is therefore severely limited in the product profiles it can generate relative to a PKS with additional tailoring steps available (Figure 1.8). Second, unlike fatty acids which are necessary for all living organisms,

polyketide synthesis can be introduced into heterologous organisms which do not have native PKS function, eliminating a level of competition for starter metabolite pools and aspects of native regulation. Of the three PKS types, only Type III utilizes the CoA carrier, meaning acyl transferase (AT) activity and a corresponding phosphopantetheinyl transferase (PPT) are not needed. Heterologous incorporation of a PKS system into a microbial host like *S. cerevisiae* can therefore be preferable to a FAS complex and can also be simplified significantly based on these attributes.

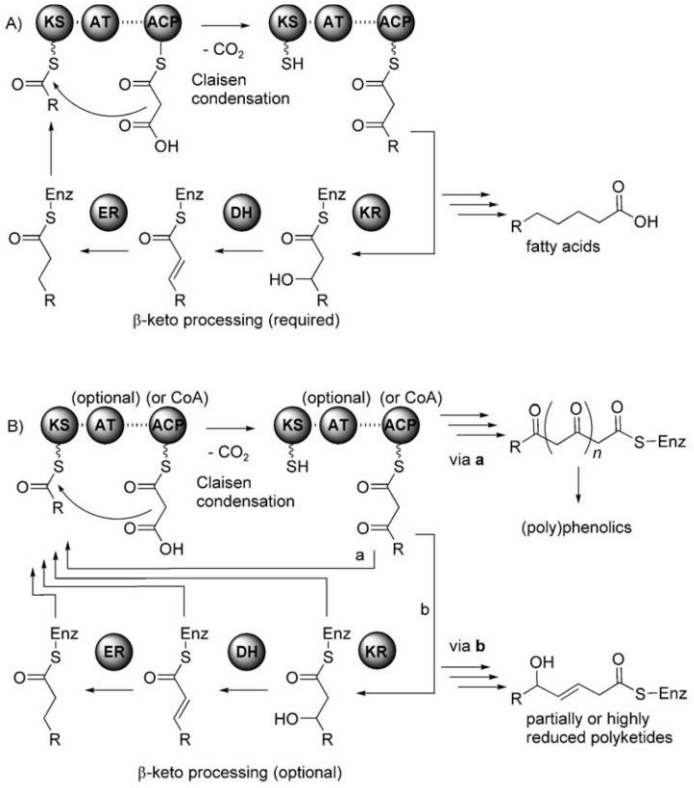


Figure 1.7 Comparison of fundamental processing steps in both (A) Fatty Acid Biosynthesis and (B) Polyketide Biosynthesis (not including cyclization steps) (Hertweck, 2009). The various steps confer: ketoacyl synthase (KS) activity, acyl transferase activity (AT), dehydrogenase (DH) activity, enoyl reductase (ER) and ketoreductase (KR) activity.

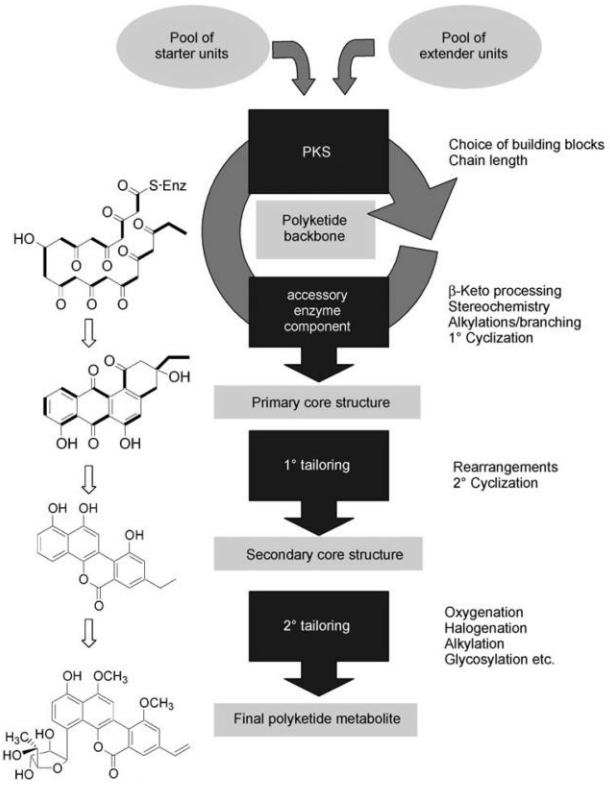


Figure 1.8 Mechanisms of diversification in the biosynthesis of polyketides (Hertweck, 2009).

1.3.2 Fine-tuning expression in *S. cerevisiae*

Apart from using well-characterized promoters to control expression, variants have been constructed and subject to truncation or random mutagenesis to create finely tuned expression profiles. As an example, a study performed error-prone PCR on the *TEF1* promoter, and using a fluorescence reporter identified *TEF1*p variants ranging from 10-fold lower to 3-fold higher fluorescence than the native *TEF1*p (Alper et al., 2005). Similarly, the *ADH1* promoter has been modified using the method of truncating segments off the promoter to identify regulatory elements in the promoter region (Ruohonen et al., 1995). Two fragments, the first with 1100 base pairs (bp) removed and the second with 800 bp deleted, were used to elucidate glucose

repressed and constitutive expression profiles, both drastically different from early-stage expression by the native promoter. Since promoters can be large (~2kb) and have many of these regulatory elements, researchers have attempted to further characterize synthetic promoters that achieve some form of a minimal set of features conferring variable expression levels. Specifically, this concept was proven for five yeast promoters using minimal promoter cores synthetically hybridized to different upstream activating sequences (UASs) (Blazeck et al., 2012). These synthetic promoters created varying expression levels consisting of controlled, step-wise increases from no detectable levels to 15% above the native promoter sequence. Once discovered, incorporating these features into metabolic engineering studies can be facilitated by having flexible yeast shuttle vectors. Recently our group constructed a vector library of ready-to-use plasmids employing features like a flexible multiple cloning site (MCS), variable promoters and origin of replication sequences, and recyclable selection markers for rapid and robust strain engineering useful in auxotrophic lab strains (Fang et al., 2011). Additional vectors were designed to expand this library (Shen et al., 2012), netting a pXP library of 44 vectors (Figure 1.9). In addition to being a shuttle vector, the features of this library allow for rapid integration and gene disruption strategies as illustrated (Figure 1.10). These are brief examples of the breadth of flexibility in heterologous expression using *S. cerevisiae* (Da Silva and Srikrishnan, 2012).

Additionally, the advent of new methods and the relative ease of PCR-based assemblies is expanding the way vectors are constructed for strain development studies, especially when using large multigene pathways. This type of gene assembly removes the dependence on

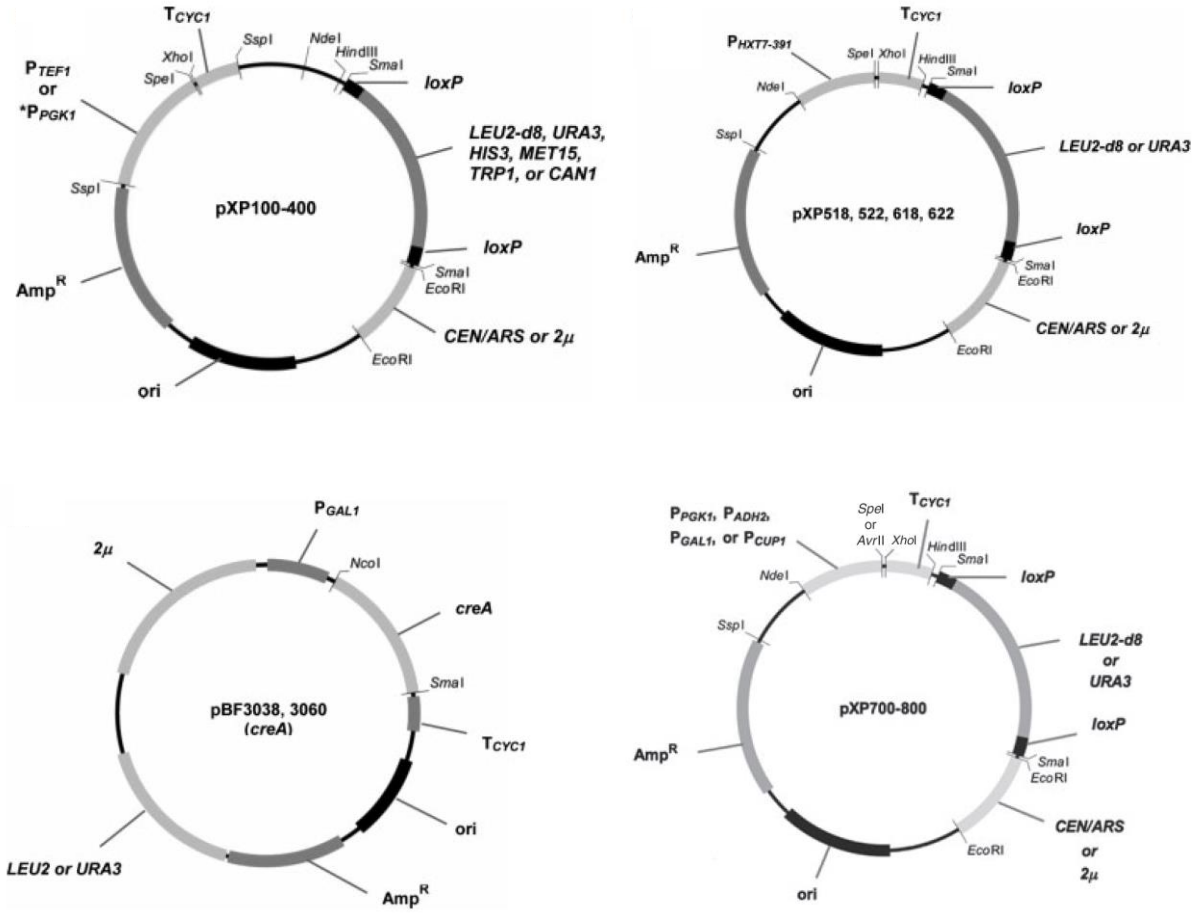


Figure 1.9 Complete pXP vector library (Fang et al., 2011; Shen et al., 2012). The plasmids have *CEN/ARS* or 2µ ori, CYC1 terminator, loxP-flanked selection markers (*LEU2-d8*, *URA3*, *HIS3*, *MET15*, *TRP1*, or *CAN1*), using one of six promoters (*TEF1*p, *PGK1*p, *HXT7-391*p, *GAL1*p, *ADH2*p, or *CUP1*p).

restriction enzymes and multiple cloning sites (MCSs) when doing traditional cloning, a benefit which is critical to promote proper expression levels (Crook et al., 2011). First, DNA Assembler uses overlap-extension PCR to quickly generate long multigene cassettes, eliminating the need of restriction enzymes and multiple cloning sites (MCSs) (Shao et al., 2009). These fragments can then be reassembled *in vivo* using the ability of *S. cerevisiae* in homologous recombination (Figure 1.11). The linearized backbone used in this strategy causes significant issues including backbone

recircularization (Kuijpers et al., 2013), leading to very low (10-20%) transformation efficiencies for an eight-gene pathway (Shao et al., 2009). In addition, chromosomal integration via single-crossover can lead to instabilities. Poor efficiency has recently been circumvented by the breakdown of the backbone into smaller elements with 60-mer homology (Figure 1.12) yielding 95% transformation efficiency for a seven-gene pathway (Kuijpers et al., 2013). Assembly of the gene fragments can additionally be simplified by the versatile genetic assembly system (VEGAS) which similarly relies on *in vivo* recombination in *S. cerevisiae* (Mitchell et al., 2015). This system, using a variation of the Golden Gate assembly (Engler et al., 2009, 2008) called yeast Golden Gate (yGG) (Agmon et al., 2015), creates overhangs by BsaI/BsmBI restriction enzymes to quickly join

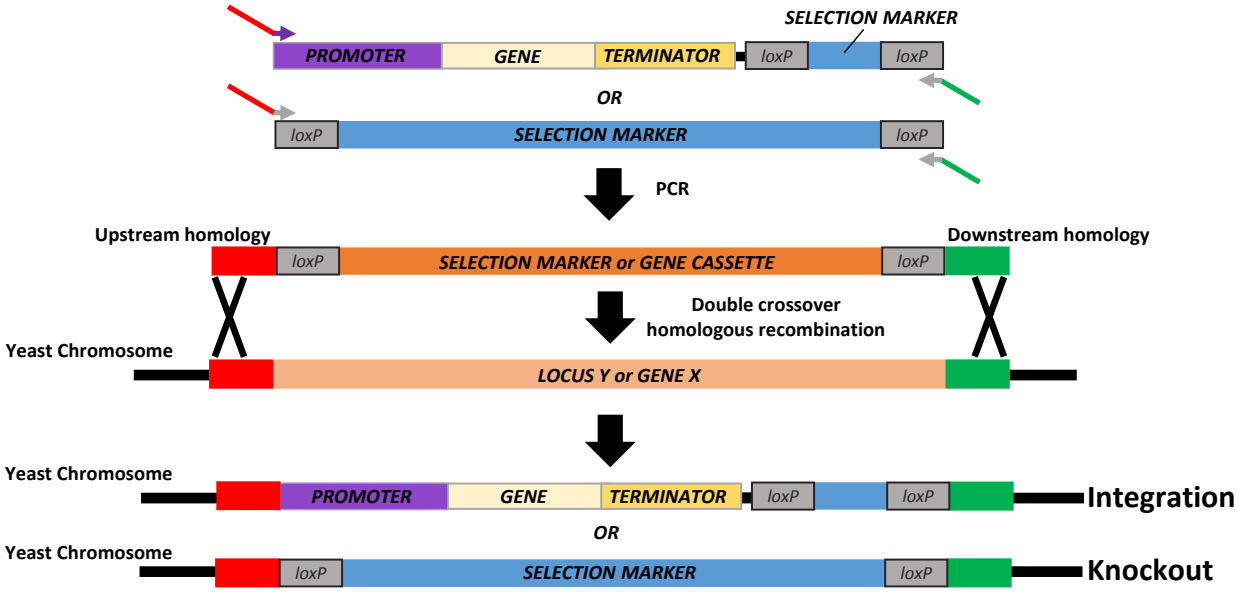


Figure 1.10 Schematic for gene integration or gene knockout, as facilitated by the pXP library. One-step PCR amplification of either the selection marker or gene cassette (Promoter-Gene-Terminiator-loxP-flanked Selection Marker) introduces approximately 50bp homology upstream (red) and downstream (green) of (i) the locus for integration or (ii) the gene being targeting for knockout. Homologus recombination allows the PCR-amplified fragments to be inserted selectively into the yeast genome, which include loxP-flanked markers for recycling.

transcription units to a coding sequence without introducing any restriction site sequences (Figure 1.13) (Mitchell et al., 2015).

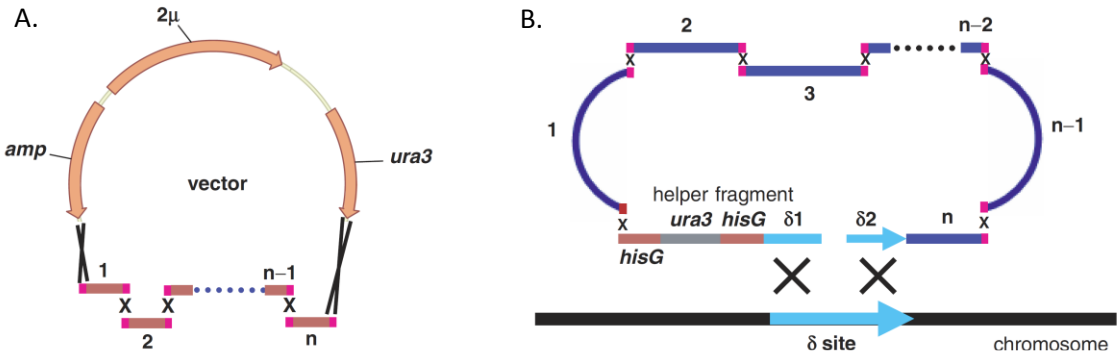


Figure 1.11 DNA Assembler method (Shao et al., 2009) creating one-step assembly of both (A) high-copy based expression and (B) delta-integration chromosomal-based expression, as illustrated for n number of DNA fragments using homologous recombination events (cross) in *S. cerevisiae*.

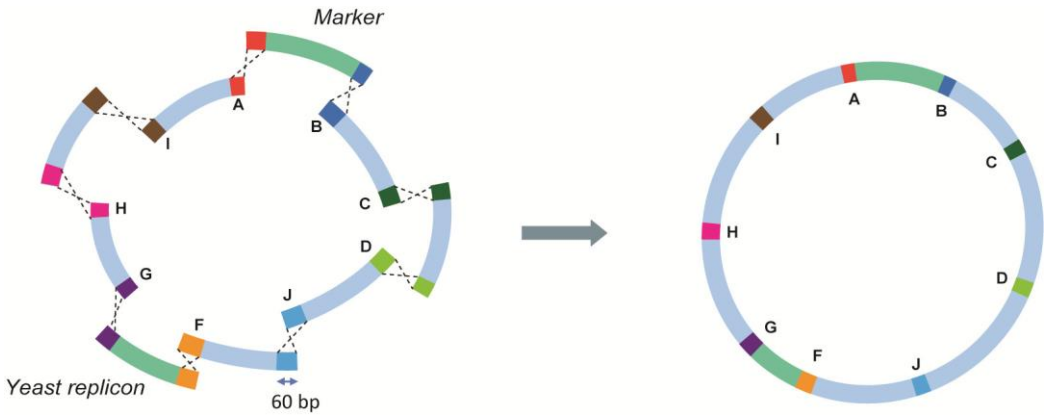
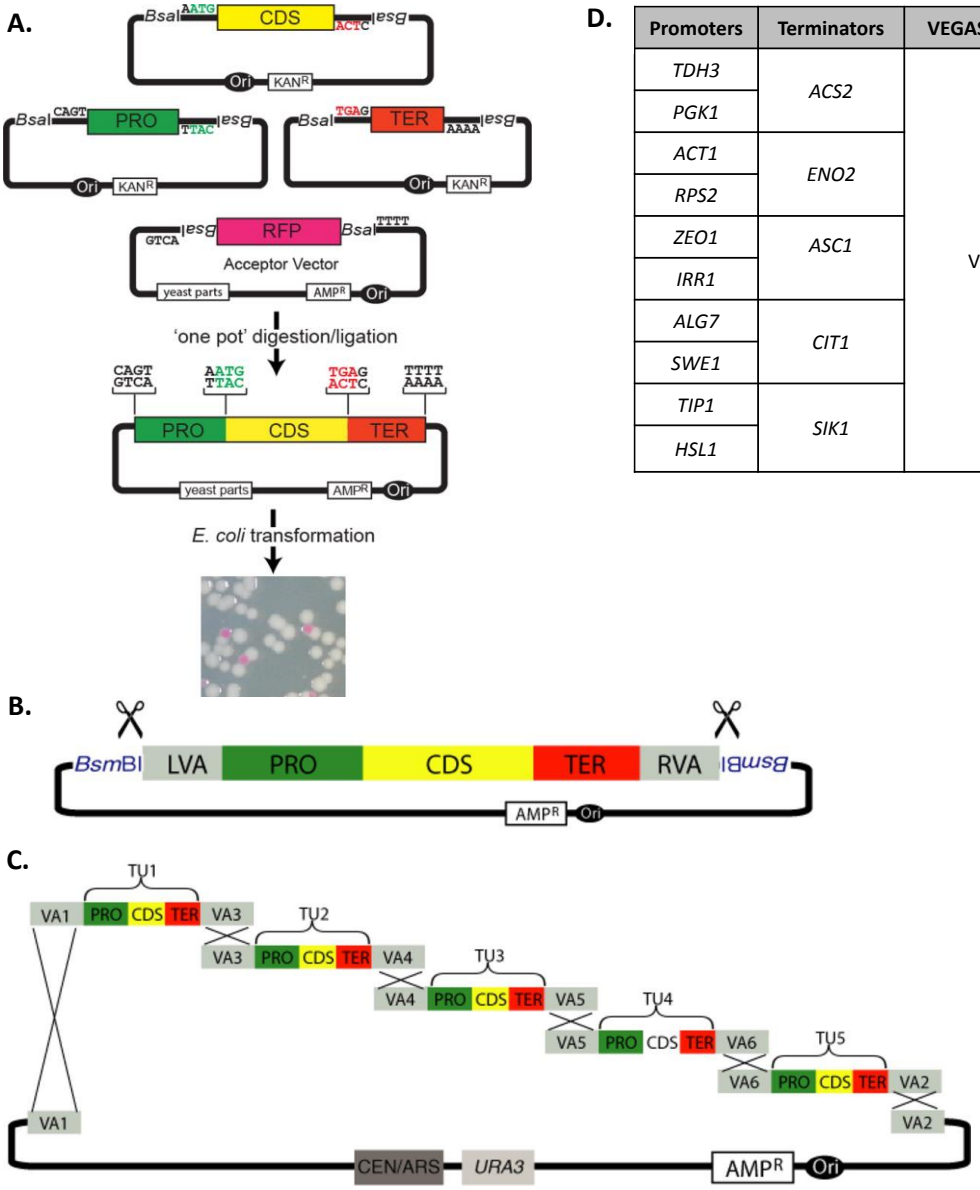


Figure 1.12 Plasmid construction using maximum available 60-mer homologous sequences (A-I) and backbone elements (Kuijpers et al., 2013). Two survival elements (Yeast replicon, Marker) are reassembled using homologous recombination (dashed crosses) in *S. cerevisiae*, reconstructing the full vector with a pathway of interest, with up to 7 genes.



D.

Promoters	Terminators	VEGAS Adapters
TDH3	*ACS2*	V1-V18
PGK1		
ACT1	*ENO2*	
RPS2		
ZEO1	*ASC1*	
IRR1		
ALG7	*CIT1*	
SWE1		
TIP1	*SIK1*	
HSL1		

Figure 1.13 Vector assembly by the versatile genetic assembly system (VEGAS) (Mitchell et al., 2015). (A) Promoter (PRO, green), gene (CDS, yellow), and terminator (TER, red) units cloned into modular constructs are treated by Type IIS restriction enzyme (RE) digestion/ligation using a red fluorescent protein (RFP) plasmid for appropriate selection of constructs by white/red screening. (B) Once each transcription unit (TU) is constructed, (C) all transcription units (TU1-TU5) can be digested and incubated with acceptor yGG backbone for recombination *in vivo* using VEGAS adapters (VAs) (homologous regions, labeled VA1-VA5). (D) A set of yGG acceptor backbone constructs have been developed with the listed promoters, terminators, and VAs to create plasmids with up to 17 transcription units.

In parallel to these *in vivo* methodologies, a very successful *in vitro* construction method has been that of Gibson Assembly (Gibson et al., 2009). This method also uses PCR-amplified fragments, but instead of using homologous recombination for joining and circularization of all fragments, Gibson Assembly utilizes three enzymes for reconstruction (Figure 1.14). First, a 5' exonuclease creates overhangs by chewing back the 5' end sequences. Using 15-20 bp of homology between fragments for annealing, a high efficiency DNA polymerase fills in gaps in order for the DNA ligase to fill in the nick left behind between the final base on each end, sealing the entire double-stranded vector construct. Combined, these strategies are able to provide the building blocks for rapid construction of a seemingly endless number of plasmids for *S. cerevisiae* with improved, fine-tuned expression capabilities.

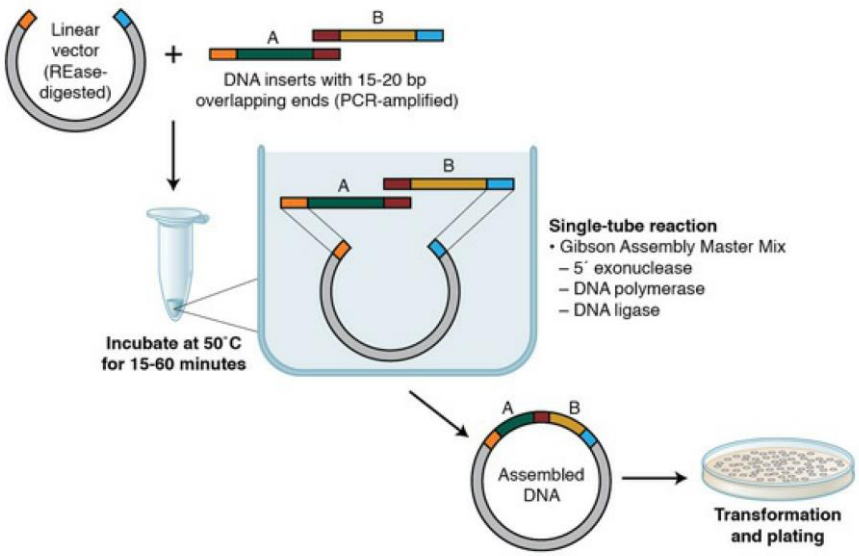


Figure 1.14 Gibson Assembly for *in vitro* plasmid construction, incubated at 50°C for up to 60 min (modified from New England Biolabs, Ipswich, MA). Restriction enzyme digested backbone is linearized and incubated with fragments for recombination, shown here for fragment A (green) and B (gold) with a minimum 18bp overlap (red). Both the backbone and fragments have complementary sequences (orange and light blue) which anneal along with the fragment overlapping region, followed by 5' exonuclease, DNA Polymerase, and DNA ligase activity to create the assembled DNA for transformation.

1.3.3 Metabolic engineering and systems biology

The constant evolution in metabolic engineering tools creates a new perspective to approaching strain development. As omics tools more readily provide information and data pertaining to global mechanisms, computational and mathematical modeling strategies gain accuracy which translates to faster metabolic engineering. Recent attempts to engineer *S. cerevisiae* for natural product synthesis using traditional metabolic engineering approaches illustrates the vast genetic engineering modifications available for incorporation (Jong et al., 2014; Krivoruchko and Nielsen, 2015; Pfleger et al., 2015; Tseng et al., 2009; Yang et al., 2015; Yun et al., 2015). In addition, systems biology is gaining momentum as an approach to gain global understanding of genetic interventions to inform better strain development and predictive strategies (Likic et al., 2011). This method aims to use computational and mathematical modeling to represent things that we know about the cell, whether it be genes that are conferring a specific reaction (Burgard et al., 2003; Schellenberger et al., 2011), or transcription level changes to expression (Kim and Reed, 2010). This strategy therefore relies heavily on the robustness of the description of the 'system,' which often times is the cell metabolic network. These genome-scale models (GEMs) become the full representation of cell activity in a computational framework (Tervo and Reed, 2014). With that in mind, these models are constantly changing to more closely represent the real-life observations found by experimental data (Zomorrodi and Maranas, 2010). The iterative nature that emerges from using both metabolic engineering and systems biology approaches together creates a new way of designing robust organisms (Tee et al., 2014). As flux data and other omics information is made available, modeling applications can tighten the networks to provide more accurate solutions, which in turn identifies more relevant

interventions that continue to guide improvements and new flux maps. This is apparent in two recent studies: the first using flux distributions to adjust the mathematical parameters and guide engineering of a microbial host for fatty acids (Ranganathan et al., 2012), and the second for introducing kinetic parameters into the stoichiometric genome-scale model to improve predicted optimizations (Chowdhury et al., 2014). The coupling of these two disciplines highlights the rapidly changing landscape for strain engineering that will expedite the development of production platforms in *S. cerevisiae*.

For computational algorithms, a few published optimizations are available through the toolbox COBRA 2.0 (Schellenberger et al., 2011). Major goals of the toolbox were to include more functionality, but also to make the algorithms usable in order to facilitate adoption by the scientific community, especially researchers not devoted to writing these large models and complex optimization programs. This MATLAB package introduced three *in silico* metabolic engineering tools: OptKnock, OptGene, and genetic design through local search (GDLS). The OptKnock algorithm is a bi-level mixed-integer linear program (MILP) designed to suggest knockout strategies for biochemical overproduction (Burgard et al., 2003). Constraints in flux distributions are designed around two possible cellular objectives: maximization of biomass, and the minimization of metabolic adjustment (MOMA). OptGene (Patil et al., 2005) instead represents genes extremely simplistically, assigning binary for the presence (1) or absence (0) of an associated reaction. The gene assignment can be input manually or designed randomly for any population (each individual model network with random genes). The networks can then be screened for an objective function (productivity, yield, etc.) and scored (high, low). The individuals who score high for the objective are 'mated' and then mutated randomly as previously

done (absent/present genes). This cycle is effectively iterated until the desired criteria is met. One benefit of OptGene over OptKnock is that non-linear objective functions can be optimized, which is useful for the function of strain productivity. Lastly, the GDLS method (Figure 1.15), also a mixed-integer linear program (MILP) like OptKnock, relies on simplification of the model (by clustering genes) in order to effectively search the space for perturbations that create a desirable outcome (target flux, etc.) (Lun et al., 2009). Overall, these algorithms behave similarly and provide useful interventions for guiding *in vivo* manipulations.

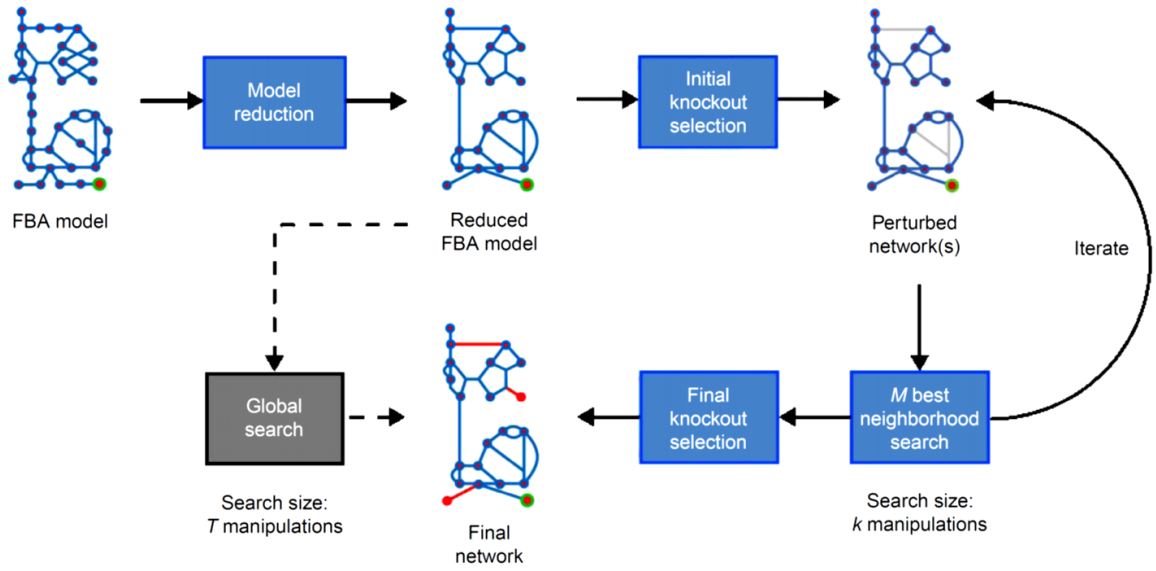


Figure 1.15 The overall process flow for GDLS method (Lun et al., 2009). The original 'wildtype' model (FBA model) is initially reduced to create a simplified network with fewer genes and reactions (Reduced FBA model) in order to initiate knockout selection. The program searches for the best manipulation strategies (*M* number of models) that differs by the parent model by *k* manipulations. The best perturbed networks get iterated for further knockout manipulations until no improvement is made or when the max *T* number of manipulations has been reached to achieve the final network.

Recently, a new algorithm called OptForce was designed to identify metabolic interventions similar to OptKnock, but instead of disrupting reactions entirely, the optimization seeks to identify necessary changes in flux distributions (Ranganathan et al., 2010). Identifying manipulations by altered flux values allows the user to screen reactions in the network relative to the wildtype strain, informing the user to either eliminate flux entirely, or instead simply increase or reduce fluxes. The analysis therefore facilitates a solution set which requires not only disruption strategies, but upregulation and / or downregulation activities to achieve optimal flux toward a target molecule. Use of this method has been validated for succinate production in *E. coli* (Ranganathan et al., 2010) as well as overproduction of fatty acids in *E. coli* (Ranganathan et al., 2012). OptForce has recently been optimized further by introducing kinetics where available for reactions in central carbon metabolism; this algorithm is known as k-OptForce (Chowdhury et al., 2014). This modification in the model network leads to very different solution sets relative to the original OptForce interventions, finding similar yield improvements as OptForce with less interventions. Improvements are possible by eliminating previous solutions which may have violated the introduced kinetic constraints as well as the limitations from metabolite concentrations (Chowdhury et al., 2014). More importantly, the new method is constructed as a mixed-integer nonlinear optimization program (MINLP) which paves the way for introduction of other information, including the possibility of dynamic temporal modeling and use of transcriptional regulation.

1.4 References

Agmon, N., Mitchell, L.A., Cai, Y., Ikushima, S., Chuang, J., Zheng, A., Choi, W.-J., Martin, J.A., Caravelli, K., Stracquadanio, G., Boeke, J.D., 2015. Yeast Golden Gate (yGG) for the Efficient Assembly of S. cerevisiae Transcription Units. ACS Synth. Biol. 4, 853–859. doi:10.1021/sb500372z

Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., 2005. Tuning genetic control through promoter engineering. Proceedings of the National Academy of Sciences of the United States of America 102, 12678 –12683. doi:10.1073/pnas.0504604102

Austin, M.B., Noel, J.P., 2003. The chalcone synthase superfamily of type III polyketide synthases. Nat. Prod. Rep. 20, 79–110. doi:10.1039/B100917F

Blazeck, J., Garg, R., Reed, B., Alper, H.S., 2012. Controlling promoter strength and regulation in Saccharomyces cerevisiae using synthetic hybrid promoters. Biotechnol. Bioeng. 109, 2884–2895. doi:10.1002/bit.24552

Borman, S., 2013. Ringing In New Drug Candidates. Chemical & Engineering News 91, 9.

Burgard, A.P., Pharkya, P., Maranas, C.D., 2003. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnology and Bioengineering 84, 647–657. doi:10.1002/bit.10803

Calo-Fernández, B., Martínez-Hurtado, J.L., 2012. Biosimilars: Company Strategies to Capture Value from the Biologics Market. Pharmaceuticals 5, 1393–1408. doi:10.3390/ph5121393

Chia, M., Schwartz, T.J., Shanks, B.H., Dumesic, J.A., 2012. Triacetic acid lactone as a potential biorenewable platform chemical. Green Chem. 14, 1850–1853. doi:10.1039/C2GC35343A

Chowdhury, A., Zomorrodi, A.R., Maranas, C.D., 2014. k-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design. PLoS Comput Biol 10, e1003487. doi:10.1371/journal.pcbi.1003487

Comfort, C., 2013. Dyslipidemia Market Will Grow to $31 Billion in 2022 - Decision Resources. Burlington, MA.

Crook, N.C., Freeman, E.S., Alper, H.S., 2011. Re-engineering multicloning sites for function and convenience. Nucleic Acids Res 39, e92. doi:10.1093/nar/gkr346

Da Silva, N.A., Srikrishnan, S., 2012. Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. FEMS Yeast Res. 12, 197–214. doi:10.1111/j.1567-1364.2011.00769.x

Engler, C., Gruetzner, R., Kandzia, R., Marillonnet, S., 2009. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. PLoS ONE 4, e5553. doi:10.1371/journal.pone.0005553

Engler, C., Kandzia, R., Marillonnet, S., 2008. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLoS ONE 3, e3647. doi:10.1371/journal.pone.0003647

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. doi:10.1002/yea.1824

Feher, M., Schmidt, J.M., 2003. Property Distributions: Differences between Drugs, Natural Products, and Molecules from Combinatorial Chemistry. J. Chem. Inf. Comput. Sci. 43, 218–227. doi:10.1021/ci0200467

Generally Recognized as Safe (GRAS), 1986. URL http://www.fda.gov/ (accessed 7.23.15).

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Meth 6, 343–345. doi:10.1038/nmeth.1318

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, 563–567.

Gokhale, R.S., Saxena, P., Chopra, T., Mohanty, D., 2007. Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. Nat Prod Rep 24, 267–277. doi:10.1039/b616817p

Herper, M., 2012. The Truly Staggering Cost Of Inventing New Drugs. Forbes. URL http://www.forbes.com/sites/matthewherper/2012/02/10/the-truly-staggering-cost-of-inventing-new-drugs/ (accessed 7.12.15).

Hertweck, C., 2009. The Biosynthetic Logic of Polyketide Diversity. Angewandte Chemie International Edition 48, 4688–4716. doi:10.1002/anie.200806121

Huigens III, R.W., Morrison, K.C., Hicklin, R.W., Flood Jr, T.A., Richter, M.F., Hergenrother, P.J., 2013. A ring-distortion strategy to construct stereochemically complex and structurally diverse compounds from natural products. Nat Chem 5, 195–202. doi:10.1038/nchem.1549

Johnson, L., 2012. Envisioning the Future of Biorenewables. Cultivation Corridor.

Jong, B.W. de, Shi, S., Siewers, V., Nielsen, J., 2014. Improved production of fatty acid ethyl esters in Saccharomyces cerevisiae through up-regulation of the ethanol degradation pathway and expression of the heterologous phosphoketolase pathway. Microbial Cell Factories 13, 39. doi:10.1186/1475-2859-13-39

Kim, J., Reed, J.L., 2010. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. BMC Systems Biology 4, 53. doi:10.1186/1752-0509-4-53

Krivoruchko, A., Nielsen, J., 2015. Production of natural products through metabolic engineering of Saccharomyces cerevisiae. Current Opinion in Biotechnology 35, 7–15. doi:10.1016/j.copbio.2014.12.004

Kuijpers, N.G., Solis-Escalante, D., Bosman, L., van den Broek, M., Pronk, J.T., Daran, J.-M., Daran-Lapujade, P., 2013. A versatile, efficient strategy for assembly of multi-fragment expression vectors in Saccharomyces cerevisiae using 60 bp synthetic recombination sequences. Microb Cell Fact 12, 47.

Langeveld, J.W.A., Dixon, J., Jaworski, J.F., 2010. Development Perspectives Of The Biobased Economy: A Review. Crop Science 50, S–142–S–151. doi:10.2135/cropsci2009.09.0529

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. doi:10.1016/j.ab.2009.07.010

Leibundgut, M., Maier, T., Jenni, S., Ban, N., 2008. The multienzyme architecture of eukaryotic fatty acid synthases. Curr. Opin. Struct. Biol. 18, 714–725. doi:10.1016/j.sbi.2008.09.008

Likic, V., McConville, M.J., Bacic, A., 2011. Systems Biology: The Next Frontier for Bioinformatics. Advances in Bioinformatics 2010, e268925. doi:10.1155/2010/268925

Lun, D.S., Rockwell, G., Guido, N.J., Baym, M., Kelner, J.A., Berger, B., Galagan, J.E., Church, G.M., 2009. Large-scale identification of genetic design strategies using local search. Molecular Systems Biology 5. doi:10.1038/msb.2009.57

Lussier, F.-X., Colatriano, D., Wiltshire, Z., Page, J.E., Martin, V.J.J., 2012. Engineering Microbes for Plant Polyketide Biosynthesis. Computational and Structural Biotechnology Journal 3, 1–11. doi:10.5936/csbj.201210020

Mander, L., Liu, H.-W., 2010. Comprehensive Natural Products II: Chemistry and Biology: 10 Volume Set, 1 edition. ed. Elsevier Science, Boston, MA.

Mitchell, L.A., Chuang, J., Agmon, N., Khunsriraksakul, C., Phillips, N.A., Cai, Y., Truong, D.M., Veerakumar, A., Wang, Y., Mayorga, M., Blomquist, P., Sadda, P., Trueheart, J., Boeke, J.D., 2015. Versatile genetic assembly system (VEGAS) to assemble pathways for expression in S. cerevisiae. Nucleic Acids Research. doi:10.1093/nar/gkv466

Newman, D.J., Cragg, G.M., 2007. Natural Products as Sources of New Drugs over the Last 25 Years. Journal of Natural Products 70, 461–477. doi:10.1021/np068054v

Nikolau, B.J., Perera, M.A.D.N., Brachova, L., Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. The Plant Journal 54, 536–545. doi:10.1111/j.1365-313X.2008.03484.x

Patil, K.R., Rocha, I., Förster, J., Nielsen, J., 2005. Evolutionary programming as a platform for in silico metabolic engineering. BMC bioinformatics 6, 308.

Pfeifer, B.A., Khosla, C., 2001. Biosynthesis of Polyketides in Heterologous Hosts. Microbiol. Mol. Biol. Rev. 65, 106–118. doi:10.1128/MMBR.65.1.106-118.2001

Pfleger, B.F., Gossing, M., Nielsen, J., 2015. Metabolic engineering strategies for microbial synthesis of oleochemicals. Metab. Eng. 29, 1–11. doi:10.1016/j.ymben.2015.01.009

Pronk, J.T., Yde Steensma, H., Van Dijken, J.P., 1996. Pyruvate Metabolism in Saccharomyces cerevisiae. Yeast 12, 1607–1633. doi:10.1002/(SICI)1097-0061(199612)12:16<1607::AID-YEA70>3.0.CO;2-4

Ranganathan, S., Suthers, P.F., Maranas, C.D., 2010. OptForce: An Optimization Procedure for Identifying All Genetic Manipulations Leading to Targeted Overproductions. PLoS Comput Biol 6. doi:10.1371/journal.pcbi.1000744

Ranganathan, S., Tee, T.W., Chowdhury, A., Zomorrodi, A.R., Yoon, J.M., Fu, Y., Shanks, J.V., Maranas, C.D., 2012. An integrated computational and experimental study for overproducing fatty acids in Escherichia coli. Metab. Eng. 14, 687–704. doi:10.1016/j.ymben.2012.08.008

Regnier, E., 2007. Oil and energy price volatility. Energy Economics 29, 405–427. doi:10.1016/j.eneco.2005.11.003

Rouhi, A.M., 2003. Rediscovering Natural Products. Chemical & Engineering News 81, 77–91.

Ruohonen, L., Aalto, M.K., Keränen, S., 1995. Modifications to the ADH1 promoter of Saccharomyces cerevisiae for efficient production of heterologous proteins. Journal of Biotechnology 39, 193–203. doi:16/0168-1656(95)00024-K

Saccharomyces cerevisiae Genome Overview, 2015. URL http://www.yeastgenome.org/ (accessed 7.23.15).

Saunders, L.P., Bowman, M.J., Mertens, J.A., Silva, N.A.D., Hector, R.E., 2015. Triacetic acid lactone production in industrial Saccharomyces yeast strains. J Ind Microbiol Biotechnol 1–11. doi:10.1007/s10295-015-1596-7

Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J., Feist, A.M., Zielinski, D., Bordbar, A., Lewis, N., Rahmanian, S., Kang, J., Hyduke, D.R., Palsson, B., 2011. COBRA Toolbox 2.0. Protocol Exchange 6, 1290–1307. doi:10.1038/protex.2011.234

Shao, Z., Zhao, H., Zhao, H., 2009. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucl. Acids Res. 37, e16–e16. doi:10.1093/nar/gkn991

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. doi:10.1002/yea.2930

Tee, T.W., Chowdhury, A., Maranas, C.D., Shanks, J.V., 2014. Systems metabolic engineering design: Fatty acid production as an emerging case study. Biotechnology and bioengineering 111, 849–857.

Tervo, C.J., Reed, J.L., 2014. Expanding metabolic engineering algorithms using feasible space and shadow price constraint modules. Metabolic Engineering Communications 1, 1–11. doi:10.1016/j.meteno.2014.06.001

Tseng, H.-C., Martin, C.H., Nielsen, D.R., Prather, K.L.J., 2009. Metabolic Engineering of Escherichia coli for Enhanced Production of (R)- and (S)-3-Hydroxybutyrate. Appl. Environ. Microbiol. 75, 3137–3145. doi:10.1128/AEM.02667-08

Vennestrøm, P.N.R., Osmundsen, C.M., Christensen, C.H., Taarning, E., 2011. Beyond Petrochemicals: The Renewable Chemicals Industry. Angew. Chem. Int. Ed. 50, 10502–10509. doi:10.1002/anie.201102117

White, S.W., Zheng, J., Zhang, Y.-M., Rock, C.O., 2005. The Structural Biology of Type II Fatty Acid Biosynthesis. Annual Review of Biochemistry 74, 791–831. doi:10.1146/annurev.biochem.74.082803.133524

Yang, Y., Lin, Y., Li, L., Linhardt, R.J., Yan, Y., 2015. Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products. Metabolic Engineering. doi:10.1016/j.ymben.2015.03.018

Yun, E.J., Kwak, S., Kim, S.R., Park, Y.-C., Jin, Y.-S., Kim, K.H., 2015. Production of (S)-3-hydroxybutyrate by metabolically engineered Saccharomyces cerevisiae. Journal of Biotechnology 209, 23–30. doi:10.1016/j.jbiotec.2015.05.017

Zomorrodi, A.R., Maranas, C.D., 2010. Improving the iMM904 S. cerevisiae metabolic model using essentiality and synthetic lethality data. BMC Syst Biol 4, 178. doi:10.1186/1752-0509-4-178

Chapter 2:

Metabolic engineering of *Saccharomyces cerevisiae* for the production of triacetic acid lactone

2.1 Abstract

Biobased chemicals have become attractive replacements for their fossil-fuel counterparts. Recent studies have shown triacetic acid lactone (TAL) to be a promising candidate, capable of undergoing chemical conversion to sorbic acid and other valuable intermediates. In this study, *Saccharomyces cerevisiae* was engineered for the high-level production of TAL by overexpression of the *Gerbera hybrida* 2-pyrone synthase (2-PS) and systematic engineering of the yeast metabolic pathways. Pathway analysis and a computational approach were employed to target increases in cofactor and precursor pools to improve TAL synthesis. The pathways engineered include those for energy storage and generation, pentose biosynthesis, gluconeogenesis, lipid biosynthesis and regulation, cofactor transport, and fermentative capacity. Seventeen genes were selected for disruption and independently screened for their effect on TAL production; combinations of knockouts were then evaluated. A combination of the pathway engineering and optimal culture parameters led to a 37-fold increase in titer to 2.2 g/L and a 50-fold increase in yield to 0.13 (g/g glucose). These values are the highest reported in the literature, and provide a 3-fold improvement in yield over previous reports using *S. cerevisiae*. Identification of these metabolic bottlenecks provides a strategy for overproduction of other acetyl-CoA-dependent products in yeast.

2.2 Introduction

The demand for renewable, biobased products as an alternative to those of petroleum origin is rapidly reshaping the scope of biotechnology. To date, the chemical industry has relied heavily on petroleum resources and chemical catalysis for the production of a wide array of end products (Nikolau et al., 2008). Using microorganisms and enzymes, for example to ferment lignocellulosic biomass (Hamelinck et al., 2005) or to produce novel compounds (Westfall et al., 2012), is an essential component to establishing a biorenewables infrastructure. By pulling together biocatalysis and chemical catalysis, valuable chemical intermediates can similarly be synthesized into numerous end products, with biorenewable feedstocks replacing those from fossil-carbon (Nikolau et al., 2008).

Triacetic acid lactone (TAL), also known as 4-hydroxy-6-methyl-2-pyrone, can be used in the production of polyketide anthraquinones by deriving the tetraphenolic acid from TAL's methyl ether (Evans et al., 1979). Recent efforts have uncovered novel catalysis methods expanding the utility of the TAL moiety. (Chia et al., 2012) illustrated that TAL can undergo chemical upgrading through a variety of reactions to generate attractive end products like sorbic acid and 1,3-pentadiene that are precursors to preservatives and plastics, respectively. The ability to readily convert TAL into numerous chemicals that are currently produced using petroleum feedstocks (Chia et al., 2012) makes TAL a relevant precursor candidate for microbial synthesis. While TAL is commercially available, the chemical manufacture of TAL significantly limits its potential for upgrading to low-cost commodity chemicals due to the presence of various byproducts and the requirement of a high-cost multi-stage process (Goel and Ram, 2009;

Taeschler, 2010). Therefore developing a biological route for converting sugars directly to TAL will facilitate its use industrially.

TAL is natively produced by the plant *Gerbera hybrida*, where it is further modified to generate anti-pathogen compounds gerberin and parsorboside (Eckermann et al., 1998). The type III polyketide synthase (PKS) 2-pyrone synthase (2-PS) encoded by the *g2ps1* gene, produces TAL via two iterative decarboxylation/condensation reactions using a starter acetyl-CoA and two extender malonyl-CoA molecules (Austin and Noel, 2003), both of which are common metabolites in bacteria and eukaryotes. Recently, Tang et al. (2013) reported the use of directed evolution to identify 2-PS mutants yielding higher expression as identified by a fluorescence reporter assay. TAL can also be produced from non-native synthases. A closely related type III PKS from *Medicago sativa*, the *chs2*-encoded chalcone synthase (CHS), was shown to produce TAL following three point mutations (T197L/G256L/S338I) resulting in altered specificity for acetyl-CoA and reduced elongations (Jez et al., 2000). The type I PKS 6-methylsalicylic acid synthase (6-MSA) from *Penicillium patulum* was modified to inactivate the ketoreductase domain and produce TAL *in vivo* (Richardson et al., 1999; Xie et al., 2006), and the *Brevibacterium ammoniagenes* fatty acid synthase B (FAS-B) was also rationally designed with an inactive ketoreductase domain (Y2226F) to produce TAL (Zha et al., 2004). A drawback to these latter two systems is that they require a phosphopantetheinyl transferase for activation, while the 2-PS and CHS type III PKSs do not (Austin and Noel, 2003).

