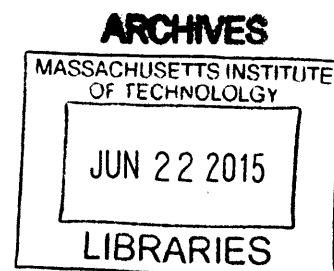


¹³C-Metabolic Flux Analysis of Recombinant Yeasts for Biofuels Applications

by

Thomas M. Wasylenko

B.S.E. Chemical Engineering, Princeton University, 2008



SUBMITTED TO THE DEPARTMENT OF CHEMICAL ENGINEERING IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY


JUNE 2015

© 2015 Massachusetts Institute of Technology. All rights reserved.

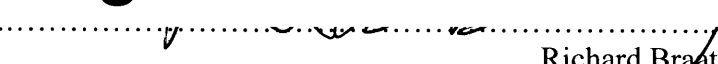
Signature redacted

Signature of author.....
Department of Chemical Engineering
March 13, 2015

Signature redacted

Certified by.....

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

Signature redacted

Accepted by.....

Richard Braatz
Edwin R. Gilliland Professor of Chemical Engineering
Chairman, Committee for Graduate Students

¹³C-Metabolic Flux Analysis of Recombinant Yeasts for Biofuels Applications

by

Thomas M. Wasylenko

Submitted to the Department of Chemical Engineering
on March 13, 2015 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Abstract

Fossil fuels have powered the transportation industry since the Industrial Revolution. However, rising transportation energy demand and new knowledge about the environmental impact of burning fossil fuels have motivated the development of technologies for sustainable production of renewable, carbon-neutral liquid fuels. To that end, biological systems may be leveraged to fix carbon dioxide and to catalyze the conversion of renewable feed stocks to fuel molecules. Today, the gasoline additive ethanol and biodiesel are produced by yeast fermentation of sugars derived from cornstarch and sucrose and transesterification of vegetable oils, respectively. However ethanol has many drawbacks as a fuel additive, and both biofuels are currently produced from edible feed stocks. For biofuels to contribute significantly to meeting total transportation energy demand, processes for production of fuel molecules from non-food feed stocks must be engineered. Two promising solutions are fermentation of sugars derived from “woody,” lignocellulosic biomass and production of fuels from volatile fatty acids (VFAs) such as acetate, which can be produced by fermentation of organics in municipal solid waste and sewage or syngas. The production of biofuels from lignocellulosic material or VFAs will require metabolic engineering of biocatalysts to improve yields, productivities, and final titers. These metabolic engineering efforts can be facilitated by ¹³C-Metabolic Flux Analysis (MFA), a method for elucidating the otherwise unobservable intracellular metabolic fluxes in biological systems.

We first developed protocols for extraction and LC-MS/MS analysis of intracellular metabolites, which provides data that may be used for metabolic flux estimation. We then performed an analysis of both the measurement and modeling errors associated with using these data for flux determination. Finally, we applied ¹³C-MFA to two industrially relevant systems: 1) Fermentation of xylose, a sugar present in lignocellulosic biomass, to ethanol in *Saccharomyces cerevisiae*, and 2) overproduction of fatty acids that may be transesterified to biodiesel from either glucose or acetate in the oleaginous yeast *Yarrowia lipolytica*. These experiments identified a potential bottleneck in xylose fermentation in *S. cerevisiae* and the primary source of NADPH for fatty acid biosynthesis in *Y. lipolytica*, and also suggested potential strategies for improving lipid yields in *Y. lipolytica*.

Thesis Supervisor: Gregory Stephanopoulos

Title: Willard Henry Dow Professor of Chemical Engineering and Biotechnology

Acknowledgments

Graduate school has been a long journey, and I'd like to thank the many, many people who have helped me along the way. My thesis advisor Prof. Gregory Stephanopoulos has provided me with ample resources and, just as importantly, a great deal of freedom to use them as I have seen fit and grow as a researcher. I am inherently a detail-oriented person, and Greg's tremendous vision has helped me learn to see the "big picture" and understand the motivation for my work. Finally his enthusiasm is unmatched; I can honestly say that every time I have had an individual meeting with Greg, I have come out more excited about my research than I was when I went in.

I am also grateful to those who have served on my thesis committee—Gerry Fink, Joanne Kelleher, Narendra Maheshri, and Kris Prather—for their support, encouragement, and helpful questions and suggestions.

I thank Rosangela dos Santos for her administrative assistance, which keeps the Stephanopoulos lab running smoothly, has simplified my life (and those of everyone else in the lab) considerably, and allows us to devote more time and attention to our research.

The Stephanopoulos Lab is a large lab with high turnover, and consequently I have had the privilege to interact with many talented individuals who have made the lab a fun and exciting place. Curt Fischer trained me when I first joined the lab. Jason Walther introduced me to the metranCL flux estimation software and showed me the basics of intracellular metabolite extractions. Takashi Yamamoto worked with me to set up the Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) system for measurement of phosphorylated metabolites. Working with Takashi, an experienced researcher, early in my graduate school career was a great opportunity which allowed me to learn first-hand how to conduct good research. Hang Zhou developed the xylose-consuming *Saccharomyces cerevisiae* strain characterized in Chapter 5 and taught me about that strain and about techniques for working with yeast. Simon Carlsen helped me with yeast molecular biology. Sarah Fendt and Marjan de Mey taught me about using ^{13}C -labeled metabolite extracts as internal standard to greatly improve the precision of my metabolite pool size measurements. Hamid Rismani-Yazdi and Brian Pereira helped me when I was struggling with culturing *S. cerevisiae* anaerobically. KJ Qiao and Andy Silverman showed me the ropes when I transitioned to working with *Yarrowia lipolytica*. I had the pleasure of working with Woo Suk Ahn and Moritz Wolf on *Y. lipolytica* flux estimation, described in Chapters 6 and 7. Each directly generated data that were used in this thesis. Mark Keibler shared with me a Flux Balance Analysis code he had written, which I used for the analysis of *Y. lipolytica* acetate metabolism in Chapter 7. The technical expertise of all of these people (and their willingness to share it) saved me countless hours and a great deal of frustration. Each one of them truly contributed to making this thesis what it is.

I also want to thank the people who made my time in the Stephanopoulos lab enjoyable, even if they didn't contribute directly to my research. Brian Pereira and Haoran Zhang were my "mentors" who helped me cope with the ups and downs of grad school. I could always count on Haoran to help me out in a pinch, and I lost track of how many Legal Seafoods dinners I owe him long ago. Vikram Yadav was my bay-mate early on, and we had good times, spirited debates, and some very competitive games of lab mini-golf. It was always a pleasure talking to

Adel Ghaderi, and I learned a great deal about science, philosophy, and many things in between in discussions with him. Mark Keibler and Sagar Chakraborty often kept me company late into the night during my 24-hour time courses, and Mark always kept me up-to-date on the latest in cancer metabolism and obscure news. It has been a pleasure to watch Steven Edgar, Ben Woolston, Zhe Zhang, David Emerson, and Alkis Chatzivasileiou grow as scientists, and although I am leaving the lab I look forward to seeing the great papers I know they will publish.