Triacetic acid lactone has been reported to have negative effects on microbial cell growth. Specifically, TAL is extremely toxic to *Escherichia coli* with an observed reduction in growth rate

of 25% and 90% in 10 mM and 20 mM TAL, respectively, with no viable growth detected at 50 mM TAL (Laura Jarboe, personal communication). However, there was no effect on the growth rate of *S. cerevisiae* at concentrations of up to 200 mM TAL (the higher concentrations exceeded the solubility limit). In the *in vivo* studies on TAL biosynthesis in *E. coli*, concentrations were much lower than 50mM, and maximal levels were not achieved until after exponential growth (Xie et al., 2006; Tang et al., 2013). Tang et al. (2013) recently reported titers of 2.06 g/L (16.4 mM), corresponding to a yield of 0.102 g/g glycerol (22% of theoretical) using an engineered 2-PS variant (L202G/M259L/L261N) in *E. coli*. This was a significant increase over the previously reported levels achieved in *E. coli* (0.47 g/L, 0.6% of theoretical yield on glucose) using the wildtype 2-PS and under fermenter-controlled conditions (Xie et al., 2006). In this latter study *S. cerevisiae* was also evaluated and produced titers of 0.37 g/L (1.0% of theoretical) using the wild-type 2-PS enzyme and 1.8 g/L (8.9% of theoretical) using an Y1572F variant 6-MSA synthase. In both cases, this was after 168 h of fed-batch fermentation. Yields thus remain relatively low during microbial production of TAL.

The above studies considered the enzyme and culture conditions used for TAL synthesis. The current study focuses on the engineering of *Saccharomyces cerevisiae* to enhance TAL production via targeted improvements in the pyruvate, acetyl-CoA, and NAD^+ pools. To our knowledge this is the only study to consider metabolic pathway engineering to increase the synthesis of triacetic acid lactone. An emphasis was placed on yeast as the host based on the TAL sensitivities reported in *E. coli*. After establishing an initial expression system, metabolic pathways were engineered for improved TAL biosynthesis. The computational tool OptKnock (Schellenberger et al., 2011) available in the COBRA Toolbox was used to aid the pathway

engineering by identifying additional gene deletions for improved flux toward acetyl-CoA and malonyl-CoA. These strain modifications were evaluated independently and in combination to determine their effect on TAL production in *S. cerevisiae*, and final cultivations were done in fed-batch mode to increase titers. The metabolic pathway interventions shown to generate higher levels of TAL should prove directly applicable to the increased synthesis of other polyketide products.

2.3 Materials and Methods

2.3.1 Strains and plasmids

Escherichia coli strain XL1-Blue (Stratagene, Santa Clara, CA) was used for amplification of plasmids. *S. cerevisiae* strains BY4741 and BY4741Δ*trp1* (Open Biosystems, Huntsville, AL) and BJ5464 (Jones, 1991) were used as the base strains for TAL production, and subsequent engineered strains were derived from these (Table 2.1). A variety of additional single gene knockouts from the BY4741 knockout library (Open Biosystems) were also utilized for preliminary strain screening. Details on the strain construction can be found in the Supplementary Data.

pXP vectors (Fang et al., 2011) were used to amplify loxP-flanked selectable markers *LEU2-d8, TRP1, MET15,* or *HIS3* for gene disruption. Using the corresponding primer sets (Table 2.2), the markers were amplified with approximately 50bp homology to the target gene location. Yeast cells were transformed as previously described (Gietz et al., 1992; Hill et al., 1991). Cell colonies were allowed to grow on selective plates for 3-5 days, and gene disruptions were verified following amplification of genomic DNA using Taq polymerase (New England Biolabs, Ipswich, MA), and confirming the presence of a loxP-flanked selection marker.

The pXP218 and pXP842 yeast shuttle vectors (Fang et al., 2011; Shen et al., 2012) were used to carry the TAL-synthesizing genes. These 2µ-based vectors harbor a *PGK1* or *ADH2* promoter, respectively, the *CYC1* terminator, and a loxP-flanked *URA3* selection marker. The *g2ps1* gene encoding the 2-PS from *Gerbera hybrida* was PCR amplified from the pHIS8 cassette (Jez et al., 2000). Following *SpeI* and *XhoI* digestion, the gene was inserted into pXP218 and pXP842 using the Rapid DNA Ligation Kit (Thermo Scientific, Waltham, MA) to construct pXP218-

Table 2.1. List of plasmids and strains

Plasmids	Description	Source
pXP218	2μ vector, *PGK1* promoter, *CYC1* terminator, *URA3* selectable marker	F. Fang (2011)
pXP214	2μ vector, *PGK1* promoter, *CYC1* terminator, *MET15* selectable marker	F. Fang (2011)
pXP216	2μ vector, *PGK1* promoter, *CYC1* terminator, *TRP1* selectable marker	F. Fang (2011)
pXP220	2μ vector, *PGK1* promoter, *CYC1* terminator, *HIS3* selectable marker	F. Fang (2011)
pXP841	2μ vector, *ADH2* promoter, *CYC1* terminator, *LEU2-d8* selectable marker	M. Shen (2012)
pXP842	2μ vector, *ADH2* promoter, *CYC1* terminator, *URA3* selectable marker	M. Shen (2012)
pXP218-2PS	pXP218 harboring the *g2ps1* insertion encoding *G. hybrida* 2-pyrone synthase	This study
pXP842-2PS	pXP842 harboring the *g2ps1* insertion encoding *G. hybrida* 2-pyrone synthase	This study
pXP842-[3X]CHS	pXP842 harboring the T197L/G256L/S338I CHS encoding the *M. sativa* chalcone synthase mutant	This study
pXP842-[1X]6MSAS	pXP842 harboring the Y1572F 6-MSAS encoding the *P. patulum* 6-MSA synthase mutant	This study
Strains	**Description**	**Source**
BY4741	*MAT*a *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Open Biosystems
BY4741Δ*trp1*	BY4741 *trp1::KanMX*	Open Biosystems
BJ5464	*MAT*α *ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL*	E.W. Jones (1991)
BYAN	BY4741 with integrated copy of *Aspergillus nidulans npgA*, *ADH2* promoter	This study
BY4741Δ*pyc1*	BY4741 *pyc1::KanMX*	Open Biosystems
BY4741Δ*pyc2*	BY4741 *pyc2::KanMX*	Open Biosystems
BY4741Δ*gpd1*	BY4741 *gpd1::KanMX*	Open Biosystems
BY4741Δ*pgm1*	BY4741 *pgm1::KanMX*	Open Biosystems
BY4741Δ*pgm2*	BY4741 *pgm2::KanMX*	Open Biosystems
BY4741Δ*nte1*	BY4741 *nte1::KanMX*	Open Biosystems
BY4741Δ*mls1*	BY4741 *mls1::KanMX*	Open Biosystems
BY4741Δ*pck1*	BY4741 *pck1::KanMX*	Open Biosystems
BY4741Δ*zwf1*	BY4741 *zwf1::KanMX*	Open Biosystems
BY4741Δ*yia6*	BY4741 *yia6::KanMX*	Open Biosystems
BY4741Δ*gsy1*	BY4741 *gsy1::KanMX*	Open Biosystems
BY4741Δ*gsy2*	BY4741 *gsy2::KanMX*	Open Biosystems
BY4741Δ*fbp1*	BY4741 *fbp1::KanMX*	Open Biosystems
BYtΔ*adh1*	BY4741Δ*trp1 adh1::LEU2*	This study
BYtΔ*pep4*	BY4741Δ*trp1 pep4::LEU2*	This study
BYtΔ*prb1*	BY4741Δ*trp1 prb1::MET15*	This study
BYtΔ*pep4*Δ*prb1*	BY4741Δ*trp1 pep4::LEU2 prb1::MET15*	This study
BYtΔ*prb1*Δ*pyc2*	BY4741Δ*trp1 prb1::MET15 pyc2::LEU2*	This study
BYtΔ*prb1*Δ*pyc2*Δ*yia6*	BY4741Δ*trp1 prb1::MET15 pyc2::LEU2 yia6::TRP1*	This study
BYtΔ*prb1*Δ*pyc2*Δ*nte1*	BY4741Δ*trp1 prb1::MET15 pyc2::LEU2 nte1::TRP1*	This study
BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1*	BY4741Δ*trp1 prb1::MET15 pyc2::LEU2 yia6::HIS3 nte1::TRP1*	This study
BJΔ*pyc2*	BJ5464 *pyc2::TRP1*	This study
BJΔ*pyc2*Δ*nte1*	BJ5464 *pyc2::TRP1 nte1::LEU2*	This study

Table 2.2 List of primers

Primer Name	Primer Sequence
2PS-F	TCCCACTAGTAAAAAAATGCATCATCATCATCATCATGGATCTTTACTCATCCGATGTGGAG
2PS-R	TGGTAATCCTCGAGTCAGTTTCCATTGGCAACCGCAGCAGTAACG
3xCHS-F	TCCCACTAGTAAAAAAATGCATCATCATCATCATCATGTGAGTG
3XCHS-R	TGGTAATCCTCGAGTTATATAGCCACGCTACGCAAAACAACTG
6MSASm-F	CGTCGGGTTCACTGGTCAAGCTTCCTTCGGCAGTGGTAACGCCTTCTTGG
6MSASm-R	CCAAGAAGGCGTTACCACTGCCGAAGGAAGCTTGACCAGTGAACCCGACG
6MSASmCh-F	CCTCTTGTGGAAACCTCGTCGGG
6MSASmCH-R	CCTAGGCGCGCACGGTGGGTGGC
PRB1KO-F	CAAACGATAGTGAAGAGGGACTCCGACTTGTAACCTCGAGACGCCTAAGGGGTCGACTCTAGAGGATCCCCGGG
PRB1KO-R	CGTCCCGTTATATTGGAGTTCTTCCCATACAAACTTAAGAGTCCAATTAGGTTGAATTCGAGCTCGGTACCCGGG
PRB1Ch-F	GACTTGTAACCTCGAGACGCCTAAG
PRB1Ch-R	CGTCCCGTTATATTGGAGTTCTTCC
ADH1KO-F	CGGCATGCCGGTAGAGGTGTGGTCAATAAGAGCGACCTCATGCTATACCGGTCGACTCTAGAGGATCCCCGGG
ADH1KO-R	GCACAATATTTCAAGCTATACCAAGCATACAATCAACTATCTCATATACAGTTGAATTCGAGCTCGGTACCCGGG
ADH1Ch-F	CGGCATGCCGGTAGAGGTGTGG
ADH1-Ch-R	CCAAGCATACAATCAACTATC
PYC2KO-F	GGGATGGGGGTAGGGTTTCTTCTTCTAGGACAACCAACAAATCTGATGCGGGTCGACTCTAGAGGATCCCCGGG
PYC2KO-R	GGCCGGTCTTAGGGACAATTTCAGTTTGCTCGGCGAAAAGAATAAGATCTGTTGAATTCGAGCTCGGTACCCGGG
PYC2Ch-F	CAGCTTGGTTTGGCGAGTGACATT
PYC2Ch-R	GGCGCAGGGAGTGTTAAGTAAGAAGTAC
YIA6KO-F	CAACACTGAAGGCAGCAGTCTTTAGCACAATGACACAGACTGATAATCC GGTCGACTCTAGAGGATCCCCGGG
YIA6KO-R	CTCAATGTGTGGCGTGTGTGAATTTTAAGGCTTTAAATTACCATAGTGCGTTGAATTCGAGCTCGGTACCCGGG
YIA6Ch-F	GTGAGATCAACACTGAAGGCAGCAG
YIA6Ch-R	CTCAATGTGTGGCGTGTGTG
NTE1KO-F	CGGTCACCAAGGCGTTGGAAGATAGAAATCTATCGGTCGACTCTAGAGGATCCCCGGG
NTE1KO-R	GAAAGGCTTGGCTCTAGTTTTTGCATGCGTTGAATTCGAGCTCGGTACCCGGG
NTE1Ch-F	CTGGGAAACTTGACACTTTTCTCG
NTE1Ch-R	GATAAAAGCACGCGACCGAAAAGG
NpgA-F	CACAGAGTTGTATTTGCGCTTCTGAGCGATGCTTCCGAGATTGTTGAAGCAGCAAAAACGTAGGGGCAAACAAAC
FF2325	GATTATTGAAGAGGGATGCGTTTGGTACAATAAAAAACATAGGTTCCCAAACCGAATTCGAGCTCGGTACCCGGG

2PS and pXP842-2PS. Similar methods were used to introduce the altered *chs2* gene encoding the T197L/G256L/S338I mutant of chalcone synthase (CHS) from *Medicago sativa* (Jez et al., 2000) and a Y1572F mutant of *6-MSAS* from *Penicillium patulum* into pXP842 (Figure 2.1). Plasmid recovery was performed using the GeneJet™ Plasmid Miniprep Kit (Thermo Scientific, Waltham, MA) and DNA sequence analysis confirmed the correct sequence of all PCR-amplified inserts (GeneWiz, South Plainfield, NJ; Eton Biosciences, San Diego, CA).

The KOD Hot-start polymerase (EMD Chemicals, San Diego, CA) was used in PCR reactions for plasmid constructs and knockout strains. PCR for site-directed mutagenesis of 6-MSAS was performed using *Pfu*Ultra II Fusion HS DNA Polymerase (Stratagene). Restriction enzymes, T4 DNA ligase, Taq DNA polymerase, and deoxynucleotides were purchased from New England Biolabs. Oligonucleotide primers were purchased from IDT DNA (San Diego, CA).

2.3.2 Media and cultivation

Luria-Bertani (LB) media was used for proliferation of XL1-Blue cells with 150 mg/L ampicillin for selection of plasmid-containing strains (Sambrook and Russell, 2001). Complex YPD media (0.5%, 1%, or 2% dextrose; 1% Bacto yeast extract; 2 % Bacto peptone), selective SDC(A) media (1% dextrose, 0.67% yeast nitrogen base, 0.5% Bacto casamino acids, 0.5% ammonium sulfate and 100 mg/L adenine), selective SDC(A,T) media (SDC(A) with 100 mg/L tryptophan), and minimal selective SD media (2% dextrose, 0.67% yeast nitrogen base, 0.5% ammonium sulfate) supplemented with 100 mg/L adenine, 100 mg/L tryptophan, 100 mg/L histidine, and 150 mg/L leucine were used for cultivation of yeast strains. Strains harboring the pXP218-based vectors

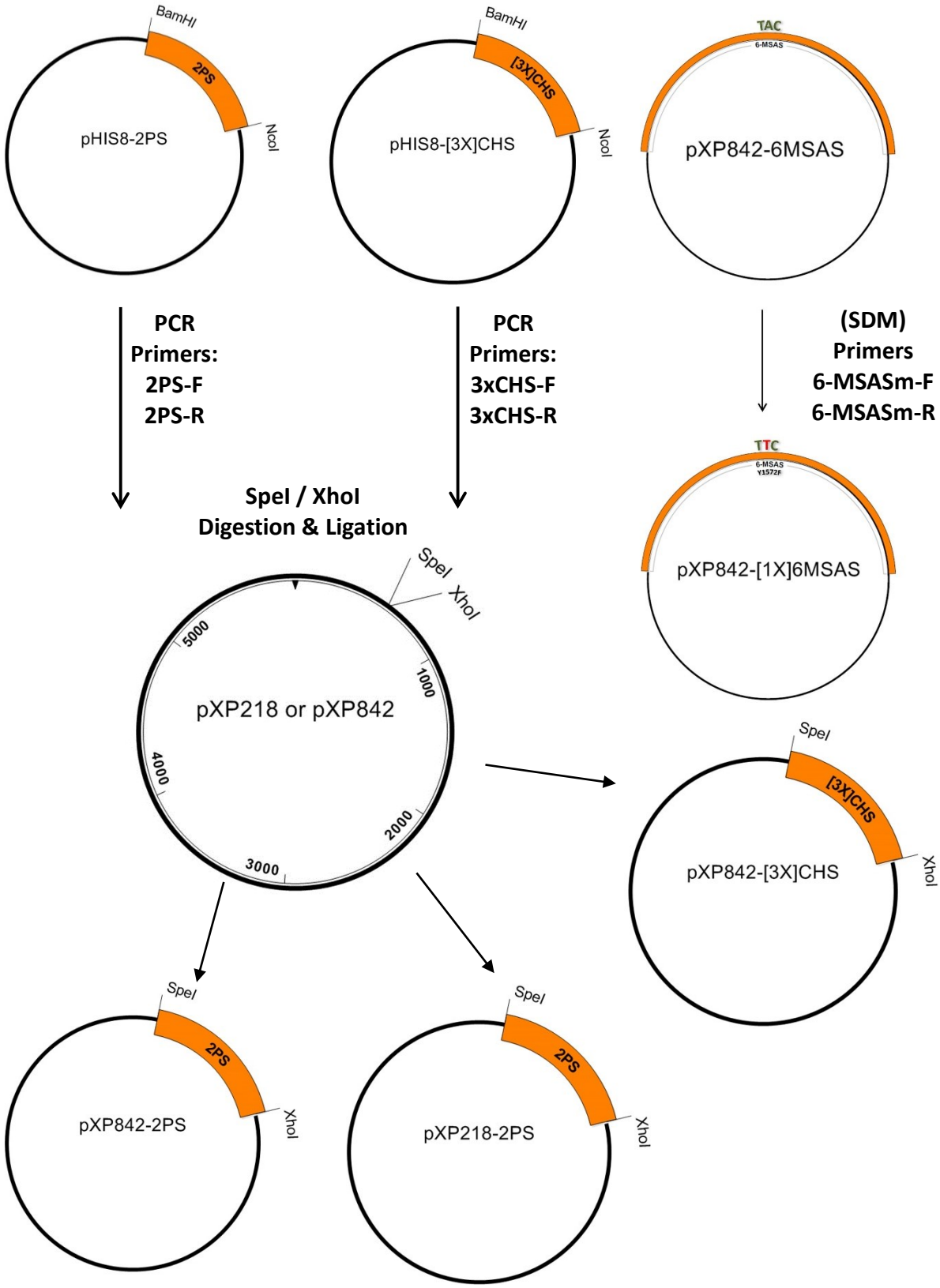


Figure 2.1. Strategy for plasmid construction of pXP218-2PS, pXP842-2PS, pXP842-[3X]CHS, pXP842-[1X]6MSAS.

(*PGK1* promoter) were grown in selective media. Strains harboring the pXP842-based vectors (*ADH2* promoter) were grown in YPD medium.

S. cerevisiae strains were grown for 16 h overnight in 5 mL selective SDC(A) media in an air shaker (New Brunswick Scientific) at 250 rpm and 30°C, and used to inoculate 5 mL tube or 50 mL flask cultures to an initial cell density (OD_{600}) of 0.3 (Shimadzu UV-2450 UV-VIS Spectrophotometer, Columbia, MD). A correlation factor was used to convert OD to dry cell weight per liter (1 OD = 0.43 gDCW/L). During cultivation, samples were taken at 24, 48, and/or 72 h, cell densities were determined, and the samples were centrifuged at 3,000 rpm (2,600 g) for 5 min at 4°C (Beckman GS-6R Centrifuge, Brea, CA). The supernatants were stored at 4°C for HPLC analysis of TAL levels in the culture broth.

2.3.3 Plasmid stability

The yeast strain was cultivated in 1% YPD medium for 48 h, diluted with sterile water, and plated onto YPD plates. Approximately 100 colonies were transferred to selective SDC-A plates and non-selective YPD plates (to verify viability). The percent plasmid-containing cells was determined as the number of colonies on the SDC-A plates divided by the number of viable colonies transferred. Three independent experiments were performed.

2.3.4 Glucose-limited fed-batch fermentations

A New Brunswick BioFlo III system equipped with a 2.5 L working capacity vessel was employed for fed-batch operation. Initial 5mL overnight cultures were used to inoculate a 50mL shake flask culture that was used as the seed inoculum. Media was YPD containing 1% glucose.

Following inoculation to an OD_{600} of 0.3 (V=1.5L), the fermentor pH was controlled at pH 6 by automatic supply of either 6M sodium hydroxide or 6M hydrochloric acid. Agitation speed was 400 rpm and sparged air was initially supplied at 0.2 vvm. To maintain the dissolved oxygen levels above 20%, the aeration rate was gradually increased to a maximum of 0.8 vvm. Following 12 h of batch cultivation in the fermentor, a glucose feed (3.6M) was initiated. The speed for the glucose feed pump was varied to maintain DO levels at 20% with a constant aeration rate of 0.8 vvm and a maximum glucose pump speed of 1.5 mL/h. The fermentation was allowed to proceed for a total of 120 h, leading to a final culture volume of approximately 1.7 L.

2.3.5 HPLC assay

The concentration of triacetic acid lactone was measured by HPLC using a Shimadzu HPLC system: LC-10AT pumps (Shimadzu), UV-VIS detector (SPD-10A VP, Shimadzu), Zorbax SB-C18 reversed-phase column (2.1x150 mm, Agilent Technologies). Acetonitrile buffered in 1% acetic acid was used as the mobile phase, while HPLC grade water buffered in 1% acetic acid was used as the aqueous phase. A gradient program using a 95-85% Pump B gradient (H_2O with 1% acetic acid) provided an elution time of approximately 12 minutes (flow rate 0.25 mL/min, column temperature 25°C).

2.3.6 Yield calculations

Actual yield was calculated as the total TAL formed divided by the total glucose supplied to the culture at inoculation. Maximum theoretical yield was calculated based on the metabolic reactions from glucose to TAL assuming all glucose was converted to product (i.e., no growth) and was determined to be 0.47 g/g glucose. (The theoretical maximum on glycerol was also found

to be 0.46 g/g glycerol.) Therefore, the percent of the theoretical yield of TAL on glucose was calculated as: [TAL (g/L) / Glucose Fed (g/L) / 0.47 (g/g) * 100%]. Although complex YPD medium was used, it does not support significant cell growth in the absence of glucose (< 6% of total growth). Therefore, other carbon sources in YPD should have little effect on the calculated yields.

2.4 Results and Discussion

2.4.1 Resistance of *Saccharomyces cerevisiae* to TAL

Previous research has demonstrated the high tolerance of *S. cerevisiae* to TAL relative to *E. coli*. To further explore the sensitivity of this yeast to TAL, we determined the effects of two TAL concentrations on yeast growth at a pH of 5.0 and 6.0. We first supplemented with increasing amounts of solid TAL and determined the solubility limits at 30°C, approximately 9.0 g/L in water and 6.5 g/L in SDC(A) and YPD media. A significant drop in solubility was observed when temperature was reduced, with a TAL solubility of 2.8 g/L at 4°C in SDC(A) medium.

To determine the effects on growth, *S. cerevisiae* strain BY4741 was cultivated in SDC(A) or complex YPD medium containing 1% glucose and spiked with TAL at 5 or 10 g/L. Medium pH was then adjusted to 5.0 or 6.0. Insoluble TAL (in the 10 g/L culture) was allowed to settle prior to measuring cell density. Growth rates and final cell densities (48 h) were determined, and are shown in Table 2.3. Increasing TAL concentration reduced both growth rate and final cell density; this effect was more pronounced at the lower pH and in SDC(A) medium. In the complex YPD medium, minimal decreases were observed, particularly at pH 6.0. YPD has a natural buffering capacity around pH 6.5 relative to SDC(A); in SDC(A), a drop in pH (up to 1 pH unit) occurs during cultivation. TAL has a pKa of 4.9 (Moreno-Mañas et al., 1988), and the protonated form will increase at lower pH. Protonated forms of fatty acids are incorporated at higher levels in the membrane resulting in membrane leakage in *S. cerevisiae* and reduced growth rate (Liu et al., 2013). Related lactones can also be incorporated into the membrane and detrimentally affect

Table 2.3. Growth rates and final cell densities for strain BY4741 cultured in the presence of varying concentrations of TAL in two media at a pH of 5.0 or 6.0. Values are mean ± standard deviation (n=3).

Media	Starting pH	Maximum Specific Growth Rate (h^{-1})			Final Cell Density at 48 h (g DCW L^{-1})		
		0 g/L TAL	5 g/L TAL	10 g/L TAL*	0 g/L TAL	5 g/L TAL	10 g/L TAL*
SDC (A,T,U)	5.0	0.42 ± 0.02	0.33 ± 0.01	0.28 ± 0.02	6.6 ± 0.5	4.3 ± 0.3	2.9 ± 0.5
	6.0	0.43 ± 0.01	0.38 ± 0.01	0.33 ± 0.01	6.2 ± 0.4	5.5 ± 0.1	5.1 ± 0.2
YPD	5.0	0.42 ± 0.01	0.37 ± 0.01	0.32 ± 0.01	9.9 ± 0.4	7.9 ± 0.1	6.2 ± 0.1
	6.0	0.42 ± 0.01	0.40 ± 0.01	0.38 ± 0.01	9.9 ± 0.3	9.1 ± 0.1	8.6 ± 0.1

* 10 g/L exceeded solubility and TAL precipitated

membrane fluidity and integrity (Aguedo et al., 2003). Therefore, at lower pH, the protonated TAL may similarly be detrimental to the membrane and thus cell growth rate.

These results were obtained for high levels of TAL added at the start of batch culture; even so, the effects on growth were relatively minor in *S. cerevisiae*. High TAL concentrations are not expected until late in the cultivation, therefore *S. cerevisiae* is an excellent host for TAL production.

2.4.2 Initial expression system and culture conditions

Our baseline strain BY4741 expressing *Gerbera hybrida* 2-pyrone synthase under the strong glycolytic *PGK1* promoter (plasmid pXP218-2PS) produced only 58 mg/L of TAL (Figure 2.2) with a yield on glucose of less than 1% (g/g). Therefore, we varied the promoter, the synthase enzyme, and culture conditions, and introduced protease knockouts to increase TAL levels prior to our pathway engineering.

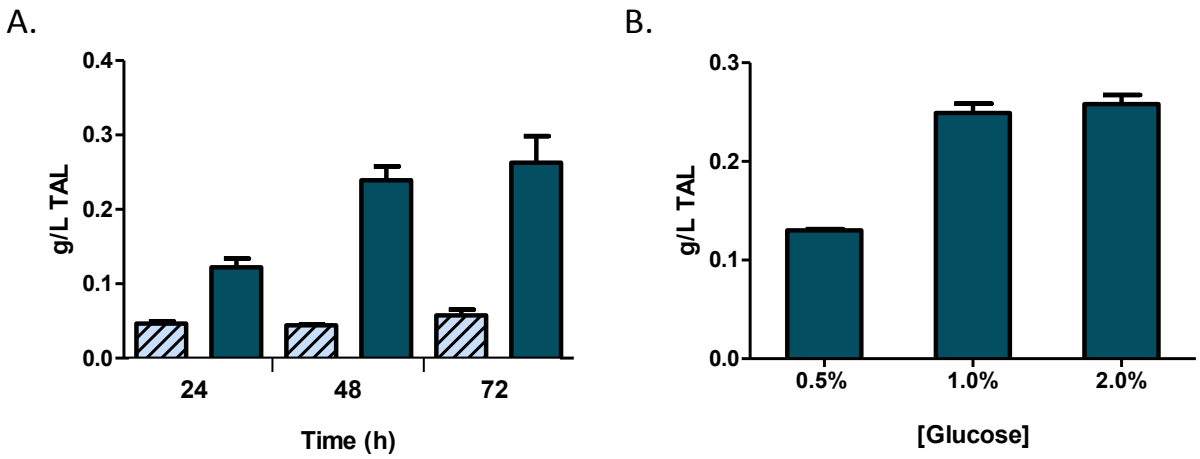


Figure 2.2. (A) Comparison of the strong glycolytic *PGK1* promoter (hatched) versus the late-phase glucose-repressible *ADH2* promoter (solid) expressing the *G. hybrida* 2-pyrone synthase. Values correspond to HPLC analysis for samples harvested at 24, 48, and 72 h. (B) Varying glucose concentration for 48 h batch cultures utilizing the *ADH2* promoter in YPD medium containing 0.5%. 1.0%, or 2.0% glucose. Bars represent mean values ± one standard deviation (n=6).

The late-phase *ADH2* promoter has shown excellent performance for the expression of other polyketides and fatty acid synthases in our lab and others (Kealey et al., 1998; Leber and Da Silva, 2013; Lee et al., 2009; Ma et al., 2009; Mutka et al., 2006). This promoter is active when glucose levels fall, and optimal induction occurs in 1% glucose YPD medium (Lee and Da Silva, 2005). Due to the late induction, plasmid stabilities remain high even in complex medium (Shen et al., 2012). Substitution of this promoter (plasmid pXP842-2PS) increased titers 4.5-fold over the *PGK1* promoter to 260 mg/L (Figure 2.3). This increase is similar to that observed by Xie et al. (2006); they reported higher levels with the *ADH2* promoter relative to the *GPD1* and *ADH1* promoters. In our batch cultures, only a 10% increase in TAL was observed from 48-72 h; therefore a 48 h cultivation period was selected for subsequent studies. As expected, a 1% glucose level resulted in the best performance.

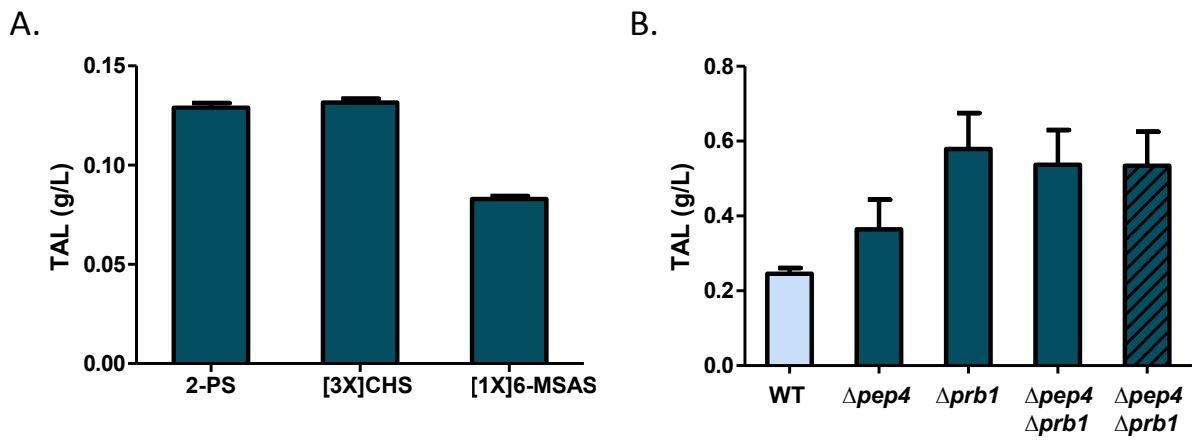


Figure 2.3. (A) TAL titers for *S. cerevisiae* strain BY4741 expressing the wildtype *G. hybrida* 2-pyrone synthase (2-PS), the triple mutant (T197L/G256L/S338I) *M. sativa* chalcone synthase ([3X]CHS), and the single mutant (Y1572F) *P. patulum* 6-MSAS ([1X]6-MSAS). The [1X]6-MSAS was expressed in strain BYAN (BY4741 with one integrated *A. nidulans npgA* gene for activation of the synthase). HPLC analysis was performed on samples collected after 48 h of cultivation in YPD (0.5% glucose). (B) Comparison of TAL titers in BY4741 protease knockout strains expressing the 2-PS under the control of the *ADH2* promoter. Expression in protease deficient strain BJ5464 is shown for comparison. Bars represent mean values ± one standard deviation (n=6 independent experiments).

The native *G. hybrida* 2-pyrone synthase (2-PS) is responsible for the synthesis of TAL, but mutant PKSs have also been shown to make TAL (Jez et al., 2000; Xie et al., 2006). These two mutants were briefly compared (T197L/G256L/S338I chalcone synthase from *M. sativa*, Y1572F 6-MSA synthase from *P. patulum*) with the native 2-PS using the same promoter and backbone vector (pXP-842). During the 48 h cultivation, the variants produced similar or lower levels of TAL relative to 2-PS (Figure 2.3, Panel A). The native 2-PS enzyme was selected for use in our strain development studies. An additional advantage of moving forward with the 2-PS relative to the 6-MSAS variant is that overexpression of a phosphopantetheinyl transferase (for activation) is not required.

Yeast proteases have been shown to act on heterologous enzymes produced in *S. cerevisiae*, and can negatively affect PKS levels and thus polyketide production. In our previous work (Lee et al., 2009), disruption of the two vacuolar protease genes (*PEP4*, *PRB1*) encoding the aspartyl protease and proteinase B, respectively, were shown to be critical to detect lovastatin nonaketide synthase (LNKS) produced in yeast. Therefore, BY4741 strains with single and double knockouts of these genes were evaluated, and TAL production was measured following a 48 h cultivation in YPD. The Δ*pep4* strain resulted in an increase in titer of 50% over the base strain, while Δ*prb1* yielded an increase of 135% (Figure 2.3, Panel B), with no added improvement when combining the deletions (Δ*pep4*Δ*prb1*). An independent strain (BJ5464) that also lacks these two proteases produced similar levels of TAL. Reduction of proteolytic degradation of 2-PS is thus important for TAL production. Titer and yield increased more than 2-fold to 580 mg/L and 12% of theoretical yield (g/g glucose).

2.4.4 Identification and disruption of competing pathways to improve TAL synthesis

The 2-PS requires one acetyl-CoA starter and two malonyl-CoA extender units to generate one triacetic acid lactone product (Jez et al., 2000). These two precursors are produced naturally by *S. cerevisiae* and are in high demand for energy production, fatty acid biosynthesis, and other cellular activities, with malonyl-CoA pools believed to be a significant limiting factor in these processes (Chen et al., 2012; Wattanachaisaereekul et al., 2008). Synthesis of acetyl-CoA / malonyl-CoA metabolites relies on cytosolic pyruvate, a critical branch point in glycolysis (Pronk et al., 1996). For this reason, engineering *S. cerevisiae* to increase acetyl-CoA and pyruvate levels could alleviate flux limitations toward TAL production.

Enzymes participating in central carbon metabolism that are responsible for producing or consuming pyruvate and acetyl-CoA were analyzed to identify potential gene deletions to increase TAL production. The pathways illustrated in Figure 2.4 focus on the utilization of glucose-6-phosphate (G6P) by the cell following glucose uptake from the medium. Significant byproducts (glycogen, pentose sugars, glycerol, and ethanol) that can be produced while generating TAL precursor pools are also shown. Using this pathway analysis, a variety of genes were selected for deletion including *GSY1/2*, *ZWF1*, *PYC1/2*, *PCK1*, *FBP1*, *YIA6*, *ADH1*, and *MLS1* (Figure 2.4, circled and underlined genes). The *GSY1/2* and *ZWF1* genes focus on an early branch point in glycolysis; *GSY1* and *GYS2* encode glycogen synthase paralogs that produce glycogen for energy storage in the cell, and *ZWF1* encodes the glucose-6-phosphate dehydrogenase important for driving the pentose phosphate pathway (Ni and LaPorte, 1995; Nogae and Johnston, 1990). The *PYC1/2*, *PCK1*, and *FBP1* genes encode activities important in the gluconeogenic pathway that support increases in biomass and glycogen using non-fermentable carbon sources (Lin et al., 2001). The yeast mitochondrial NAD^+ transporter encoded by *YIA6* was recently identified to perform a unidirectional transport of NAD^+ out of the cytoplasm and into the mitochondria (Todisco et al., 2006). The *ADH1*-encoded alcohol dehydrogenase is known to convert acetaldehyde to ethanol and also plays an important role in redox balance by generating NAD^+ as a byproduct (Denis et al., 1983). Malate synthase is encoded by *MLS1* and is responsible for generating malate from acetyl-CoA in the cytosol (Chen et al., 2012).

In order to eliminate their native activities and thus improve precursor (acetyl-CoA and malonyl-CoA) pools, strains were constructed with single gene disruptions (Table 2.1) and then transformed with pXP842-2PS. This set of single knockout strains was cultivated in YPD medium

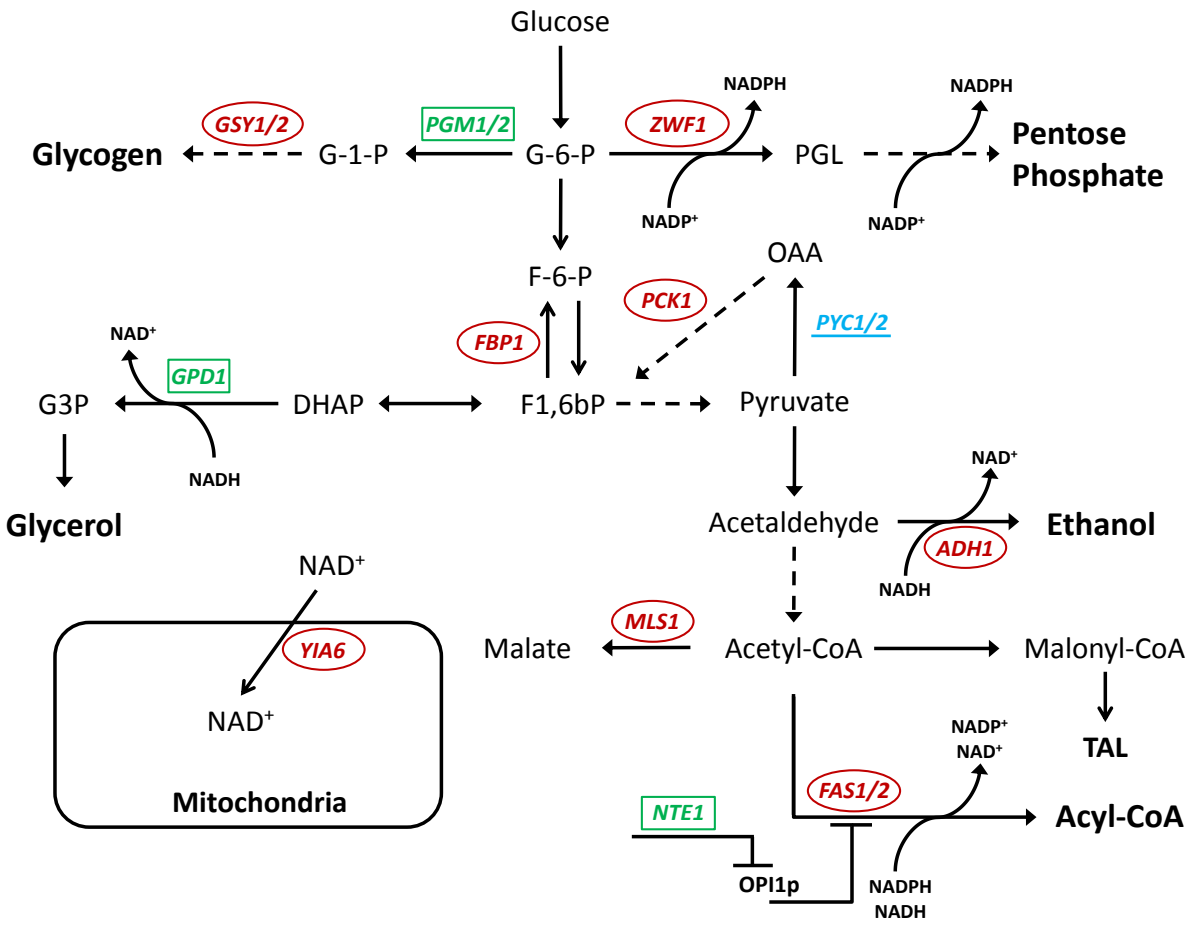


Figure 2.4. Metabolic pathways identified for optimizing TAL precursors in *S. cerevisiae*. The highlighted metabolites (bold) identify major by-products generated in central carbon metabolism leading to reduced pyruvate and acetyl-CoA pools. The genes corresponding to enzymes catalyzing the reactions are shown in italics and were determined by pathway analysis (circled), computational predictions (boxed), or both (underlined). In this schematic, reactions not directly used in our analysis have been consolidated as dashed arrows.

(1% glucose) for 48 h and TAL production was compared on both a specific (per g cell) basis (Figure 2.5, Panel A) and a volumetric (titer) basis (Figure 2.5, Panel B). The final cell densities of the cultures are shown in Figure 2.5, Panel C (growth curves can be found in Figure 2.6). Knockouts of *GSY1/2*, *ZWF1*, *PYC1/2*, or *YIA6* gave greater than 2-fold increases in specific TAL

levels and approximately 50% increases in final biomass. In combination, this led to ca. 3-fold increases in TAL with titers of up to 0.81 g/L. Knockouts of *PCK1*, *ADH1*, and *MLS1* did not increase TAL titers or specific production, with levels comparable or lower than the control strain (WT). The knockout of *FBP1* increased titer by 2.2-fold and per cell levels by 2.4-fold, one of the highest specific production levels despite a slight reduction in growth. The significance of this deletion is that Fbp1p is responsible for conversion of fructose 1,6-biphosphate, which is the only reversible step (Pfk1p acts in the forward direction) of glycolysis that uses an independent enzyme (Lin et al., 2001). The results for the best strains (Figure 2.5) show that the most important factor was the effect of the gene disruptions on the TAL produced on a per cell basis. However, the highest titers came from strains that were also able to generate increased biomass. By exploiting the well-known metabolic pathways, our analysis suggested ten genes for disruption to redirect the metabolic flux toward TAL precursors. This strategy proved successful as six of the ten knockouts resulted in 3-fold higher TAL titers and >2-fold increases in specific production.

2.4.5 *In silico* prediction of pathway gene knockouts via OptKnock

An *in silico* metabolic engineering approach was also used to identify additional promising single gene disruptions, where identification could be facilitated using a genome-scale model. Using the iMM904 yeast metabolic network, we employed the OptKnock algorithm packaged within the Cobra Toolbox (Mo et al., 2009; Schellenberger et al., 2011). This algorithm has been designed to identify gene knockouts in an organism core model that will increase a target flux value as specified by the user (Burgard et al., 2003), and uses a bilevel optimization framework

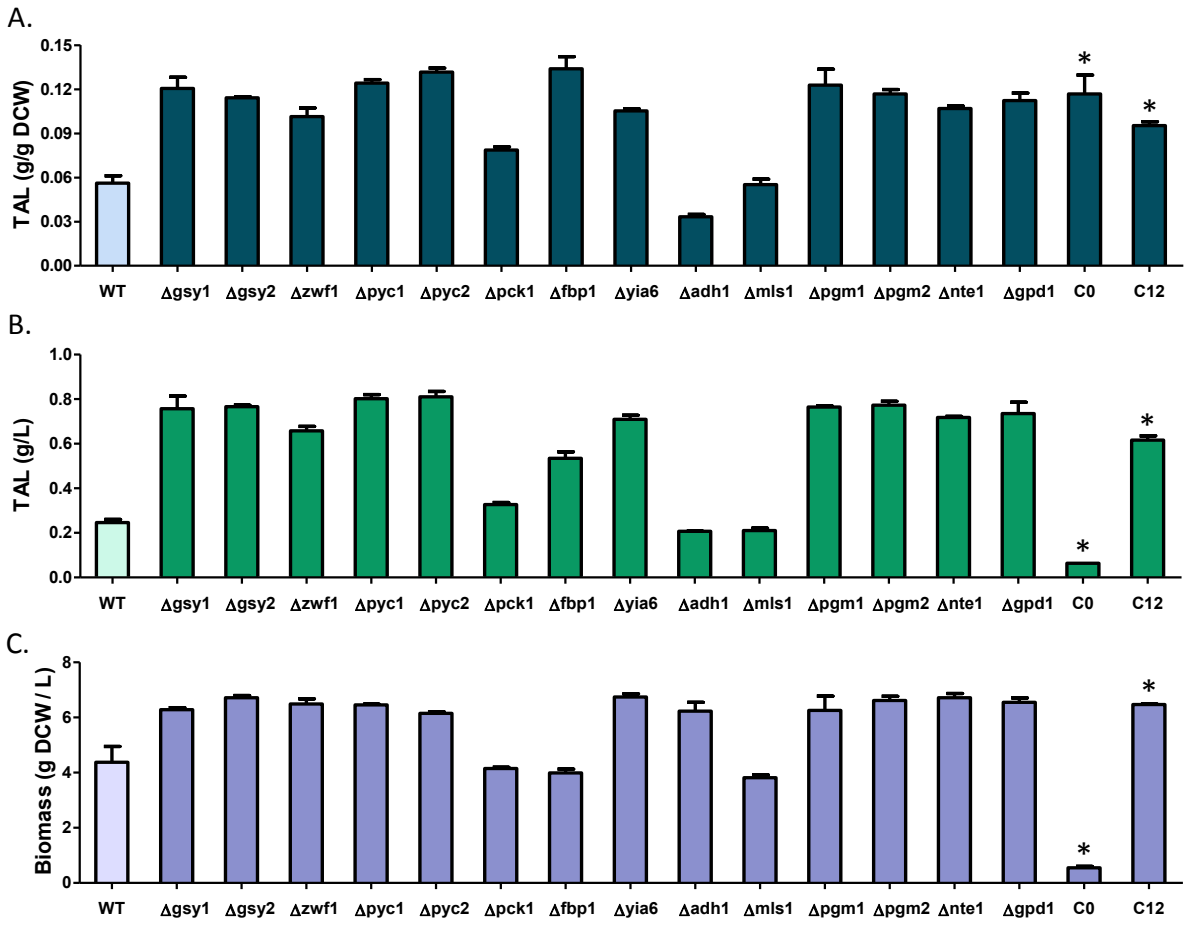


Figure 2.5. Comparison of TAL produced in *S. cerevisiae* strains with single gene knockouts after 48 h of cultivation: (A) production on a per cell mass basis [mg/g DCW], (B) titers [g/L], and (C) final cell densities (FCD, g DCW/L). WT refers to the base strain BY4741. The asterisk denotes the WT strains supplemented with 10 µM cerulenin at 0 h (bar C0) and 12 h (bar C12). Bars represent mean values ± one standard deviation (n=6 independent experiments).

such that maximization of the target metabolite is subject to maximizing biomass flux. In doing so, solutions should provide gene disruptions free of growth defects, adding a layer of robustness relative to the metabolic engineering approach described above. The procedure has been validated extensively *in vivo* using *E. coli*, and a recent report achieved nearly a 60-fold

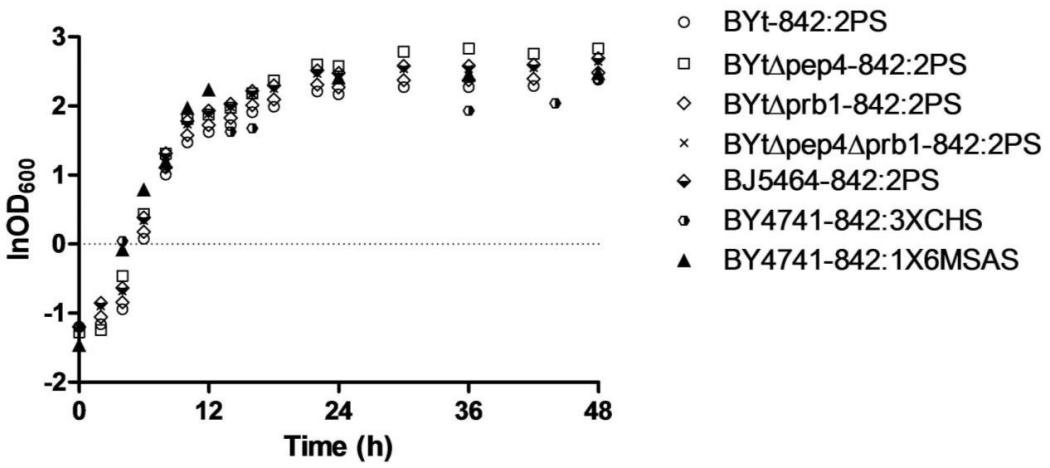


Figure 2.6. Growth curves for various strains in this study, as denoted above. Single curves were chosen as representative data for each strain.

improvement in 2, 3-butanediol titers using *S. cerevisiae* (Ng et al., 2012). Our optimization was constrained to target the flux of either acetyl-CoA or pyruvate, and limited to up to five deletions in any given solution set. Two solution sets were generated for acetyl-CoA and one unique set for pyruvate, made possible by modulating both biomass and ATP maintenance in the optimization constraints. The reactions and their corresponding genes identified in the optimization problem using these two targets are summarized in Table 2.4. The metabolic pathways associated with these interventions can be found in Figure 2.4. For acetyl-CoA flux, three reactions were identified and had genes associated with glycogen biosynthesis (phosphoglucomutase; *PGM1/2*), lipid metabolism (serine esterase; *NTE1*), or glycerol biosynthesis (glycerol-3-phosphate dehydrogenase; *GPD1*). Figure 2.4 illustrates how these disruptions benefit TAL production by increasing acetyl-CoA levels. With these deletions there will be a reduction in flux toward significant branch points in glycogen and pentose biosynthesis, as well as fatty acid biosynthesis that is known to consume large quantities of acetyl-CoA and

Table 2.4. List of reactions and gene knockouts identified by the OptKnock algorithm for increased TAL production.

Target	Acetyl-CoA	Pyruvate
Reactions Identified	PLBPC_Sce, PGPPAm_Sc, G3PD1ir	PC
Corresponding Genes	*NTE1, PGM1/2, GPD1*	*PYC1/2*

malonyl-CoA. In the case of pyruvate flux as the target, OptKnock selected the PC reaction for disruption, corresponding to pyruvate carboxylase activity encoded by the *PYC1/2* genes, and the first committed step of gluconeogenesis. These two deletions were already evaluated as discussed above. Strains harboring a single deletion in *PGM1*, *PGM2*, *NTE1*, or *GPD1* were constructed, transformed with pXP842-2PS, cultivated for 48 h in YPD, and then screened for TAL biosynthesis. All six strains corresponding to interventions identified by the *in silico* method showed robust growth and gave significant improvements (Figure 2.5) with increases that resulted in 3-fold higher TAL titers and >2-fold increases in specific production.

It was not surprising that interventions selected by OptKnock were in close alignment to those previously identified using a pathway analysis. A clear example is the identification of the *PYC1/2* genes by both strategies. Additionally, both strategies sought to manipulate glycogen production by either the *GSY1/2* paralog system or the *PGM1/2* system. The deletion of *GSY1/2* influences the final process in glycogen synthesis (Figure 2.4), where *PGM1/2* affects the pathway further upstream, preventing synthesis of unnecessary metabolites. A final feature of the OptKnock strategy is that disruptions not obvious via traditional pathway analysis can be

elucidated through the use of genome-scale analysis. Using the core model with OptKnock suggested the regulator of lipid metabolism (*NTE1*) as an avenue for increased TAL precursor pools in *S. cerevisiae*, a logical result that was not readily identified in our previous analysis. As shown in Figure 2.4, *NTE1* encodes the serine esterase responsible for, among other things, repression of the *OPI1*-encoded Opi1p repressor protein (Fernández-Murray et al., 2009). The enzyme silences the yeast fatty acid synthase, helping modulate phospholipid biosynthesis in *S. cerevisiae* (Schweizer and Hofmann, 2004). Disrupting *NTE1* potentially relieves the repression of Opi1p and reduces flux of TAL precursors toward competing products like fatty acids. With Nte1p participating in the global regulation of lipid metabolism, further experiments need to be performed to pinpoint exactly how Δ*nte1* is improving TAL production.

2.4.6 Downregulation of fatty acid synthesis via cerulenin addition

Yeast fatty acid biosynthesis is an obvious choice for downregulation to improve the levels of TAL and other polyketides. The FAS1/FAS2 complex uses acetyl-CoA, malonyl-CoA, and NADPH for the creation of long chain fatty acids, primarily the 16-carbon palmitate (Wakil et al., 1983). An initial experiment was performed to determine the effect of fatty acid downregulation on TAL levels. Strain BY4741 (carrying pXP842-2PS) was cultivated in YPD medium containing 10 µM cerulenin for 48 h. Cerulenin is a known fatty acid synthase inhibitor, and at this concentration the strain would be subject to a significantly repressed FAS complex (Omura, 1976). The cerulenin was added at time zero (C0) or 12 hours post-inoculation (C12); the latter time corresponds to glucose depletion in the batch culture. The specific production of TAL for culture C0 was approximately 2-fold higher than the control (Figure 2.5, Panel A). However, due to the inhibitory

effects of this compound, final cell density reached only 10% of the control strain and therefore volumetric (titer) production was very low (Figure 2.5, Panels B and C). The outcome was very different when the cerulenin was spiked at a later time (C12). By postponing the inhibition of fatty acid biosynthesis, a normal cell density was achieved (Figure 2.5, Panel B). Specific TAL production was reduced only slightly relative to C0, while titers were 10-fold higher relative to C0. Final titer for the C12 culture was comparable to the other interventions, and 2.5-fold greater than the base strain. This result, and that for the *NTE1* deletion, suggests that modulating fatty acid synthesis will be instrumental for increased production of triacetic acid lactone in *S. cerevisiae*.

2.4.7 Combining pathway interventions

A total of 17 single gene deletions were identified and tested for improved TAL synthesis in *S. cerevisiae*. To determine whether further increases could be achieved by combining multiple deletions, a subset of the 17 genes was selected that would potentially provide the best gains, with minimal deleterious effects. We selected strong performers from Figure 2.5 and introduced the following disruptions: protease activity (*PRB1*), reverse glycolysis (*PYC2*), cofactor transport (*YIA6*), and lipid biosynthesis (*NTE1*). The genes were selected to affect significantly different pathways in order to ensure minimal overlap with the hope of producing substantial increases in TAL production. The strains were constructed first by sequential deletion of genes *PRB1* and *PYC2* in the base strain BY4741Δ*trp1*. *YIA6* or *NTE1* was then knocked out to create strains BYtΔ*prb1*Δ*pyc2*Δ*yia6* and BYtΔ*prb1*Δ*pyc2*Δ*nte1*, respectively. Lastly, *YIA6* was deleted from BYtΔ*prb1*Δ*pyc2*Δ*nte1* to construct the quadruple knockout strain (BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1*).

The Δ*pyc2* and Δ*pyc2*Δ*nte1* deletions were also introduced into the alternate protease-deficient strain BJ5464 for comparison. All strains were then transformed with pXP842-2PS, cultivated for 48 h in YPD (1% glucose), and TAL levels were measured (Figure 2.7).

The TAL titer for BYtΔ*prb1*Δ*pyc2* strain was 0.91 g/L, nearly three times higher than that of the base strain and 57% higher than for BYtΔ*prb1*. Subsequent deletion of either *YIA6* or *NTE1* further increased TAL levels, with the latter exceeding 1.0 g/L. The combination of all four deletions in strain BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1* resulted in a final titer of 1.3 g/L (29% of theoretical yield) corresponding to an overall 5-fold improvement over the base strain. If the increases for the single gene deletions were cumulative, titer would have been approximately 1.9 g/L, suggesting new bottlenecks might be present. The Δ*pyc2* and Δ*pyc2*Δ*nte1* deletions were also introduced into the alternate strain BJ5464 to see if similar behavior was observed. Sequential gene knockouts increased TAL levels in this strain as well (Figure 2.7). Interestingly, this strain generally showed higher levels of TAL production relative to the comparable BY4741-based strains. BJΔ*pyc2* produced 23% more TAL than BYtΔ*prb1*Δ*pyc2*, and BJΔ*pyc2*Δ*nte1* produced 29% more than BYtΔ*prb1*Δ*pyc2*Δ*nte1*. We have observed similar results with other polyketides in our laboratory.

All strains showed similar growth behaviors and final cell densities. Therefore, specific TAL production (per g cell basis) showed the same trend as that in Figure 2.7 (see Figure 2.8). In addition, although the strains were cultivated in YPD medium, plasmid stability was not an issue due to the late-phase induction, and was routinely measured at over 85% for all strains

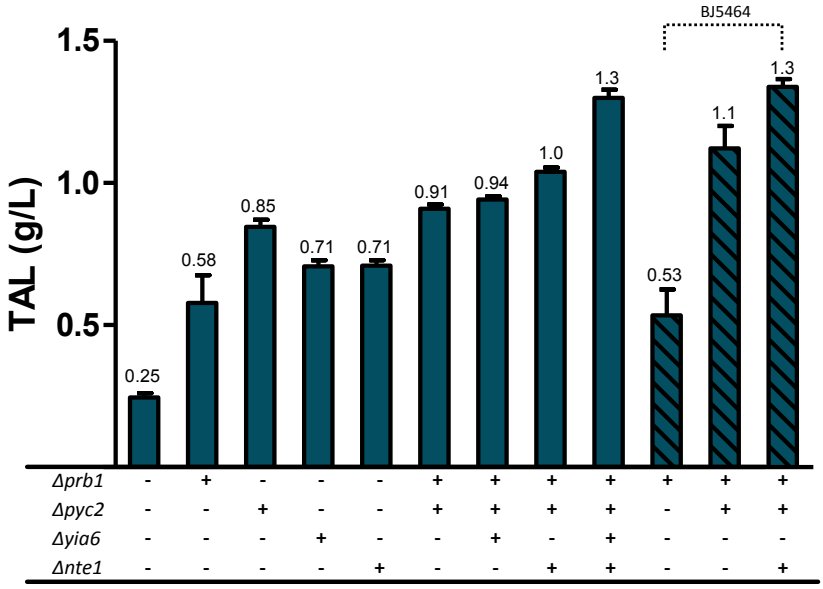


Figure 2.7. Improvements in TAL levels following combination of protease and pathway gene knockouts. The open bars represent strain BY4741 and the hatched bars are for strain BJ5464 (this base strain has the *PRB1* gene disrupted). Strains either have (+) or lack (-) the gene deletions shown. Bars represent means ± one standard deviation (n=6 independent experiments).

tested (e.g., for strain BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1*, 87 ± 2 % of the cells retained the 2-PS plasmid after 48 h).

2.4.8 Glucose fed-batch fermentations

In order to further evaluate the engineered strain BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1* for the high-level production of TAL, fed-batch fermentations were performed. The seed and batch cultures were in YPD with 1% glucose. Based on the toxicity data (Section 3.1), we chose to maintain pH at 6.0. After allowing the initial glucose to be consumed for approximately 12 h, pure glucose was continuously fed while ensuring that glucose levels were kept low to maximize expression of the

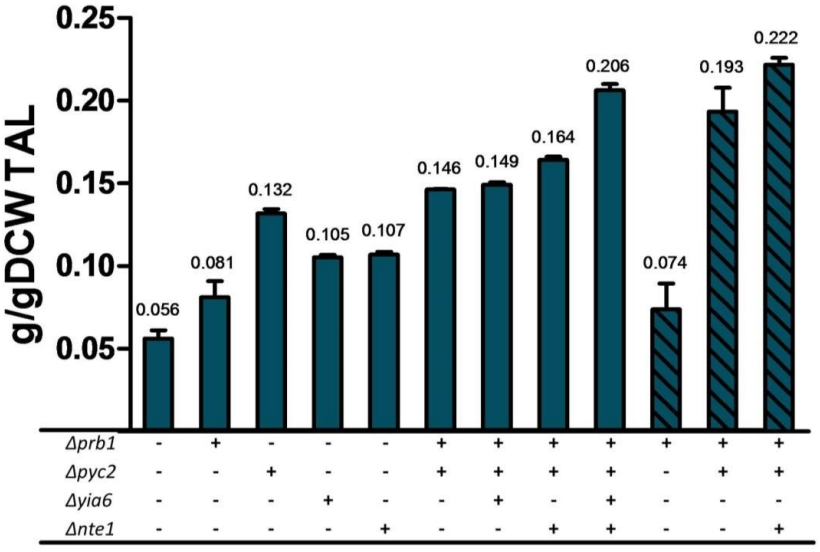


Figure 2.8. Improvements for strains with combined disruptions on a per gram cell basis (g/g DCW TAL), corresponding directly to the volumetric basis data shown in Figure 2.7. Bars represent mean values ± one standard deviation (n=6 independent experiments).

2-PS under the P_{ADH2} promoter. To do this, glucose feed rate was allowed to vary in order to limit dissolved oxygen levels to 20% throughout the fermentation. Additionally, off-line HPLC analysis was performed to quantify glucose levels during the fed-batch stage; glucose concentration never exceeded 0.35 g/L. From 24 to 48 h, triacetic acid lactone levels increased 3-fold to over 1 g/L, and by 72 h began to level off at greater than 2 g/L (Figure 2.9). Biomass concentration and TAL production on a per g cell basis are shown in Figure 2.10. By 96 h, supplementation of glucose to maintain DO levels was minimal, and culture density was unchanged. This is likely due to other nutrient limitations after 48 h using a glucose-only feed. Using our engineered strain, a final titer of 2.2 g/L was achieved in the fed-batch culture (versus 1.3 g/L in batch). It is noteworthy that a similar yield was obtained (0.12 g/g glucose) as in the small scale batch cultures (0.13 g/g glucose).

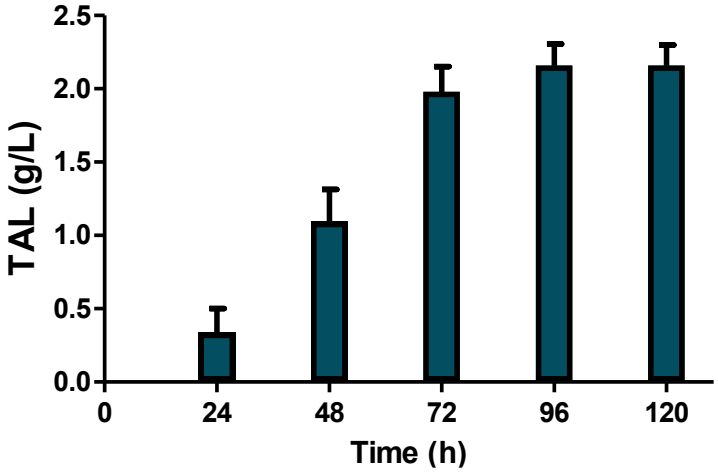


Figure 2.9. TAL titers during fermentor-controlled (New Brunswick BioFlo III) glucose fed-batch cultivation using the engineered strain BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1*. Following inoculation, the cells were grown in batch culture for 12 hours. The glucose feed (3.6M) was then initiated, and feed rate was controlled (0-1.5 mL/h) to maintain DO at 20% with an aeration rate of 0.8 vvm. Bars represent means ± one standard deviation (n=3 independent experiments).

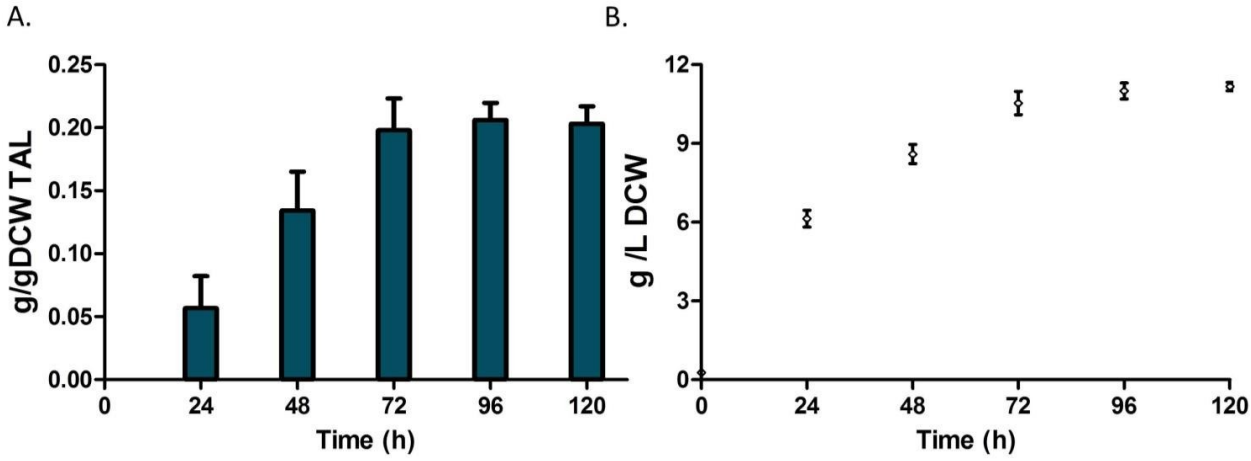


Figure 2.10. Additional characteristics of fed-batch fermentation of BYtΔ*prb1*Δ*pyc2*Δ*nte1*Δ*yia6* (P*ADH2*-2PS). (A) TAL production on a per gram cell basis (g/g DCW TAL), and (B) biomass production (g/L DCW). Bars represent mean values ± one standard deviation (n=3 independent experiments).

A summary of the substantial improvements in TAL titer and yield obtained with the various interventions is shown in Figure 2.11. The highest TAL production was observed with strain BYtΔ*prb1*Δ*pyc2*Δ*yia6*Δ*nte1* reaching a maximum titer of 2.2 g/L and a maximum yield of 0.12-0.13 g/g glucose (26-29% of theoretical). These are the highest yields and titers reported to date, with the yield 13-fold and 3.0-fold higher than reported by Xie et al. (2006) for *S. cerevisiae* expressing the 2-PS or 6-MSAS variant, respectively.

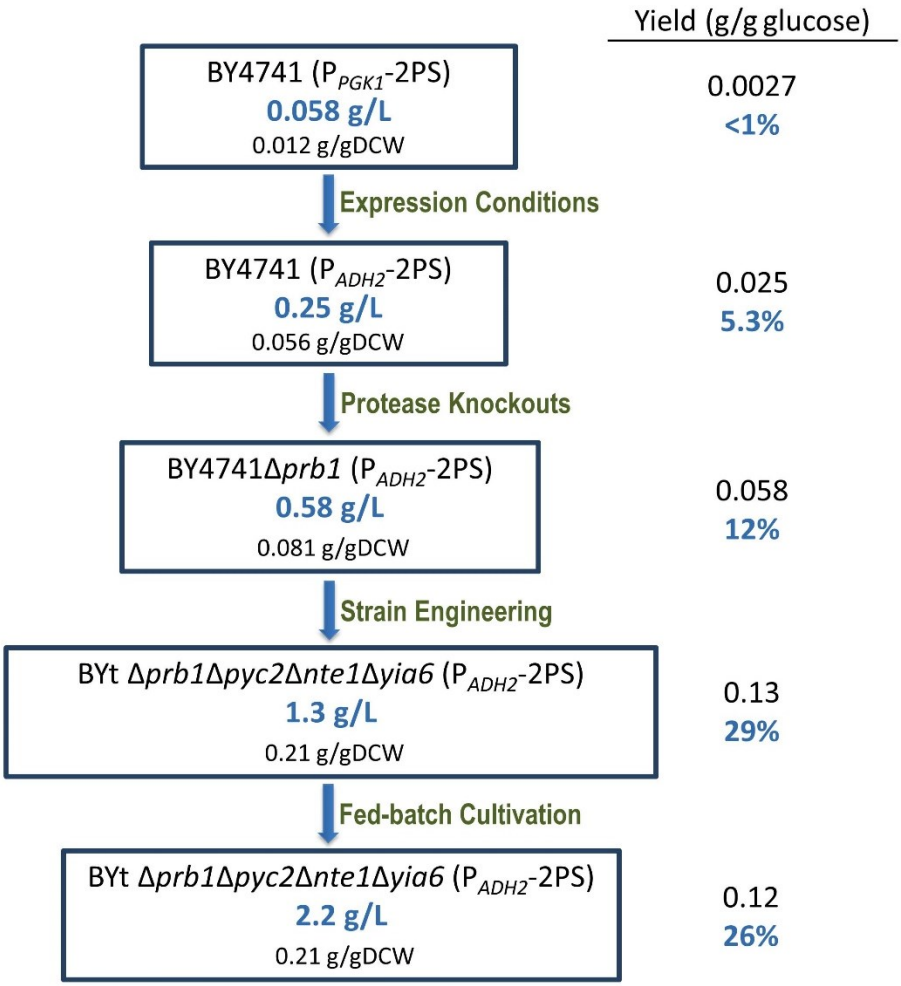


Figure 2.11. Final titer and yield summary for TAL production in *S. cerevisiae*. Volumetric titer (g/L), specific production (cell basis; g/g DCW), and corresponding yield on glucose (g/g) with percent theoretical yield are shown.

2.5 Conclusions

S. cerevisiae is an excellent host for the production of triacetic acid lactone due to the minimal toxicity. Our initial baseline strain produced very low titers (58 mg/L) of TAL with a yield on glucose of less than 1% (g/g). Initial changes in the expression system (e.g., promoter, media) increased TAL titer by approximately 4-fold to 0.25 g/L and increased yield by approximately 10-fold. The introduction of protease knockouts to improve *in vivo* stability of the 2-pyrone synthase produced an additional 2-fold improvement in titer to 0.58 g/L and a further 2-fold increase in yield. Pathway engineering of central carbon metabolism, including the evaluation of computational predictions, identified a total of 15 potential gene knockouts. Up to 3-fold increases in TAL titer were observed for the single gene deletion strains. By combining multiple gene knockouts and performing batch and fed-batch culture, our highest titer was 2.2 g/L and our highest yield was 0.13 (g/g glucose), or 29% of theoretical. To our knowledge, this is the first study implementing metabolic pathway engineering to increase the synthesis of TAL. This study also provides the highest titer and yield of TAL reported to date. Overall, incorporation of optimal culture parameters and combined pathway gene knockouts led to a 37-fold increase in final titer and an approximately 50-fold increase in yield. Additional studies are underway to evaluate the upregulation of specific pathway genes and rational enzyme engineering approaches for improved TAL synthesis in *S. cerevisiae*.

2.6 Acknowledgements

This research was supported by the National Science Foundation (Grant No. EEC-0813570) through the Engineering Research Center CBiRC (Center for Biorenewable Chemicals). The authors would like to thank Dr. Joseph Noel (Salk Institute) for *g2ps1* (encoding 2-PS) and the *chs2* variant (encoding the T197L/G256L/S338I mutant of CHS), and Dr. Laura Jarboe (Iowa State University) for helpful input on TAL toxicity.

2.7 References

Aguedo, M., Beney, L., Waché, Y., Belin, J.-M., 2003. Mechanisms underlying the toxicity of lactone aroma compounds towards the producing yeast cells. Journal of Applied Microbiology 94, 258–265. doi:10.1046/j.1365-2672.2003.01828.x

Austin, M.B., Noel, J.P., 2003. The chalcone synthase superfamily of type III polyketide synthases. Nat. Prod. Rep. 20, 79–110. doi:10.1039/B100917F

Burgard, A.P., Pharkya, P., Maranas, C.D., 2003. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. 84, 647–657. doi:10.1002/bit.10803

Chen, Y., Siewers, V., Nielsen, J., 2012. Profiling of Cytosolic and Peroxisomal Acetyl-CoA Metabolism in Saccharomyces cerevisiae. PLoS ONE 7, e42475. doi:10.1371/journal.pone.0042475

Chia, M., Schwartz, T.J., Shanks, B.H., Dumesic, J.A., 2012. Triacetic acid lactone as a potential biorenewable platform chemical. Green Chem. 14, 1850–1853. doi:10.1039/C2GC35343A

Denis, C.L., Ferguson, J., Young, E.T., 1983. mRNA levels for the fermentative alcohol dehydrogenase of Saccharomyces cerevisiae decrease upon growth on a nonfermentable carbon source. J. Biol. Chem. 258, 1165–1171.

Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Proksch, P., Teeri, T.H., Schröder, J., 1998. New pathway to polyketides in plants. Nature 396, 387–390. doi:10.1038/24652

Evans, G.E., Leeper, F.J., Murphy, J.A., Staunton, J., 1979. Triacetic acid lactone as a polyketide synthon: synthesis of toralactone and polyketide-type anthracene derivatives. J. Chem. Soc., Chem. Commun. 205–206. doi:10.1039/C39790000205

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. doi:10.1002/yea.1824

Fernández-Murray, J.P., Gaspard, G.J., Jesch, S.A., McMaster, C.R., 2009. NTE1-encoded phosphatidylcholine phospholipase b regulates transcription of phospholipid biosynthetic genes. J. Biol. Chem. 284, 36034–36046. doi:10.1074/jbc.M109.063958

Gietz, D., Jean, A.S., Woods, R.A., Schiestl, R.H., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucl. Acids Res. 20, 1425–1425. doi:10.1093/nar/20.6.1425

Goel, A., Ram, V.J., 2009. Natural and synthetic 2H-pyran-2-ones and their versatility in organic synthesis. Tetrahedron 65, 7865–7913. doi:10.1016/j.tet.2009.06.031

Hamelinck, C.N., Hooijdonk, G. van, Faaij, A.P., 2005. Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. Biomass and Bioenergy 28, 384–410. doi:10.1016/j.biombioe.2004.09.002

Hill, J., Donald, K.A.I.G., Griffiths, D.E., 1991. DMSO-enhanced whole cell yeast transformation. Nucl. Acids Res. 19, 5791–5791. doi:10.1093/nar/19.20.5791

Jez, J.M., Austin, M.B., Ferrer, J.-L., Bowman, M.E., Schröder, J., Noel, J.P., 2000. Structural control of polyketide formation in plant-specific polyketide synthases. Chemistry & Biology 7, 919–930. doi:10.1016/S1074-5521(00)00041-7

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.

Leber, C., Da Silva, N.A., 2013. Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids. Biotechnology and Bioengineering. doi:10.1002/bit.25021

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. doi:10.1016/j.ab.2009.07.010

Lee, M.K., Da Silva, N.A., 2005. Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis. Yeast 22, 431–440. doi:10.1002/yea.1221

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, 36000–36007. doi:10.1074/jbc.M103509200

Liu, P., Chernyshov, A., Najdi, T., Fu, Y., Dickerson, J., Sandmeyer, S., Jarboe, L., 2013. Membrane stress caused by octanoic acid in Saccharomyces cerevisiae. Appl Microbiol Biotechnol 97, 3239–3251. doi:10.1007/s00253-013-4773-5

Ma, S.M., Li, J.W.-H., Choi, J.W., Zhou, H., Lee, K.K.M., Moorthie, V.A., Xie, X., 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. doi:10.1126/science.1175602

Mo, M.L., Palsson, B.O., Herrgård, M.J., 2009. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Syst Biol 3, 37. doi:10.1186/1752-0509-3-37

Moreno-Mañas, M., Prat, M., Ribas, J., Virgili, A., 1988. Palladium catalyzed allylic C-alkylation of highly acidic and enolic heterocyclic substrates: tetronic acid and triacetic acid lactone. Tetrahedron Letters 29, 581–584. doi:10.1016/S0040-4039(00)80156-2

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 Res 6, 40–47. doi:10.1111/j.1567-1356.2005.00001.x

Ng, C.Y., Jung, M., Lee, J., Oh, M.-K., 2012. Production of 2,3-butanediol in Saccharomyces cerevisiae by in silico aided metabolic engineering. Microbial Cell Factories 11, 68. doi:10.1186/1475-2859-11-68

Ni, H.T., LaPorte, D.C., 1995. Response of a yeast glycogen synthase gene to stress. Mol. Microbiol. 16, 1197–1205.

Nikolau, B.J., Perera, M.A.D.N., Brachova, L., Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. Plant J. 54, 536–545. doi:10.1111/j.1365-313X.2008.03484.x

Nogae, I., Johnston, M., 1990. Isolation and characterization of the ZWF1 gene of Saccharomyces cerevisiae, encoding glucose-6-phosphate dehydrogenase. Gene 96, 161–169.

Omura, S., 1976. The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. Bacteriol Rev 40, 681–697.

Pronk, J.T., Yde Steensma, H., Van Dijken, J.P., 1996. Pyruvate Metabolism in Saccharomyces cerevisiae. Yeast 12, 1607–1633. doi:10.1002/(SICI)1097-0061(199612)12:16<1607::AID-YEA70>3.0.CO;2-4

Richardson, M.T., Pohl, N.L., Kealey, J.T., Khosla, C., 1999. Tolerance and Specificity of Recombinant 6-Methylsalicylic Acid Synthase. Metabolic Engineering 1, 180–187. doi:10.1006/mben.1999.0113

Royce, L.A., Liu, P., Stebbins, M.J., Hanson, B.C., Jarboe, L.R., 2013. The damaging effects of short chain fatty acids on Escherichia coli membranes. Appl Microbiol Biotechnol 97, 8317–8327. doi:10.1007/s00253-013-5113-5

Sambrook, J.J., Russell, D.D.W., 2001. Molecular cloning: a laboratory manual. Vol. 2. CSHL Press.

Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J., Feist, A.M., Zielinski, D., Bordbar, A., Lewis, N., Rahmanian, S., Kang, J., Hyduke, D.R., Palsson, B., 2011. COBRA Toolbox 2.0. Protocol Exchange 6, 1290–1307. doi:10.1038/protex.2011.234

Schweizer, E., Hofmann, J., 2004. Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. Microbiol. Mol. Biol. Rev. 68, 501–517. doi:10.1128/MMBR.68.3.501-517.2004

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. doi:10.1002/yea.2930

Taeschler, C., 2010. Ketenes, Ketene Dimers, and Related Substances, in: Kirk-Othmer Encyclopedia of Chemical Technology. John Wiley & Sons, Inc.

Tang, S.-Y., Qian, S., Akinterinwa, O., Frei, C.S., Gredell, J.A., Cirino, P.C., 2013. Screening for Enhanced Triacetic Acid Lactone Production by Recombinant Escherichia coli Expressing a Designed Triacetic Acid Lactone Reporter. Journal of the American Chemical Society 135, 10099–10103. doi:10.1021/ja402654z

Todisco, S., Agrimi, G., Castegna, A., Palmieri, F., 2006. Identification of the Mitochondrial NAD+ Transporter in Saccharomyces cerevisiae. J. Biol. Chem. 281, 1524–1531. doi:10.1074/jbc.M510425200

Wakil, S.J., Stoops, J.K., Joshi, V.C., 1983. Fatty Acid Synthesis and its Regulation. Annual Review of Biochemistry 52, 537–579. doi:10.1146/annurev.bi.52.070183.002541

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. doi:10.1016/j.ymben.2008.04.005

Westfall, P.J., Pitera, D.J., Lenihan, J.R., Eng, D., Woolard, F.X., Regentin, R., Horning, T., Tsuruta, H., Melis, D.J., Owens, A., Fickes, S., Diola, D., Benjamin, K.R., Keasling, J.D., Leavell, M.D., McPhee, D.J., Renninger, N.S., Newman, J.D., Paddon, C.J., 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. PNAS 109, 111–118. doi:10.1073/pnas.1110740109

Xie, D., Shao, Z., Achkar, J., Zha, W., Frost, J.W., Zhao, H., 2006. Microbial synthesis of triacetic acid lactone. Biotechnology and Bioengineering 93, 727–736. doi:10.1002/bit.20759

Zha, W., Shao, Z., Frost, J.W., Zhao, H., 2004. Rational Pathway Engineering of Type I Fatty Acid Synthase Allows the Biosynthesis of Triacetic Acid Lactone from d-Glucose in Vivo. J. Am. Chem. Soc. 126, 4534–4535. doi:10.1021/ja0317271

Chapter 3:

A combined approach for engineering a heterologous Type III PKS and enhancing polyketide biosynthesis

3.1 Abstract

Polyketides are attractive compounds for uses ranging from biorenewable commodity chemicals to high-value therapeutics. In the vast majority of cases, heterologous enzyme production is necessitated in order to produce these compounds in industrially relevant host organisms. Various stages of expression engineering focus on transcription, translation, mRNA stability, or post translational modifications for improving biocatalyst performance. In this study, optimal functionality of the synthase during *in vivo* production by *Saccharomyces cerevisiae* was instead achieved using an iterative design by incorporating both *in vitro* kinetics and *in vivo* performance data. The Type III polyketide synthase (PKS) 2-pyrone synthase (2-PS) from *Gerbera hybrida* was used to produce triacetic acid lactone (TAL) as a model system. Rational design using steric hindrance of the active site led to initial *in vitro* characterization of variants with improved k_{cat} and / or K_M values relative to wildtype, and up to a 30-fold improvement in catalytic efficiency. Further evaluation *in vivo* suggested cysteines either proximal to the surface or CoA binding domain had a significant impact on enzyme performance during fermentation, and could protect the enzyme from proteolytic degradation mechanisms. Therefore, additional rounds of enzyme design were performed with incorporation of up to 6 amino acid changes, and a total of 41 variants were evaluated. The E31 2-PS (C35S, C372S) variant was the most effective biocatalyst of all enzymes tested *in vivo* which had final product titers ranging from 40 mg/L to 1.23 g/L. This optimal enzyme variant was coupled to the engineered strain BJ5464Δ*pep4*Δ*prb1*Δ*pyc2*Δ*nte1* which led to nearly 8 g/L TAL at high yield (44% of theoretical) following fed-batch fermentation, both representing the highest reported values to date. Our studies therefore provide insight towards removing 2-PS surface cysteines for added stability *in vivo* and suggest a new strategy for improving biocatalyst performance in non-native microbial hosts including *S. cerevisiae*.

3.2. Introduction

Biocatalysts, or enzymes, are extremely useful and can be found readily in most household items like detergents, but are also relevant for more advanced purposes like facilitating conversion in biotechnology and biorenewable applications (Nikolau et al., 2008). Other materials also rely on biocatalysts, including but not limited to the microbial production of large volume commodity chemicals styrene and 1-4 butanediol as well as the therapeutic atorvastatin (Lipitor) which is a cholesterol-lowering agent (Ma et al., 2010; McKenna and Nielsen, 2011; Yim et al., 2011). This small sample of products which rely on enzymes for their manufacture contributes to the enzyme business, reported to be worth more than $7 billion (USD) worldwide (Sarrouh, 2012). The ability to leverage catalysts from nature has recently been enhanced by the advancement of directed evolution techniques (Bornscheuer et al., 2012; Esvelt et al., 2011). This approach expedites the iterative process by which improvements in stability or specificity can be screened and selected very rapidly. Nonetheless, advances in techniques including directed evolution, random mutagenesis, or combinations of them still create extremely large libraries that need to be screened for identification of novel enzyme variants. Once the variants are found it is difficult to construct a roadmap for rational construction. Therefore as the field of biocatalysis evolves, new design principles need to be established in order to continue advancing biocatalyst design.

A recent report suggests that the goal of biocatalysis will soon be to routinely employ enzymes in a complex fashion using multistep biocatalysis (Bornscheuer et al., 2012). This refers to the process by which natural products as the starting material will be converted into modified non-natural derivatives (both chemically and biologically) using reactions and pathways that are

not found naturally. This is consistent with the notion that the biocatalysis and traditional catalysis fields are envisioned to work cooperatively for the synthesis of chemicals from renewable means (Nikolau et al., 2008). On the biological side, the creation of complex non-natural pathways for implementation in a microbial platform will require a high degree of understanding and characterizing of enzymes. Unfortunately, a significant amount of effort is still needed to achieve this level of understanding for the creation of robust platforms. Many studies still rely on random mutagenesis of an enzyme which may not readily specify what exact modification led to the desirable trait, making it difficult to engineer the protein in a rational manner. Also, proteins are often studied in crystalline form, that unfortunately is far removed from the accurate fluid protein dynamics which actually take place in the cell (Bornscheuer et al., 2012; Steiner and Schwab, 2012). Therefore, rational approaches using known design rules allow for very selective and specific modifications which further expedite improved biocatalyst design.

In enzyme engineering and the field of biocatalysis, the natural product class of polyketides represents a great opportunity. Due to their diverse nature and often times promiscuity for various substrate metabolites, polyketide biosynthesis is the biological equivalent of combinatorial chemistry, denoted as 'combinatorial biosynthesis' (Nikolau et al., 2008). Polyketide synthases (PKSs), which catalyze the functions required for the biosynthesis of a polyketide, tend to have a wide distribution in their substrate specificity and catalytic activity as a whole (Hertweck, 2009). Additionally, subdivided into types similar to fatty acid synthase systems, PKSs are grouped into three categories as designated by differing iterative functions or carrier molecules used. Therefore a significant amount is known and well characterized for these enzymes. Nonetheless, many polyketides have no known biosynthetic pathway associated with

their synthesis (Shen, 2003), making it difficult to construct microbial cell factories purposed toward producing these particular polyketide structures. One method to reduce this complexity is to formulate design rules around very simple structures and systems as a starting point.

Type III PKSs are a good candidate for this endeavor since these enzymes are iterative multifunctional complexes often needing only one gene with no additional activities required (Hertweck, 2009). In addition, significant structural information has been gathered on these enzymes, leading to an opportunity for guided rational engineering (Steiner and Schwab, 2012). Specifically, crystallographic studies on *Medicago sativa* chalcone synthase (CHS2) and *Gerbera hybrida* 2-pyrone synthase (2-PS) suggest that a very simple organization of few residues can dictate the entire specificity and selectivity of the enzyme (Jez et al., 2002). The 2-PS in particular has also gained traction for the production of a useful chemical intermediate triacetic aid lactone (TAL) (Cardenas and Da Silva, 2014; Chia et al., 2012; Tang et al., 2013; Xie et al., 2006). Using novel chemical catalysis routes, TAL can act as intermediary to many useful materials including the food preservative sorbic acid, currently synthesized using non-renewable fossil fuel derivatives (Chia et al., 2012). Recent studies have made TAL biologically using wildtype 2-PS and 2-PS variants in the microbial host *Escherichia coli* (Tang et al., 2013; Xie et al., 2006) and 2-PS in *Saccharomyces cerevisiae* (Cardenas and Da Silva, 2014), as well as modified PKSs (Type I 6MSAS, Type III CHS) with 2-PS activity (Cardenas and Da Silva, 2014; Xie et al., 2006). In the study using 2-PS variants, random mutagenesis and fluorescence sorting were used to identify the variants providing increased TAL levels in *E. coli* (Tang et al., 2013). After constructing the library and screening, selection of strong performers from high producing cultures only identified six residue modifications that conferred a significant change in TAL production, of which two were silent

mutations. This illustrates once again that while directed evolution approaches are rapid, they are also limited in the information gained. Therefore, having a rational approach to aid design in an attractive and simple platform may benefit the engineering of other related enzymes.

In this study, structural information on this class of PKS enzymes was employed to guide rational design of the 2-pyrone synthase for increased catalytic efficiency and stability. As enzyme variants were constructed and tested, subsequent variants were developed using the rules which were learned from the prior round using both *in vitro* and *in vivo* data as output. In total, four generations of 2-PS variants were created, and structure-function relationships relevant to biological conditions were elucidated. This combined approach of using *in vitro* and *in vivo* analysis for guided design led to the creation of many generations of 2-PS variants for evaluation for high level polyketide production. A high-producing variant was then selected for use in an engineered *S. cerevisiae* strain (BJ5464Δ*pep4*Δ*prb1*Δ*pyc2*Δ*nte1*) under fed-batch operation. Three fermentation feed strategies were evaluated and changes in productivity, yield, and titer were determined. This novel approach in combining *in vitro* and *in vivo* methods for biocatalyst design highlights the importance of a rational strategy for the development of protein design rules that could be applied to other enzyme systems.

3.3. Materials and Methods

3.3.1 Strains and plasmids

Escherichia coli BL21 (DE3) (EMD Millipore, Billerica, MA) cells were used to subclone *G. hybrida* 2-PS and all variants (provided by Bowman, M.; Salk Institute) into the pHIS8 expression vectors (Jez et al., 2000b). Heterologous expression and purification of the 2-PS and all variants for *in vitro* assays were performed as previously described (Jez et al., 2000b). *E. coli* strain XL1-Blue (Stratagene, Santa Clara, CA) was used for amplification of yeast shuttle vector pXP842 (Shen et al., 2012). Protease-limited *S. cerevisiae* strain BJ5464 (Jones, 1991) was used as the base strain for TAL production. *S. cerevisiae* strain BY4741Δ*trp1* (Open Biosystems, Huntsville, AL) and related protease-limited strains were also used for comparison.

Our pXP vector system (Shen et al., 2012) was used to carry all 2-PS enzyme variants. Specifically, the high copy 2µ-based pXP842 vector was used, harboring the glucose-repressed *ADH2* promoter and *CYC1* terminator with loxP-flanked *URA3* selection marker. Yeast cells were transformed as previously described (Gietz et al., 1992; Hill et al., 1991) using selective SDC(A) plates. Cell colonies were allowed to grow for 3-5 days before generating inoculum cultures or glycerol stocks for long-term storage. A list of all strains and plasmids can be found in Table 3.1 and Table 3.2.

The *Gerbera hybrida g2ps1* gene encoding the 2-PS and all variants were PCR amplified from the pHIS8 cassettes (Jez et al., 2000a). Following *SpeI* and *XhoI* digestion, the gene was inserted into pXP842 using the Rapid DNA Ligation Kit (Thermo Scientific, Waltham, MA) to

Table 3.1. List of plasmids and strains in *E. coli*

Plasmids	Description	Source
pHIS8-WT	pET-28a(+), 8x Histag-2-PS	Bowman, M. (Salk Institute)
pHIS8-E1	pET-28a(+), 8x Histag-2-PS (C35A)	Bowman, M. (Salk Institute)
pHIS8-E2	pET-28a(+), 8x Histag-2-PS (C35S)	Bowman, M. (Salk Institute)
pHIS8-E3	pET-28a(+), 8x Histag-2-PS (C35S, L268F)	Bowman, M. (Salk Institute)
pHIS8-E4	pET-28a(+), 8x Histag-2-PS (C35S, L202M)	Bowman, M. (Salk Institute)
pHIS8-E5	pET-28a(+), 8x Histag-2-PS (C35S, L202F)	Bowman, M. (Salk Institute)
pHIS8-E6	pET-28a(+), 8x Histag-2-PS (C35S, L268M)	Bowman, M. (Salk Institute)
pHIS8-E7	pET-28a(+), 8x Histag-2-PS (C35S, L268F)	Bowman, M. (Salk Institute)
pHIS8-E8	pET-28a(+), 8x Histag-2-PS (C35S, I343M)	Bowman, M. (Salk Institute)
pHIS8-E9	pET-28a(+), 8x Histag-2-PS (C35S, I343F)	Bowman, M. (Salk Institute)
pHIS8-E10	pET-28a(+), 8x Histag-2-PS (C35S, M259F)	Bowman, M. (Salk Institute)
pHIS8-E11	pET-28a(+), 8x Histag-2-PS (C35S, T137L)	Bowman, M. (Salk Institute)
pHIS8-E12	pET-28a(+), 8x Histag-2-PS (C35S, T137M)	Bowman, M. (Salk Institute)
pHIS8-E13	pET-28a(+), 8x Histag-2-PS (C35S, T137F)	Bowman, M. (Salk Institute)
pHIS8-E14	pET-28a(+), 8x Histag-2-PS (C35S, I201L)	Bowman, M. (Salk Institute)
pHIS8-E15	pET-28a(+), 8x Histag-2-PS (C35S, I201M)	Bowman, M. (Salk Institute)
pHIS8-E16	pET-28a(+), 8x Histag-2-PS (C35S, I201F)	Bowman, M. (Salk Institute)
pHIS8-E17	pET-28a(+), 8x Histag-2-PS (C35S, L202F, L268M)	Bowman, M. (Salk Institute)
pHIS8-E18	pET-28a(+), 8x Histag-2-PS (C35S, L202F, L268F)	Bowman, M. (Salk Institute)
pHIS8-E19	pET-28a(+), 8x Histag-2-PS (C35S, L268M, I343M)	Bowman, M. (Salk Institute)
pHIS8-E20	pET-28a(+), 8x Histag-2-PS (C35S, C65A)	Bowman, M. (Salk Institute)
pHIS8-E21	pET-28a(+), 8x Histag-2-PS (C35S, C65S)	Bowman, M. (Salk Institute)
pHIS8-E22	pET-28a(+), 8x Histag-2-PS (C35S, C89A)	Bowman, M. (Salk Institute)
pHIS8-E23	pET-28a(+), 8x Histag-2-PS (C35S, C89S)	Bowman, M. (Salk Institute)
pHIS8-E24	pET-28a(+), 8x Histag-2-PS (C35S, C135A)	Bowman, M. (Salk Institute)
pHIS8-E25	pET-28a(+), 8x Histag-2-PS (C35S, C135S)	Bowman, M. (Salk Institute)
pHIS8-E26	pET-28a(+), 8x Histag-2-PS (C35S, 195A)	Bowman, M. (Salk Institute)
pHIS8-E27	pET-28a(+), 8x Histag-2-PS (C35S, 195S)	Bowman, M. (Salk Institute)
pHIS8-E28	pET-28a(+), 8x Histag-2-PS (C35S, 346A)	Bowman, M. (Salk Institute)
pHIS8-E29	pET-28a(+), 8x Histag-2-PS (C35S, 346S)	Bowman, M. (Salk Institute)
pHIS8-E30	pET-28a(+), 8x Histag-2-PS (C35S, 372A)	Bowman, M. (Salk Institute)
pHIS8-E31	pET-28a(+), 8x Histag-2-PS (C35S, 372S)	Bowman, M. (Salk Institute)
pHIS8-E32	pET-28a(+), 8x Histag-2-PS (C89S, C372S)	Bowman, M. (Salk Institute)
pHIS8-E33	pET-28a(+), 8x Histag-2-PS (C35S, C89S, C372S)	Bowman, M. (Salk Institute)
pHIS8-E34	pET-28a(+), 8x Histag-2-PS (C89S, C346A, 372S)	Bowman, M. (Salk Institute)

Table 3.1. List of plasmids and strains in *E. coli* (cont.)

Plasmids	**Description**	**Source**
pHIS8-E35	pET-28a(+), 8x Histag-2-PS (C35S, C89S, C346A, 372S)	Bowman, M. (Salk Institute)
pHIS8-E36	pET-28a(+), 8x Histag-2-PS (C65, C89S, 372S)	Bowman, M. (Salk Institute)
pHIS8-E37	pET-28a(+), 8x Histag-2-PS (C35S, C65, C89S, 372S)	Bowman, M. (Salk Institute)
pHIS8-E38	pET-28a(+), 8x Histag-2-PS (C89S, C346A, C195S, 372S)	Bowman, M. (Salk Institute)
pHIS8-E39	pET-28a(+), 8x Histag-2-PS (C35S, C89S, C195S, C346A 372S)	Bowman, M. (Salk Institute)
pHIS8-E40	pET-28a(+), 8x Histag-2-PS (C65S, C89S, C195S, C346A, 372S)	Bowman, M. (Salk Institute)
pHIS8-E41	pET-28a(+), 8x Histag-2-PS (C35S, C65S, C89S, C195S, C346A, 372S)	Bowman, M. (Salk Institute)
Strains	**Description**	**Source**
E. coli BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	EMD Millipore

Table 3.2. List of plasmids and strains in *S. cerevisiae*

Plasmids	Description	Source
pXP842	2μ vector, *ADH2* promoter, *CYC1* terminator, *URA3* selectable marker	Shen et al. (2012)
pXP842-WT	pXP842 harboring 2-PS	This study
pXP842-E1	pXP842 harboring 2-PS (C35A)	This study
pXP842-E2	pXP842 harboring 2-PS (C35S)	This study
pXP842-E3	pXP842 harboring 2-PS (C35S, L268F)	This study
pXP842-E4	pXP842 harboring 2-PS (C35S, L202M)	This study
pXP842-E5	pXP842 harboring 2-PS (C35S, L202F)	This study
pXP842-E6	pXP842 harboring 2-PS (C35S, L268M)	This study
pXP842-E7	pXP842 harboring 2-PS (C35S, L268F)	This study
pXP842-E8	pXP842 harboring 2-PS (C35S, I343M)	This study
pXP842-E9	pXP842 harboring 2-PS (C35S, I343F)	This study
pXP842-E10	pXP842 harboring 2-PS (C35S, M259F)	This study
pXP842-E11	pXP842 harboring 2-PS (C35S, T137L)	This study
pXP842-E12	pXP842 harboring 2-PS (C35S, T137M)	This study
pXP842-E13	pXP842 harboring 2-PS (C35S, T137F)	This study
pXP842-E14	pXP842 harboring 2-PS (C35S, I201L)	This study
pXP842-E15	pXP842 harboring 2-PS (C35S, I201M)	This study
pXP842-E16	pXP842 harboring 2-PS (C35S, I201F)	This study
pXP842-E17	pXP842 harboring 2-PS (C35S, L202F, L268M)	This study
pXP842-E18	pXP842 harboring 2-PS (C35S, L202F, L268F)	This study
pXP842-E19	pXP842 harboring PS (C35S, L268M, I343M)	This study
pXP842-E20	pXP842 harboring 2-PS (C35S, C65A)	This study
pXP842-E21	pXP842 harboring 2-PS (C35S, C65S)	This study
pXP842-E22	pXP842 harboring 2-PS (C35S, C89A)	This study
pXP842-E23	pXP842 harboring 2-PS (C35S, C89S)	This study
pXP842-E24	pXP842 harboring 2-PS (C35S, C135A)	This study
pXP842-E25	pXP842 harboring 2-PS (C35S, C135S)	This study
pXP842-E26	pXP842 harboring 2-PS (C35S, 195A)	This study
pXP842-E27	pXP842 harboring 2-PS (C35S, 195S)	This study
pXP842-E28	pXP842 harboring 2-PS (C35S, 346A)	This study
pXP842-E29	pXP842 harboring 2-PS (C35S, 346S)	This study
pXP842-E30	pXP842 harboring 2-PS (C35S, 372A)	This study
pXP842-E31	pXP842 harboring 2-PS (C35S, 372S)	This study
pXP842-E32	pXP842 harboring 2-PS (C89S, C372S)	This study
pXP842-E33	pXP842 harboring 2-PS (C35S, C89S, C372S)	This study
pXP842-E34	pXP842 harboring 2-PS (C89S, C346A, 372S)	This study

Table 3.2. List of plasmids and strains in *S. cerevisiae* (contd.)