I shared a lot of fun times with my Davis Square roommates Andrew Hilmer and Nate Vacanti (and Andrew's now-wife Jingqing) and my lunch crew Caroline Ahrens, Rose Kanasty, Yuko Kida, Micah Sheppard, and Nate. They also commiserated with me during the inevitable lows of research and helped me to keep things in perspective.

I thank my parents Michael and Lydia Wasylenko for their unconditional love, guidance, and support which made it possible for me to get to and get through MIT, and my "golden ticket," my wife Theresa, for putting up with my probably insufferable ego during the highs and my definitely insufferable moodiness and complaining during the (much more common) lows. I cannot imagine the last four years without her.

When I think about all those who have helped me during my time at MIT, I believe it is truly a miracle that I have had so many wonderful people come into my life in such a short time. And so finally, I thank God, who has sustained me over the past six-plus years. Scripture assures me that all He has created is good (Gen 1:31). Even *Yarrowia*.

Contents

1 Introduction	20
1.1 The Need for New Biofuel Technologies.....	20
1.1.1 Lignocellulosic Feed Stocks.....	21
1.1.2 Alternative Fuel Molecules.....	23
1.2 ¹³ C-Metabolic Flux Analysis.....	25
1.3 ¹³ C-MFA Mathematical Formulation.....	28
1.4 Thesis Overview: ¹³ C-MFA of Recombinant Yeasts for Biofuels Applications.....	32
1.5 References.....	33
2 Methods for Metabolite Extraction and LC-MS/MS Analysis	41
2.1 Introduction.....	41
2.2 Materials and Methods.....	46
2.2.1 Optimization of Compound Parameters.....	46
2.2.2 Optimization of Source/Gas Parameters.....	46
2.2.3 Liquid Chromatography-Tandem Mass Spectrometry.....	47
2.2.4 Strain and Culture Conditions.....	49
2.2.5 Comparison of Quenching Methods.....	50
2.2.6 Comparison of Metabolite Extraction Protocols with GC-MS.....	51
2.2.6.1 Hot Ethanol Extraction.....	51
2.2.6.2 Chloroform-Methanol-Water Extraction.....	51
2.2.6.3 Acidic Acetonitrile Methanol Extraction.....	52

2.2.7 Comparison of Hot Ethanol and Chloroform-Methanol-Water Extraction Protocols with LC-MS/MS.....	53
2.2.8 Measurement of Mass Isotopomer Distributions (MIDs) of Metabolites Labeled to Natural Abundance.....	53
2.2.9 Validation of Protocols by Measurement of Metabolite Ratios.....	54
2.2.10 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of TMS Derivatives.....	54
2.2.11 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of TBDMS Derivatives.....	55
2.3 Results and Discussion.....	55
2.3.1 Optimization of Compound Parameters.....	55
2.3.2 Optimization of Source/Gas Parameters.....	58
2.3.3 Liquid Chromatography-Tandem Mass Spectrometry.....	62
2.3.4 Comparison of Quenching Methods.....	63
2.3.5 Comparison of Metabolite Extraction Protocols with GC-MS.....	67
2.3.6 Comparison of Hot Ethanol and Chloroform-Methanol-Water Extraction Protocols with LC-MS/MS.....	70
2.3.7 Measurement of Mass Isotopomer Distributions (MIDs) of Metabolites Labeled to Natural Abundance.....	72
2.3.8 Validation of Quenching and Extraction Protocols by Measurement of Metabolite Ratios.....	73
2.4 Conclusions.....	75
2.5 References.....	76

3 Modeling of Propagation of Error in Mass Spectrometry Mass Isotopomer

Distribution Data	80
3.1 Introduction.....	80
3.2 Error Model.....	83
3.3 Materials and Methods.....	87
3.3.1 Simulation of Noisy MS Labeling Data.....	87
3.3.2 Strain and Culture Conditions.....	88
3.3.3 Protein Hydrolysis.....	88
3.3.4 Gas Chromatography/Mass Spectrometry (GC-MS).....	89
3.4 Results.....	89
3.4.1 Simulation of Noisy MS Labeling Data.....	89
3.4.2 Analysis of <i>E. coli</i> Protein Hydrolysates by GC-MS.....	95
3.5 Discussion.....	97
3.6 References.....	100

4 Kinetic Isotope Effects Significantly Influence Intracellular Metabolite ¹³C Labeling

Patterns and Flux Determination	103
4.1 Introduction.....	103
4.2 Modeling Isotopic Fractionation at the Pyruvate Node.....	104
4.3 Results.....	116
4.4 Discussion.....	122
4.5 References.....	125

5 Metabolomic and ¹³C-Metabolic Flux Analysis of a Xylose-Consuming <i>Saccharomyces cerevisiae</i> Strain Expressing Xylose Isomerase	131
5.1 Introduction.....	131
5.2 Materials and Methods.....	135
5.2.1 Strain and Culture Conditions.....	135
5.2.2 Analytical Methods.....	136
5.2.3 Determination of Optical Density-Dry Cell Weight Correlations.....	137
5.2.4 Estimation of Extracellular Fluxes.....	137
5.2.5 Preparation of Uniformly ¹³ C-Labeled Cell Extract for Internal Standard.....	139
5.2.6 Metabolite Extractions for Pool Size Measurements and ¹³ C-Metabolic Flux Analysis.....	139
5.2.7 Gas Chromatography/Mass Spectrometry (GC-MS).....	140
5.2.8 Liquid Chromatography/Tandem Mass Spectrometry (LC-MS/MS).....	141
5.2.9 Quantification of Metabolite Pool Sizes.....	141
5.2.10 Metabolic Flux Estimation.....	142
5.3 Results and Discussion.....	144
5.3.1 Determination of Optical Density-Dry Cell Weight Correlations.....	144
5.3.2 Estimation of Extracellular Fluxes.....	145
5.3.3 Metabolite Pool Size Measurements and Metabolic Flux Estimation.....	147
5.3.4 Pentose Phosphate Pathway.....	154
5.3.5 Glycolysis.....	155
5.3.6 Tricarboxylic Acid Cycle.....	156

5.3.7 Incomplete Activation of Fermentation Program.....	157
5.3.8 Lower Glycolysis as a Potential Bottleneck in Xylose Metabolism.....	161
5.4 References.....	165

6 The Oxidative Pentose Phosphate Pathway is the Primary Source of NADPH for Lipid

Overproduction from Glucose in *Yarrowia lipolytica* 180

6.1 Introduction.....	180
6.2 Material and Methods.....	185
6.2.1 Strains and Culture Conditions.....	185
6.2.2 Extracellular Metabolite Quantification.....	186
6.2.3 Dry Cell Weight Measurements.....	187
6.2.4 Lipid Analysis.....	187
6.2.5 Estimation of Extracellular Fluxes.....	188
6.2.6 Metabolite Extractions.....	189
6.2.7 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of Metabolite Extracts.....	190
6.2.8 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS).....	191
6.2.9 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of Glucose.....	191
6.2.10 Metabolic Flux Estimation.....	192
6.3 Results.....	194
6.3.1 Fermentation Profiles and Establishment of Metabolic Steady State.....	194
6.3.2 Estimation of Extracellular Fluxes.....	199