Plasmids	Description	Source
pXP842-E35	pXP842 harboring 2-PS (C35S, C89S, C346A, 372S)	This study
pXP842-E36	pXP842 harboring 2-PS (C65, C89S, 372S)	This study
pXP842-E37	pXP842 harboring 2-PS (C35S, C65, C89S, 372S)	This study
pXP842-E38	pXP842 harboring 2-PS (C89S, C346A, C195S, 372S)	This study
pXP842-E39	pXP842 harboring 2-PS (C35S, C89S, C195S, C346A 372S)	This study
pXP842-E40	pXP842 harboring 2-PS (C65S, C89S, C195S, C346A, 372S)	This study
pXP842-E41	pXP842 harboring 2-PS (C35S, C65S, C89S, C195S, C346A, 372S)	This study
Strains	**Description**	**Source**
BY4741	*MAT*a *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Open Biosystems
BY4741Δ*trp1*	BY4741 *trp1::KanMX*	Open Biosystems
BJ5464	*MAT*α *ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL*	E.W. Jones (1991)
BJ*Δpyc2Δnte1*	BJ5464 *pyc2::TRP1 nte1::LEU2*	Cardenas and Da Silva (2014)

Table 3.3. List of primers used

Primer Name	Primer Sequence
2PS-F	TCCCACTAGTAAAAAAATGCATCATCATCATCATCATGGATCTTTACTCATCCGATGTGGAG
2PS-R	TGGTAATCCTCGAGTCAGTTTCCATTGGCAACCGCAGCAGTAACG

construct pXP218-2PS and pXP842-2PS. A list of primers used for plasmid construction can be found in Table 3.3. Plasmid recovery was performed using the GeneJet™ Plasmid Miniprep Kit (Thermo Scientific, Waltham, MA) and DNA sequence analysis confirmed the correct sequence of all PCR-amplified inserts (GeneWiz, South Plainfield, NJ; Eton Biosciences, San Diego, CA).

The KOD Hot-start polymerase (EMD Chemicals, San Diego, CA) was used in PCR reactions for plasmid constructs. Restriction enzymes, T4 DNA ligase, Taq DNA polymerase, and deoxynucleotides (dNTPs) were purchased from New England Biolabs (Ipswich, MA). Oligonucleotide primers were purchased from IDT DNA (San Diego, CA).

3.3.2 Media and cultivation

Luria-Bertani (LB) media was used for proliferation of strain XL1-Blue with 150 mg/L ampicillin added for selection of plasmid-containing cells (Sambrook and Russell, 2001). For yeast, complex YPD media (1% dextrose; 1% Bacto yeast extract; 2% Bacto peptone) was used for 48h cultivations. For inoculum cultures, *S. cerevisiae* strains were grown for 16 h overnight in 5 mL selective SDC(A) (1% dextrose, 0.67% yeast nitrogen base, 0.5% Bacto casamino acids, 0.5% ammonium sulfate and 100 mg/L adenine) media in an air shaker (New Brunswick Scientific) at 250 rpm and 30°C, and used to inoculate cultures to an initial cell density (OD_{600}) of 0.3 (Shimadzu UV-2450 UV-VIS Spectrophotometer, Columbia, MD). During the cultivation, a 48h sample was taken where cell densities were determined and the samples were centrifuged at 3,000 rpm (2,600 g) for 5 min at 4°C (Beckman GS-6R Centrifuge, Brea, CA). The supernatants were stored at 4°C for HPLC analysis of TAL levels in the culture broth.

3.3.3 *In vitro* 2-PS activity assays

Standard activity assay conditions for screening included 100mM potassium phosphate (pH 7.0), 60µM [2-^{14}C]malonyl-CoA, 60µM acetyl-CoA, and 250 nM protein in a 500µL reaction volume. Reactions were quenched after 30 minutes with 5% acetic acid, and extracted with ethyl acetate at 25°C. Steady state kinetic parameters were determined from initial velocity measurements in which product formation was linear over the time periods monitored, using standard assay conditions with 60µM malonyl-CoA and varying acetyl-CoA from 0.5 to 64 µM. Reactions were similarly quenched as stated above. Quantitation and fitting of data were performed as previously described (Jez et al., 2000b).

3.3.4 2-PS oxidation assays

Preliminary 2-PS oxidation analysis was performed *in vitro* on purified enzyme isolates from *E. coli* incubated with 100mM potassium phosphate (pH 7.0), in the presence or absence of 10mM reducing agent tris (2-carboxyethyl) phosphine (TCEP). Activity assays were performed as above after purified enzymes were allowed to incubate at room temperature for a period between 0-2 h in assay buffer. The half-life (time to reach 50% activity) was determined from this data. Activity measurements were normalized to the percent activity at time zero (T = 0).

For 2-PS oxidation studies using TAL-producing yeast, standard DTNB assay conditions included 150 mM tris buffer (pH 8.0), 4 mM DTNB, and 2nmol of purified protein. The purified protein was extracted from yeast cultures at 12 h intervals from 24 to 48 h using a modified purification scheme as described (Jez et al., 2000b). Samples were taken from cultivations done aerobically and anaerobically, and protein extracts were protected from further environmental

oxidation via nitrogen-sparging until performing the assay. No reducing agents were added during the purification which was terminated following elution through Ni-NTA columns. The free thiol levels were reported as a concentration (μM) of free thiols after converting the absorbance of samples at 412nm (Shimadzu UV-2450 spectrophotometer; Columbia, MD) using a standard curve from dithiothreitol (DTT) standards with concentrations ranging from 0-20 μM.

3.3.5 HPLC assay

The concentration of triacetic acid lactone was measured by HPLC using a Shimadzu HPLC system: LC-10AT pumps (Shimadzu), UV-VIS detector (SPD-10A VP, Shimadzu), Zorbax SB-C18 reversed-phase column (2.1x150 mm, Agilent Technologies). Acetonitrile buffered in 1% acetic acid was used as the mobile phase, while HPLC grade water buffered in 1% acetic acid was used as the aqueous phase. A gradient program using a 95-85% Pump B gradient (H_2O with 1% acetic acid) provided an elution time of approximately 12 minutes (flow rate 0.25 mL/min, column temperature 25°C).

3.3.6 Batch and fed-batch fermentations

A New Brunswick BioFlo III system equipped with a 2.5 L vessel was employed to maintain cultivation parameters during batch and fed-batch operation. Initial 5mL overnight cultures of the engineered strain were grown to inoculate a 50mL shake flask culture to be used as the seed inoculum. Following inoculation to an OD_{600} of 0.3 in YPD, the fermentor pH was controlled at pH 6 by automatic supply of either 6M sodium hydroxide or 6M hydrochloric acid. Throughout the fermentation, Antifoam SE-15 (Sigma) was pumped for maintenance of foaming and controlled by the antifoam probe. Dissolved oxygen was initially maintained by using a 400 rpm agitation

speed, and supplying sparged air at a rate of 0.2 vvm in order to maintain dissolved oxygen levels at 20%. While agitation was kept constant, the aeration rate was allowed to increase in order to ensure DO levels could be maintained. No additional control was employed for batch operation. The cultivation was allowed to proceed under these conditions for approximately 12 h for glucose-fed and 24 h for ethanol-fed operation before the selected carbon source feed was initiated at 1.5 mL/h (glucose feed, 3.6M) and 2.5 mL/h (ethanol feed, 16% v/v). For use of a fortified glucose feed, nitrogen-containing amino acids were supplemented into the 3.6M glucose feed at a carbon-to-nitrogen ratio of 35:1 (g/g) for nitrogen-limited conditions using a modified amino acid profile from the literature (Albers et al., 1996; Manikandan et al., 2010) The amino acid supplementation was: 869 mg/L L-arginine, 960 mg/L L-aspartic acid, 12.4 mg/L L-glutamine, 360 mg/L L-glutamic acid, 84.5 mg/L L-lysine-HCL, 100 mg/L adenine sulfate. The glucose feed rate varied to maintain DO levels at 20% with constant agitation (400 rpm) and aeration (0.8 vvm). The ethanol feed rate was kept constant throughout the fermentation; therefore, DO levels were maintained using an agitation-coupled DO control using a maximum agitation rate of 425 rpm and constant aeration rate of 0.2 vvm.

3.3.7 Yield calculations

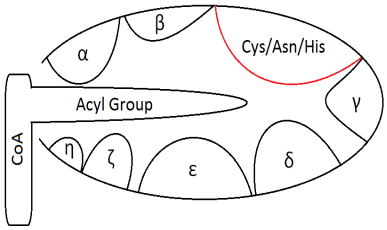
Yield was calculated as the total TAL formed divided by the total carbon supplied to the culture. Maximum theoretical yield was calculated based on the metabolic reactions from the carbon source to TAL assuming all carbon was converted to product, and was determined to be 0.47 g/g on glucose, and 0.91 g/g on ethanol. Therefore, the percent of the theoretical yield of TAL on glucose was calculated as: [TAL (g/L) / Carbon Fed (g/L) / X (g/g Carbon) * 100%].

3.4. Results and Discussion

3.4.1 Modulating activity of the active site

Prior work in our collaborator's lab showed that modulation of a Type III PKS active site pocket is enough to alter final product formation (Jez et al., 2000a). Enzyme promiscuity is well-known for PKSs in nature (Pfeifer and Khosla, 2001); therefore steric modulation was performed in order to improve specificity, and performance was verified by changes in enzyme activity and kinetic parameters. Based on the cavity conformation of the active site, specific amino acids were selected for modification (Figure 3.1). As a means to reduce promiscuity in substrates accommodated by the 2-pyrone synthase (data not shown), a constriction of the active site was performed by introduction of bulkier residues as substitutions. To this end, amino acids were modified to either a leucine, methionine, or phenylalanine residue. For amino acid residues that are one of these three, only larger residues if available were substituted. To perform *in vitro* assays, the highly exposed C35 cysteine was also substituted for a serine (C35S) to reduce thiol oxidation during execution of the assay.

A)



B)

- α = 343 (CHS 338) 3X mutation site
- β = 137 (CHS 132)
- γ = 259 (CHS 254)
- δ = 261 (CHS 256) 3X mutation site
- ε = 268 (CHS 263)
- ζ = 201 (CHS 196)
- η = 202 (CHS 197) 3X mutation site

C)

Mutation Site	Label	Function
I201, L202	η, ζ	Prevent long extensions
L268	ε	Limit loading specificity
I343, M259	α , γ	Reduce pocket size
T137	β	Derailment pocket
C35	---	Reduce oxidation sensitivity

Figure 3.1 (A) Illustration of amino acid positions conferring size in the active site, as denoted by Greek symbols and (C) corresponding function. (B) Amino acid residue positions, highlighting CHS-conserved residues in the naringenin-producing chalcone synthase leading to 2-PS activity (Jez et al., 2000a)

Prior to testing the enzyme performance in *S. cerevisiae*, preliminary *in vitro* assays were performed on a selection of variants that purified well, and steady state kinetic parameters were determined (Figure 3.2, Table 3.4). As discussed, variations in the active site were coupled to the C35S substitution. From this subset tested (Gen-1), a significant change in both k_{cat} and K_M was possible. Modifying at C35 to C35S drastically reduced the K_M without altering the catalytic efficiency, suggesting that substrate binding had been altered in this variant. This was unexpected since C35S is not near the active site nor substrate binding region. The improvements in K_M values were significantly larger (up to 16X) than those of the k_{cat} (up to 1.9X), suggesting that turnover is not limiting, and substrate specificity is the biggest contribution to improvements in engineered 2-PS variants. 2-PS enzyme variant E5 had the highest change in both k_{cat} and K_M, resulting in a k_{cat} / K_M nearly 30-fold better than the wildtype (WT) 2-PS. Variant E15 had a significant reduction in k_{cat} (15% of WT) and K_M (15% of WT) which implies this I201M modification is severely affecting the turnover of the enzyme. These 2-PS variants were evaluated in order to have a baseline for comparing performance *in vitro* and *in vivo*.

Following these initial studies, TAL was produced *in vivo* in by transforming the yeast strain BJ5464 using yeast shuttle vectors carrying the complete Gen-1 set of 2-PS variants (Table 3.2). The transformed strains were then cultivated in 1% YPD, and titers measured after 48 h of cultivation, as shown for variants E2-E19 in Figure 3.3. As a baseline for *in vivo* TAL levels, the wildtype 2-PS was expressed and was found to produce 0.48 g/L TAL. First, we compared *in vivo* data to what was observed *in vitro* (Figure 3.3). Changes in E2 are consistent in

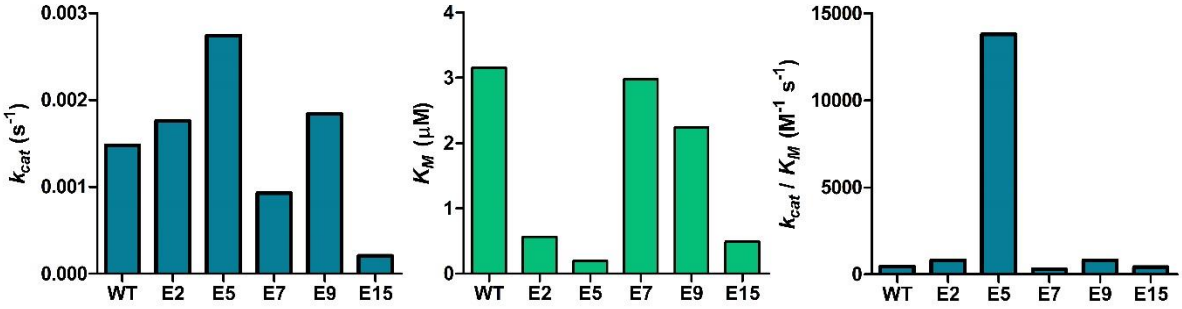


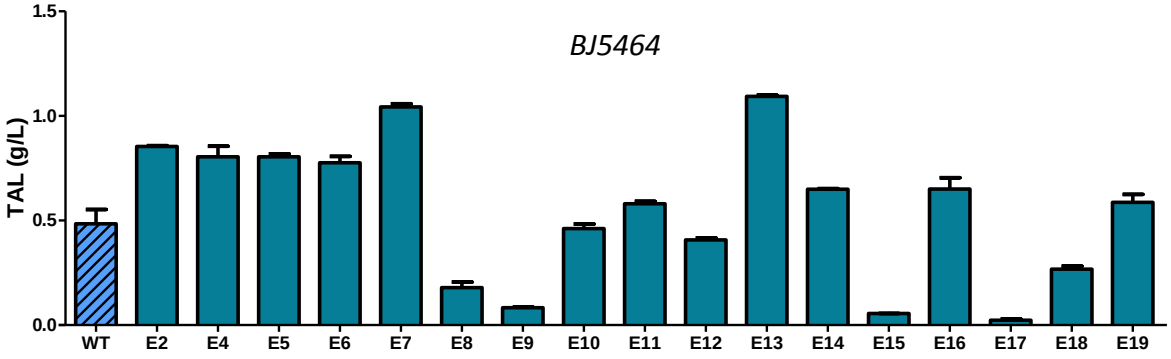
Figure 3.2 Steady-state kinetic parameters for 2-PS and initial 2-PS variants (Data from Table 3.4).

Table 3.4 Steady state kinetic parameters for wildtype 2-PS and 2-PS variants (Gen-1 subset)

	k_{cat} (s^{-1})	K_M (μM)	k_{cat} / K_M ($M^{-1}s^{-1}$)
WT	0.00148	3.15	471
E2 [C35S]	0.00176	0.565	3108
E5 [C35S,L202F]	0.00274	0.199	13,790
E7 [C35S,L268F]	0.000928	2.981	311
E9 [C35S,I343F]	0.00184	2.246	819
E15 [C35S,I201M]	0.000209	0.488	430

Values for k_{cat}, K_M, and k_{cat} / K_M are s^{-1}, μM, and $M^{-1}s^{-1}$, respectively, and represent mean values (n=3)

both studies, where reduced thiol oxidation translated to increased TAL titer (0.85 g/L) and appears to support improved stability inside the cell. In E5, which was the best performer by far *in vitro* (30-fold higher k_{cat} / K_M), a nearly 70% increase was found (0.80 g/L), although slightly lower than E2 (C35S alone); this suggests that *in vivo*, E5 no longer is an ideal candidate for TAL synthesis. In regards to E7, E9, and E15, *in vivo* data was more inconsistent with the *in vitro* results. From the data shown in Table 3.4, E7 (C35S, L268F) was found to have a slightly lower catalytic efficiency than wildtype, yet our *in vivo* data indicates that E7 is one of the best enzyme



C35 Residue	C	S	S		S		S		S	S			S			S	S	S
Position Modified	---	C35	L202		L268		I343		M259	T137			I201			L202, L268	L202, L268	L268, I343
New Residue	---	S	M	F	M	F	M	F	F	L	M	F	L	M	F	F,M	F,F	M, M

Figure 3.3 Triacetic acid lactone titers (g/L) *in vivo* for 1st generation (Gen-1) variants in strain BJ5464. All variants have C35S in addition to one or multiple active site modifications as shown. Wildtype 2-PS is shown for comparison (light blue, hatched). Bars represent mean values ± standard deviation (n=6 independent experiments).

variants for TAL, with a titer increase over 2-fold that of the wildtype (1.04 g/L) and over 20% higher than E2 (C35S). In addition, E9 had a nearly 2-fold higher k_{cat} / K_M *in vitro* and only made ~80 mg/L TAL (17% of WT) *in vivo*, suggesting *in vivo* processes have rendered this variant a poor performer in the cell. With respect to E15, only ~50 mg/L TAL was made in strain BJ5464 (Figure 3.3) even though catalytic efficiency was similar to wildtype. Since E9 had a severely low turnover, the reduced K_M appears to have had no impact in our substrate-limited cultures, causing poor *in vivo* performance.

Analyzing the rest of Gen-1 variants provides further guidance in devising engineered enzyme constructs. For modifiers of L202 (E4 and E5) which is in close proximity to docking of the acyl chain (Figure 3.1), either M or F substitution gives similar enhancement to *in vivo* TAL levels. A similar trend was observed for the L268 variants (E6 and E7) in which both substitutions allowed

for increased TAL over the wildtype, although E7 was the only one able to improve beyond E2. Bulkier residues at I343 (E8 and E9) both performed poorly, and in conjunction with *in vitro* data for E9, suggests that there is no room for improvement at this position, and any cavity reduction here will render the enzyme with low activity. For changes at position M259 (E10), a reduction in TAL titer was found relative to E2, suggesting that bulkier residues are detrimental at this position. The T137 position was also interesting in that while E11 and E12 variants truncated the pocket further than normal, only T137F (E13) exceeded E2 TAL titers to 1.1 g/L (30% above E2). This result implies that the nucleophilic threonine is more accommodating of bulky aromatics than hydrophobic residues like leucine and methionine. In addition, since T137M was the worst of the E11-E13 cluster, it is likely that sulfur chemistry with neighboring side chains may cause an issue with TAL formation. At I201, neither leucine nor phenylalanine substitutions were capable of conferring increased TAL over E2, with additional poor performance in I201M in both *in vitro* kinetics and *in vivo* titers. This would suggest that I201 is not amenable to modification for improving enzyme efficiency and should be left unmodified. Further combinations of these substitutions were performed (E17-E19), and it appears coupling of multiple bulky residues into one enzyme cavity restricts the flexibility too much. From these results, it was determined that E2 was profoundly critical in conferring increased activity both *in vitro* and *in vivo*. Additionally, based on the predicted function at each residue (Figure3.1), 2-PS cavity modifications at T137, L202, and L268 likely favored a restricted specificity of both substrates and intermediates for increased activity and TAL titers *in vivo*.

3.4.2 Importance of C35 and role of C35S variant

In the process of evaluating the previously discussed subset of enzyme variants *in vitro*, the C35S substitution was revealed to have a significant role in 2-PS catalytic efficiency. The data suggests C35S is affecting K_M to a greater degree than k_{cat}, and therefore altering the amount of substrate required to reach half of the V_{max}. Based off of the known crystal structure of 2-PS (Figure 3.4), C35 is not proximal to the active site, but is exposed to the surface of the enzyme. We therefore sought to investigate a possible mechanism for why C35 was so important. To do this, new variants (Gen-2) were constructed for evaluation of each E7 modification independent of the other (E2 versus E3), as well as using a C35A modifier (E1). The C35A substitution allows us to introduce a similar-sized residue to the wildtype but without a reactive side group. In addition to wildtype, the E1, E2, E3, and E7 variants (note that E2 and E7 were already

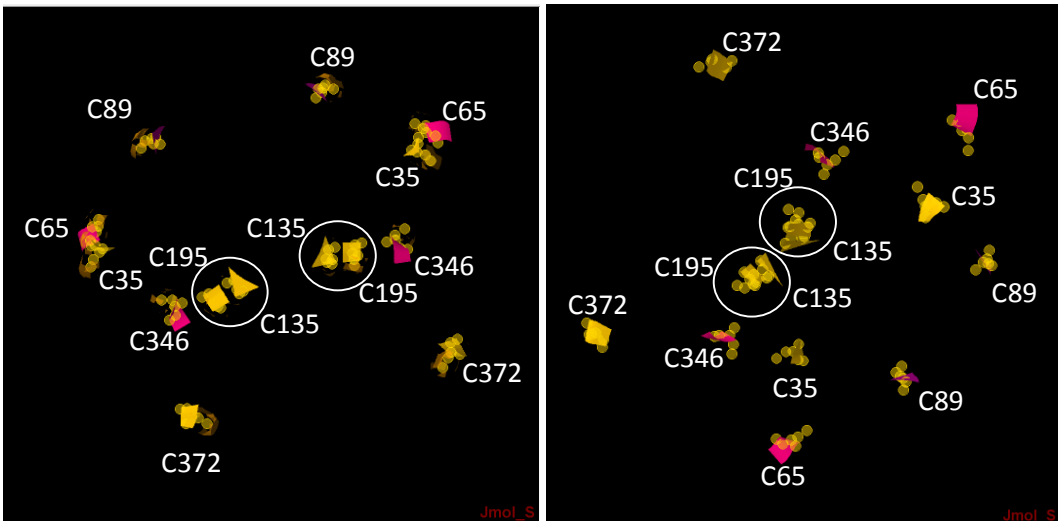
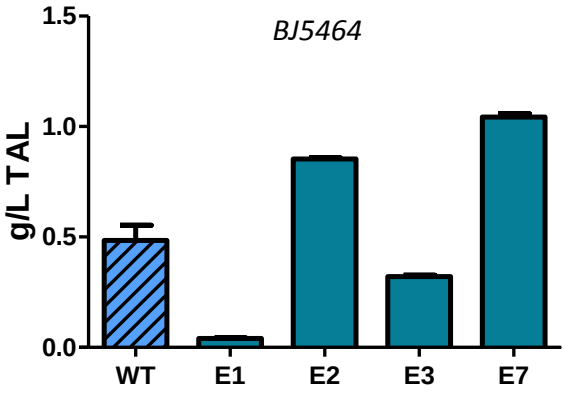


Figure 3.4 Jmol representation of 2-pyrone synthase crystal structure (E.C.2.3.1.74, pdbj.org). Structure illustrated as active homodimer with all 14 cysteine residues (active site omitted) highlighted (yellow) along with side chain thiol groups (pink) and the disulfide-bonded sites (circle, white).

constructed previously) were expressed in *S. cerevisiae* using pXP842, and cultivated for 48 h before determining *in vivo* TAL levels (Figure 3.5). In contrast to E2 and E7 which previously showed increased levels, we see that *in vivo* TAL for E1 and E3 (with C35A) were much lower than the wildtype 2-PS. The E1 variant made ~40 mg/L TAL and E3 produced ~320 mg/L, corresponding to a 90% and 50% loss in titer relative to wildtype, respectively. Kinetic parameters were also determined for the E1 and E3 variants (Table 3.5). Using *in vitro* kinetic measurements, E1 increased 6-fold and E3 decreased 2-fold in catalytic efficiency (k_{cat}/K_M), respectively. The strong performance of C35S in yeast relative to C35A in conjunction with the high k_{cat}/K_M of C35A relative to C35S suggests a discrepancy between what determines 2-PS performance at the kinetics level *in vitro* versus *in vivo*. Additionally, the 2-PS crystal structure shows that C35 is fully exposed to the surface (Figure 3.4). Therefore substituting a polar side chain like serine at C35 is



Original Residue	---	C35		L268	C35, L268
New Residue	---	A	S	F	S, F

Figure 3.5 Triacetic acid lactone titers (g/L) *in vivo* for Gen-2 variants and comparison variants E2 and E7 along with the wildtype (light blue, hatched) in strain BJ5464. Bars represent mean values ± standard deviation (n=6 independent experiments).

Table 3.5 Steady state kinetic parameters for wildtype 2-PS and 2-PS variants (Gen-2)

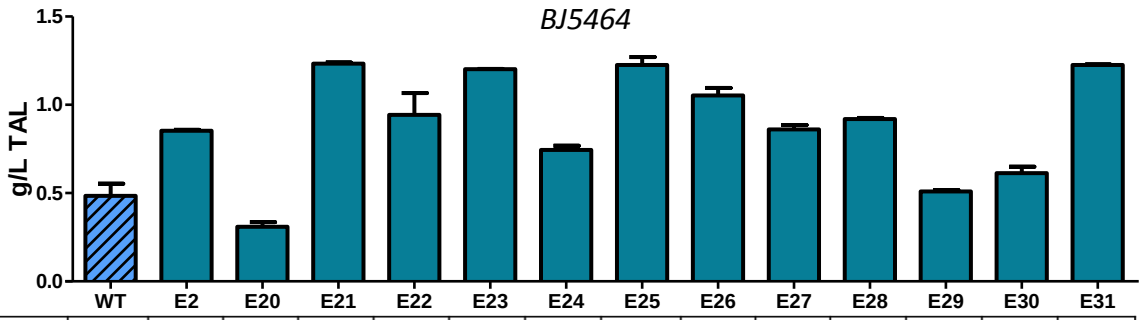
	k_{cat}	K_M	k_{cat} / K_M
WT	0.00148	3.15	471
E1 [C35A]	0.000972	0.336	2891
E3 [L268F]	0.00205	8.56	240

Values for k_{cat}, K_M, and k_{cat} / K_M are s^{-1}, µM, and $M^{-1}s^{-1}$, respectively and represent mean values (n=3)

more beneficial than a hydrophobic alanine and will be tolerant of the hydrophilic environment of the cytosol. With respect to E3, it appears that the reduced catalytic efficiency resulted in a poor performance *in vivo*. In addition, Gen-2 variants show that catalytic turnover is not playing a significant role, whereas K_M shows the biggest changes. These results would suggest that C35S and L268F work cooperatively *in vivo*, likely related to their proximity to the acyl-loading site. This new evidence highlights the importance of C35, a residue distant to the active site and exposed to the enzyme's surface.

3.4.3 Role of surface exposed 2-PS cysteine residues *in vivo*

The 2-pyrone synthase is a cysteine-rich protein, harboring 16 thiol side chains in an active homodimer conformation. Following the identification of the E2 variant and the robustness in both *in vitro* and *in vivo* environments, single cysteine residues outside of the conserved catalytic triad (i.e., C169 omitted) were modified by substitution with an A or S. These amino acids were chosen due to their comparable side chain size as well as the ability to introduce similar (serine, OH side group) or different (alanine, methyl group) properties. This allows determination of



	WT	E2	E20	E21	E22	E23	E24	E25	E26	E27	E28	E29	E30	E31
Modified Residues	---	35S	35S	35S	35S	35S	35S	35S	35S	35S	35S	35S	35S	35S
	---	---	65A	65S	89A	89S	135A	135S	195A	195S	346A	346S	372A	372S

Figure 3.6 Triacetic acid lactone titers (g/L) *in vivo* for Gen-3 variants as compared to wildtype (light blue, hatched) and E2 (C35S) in strain BJ5464. All variants have C35S in addition to one other cysteine residue as shown. Bars represent mean values ± standard deviation (n=6 independent experiments).

interactions that are most important at the different residue positions. All variants (Gen-3) additionally harbored the C35S modification, and *in vivo* production of TAL was once again evaluated (Figure 3.6). The percent increase in the titer relative to the E2 variant (C35S) was also tabulated (Table 3.6). The C65 residue had the largest difference between any A/S substitution pair evaluated. This residue is also the closest cysteine residue proximal to the CoA binding domain. Since the hydrophobic residue had such a negative effect *in vivo*, it is likely that side chain chemistry at this residue alters substrate recruitment in our strains. Three cysteines (C65, C89, and C372) are exposed to the cell surface similarly to C35. In all three cases, the C→S substitution was much more favorable (E21, E23, E31) than the alanine substitution. This parallels the result with the exposed C35, where the hydrophobic alanine was extremely detrimental to TAL formation *in vivo*.

For C135 and C195, crystal structure suggests they form a disulfide bond with each other on each subunit (Figure 3.4). Disruption of this disulfide bond can happen at either one of the two cysteines via A or S substitution (C135A, C135S, C195A, or C195S). Out of the four possible arrangements, the C135S/C195 conformation led to a substantial increase in TAL titer relative to C35S alone (Table 3.6). The C135/C195A was also able to increase performance *in vivo* (23% on titer basis). This data therefore suggests that this disulfide bridge is not strictly needed for activity, and can be manipulated to improve TAL production in *S. cerevisiae*. Specifically, C135 is slightly exposed and appears to prefer the polar serine modification. Regardless, multiple variants with changes to the highly exposed surface cysteines increased TAL levels and therefore should be the focus of future modifications. For C346, neither A or S substitutions yielded any improvement in TAL formation *in vivo* (Figure 3.6). This cysteine is highly buried inside the 2-PS monomer, which correlates with a preference to the hydrophobic alanine as seen in the buried C195. From these *in vivo* observations, cysteine modifications are important for *in vivo* performance regardless of spatial distance from the known catalytic machinery and substrate recruitment domain. Careful selection of serine and alanine substitutions can be a promising strategy to make better biocatalysts and improve performance of heterologous PKS enzymes in *S. cerevisiae.*

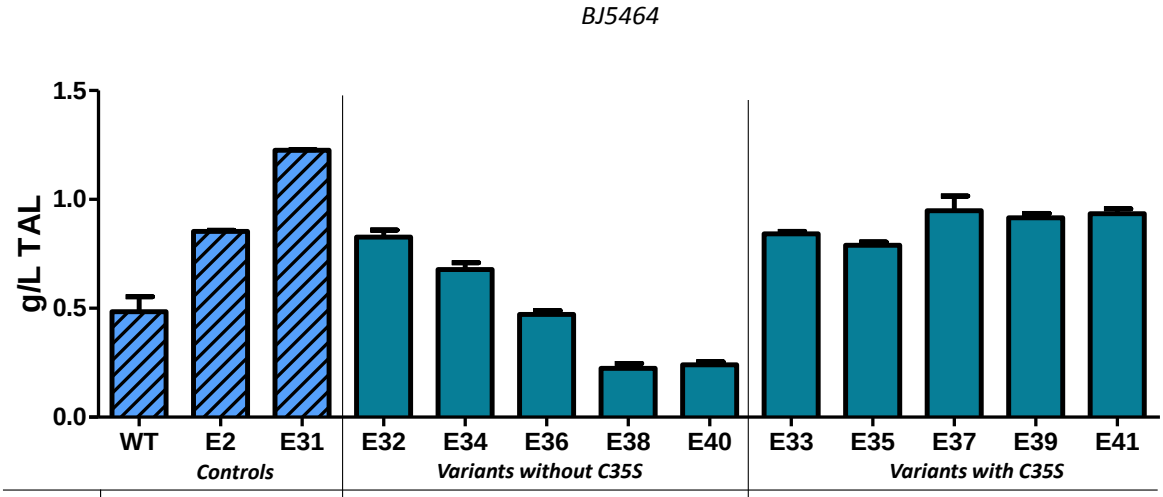
Table 3.6 Percent increased TAL titer (g/L) for Gen-3 variants, relative to E2 (C35S). Cysteine positions exposed to the surface are highlighted (bold, italicized text). Colored shading implies positive (green) versus negative (red) improvements. Values represent mean values (n = 6).

Exposed cysteines (beyond C35S)	TAL titer (g/L)	% above C35S
---	0.48	---
C65S	***1.23***	***44%***
C372S	***1.23***	***44%***
C135S	1.23	44%
C89S	***1.20***	***41%***
C195A	1.05	23%
C89A	0.94	10%
C346A	0.92	8%
C195S	0.86	1%
C135A	0.74	-13%
C372A	0.61	-28%
C346S	0.51	-40%
C65A	0.31	-64%

3.4.4 Combining cysteine substitutions in one 2-PS variant

The evaluation of cysteine residue substitutions not associated with the catalytic machinery of 2-PS has demonstrated that thiol side chains are instrumental in how the enzyme performs inside a biological environment. To understand the mechanisms that are responsible for this behavior, additional variants (Gen-4) were constructed. First, cysteines at varying positions which accommodate desired *in vivo* performance were selected, including C65, C89, C195, C346, and C372. These positions were modified in the presence or absence of the critical C35S substitution, and *in vivo* TAL levels were compared (Figure 3.7). For all variants incorporating C35S (E33-E41, odd only), TAL production was maintained at similar levels (0.80-0.95 g/L), although slightly lower than E31 (C35S, C372S). In the case of the same constructs not

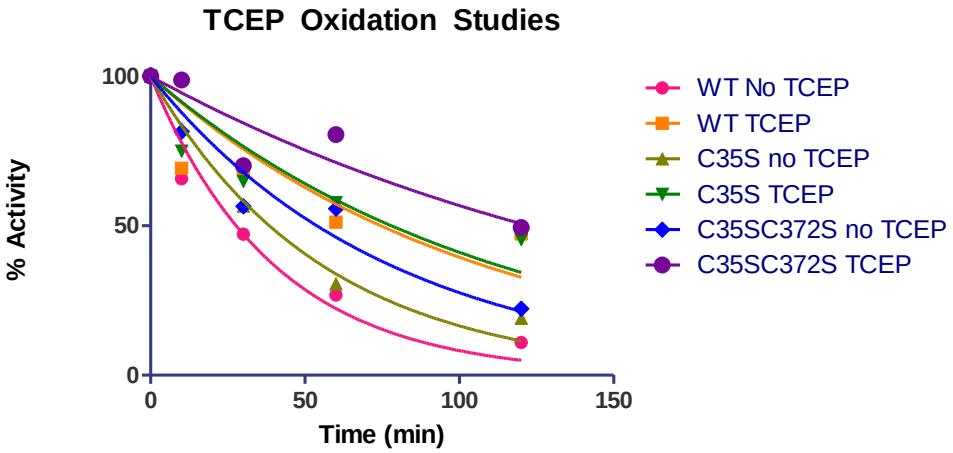
harboring C35 (E32-40, even only), subsequent introduction of favorable cysteine substitutions led to increasingly reduced levels of TAL *in vivo* (Figure 3.7). Knowing that C35S dictated whether or not improvements were achieved when using the L268F modification (Figure 3.5), this data further supports an active role of C35S on 2-PS performance. The kinetics data (Table 3.4) also demonstrate that C35 can have an effect on catalytic efficiency, primarily in substrate affinity. In addition, data from Gen-4 variants supports a concerted mechanism that selectively alters enzyme performance when C35S is absent. This suggests that cellular machinery native to the cytosol might act on the highly reactive thiol side groups, and should the C35 become a preferred site (i.e. when all other surface exposed cysteines are removed), TAL-synthesizing potential is greatly reduced.



	WT	E2	E31	E32	E34	E36	E38	E40	E33	E35	E37	E39	E41
C35S	-	+	+	-	-	-	-	-	+	+	+	+	+
C65S	-	-	-	-	-	+	-	+	-	-	+	-	+
C89S	-	-	-	+	+	+	+	+	+	+	+	+	+
C195S	-	-	-	-	-	-	+	+	-	-	-	+	+
C346A	-	-	-	-	+	-	+	+	-	+	-	+	+
C372S	-	-	+	+	+	+	+	+	+	+	+	+	+

Figure 3.7 Triacetic acid lactone titers (g/L) *in vivo* for Gen-4 variants as compared to wildtype (light blue, hatched), E2 (C35S), and E31 (C35S, C372S) in strain BJ5464. Bars represent mean values ± standard deviation (n=6 independent experiments).

Many enzymes have cysteines on the surface like 2-PS, often times as a defensive mechanism to oxidation which might render the enzyme inactive (Requejo et al., 2010). To determine any role in 2-PS, oxidation of the enzyme was evaluated. First, oxidation of the enzyme was evaluated by an *in vitro* assay using *E. coli* purified 2-PS enzymes WT, E2 (C35S), and E31 (C35S, C372S). In this assay, the purified enzymes were incubated in the presence or absence of the common reducing agent tris (2-carboxyethyl) phosphine (TCEP). Enzymes were allowed to undergo oxidation on the lab bench for up to 2 h, and specific activity was evaluated at different intervals (Figure 3.8). Relative to their starting activity, the percent activity of the enzyme as it was oxidized was determined and the enzyme half-life calculated. In reactions performed in the



	WT No TCEP	WT TCEP	C35S no TCEP	C35S TCEP	C35SC372S no TCEP	C35SC372S TCEP
Half Life	27.73	74.56	38.46	77.94	53.84	122.3

Figure 3.8 Percent activity levels as a function of time following incubation at room temperature (aerobic environment), performed on wildtype 2-PS, E2 (C35S), and E31 (C35S, C372S). Enzymes were simultaneously incubated in the presence or absence of reducing agent TCEP. Values shown represent mean values (n = 3).

presence of TCEP, we see that the wildtype half-life was 27 min, versus 38 min and 54 min for E2 and E31 variants. Additionally, all three enzyme half-lives increased when TCEP was introduced. For wildtype and E2, half-life improved to similar levels (75 and 78 min, respectively), but E31 increased further to 122 min. These results confirm that 2-PS is sensitive to oxidation. It also provides a measure for the amount of protection from oxidation that a variant might provide, as 10mM TCEP was enough reducing power to make wildtype comparably as good as C35S. This assay therefore confirms 2-PS can be engineered to have reduced oxidation which has been found to support enhanced TAL levels in yeast.

While this assay demonstrated 2-PS oxidation susceptibility using purified *E. coli* extracts following exposure to an oxygen rich environment (lab bench), we also chose to evaluate oxidation susceptibility during TAL biosynthesis by the wildtype 2-PS in *S. cerevisiae*. Following aerobic and anaerobic cultivation in 1% YPD media for 24, 36, and 48 h, 2-PS was purified out of the culture and assayed using Ellman's reagent (DTNB). By exposing DTNB to 2-PS, free thiol groups attack DTNB and form the TNB^{2-} anion that is readily measured by absorbance at 412nm. To ensure that the oxidation state of each thiol was unaltered from the time of cultivation, purifications proceeded using native conditions and in the absence of oxygen. Once the DTNB reaction was allowed to proceed completely, free thiol levels were determined (Figure 3.9). For aerobic cultivation of *S. cerevisiae*, thiol levels increased from 24 to 36 h per mg of 2-PS enzyme assayed. A similar trend was observed for anaerobic cultivation, although overall lower free thiols were present under these conditions. In all cases, the free thiol levels are extremely low relative to an active 2-PS homodimer having all thiol side groups available for DTNB attack. Using these concentrations, we can calculate that on a relative percent basis to a fully reduced 2-PS, only 2-

30% of the available thiols are in the free, non-oxidized form. Therefore this data suggests there might be enzyme oxidation taking place during cultivation conditions in *S. cerevisiae*. One factor to the overall low levels could be the known oxygen-stress that happens when oxygen is absent during cultivation (Koc et al., 2004; Steels et al., 1994). First, the overall lifespan of a culture grown anaerobically is approximately half that of a fully aerobic culture (Koc et al., 2004), but also cells grown anaerobically are unable to tolerate heat and peroxide stresses (Steels et al., 1994). The cultivation conditions alone will result in more oxidized proteins in general as a response to the accelerated aging that takes place anaerobically. The increasing / decreasing thiol levels (Figure 3.9) in both cultivation conditions also suggests that an oxygen-insensitive mechanism may be oxidizing the protein. This may be explained by reactive oxygen species (ROS)

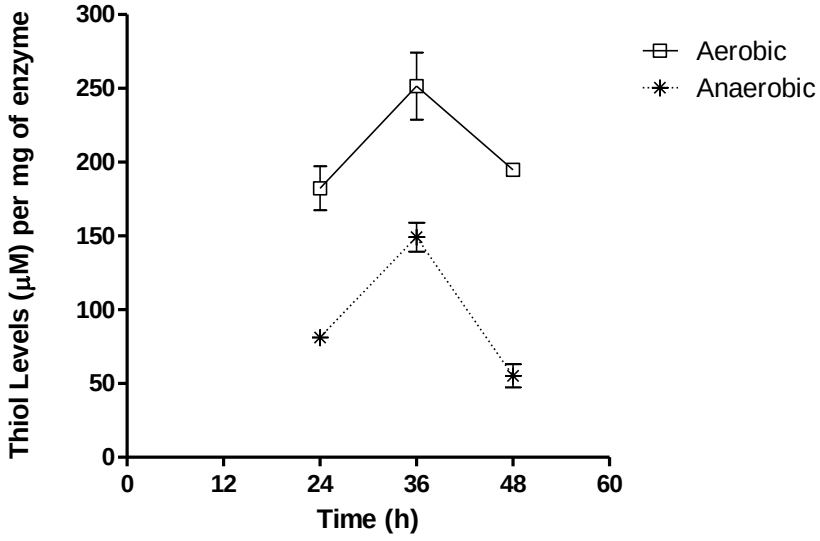


Figure 3.9 Thiol concentrations in wildtype 2-PS purified following aerobic (open square) and anaerobic (asterisk) cultivation in *S. cerevisiae* at 24, 36, and 48 h. Values shown represent mean values ± one standard deviation (n = 6 independent experiments).

which trigger oxidation in later stages of growth (Costa et al., 2007). Since one of yeast's ROS response systems utilizes the antioxidant glutathione to balance its reduced (GSH) and oxidized (GSSG) forms, it is possible that covalent attachment of glutathione to a protein thiol, like an exposed cysteine on the 2-PS surface, could diminish protein function (Grek et al., 2013). Observations therefore support a temporal change in free thiols for 2-PS cultivated in *S. cerevisiae* which could be associated with protein oxidation.

3.4.5 Variant 2-PS enzymes produce high TAL levels in presence of endogenous proteases

The experimental data presented suggests that surface cysteines should be altered for increased performance. The E2 (C35S) variant was capable of improved catalytic efficiency, and necessary to maintain 2-PS activity with fewer surface cysteines (Figure 3.7). The initial TCEP assay results (Figure 3.8) suggest that removal of reactive thiols supports better stability (increased half-life) by preventing oxidation. The temporal changes in free thiol levels when 2-PS is expressed in yeast (Figure 3.9) also suggests eliminating cysteines could be advantageous. It is well known that proteolytic activity is useful for repairing or recycling damaged protein (Costa et al., 2007). Previous studies on wildtype 2-pyrone synthase (2-PS) suggested that heterologous expression of 2-PS was indeed targeted for proteolytic activities in the vacuole (Chapter 2). Therefore 2-PS variants were similarly tested to determine how cysteine modifications might affect proteolytic targeting of 2-PS. To test the impact proteases have on 2-PS variants, E2 and E31 variants were selected for transformation into four strains with a varying level of disruption of the two vacuolar proteases Pep4 and Prb1 (Table 2.1). Following cultivation in 1% YPD after 48 h, samples were collected and TAL titer (g/L) determined (Figure 3.10). For the

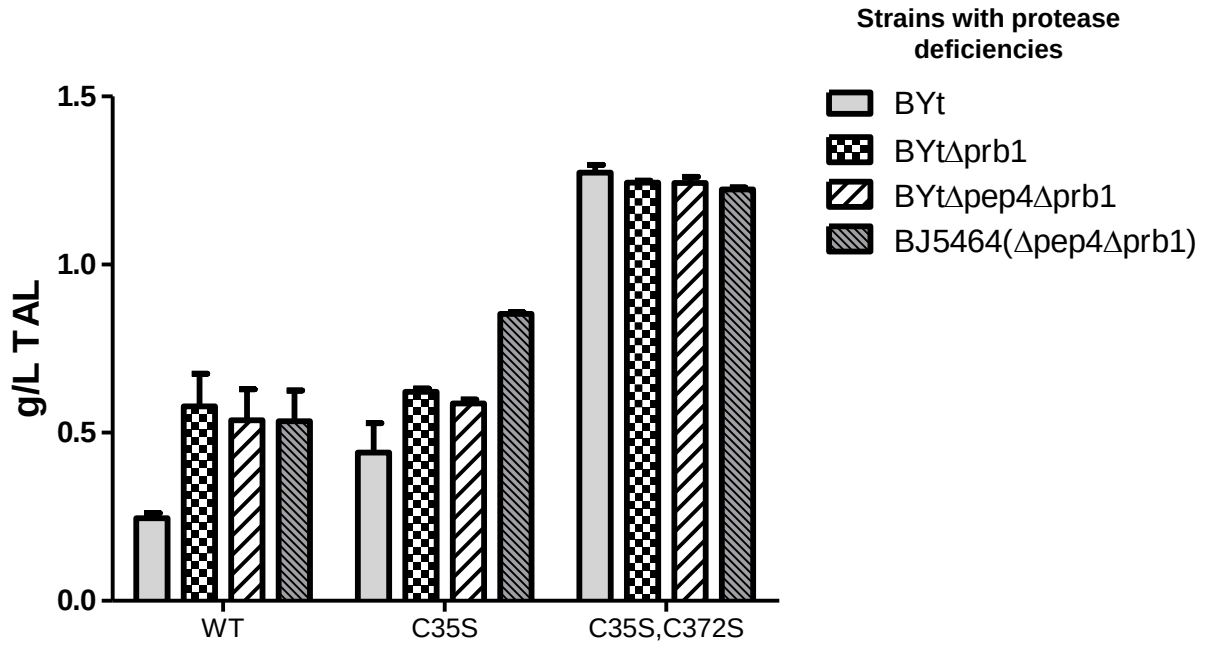


Figure 3.10 *In vivo* TAL levels (g/L) in protease-limited strains expressing WT, E2 (C35S), and E31 (C35S, C372S) 2-PS. The BY4741-based protease deletion strains BYtΔ*prb1* (checkered) and BYtΔ*pep4*Δ*prb1* (hatched, white) were compared against the base strain BYt (light grey) and the independent protease-limited BJ5464 (hatched, grey). All values reported represent mean values ± one standard deviation (n=6 independent experiments).

strains expressing WT 2-PS, the removal of vacuolar proteases improved TAL synthesis (also see Chapter 2). When the E2 variant was expressed, the removal of proteases still increased TAL levels, although the percent increase was lower. This suggests that protease activity is still limiting production when using the E2 variant. For E31 (C35S, C372S) titers significantly improved in all strains (~1.3 g/L) including over 5-fold for strain BYt with both proteases intact. Therefore, under these conditions the E31 variant was ideal for producing increased levels of TAL, consistent with screening of the 2-PS variants done previously (Figure 3.6). These cysteine residue changes provide an avenue for creating a stable and more active polyketide synthase in *S. cerevisiae*.

3.4.6 Coupling 2-PS variants to engineered strains and fed-batch fermentation

In Chapter 2, metabolic engineering of *S. cerevisiae* for the production of triacetic acid lactone used various genetic interventions in significant pathways including central carbon metabolism. The strains BYtΔ*pep4*Δ*prb1*Δ*nte1*Δ*yia6* and BJ5464Δ*pyc2*Δ*nte1* were able to achieve 1.3 g/L titers, and subsequent glucose-fed batch fermentation was demonstrated using BYtΔ*pep4*Δ*prb1*Δ*nte1*Δ*yia6* which elevated TAL titers to over 2 g/L. We thus decided to combine improved 2-PS variants with the engineered strain BJ5464Δ*pyc2*Δ*nte1*, followed by the implementation of fed-batch fermentation. The 2-PS variants were selected to have a range of increased TAL titers relative to WT, from 0.65 g/L to 1.2 g/L (using E13, E14, E31, E39, and E41) (Figures 3.3 and 3.7). By incorporating these variants in the high-producing strain BJ5464Δ*pyc2*Δ*nte1*, an increase to approximately 1.8 g/L TAL titer was achieved, corresponding to a 39% theoretical yield on glucose (Figure 3.11). Use of the variants in engineered BJ5464 (red bars, Figure 3.11) provided similar titers, in contrast to the range of levels observed in the base BJ5464 strain (green bars, Figure 3.11). The result suggests that the yeast strain is limiting rather than the enzyme. The E31 variant had the highest levels in the base strain, therefore subsequent fed-batch studies used this E31 variant for expression.

For fed-batch operation, multiple carbon feeds were supplied to the cell in order to extend productivity further. First, an ethanol feed modeled from glucose batch fermentations was designed (Figure 3.12). The maximum ethanol consumption rate in microaerobic fermentation in our system is 0.16 g/L/h. Therefore, a similar feed rate was implemented when

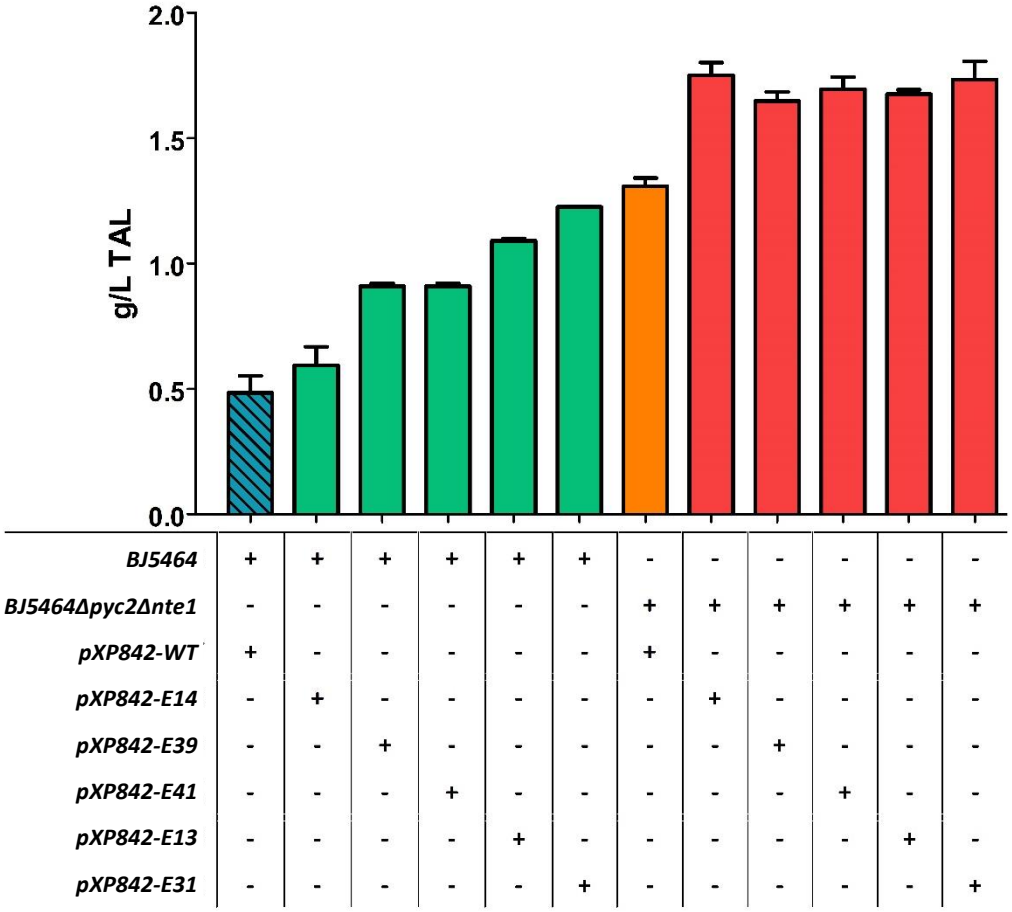


Figure 3.11 Maximum observed *in vivo* TAL levels (g/L) in batch cultivation of BJ5464 and BJ5464Δ*pyc2*Δ*nte1* using pXP842-based expression of WT and 2-PS enzymes (WT, E14, E39, E41, E13, E31). The base strain using WT 2-PS (blue, hatched) was compared to the use of engineered variants (green), the engineered strain BJ5464Δ*pyc2*Δ*nte1* (orange), and the combination of both (red). Bars represent mean values ± one standard deviation (n=6 independent experiments).

50% of endogenously formed ethanol was consumed (~24 h after inoculation, Figure 3.12) to ensure a constant supply of ethanol. Second, a glucose feed was modified to support an appropriate carbon:nitrogen ratio (C/N) (see section 3.3.6). The optimal nitrogen source was selected as a mixture of nitrogen-containing amino acids based on literature comparing to

ammonium sulfate or glutamic acid (Albers et al., 1996). To initiate the fermentations, an overnight shake-flask culture was inoculated in selective SDC(A) from 5mL tube cultures. This seed flask culture was used to inoculate the fermentor to 1.5L starting volume and a cell density (OD_{600}) of 0.3. Depending on the feed strategy used, the volume was 1.7-2.0 L at the end of the experiment. Following fed-batch operation for up to 144 h, maximum theoretical yield, productivity, and TAL titer were determined (Figure 3.13). Incorporation of the E31 2-PS variant led to a 20% improvement (2.6 g/L) when using glucose fed-batch operation relative to the WT 2-PS. The titer was increased 1.6-fold to 4.2 g/L for the ethanol feed. An additional fed-batch fermentation using a supplemented glucose / nitrogen feed was also used to ensure the carbon

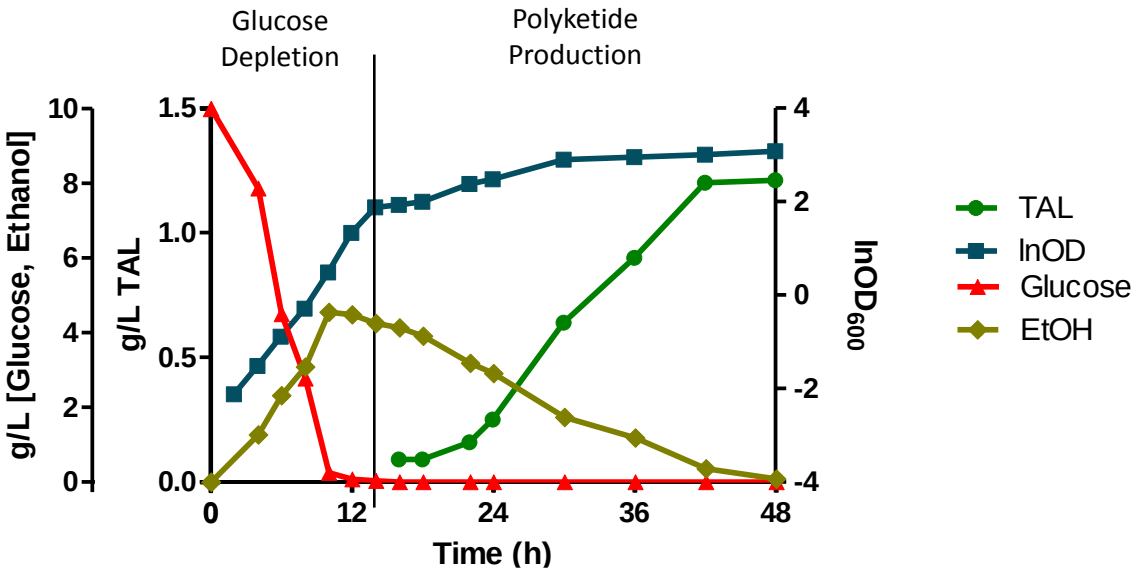
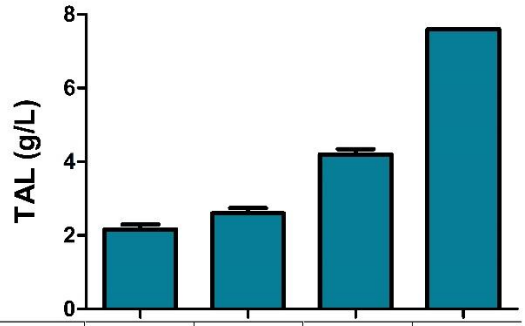


Figure 3.12 Profiling of glucose batch cultivation (1% YPD) in BJ5464Δ*pyc2*Δ*nte1* transformed with pXP842-E31. Polyketide formation (green, circles), lnOD (blue, squares), glucose levels (red, triangles), and endogenous ethanol (gold, diamonds) during stages of glucose depletion and polyketide production. All values reported represent mean values (n=3 independent experiments).



pXP842-WT	+	-	-	-
pXP842-E31	-	+	+	+
Glucose	+	+	-	-
Ethanol	-	-	+	-
Glucose / Nitrogen	-	-	-	+
% of Theoretical Yield	26	29	30	44
Productivity	0.023	0.038	0.044	0.08

Figure 3.13 Maximum observed *in vivo* TAL levels (g/L) for fed-batch fermentations of BJ5464Δ*pep4*Δ*prb1*Δ*pyc2*Δ*nte1* using pXP842-based expression of WT and E31 2-PS enzymes. Carbon feeds are denoted as glucose, ethanol, and glucose / nitrogen (for glucose feed supplemented with nitrogen sources). Bars not including the final value (single run) represent mean values ± one standard deviation (n=3 independent experiments). Units: Percent of theoretical yield [g TAL observed/g TAL theoretical *100%]; productivity is maximum observed [g/L/h].

feed supplied a stable control of the C/N ratio to 35:1. Using previous studies as a reference, this C/N ratio was implemented using an amino acid-based nitrogen source (see Materials and Methods) found to support high ethanol yield on glucose (Albers et al., 1996; Larsson et al., 1993; Manikandan et al., 2010). An additional 3-fold improvement in titer was observed relative to the strain using a glucose feed. This change was primarily due to the extended time during which glucose was consumed, and subsequent boost in biomass formation (Figure 3.14). This strain (BJ5464Δ*pep4*Δ*prb1*Δ*pyc2*Δ*nte1* + pXP842-E31) achieved 30 g/L DCW biomass and 7.6 g/L TAL

titer, a nearly 4-fold increase relative to wildtype 2-PS expressed in the same strain by glucose fed- batch fermentation. From these results, further optimization could be designed to suppress biomass formation in favor of product formation.

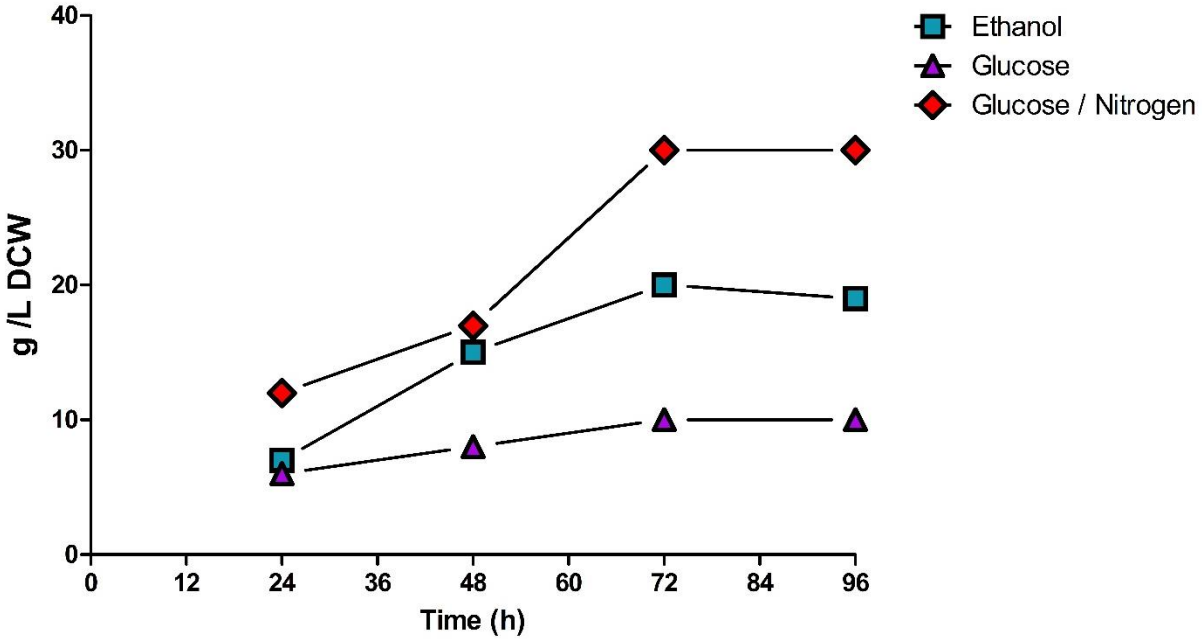


Figure 3.14 Biomass profiles (g/L DCW) for fed-batch fermentation using 3 feed strategies: glucose (purple triangles), ethanol (light blue squares), and glucose / nitrogen (red diamonds). Points not including the glucose / nitrogen feed (single run) represent mean values ± one standard deviation (n=3 independent experiments).

3.5 Conclusions

Various efforts are underway to improve enzyme efficiency and employ them in biological systems. In this study, a combined *in vitro* / *in vivo* approach was used to design an enzyme from both an activity and stability standpoint. Through *in vitro* characterization, novel 2-PS variants were constructed using steric modulation to reduce cavity size in the active site and prevent improper chain extensions or promiscuity in the loading moieties. For a significant set of enzyme variants, both *in vitro* and *in vivo* results correlated well, but various enzymes showed a difference between the catalytic efficiency changes and subsequent TAL titers produced in *S. cerevisiae*. Further investigation led to the identification of surface-exposed cysteine residues that were highly reactive in the microbial host. Specifically, C35S improved k_{cat} / K_M almost 2-fold relative to the wildtype, but also appeared to be critical for stability of 2-PS variants *in vivo*. The *in vitro* oxidation assays suggest that oxidation machinery is responsible for the perceived importance of cysteine elimination, and removal of cysteines at specific locations with high exposure to the protein surface had the largest increases in TAL formation by *S. cerevisiae*. Combining the high-performing variants with the engineered BJ5464Δ*pyc2*Δ*nte1* strain provided an opportunity to employ a robust biocatalyst enzyme in a host designed for increased TAL production. Fed-batch fermentation employing this engineered strain (BJ5464Δ*pyc2*Δ*nte1*) with variant E31 (C35S, C372S) using a glucose, ethanol, or glucose/nitrogen amino acids feed, led to TAL titers over 2.5 g/L, 4 g/L, and 7.6 g/L at 29%, 30%, and 44% of the theoretical yield, respectively. These are the highest demonstrated TAL levels reported to date. The 2-PS modifications evaluated in this work should provide successful strategies for other proteins used as biocatalyst systems in *S. cerevisiae*.

3.6 Acknowledgements

This research was supported by the National Science Foundation (Grant No. EEC-0813570) through the Engineering Research Center CBiRC (Center for Biorenewable Chemicals). Thanks to Chris Vickery, Kate Woods, Marianne Bowman, and Joe Noel at the Salk Institute for constructing the 2-PS variants and for performing *in vitro* kinetic analysis and TCEP assays for 2-PS and variants produced using *E. coli*.

3.7 References

Albers, E., Larsson, C., Lidén, G., Niklasson, C., Gustafsson, L., 1996. Influence of the nitrogen source on Saccharomyces cerevisiae anaerobic growth and product formation. Applied and Environmental Microbiology 62, 3187–3195.

Bornscheuer, U.T., Huisman, G.W., Kazlauskas, R.J., Lutz, S., Moore, J.C., Robins, K., 2012. Engineering the third wave of biocatalysis. Nature 485, 185–194. doi:10.1038/nature11117

Cardenas, J., Da Silva, N.A., 2014. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. Metabolic Engineering 25, 194–203. doi:10.1016/j.ymben.2014.07.008

Chia, M., Schwartz, T.J., Shanks, B.H., Dumesic, J.A., 2012. Triacetic acid lactone as a potential biorenewable platform chemical. Green Chem. 14, 1850–1853. doi:10.1039/C2GC35343A

Costa, V., Quintanilha, A., Moradas-Ferreira, P., 2007. Protein oxidation, repair mechanisms and proteolysis in Saccharomyces cerevisiae. IUBMB Life 59, 293–298. doi:10.1080/15216540701225958

Esvelt, K.M., Carlson, J.C., Liu, D.R., 2011. A system for the continuous directed evolution of biomolecules. Nature 472, 499–503. doi:10.1038/nature09929

Gietz, D., Jean, A.S., Woods, R.A., Schiestl, R.H., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucl. Acids Res. 20, 1425–1425. doi:10.1093/nar/20.6.1425

Grek, C.L., Zhang, J., Manevich, Y., Townsend, D.M., Tew, K.D., 2013. Causes and Consequences of Cysteine S-Glutathionylation. Journal of Biological Chemistry 288, 26497–26504. doi:10.1074/jbc.R113.461368

Hertweck, C., 2009. The Biosynthetic Logic of Polyketide Diversity. Angewandte Chemie International Edition 48, 4688–4716. doi:10.1002/anie.200806121

Hill, J., Donald, K.A.I.G., Griffiths, D.E., 1991. DMSO-enhanced whole cell yeast transformation. Nucl. Acids Res. 19, 5791–5791. doi:10.1093/nar/19.20.5791

Jez, J.M., Austin, M.B., Ferrer, J.-L., Bowman, M.E., Schröder, J., Noel, J.P., 2000a. Structural control of polyketide formation in plant-specific polyketide synthases. Chemistry & Biology 7, 919–930. doi:10.1016/S1074-5521(00)00041-7

Jez, J.M., Bowman, M.E., Noel, J.P., 2002. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. Proc Natl Acad Sci U S A 99, 5319–5324. doi:10.1073/pnas.082590499

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., Noel, J.P., 2000b. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. Biochemistry 39, 890–902.

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

Koc, A., Gasch, A.P., Rutherford, J.C., Kim, H.-Y., Gladyshev, V.N., 2004. Methionine sulfoxide reductase regulation of yeast lifespan reveals reactive oxygen species-dependent and -independent components of aging. Proc Natl Acad Sci U S A 101, 7999–8004. doi:10.1073/pnas.0307929101

Larsson, C., von Stockar, U., Marison, I., Gustafsson, L., 1993. Growth and metabolism of Saccharomyces cerevisiae in chemostat cultures under carbon-, nitrogen-, or carbon-and nitrogen-limiting conditions. Journal of bacteriology 175, 4809–4816.

Manikandan, K., Viruthagiri, T., others, 2010. Optimization of C/N ratio of the medium and fermentation conditions of ethanol production from tapioca starch using co–culture of Aspergillus niger and Saccharomyces cerevisiae. Int J Chem Tech Res 2, 947–955.

Ma, S.K., Gruber, J., Davis, C., Newman, L., Gray, D., Wang, A., Grate, J., Huisman, G.W., Sheldon, R.A., 2010. A green-by-design biocatalytic process for atorvastatin intermediate. Green Chem. 12, 81–86. doi:10.1039/B919115C

McKenna, R., Nielsen, D.R., 2011. Styrene biosynthesis from glucose by engineered E. coli. Metabolic Engineering 13, 544–554. doi:10.1016/j.ymben.2011.06.005

Nikolau, B.J., Perera, M.A.D.N., Brachova, L., Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. The Plant Journal 54, 536–545. doi:10.1111/j.1365-313X.2008.03484.x

Pfeifer, B.A., Khosla, C., 2001. Biosynthesis of Polyketides in Heterologous Hosts. Microbiol. Mol. Biol. Rev. 65, 106–118. doi:10.1128/MMBR.65.1.106-118.2001

Requejo, R., Hurd, T.R., Costa, N.J., Murphy, M.P., 2010. Cysteine residues exposed on protein surfaces are the dominant intramitochondrial thiol and may protect against oxidative damage. FEBS J 277, 1465–1480. doi:10.1111/j.1742-4658.2010.07576.x

Sambrook, J.J., Russell, D.D.W., 2001. Molecular cloning: a laboratory manual. Vol. 2. CSHL Press.

Sarrouh, B., 2012. Up-To-Date Insight on Industrial Enzymes Applications and Global Market. Journal of Bioprocessing & Biotechniques s1. doi:10.4172/2155-9821.S4-002

Shen, B., 2003. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. Current Opinion in Chemical Biology 7, 285–295. doi:10.1016/S1367-5931(03)00020-6

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. doi:10.1002/yea.2930

Steels, E.L., Learmonth, R.P., Watson, K., 1994. Stress tolerance and membrane lipid unsaturation in Saccharomyces cerevisiae grown aerobically or anaerobically. Microbiology 140, 569–576. doi:10.1099/00221287-140-3-569

Steiner, K., Schwab, H., 2012. Recent advances in rational approaches for enzyme engineering. Comput Struct Biotechnol J 2. doi:10.5936/csbj.201209010

Tang, S.-Y., Qian, S., Akinterinwa, O., Frei, C.S., Gredell, J.A., Cirino, P.C., 2013. Screening for Enhanced Triacetic Acid Lactone Production by Recombinant Escherichia coli Expressing a Designed Triacetic Acid Lactone Reporter. J. Am. Chem. Soc. 135, 10099–10103. doi:10.1021/ja402654z

Xie, D., Shao, Z., Achkar, J., Zha, W., Frost, J.W., Zhao, H., 2006. Microbial synthesis of triacetic acid lactone. Biotechnology and Bioengineering 93, 727–736. doi:10.1002/bit.20759

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H.B., Andrae, S., Yang, T.H., Lee, S.Y., Burk, M.J., Van Dien, S., 2011. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat Chem Biol 7, 445–452. doi:10.1038/nchembio.580

Chapter 4:

Engineering cofactor and transport mechanisms in *Saccharomyces cerevisiae* for enhanced acetyl-CoA and polyketide biosynthesis

4.1 Abstract

Industrial large-scale synthesis of polyketides at high yields is important for producing pharmaceuticals and biorenewable precursors. In this work, we engineered cofactor and transport pathways to increase acetyl-CoA in *Saccharomyces cerevisiae*, an important polyketide building block. The highly regulated yeast pyruvate dehydrogenase bypass pathway was supplemented by overexpressing a modified *Escherichia coli* pyruvate dehydrogenase complex (PDHm) that accepts $NADP^+$ for acetyl-CoA production. After 24 h of cultivation, a 3-fold increase in NADPH/$NADP^+$ ratio was observed relative to the base strain, and a 2-fold increase relative to introduction of the native *E. coli* PDH. Both *E. coli* pathways increased acetyl-CoA levels 2-fold relative to the yeast base strain. Combining the modified PDH with a *ZWF1* deletion to block the major yeast NADPH synthesis pathway resulted in a 12-fold NADPH boost and 2-fold increase in acetyl-CoA levels. By 48 h, only this coupled approach showed increased acetyl-CoA levels, 3-fold that of the base strain. The impact on polyketide synthesis was evaluated in a *S. cerevisiae* strain expressing the *Gerbera hybrida* 2-pyrone synthase (2-PS) for the production of triacetic acid lactone (TAL). Titers of TAL relative to the base strain improved only 30% when overexpressing the unmodified PDH, but 3-fold for expression of the modified PDH. In combination with the *ZWF1* deletion, the modified PDH showed a further increase in TAL titer of approximately 50%. Carbon was further routed toward TAL production by minimizing mitochondrial transport of pyruvate and acetyl-CoA; deletions in *POR2*, *MPC2*, *PDA1*, and *YAT2* genes each enhanced titer up to 3-fold (0.8 g/L) over the base strain, and in combination to 1.4 g/L. Combining the two approaches (NADPH-generating acetyl-CoA pathway with reduced metabolite flux into the mitochondria) resulted in a final TAL titer of 1.6 g/L, a 6-fold increase over the non-engineered

yeast strain, and 36% theoretical yield (0.16 g/g glucose), the highest reported to date. The creation of these biological driving forces for product formation presents additional avenues for improving high-yield production of acetyl-CoA derived compounds.

4.2 Introduction

Polyketides are a diverse group of natural products with valuable properties useful for the manufacture of biologics and pharmaceuticals as well as replacement of fossil-fuel derived materials (Jeandet et al., 2013; Li and Vederas, 2009; Nikolau et al., 2008). The synthesis of these molecules require polyketide synthase (PKS) systems which are found in all forms of life (Nikolau et al., 2008), and catalyze a complex set of reactions including decarboxylation, condensation, cyclization, and aromatization steps (Hertweck, 2009; Jez et al., 2002). Limitations in the native host include biosynthesis of small quantities, and a lack of genetic tools for strain manipulations (Jeandet et al., 2012; Lussier et al., 2012; Santos et al., 2011); therefore, engineering of heterologous microbial cell factories is of high interest (Krivoruchko and Nielsen, 2015). The yeast *Saccharomyces cerevisiae* is advantageous as it is capable of synthesizing fungal and plant polyketides that are toxic to or difficult to produce in bacteria (Cardenas and Da Silva, 2014; Ma et al., 2009), and also allow the synthesis of complex polyketides relying on cytochrome P450 activities (Howat et al., 2014). Various types of PKSs are active in yeast, and the synthesis of fungal (6-methylsalicylic acid, dihydromonacolin L) and plant (triacetic acid lactone, TAL) as functional replacements in biologics (Ma et al., 2009; Wattanachaisaereekul et al., 2007; Xu et al., 2013) and biorenewable chemical precursors (Cardenas and Da Silva, 2014; Schwartz et al., 2014) demonstrate the promise of yeast for polyketide biosynthesis.

Many polyketides are built from the simple building blocks acetyl-CoA and malonyl-CoA. These starter metabolites are in high demand in yeast, to for example make fatty acids (Jong et al., 2014) or energy via ATP (Pronk et al., 1996). In order to synthesize polyketides at high levels, it is therefore necessary to improve availability of these starter metabolites and relevant cofactor

pools. One of the most direct methods for improving synthesis of acetyl-CoA and malonyl-CoA is by overexpressing the enzymes responsible for their formation and eliminating competing pathways. Using various overexpression and pathway disruption strategies, several studies have reported interventions aimed at increasing acetyl-CoA for the synthesis of sesquiterpenes (Chen et al., 2010, 2013; Shiba et al., 2007), butanol (Krivoruchko et al., 2013; Lian et al., 2014), polyhydroxybutyrate (PHB) (Kocharin et al., 2013, 2012), as well as fatty acids (Li et al., 2014; Tang et al., 2013). For malonyl-CoA formation, acetyl-CoA carboxylase (Acc1) is important and can limit malonyl-CoA synthesis. This can be circumvented by overexpression (Runguphan and Keasling, 2014; Wattanachaisaereekul et al., 2008) or prevention of deactivation of the Acc1 enzyme following glucose consumption (Choi and Da Silva, 2014; Hofbauer et al., 2014; Shi et al., 2014). In this recent work in our lab, the overexpression of a deactivation-resistant Acc1 yielded a 9-fold increase in activity *in vivo* leading to a 3-fold increase in both fatty acid and polyketide levels.

Apart from direct increases in starter metabolite pools, cofactors are also useful in driving carbon flow. Since redox is a significant factor in specifying flux through key metabolic nodes (Ying, 2008), levels of NAD^+ and NADPH can be improved to increase the cell's reducing power and support product formation. It has been shown in yeast that increasing cytosolic NAD^+ gave 2-fold higher polyketide levels (Cardenas and Da Silva, 2014) and overexpressing *ALD6* improved NADPH availability and increased synthesis of fatty acid ethyl esters (Jong et al., 2014). In another study, overexpression of *gapN*-encoded $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase increased NADPH pools which improved polyhydroxybutyrate (PHB) (Kocharin et al., 2013). A recent report increased NADH oxidation using heterologous expression of two

known oxidases (NOX, AOX1) to increase 7-dehydrocholesterol (Su et al., 2015). These studies demonstrate that altering cofactor levels can be beneficial toward product formation. In addition, by first eliminating the cell's natural cofactor regeneration pathways, introduction of a novel pathway to regenerate levels through precursor synthesis can be a much more effective driving force. This strategy was employed recently in *E. coli* which saw improved butanol synthesis when coupled to NADH oxidation (Shen et al., 2011). It is therefore possible to envision a redesigned cell network where cofactors are reduced and oxidized in a manner conducive towards a pathway of interest.

To engineer *S. cerevisiae* for increases in the starter metabolites acetyl-CoA and malonyl-CoA in the cytosol, it is also important to focus on the availability of pyruvate in the cell. Pyruvate is a major branch point for carbon following glycolysis (Pronk et al., 1996), one of which is the synthesis of acetyl-CoA via the PDH bypass (see Figure 4.2). Pyruvate is first decarboxylated, then dehydrogenase activity generates acetate that acetyl-CoA synthetase (Acs1/2) uses to produce acetyl-CoA. Recent computational optimizations rerouted yeast flux from the PDH bypass through an *E. coli* PDH to increase TAL; this also conserves ATP by eliminating use of the ATP-dependent Acs1/2 (Chowdhury et al., 2014). In this study, the native PDH bypass was downregulated, and only use of an engineered $NADP^+$-dependent *E. coli* PDH predicted an improved yield of TAL. The redox imbalance caused by the NAD^+-dependent PDH from *E. coli* resulted in no yield increase. Pyruvate can also be utilized by other pathways via its carboxylation or transport to the mitochondria. The yeast mitochondrial pyruvate carrier (MPC) has been shown to transport pyruvate into the mitochondria (see Figure 4.6) (Bricker et al., 2012; Herzig et al., 2012). Similarly, acetyl-CoA can be transported using the carnitine shuttle system (Figure

4.6) where carnitine acetyltransferases (CATs) are responsible for first priming acetyl-CoA for transport (via Yat2) before facilitating movement into the mitochondria (Yat1) where acetyl-CoA is formed (Cat2) (Franken et al., 2008). Pyruvate and acetyl-CoA are therefore invaluable to the cell, and rely on these systems for transport into the mitochondria, causing a significant drain on their cytosolic availability. Impeding this transport may result in larger cytosolic pools and thus increase polyketide synthesis in yeast.

The overall aim of our work was to engineer novel cofactor and transport pathways in *S. cerevisiae* to directly increase acetyl-CoA levels and thus polyketide synthesis. To increase acetyl-CoA, the $NADP^+$ variant of the *E. coli* pyruvate dehydrogenase complex (PDHm) was expressed and NADPH/$NADP^+$ ratios and acetyl-CoA were compared to those for the base yeast strain and one expressing the native *E. coli* PDH. This is the first study expressing this functional variant of PDH in yeast. Construction of a Δ*ZWF1* yeast strain (with the major native NADPH-producing pathway blocked) was coupled to PDHm overexpression to further increase carbon flux to acetyl-CoA, and the effects on the synthesis of the polyketide triacetic acid lactone (TAL) were determined. Additional pathway design incorporated key gene disruptions that reduced mitochondrial transport of pyruvate and acetyl-CoA to increase their availability. This novel strategy to increase cytosolic pyruvate and acetyl-CoA levels in *S. cerevisiae* was assessed by evaluating the effects on TAL synthesis. The hindered transport and NADPH/acetyl-CoA upregulation strategies were then combined to further increase TAL production in the host. We demonstrate how increasing metabolite levels via these strategies can help alleviate the current challenges in high-yield production of polyketides.

4.3. Materials and Methods

4.3.1 Strains and plasmids

Escherichia coli strain XL1-Blue (Stratagene, Santa Clara, CA) was used for plasmid amplification, and *E. coli* NEB 10-beta cells (New England Biolabs, Ipswich, MA) were used for amplification of Gibson-assembled vectors. *Saccharomyces cerevisiae* strain BY4741 and various single gene deletion strains were obtained from Open Biosystems (Huntsville, AL). Further deletion strains were constructed using loxP-flanked *MET15*, *LEU2-d8, TRP1,* and *HIS3* selection markers amplified from our pXP vector library (Fang et al., 2011). The amplification primers contained approximately 50bp homology to locations flanking the target gene location. Yeast cells were transformed as described previously (Gietz et al., 1992; Hill et al., 1991) and gene disruptions were verified following amplification of genomic DNA using Taq polymerase (New England Biolabs, Ipswich, MA). The strains and vectors, and all primer sequences can be found in Table 4.1 and 4.2, respectively.

Plasmid pXP842-2PS harbors the *G. hybrida g2ps1* gene (Austin and Noel, 2003) encoding 2-pyrone synthase (Cardenas and Da Silva, 2014). To express the *E. coli* PDH complex, a 2µ-based trigenic vector was constructed. DNA fragments were created to flank the three PDH genes with promoters and terminators as listed: *ADH2p-lpd1-ADH2t*, *ADH1Mp-aceF-ADH2t*, and *HXT7p-aceE-CYCt*. The three genes were isolated by PCR from *E. coli* chromosomal DNA preparations (Dale and Greenaway, 1985). The *lpd1* and *aceF* fragments were placed in opposite orientation to prevent recombination of the terminator sequences. *ADH1Mp* refers to the medium truncated *ADH1* promoter which alters glucose-induced expression to approximately constitutive

Table 4.1. List of plasmids and strains

Plasmids	**Description**	**Source**
pXP842	2µ vector, *ADH2* promoter, *CYC1* terminator, *URA3* selectable marker	Shen et al. (2012)
pXP843	2µ vector, *ADH2* promoter, *CYC1* terminator, *TRP1* selectable marker	Leber and Da Silva (2014)
pXP842-2PS	pXP218 harboring the *g2ps1* insertion encoding *G. hybrida* 2-pyrone synthase	Cardenas and Da Silva (2014)
pJCT-PDH	2µ vector, *ADH2p-lpd1-ADH2t*, *HXT7p-aceE-CYC1t*, *ADH1Mp-aceF-ADH2t*, *TRP1* selectable marker	This study
pJCT-PDHm	2µ vector, *ADH2p-lpd1*m-*ADH2t*, *HXT7p-aceE-CYC1t*, *ADH1Mp-aceF-ADH2t*, *TRP1* selectable marker	This study
Strains	**Description**	**Source**
BY4741	*MAT*a *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Open Biosystems
BY4741Δ*trp1* (BYt)	BY4741 *trp1::KanMX*	Open Biosystems
BYΔ*por1*	BY4741 *por1::KanMX*	Open Biosystems
BYΔ*por2*	BY4741 *por2::KanMX*	Open Biosystems
BYΔ*mpc1*	BY4741 *mpc1::KanMX*	Open Biosystems
BYΔ*mpc2*	BY4741 *mpc2::KanMX*	Open Biosystems
BYΔ*mpc3*	BY4741 *mpc3::KanMX*	Open Biosystems
BYΔ*pda1*	BY4741 *pda1::KanMX*	Open Biosystems
BYΔ*yat1*	BY4741 *yat1::KanMX*	Open Biosystems
BYΔ*yat2*	BY4741 *yat2::KanMX*	Open Biosystems
BYtΔ*zwf1*	BYt *zwf1::MET15*	This study
BYtΔ*por2*	BYt *por2::TRP1*	This study
BYtΔ*por2*Δ*mpc2*	BYt *por2::TRP1 mpc2::MET15*	This study
BYtΔ*por2*Δ*mpc2*Δ*pda1*	BYt *por2::TRP1 mpc2::MET15 pda1::HIS3*	This study
BYtΔ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*	BYt *por2::TRP1 mpc2::MET15 pda1::HIS3 yat2::LEU2*	This study
BYtAΔ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*	BYt *por2::loxP mpc2::MET15 pda1::HIS3 yat2::LEU2*	This study
BYtΔ*zwf1*Δ*por2*	BYt *zwf1::MET15 por2::LEU2*	This study
BYtΔ*zwf1*Δ*por2*Δ*mpc2*	BYt *zwf1::MET15 por2::LEU2 mpc2::HIS3*	This study
BYtΔ*zwf1*Δ*por2*Δ*mpc2*Δ*pda1*	BYt *zwf1::MET15 por2::LEU2 mpc2::HIS3 pda1::TRP1*	This study
BYtΔ*zwf1*Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*	BYt *zwf1::MET15 por2::LEU2 mpc2::HIS3 pda1::TRP1 yat2::URA3*	This study
BYtBΔ*zwf1*Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*	BYt *zwf1::MET15 por2::LEU2 mpc2::HIS3 pda1::loxP yat2:loxP*	This study

Table 4.2. List of primers used

Primer Name	Primer Sequence
ZWF1_KO_Fwd	GAGGAGTTGGTGGGGGGAAGATGCCATTATAGAGGAGAAAAGAACAAAGGGTCGACTCTAGAGGATCCCCGGG
ZWF1_KO_Rev	CTGGAAGCACCACGTAATAGTGGAAAAGAACTGGAAAAACCGCGTTGAATTCGAGCTCGGTACCCGGG
POR2_KO_Fwd	GGTCGTAGGTCACGATAAAAGTTGCGATGTCCCCTCCCACCTAGTTTCCGGTCGACTCTAGAGGATCCCCGGG
POR2_KO_Rev	CAACCTAATTAGGAAGTTGCCTGTTCCAGAAGGGGAGTTGCCAAGGCCGCGTTGAATTCGAGCTCGGTACCCGGG
MPC2_KO_Fwd	CATCGTTGACATCGCTGACTGCAATAGGAAACTGAAATAGACGGCAAACCGGTCGACTCTAGAGGATCCCCGGG
MPC2_KO_Rev	CTTTGACAAAGCAAAGTTTCGAGTTGTAAATTCGGCGTTTATCTGCCCGGTTGAATTCGAGCTCGGTACCCGGG
PDA1_KO_Fwd	CGACAATAATCCCCTCTGGTATAGCGAGAAGCAACTTTAGCTTCTTAACGGGTCGACTCTAGAGGATCCCCGGG
PDA1_KO_Rev	CTTGCTATAAAATGGCGCTTCCAGGGAAGAATATCATGCGATCACAGCACGTTGAATTCGAGCTCGGTACCCGGG
YAT2_KO_Fwd	CCCAAGTTACCATTACCTAAACTCTGTGACACTTTGCAGCGTCTGAAGGGGTCGACTCTAGAGGATCCCCGGG
YAT2_KO_Rev	GCTTAGCTCCTGGTACAGGTCGAGATGAAATCTCTGTCTTTTGTACTGATGGTTGAATTCGAGCTCGGTACCCGGG
F1_lpd1m_Fwd	GTCCCACTAGTAAAAAAATGAGTACTGAAATCAAAACTCAGGTCGTGG
F1_lpd1m_Rev	AGCCATTTCCAGAGCGATGATACCGCCACCC
F2_lpd1m_Fwd	GCTCTGGAAATGGCTACCGTTTACCACGCGC
F2_lpd1m_Rev	ACGGATAACCTGATGTTTACGAACAACCACGTCAAT
F3_lpd1m_Fwd	GTTCGTAAACATCAGGTTATCCGTGCAGCTGACA
F3_lpd1m_Rev	TGGTAATCCTCGAGTTACTTCTTCTTCGC
aceE_Fwd	ATGTCAGAACGTTTCCCAAATGACG
aceE_Rev	TTACGCCAGACGCGGGTTAACTTTATCTGC
aceF_Fwd	ATG GCT ATC GAA ATC AAA GTA CCG GAC ATC GGG GCT G
aceF_Rev	TTACATCACCAGACGGCGAATGTCAGACAGCG
OE1_lpd1m_Fwd	GTCCATATTGTACACCCGGAAACAACAAAAAAACGTAGGGGCAAACAAACGGAAAAATCG
lpd1_ADH2p_Rev	CCTGAGTTTTGATTTCAGTACTCATTTTTTTATTACGATATAG
ADH2p_lpd1m_Fwd	CTATTAACTATATCGTAATAAAAAAATGAGTACTGAAATCAAAAC

Table 4.2. List of primers used (contd.)

Primer Name	Primer Sequence
ADH2t_lpd1m_Rev	AAATCGTGAAGGCATTTACTTCTTCTTCGCTTTC
lpd1_ADH2t_Fwd	CTGCCGAACCCGAAAGCGAAGAAGAAGTAAATGCCTTCACG
Tri_hx_2	TTCCACGGGGGGAGCAAACGCGCTGG
OE1_aceF_Rev	CGATTTTTCCGTTTGTTTGCCCCTACGTTTTTTTGTTGTTTCCGGGTGTACAATATGGAC
aceF_ADH1Mp_Rev	CCGGTACTTTGATTTCGATAGCCATTTTTTTAGTTGATTGTATGCTTGG
ADH1Mp_aceF_Fwd	GCATACAATCAACTAAAAAAATGGCTATCGAAATCAAAG
ADH2t_aceF_Rev	AAATCGTGAAGGCATTTACATCACCAGACGGCG
aceF_ADH2t_Fwd	CGCCGTCTGGTGATGTAAATGCCTTCACG
Tri_bb_1	AGTGCCACCTGACGTCGGGGAGCAAACGCGCTGG
Tri_oe1_3	TGCTCCCCCCGTGGAAATGAGGGGTATG
aceE_HXT7p_Rev	GGGAAACGTTCTGACATTTTTTTTTTTTGATTAAAATTAAAAAAAC
HXT7p_aceE_Fwd	AATTTTAATCAAAAAAAAAAAATGTCAGAACGTTTCCC
CYC1t_aceE_Rev	TAACTAATTACATGATTACGCCAGACGCGGGTTAAC
aceE_CYCt_Fwd	GTTAACCCGCGTCTGGCGTAATCATGTAATTAGTTATG
Tri_bb_4	CGAAGTTATCCCGGGGATCCGGCCGCAAATTAAAGCCTTC
Tri_bb_1	AGTGCCACCTGACGTCGGGGAGCAAACGCGCTGG

expression (Ruohonen et al., 1995). Template vector pXP843 (Leber and Da Silva, 2013) was linearized by BamHI / AatII (New England Biolabs, Ipswich, MA). Gibson assembly (Gibson et al., 2009) using the Phusion Hot Start II DNA Polymerase (New England Biolabs, Ipswich, MA) was performed using this linearized backbone and the three gene fragments listed above to generate pJCT-PDH. For construction of the alternate trigenic vector pJCT-PDHm for expression of the variant *E. coli* PDHm complex, the wildtype sequence for the *lpd1* gene was replaced with a variant incorporating 15 base changes using insertion mutagenesis (Lee et al., 2010). These modifications coincide with the amino acid changes previously identified for $NADP^+$ binding (Bocanegra et al., 1993) as well as optimal codon usage preferences in *S. cerevisiae*. Further details for construction of pJCT-PDH and pJCT-PDHm can be found in Figure 4.1. Unless otherwise stated, all PCR-based amplifications were performed using KOD Hot Start DNA Polymerase (EMD Chemicals, San Diego, CA). Oligonucleotides used in this study were provided by IDT DNA (San Diego, CA), the GeneJet™ Plasmid Miniprep Kit (Thermo Scientific, Waltham, MA) was used for plasmid preparation, and DNA sequences were confirmed for all PCR-amplified inserts (Eton Biosciences, San Diego, CA).

4.3.2 Media and cultivation

Luria-Bertani (LB) media was used for cultivation of XL1-Blue cells, with 150 mg/L ampicillin added for selection of plasmid-containing strains (Sambrook and Russell, 2001). Yeast selective SDC medium (2% dextrose, 0.5% casamino acids, 0.67% yeast nitrogen base, 0.5% ammonium sulfate supplemented with 100 mg/L adenine sulfate) and SD minimal medium (2%

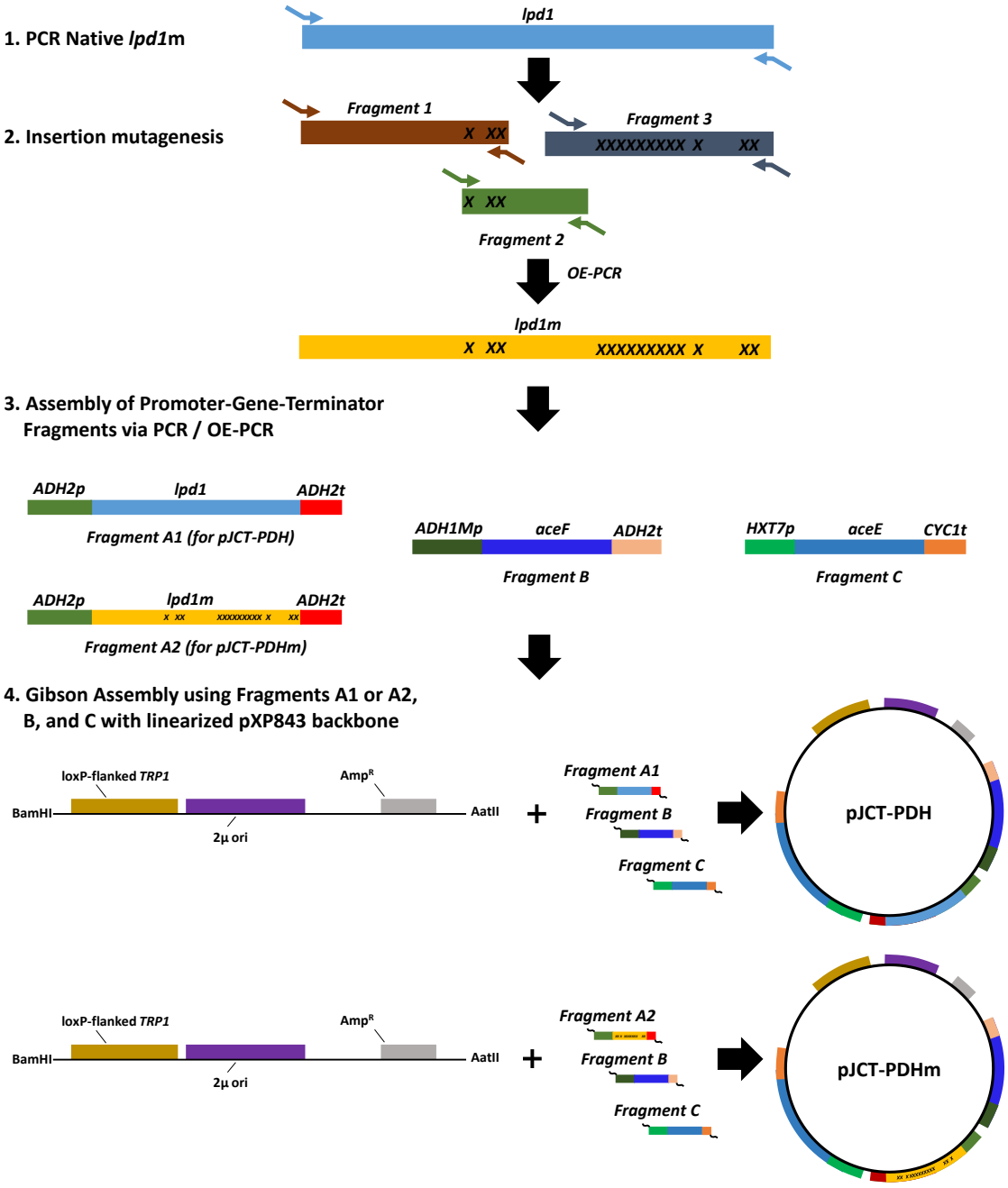


Figure 4.1. Vector construction of pJCT-PDH and pJCT-PDHm. The *lpd1* gene was first PCR amplified from the *E. coli* genome (Step 1). This was used as template (Step 2) to generate Fragments 1-3 harboring all 15 base changes (denoted by an ***X***), yielding the full *lpd1*m variant following overlap extension PCR (OE-PCR). Both PCR and OE-PCR were performed to amplify promoter-gene-terminator cassettes (Step 3) including PCR amplification of additional *aceF* and *aceE* genes, and homology introduced to fragment ends for full assembly by Gibson reaction using pXP843 as a backbone vector (Step 4). Use of Fragment A1 or A2 resulted in pJCT-PDH or pJCT-PDHm, respectively. Table 4.2 details all primers used in this construction.

dextrose, 0.67% yeast nitrogen base, 0.5% ammonium sulfate and supplemented with 100 mg/L adenine sulfate, 100 mg/L L-tryptophan, 100 mg/L L-histidine-HCL, 100 mg/L L-methionine, and/or 150 mg/L L-leucine) were used for seed cultures. Complex YPD medium (1% dextrose, 1% yeast extract, 2% peptone) was used for the NADPH/$NADP^+$ and acetyl-CoA quantification studies and for TAL expression. The *g2ps1*, *lpd1, and lpd1*m genes were expressed from the *S. cerevisiae* late-phase *ADH2* promoter that is active after glucose depletion and has highest expression levels in 1% glucose YPD medium (Lee and Da Silva, 2005). Due to the late expression, plasmid stabilities remain high even in this complex medium (Shen et al., 2012).

S. cerevisiae strains were grown overnight in 5 mL selective media in an air shaker (New Brunswick Scientific) at 250 rpm and 30°C, and used to inoculate 5 mL tube cultures to an initial cell density (OD_{600}) of 0.3 (Shimadzu UV-2450 UV-VIS Spectrophotometer, Columbia, MD). A correlation factor was used to convert OD measurements to dry cell weight per liter (Chapter 2). For NADPH/$NADP^+$ and acetyl-CoA quantification, samples were taken at 24, and 48 h. For measuring polyketide TAL titers in the culture broth, the supernatant at 48 h was collected for HPLC analysis.

4.3.3 Quantitative NADPH / $NADP^+$ assay

Both NADPH and $NADP^+$ were quantified using a colorimetric 96-well plate assay (Sigma-Aldrich, St Louis, MO). Following a cold buffer (PBS) wash of the cells, a combination of 2 freeze / thaw cycles and homogenization was performed prior to deproteinization using a 10 kD cut-off spin filter (Sigma-Aldrich, St. Louis, MO). The cycling reaction was performed, and absorbance taken at 450 nm. A SpectraMax plate reader using SoftMax software was used for data

acquisition (Molecular Devices, Sunnyvale, CA). The reported values are averaged measurements taken at 24 and 48 for each duplicate sample set.