6.3.3 Metabolic Flux Estimation.....	200
6.4 Discussion.....	207
6.5 References.....	210
7 ¹³C-Metabolic Flux Analysis and Flux Balance Analysis of Lipid Production from	
Acetate in <i>Yarrowia lipolytica</i>	220
7.1 Introduction.....	220
7.2 Materials and Methods.....	222
7.2.1 Strain and Culture Conditions.....	222
7.2.2 Extracellular Metabolite Quantification.....	223
7.2.3 Dry Cell Weight Measurements.....	223
7.2.4 Analytical Methods.....	223
7.2.5 Estimation of Extracellular Fluxes.....	223
7.2.6 Metabolic Flux Estimation.....	224
7.2.7 Flux Balance Analysis.....	225
7.3 Results and Discussion.....	226
7.3.1 Fermentation Profiles and Metabolic Steady State.....	226
7.3.2 Estimation of Extracellular Fluxes.....	229
7.3.3 Metabolic Flux Estimation.....	229
7.3.4 Flux Balance Analysis.....	232
7.3.5 Stoichiometric Analysis of NADPH Production Pathways.....	242
7.4 Conclusions.....	246
7.5 References.....	248

8 Conclusions and Recommendations for Future Work	252
8.1 Conclusions.....	252
8.2 Recommendations for Future Work.....	254
8.3 References.....	256
A <i>Saccharomyces cerevisiae</i> Xylose Flux Estimation Details	258
B <i>Yarrowia lipolytica</i> Glucose Flux Estimation Details	286
C <i>Yarrowia lipolytica</i> Acetate Flux Estimation Details	305

List of Figures

2-1	Curtain Gas (CUR) optimization.....	59
2-2	IonSpray Voltage (IS) optimization.....	60
2-3	Ion Source Gas 1 (GS1) optimization.....	60
2-4	Ion Source Gas 2 (GS2) optimization.....	61
2-5	Collision Gas (CAD) optimization.....	61
2-6	Comparison of quenching solutions using the pipetting method for washing pellets.....	65
2-7	Comparison of methods for washing cell pellets with pure methanol quenching solution.....	66
2-8	Comparison of extraction protocols by GC-MS with TBDMS derivatives.....	68
2-9	Comparison of extraction protocols by GC-MS with TMS derivatives.....	69
2-10	Comparison of hot ethanol and chloroform-methanol-water extraction protocols by LC-MS/MS.....	71
2-11	Validation of LC-MS/MS method for measuring metabolite labeling patterns.....	73
3-1	Results of Monte Carlo simulations used to validate MS data error model in the case of proportional error only in ion counts ($a_1 = 0, a_2 = 0.01, N = 3$).....	91
3-2	Results of Monte Carlo simulations used to validate MS data error model in the case of constant error only in ion counts ($a_1 = 0.003, a_2 = 0, N = 3$).....	92
3-3	Results of Monte Carlo simulations used to validate MS data error model in the case of both proportional and constant errors in ion counts ($a_1 = 0.003, a_2 = 0.01, N = 3$)...	92
3-4	Results of Monte Carlo simulations used to validate MS data error model with double proportional error only in ion counts ($a_1 = 0, a_2 = 0.02, N = 3$).....	93

3-5	Results of Monte Carlo simulations used to validate MS data error model with double constant error only in ion counts ($a_1 = 0.006, a_2 = 0, N = 3$).....	93
3-6	Results of Monte Carlo simulations used to validate MS data error model with double proportional and constant errors in ion counts ($a_1 = 0.006, a_2 = 0.02, N = 3$).....	94
3-7	Results of Monte Carlo simulations used to validate MS data error model in the case of proportional error only in ion counts ($a_1 = 0, a_2 = 0.01, N = 6$).....	94
3-8	Results of Monte Carlo simulations used to validate MS data error model in the case of constant error only in ion counts ($a_1 = 0.003, a_2 = 0, N = 6$).....	95
3-9	GC-MS measurement errors in MID mole fractions of proteinogenic amino acids labeled to natural abundance.....	96
4-1	The pyruvate node model system.....	106
4-2	Typical results for errors associated with neglecting isotope effects as a function of f .	117
4-3	Expected errors associated with isotope effects at low values of f	121
5-1	DCW vs. OD ₆₀₀ for each set of culture conditions.....	145
5-2	Metabolite pool sizes for metabolites in Glycolysis and the PPP.....	148
5-3	Metabolite pool sizes for metabolites in the TCA Cycle.....	149
5-4	Metabolite pool sizes and ratios for select cofactors.....	150
5-5	Best-fit metabolic flux distributions in xylose-consuming <i>S. cerevisiae</i> strain under four sets of conditions.....	151
5-6	Weighted Sums of Squared Residuals from flux estimations with different models.....	152
5-7	Confidence intervals for select fluxes.....	153
6-1	Cycle for shuttling acetyl-CoA from the mitochondria to the cytosol for fatty acid synthesis.....	184

6-2	Extracellular metabolite fermentation profiles.....	197
6-3	Fatty acid fermentation profiles.....	198
6-4	Final fatty acid distribution profiles.....	199
6-5	Estimated metabolic flux distributions.....	203
6-6	Select flux confidence intervals.....	205
6-7	Mass isotopomer distributions (MIDs) of illustrative metabolites from the 1,2- ¹³ C ₂ - glucose labeling experiment.....	206
7-1	Extracellular metabolite fermentation profiles.....	227
7-2	Fatty acid fermentation profiles.....	228
7-3	Final fatty acid distribution profile.....	228
7-4	Estimated metabolic flux distribution in strain MTYL065 on acetate.....	233
7-5	Results of Flux Balance Analysis with base model.....	237
7-6	Results of Flux Balance Analysis with malic enzyme as only source of NADPH.....	238
7-7	Results of Flux Balance Analysis with malic enzyme as only source of NADPH and malate/oxaloacetate shuttling.....	239
7-8	Results of Flux Balance Analysis with addition of cytosolic NADP-dependent isocitrate dehydrogenase.....	240
7-9	Results of Flux Balance Analysis with cytosolic NADP-dependent isocitrate dehydrogenase and irreversible mitochondrial NAD-dependent isocitrate dehydrogenase.....	241
7-10	Maximum theoretical oleate yields as calculated by Flux Balance Analysis under different assumptions about NADPH sources.....	242

B-1	Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL037, natural abundance glucose culture.....	287
B-2	Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL037, 20% U- ¹³ C ₆ -glucose culture.....	288
B-3	Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL037, 1,2- ¹³ C ₂ -glucose culture.....	289
B-4	Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL065, natural abundance glucose culture.....	290
B-5	Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL065, 20% U- ¹³ C ₆ -glucose culture.....	291
B-6	Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL065, 1,2- ¹³ C ₂ -glucose culture.....	292
C-1	Linear regressions for Estimation of Extracellular Fluxes.....	306

List of Tables

2-1	Solvent gradient for slow LC-MS/MS method.....	48
2-2	Solvent gradient for fast LC-MS/MS method.....	49
2-3	Optimal Metabolite Compound Parameters.....	57
2-4	Optimal Source/Gas parameter values.....	62
2-5	Typical metabolite retention times (in minutes) for fast and slow HPLC methods.....	63
2-6	Measurement of key metabolite ratios for validation of quenching, extraction, and LC-MS/MS methods.....	75
4-1	Predicted Isotopomer Flux Fractions.....	113
5-1	Conversion factors for converting OD ₆₀₀ to DCW, k_j^{DCW}	144
5-2	Specific growth rates and specific metabolite consumption and production rates under different conditions.....	146
6-1	Acetyl-CoA, Glycerol-3-Phosphate, and NADPH requirements for biosynthesis of each of the five principal fatty acids in <i>Y. lipolytica</i>	189
6-2	Extracellular flux values used for ¹³ C-MFA.....	200
7-1	Extracellular flux values used for ¹³ C-MFA.....	229
7-2	ATP balance.....	231
A-1	Model Metabolic Reaction Network for ¹³ C-MFA.....	258
A-2	Measured and Simulated Mass Spectrometry Data.....	262
A-3	Best-Fit Fluxes and Flux Confidence Intervals.....	275
B-1	Model Metabolic Reaction Network for ¹³ C-MFA.....	293
B-2	Measured and Simulated Mass Spectrometry Data.....	295
B-3	Best-Fit Fluxes and Flux Confidence Intervals.....	301

C-1	Model Metabolic Reaction Network for ^{13}C -MFA.....	307
C-2	Measured and Simulated Mass Spectrometry Data.....	309
C-3	Best-Fit Fluxes and Flux Confidence Intervals.....	313
C-4	List of Reactions Included in FBA Model.....	316

Chapter 1

Introduction

1.1 The Need for New Biofuel Technologies

Today the global economy relies heavily on fossil fuels to meet demand for transportation energy. However in coming years this demand is expected to increase dramatically due to world population growth and improved standard of living in nations with rapidly growing economies, including China and India. Moreover, the burning of fossil fuels increases the concentration of carbon dioxide in the atmosphere and may contribute to climate change. Consequently, there is great interest in development of technologies for production of liquid fuels from plant biomass. Such technologies could help meet rising demand for transportation fuel and diversify the transportation energy supply to protect against volatility in the market. Moreover, since the carbon in plant biomass is derived from atmospheric carbon dioxide through photosynthesis, biofuels are renewable and can theoretically be “carbon neutral” in so far as the burning of biofuels does not result in a net increase in the amount of carbon dioxide in the atmosphere.

Ethanol is presently the most widely used biofuel, because it can be produced at high yield, productivity, and titer by fermentation of hexose sugars obtained from cornstarch or sucrose (sugarcane). The yield, productivity, and titer metrics are important because they determine the cost of feed stock, the capital costs (i.e. number and size of reactors and other units), and cost of

downstream processing (i.e. separations and purifications) to produce a given amount of ethanol per unit time. However, there are two major drawbacks associated with these “first generation” processes for bioethanol production. The first is that corn and sugarcane are important food crops. Consequently they are relatively expensive feed stocks and represent a major fraction of the total cost of ethanol production. The high feed stock price makes it difficult for first generation bioethanol to be economically competitive with fossil fuels. Moreover arable land is a finite resource, and the use of food crops for fuel production could potentially compete with the food supply and drive up food prices. The second major drawback is that ethanol itself is not an ideal fuel molecule. Ethanol has relatively low energy density and is hygroscopic. The consequences are that ethanol-gasoline blends have inferior fuel economy and can absorb water, which can result in phase separation of the fuel. Phase separation leads to a whole host of problems in the existing fuel infrastructure, as does ethanol’s corrosive nature. Thus, broadly speaking there are (at least) two major challenges to improving on first generation ethanol processes. The first is to develop strategies for economical production of liquid biofuels from non-food crop feed stocks. The second is to engineer the production of fuel molecules superior to ethanol at high yield, productivity, and titer.

1.1.1 Lignocellulosic Feed Stocks

An alternative to cornstarch and sucrose feed stocks is lignocellulosic biomass. Lignocellulosic material is “woody” biomass which can be obtained from agricultural residues (e.g. bagasse, corn stover, rice straw, and wheat straw) or dedicated fuel crops (e.g. miscanthus and switchgrass) which can be grown to yields of biomass per acre far superior to those of corn and sugarcane and can often be cultivated on low quality soil that is not suitable for production of

food crops. Consequently, their production is not expected to compete with the food supply (at least not directly). The United States Department of Energy has estimated that by mid-century the United States could sustainably produce 1.3 billion tons of dry biomass, equivalent to approximately one-third of the nation's current petroleum consumption, while still meeting food, feed, and export needs (Perlack et al., 2005).

However, conversion of lignocellulosic biomass to liquid fuels is difficult. This material consists primarily of three fractions: cellulose (a polymer of glucose), hemicellulose (polysaccharides composed of many sugars but especially rich in the pentose sugar xylose), and lignin (a polymer of aromatics). Production of liquid fuels requires the hydrolysis of the cellulose and hemicellulose fractions to their constituent sugars, which can be converted to fuel molecules by fermentation. However, the three primary components of woody biomass are interwoven and covalently cross-linked into a complex structure, making lignocellulosic material notoriously recalcitrant. In addition, the structure of cellulose makes its hydrolysis significantly more difficult than that of cornstarch, even in a purified form. Thus, while lignocellulosic feed stocks are much less expensive than food crops, their conversion to fermentable sugars requires significantly more time and money.

There are also obstacles to the cost-effective fermentation of sugar monomers derived from lignocellulosic biomass to fuel molecules. Hydrolysis of lignocellulosic material generates compounds such as furan derivatives, weak acids, and phenolics which inhibit microbial biocatalysts (Almeida et al., 2007). In addition, the yeasts used in first generation processes to effect the conversion of hexose sugars to ethanol generally cannot metabolize pentose sugars

including xylose, which is the major component of hemicellulose. Production of lignocellulosic fuels cannot be cost-effective if the pentose sugars are not utilized (Stephanopoulos, 2007; Carroll and Somerville, 2009). Yeasts capable of pentose fermentation have been engineered in recent years (Hahn-Hägerdal et al., 2007; van Maris et al., 2007; Wisselink et al., 2009), but productivities are still significantly lower than those obtained in glucose fermentations. Thus, yeast strains with superior ability to utilize xylose are still needed. Xylose fermentation in *Saccharomyces cerevisiae* is the subject of Chapter 5 of this thesis.