4.3.4 Acetyl-CoA assay

A boiling ethanol extraction was used on collected cell pellets for acetyl-CoA analysis, based on published procedures (Gonzalez et al., 1997; Villas-Bôas et al., 2005). Following release of the metabolites, nitrogen-supported solvent evaporation provided rapid sample concentration at room temperature. Following complete solvent evaporation, 200 µL Sigma acetyl-CoA assay buffer was added for resuspension. A fluorescence-based 96-well acetyl-CoA plate assay was used (Sigma-Aldrich, St. Louis, MO), and concentrated acetyl-CoA identified using wavelengths 535 nm (excitation) and 587 nm (emission) on a SpectraMax plate reader using SoftMax software for data acquisition. The reported values are average measurements for each duplicate sample set.

4.3.5 Polyketide detection by HPLC

The concentration of triacetic acid lactone was measured by HPLC using the following Shimadzu configuration: LC-10AT pumps (Shimadzu), UV-VIS detector (SPD-10A VP, Shimadzu), and Zorbax SB-C18 reversed-phase column (2.1x150 mm, Agilent Technologies). Acetic acid-buffered (1%) acetonitrile and water were used as the mobile and aqueous phases, respectively. A gradient program using a 95-85% Pump B gradient (H_2O with 1% acetic acid) provided an elution time of approximately 12 minutes (flow rate 0.25 mL/min, column temperature 25°C).

4.4. Results and Discussion

4.4.1 Enhancement of NADPH/NADP$^+$ and acetyl-CoA using PDH overexpression

The synthesis of polyketides requires the availability of starter metabolites that react to form the final product. Triacetic acid lactone (TAL) is an acetyl-CoA derived polyketide which requires one acetyl-CoA and two malonyl-CoA units. Since cytosolic malonyl-CoA is formed directly from acetyl-CoA, the pool of acetyl-CoA was increased through a non-native pathway. The *E. coli* pyruvate dehydrogenase complex (PDH), responsible for converting pyruvate to acetyl-CoA, was introduced to supplement the highly-regulated and ATP-requiring yeast pathway (Figure 4.2). To further increase utilization of this heterologous pathway, a modified NADP$^+$-utilizing *E. coli* PDH variant (PDHm) (Bocanegra et al., 1993) was constructed and introduced to simultaneously increase NADPH and acetyl-CoA levels.

The impact of the cofactor preference of PDHm and PDH was first characterized by measuring the NADPH/NADP$^+$ ratios for base strain BYt+pXP843 (carrying an empty vector), strain BYt+pJCT-PDH, and strain BYt+pJCT-PDHm during batch cultivation. Samples were taken at 12, 24, and 48 h and the measured ratios are shown in Figure 4.3. The PDH genes were expressed from the *S. cerevisiae* late-phase *ADH2* promoter that is active after glucose depletion (t >12 h). At 12 h (data not shown) and 48 h (late stationary), the three strains had similar ratios. This is consistent with the fact that redox systems become active at later stages of cultivation for maintenance of reactive oxygen species that form (Ying, 2008). However, at 24 h, the strain expressing the modified PDH had an increased cofactor ratio, with NADPH/NADP$^+$ ratio 3-fold

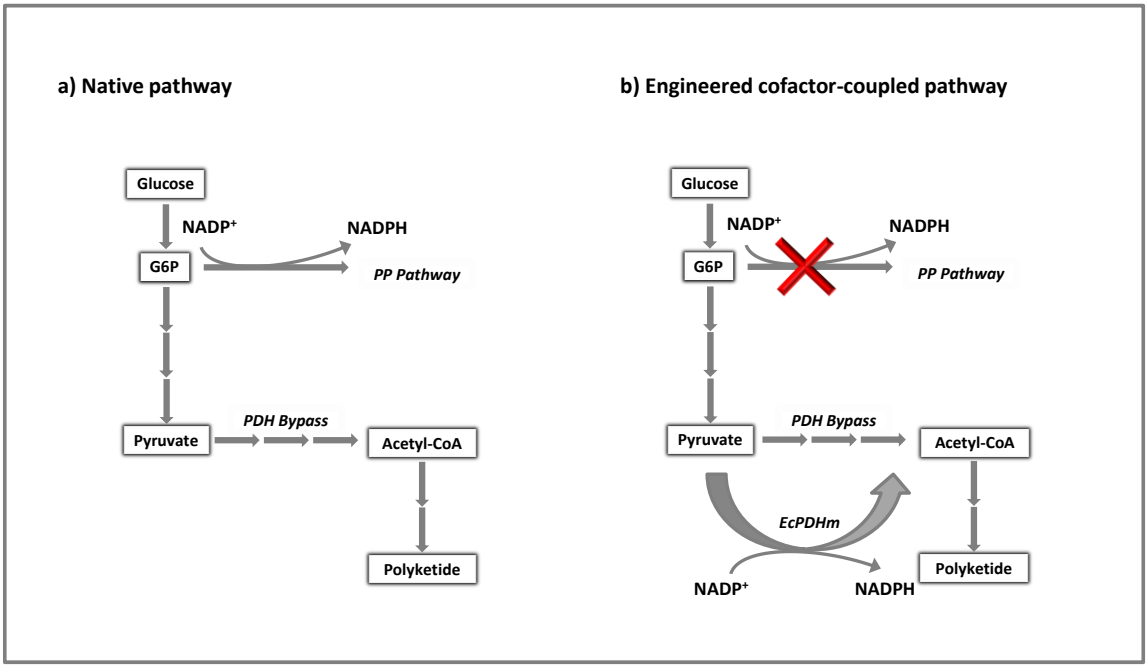


Figure 4.2. General engineering strategy applied for introducing biological driving forces. (A) Yeast relies on the pentose phosphate biosynthetic pathway (*PP Pathway*) as the major route for NADPH generation. To make acetyl-CoA in the cytoplasm, the native pyruvate dehydrogenase bypass *(PDH bypass)* converts pyruvate to acetyl-CoA in three steps. The acetyl-CoA can then be used for polyketide synthesis. (B) Our engineered cofactor-coupled pathway blocks flux through the pentose phosphate pathway, and includes the modified *E. coli* pyruvate dehydrogenase complex (PDHm) to produce acetyl-CoA and NADPH.

greater than the base strain and 2-fold greater than the strain expressing the native *E. coli* PDH. The modified PDH thus increased NADPH availability in the cell. The similar ratios at 48 h therefore suggest strong utilization of NADPH as a reducing agent in stationary cells.

Expression of the *E. coli* PDH complex may also improve acetyl-CoA levels as this pathway supplements the native yeast pyruvate dehydrogenase bypass pathway. With *in vivo* PDH activity indirectly verified by the improved cofactor ratios, the levels of acetyl-CoA were then measured for base strain BYt+pXP843 (empty vector), strain BYt+pJCT-PDH, and strain BYt+pJCT-PDHm

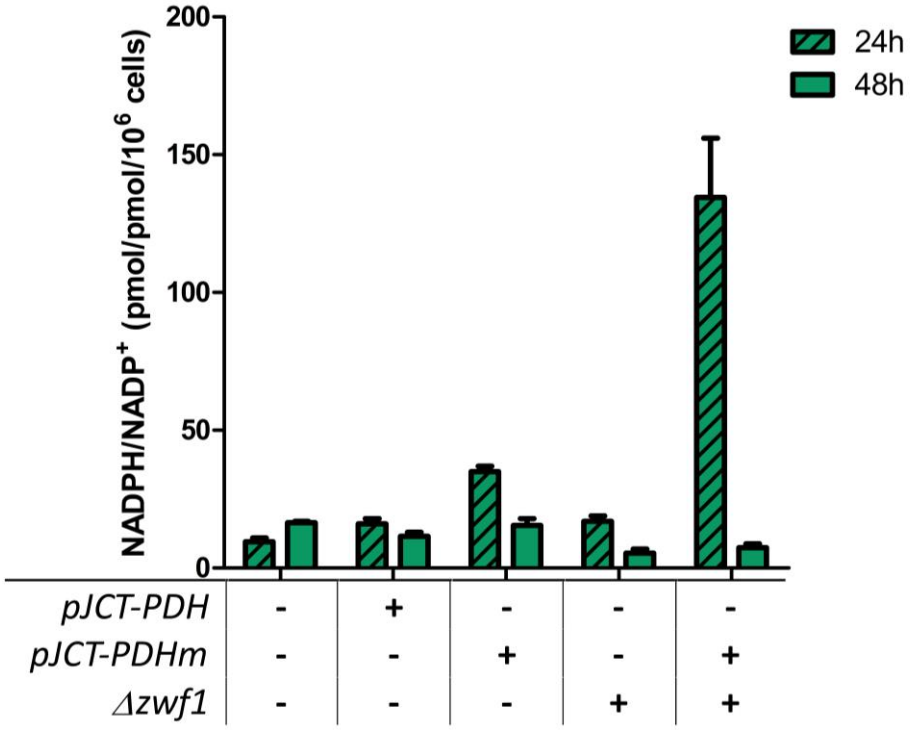


Figure 4.3. NADPH/NADP$^+$ ratios for BYt strains harboring *E. coli* native PDH (pJCT-PDH), NADPH-generating variant PDH (pJCT-PDHm), or Δ*zwf1*. Values at 24 h (hatched) and 48 h (solid) correspond to mean values ± SEM (n=2 independent experiments).

during batch culture (Figure 4.4). At 24 h, the strains expressing PDH and PDHm showed increased acetyl-CoA, similar to observations at 12 h (data not shown). By 48 h the PDH overexpression system (pJCT-PDH, Figure 4.4) was shown to have a 58% reduction in acetyl-CoA levels. This may be attributed to the NAD$^+$ consumption associated with the native *E. coli* complex and corresponding redox imbalance (Chowdhury et al., 2014), causing difficulty in maintaining reducing equivalents. Recent computational studies demonstrated the importance of cofactor preference when introducing the PDH complex in *S. cerevisiae* for TAL synthesis (Chowdhury et al., 2014), and our experimental results corroborate this requirement. While PDH overexpression

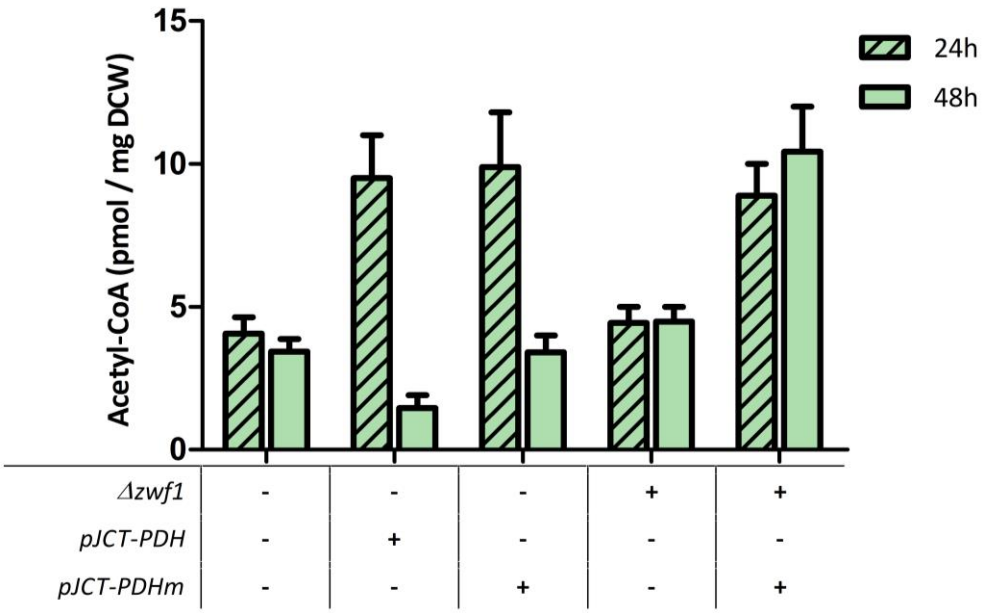


Figure 4.4. Acetyl-CoA levels for BYt strains harboring *E. coli* native PDH (pJCT-PDH), NADPH-generating variant PDH (pJCT-PDHm), or Δ*zwf1*. Values at 24 h (hatched) and 48 h (solid) correspond to mean values ± SEM (n=2 independent experiments).

can improve acetyl-CoA levels, introducing favorable redox via the NADPH-generating variant (PDHm) is more beneficial, giving rise to increases in both the NADPH/$NADP^+$ ratio and acetyl-CoA levels.

4.4.2 NADPH driving force further increases cofactor and acetyl-CoA pools

Introduction of the NADPH-generating PDH variant (PDHm) alone was capable of increasing flux toward acetyl-CoA. Elimination of the pentose phosphate pathway (the major NADPH-generating pathway in yeast) via deletion of *ZWF1* (encoding glucose-6-phosphate dehydrogenase) should further improve utilization of this new route for acetyl-CoA, as the new pathway will now be a major source of NADPH for the cell (Figure 4.2).

Both NADPH/$NADP^+$ and acetyl-CoA levels were determined for strains BYtΔ*zwf1* (carrying an empty vector) and strain BYtΔ*zwf1*+pJCT-PDHm (Figures 4.3 and 4.4). At 12 h, the coupled strain had 3-fold higher cofactor ratio levels than the base strain (data not shown). By 24 h, this strain had further elevated its NADPH/$NADP^+$ ratio to over 10-fold that of the base strain. This result is consistent with our objective to divert carbon through the heterologous PDH pathway to simultaneously support healthy NADPH formation and increase acetyl-CoA, the important starter metabolite for polyketide biosynthesis. The most notable outcome of this coupled strain was the observed change in acetyl-CoA levels (Figure 4.4). All strains with PDH overexpression (pJCT-PDH, pJCT-PDHm, and Δ*zwf1*+pJCT-PDHm) exhibited over 2-fold increases in acetyl-CoA after 24 h over the base strain. However, strain BYtΔ*zwf1*+pJCT-PDHm was the only one that had high acetyl-CoA levels at 48 h. This indicates that a large change in yeast cofactor and acetyl-CoA levels takes place only when a driving force (knockout of the pentose phosphate pathway) has been implemented. The increased levels of NADPH and acetyl-CoA can therefore be applied toward the biosynthesis of polyketides or other acetyl-CoA derived molecules.

4.4.3 Improved polyketide levels using the NADPH-generating PDH variant in a cofactor-deficient strain

The overexpression of PDHm in the Δ*zwf1* strain increased both NADPH and acetyl-CoA availability, with acetyl-CoA levels higher at late stages of the cultivation. To determine the effect on polyketide synthesis, we transformed these previously tested strains with pXP842-2PS, carrying the *G. hybrida g2ps1* (encoding 2-pyrone synthase; 2-PS) under the late-phase *ADH2* promoter. TAL levels were measured after 48 h of batch cultivation (Figure 4.5). Co-expression of

2-PS with the native *E. coli* PDH generated 0.32 g/L TAL, a 30% increase relative to the base yeast strain. This was the smallest increase out of all strains tested, indicating that the cofactor imbalance introduced with the NAD^+-consuming PDH was not favorable. Co-expression of 2-PS with the modified *E. coli* PDH (PDHm), resulted in a 3-fold increase in titer to 0.75 g/L TAL. To create a driving force for acetyl-CoA, the *ZWF1* deletion blocking the major NADPH-generating pathway was important (Figure 4.4), and should lead to increased TAL production. This Δ*zwf1* modification alone was previously shown to result in TAL increases, as pentose biosynthesis is a significant carbon drain from glycolysis (Cardenas and Da Silva, 2014). In the current study, knocking out *ZWF1* alone resulted in a 2.6-fold improvement over the base strain TAL. Expression of the native *E. coli* PDH in the Δ*zwf1* strain gave similar TAL levels as in the Δ*zwf1* strain alone. In contrast, coupling the Δ*zwf1* strain with PDHm overexpression increased TAL levels an additional 70% relative to BYtΔ*zwf1*, and 4.5-fold higher relative to the base yeast strain. Therefore, introducing a biological driving force for acetyl-CoA production via PDHm improved the acetyl-CoA pools (Figure 4.4), and thus increased TAL synthesis (Figure 4.5) in late batch culture. This strategy may prove even more useful for polyketides that utilize NADPH for their synthesis, and thus where an even greater NADPH pool is required.

4.4.4 Eliminating mitochondrial precursor transport pathways

Our overall aim is to direct carbon toward cytosolic acetyl-CoA for polyketide production. In *S. cerevisiae*, various shuttling events involving acetyl-CoA and pyruvate can limit the availability of this metabolite. We have considered two main routes, the mitochondrial pyruvate carrier (MPC) and the carnitine shuttle system, as well as the common mitochondrial porins. For

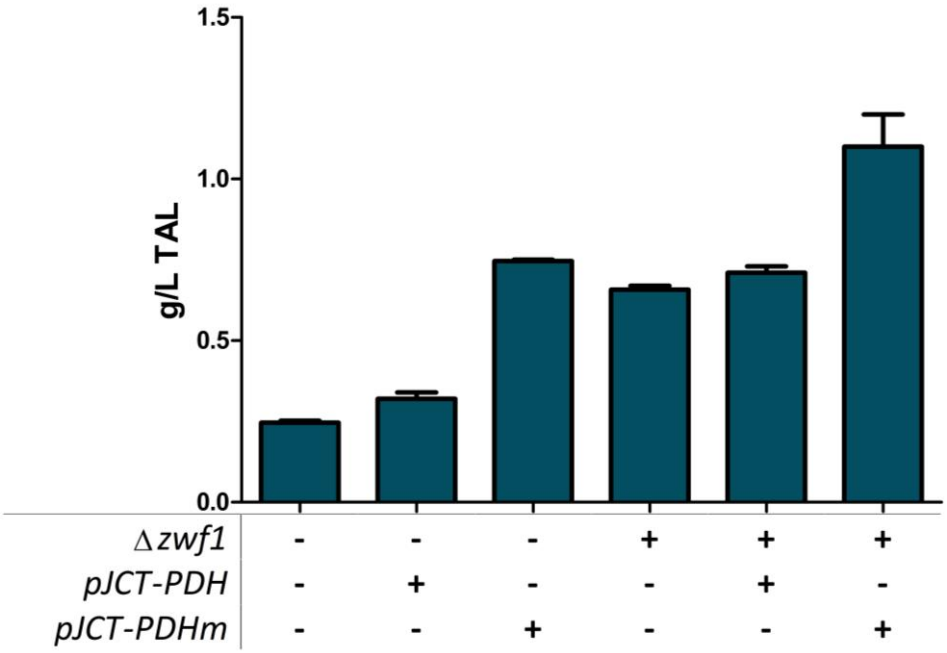


Figure 4.5. TAL titers (g/L) for BYt strains after 48 cultivation harboring *E. coli* native PDH (pJCT-PDH), NADPH-generating variant PDH (pJCT-PDHm), and / or Δ*zwf1*. Values correspond to mean values ± SEM (n=6 independent experiments).

the MPC, gene products Mpc1, Mpc2, and Mpc3 participate in active transport, where Mpc1/Mpc2 and Mpc1/Mpc3 complex to form part of the active transport mechanism (Mpc2 and Mpc3 are paralogs) (Herzig et al., 2012). Using the MPC, pyruvate transport into the mitochondrial matrix and subsequent conversion by decarboxylation activity commits this carbon flow toward mitochondrial acetyl-CoA (Figure 4.6). Replenishing pyruvate will be important to effectively use our engineered PDHm platform. Similarly, cytosolic acetyl-CoA can be transported to the mitochondria via the carnitine shuttle system (Franken et al., 2008) instead of being utilized in the cytosol (Figure 4.6). Therefore, limiting transport of pyruvate and acetyl-

CoA from the cytoplasm while minimizing deleterious effects on the cell may be an effective means of increasing polyketide production.

We first identified a group of eight proteins involved in pyruvate and acetyl-CoA transport (Figure 4.6), and evaluated the effects of their single gene knockouts on both growth and TAL production. This set of genes encodes for pathways which facilitate transport of various metabolites across the outer membrane by porins (Por1, Por2), specific transport of pyruvate from the intermembrane space into the mitochondrial matrix (Mpc1, Mpc2, Mpc3), and transport of acetyl-CoA from the cytosol into intermembrane space (Yat1, Yat2). The knockout strains were transformed with pXP842-2PS and samples were taken at 48 h. Por1 and Por2 are responsible for the mitochondrial permeability of charged solutes including pyruvate, ATP, and phosphate by forming voltage-dependent, anion-selective channels (VDACs) in the outer membrane (Blachly-Dyson et al., 1997). Due to differences in their activity, the Δ*por1* and Δ*por2* strains produced very different levels of TAL (Figure 4.7). While Δ*por2* increased TAL 2.5-fold to 0.62 g/L, Δ*por1* nearly abolished product formation with a 94% reduction in titer relative to the parent strain. This is consistent with reports that Por1 maintains mitochondrial osmotic stability (Galganska et al., 2008). Por2 is not associated with this activity, and instead represses Sod1 (superoxide dismutase) which in turn alleviates respiration repression. Additionally, Δ*por1* has been found to suffer a growth defect due to reduced respiratory function on non-fermentable carbon, in contrast to the increase under the same conditions for Δ*por2* (Dihanich et al., 1987; Sánchez et al., 2001). This is consistent with the growth observed for the culture; the Δ*por1* strain reached a final density equal to one-third (4.2 g DCW/L) that of the base strain (Figure 4.7).

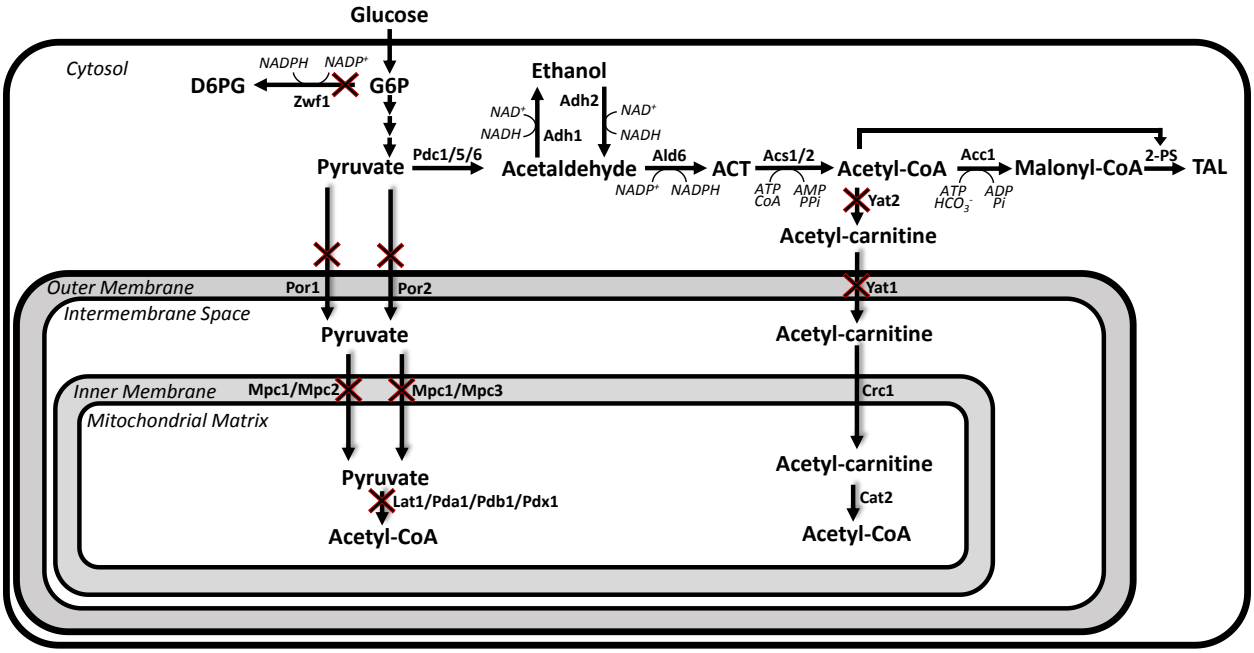


Figure 4.6. Metabolic pathways related to the transport and utilization of pyruvate and acetyl-CoA, the starter metabolites for subsequent TAL formation (catalyzed by 2-PS). Pathways highlighted include the major NADPH supply (Zwf1), mitochondrial pyruvate carrier (MPC) system (Mpc1/Mpc2 and Mpc1/Mpc3), mitochondrial porins (Por1 and Por2), and carnitine shuttle system (Yat1, Yat2, Crc1, Cat2). Reactions chosen for disruption in this study have been crossed out.

The disruption of the mitochondrial pyruvate carrier (Mpc1, Mpc2, Mpc3) also showed varying polyketide levels following the single gene knockouts. The Δ*mpc1* strain showed no significant effect on TAL titer relative to the base strain, while Δ*mpc2* and Δ*mpc3* strains provided approximately 3-fold and 1-fold increases in TAL titer, respectively. Final cell densities were 86%, 89%, and 74% of the base strain for Δ*mpc1*, Δ*mpc2*, and Δ*mpc3*, respectively. The Mpc1p/Mpc3p complex is necessary for retaining strong respiration and its expression is closely tied to stress responses (Timón-Gómez et al., 2013), consistent with the data showing the lowest cell density by Δ*mpc3* for strains of altered MPC function. The data therefore indicates an intact Mpc1/Mpc3

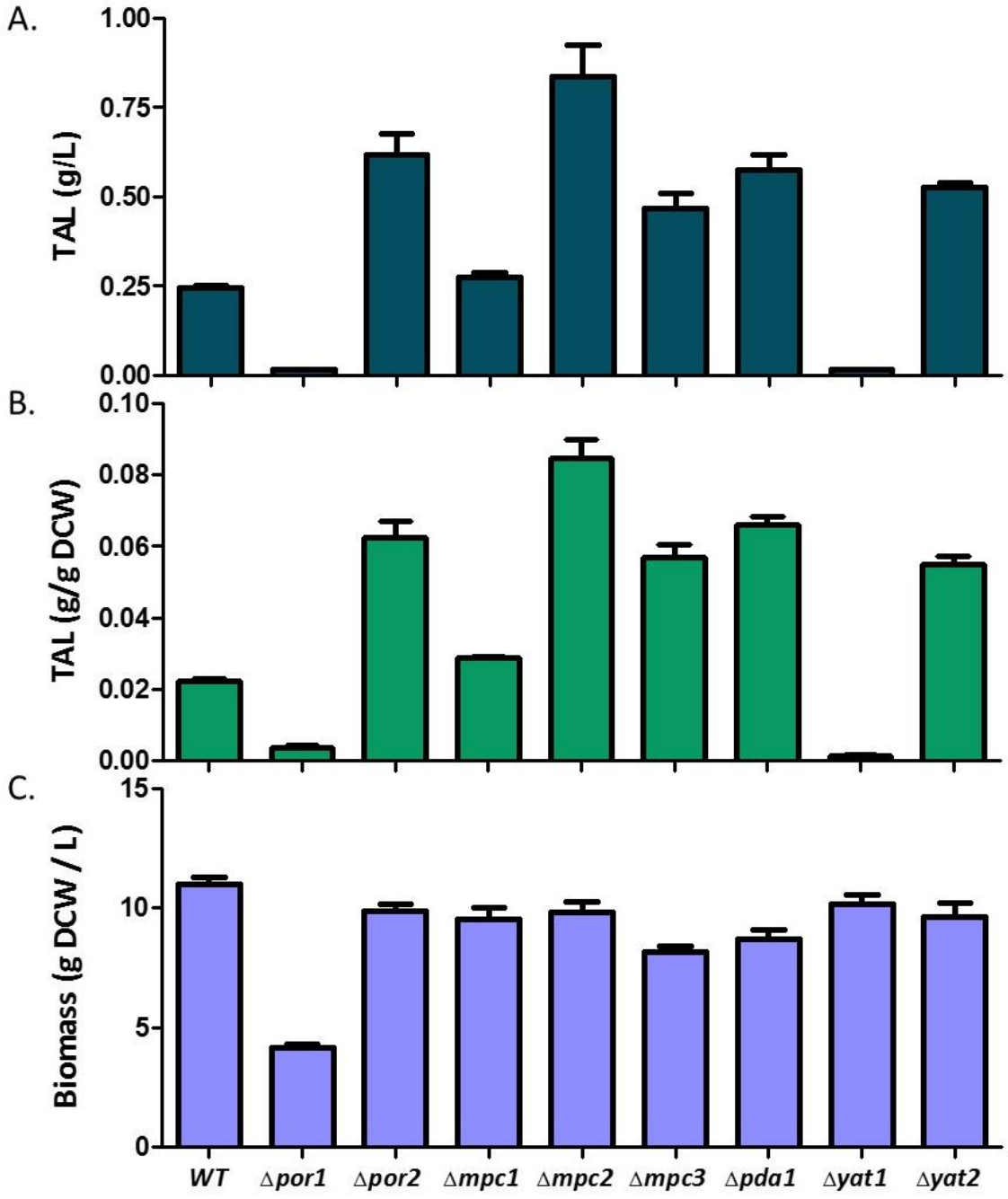


Figure 4.7. TAL titers (g/L, Panel A), specific levels (g/g DCW, Panel B), and biomass (g DCW/L, Panel C) following single gene knockouts after 48 h cultivation. Strains correspond to genes selected for their roles in transport and utilization of pyruvate and acetyl CoA (refer to Figure 4.6). Increases in TAL guided the subsequent selection of gene disruptions for combination with the cofactor-coupled PDHm pathway. Bars represent means ± SEM (n=6 independent experiments).

pyruvate carrier and disruption of the Mpc1/Mpc2 complex supports respiratory function while likely increasing cytosolic pyruvate levels necessary for significant improvement in TAL synthesis. The ability to convert pyruvate to acetyl-CoA in the mitochondria was also eliminated by deleting *PDA1* (BYtΔ*pda1*); this deletion results in no detectable pyruvate decarboxylase activity (Gey et al., 2008). This strain produced greater than 2-fold more TAL than the base yeast strain (0.58 g/L), and had a final cell density at 78% of the base strain. This indicates that eliminating this pathway can improve titers of acetyl-CoA derived compounds like TAL.

For cytosolic acetyl-CoA transport to the mitochondria, the first two committed steps from the carnitine shuttle system were evaluated. Yat1 and Yat2 are responsible for committing acetyl-CoA to the carnitine shuttle system and transporting it to the mitochondrial matrix, respectively (Figure 4.6). In the Δ*yat1* strain, a 94% drop in TAL was observed, while in the Δ*yat2* strain TAL levels increased over 2-fold relative to the base strain (Figure 4.7). The data suggests the cell is much more tolerant of eliminating the cytosolic Yat2 than the mitochondrial Yat1; the mitochondrial Yat1 is responsible for shuttling acetyl-CoA from other compartments like the peroxisome and may be more critical for proper cell function. In addition, Yat2 is highly active when ethanol is used as the carbon source (Franken et al., 2008; Swiegers et al., 2001), the same time that TAL synthesis is occurring with the *ADH2* promoter, and thus deleting this function may be advantageous.

4.4.5 Combining transport gene disruptions with engineered NADPH driving force

From the eight genes tested for disruption, four were chosen for combination to further reduce the flow of pyruvate and acetyl-CoA transport into the mitochondria. With the single

knockout strains, was achieved and significant improvements in polyketide levels were observed by disrupting Δ*por2*, Δ*mpc2*, Δ*pda1*, and Δ*yat2*. Therefore, we deleted all four genes in a single strain and measured TAL levels (Figure 4.8, Panels A-C). As before, the constructed Δ*por2* yielded 0.62 g/L TAL and reduced biomass by 10% relative to the base strain. Knockout of *MPC2* further elevated titers to 0.98 g/L TAL, while also reducing biomass by an additional 20%. Further interventions had no negative effect on biomass, and incorporation of Δ*pda1* and Δ*yat2* led to additional titer increases to 1.27 and 1.33 g/L, corresponding to a final yield on glucose of 0.13 g/g (28% of theoretical). The diminished improvements observed in the four-knockout strain suggest that other pathways are now limiting the production of TAL. This strain was then transformed with pJCT-PDHm to determine if coupling PDHm expression with the transport knockouts would further increase TAL levels (Figure 4.8). However, no significant increase was observed. Both the absence of a driving force and the presence of new limiting factors are likely causes for this result.

To also introduce a NADPH driving force, the four deletions were incorporated in a Δ*zwf1* strain. TAL levels were measured after 48 h of batch culture (Figure 4.8, Panels D-F). As shown previously (Figure 4.5), deletion of *ZWF1* alone led to increased TAL levels, likely due to the improved carbon flux through glycolysis. Similar increases were seen for the strains with the sequential transport knockouts; a 2-fold improvement was achieved by incorporating this four-gene disruption (Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*) into the Δ*zwf1* strain. However, in this case, subsequent transformation with pJCT-PDHm led to a further significant increase in specific TAL production ($p<0.001$) and in titer. The combination of approaches led to a final titer of 1.6 g/L

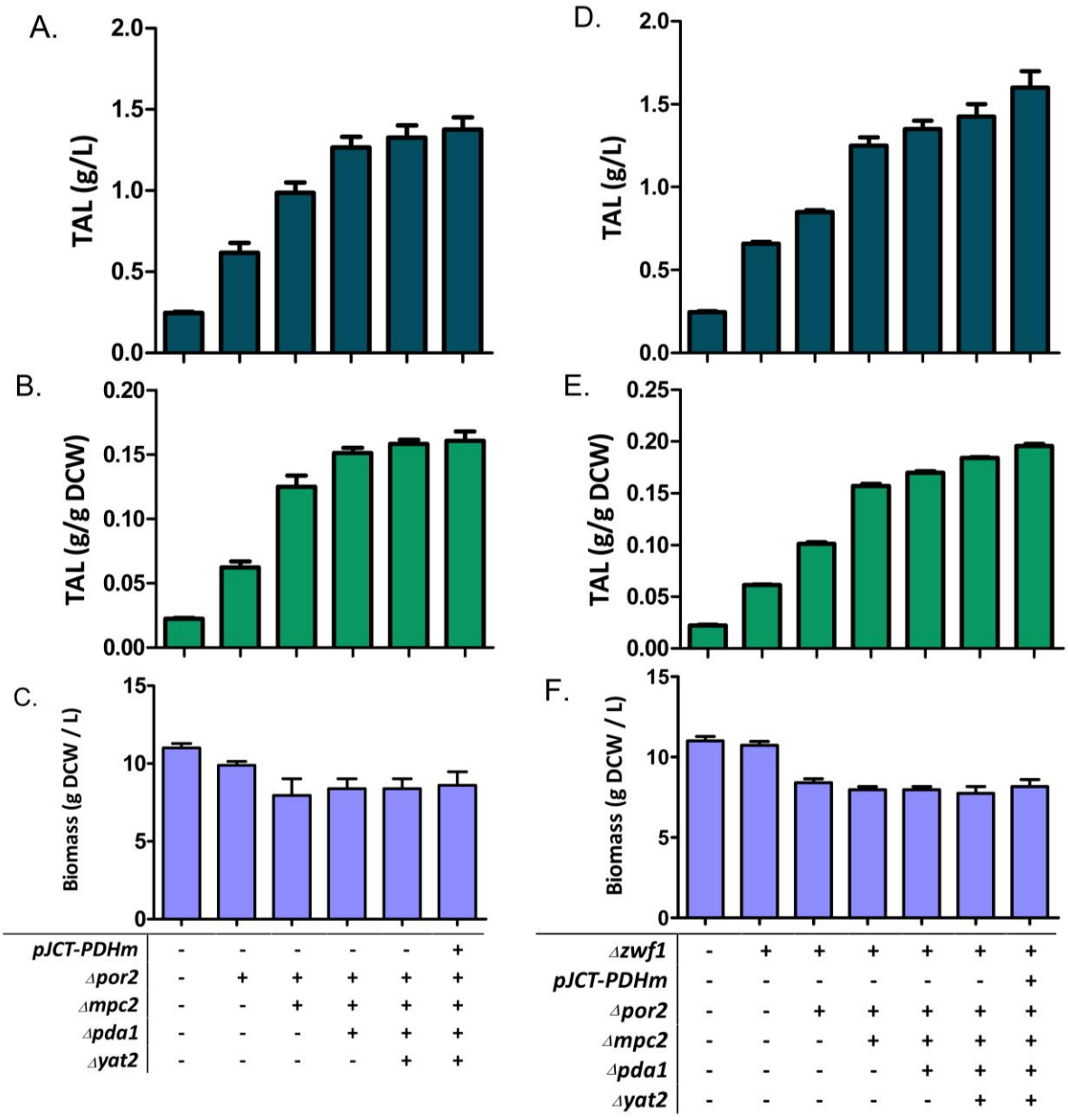


Figure 4.8. TAL titers (g/L, Panel A or D), specific levels (g/g DCW, Panel B or E), and biomass (g DCW/L, Panel C or F) in strains using pJCT-PDHm engineered for transport alone (Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*, Panels A-C), or in addition to Δ*zwf1* (Panels D-F), after 48 h cultivation. All strains express 2-PS on pXP842 for TAL synthesis. Bars represent means ± SEM (n=6 independent experiments).

TAL, corresponding to 35% of the theoretical yield on glucose (0.16 g/g), the highest reported value for batch fermentation. This strategy therefore proved useful in maintaining high pyruvate and acetyl-CoA levels in the cytosol, which was in turn used for polyketide formation.

4.5 Conclusions

S. cerevisiae has been shown to be a model host organism capable of significant metabolic rewiring for producing acetyl-CoA derived compounds. This study considered creating a biological driving force that could increase acetyl-CoA pools applicable toward producing polyketides. Characterization of the modified NADPH-generating *E. coli* PDH (PDHm) relative to the native *E. coli* PDH indicated that NADPH synthesis was beneficial even though TAL biosynthesis does not require it directly. Coupling this with a Δ*zwf1* strain (to create additional driving force for acetyl-CoA synthesis) led to TAL levels more than 4-fold higher than the parent strain. This improvement was attributed to increased NADPH/$NADP^+$ and acetyl-CoA levels, up to 10-fold and 3-fold higher than in the parent. To further increase the availability of acetyl-CoA for polyketide synthesis, gene deletions were used to reduce transport of acetyl-CoA and pyruvate into the mitochondria. The single gene deletions in *POR2, MPC2, PDA1,* and *YAT2* each provided over a 3-fold increase in titer (up to 0.84 g/L) relative to the parent strain. Coupling these four in a single knockout strain (Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*) provided increased titer to 1.33 g/L and a reduction of biomass by 20%. Introducing this four-gene disruption into the coupled Δ*zwf1* / pJCT-PDHm strain provided an additional 50% improvement in titer (1.6 g/L), creating 35% of theoretical yield of TAL on glucose (0.16 g/g), the highest reported value for batch cultivation. This coupled strategy (summarized in Figure 4.9) illustrates a promising step toward improving strain development for synthesis of acetyl-CoA derived compounds. Additional features can also be incorporated to further enhance acetyl-CoA pools, including additional knockouts for reduced byproduct formation (Chapter 2), better heterologous enzyme expression including the use of synthetic promoters (Nevoigt et al., 2006) and stable chromosomal integrations (Fang et al., 2011), as well

as downregulation of competing pathways for yeast essential genes like those in fatty acid biosynthesis.

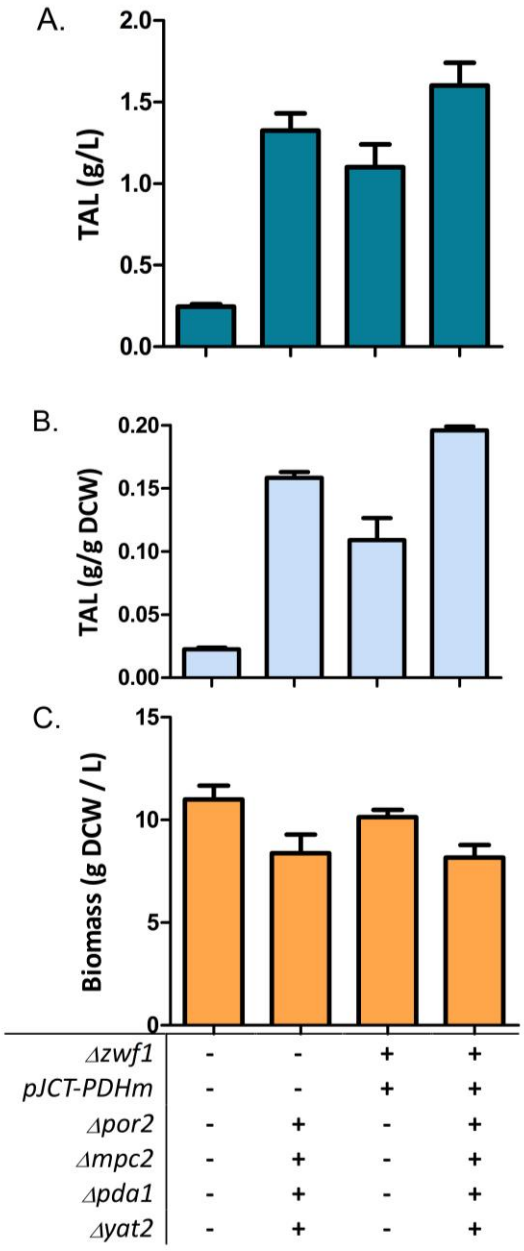


Figure 4.9. Final titer summary for polyketide production (g/L TAL) in *S. cerevisiae*. The introduction of transport knockouts (Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*) and a NADPH driving force (Δ*zwf1*+pJCT-PDHm) led to increased titers, and combining them (Δ*zwf1*Δ*por2*Δ*mpc2*Δ*pda1*Δ*yat2*+pJCT-PDHm) provided the largest increase over all strains tested. Bars represent means ± SEM (n=6 independent experiments).

4.6 Acknowledgements

This research was supported by the National Science Foundation (Grant No. EEC-0813570) through the Engineering Research Center CBiRC (Center for Biorenewable Chemicals).

4.7 References

Austin, M.B., Noel, J.P., 2003. The chalcone synthase superfamily of type III polyketide synthases. Nat. Prod. Rep. 20, 79–110. doi:10.1039/B100917F

Blachly-Dyson, E., Song, J., Wolfgang, W.J., Colombini, M., Forte, M., 1997. Multicopy suppressors of phenotypes resulting from the absence of yeast VDAC encode a VDAC-like protein. Mol. Cell. Biol. 17, 5727–5738.

Bocanegra, J.A., Scrutton, N.S., Perham, R.N., 1993. Creation of an NADP-dependent pyruvate dehydrogenase multienzyme complex by protein engineering. Biochemistry 32, 2737–2740.

Bricker, D.K., Taylor, E.B., Schell, J.C., Orsak, T., Boutron, A., Chen, Y.-C., Cox, J.E., Cardon, C.M., Vranken, J.G.V., Dephoure, N., Redin, C., Boudina, S., Gygi, S.P., Brivet, M., Thummel, C.S., Rutter, J., 2012. A Mitochondrial Pyruvate Carrier Required for Pyruvate Uptake in Yeast, Drosophila, and Humans. Science 337, 96–100. doi:10.1126/science.1218099

Cardenas, J., Da Silva, N.A., 2014. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. Metabolic Engineering 25, 194–203. doi:10.1016/j.ymben.2014.07.008

Chen, F., Zhou, J., Shi, Z., Liu, L., Du, G., Chen, J., 2010. [Effect of acetyl-CoA synthase gene overexpression on physiological function of Saccharomyces cerevisiae]. Wei Sheng Wu Xue Bao 50, 1172–1179.

Chen, Y., Daviet, L., Schalk, M., Siewers, V., Nielsen, J., 2013. Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. Metabolic Engineering 15, 48–54. doi:10.1016/j.ymben.2012.11.002

Choi, J.W., Da Silva, N.A., 2014. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase. Journal of Biotechnology 187, 56–59. doi:10.1016/j.jbiotec.2014.07.430

Chowdhury, A., Zomorrodi, A.R., Maranas, C.D., 2014. k-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design. PLoS Comput Biol 10, e1003487. doi:10.1371/journal.pcbi.1003487

Dale, J.W., Greenaway, P.J., 1985. Preparation of Chromosomal DNA from E. coli. Methods Mol. Biol. 2, 197–200. doi:10.1385/0-89603-064-4:197

Dihanich, M., Suda, K., Schatz, G., 1987. A yeast mutant lacking mitochondrial porin is respiratory-deficient, but can recover respiration with simultaneous accumulation of an 86-kd extramitochondrial protein. EMBO J 6, 723–728.

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. doi:10.1002/yea.1824

Franken, J., Kroppenstedt, S., Swiegers, J.H., Bauer, F.F., 2008. Carnitine and carnitine acetyltransferases in the yeast Saccharomyces cerevisiae: a role for carnitine in stress protection. Curr Genet 53, 347–360. doi:10.1007/s00294-008-0191-0

Galganska, H., Budzinska, M., Wojtkowska, M., Kmita, H., 2008. Redox regulation of protein expression in Saccharomyces cerevisiae mitochondria: Possible role of VDAC. Archives of Biochemistry and Biophysics 479, 39–45. doi:10.1016/j.abb.2008.08.010

Gey, U., Czupalla, C., Hoflack, B., Rödel, G., Krause-Buchholz, U., 2008. Yeast Pyruvate Dehydrogenase Complex Is Regulated by a Concerted Activity of Two Kinases and Two Phosphatases. J. Biol. Chem. 283, 9759–9767. doi:10.1074/jbc.M708779200

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Meth 6, 343–345. doi:10.1038/nmeth.1318

Gietz, D., Jean, A.S., Woods, R.A., Schiestl, R.H., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucl. Acids Res. 20, 1425–1425. doi:10.1093/nar/20.6.1425

Gonzalez, B., François, J., Renaud, M., 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. Yeast 13, 1347–1355. doi:10.1002/(SICI)1097-0061(199711)13:14<1347::AID-YEA176>3.0.CO;2-O

Hertweck, C., 2009. The Biosynthetic Logic of Polyketide Diversity. Angewandte Chemie International Edition 48, 4688–4716. doi:10.1002/anie.200806121

Herzig, S., Raemy, E., Montessuit, S., Veuthey, J.-L., Zamboni, N., Westermann, B., Kunji, E.R.S., Martinou, J.-C., 2012. Identification and Functional Expression of the Mitochondrial Pyruvate Carrier. Science 337, 93–96. doi:10.1126/science.1218530

Hill, J., Donald, K.A.I.G., Griffiths, D.E., 1991. DMSO-enhanced whole cell yeast transformation. Nucl. Acids Res. 19, 5791–5791. doi:10.1093/nar/19.20.5791

Hofbauer, H.F., Schopf, F.H., Schleifer, H., Knittelfelder, O.L., Pieber, B., Rechberger, G.N., Wolinski, H., Gaspar, M.L., Kappe, C.O., Stadlmann, J., Mechtler, K., Zenz, A., Lohner, K., Tehlivets, O., Henry, S.A., Kohlwein, S.D., 2014. Regulation of Gene Expression through a Transcriptional Repressor that Senses Acyl-Chain Length in Membrane Phospholipids. Developmental Cell 29, 729–739. doi:10.1016/j.devcel.2014.04.025

Howat, S., Park, B., Oh, I.S., Jin, Y.-W., Lee, E.-K., Loake, G.J., 2014. Paclitaxel: biosynthesis, production and future prospects. New Biotechnology 31, 242–245. doi:10.1016/j.nbt.2014.02.010

Jeandet, P., Delaunois, B., Aziz, A., Donnez, D., Vasserot, Y., Cordelier, S., Courot, E., 2012. Metabolic Engineering of Yeast and Plants for the Production of the Biologically Active Hydroxystilbene, Resveratrol. BioMed Research International 2012, e579089. doi:10.1155/2012/579089

Jeandet, P., Vasserot, Y., Chastang, T., Courot, E., 2013. Engineering Microbial Cells for the Biosynthesis of Natural Compounds of Pharmaceutical Significance. BioMed Research International 2013, e780145. doi:10.1155/2013/780145

Jez, J.M., Bowman, M.E., Noel, J.P., 2002. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. Proc Natl Acad Sci U S A 99, 5319–5324. doi:10.1073/pnas.082590499

Jong, B.W. de, Shi, S., Siewers, V., Nielsen, J., 2014. Improved production of fatty acid ethyl esters in Saccharomyces cerevisiae through up-regulation of the ethanol degradation pathway and expression of the heterologous phosphoketolase pathway. Microbial Cell Factories 13, 39. doi:10.1186/1475-2859-13-39

Kocharin, K., Chen, Y., Siewers, V., Nielsen, J., 2012. Engineering of acetyl-CoA metabolism for the improved production of polyhydroxybutyrate in Saccharomyces cerevisiae. AMB Express 2, 52. doi:10.1186/2191-0855-2-52

Kocharin, K., Siewers, V., Nielsen, J., 2013. Improved polyhydroxybutyrate production by Saccharomyces cerevisiae through the use of the phosphoketolase pathway. Biotechnology and Bioengineering n/a–n/a. doi:10.1002/bit.24888

Krivoruchko, A., Nielsen, J., 2015. Production of natural products through metabolic engineering of Saccharomyces cerevisiae. Current Opinion in Biotechnology 35, 7–15. doi:10.1016/j.copbio.2014.12.004

Krivoruchko, A., Serrano-Amatriain, C., Chen, Y., Siewers, V., Nielsen, J., 2013. Improving biobutanol production in engineered Saccharomyces cerevisiae by manipulation of acetyl-CoA metabolism. Journal of Industrial Microbiology & Biotechnology 40, 1051–1056. doi:10.1007/s10295-013-1296-0

Leber, C., Da Silva, N.A., 2013. Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids. Biotechnology and Bioengineering. doi:10.1002/bit.25021

Lee, J., Shin, M.-K., Ryu, D.-K., Kim, S., Ryu, W.-S., 2010. Insertion and deletion mutagenesis by overlap extension PCR. Methods Mol. Biol. 634, 137–146. doi:10.1007/978-1-60761-652-8_10

Lee, M.K., Da Silva, N.A., 2005. Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis. Yeast 22, 431–440. doi:10.1002/yea.1221

Lian, J., Si, T., Nair, N.U., Zhao, H., 2014. Design and construction of acetyl-CoA overproducing Saccharomyces cerevisiae strains. Metabolic Engineering 24, 139–149. doi:10.1016/j.ymben.2014.05.010

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.