1.1.2 Alternative Fuel Molecules

The disadvantages of ethanol as a fuel molecule were discussed above. In recent years there have been efforts to engineer microbial biocatalysts for production of superior targets, including higher alcohols (Atsumi et al., 2008; Avalos et al., 2013; Sheppard et al., 2014), fatty acids (Steen et al., 2010), and alkanes (Schirmer et al., 2010). Relative to ethanol these molecules are more energy dense, more hydrophobic, and less corrosive. However at present the most widely used biofuel after ethanol is biodiesel. Biodiesel is today produced by transesterification of vegetable oils (e.g. rapeseed, palm, and soybean), animal fats, and waste oils, with vegetable oils accounting for the majority of biodiesel production (Gui et al., 2008). However, these first generation technologies for production of biodiesel from food oils have the same drawbacks as processes for production of ethanol from food crops described above. In order to produce enough biodiesel to meet a significant portion of demand for transportation energy without competing with the food supply, alternative sources of oil will be required.

It has been proposed that oils for biodiesel production could be derived from oleaginous microbes capable of accumulating lipids to high fractions of total biomass (Li et al., 2008; Meng et al., 2009; Ageitos et al., 2011). Some algae are capable of producing lipids from carbon dioxide via photosynthesis (Chisti, 2007), however because these microbes require light only low cell densities can be achieved. “Single cell oil” could also be produced by an oleaginous yeast, such as *Yarrowia lipolytica* (Tai and Stephanopoulos, 2013). A cost-effective process will require a biocatalyst capable of producing lipids at high yield, productivity, and titer, with the titer being determined by both cell density and lipid content (i.e. the fraction of the total cell mass which is lipid). Lipid overproduction from glucose in *Y. lipolytica* is the subject of Chapter 6 of this thesis.

As with ethanol, lipids could be produced by conversion of hexose sugars, but other feed stocks can also be envisioned. Several recent studies (Fei et al., 2011a, 2011b; Christophe et al., 2012; Fontanille et al., 2012) have explored the possibility of producing single cell oil from volatile fatty acids (VFAs), which can be obtained by fermentation of organics present in municipal solid waste and sewage sludge (Sans et al., 1995a, 1995b; Morgan-Sagastume et al., 2011) or syngas (Hu et al., 2013). A process for conversion of VFAs to biodiesel would address both challenges discussed above, generating a superior fuel molecule from an inexpensive feed stock. Production of lipids from acetate, the most abundantly available VFA, is the subject of Chapter 7.

1.2 ¹³C-Metabolic Flux Analysis

The following section is adapted from (Wasylenko and Stephanopoulos, 2013).

As discussed above, economical production of biofuels will require engineering of microbes with novel capabilities such as rapid utilization of xylose and overproduction of lipids at high yield, productivity, and final titer. Strain development for biofuels applications can be informed by metabolic flux estimation. Cells assimilate carbon, energy, and reducing power for growth and other essential processes through the reactions of central carbon metabolism. The magnitudes of the fluxes through these central metabolic reactions are tightly controlled, and represent the functional output of transcriptional, translational, and post-translational regulatory processes (Nielsen, 2003; Sauer, 2006). Consequently, in microbial systems knowledge of the distributions of metabolic flux in different strains or under different conditions has proven to be extremely powerful in basic physiological studies (Fischer and Sauer, 2005; Gombert et al., 2001). More importantly for the work in this thesis, experimentally determined metabolic flux distributions can be informative in biotechnological applications, for instance in understanding the genetic manipulations and metabolic rearrangements that might lead to increased production of a desired compound (Wittmann and Heinzle, 2002) or how metabolism changes over the course of a fermentation (Antoniewicz et al., 2007b).

Determination of intracellular metabolic fluxes is experimentally challenging. Direct measurement of these metabolic fluxes is impossible so that mathematical models are required to estimate the intracellular flux distribution from other directly measurable quantities (Schmidt et

al., 1999a). In the metabolite balancing approach, measured rates of consumption and production of extracellular metabolites (referred to as “extracellular flux measurements”), known stoichiometry of intracellular reactions, and a pseudo-steady state assumption on intracellular metabolite concentrations are used to construct a mass balance constraint for each metabolite in the metabolic network (Vallino and Stephanopoulos, 1993). However, in most systems of interest the number of unknown fluxes to be estimated exceeds the number of metabolites for which a mass balance equation can be written so that the resulting system of linear equations is underdetermined and an infinite number of solutions for the metabolic flux distribution exist. Optimization methods can be used to find the flux distribution that optimizes a specified objective function, the most commonly chosen objective function being maximum growth rate (Varma and Palsson, 1994). However, in many instances it is unclear that the flux distribution that maximizes growth rate will approximate the true flux distribution in the cell, especially for genetically engineered strains (Segrè et al., 2002; Fischer and Sauer, 2005), so that determining an appropriate objective function can be difficult. Moreover, the purely stoichiometric models used in metabolite balancing often rely on assumptions about redox cofactor usage and energy yields which are difficult to verify and can introduce significant errors into the estimated flux distributions (Marx et al., 1996; Schmidt et al., 1998).

Additional constraints on the metabolic flux distribution can be obtained through the application of isotopic tracers, substrates which are selectively “labeled” with rare, stable isotopes. The most commonly used stable isotope label is ^{13}C . The ^{13}C atoms from the tracer may be incorporated into any of the carbon atom positions of an intracellular metabolite. For a metabolite with n carbons in its backbone, there are 2^n possible carbon labeling patterns, each of

which corresponds to a different combination of ^{12}C and ^{13}C atoms at the various carbon atom positions. These 2^n species are referred to as the “isotopic isomers,” or “isotopomers,” of the metabolite. For each metabolite, the set of 2^n mole fractions describing the fraction of the total metabolite pool with each possible ^{13}C labeling pattern is referred to as the “isotopomer distribution” of that metabolite. If an appropriate choice is made for the tracer, these isotopomer distributions will be functions of the intracellular metabolic fluxes, so that information about metabolite labeling patterns can be used to add constraints to the stoichiometric models described above. Such information can in principle be obtained using ^1H nuclear magnetic resonance (NMR) (Marx et al., 1996), ^{13}C NMR (Sauer et al., 1997; Schmidt et al., 1999a; Szyperski, 1995), or mass spectroscopy (MS) (Christensen and Nielsen, 1999; Dauner and Sauer, 2000), although in practice MS has become the technology of choice for determination of ^{13}C labeling patterns because its sensitivity and precision greatly exceed those of NMR (Christensen et al., 2002; Wittmann and Heinzle, 1999). The supplementation of a stoichiometric model with additional constraints derived from measured isotopic labeling patterns typically yields a system that is significantly overdetermined, allowing for the determination of a unique flux solution without identification of an appropriate objective function or inclusion of often unreliable cofactor balances and assumed energy yields.

Over the past 20 years, rigorous quantitative methods have been developed for the estimation of the metabolic fluxes through central carbon metabolism using such stoichiometric models augmented with ^{13}C labeling data (Marx et al., 1996; Möllney et al., 1999; Schmidt et al., 1999a, 1999b; Wiechert et al., 1999, 1997; Wiechert and de Graaf, 1997). Flux estimation is achieved through an iterative procedure in which a putative metabolic flux distribution is generated and

the ^{13}C label distributions in intracellular metabolites that would result from this flux distribution are predicted. The putative metabolic fluxes are refined until the predicted ^{13}C metabolite labeling data match as closely as possible the ^{13}C labeling data obtained from experiment. The metabolic flux distribution that minimizes the lack-of-fit between simulated and measured ^{13}C metabolite labeling data is considered to be the best estimate for the true intracellular metabolic fluxes. This process of estimating intracellular metabolic fluxes from stoichiometric relations and ^{13}C -labeling data is referred to as ^{13}C -Metabolic Flux Analysis (MFA).

The key step in these flux estimation algorithms is the prediction of the ^{13}C labeling state that will result from a putative flux distribution. This typically involves the solution of a large set of isotopomer balance equations. Although the presence of bimolecular reactions results in non-linearities in these isotopomer balance equations, methods have been developed for the transformation of systems of non-linear isotopomer balance equations to cascades of systems of linear equations with the same information content so that the labeling state resulting from a specified flux distribution can always be determined (Antoniewicz et al., 2007a; Wiechert et al., 1999).

1.3 ^{13}C -MFA Mathematical Formulation

This section presents the mathematical formulation of ^{13}C -MFA and the notation used throughout this thesis. For the work described in this thesis, MS was used for measurement of metabolite ^{13}C labeling patterns. In some cases (especially with GC-MS), carbon-carbon bonds are broken during the MS ionization process and a fragment with only a subset of the carbons in

the metabolite backbone is detected. In this section, the term “metabolite” will be used to describe a metabolite fragment detected by MS, whether or not the fragment contains all of the carbon atoms present in the unfragmented metabolite.

MS resolves the isotopomers of a metabolite by molecular mass. In general, a metabolite isotopomer with i ^{13}C atoms will have a mass of $M + i$, where M is the molecular mass of the metabolite when all the carbon atoms are ^{12}C . Isotopomers of a metabolite with the same number of ^{13}C atoms cannot be distinguished by MS, even if the ^{13}C atoms are in different positions. Consequently, for MS ^{13}C labeling analysis it is useful to define the concept of a “mass isotopomer,” which is a set of all isotopomers of a metabolite with the same molecular mass (and equivalently, the same number of ^{13}C atoms). In general, a metabolite with n carbons will have $n + 1$ mass isotopomers ($i = 0, 1, 2, \dots, n$). For each metabolite, the set of $n + 1$ mole fractions describing the fraction of the total metabolite pool with each possible molecular mass will be referred to as the “Mass Isotopomer Distribution” (MID). (As an aside, the foregoing analysis ignores the effects of naturally occurring heavy isotopes of carbon and other elements in the metabolite. The presence of these naturally occurring heavy isotopes results in a “skewing” of the measured metabolite MIDs. Methods for correction for these effects are well-known, and so in the following these effects can be ignored without loss of generality (Fernandez et al., 1996; Wittmann and Heinzle, 1999; van Winden et al., 2002).)

The intracellular metabolite labeling state can be represented by a set of mole fractions \vec{y} in which each mole fraction y_i is equal to the fraction of a metabolite pool with a certain ^{13}C -labeling pattern. In the work in this thesis \vec{y} will represent the set of all metabolite MID mole

fractions measured by MS in a given experiment, but the following analysis could also apply to labeling data acquired by NMR. In the previous section it was noted that for a given tracer $\bar{\mathbf{y}}$ is a single-valued function of the metabolic fluxes $\bar{\mathbf{v}}$:

$$\bar{\mathbf{y}} = f(\bar{\mathbf{v}}) \quad (1.1)$$

and this function can always be evaluated by solving a set of balance equations. ^{13}C -MFA requires that the labeling state of the intracellular metabolites be measured in a Carbon Labeling Experiment (CLE). The metabolite labeling measurements will be referred to as $\bar{\mathbf{y}}^{meas}$. The flux estimation algorithm seeks to determine the flux distribution $\bar{\mathbf{v}}$ which minimizes the difference between $\bar{\mathbf{y}}$ and $\bar{\mathbf{y}}^{meas}$.

At metabolic steady state the rate of consumption must be equal to the rate of production for every intracellular metabolite. Thus the metabolic flux distribution is constrained by a set of metabolite mass balances which can be expressed using the stoichiometric matrix \mathbf{S} :

$$\mathbf{S} \cdot \bar{\mathbf{v}} = \bar{\mathbf{0}} \quad (1.2)$$

As noted above, the number of fluxes to be estimated generally exceeds the number of mass balance constraints (i.e. \mathbf{S} has more columns than rows) so that this system of equations is underdetermined. A unique flux distribution can be obtained by solving an optimization problem:

$$\min_{\bar{\mathbf{v}}} g(\bar{\mathbf{v}}) \quad (1.3)$$

$$\text{s. t. } \mathbf{S} \cdot \bar{\mathbf{v}} = \bar{\mathbf{0}}$$

$$\mathbf{A} \cdot \bar{\mathbf{v}} = \bar{\mathbf{b}}$$

$$\mathbf{C} \cdot \bar{\mathbf{v}} \leq \bar{\mathbf{d}}$$

where \mathbf{A} , $\vec{\mathbf{b}}$, \mathbf{C} , and $\vec{\mathbf{d}}$ can be chosen to impose equality or inequality constraints on the flux distribution (e.g. fixing a flux to a specified value or placing upper and lower bounds on specified fluxes). In Flux Balance Analysis (FBA) (Orth et al., 2010) the objective function $g(\vec{\mathbf{v}})$ is obtained from theory and is often chosen to solve for the flux distribution which maximizes the biomass yield on the carbon source. For instance, if the specific growth rate μ is fixed with an equality constraint, the objective function could be the carbon source uptake flux. Conversely, if the carbon source uptake flux is fixed then the objective function could be chosen to be $g(\vec{\mathbf{v}}) = -\mu$.