Li, X., Guo, D., Cheng, Y., Zhu, F., Deng, Z., Liu, T., 2014. Overproduction of fatty acids in engineered Saccharomyces cerevisiae. Biotechnology and bioengineering 111, 1841–1852.

Lussier, F.-X., Colatriano, D., Wiltshire, Z., Page, J.E., Martin, V.J.J., 2012. Engineering Microbes for Plant Polyketide Biosynthesis. Computational and Structural Biotechnology Journal 3, 1–11. doi:10.5936/csbj.201210020

Ma, S.M., Li, J.W.-H., Choi, J.W., Zhou, H., Lee, K.K.M., Moorthie, V.A., Xie, X., Kealey, J.T., Silva, N.A.D., Vederas, J.C., Tang, Y., 2009. Complete Reconstitution of a Highly Reducing Iterative Polyketide Synthase. Science 326, 589–592. doi:10.1126/science.1175602

Nevoigt, E., Kohnke, J., Fischer, C.R., Alper, H., Stahl, U., Stephanopoulos, G., 2006. Engineering of Promoter Replacement Cassettes for Fine-Tuning of Gene Expression in Saccharomyces cerevisiae. Appl Environ Microbiol 72, 5266–5273. doi:10.1128/AEM.00530-06

Nikolau, B.J., Perera, M.A.D.N., Brachova, L., Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. Plant J. 54, 536–545. doi:10.1111/j.1365-313X.2008.03484.x

Pronk, J.T., Yde Steensma, H., Van Dijken, J.P., 1996. Pyruvate Metabolism in Saccharomyces cerevisiae. Yeast 12, 1607–1633. doi:10.1002/(SICI)1097-0061(199612)12:16<1607::AID-YEA70>3.0.CO;2-4

Runguphan, W., Keasling, J.D., 2014. Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. Metabolic Engineering 21, 103–113. doi:10.1016/j.ymben.2013.07.003

Ruohonen, L., Aalto, M.K., Keränen, S., 1995. Modifications to the ADH1 promoter of Saccharomyces cerevisiae for efficient production of heterologous proteins. Journal of Biotechnology 39, 193–203. doi:16/0168-1656(95)00024-K

Sambrook, J.J., Russell, D.D.W., 2001. Molecular cloning: a laboratory manual. Vol. 2. CSHL Press.

Sánchez, N.S., Pearce, D.A., Cardillo, T.S., Uribe, S., Sherman, F., 2001. Requirements of Cyc2p and the Porin, Por1p, for Ionic Stability and Mitochondrial Integrity in Saccharomyces cerevisiae. Archives of Biochemistry and Biophysics 392, 326–332. doi:10.1006/abbi.2001.2465

Santos, C.N.S., Koffas, M., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metabolic Engineering 13, 392–400. doi:16/j.ymben.2011.02.002

Schwartz, T.J., Johnson, R.L., Cardenas, J., Okerlund, A., Da Silva, N.A., Schmidt-Rohr, K., Dumesic, J.A., 2014. Engineering Catalyst Microenvironments for Metal-Catalyzed Hydrogenation of Biologically Derived Platform Chemicals. Angew. Chem. Int. Ed. n/a–n/a. doi:10.1002/anie.201407615

Shen, C.R., Lan, E.I., Dekishima, Y., Baez, A., Cho, K.M., Liao, J.C., 2011. Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia coli. Appl. Environ. Microbiol. 77, 2905–2915. doi:10.1128/AEM.03034-10

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. doi:10.1002/yea.2930

Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.-K., Keasling, J.D., 2007. Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. Metabolic Engineering 9, 160–168. doi:10.1016/j.ymben.2006.10.005

Shi, S., Chen, Y., Siewers, V., Nielsen, J., 2014. Improving Production of Malonyl Coenzyme A-Derived Metabolites by Abolishing Snf1-Dependent Regulation of Acc1. mBio 5, e01130–14. doi:10.1128/mBio.01130-14

Su, W., Xiao, W.-H., Wang, Y., Liu, D., Zhou, X., Yuan, Y.-J., 2015. Alleviating Redox Imbalance Enhances 7-Dehydrocholesterol Production in Engineered Saccharomyces cerevisiae. PLOS ONE 10, e0130840. doi:10.1371/journal.pone.0130840

Swiegers, J.H., Dippenaar, N., Pretorius, I.S., Bauer, F.F., 2001. Carnitine-dependent metabolic activities in Saccharomyces cerevisiae: three carnitine acetyltransferases are essential in a carnitine-dependent strain. Yeast 18, 585–595. doi:10.1002/yea.712

Tang, X., Feng, H., Chen, W.N., 2013. Metabolic engineering for enhanced fatty acids synthesis in Saccharomyces cerevisiae. Metabolic Engineering 16, 95–102. doi:10.1016/j.ymben.2013.01.003

Timón-Gómez, A., Proft, M., Pascual-Ahuir, A., 2013. Differential Regulation of Mitochondrial Pyruvate Carrier Genes Modulates Respiratory Capacity and Stress Tolerance in Yeast. PLoS One 8. doi:10.1371/journal.pone.0079405

Villas-Bôas, S.G., Højer-Pedersen, J., Åkesson, M., Smedsgaard, J., Nielsen, J., 2005. Global metabolite analysis of yeast: evaluation of sample preparation methods. Yeast 22, 1155–1169. doi:10.1002/yea.1308

Wattanachaisaereekul, S., Lantz, A.E., Nielsen, M.L., Andrésson, O.S., Nielsen, J., 2007. Optimization of heterologous production of the polyketide 6-MSA in Saccharomyces cerevisiae. Biotechnology and Bioengineering 97, 893–900.

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. doi:10.1016/j.ymben.2008.04.005

Xu, W., Chooi, Y.-H., Choi, J.W., Li, S., Vederas, J.C., Da Silva, N.A., Tang, Y., 2013. Finding the Missing Link in Lovastatin Biosynthesis: LovG is the Thioesterase Required for Dihydromonacolin L Release and Lovastatin Nonaketide Synthase Turnover. Angew Chem Int Ed Engl 52. doi:10.1002/anie.201302406

Ying, W., 2008. NAD^+/NADH and $NADP^+$/NADPH in Cellular Functions and Cell Death: Regulation and Biological Consequences. Antioxidants & Redox Signaling 10, 179–206. doi:10.1089/ars.2007.1672

Chapter 5:

Conclusion and Future Perspectives

5.1 Conclusion

The overarching goal of this research was to determine strategies for modifying the metabolism of yeast for improved polyketide synthesis using both metabolic and enzyme engineering. These efforts were applied in *Saccharomyces cerevisiae* toward the high-level synthesis of triacetic acid lactone (TAL), a valuable precursor for the chemical industry (Chia et al., 2012). Rational engineering of both the strain and enzyme were performed in an integrated approach. First, the native strain was designed with gene disruptions to support carbon flow towards acetyl-CoA and in turn TAL. This strain relied on traditional metabolic engineering and computational modeling to improve carbon utilization and eliminate losses to competing pathways. Second, the 2-pyrone synthase from *G. hybrida* was designed for increased catalytic efficiency (k_{cat}/K_M) and stability using engineering of the both the active site and surface exposed sites as characterized by enzyme kinetics and oxidation studies. Directly forcing the cell to make more acetyl-CoA in order to maintain redox was a parallel approach to further enhance the availability of this starter metabolite and thus polyketide production. This strategy was enhanced by evaluating a combination of disruptions in genes for pyruvate and acetyl-CoA transport to the mitochondria. Finally, fermentation strategies were employed to leverage the stabilized, more catalytically active enzymes expressed by strains with diverted carbon flow toward the target compound. This overall strategy has led to significant increases in the performance of our strains over the course of this study (Figure 5.1). The emphasis on improving acetyl-CoA pools for TAL production allows these efforts to be directly applicable to production of other relevant acetyl-CoA derived compounds, including other polyketides and fatty acids.

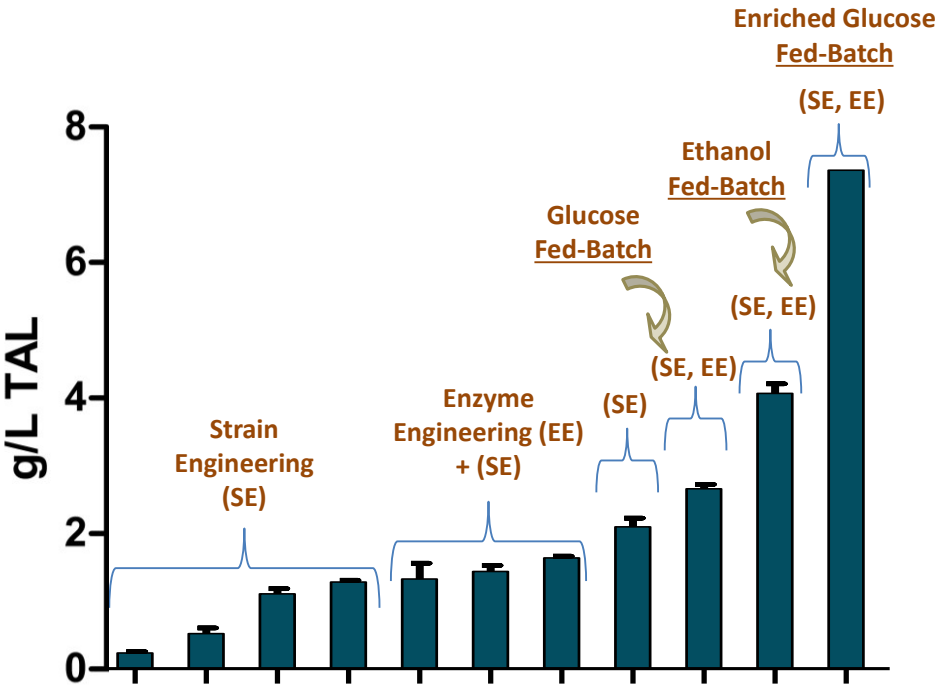


Figure 5.1 Overall increases in various platform strains for TAL synthesis in *S. cerevisiae*, showcasing the development of the strain over the course of this study. Strains correspond to strain engineering (SE) or enzyme engineering (EE) approaches, or both (SE, EE), with or without the use of fed-batch operation. Values (excluding final bar, single run) represent mean ± one standard deviation (n=3 independent experiments).

5.2 Future Perspectives

Although this work made major advances, the ever-increasing need for biobased materials will continue to drive pressure for better, more productive microbial hosts. A transformed chemical manufacturing landscape will require low cost, readily accessible feedstocks like biomass, while co-existing with current agricultural demands for human consumption (Carus et al., 2014). Fortunately, advances in metabolic engineering are also taking place concurrently, keeping the door open for a future where sustainable biobased production of previously petroleum-derived materials is possible. The experiments performed highlight various interventions for high-level synthesis in the yeast metabolic framework, which can be expanded on further. As an example, our recent work in the breakdown of cellulosic materials towards ethanol by *S. cerevisiae* can be implemented (Srikrishnan, 2012). A microbial system that can make valuable intermediates like TAL from biomass waste is well positioned for industrial applications. To that end, we are currently developing a cellulosome system for TAL production from cellulose.

The ease of use of computational modeling has significantly contributed to its rapid adoption as an approach for metabolic engineering (Likic et al., 2011). Efforts to standardize the language (SBML) in creating models and their associated matrices and make the models free and open-access have made it possible for user-friendly interfaces to let one build, view, and edit the networks (Likic et al., 2011). Additionally, the construction of open-access optimization tools and user-friendly procedures allows non-computational engineers and scientists to run simulations in a routine manner almost effortlessly (Tervo and Reed, 2014). With respect to available genome-scale models, the first eukaryotic model, *S. cerevisiae*, was constructed in 2003

harboring 708 genes (Feist et al., 2009). There are now 94 bacterial models and 39 eukaryotic models available, all of which were developed by different research groups with different motivations. This wealth of data and tools has already spawned a rapidly evolving field, and will change the way 'design-test-build' cycles are used in metabolic engineering. For example, the strain construct developed for improved TAL synthesis (see Chapter 2) was recently employed as the new model network in a simulation with ethanol as a carbon feed (relevant for *ADH2*-based expression). This was carried out in collaboration with the CBiRC Members in the Maranas group (Penn State) using their program OptKnock. Both the wildtype and engineered strain predicted disruption of ammonia / ammonium transport reactions during growth on ethanol. In a feasibility test, our results with knockout library strains suggested that this was indeed a valid approach for further TAL enhancement (Figure 5.2). Deletion of the *ATO2* or *MEP3* gene increased TAL titer by greater than 3-fold. Additionally, a new optimization algorithm has recently been developed by the Maranas group called k-OptForce, a variation of OptForce with integrated kinetic parameters for glycolytic reactions (Chowdhury et al., 2014; Ranganathan et al., 2010). OptForce was designed to predict up- and down-regulations as well as knockout interventions for ideal product formation. k-OptForce is an improved network with more input data, leading to more accurate predictions. This optimized algorithm was the first to suggest heterologous PDH overexpression for increased TAL, as well as the downregulation of native pathways like the PDH bypass (Chowdhury et al., 2014).

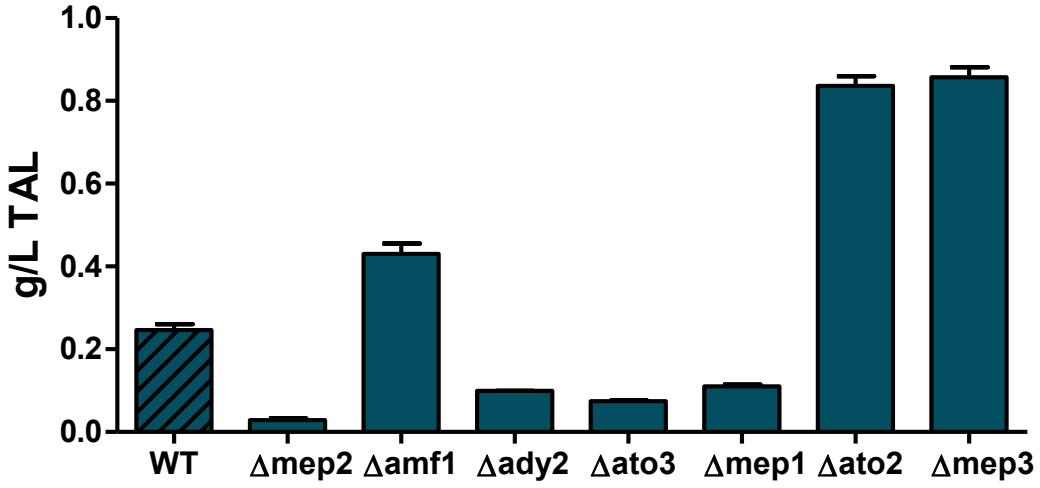


Figure 5.2 Triacetic acid lactone titers for strains of the Knockout Library collection. All strains denote single gene disruption strain as selected from OptKnock data (all genes implicated in ammonia or ammonium transport). Values are mean ± one standard deviation (n=6 independent experiments).

To successfully extend our work, the ability to suppress carbon flow through a competing pathway will be needed, where the most obvious strategy would be altering promoter control. Although vector sets are available with very tailored needs, having additional constructs available for different expression patterns would facilitate these types of downregulation / orverexpression studies. Our recent work explored a selection of different promoters, leveraging the ease of TAL biosynthesis for rapid evaluation that allows us to characterize relative expression levels under different culture conditions (Figure 5.3). A recent application of these promoters has been for downregulation of fatty acid biosynthesis (Figure 5.4), a major competitor for acetyl-CoA and malonyl-CoA. From these preliminary efforts, it is clear that there remain multiple promising metabolic engineering strategies to increase polyketide levels. Such strategies should lead to expedited development of *S. cerevisiae* as a platform organism for polyketide synthesis.

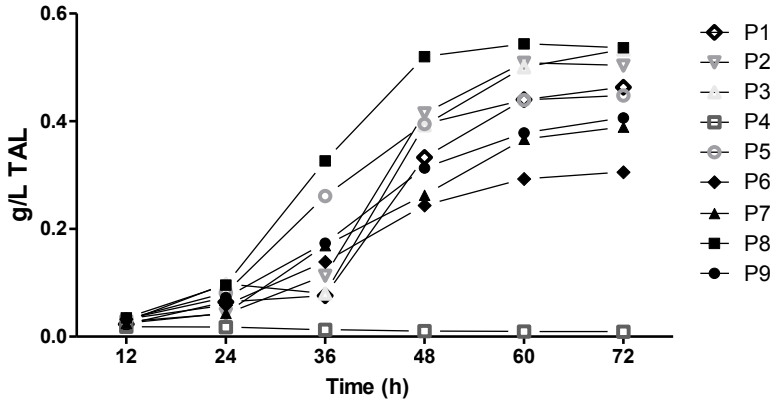


Figure 5.3 Triacetic acid lactone profiles for 9 promoters using high-copy expression plasmids in optimized SDC medium: (P1) *HXT7p* ^; (P2) *TEF1p* *; (P3) *TEF1p-mut4* *; (P4) *TEF1p-mut6* *; (P5) *TEF1p-mut7* *; (P6) *ADH1p* **; (P7) *ADH1Sp* **; (P8) *ADH1Mp* **; (P9) $HXT7_{391}p$ # (*Nevoigt et al., 2006; **Ruohonen et al.,1995; ^Thierfelder et al., 2010; #Fang et al., 2011).

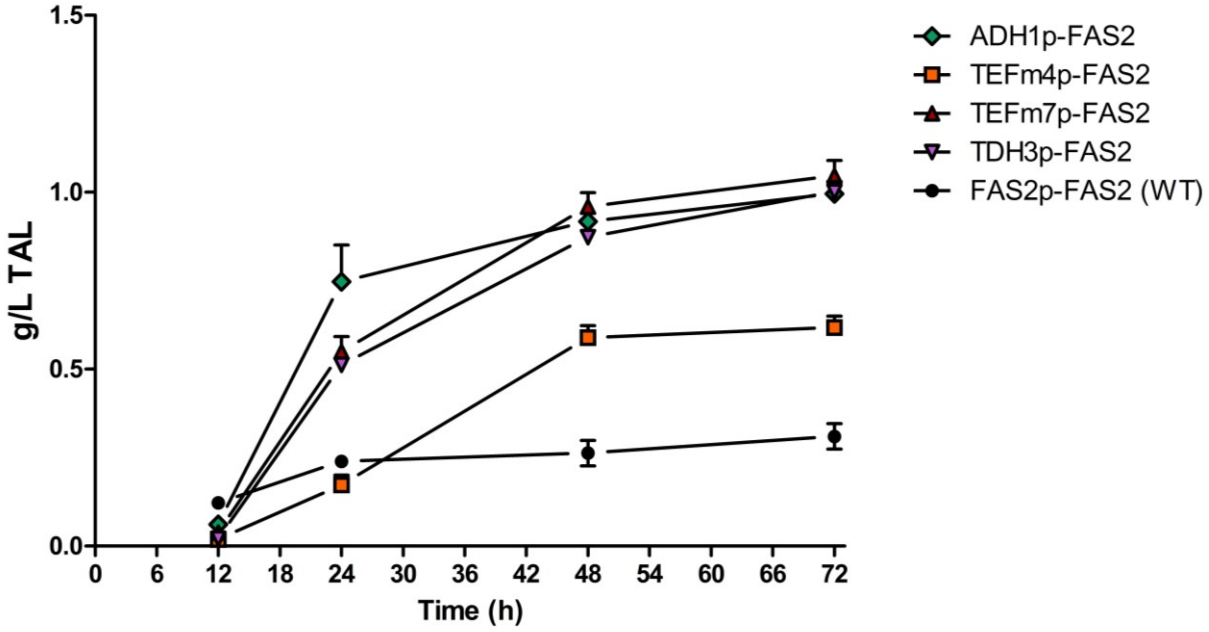


Figure 5.4 Triacetic acid lactone profiles after 48 h cultivation in 1% glucose YPD for four strains harboring non-native promoters (*ADH1p*, *TEF1p-mut4*, *TEF1p-mut7*, *TDH3p*) upstream of the *FAS2* gene (encoding one of the yeast FAS subunits). Values represent mean ± one standard deviation (n=6 independent experiments).

5.3 References

Carus, M., Dammer, L., Hermann, A., Essel, R., 2014. Proposals for a reform of the Renewable Energy Directive to a Renewable Energy and Materials Directive (REMD). Going to the next level: Integration of bio-based chemicals and materials in the incentive scheme. Nova paper 2014–05.

Chia, M., Schwartz, T.J., Shanks, B.H., Dumesic, J.A., 2012. Triacetic acid lactone as a potential biorenewable platform chemical. Green Chem. 14, 1850–1853. doi:10.1039/C2GC35343A

Chowdhury, A., Zomorrodi, A.R., Maranas, C.D., 2014. k-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design. PLoS Comput Biol 10, e1003487. doi:10.1371/journal.pcbi.1003487

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. doi:10.1002/yea.1824

Feist, A.M., Herrgård, M.J., Thiele, I., Reed, J.L., Palsson, B.Ø., 2009. Reconstruction of Biochemical Networks in Microbial Organisms. Nat Rev Microbiol 7, 129–143. doi:10.1038/nrmicro1949

Likic, V., McConville, M.J., Bacic, A., 2011. Systems Biology: The Next Frontier for Bioinformatics. Advances in Bioinformatics 2010, e268925. doi:10.1155/2010/268925

Nevoigt, E., Kohnke, J., Fischer, C.R., Alper, H., Stahl, U., Stephanopoulos, G., 2006. Engineering of Promoter Replacement Cassettes for Fine-Tuning of Gene Expression in Saccharomyces cerevisiae. Appl Environ Microbiol 72, 5266–5273. doi:10.1128/AEM.00530-06

Ranganathan, S., Suthers, P.F., Maranas, C.D., 2010. OptForce: An Optimization Procedure for Identifying All Genetic Manipulations Leading to Targeted Overproductions. PLoS Comput Biol 6. doi:10.1371/journal.pcbi.1000744

Ruohonen, L., Aalto, M.K., Keränen, S., 1995. Modifications to the ADH1 promoter of Saccharomyces cerevisiae for efficient production of heterologous proteins. Journal of Biotechnology 39, 193–203. doi:16/0168-1656(95)00024-K

Srikrishnan, S., 2012. Engineering of Enzymes and Yeast for the Hydrolysis of Biomass.

Tervo, C.J., Reed, J.L., 2014. Expanding metabolic engineering algorithms using feasible space and shadow price constraint modules. Metabolic Engineering Communications 1, 1–11. doi:10.1016/j.meteno.2014.06.001

Thierfelder, S., Ostermann, K., Göbel, A., Rödel, G., 2010. Vectors for Glucose-Dependent Protein Expression in Saccharomyces cerevisiae. Appl Biochem Biotechnol 163, 954–964. doi:10.1007/s12010-010-9099-5

Appendix A

Functional replacement of the *Saccharomyces cerevisiae* fatty acid synthase with a bacterial type II system allows flexible product profiles

This appendix has been modified from the accepted publication: Fernandez-Moya, R., Leber, C., Cardenas, J., Da Silva, N.A., 2015. Functional replacement of the Saccharomyces cerevisiae fatty acid synthase with a bacterial type II system allows flexible product profiles. Biotechnol. Bioeng. n/a–n/a. doi:10.1002/bit.25679

Author Contributions: See List of Author Contributions, pg. xv

Abstract

The native yeast type I fatty acid synthase (FAS) is a complex, rigid enzyme, and challenging to engineer for the production of medium- or short-chain fatty acids. Introduction of a type II FAS is a promising alternative as it allows expression control for each discrete enzyme and the addition of heterologous thioesterases. In this study, the native *Saccharomyces cerevisiae* FAS was functionally replaced by the *Escherichia coli* type II FAS (eFAS) system. The *E. coli acpS* + *acpP* (together), *fabB*, *fabD*, *fabG*, *fabH*, *fabI*, *fabZ* and *tesA* were expressed in individual *S. cerevisiae* strains, and enzyme activity was confirmed by *in vitro* activity assays. Eight genes were then integrated into the yeast genome, while *tesA* or an alternate thioesterase gene, *fatB* from *Ricinus communis* or TEII from *Rattus novergicus*, was expressed from a multi-copy plasmid. Native FAS activity was eliminated by knocking out the yeast *FAS2* gene. The strains expressing only the eFAS as *de novo* fatty acid source grew without fatty acid supplementation demonstrating that this type II FAS is able to functionally replace the native yeast FAS. The engineered strain expressing the *R. communis fatB* thioesterase increased total fatty acid titer 1.7-fold and shifted the fatty acid profile towards C_{14} production, increasing it from <1% in the native strain to more than 30% of total fatty acids, and reducing C_{18} production from 39% to less than 9%.

Key Words: *Saccharomyces cerevisiae*, *E. coli* fatty acid synthase, medium-chain fatty acids, thioesterase, biorenewable chemicals

The United States Department of Agriculture (USDA) has projected a 3-fold growth of the global biorenewable chemicals market from 2010 to 2025, reaching approximately $500 billion (USDA, Market Potential and Projections Through 2025, 2008). Attention has focused on biofuels, but near-term opportunities for biobased products are particularly promising for chemicals (Bozell and Petersen, 2010). Functional replacement of the platform chemicals produced from petroleum should include molecules with highly reduced carbon backbones in conjunction with highly reactive functional groups, and polyketide and fatty acid (FA) biosynthesis provide a promising solution route to generate a vast library of such compounds (Nikolau et al., 2008).

S. cerevisiae produces cytosolic *de novo* fatty acids using a complex and iterative fatty acid synthase (FAS). The yeast type I FAS is composed of Fas1 and Fas2 subunits in an □$_6$□$_6$ complex (Tehlivets et al., 2007), and the reactions are catalyzed by separate domains on the polypeptide chains. The primary products are C_{16} and C_{18} acyl-CoA esters, and the yeast FAS complexity and rigidity makes it a challenging enzyme to engineer to reduce this chain length (Leibundgut et al., 2008). As an alternative, our laboratory recently expressed, characterized and engineered the human type I FAS to produce short-chain fatty acids in *S. cerevisiae* (Leber and Da Silva, 2013). The higher flexibility of the hFAS structure allows replacement of the native thioesterase (TE) domain with one cleaving at the desired chain length. One disadvantage of Type I FAS systems is the iterative use of domains in a single enzyme complex, and the limited ability to intervene in the series of reaction steps. In contrast, bacterial, plant, and yeast mitochondrial fatty acid synthases are type II FASs, dissociated systems consisting of individual enzymes of the catalytic pathway codified by separate gene sequences (White et al., 2005). These systems allow the flexibility to engineer the pathway to produce alternate final FA–based products and to facilitate the metabolic optimization of the biosynthesis pathway. A good example of the flexibility to produce different end-products by the dissociated type II FAS is the production of methyl ketones in anengineered *E. coli* strain by the introduction of a 3-ketoacyl-ACP thioesterase (shmks2) from *Solanum habrochaites* (Park et al., 2012).

In this study, we replaced for the first time the native FAS system of *S. cerevisiae* with a type II FAS (from *E. coli*) and heterologous thioesterases, and then demonstrated the use of this system for the production of medium-chain fatty acids. Competition between the native and the heterologous FAS was removed by knocking-out the yeast *FAS2* gene, leaving the final strain with only the *E.coli* type II FAS as a source of *de novo* fatty acid synthesis. As a first step, we chose a minimum set of required enzymes, and then tested the expression and *in vitro* activity of the enzymes synthesized in yeast. Nine genes (*acpS, acpP, fabB, fabD, fabG, fabH*, *fabI, fabZ*, *tesA*) coding for nine proteins are required to produce fatty acids in yeast using this type II FAS. The genes, enzymes, and relevant reactions are shown in Table 3S-1, and have been extensively reviewed in the literature, e.g., White et al. (2005).

Each of the genes (except *acpP*) were introduced into an independent yeast strain on a 2μ-based plasmid (Table 1). AcpS activates apo-ACP (encoded by *acpP*), transferring to it the 4'-phosphopantetheine from coenzyme A, resulting in holo-ACP (Flugel et al., 2000). Active ACP was obtained using *S. cerevisiae* strain BY1G (Table 1), with one-copy of the *acpS* gene integrated into the genome, and carrying a plasmid harboring *acpP*. Yeast strains expressing the different type II FAS proteins were cultivated in selective SDC-A medium, the cells were harvested, and the proteins were extracted and purified using Ni-NTA columns. Western blots of the purified proteins were used to verify their expression in yeast (Figure S-1), and two *in vitro* activity assays were performed to confirm their activity (Figure 1). A NADPH colorimetric assay (Bays et al., 2009) was used to directly verify activity of FabG, and thus the activity of the upstream enzymes AcpP, AcpS, FabD, FabH (Figure S-2). The assay results (Figure 1A) demonstrate that these five enzymes were actively expressed in yeast. To determine if all of the type II FAS proteins are active and can work together when expressed in *S. cerevisiae*, an *in vitro* acrylodan-labeled intestinal fatty acid binding protein (ADIFAB) assay was used. The purified enzymes were combined with acetyl-CoA, malonyl-CoA and NADPH, and the concentrations of free fatty acids (FFAs) were measured (Figure 1B). The negative controls, lacking either TesA (thioesterase activity) or holo-ACP (AcpP and AcpS) (essential prosthetic

group), had a background concentration of FFA, likely due to carryover from the protein elutes and non-specific binding of the ADIFAB protein. However, the FFA level for the mixture with all enzymes was higher than the controls ($p<0.05$), demonstrating that the *E. coli* type II enzymes expressed in yeast are able to produce FFAs *in vitro* from malonyl-CoA and acetyl-CoA. The trends shown in Figure 1B were seen in each of the three independent experiments.

After confirming activity of the enzymes produced in *S. cerevisiae*, we introduced all nine genes into one strain. Eight genes (*acpS, acpP, fabB, fabD, fabG, fabH*, *fabI, fabZ*) were integrated into the BY4741 genome and the thioesterase gene was carried on a 2□-based plasmid. The rationale for the plasmid-based TE was two-fold: (1) the ability to quickly introduce different TEs and (2) the observed overproduction of free fatty acids in *E. coli* in the presence of high levels of cytosolic thioesterase activity (Lu et al., 2008), as this reaction is the bottleneck for FA production. All genes were integrated via double crossover to ensure stability (Da Silva and Srikrishnan, 2012). Single copies of *fabG* and *fabD* were sequentially integrated; the remaining six genes (*acpP*, *acpS*, *fabB*, *fabH*, *fabI*, and *fabZ*) were integrated in pairs as a single PCR product (Supplementary Material, Figure S-3). For the latter integrations, different promoters and terminators were used to avoid repeat sequences and thus avoid loss by homologous recombination. The nine genes were under either the *PGK1* or the *TEF1* promoter.

Three strains were used for our expression studies: BY4741 (parent strain), BY6G (all genes except *fabB* and *fabI* integrated), and BY8G (all eight genes integrated) (Table I). BY4741 and BY6G are control strains that do not carry the full type II FAS. Three different thioesterase genes (*tesA*, RC, and TEII) were cloned into pXP219 vectors (Table I). TesA is natively located in the periplasm of *E. coli* and shows substrate preference for C_{16} and C_{18} acyl intermediates (Cho and Cronan, 1993). TEII (thioesterase II), from *Rattus novergicus* mammary gland, shifted the *in vitro* FA production profile of the rat FAS from long chain fatty acids (C_{16} and C_{18}) to short-chain fatty acids (C_8, C_{10}, C_{12}) (Libertini and Smith, 1978). In our earlier studies, this TE proved effective for producing C_8 and C_{10} when overexpressed with the human FAS in yeast

(Leber and Da Silva, 2013). RC is a palmitoyl-ACP thioesterase from *Ricinus communis* (*fatB*); when overexpressed in *E.coli*, the production of myristic acid (C_{14}) increased more than 4-fold, accounting for 40% of the total fatty acids and increasing the total free fatty acid synthesized (Zhang et al., 2011).

Fatty acid production was first characterized in strain BY8G carrying pXP219-tesA, pXP219-TEII, or pXP219-RC and the two control strains. Growth rates were similar for all strains, indicating minimal growth inhibition due to the expression of the heterologous type II FAS and the overexpression of the thioesterases. No difference was seen in the distribution of fatty acid chain lengths or total fatty acid produced by the strains expressing TEII or RC relative to the controls (Figure S-4). However, strain BY8G+pXP219-tesA produced 201 mg/L total fatty acids, more than twice that for the control BY4741+pXP219 (86 mg/L). Overexpression of enzymes with acyl-CoA thioesterase activity has previously been shown to increase total FFA production in *S. cerevisiae* (Runguphan and Keasling, 2013). Therefore, the observed increase in fatty acids could be due to overexpression of the thioesterase rather than an increase due to the *E. coli* FAS system.

The lack of change in fatty acid distribution suggests that the native yeast FAS is dominant, limiting the impact of the heterologous eFAS system. Therefore, the native yeast FAS activity was eliminated to remove competition for substrates, cofactors, and reducing agents. We knocked out the *FAS2* gene in strains BY4741 and BY8G to create BY4741ΔFAS2 and BY8GΔFAS2. Deletion of the *FAS2* gene prevents *de novo* fatty acid production and is lethal to the yeast cells when fatty acids are not supplemented. The two *FAS2* deletion strains were plated on SDC-A supplemented with myristic acid to allow growth. Control strain BY4741ΔFAS2 was transformed with the empty vector pXP219, and strain BY8GΔFAS2 was transformed with vectors pXP219-tesA, pXP219-TEII, or pXP219-RC, creating three strains expressing the *E. coli* FAS system and different heterologous thioesterases.

Single colonies of the four strains were plated on SDC-A without FA supplementation to test for growth, and were evaluated after 6 and 25 days (Figure S-5). The control BY4741ΔFAS2+pXP219 did not

grow, as this strain cannot produce fatty acids. Similarly, the strain with the integrated *E. coli* FAS (BY8GΔFAS2) but no thioesterase was unable to grow (data not shown). However, the three strains with the integrated *E. coli* FAS and expressing the heterologous thioesterases were viable. With the RC and TEII thioesterases, colonies grew to significant size; with TesA, the cell grew poorly generating only small colonies. It is noteworthy that the type II *E. coli* FAS system and heterologous TEs were capable of functionally replacing the native yeast FAS, producing sufficient fatty acids for growth. The RC acyl-ACP thioesterase, FatB from *R. communis,* has been previously expressed in *E. coli*, showing the highest production of FFAs in this bacteria compared to other TEs tested (Zhang et al., 2011). This result agrees with the growth of the *FAS2* knockout strains observed in Figure S-5. While all three strains grew on plates, only BY8GΔFAS2+pXP219-RC grew well in liquid SDC-A and complex YPD media. We thus chose to focus on this strain for evaluation of the fatty acid profile and total fatty acid levels.

Control BY4741+pXP219 and strain BY8GΔFAS2+pXP219-RC were cultivated in selective SCD-A medium, harvested at the same stage of growth (early stationary phase: 20h and 45h, respectively (Figure S-6)), and FA levels were measured via GC-MS. BY4741+pXP219 uses the native *S. cerevisiae* FAS to produce *de novo* fatty acids, while BY8GΔFAS2+pXP219-RC has no yeast FAS and uses only the *E. coli* FAS with the *R. communis* FatB thioesterase. Total fatty acids were 1.7-fold higher for the heterologous eFAS system (160 mg/L) relative to the control strain with the native yeast FAS (93 mg/L) (Figure 2A), and 3.2-fold higher if normalized to OD. High concentrations of C_{14} fatty acids (33% of total FA) were also found in the type II eFAS strain (Figure 2B), and C_{18} levels were significantly reduced (8% of total FA). Most of the C_{14} produced was monounsaturated (20% of total FA). In contrast, total C_{14} accounted for only 0.2% of total FAs (nearly undetectable) in the control strain using the native yeast FAS.

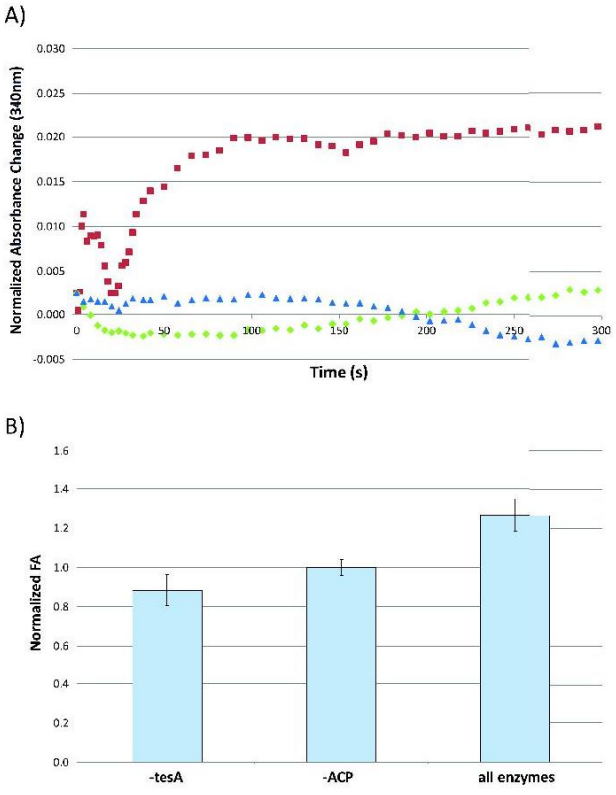


Figure 1. **A)** Confirmation of the activity of the AcpP, AcpS, FabD, FabH, and FabG proteins expressed in *S. cerevisiae*. An *in vitro* NADPH absorbance assay demonstrated the successful expression and activity of five of the nine genes needed to express the type II FAS in yeast (Figure S-2). Complete mix (squares) consisted of buffered ACP, FabD, FabH, FabG, acetyl-CoA, NADPH, and malonyl-CoA. Negative controls consisted of a complete mix minus FabG (triangles), and complete mix minus ACP (diamond). Three independent experiments showed similar profiles. **B)** Confirmation of type II FAS activity by *in vitro* fatty acid synthesis. An *in vitro* ADIFAB assay demonstrated fatty acid synthesis when ACP, FabB, FabD, FabG, FabH, FabI, FabZ, and TesA were mixed with acetyl-CoA, malonyl-CoA, and NADPH. Negative controls were the reaction mix lacking *E. coli* thioesterase I TesA and the reaction mix lacking the acyl carrier protein ACP. Three independent experiments (with two independent replicates per experiment) were performed. Values were normalized to the -ACP control to allow comparisons across experiments. Results are expressed as mean ± SEM (n=6). The negative controls were statistically lower ($p<0.05$).

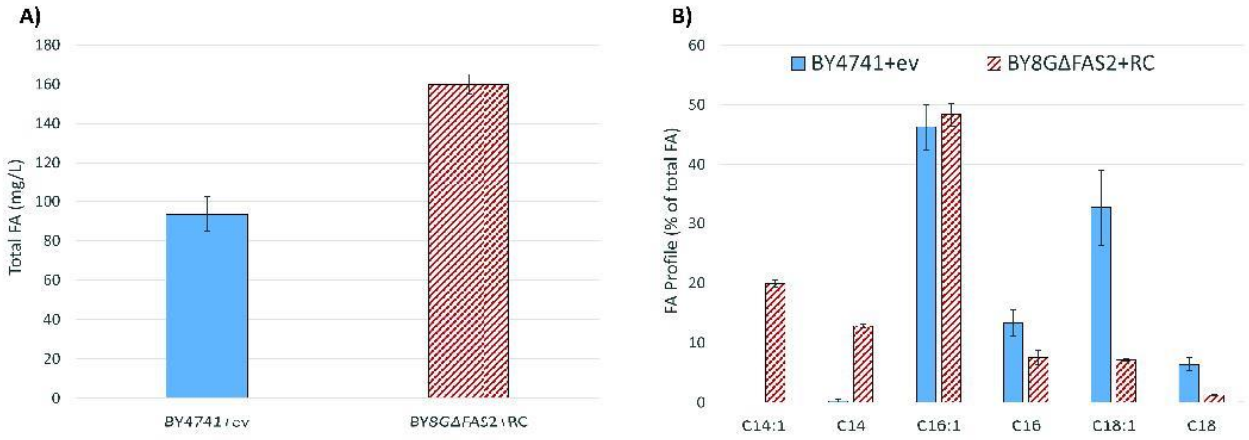


Figure 2. Fatty acid levels and chain length profile. **A)** Total fatty acids produced by strain BY8GΔFAS2 carrying pXP219-RC (blue solid bar), and control strain BY4741 carrying an empty vector (ev) pXP219 (red hatched bar). **B)** Profile of total fatty acids produced. Cells were harvested at early stationary phase: 20h for control and 45h for BY8GΔFAS2+pXP219-RC (Figure S-6). No significant changes in the extracellular FA levels were measured for the control and FA production strains. Results are expressed as mean ± SEM (n=3). The differences in total fatty acid, C14:1, C14, C18:1, and C18 levels were statistically significant ($p<0.01$).

The fatty acid profile for strain BY8GΔFAS2+pXP219-RC thus showed both higher overall FA levels and significant production of C_{14} (and lower production of C_{18}) relative to the native *S. cerevisiae*. Fatty acid chain length is determined mainly by the acyl-ACP thioesterase in a type II FAS system. Interestingly, our results are similar to those observed in *E. coli*; Zhang et al., (2011) found that overexpression of *R. communis fatB* in *E. coli* increased total fatty acid levels and altered the normal FA distribution, increasing C_{14} production and decreasing C_{18}, without removing the native thioesterase. The FA profile in yeast is also dependent on elongases, especially Elo1, which can actively elongate C_{12}-C_{14} acyl-CoA primers to C_{16} and C_{18} (Dittrich et al., 1998).

The final OD of strain BY8GΔFAS2+pXP219-RC was half that of the native BY4741 (Table 2, Figure S-6). In contrast, BY8G+pXP219-RC (with the intact native FAS) had a growth profile similar to the native

strain. Therefore, deletion of the native *FAS2* gene and reliance on only the heterologous eFAS pathway reduced the growth of the cells. This also decreased the FA productivity (mg/L h); although titers were substantially higher for BY8GΔFAS2+pXP219-RC, a longer growth period was needed to reach stationary phase. The heterologous pathway and the high percentage of C_{14} relative to C_{18} may be contributing factors. Better growth and productivity should be possible if the native yeast FAS is down-regulated instead of eliminated, supplementing sufficient native FAS function for healthy growth without excessive competition with the heterologous FAS system.

These results demonstrate the promise of the heterologous *E. coli* FAS system for engineering yeast to produce alternate fatty acid products. The flexibility offered by this dissociated system allowed us to introduce heterologous thioesterases with the ability to cleave the ACP prosthetic group of fatty acids at different chain lengths. This system can also be used for the introduction of novel thioesterases cleaving the ACP from FA intermediates, potentially producing a large number of different final products with new chemical functionalities (Kim et al., 2015). In addition, the discrete type II FAS system allows expression to be optimized or even eliminated for each independent catalytic step.

Table 2. Comparison of growth and FA production for the final engineered strain, BY8GΔFAS2+RC, with the control, BY4741+pXP219 at stationary phase (45 h for BY8GΔFAS2+RC and 20h for BY4741) in SDC-A medium. Results are expressed as mean ± SEM from triplicates.

	Final OD_{600}	C_{14} concentration (mg/L)	Total FA concentration (mg/L)	Total FA productivity mg/(L·h)
BY4741+ev	12.2±0.2	0.20±0.02	93.5±8.7	4.7±0.4
BY8GΔFAS2+RC	6.5±0.1	52.3±1.4	160±5	3.6±0.1

Materials and Methods

Vector and strain construction

The *E. coli type II* FAS gene sequences were amplified from *E. coli* XL1-Blue (Stratagene, Santa Clara, CA) with a C- or N-terminal 6xHis-tag and inserted into the pXP219 vector (Table 1). The thioesterase gene sequence from *R. communis* (RC) was amplified from pXZ18 (Zhang et al., 2011), and the thioesterase gene sequence from *R. novergicus* (TEII) was amplified from pJLA502-TEII (Naggert et al., 1991). Both sequences were independently inserted into the pXP219 vector. The acpS, *fabI*, and *fabZ* gene sequences were excised from pXP219-acpS, pXP219-fabI, and pXP219-fabZ, respectively, and inserted into the pRF320 vector.

S. cerevisiae strain BY4741 (Open Biosystems, Huntsville, AL) was the base strain. Single and double gene integrations were performed using two-piece (gene cassette, selection marker) transformation as previously described (Fang et al., 2011). The *FAS2* gene in strains BY4741 and BY8G (Table 1) was deleted by replacement with the *MET15* marker (Supplementary Material, Table S-6) and strains were selected on methionine-deficient plates supplemented with 2 mM myristic acid. After transformation of pXP219 in BY4741ΔFAS2, and pXP219-tesA, pXP219-RC and pXP219-TEII in BY8GΔFAS2, cells were selected on SDC-A + 2 mM myristic acid plates. Colonies on the previous plates were then plated on SDC-A without fatty acid supplementation. Colonies were restreaked 2 times onto new SDC-A plates and healthy cells were selected.

The strains and vectors are described in Table I, and construction details can be found in the Supplementary Material.

Media and cultivation

Luria-Bertani (LB) medium containing 100 µg/mL ampicillin was used for *E. coli* cultivation. *S. cerevisiae* strains were cultivated in complex YPD, selective SDC-A, or SD minimal medium containing 2%

Table 1. List of plasmids and *S. cerevisiae* strains used in this study

Plasmids	Description	Source
pXP219	2μ vector, *PGK1* promoter, *CYC1* terminator, *URA3* selectable marker	Fang et al., 2011
pXP219-acpP	pXP219 harboring *acpP* with 6xHis-tag on C-terminus	This study
pXP219-acpS	pXP219 harboring *acpS* with 6xHis-tag on C-terminus	This study
pXP219-fabB	pXP219 harboring *fabB* with 6xHis-tag on N-terminus	This study
pXP219-fabD	pXP219 harboring *fabD* with 6xHis-tag on N-terminus	This study
pXP219-fabG	pXP219 harboring *fabG* with 6xHis-tag on N-terminus	This study
pXP219-fabH	pXP219 harboring *fabH* with 6xHis-tag on N-terminus	This study
pXP219-fabI	pXP219 harboring *fabI* with 6xHis-tag on N-terminus	This study
pXP219-fabZ	pXP219 harboring *fabZ* with 6xHis-tag on N-terminus	This study
pXP219-tesA	pXP219 harboring *tesA* with 6xHis-tag on C-terminus	This study
pXP219-RC	pXP219 harboring RC with 6xHis-tag on C-terminus	This study
pXP219-TEII	pXP219 harboring TEII with 6xHis-tag on C-terminus	This study
pJF702	*CEN/ARS* vector, *ADH2* promoter, *ADH2* terminator, *LEU2* selectable marker	This study
pXP320	*CEN/ARS* vector, *TEF1* promoter, *CYC1* terminator, *HIS3* selectable marker	Fang et al., 2011
pRF320	*CEN/ARS* vector, *TEF1* promoter, *ADH2* terminator, *LEU2* selectable marker	This study
pRF320-acpS	pRF320 harboring *acpS* with 6xHis-tag on C-terminus	This study
pRF320-fabI	pRF320 harboring *fabI* with 6xHis-tag on N-terminus	This study
pRF320-fabZ	pRF320 harboring *fabZ* with 6xHis-tag on N-terminus	This study

Strains	Description	Source
BY4741	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Open Biosystems
BY1G	BY4741 *ura3*::P_{PGK1}-*acpS*-T_{CYC1}	This study
BY6G	BY4741 *met15*::P_{PGK1}-*fabG*-T_{CYC1} *ura3*::P_{PGK1}-*fabD*-T_{CYC1} *leu2*::P_{PGK1}-*fabH*-T_{CYC1}-P_{TEF1}-*acpS*-T_{ADH2} *ty1-5*::P_{PGK1}-*acpP*-T_{CYC1}-P_{TEF1}-*fabZ*-T_{ADH2}	This study
BY8G	BY6G *ty1-1*::P_{PGK1}-*fabB*-T_{CYC1}-P_{TEF1}-*fabI*-T_{ADH2}	This study
BY4741ΔFAS2	BY4741 fas2Δ0	This study
BY8GΔFAS2	BY8G fas2Δ0	This study

dextrose (see Supplementary Material). For *FAS2* knockout strains without the integrated eFAS, the medium was supplemented with 2 mM myristic acid and tween 40 (0.5%). *S. cerevisiae* strains were cultivated at 30˚C in an air shaker (New Brunswick Scientific, Enfield, CT) at 250 rpm. Cells were inoculated from -80˚C stock and grown overnight in 5 mL medium in 16 mm x 125 mm culture tubes for inoculation to an OD of 0.1. Cells for fatty acid measurement studies were grown in shake flasks containing 50, 200 or 250 mL medium.