In ^{13}C -MFA, the objective function $g(\vec{\mathbf{v}})$ is the lack-of-fit between measured and simulated datasets. If the lack-of-fit is quantified by the sum of squared residuals between the model and data, the ^{13}C -MFA problem formulation can then be written as:

$$\begin{aligned} \min_{\vec{\mathbf{v}}} \quad & \|\vec{\mathbf{y}}^{meas} - f(\vec{\mathbf{v}})\|^2 & (1.4) \\ \text{s. t.} \quad & \mathbf{S} \cdot \vec{\mathbf{v}} = \vec{\mathbf{0}} \\ & \mathbf{A} \cdot \vec{\mathbf{v}} = \vec{\mathbf{b}} \\ & \mathbf{C} \cdot \vec{\mathbf{v}} \leq \vec{\mathbf{d}} \end{aligned}$$

Generally the rates of consumption and production of extracellular metabolites can also be calculated in the CLE by measuring the extracellular concentrations of substrates and products over time (e.g. by HPLC) and computing the time derivatives of the extracellular concentrations. In this case, the calculated ‘‘extracellular flux’’ data are also used in determining the most likely flux distribution. The formulation of the ^{13}C -MFA problem including both metabolite labeling and extracellular flux data is:

$$\begin{aligned}
\min_{\vec{v}} \quad & \sum_i \frac{(y_i^{meas} - y_i(\vec{v}))^2}{\sigma_i^2} + \sum_{k \in EF} \frac{(v_k^{meas} - v_k)^2}{\sigma_k^2} & (1.5) \\
\text{s. t.} \quad & \mathbf{S} \cdot \vec{v} = \vec{\mathbf{0}} \\
& \mathbf{A} \cdot \vec{v} = \vec{\mathbf{b}} \\
& \mathbf{C} \cdot \vec{v} \leq \vec{\mathbf{d}}
\end{aligned}$$

where EF is the subset of fluxes which are extracellular fluxes (i.e. fluxes in which a substrate crosses the cell membrane from the extracellular medium to the cytosol or a product crosses the cell membrane from the cytosol to the extracellular medium) measured in the CLE, v_k^{meas} is the value of flux k measured in the CLE, σ_i^2 is the variance (squared standard deviation) of y_i^{meas} , and σ_k^2 is the variance of v_k^{meas} . The formulation in equation (1.5) is used for all the ^{13}C -MFA in this thesis.

1.4 Thesis Overview: ^{13}C -MFA of Recombinant Yeasts for Biofuels Applications

This thesis describes the application of ^{13}C -MFA to two yeast systems relevant for biofuel production. First a method for Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) analysis of phosphorylated metabolites and an optimized method for extraction of intracellular metabolites were developed in order to facilitate acquisition of metabolite pool size and ^{13}C -labeling data for the ^{13}C -MFA studies (Chapter 2). Next an analysis of the expected measurement (Chapter 3) and modeling (Chapter 4) errors in ^{13}C -labeling data was undertaken in order to determine the precision with which metabolic fluxes can be modeled with these data.

Finally, using the methods described in Chapter 2 and insights gained from the work described in Chapters 3 and 4, ^{13}C -MFA was applied to the two systems of interest. In Chapter 5, the metabolism of a xylose-consuming *S. cerevisiae* strain was investigated on glucose and on xylose under both aerobic and anaerobic conditions in order to identify potential bottlenecks in the anaerobic fermentation of xylose to ethanol. In Chapters 6 and 7, ^{13}C -MFA was applied to the oleaginous yeast *Y. lipolytica* on glucose and acetate carbon sources, respectively, in order to study the metabolic rearrangements required for lipid overproduction in biodiesel applications and to determine the primary source of NADPH for lipid biosynthesis. FBA was also used to explore potential strategies for improving lipid yield on acetate. Conclusions and recommendations for future work are described in Chapter 8.

1.5 References

- Ageitos, J.M., Vallejo, J.A., Veiga-Crespo, P., Villa, T.G., 2011. Oily yeasts as oleaginous cell factories. *Appl. Microbiol. Biotechnol.* 90, 1219–1227. doi:10.1007/s00253-011-3200-z
- Almeida, J.R., Modig, T., Petersson, A., Hähn-Hägerdal, B., Lidén, G., Gorwa-Grauslund, M.F., 2007. Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biotechnol.* 82, 340–349. doi:10.1002/jctb.1676
- Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007a. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng.* 9, 68–86. doi:10.1016/j.ymben.2006.09.001

Antoniewicz, M.R., Kraynie, D.F., Laffend, L.A., González-Lergier, J., Kelleher, J.K., Stephanopoulos, G., 2007b. Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol. *Metab. Eng.* 9, 277–292. doi:10.1016/j.ymben.2007.01.003

Atsumi, S., Hanai, T., Liao, J.C., 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451, 86–89. doi:10.1038/nature06450

Avalos, J.L., Fink, G.R., Stephanopoulos, G., 2013. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* 31, 335–341. doi:10.1038/nbt.2509

Carroll, A., Somerville, C., 2009. Cellulosic biofuels. *Annu. Rev. Plant Biol.* 60, 165–182. doi:10.1146/annurev.arplant.043008.092125

Chisti, Y., 2007. Biodiesel from microalgae. *Biotechnol. Adv.* 25, 294–306. doi:10.1016/j.biotechadv.2007.02.001

Christensen, B., Gombert, A.K., Nielsen, J., 2002. Analysis of flux estimates based on (13)C-labelling experiments. *Eur. J. Biochem. FEBS* 269, 2795–2800.

Christensen, B., Nielsen, J., 1999. Isotopomer analysis using GC-MS. *Metab. Eng.* 1, 282–290. doi:10.1006/mben.1999.0117

Christophe, G., Deo, J.L., Kumar, V., Nouaille, R., Fontanille, P., Larroche, C., 2012. Production of Oils from Acetic Acid by the Oleaginous Yeast *Cryptococcus curvatus*. *Appl. Biochem. Biotechnol.* 167, 1270–1279. doi:10.1007/s12010-011-9507-5

Dauner, M., Sauer, U., 2000. GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing. *Biotechnol. Prog.* 16, 642–649. doi:10.1021/bp000058h

Fei, Q., Chang, H.N., Shang, L., Choi, J., 2011a. Exploring low-cost carbon sources for microbial lipids production by fed-batch cultivation of *Cryptococcus albidus*. *Biotechnol. Bioprocess Eng.* 16, 482–487. doi:10.1007/s12257-010-0370-y

Fei, Q., Chang, H.N., Shang, L., Choi, J., Kim, N., Kang, J., 2011b. The effect of volatile fatty acids as a sole carbon source on lipid accumulation by *Cryptococcus albidus* for biodiesel production. *Bioresour. Technol.* 102, 2695–2701. doi:10.1016/j.biortech.2010.10.141

Fernandez, C.A., Des Rosiers, C., Previs, S.F., David, F., Brunengraber, H., 1996. Correction of ¹³C mass isotopomer distributions for natural stable isotope abundance. *J. Mass Spectrom.* JMS 31, 255–262. doi:10.1002/(SICI)1096-9888(199603)31:3<255::AID-JMS290>3.0.CO;2-3

Fischer, E., Sauer, U., 2005. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat. Genet.* 37, 636–640. doi:10.1038/ng1555

Fontanille, P., Kumar, V., Christophe, G., Nouaille, R., Larroche, C., 2012. Bioconversion of volatile fatty acids into lipids by the oleaginous yeast *Yarrowia lipolytica*. *Bioresour. Technol.* 114, 443–449. doi:10.1016/j.biortech.2012.02.091

Gombert, A.K., Moreira dos Santos, M., Christensen, B., Nielsen, J., 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* 183, 1441–1451. doi:10.1128/JB.183.4.1441-1451.2001