Analytical Methods

The protein purification, NADPH and ADIFAB assay protocols were modified from Leber and Da Silva (2013). Activity of AcpS, AcpS, FabD, FabH, and FabG was determined via a NADPH assay using equimolar amounts of purified *E. coli* FAS enzymes in the NADPH reaction buffer. For the ADIFAB assay, purified type II FAS enzymes (5-15 µM final conc.) were combined with acetyl-CoA (50 µM final conc.), malonyl-CoA (150 µM final conc.), and NADPH (150 µM final conc.) to a final volume of 500 µL. The reaction mixture was placed on a microfuge tube rotator at 37˚C for 1 h. 500 µL of 0.2 µM ADIFAB suspended in measuring buffer was then added and the 1 mL solution was incubated at 37˚C for 20 min. The mixture was transferred to a quartz cuvette and a Bowman Series 2 luminescence spectrometer was used to take readings at wavelengths of 432 and 505 nm. FA extraction and quantification via GC-MS followed that in Leber and Da Silva (2013).

Statistical Analysis

Three independent enzyme isolation/ADIFAB experiments were performed, with two independent assays per experiment (n=6). For the fatty acid analysis, three independent experiments were performed (n=3). A one-way analysis of variance (ANOVA) was used, followed by Dunnett's multiple comparison test. Values were considered statistically significant when p-values were <0.05. Data are reported as mean $\pm$ standard error of the mean (SEM), unless otherwise stated.

Acknowledgements

This research was supported by NSF (grant no. EEC-0813570) through the Engineering Research Center CBiRC (Center for Biorenewable Chemicals). The authors would like to thank Dr. Stuart Smith (Oakland Children's Hospital) for providing thioesterase TEII and Prof. Ka-Yiu San (Rice University) for providing thioesterase RC.

References

Bays, N.W., Hill, A.D., and Kariv, I. (2009). A Simplified Scintillation Proximity Assay for Fatty Acid Synthase Activity: Development and Comparison with Other FAS Activity Assays. J. Biomol. Screen. *14*, 636–642.

Bozell, J.J., and Petersen, G.R. (2010). Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. Green Chem. *12*, 539–554.

Cho, H., and Cronan, J.E. (1993). Escherichia coli thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme. J. Biol. Chem. *268*, 9238–9245.

Dittrich, F., Zajonc, D., Hühne, K., Hoja, U., Ekici, A., Greiner, E., Klein, H., Hofmann, J., Bessoule, J.-J., Sperling, P., et al. (1998). Fatty acid elongation in yeast. Eur. J. Biochem. *252*, 477–485.

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., and Sandmeyer, S. (2011). A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. Yeast *28*, 123–136.

Flugel, R.S., Hwangbo, Y., Lambalot, R.H., Cronan, J.E., and Walsh, C.T. (2000). Holo-(Acyl Carrier Protein) Synthase and Phosphopantetheinyl Transfer in Escherichia coli. J. Biol. Chem. *275*, 959–968.

Kim, S., Clomburg, J.M., and Gonzalez, R. (2015). Synthesis of medium-chain length (C6–C10) fuels and chemicals via β-oxidation reversal in Escherichia coli. J. Ind. Microbiol. Biotechnol. *42*, 465–475.

Leber, C., and Da Silva, N. a (2013). Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids. Biotechnol. Bioeng. *9999*, 1–12.

Leibundgut, M., Maier, T., Jenni, S., and Ban, N. (2008). The multienzyme architecture of eukaryotic fatty acid synthases. Curr. Opin. Struct. Biol. *18*, 714–725.

Libertini, L.J., and Smith, S. (1978). Purification and properties of a thioesterase from lactating rat mammary gland which modifies the product specificity of fatty acid synthetase. J. Biol. Chem. *253*, 1393–1401.

Lu, X., Vora, H., and Khosla, C. (2008). Overproduction of free fatty acids in E. coli: implications for biodiesel production. Metab. Eng. *10*, 333–339.

Naggert, J., Witkowski, A., Wessa, B., and Smith, S. (1991). Expression in Escherichia coli, purification and characterization of two mammalian thioesterases involved in fatty acid synthesis. Biochem. J. *273*, 787–790.

Nikolau, B.J., Perera, M.A.D.N., Brachova, L., and Shanks, B. (2008). Platform biochemicals for a biorenewable chemical industry. Plant J. Cell Mol. Biol. *54*, 536–545.

Park, J., Rodríguez-Moyá, M., Li, M., Pichersky, E., San, K.-Y., and Gonzalez, R. (2012). Synthesis of methyl ketones by metabolically engineered Escherichia coli. J. Ind. Microbiol. Biotechnol. *39*, 1703–1712.

Runguphan, W., and Keasling, J.D. (2013). Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. Metab. Eng.

Da Silva, N. a, and Srikrishnan, S. (2012). Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. FEMS Yeast Res. *12*, 197–214.

Tehlivets, O., Scheuringer, K., and Kohlwein, S.D. (2007). Fatty acid synthesis and elongation in yeast. Biochim. Biophys. Acta *1771*, 255–270.

USDA, United States Department of Agriculture (2008). U. S. Biobased Products. Market Potential and Projections Through 2025 (USDA, United States Department of Agriculture).

White, S.W., Zheng, J., Zhang, Y.-M., and Rock CO (2005). The structural biology of type II fatty acid biosynthesis. Annu. Rev. Biochem. *74*, 791–831.

Zhang, X., Li, M., Agrawal, A., and San, K.-Y. (2011). Efficient free fatty acid production in Escherichia coli using plant acyl-ACP thioesterases. Metab. Eng. *13*, 713–722.

APPENDIX B

Supplementary Material to:

Functional replacement of the *Saccharomyces cerevisiae* fatty acid synthase with a bacterial type II system allows flexible product profiles (Appendix A)

Table S-1. List of genes, proteins, and reactions.

Gene	Protein	Reaction	Ref.
acpP	Acyl carrier protein (ACP)		White et al., 2005
acpS	ACP synthase	apo-ACP + 4'-Phosphopantetheine Group → ACPSH	White et al., 2005
fabB*	β-ketoacyl-ACP synthase I	(n)acyl-ACP + malonyl-ACP → (n+2)β-ketoacyl-ACP acetoacetyl-ACP + malonyl-ACP → β-ketobutyryl-ACP + CO_2 + ACPSH	White et al., 2005
fabD	Malonyl-CoA:ACP transacylase	malonyl-CoA + ACPSH → malonyl-ACP	White et al., 2005
fabG	β-ketoacyl-ACP reductase	β-ketoacyl-ACP + NADPH + H^+ → β-hydroxyacyl-ACP + $NADP^+$	White et al., 2005
fabH	β-ketoacyl-ACP synthase III	malonyl-ACP + acetyl-CoA → acetoacetyl-ACP	White et al., 2005
fabI	Enoyl-ACP reductase	enoyl-ACP + NADH+H^+ → acyl-ACP + NAD^+	White et al., 2005
fabZ†	β-hydroxyacyl-ACP dehydratase	β-hydroxyacyl-ACP → enoyl-ACP + H_2O	White et al., 2005
tesA	Thioesterase I	acyl-ACP → free fatty acid + ACPSH	Cho and Cronan, 1993

* Two *E. coli* enzymes FabB (β-ketoacyl-ACP synthase I) and FabF (β-ketoacyl-ACP synthase II) catalyze the condensation of malonyl-CoA with the fatty acyl-ACP growing chain. We expressed only β-ketoacyl-ACP synthase (FabB), since expression of FabF in active form has proven challenging (Magnuson et al., 1995).

† There are two genes in *E. coli* that encode β-hydroxyacyl-ACP dehydrases, FabA and FabZ. The broad substrate specificity of FabZ and the low activity of FabA for β-hydroxybutyryl-ACP (Heath and Rock, 1996) made FabZ a better choice for our work.

Vector and strain construction

E. coli strain XL1-Blue (Stratagene, Santa Clara, CA) was used for plasmid preparation and storage. *S. cerevisiae* host strain BY4741 (Open Biosystems, Huntsville, AL) was used for *in vitro* and *in vivo* studies. Standard molecular biology procedures were carried out as described in Sambrook and Russell (2001). Restriction enzymes, T4 DNA ligase, Taq DNA polymerase, and deoxynucleotides were purchased from New England Biolabs (Ipswich, MA). KOD Hot-start DNA polymerase was obtained from Novagen (San Diego, CA). Oligonucleotide primers were purchased from IDT DNA (San Diego, CA). All sequences of gene fragments amplified by PCR were verified by DNA sequence analysis (GeneWiz, South Plainfield, NJ). *E. coli* mini-prep DNA was prepared using a Spin Miniprep Kit (Qiagen, Germantown, MD), and plasmid transformation of *E. coli* competent cells was done using a standard heat shock method (Sambrook and Russell, 2001). Plasmid and integrative transformations in *S. cerevisiae* were performed using a high efficiency LiAc method using DMSO (Hoskins, 2000). Isolation of total genomic yeast DNA for integration and checking was performed as described in Sambrook and Russell (2001).

acpP, *acpS*, *fabB*, *fabD*, *fabG*, *fabH*, *fabI*, *fabZ*, and *tesA* were amplified from *E. coli* XL1-Blue using the primers shown in Table S-2. The placement of the 6x-histidine tags was based on prior reports in the literature. If this information was not found (i.e., for *acpP*, *acpS*, and *tesA*), the C-terminus was chosen. Thioesterase RC was amplified from pXZ18 and TEII was amplified from pJLA502-TEII using the primers shown in Table S-2. PCR fragments were digested with restriction enzymes XhoI and SpeI and isolated on a 1% agar gel. DNA was extracted using a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Backbone DNA (pXP219) and genes were ligated at a 1:3 molar ratio for 20 minutes at room temperature using Thermo Scientific Fermentas Rapid DNA Ligation Kit (Thermo Fisher, San Jose, CA) to generate pXP219-acpP, pXP219-acpS, pXP219-fabB, pXP219-fabD, pXP219-fabG, pXP219-fabH, pXP219-fabI, pXP219-fabZ, pXP219-tesA, pXP219-RC, and pXP219-TEII.

Table S-2. Primers used to amplify, 6xHis-tag, and clone FAS enzymes into pXP219

Gene	Direction	Primer name	Primer sequence, 5' → 3'
acpP	Forward	acpP-for-speI	GACCTCCC-ACTAGT-AAAAAA-ATGAGCACTATCGAAGAACG
	Reverse	acpP-rev-histag	TGGTAATC-CTCGAG-TTA-ATGATGATGATGATGATG-CGCCTGGTGGCCGTTGA
acpS	Forward	acpS-for-speI	GACCTCCC-ACTAGT-AAAAAA-ATGGCAATATTAGGTTTAGG
	Reverse	acpS-rev-histag	TGGTAATC-CTCGAG-TTA-ATGATGATGATGATGATG-ACTTTCAATAATTACCG
fabB	Forward	fabB-for-speI-histag	GACCTCCC-ACTAGT-AAAAAA-ATG-CATCATCATCATCATCAT-AAACGTGCAGTGATTAC
	Reverse	fabB-rev	TGGTAATC-CTCGAG- TTAATCTTTCAGCTTGCGCA
fabD	Forward	fabD-for-pmeI-histag	GACCTCCC-GTTTAAAC-AAAAAA-ATG-CATCATCATCATCATCAT-ACGCAATTTGCATTTGT
	Reverse	fabD-rev	TGGTAATC-CGGTCCG-TTAAAGCTCGAGCGCCGCTG
fabG	Forward	fabG-for-speI-histag	GACCTCCC-ACTAGT-AAAAAA-ATG-CATCATCATCATCATCAT-AATTTTGAAGGAAAAAT
	Reverse	fabG-rev	TGGTAATC-CTCGAG- TCAGACCATGTACATCCCGC
fabH	Forward	fabH-for-speI-histag	GACCTCCC-ACTAGT-AAAAAA-ATG-CATCATCATCATCATCAT-TATACGAAGATTATTGG
	Reverse	fabH-rev	TGGTAATC-CTCGAG- CTAGAAACGAACCAGCGCGG
fabI	Forward	fabI-for-speI-histag	GACCTCCC-ACTAGT-AAAAAA-ATG-CATCATCATCATCATCAT-GGTTTTCTTTCCGGTAA
	Reverse	fabI-rev	TGGTAATC-CTCGAG- TTATTTCAGTTCGAGTTCGT
fabZ	Forward	fabZ-for-speI-histag	GACCTCCC-ACTAGT-AAAAAA-ATG-CATCATCATCATCATCAT-ACTACTAACACTCATAC
	Reverse	fabZ-rev	TGGTAATC-CTCGAG- TCAGGCCTCCCGGCTACGAG
tesA	Forward	tesA-for-speI	GACCTCCC-ACTAGT-AAAAAA-ATGATGAACTTCAACAATGT
	Reverse	tesA-rev-histag	TGGTAATC-CTCGAG-TTA-ATGATGATGATGATGATG-TGAGTCATGATTTACTA
RC	Forward	RC-for-speI	TCCC-ACTAGT-AAAAAA-ATGGTAGCAACAGCAGCGGCAGCA
	Reverse	RC-rev-histag	GGGACTCGAGTTAATGATGATGATGATGATGGGCGCTTTCAACCGGAATTTG
TEII	Forward	TEII-for-speI	GACCTCCC-ACTAGT-AAAAAA-ATGGAGACAGCAGTCAATGCTAAGAGTCCCAGGAATGAAAAGGTTTTGAACTGT
	Reverse	TEII-rev-histag	TGGTAATC-CTCGAG-TCA-ATGATGATGATGATGATG-AGTGAGTGACGAGAGTTCCAA

acpS under the control of P_{PGK1} was integrated at one copy into the genome of strain BY4741, creating strain BY-aS. PCR was used to generate fragment P_{PGK1}-*ACPS*-T_{CYC1} from vector pXP219-acpS and fragment loxP-*LEU2*-loxP from plasmid pXP222 using the primers shown in Table S-3. The P_{PGK1}-*ACPS*-T_{CYC1} fragment had 50 bp homologous sequences on the 3' end to fragment loxP-*LEU2*-loxP on the 5' end. The

flanking 5' and 3' ends of the adjoined segment had 50 bp overlapping sequences up- and down-stream of the target *URA3* locus. *acpS* integration into the *URA3* locus was confirmed by PCR analysis using primers that annealed upstream and downstream of the chromosomal site (Table S-7).

Table S-3. Primers used to integrate *acpS* into the *URA3* locus to create strain BY-aS**.**

Gene	Locus	Marker	Direction	Primer Sequence
acpS	***URA3***	***LEU2***	Forward1	GTTTTGACCATCAAAGAAGGTTAATGTGGCTGTGGTTTCAGGGTCCATAA GGCATTTGCAAGAATTACTCGTG
			Reverse1	CCCGGGGATCCTCTAGAGTCGACCGGCCGCAAATTAAAGCCTTCGAGCG
			Forward2	CGCTCGAAGGCTTTAATTTGCGGCCggtcgactctagaggatcCCCGGG
			Reverse2	GCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATTTCGTG TCGAATTCGAGCTCGGTACCCGGG

The *CYC1* terminator on pXP320 was replaced by the *ADH2* terminator. Primers For-ADH2t and Rev-ADH2t (Table S-4) were used to amplify the *ADH2* terminator from pJC702. The PCR product and pXP320 were digested with XbaI and XhoI and isolated on a 1% agar gel. DNA was extracted using Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Backbone DNA and terminator were ligated at a 1:3 molar ratio for 20 min at room temperature using a Thermo Scientific Fermentas Rapid DNA Ligation Kit (Thermo Fisher, San Jose, CA), generating plasmid pRF320. Plasmids pRF320-acpS, pRF320-fabI, and pRF320-fabZ were constructed using the primers shown in Table S-1 and the same procedure as for pXP219-versions described above.

Table S-4. Primers used to amplify *ADH2* terminator from pJC720, and clone into pXP320 (replacing CYC1 terminator), creating pRF320.

Terminator	Direction	Primer name	Primer sequence, 5' → 3'
ADH2	Forward	For-ADH2t	5'-TCCCctcgagAGTCGACATGCATCCTAGGTTTAAACATGC-3'
	Reverse	Rev-ADH2t	5'-TCCCTCTAGAGTCGAGGGGAGCAAACGCGCTGGGAGCAAA-3'

fabG under the control of P_{PGK1} was integrated at one copy into the genome of strain BY4741. PCR was used to generate fragment P_{PGK1}-*fabG*-T_{CYC1} from vector pXP219-fabG and fragment loxP-URA3-loxP from plasmid pXP219 using the primers shown in Table S-4. The P_{PGK1}-*fabG*-T_{CYC1} fragment had 50 bp homologous sequences on the 3' end to the fragment loxP-*URA3*-loxP on the 5' end. The flanking 5' and

3’ ends of the adjoined segment had 50 bp overlapping sequences up- and down-stream of the target *MET15* locus. *fabG* integration into the *MET15* locus was confirmed by PCR analysis using primers that annealed upstream and downstream of the chromosomal site (Table S-7).

fabD under the control of P_{PGK1} was integrated at one copy into the genome of strain BY-G (BY4741 *met15::*P_{PGK1}-*fabG*-T_{CYC1}). PCR was used to generate fragment P_{PGK1}-*fabD*-T_{CYC1} from vector pXP219-fabD and fragment loxP-*URA3*-loxP from plasmid pXP219, using the primers shown in Table S-4. The flanking 5’ and 3’ ends of the adjoined segment had 50 bp overlapping sequences up- and down-stream of the target URA3 locus. *fabD* integrations into the *URA3* locus was confirmed by PCR analysis using primers that annealed upstream and downstream of the chromosomal site.

fabH under the control of P_{PGK1} and *acpS* under the control of P_{ADH2} were simultaneously integrated at one copy each into the genome of strain BY-DG (BY4741 *met15::*P_{PGK1}-*fabG*-T_{CYC1} *ura3::*P_{PGK1}-*fabD*-T_{CYC1}). PCR was used to generate fragment P_{PGK1}-*fabH*-T_{CYC1} from vector pXP219-fabH, using primers Forward1-LEU2 and Reverse1-LEU2 (Table S-4). Primers Forward2-LEU2 and Reverse2-LEU2 were used to generate fragment P_{TEF1}-*acpS*-T_{ADH2}-loxP-*HIS3*-loxP from plasmid pRF320-acpS. The first fragments had 50 bp homologous sequences on the 3’ end to fragment P_{TEF1}-*acpS*-T_{ADH2}-loxP-*HIS3*-loxP on the 5’ end. The flanking 5’ and 3’ ends of the adjoined segment had 50 bp overlapping sequences up- and down-stream of the target *LEU2* locus. *fabH* and *acpS* integrations into the *LEU2* locus were confirmed by PCR analysis using primers that annealed upstream and downstream of the chromosomal site.

acpP under the control of P_{PGK1}, and *fabZ* under the control of P_{ADH2}, were simultaneously integrated at one copy each into the genome of strain BY-aSDGH (BY4741 *met15::*P_{PGK1}-*fabG*-T_{CYC1} *ura3::*P_{PGK1}-*fabD*-T_{CYC1} *leu2::*P_{PGK1}-*fabH*-T_{CYC1}-P_{TEF1}-*acpS*-T_{ADH2}). PCR was used to generate fragment P_{PGK1}-*acpP*-T_{CYC1} from vector pXP219-acpP, using primers Forward1-Ty15 and Reverse1-Ty15 (Table S-4). Primers Forward2-Ty15 and Reverse2-Ty15 were used to generate fragment P_{TEF1}-*fabZ*-T_{ADH2}-loxP-HIS3-loxP from

plasmid pRF320-fabZ. The first fragments had 50 bp homologous sequences on the 3' end to fragment P_{TEF1}-*fabZ*-T_{ADH2}-loxP-*HIS3*-loxP on the 5' end. The flanking 5' and 3' ends of the adjoined segment had 50 bp overlapping sequences up- and down-stream of the target *LEU2* locus. *fabZ* and *acpP* integrations into the *Ty1-5* locus were confirmed by PCR analysis using primers that annealed upstream and downstream of the chromosomal site.

fabB under the control of P_{PGK1}, and *fabI* under the control of P_{ADH2}, were simultaneously integrated at one copy each into the genome of strain BY4741. PCR was used to generate fragment P_{PGK1}-*fabB*-T_{CYC1} from vector pXP219-fabB, using primers Forward1-Ty11 and Reverse1-Ty11 of Table S-4. Primers Forward2-Ty11 and Reverse2-Ty11 shown on Table S-4 were used to generate fragment P_{TEF1}-*fabI*-T_{ADH2}-loxP-*HIS3*-loxP from plasmid pRF320-fabI. The first fragments had 50 bp homologous sequences on the 3' end to fragment P_{TEF1}-*fabI*-T_{ADH2}-loxP-*HIS3*-loxP on the 5' end. The flanking 5' and 3' ends of the adjoined segment had 50 bp overlapping sequences up- and down-stream of the target *Ty15* locus. *fabB* and *fabI* integrations into the *Ty1-5* locus were confirmed by PCR analysis using primers that annealed upstream and downstream of the chromosomal site.

To knockout the *FAS2* gene, nested polymerase chain reaction (PCR) was used to increase the homologous sequence, upstream and downstream, to approximately 100 bp. PCR was run using primers FAS2-KO-M-for and FAS2-KO-M-rev with pXP214 (*MET15* marker) to generate a piece with approximately 25 bp of homology upstream and downstream the *FAS2* gene sequence in the yeast genome. Gel electrophoresis with 1% agar gel was used to separate and to visualize DNA products. In order to increase the homologous sequence of the inserting DNA with the yeast genome, the DNA product from the previous step was used as template in a PCR reaction with primers Nes1-F-FAS2KO and Nes1-R-FAS2KO, which have 24 bp of homology with the template and add 36 bp of homology with the yeast genome on each end of the insert. This process was repeated with primers Nes2-F-FASKO and Nes2-R-FAS2KO to have a total homology of approximately 100 bp on each side.

Table S-5. Primers used to integrate the *E. coli* FAS genes into the *S. cerevisiae* genome.

Locus	Type of Integration	Gene/s	Direction	Primer Sequence 5' → 3'
MET15	Single, 2-piece, double crossover	**fabG**	forward1	GGCAAATGGCACGTGAAGCTGTCGATATTGGGGAACTGTGGTG GTTGGCAAGGCATTTGCAAGAATTACTCGTG
			reverse1	CCCGGGGATCCTCTAGAGTCGACCGGCCGCAAATTAAAGCCTTC GAGCG
			forward2	CGCTCGAAGGCTTTAATTTGCGGCCggtcgactctagaggatcCCCGG G
			reverse2	GCAAATAAAACACTATTGATTGCTTAAAAGGGCAATCCGACTAT ATCTGGAATTCGAGCTCGGTACCCGGG
URA3	Single, 2-piece, double crossover	**fabD**	forward1	GTTTTGACCATCAAAGAAGGTTAATGTGGCTGTGGTTTCAGGGT CCATAAGGCATTTGCAAGAATTACTCGTG
			reverse1	CCCGGGGATCCTCTAGAGTCGACCGGCCGCAAATTAAAGCCTTC GAGCG
			forward2	CGCTCGAAGGCTTTAATTTGCGGCCggtcgactctagaggatcCCCGG G
			reverse2	CCAATTTTTTTTTTTTCGTCATTATAGAAATCATTACGACCGAGAT TCCCGGGAATTCGAGCTCGGTACCCGGG
LEU2	Tandem, 2-piece, double crossover	**fabH acpS**	forward1	CCATGTATAATCTTCATTATTACAGCCCTCTTGACCTCTAATCATG AATGTTAGGCATTTGCAAGAATTACTCGTG
			reverse1	aaggattcgcggtaatattg**aaaaagg** GGCCGCAAATTAAAGCCTTCGAGCG
			forward2	CGCTCGAAGGCTTTAATTTGCGGCC ccttttttcaatattaccgcgaatcctt
			reverse2	GCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAAT TTCGTGTCGAATTCGAGCTCGGTACCCGGG
Ty 1-5	Tandem, 2-piece, double crossover	**acpP fabZ**	forward1	CACAGAGTTGTATTTGCGCTTCTGAGCGATGCTTCCGAGATTGTT GAAGCAAGGCATTTGCAAGAATTACTCGTG
			reverse1	aaggattcgcggtaatattg**aaaaagg** GGCCGCAAATTAAAGCCTTCGAGCG
			forward2	CGCTCGAAGGCTTTAATTTGCGGCCccttttttcaatattaccgcgaatcctt
			reverse2	GATTATTGAAGAGGGATGCGTTTGGTACAATAAAAAACATAGGT TCCCAAACCGAATTCGAGCTCGGTACCCGGG
Ty 1-1	Tandem, 2-piece, double crossover	**fabB fabI**	forward1	GACTTCTAGTATTATCTGTATATCTAATATTATAGTCTCTAACAAC AGTGGAATAGGCATTTGCAAGAATTACTCGTG
			reverse1	aaggattcgcggtaatattg**aaaaagg** GGCCGCAAATTAAAGCCTTCGAGCG
			forward2	CGCTCGAAGGCTTTAATTTGCGGCC ccttttttcaatattaccgcgaatcctt
			reverse2	CATGTATGAAACTGGGAATTCTGATAAATTTTGTCATAACTGTTG GGATTCCGAATTCGAGCTCGGTACCCGGG

Table S-6. Primers to knockout *FAS2* gene from BY4741 and BY8G strains.

Gene	Direction	Primer name	Primer Sequence
FAS2	Forward	FAS2-KO-M-for	CAACTTTGGCTGGGATGGCTCAAACCCGGGataacttcgtataGCATACAT
	Reverse	FAS2-KO-M-rev	ATGTCTTTCAATGCAGCACCACCGCCCGGGataacttcgtataATGTATGC
	Forward	Nes1-F-FAS2KO	TTAACACTGAAAGGGTTGTTGAAATCGGTCCTTCTCCAACTTTGGCTGGGATGGCTCAAA
	Reverse	Nes1-R-FAS2KO	ACGGCTGGAGCGTTTTTGTTAACGCGTACGATTTCGATGTCTTTCAATGCAGCACCACCG
	Forward	Nes2-F-FAS2KO	GATGGATTGAAACTCAAGATGTTTTTTTGAAGGATTTTAACACTGAAAGGGTTGTTGAAA
	Reverse	Nes2-R-FAS2KO	CTTCGGCAGCCTTTTTGGCGTTACCGTGCAGTTCAACGGCTGGAGCGTTTTTGTTAACGC

Table S-7. Loci check integration/knockout primers.

Locus	Direction	Primer Sequence
FAS2	Forward	TGTTGTTGTCGTTGTTGTCCCAGC
	Reverse	CCAGTAAGTCGCTCCAGATTT
LEU2	Forward	GGAATTCTAACAATTATCAAATTGTCCG
	Reverse	CAGCCTGTACATCTGCTTCCC
MET15	Forward	CTCCTCGAGGATTTAGGAATCC
	Reverse	GCGGGATCGAACCGCTGATCC
Ty 1-1	Forward	GGCCTGTGCGTGTTCATAAGGG
	Reverse	ACATGATATCATAACTTACCAG
Ty 1-5	Forward	GCTAATCTTTGGATATAGGGC
	Reverse	CAGATCACTTATACAGCTTTACACAG
URA3	Forward	CTAGGGAAGACAAGCAACG
	Reverse	GTGAAGTCATTGACACAGTCTG3

Media and cultivation

Luria-Bertani (LB) medium containing 100 µg/mL ampicillin was used for *E. coli* cultivation (Sambrook and Russell, 2001). *S. cerevisiae* strains were cultivated in complex YPD (20 g/L dextrose, 20 g/L peptone, 10 g/L yeast extract (BD Biosciences, Sparks, MD)), selective SDC-A (20 g/L dextrose, 5 g/L

casamino acids, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids, 100 mg/L adenine sulfate), or SD minimal medium (20 g/L dextrose, 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without amino acids, and supplementary constituents as needed: 100 mg/L adenine sulfate, 100 mg/L uracil, 100 mg/L histidine-HCl, 100 mg/L methionine, 100 mg/L leucine) (Amberg et al., 2005). Plates contained 2% Bacto-agar. For *FAS2* knockout strains without the full integrated heterologous FAS, the medium contained 2 mM myristic acid with tween 40 (0.5 %) to supplement growth (Chirala et al., 1987).

References

Amberg, D., Burke, D., and Strathern, J. (2005). Methods in yeast genetics. (New York: Cold Spring Harbor Laboratory Press).

Chirala, S.S., Kuziora, M.A., Spector, D.M., and Wakil, S.J. (1987). Complementation of mutations and nucleotide sequence of FAS1 gene encoding beta subunit of yeast fatty acid synthase. J. Biol. Chem. *262*, 4231–4240.

Cho, H., and Cronan, J.E. (1993). Escherichia coli thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme. J. Biol. Chem. *268*, 9238–9245.

Heath, R.J., and Rock, C.O. (1996). Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in Escherichia coli fatty acid biosynthesis. J. Biol. Chem. *271*, 27795–27801.

Hoskins, L. (2000). Yeast Transformation (high efficiency).

Magnuson, K., Carey, M.R., and Cronan, J.E. (1995). The putative fabJ gene of Escherichia coli fatty acid synthesis is the fabF gene. J. Bacteriol. *177*, 3593–3595.

Sambrook, J., and Russell, D. (2001). Molecular cloning: A laboratory manual (NY: Cold Spring Harbor Laboratory Press).

White, S.W., Zheng, J., Zhang, Y.M., and Rock CO (2005). The structural biology of type II fatty acid biosynthesis. Annu. Rev. Biochem. *74*, 791–831.

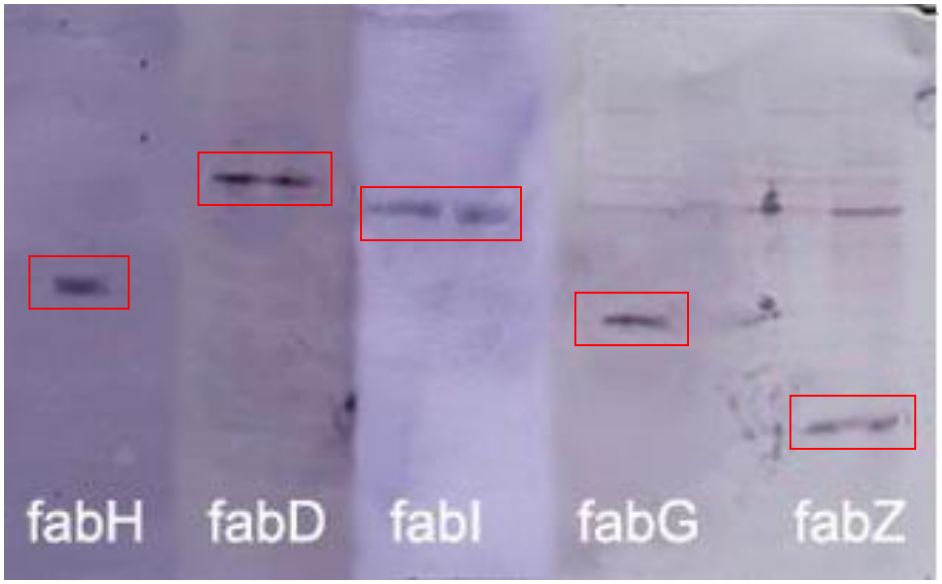


Figure S-1. Western blot analysis of purified fabH-6xHis, fabD-6xHis, fabI-6xHis, fabG-6xHis and fabZ-6xHis. The image was created from independent Western blot images. Relative sizes among enzymes are not to scale.

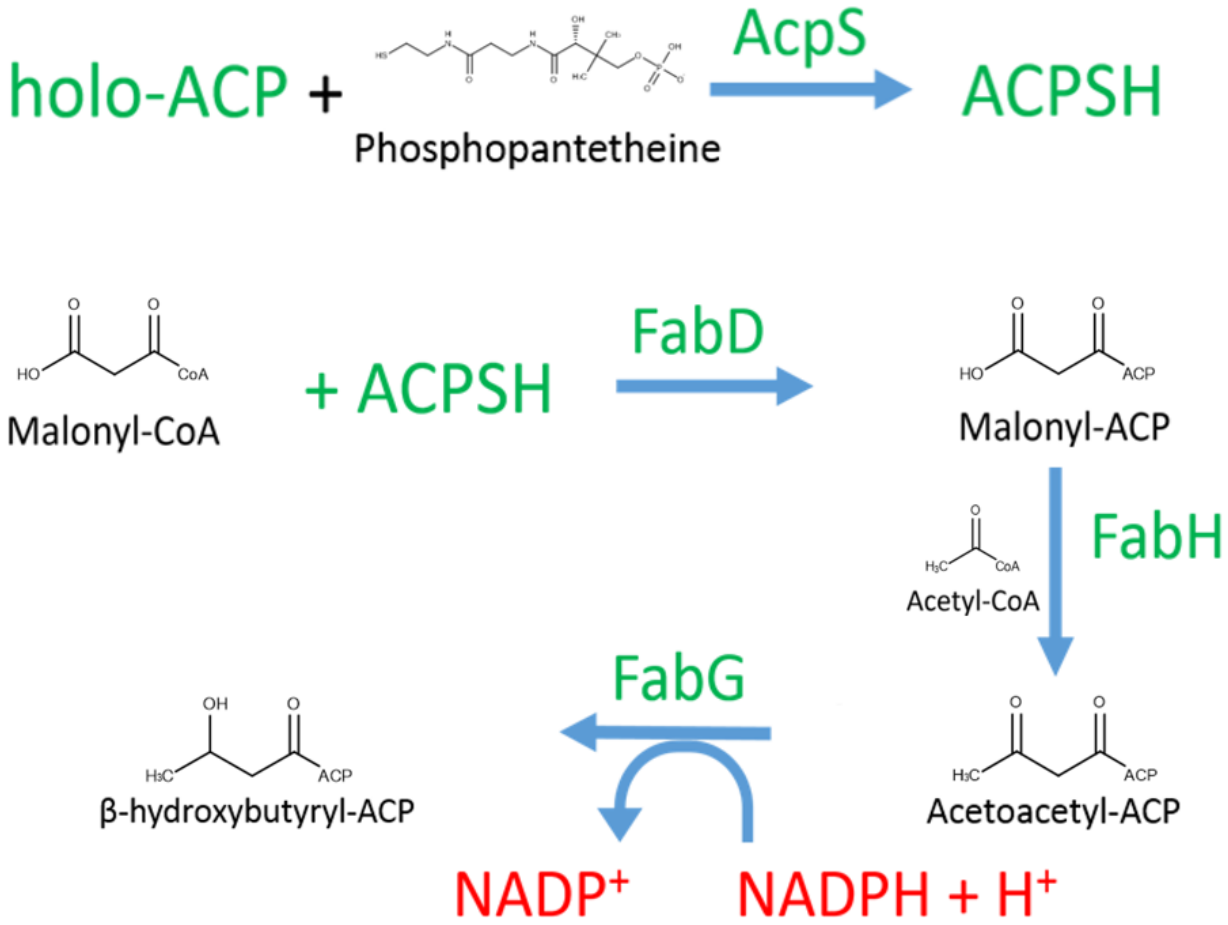


Figure S-2. Pathway for the production of β-hydroxybutyryl-ACP. The reduction of the keto group to alcohol requires NADPH as a reducing agent, and its depletion can be monitored.

Integration Strategy

A) B)

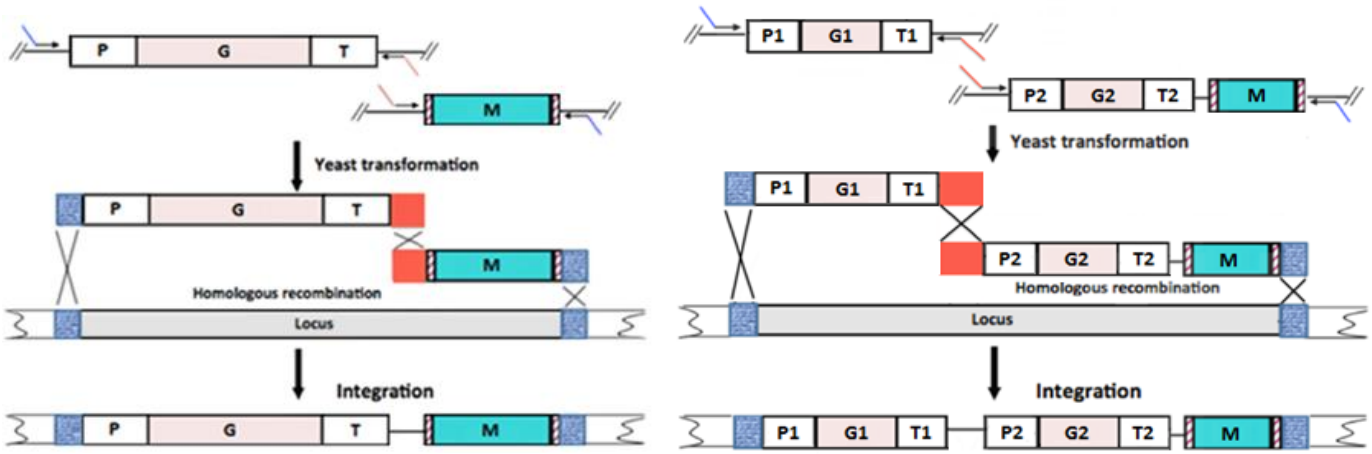
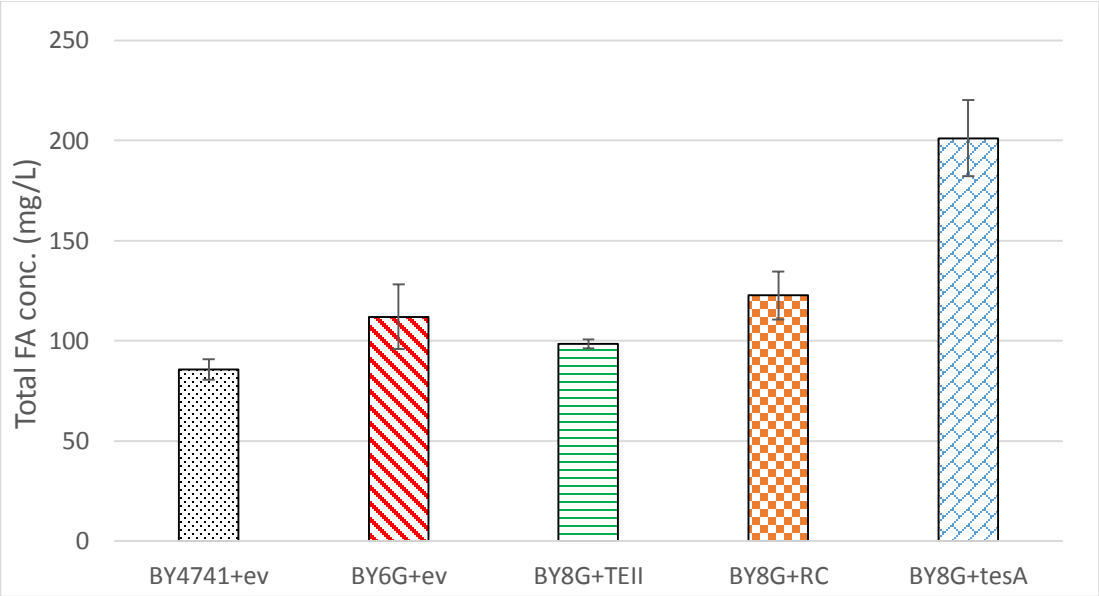


Figure S-3. Schematic comparison of the single gene and tandem gene integration. **A)** Double crossover integration of gene G into one locus. P1 is promoter P_{PGK1} and T1 is terminator T_{CYC1}. **B)** Integration of gene G1 and gene G2. P1 is P_{PGK1}, P2 is P_{TEF1}, T1 is T_{CYC1}, and T2 is T_{ADH2}. The construct P1-G1-T1 was obtained through PCR using a pXP219-based plasmid as template. The construct P2-G2-T2-M (e.g., M=*LEU2* marker) was obtained through PCR using a pRF320-based plasmid as template.

A)



B)

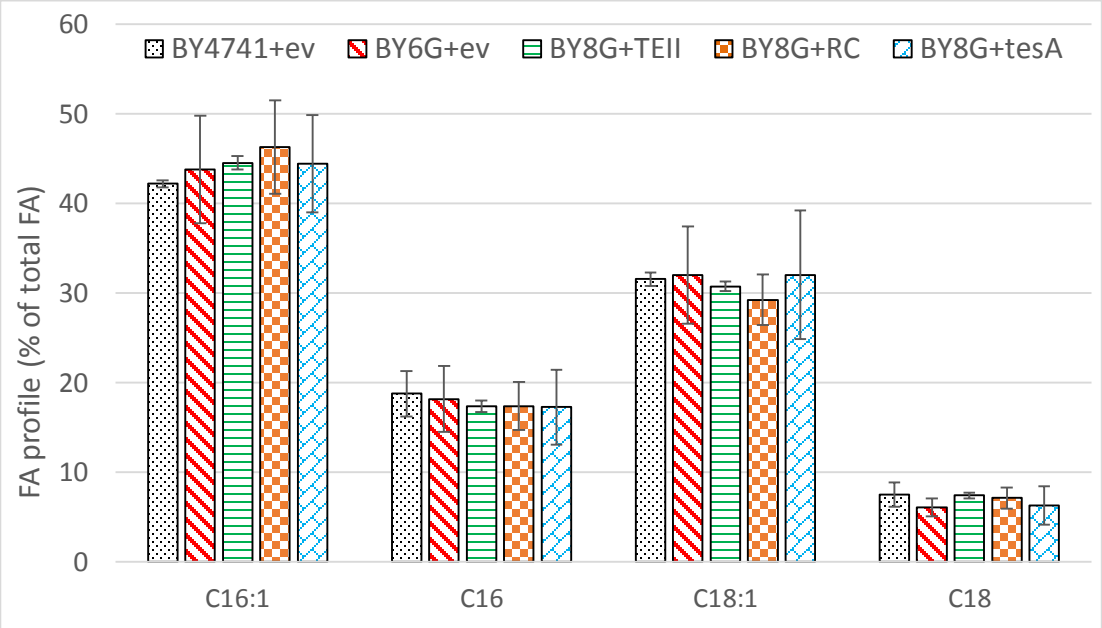


Figure S-4. A) Total fatty acids produced by strain BY8G after 24 h using thioesterase TEII, RC, and TesA. Control strains with empty vector (ev), BY4741+pXP219 (dotted) and BY6G+pXP219 (diagonal striped) are compared to strains BY8G+pXP219-TEII (horizontal striped), BY8G+pXP219-RC (squared), and BY8G+pXP219-tesA (brick). **B)** Fatty acid profile as % of the total FA produced by each strain described above. No C14 was detected for any of the samples. Results are expressed as mean ± error from two independent experiments.

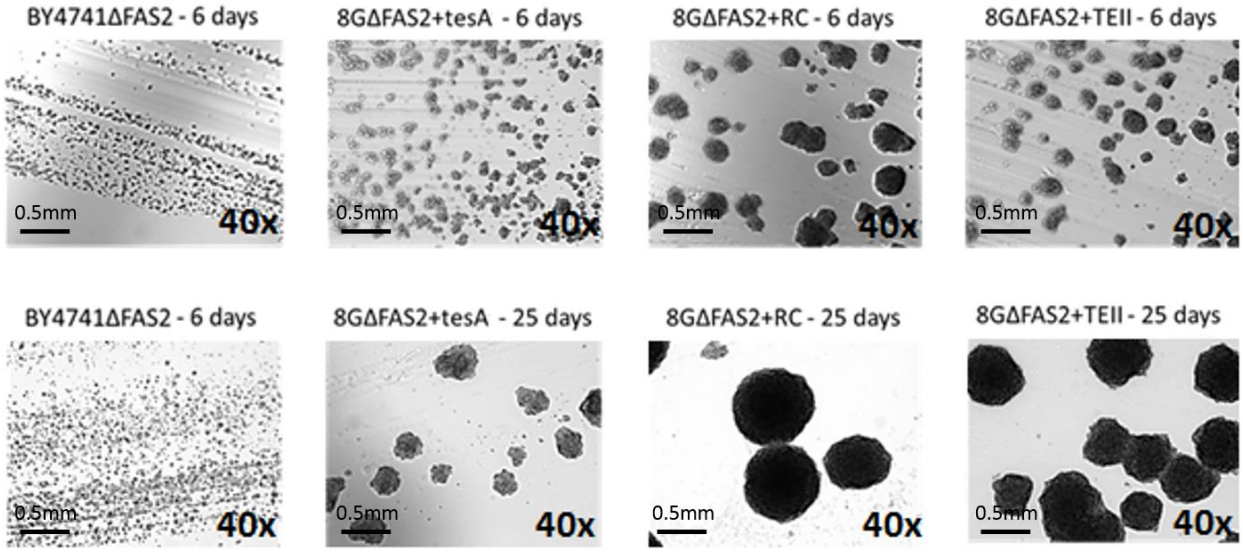


Figure S-5. *In vivo* confirmation of heterologous *E.coli* FAS activity in strain BY8GΔFAS2 with different thioesterases: TesA, RC, and TEII. The negative control was strain BY4741ΔFAS2 carrying an empty plasmid pXP219. The ΔFAS2 strains were streaked on synthetic minimal media plates supplemented with fatty acids. Subsequent colonies were re-suspended in water, plated in the absence of FAs, and allowed to grow for 6 days (first row) and 25 days (second row). The images are 40x magnification bright-field microscopy pictures (Olympus, BX51).

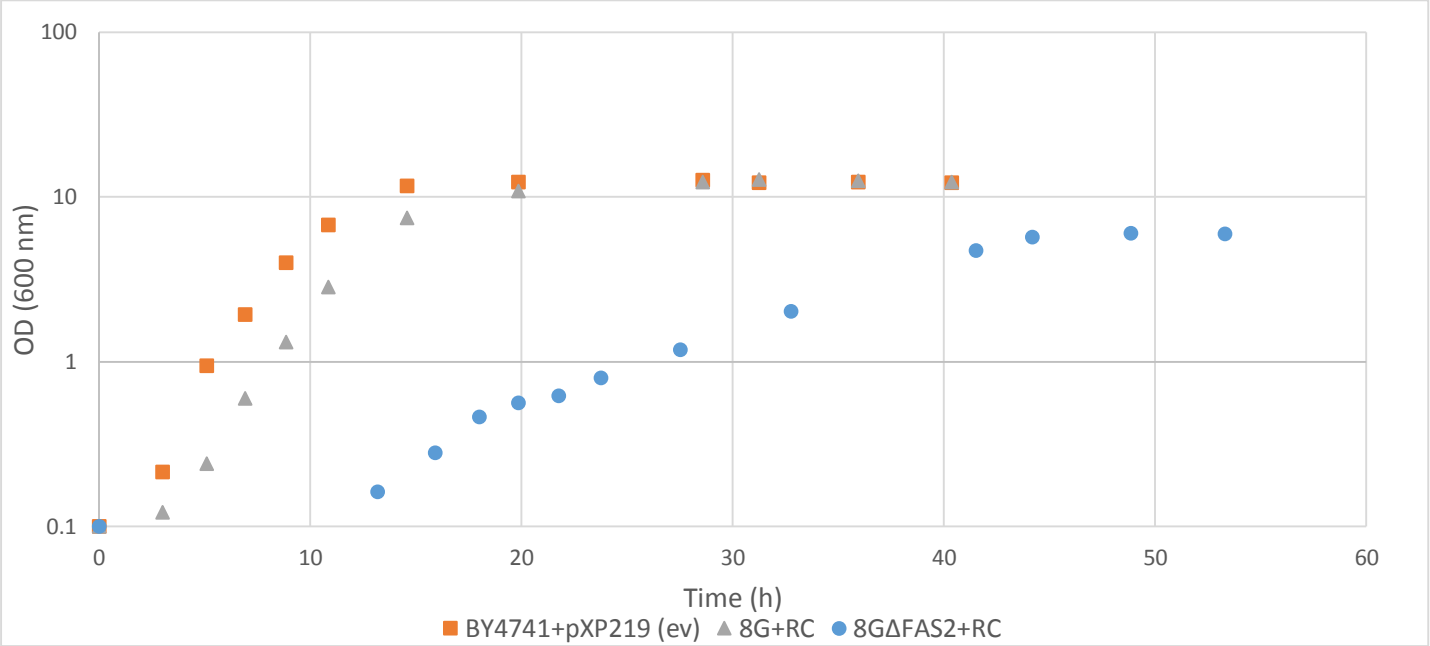


Figure S-6. Growth curves for 8GΔFAS2+pXP219-RC (blue circles), 8G+pXP219-RC (grey triangles), and BY4741+pXP219 (orange squares) as a control. Cells were inoculated overnight in 5mL SDC(A), cells from those cultures were used to inoculate 50 mL SDC(A) at an initial OD_{600} of 0.1.