- Gui, M.M., Lee, K.T., Bhatia, S., 2008. Feasibility of edible oil vs. non-edible oil vs. waste edible oil as biodiesel feedstock. *Energy* 33, 1646–1653. doi:10.1016/j.energy.2008.06.002
- Hahn-Hägerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I., Gorwa-Grauslund, M.F., 2007. Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.* 74, 937–953. doi:10.1007/s00253-006-0827-2
- Hu, P., Rismani-Yazdi, H., Stephanopoulos, G., 2013. Anaerobic CO₂ fixation by the acetogenic bacterium *Moorella thermoacetica*. *AIChE J.* 59, 3176–3183. doi:10.1002/aic.14127
- Li, Q., Du, W., Liu, D., 2008. Perspectives of microbial oils for biodiesel production. *Appl. Microbiol. Biotechnol.* 80, 749–756. doi:10.1007/s00253-008-1625-9
- Marx, A., de Graaf, A.A., Wiechert, W., Eggeling, L., Sahm, H., 1996. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 49, 111–129. doi:10.1002/(SICI)1097-0290(19960120)49:2<111::AID-BIT1>3.0.CO;2-T
- Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., Xian, M., 2009. Biodiesel production from oleaginous microorganisms. *Renew. Energy* 34, 1–5. doi:10.1016/j.renene.2008.04.014
- Möllney, M., Wiechert, W., Kownatzki, D., de Graaf, A.A., 1999. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. *Biotechnol. Bioeng.* 66, 86–103.
- Morgan-Sagastume, F., Pratt, S., Karlsson, A., Cirne, D., Lant, P., Werker, A., 2011. Production of volatile fatty acids by fermentation of waste activated sludge pre-treated in full-scale thermal hydrolysis plants. *Bioresour. Technol.* 102, 3089–3097. doi:10.1016/j.biortech.2010.10.054

- Nielsen, J., 2003. It is all about metabolic fluxes. *J. Bacteriol.* 185, 7031–7035.
- Orth, J.D., Thiele, I., Palsson, B.Ø., 2010. What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi:10.1038/nbt.1614
- Perlack, R.D., Wright, L.L., Turhollow, A.F., Stokes, B.J., Erbach, D.C., 2005. Biomass as a Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply. Oak Ridge National Laboratory.
- Sans, C., Mata-Alvarez, J., Cecchi, F., Pavan, P., Bassetti, A., 1995a. Volatile fatty acids production by mesophilic fermentation of mechanically-sorted urban organic wastes in a plug-flow reactor. *Bioresour. Technol.* 51, 89–96. doi:10.1016/0960-8524(95)95866-Z
- Sans, C., Mata-Alvarez, J., Cecchi, F., Pavan, P., Bassetti, A., 1995b. Acidogenic fermentation of organic urban wastes in a plug-flow reactor under thermophilic conditions. *Bioresour. Technol.* 54, 105–110. doi:10.1016/0960-8524(95)00098-4
- Sauer, U., 2006. Metabolic networks in motion: ¹³C-based flux analysis. *Mol. Syst. Biol.* 2. doi:10.1038/msb4100109
- Sauer, U., Hatzimanikatis, V., Bailey, J.E., Hochuli, M., Szyperski, T., Wüthrich, K., 1997. Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat. Biotechnol.* 15, 448–452. doi:10.1038/nbt0597-448
- Schirmer, A., Rude, M.A., Li, X., Popova, E., del Cardayre, S.B., 2010. Microbial biosynthesis of alkanes. *Science* 329, 559–562. doi:10.1126/science.1187936

Schmidt, K., Nielsen, J., Villadsen, J., 1999a. Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models. *J. Biotechnol.* 71, 175–189.

Schmidt, K., Nørregaard, L.C., Pedersen, B., Meissner, A., Duus, J.O., Nielsen, J.O., Villadsen, J., 1999b. Quantification of intracellular metabolic fluxes from fractional enrichment and ¹³C-¹³C coupling constraints on the isotopomer distribution in labeled biomass components. *Metab. Eng.* 1, 166–179. doi:10.1006/mben.1999.0114

Schmidt, Marx, de Graaf AA, Wiechert, Sahm, Nielsen, Villadsen, 1998. ¹³C tracer experiments and metabolite balancing for metabolic flux analysis: comparing two approaches. *Biotechnol. Bioeng.* 58, 254–257.

Segrè, D., Vitkup, D., Church, G.M., 2002. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15112–15117. doi:10.1073/pnas.232349399

Sheppard, M.J., Kunjapur, A.M., Wenck, S.J., Prather, K.L.J., 2014. Retro-biosynthetic screening of a modular pathway design achieves selective route for microbial synthesis of 4-methyl-pentanol. *Nat. Commun.* 5, 5031. doi:10.1038/ncomms6031

Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., Del Cardayre, S.B., Keasling, J.D., 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463, 559–562. doi:10.1038/nature08721

Stephanopoulos, G., 2007. Challenges in engineering microbes for biofuels production. *Science* 315, 801–804. doi:10.1126/science.1139612

Szyperski, T., 1995. Biosynthetically directed fractional ¹³C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *Eur. J. Biochem. FEBS* 232, 433–448.

Tai, M., Stephanopoulos, G., 2013. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metab. Eng.* 15, 1–9.
doi:10.1016/j.ymben.2012.08.007

Vallino, J.J., Stephanopoulos, G., 1993. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnol. Bioeng.* 41, 633–646.
doi:10.1002/bit.260410606

Van Maris, A.J.A., Winkler, A.A., Kuiper, M., de Laat, W.T.A.M., van Dijken, J.P., Pronk, J.T., 2007. Development of efficient xylose fermentation in *Saccharomyces cerevisiae*: xylose isomerase as a key component. *Adv. Biochem. Eng. Biotechnol.* 108, 179–204.
doi:10.1007/10_2007_057

Van Winden, W.A., Wittmann, C., Heinzle, E., Heijnen, J.J., 2002. Correcting mass isotopomer distributions for naturally occurring isotopes. *Biotechnol. Bioeng.* 80, 477–479.
doi:10.1002/bit.10393

Varma, A., Palsson, B.O., 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* 60, 3724–3731.

Wasylenko, T.M., Stephanopoulos, G., 2013. Kinetic isotope effects significantly influence intracellular metabolite (¹³C) labeling patterns and flux determination. *Biotechnol. J.* 8, 1080–1089. doi:10.1002/biot.201200276

Wiechert, W., de Graaf, A.A., 1997. Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments. *Biotechnol. Bioeng.* 55, 101–117. doi:10.1002/(SICI)1097-0290(19970705)55:1<101::AID-BIT12>3.0.CO;2-P

Wiechert, W., Möllney, M., Isermann, N., Wurzel, M., de Graaf, A.A., 1999. Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. *Biotechnol. Bioeng.* 66, 69–85.

Wiechert, W., Siefke, C., de Graaf, A.A., Marx, A., 1997. Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis. *Biotechnol. Bioeng.* 55, 118–135. doi:10.1002/(SICI)1097-0290(19970705)55:1<118::AID-BIT13>3.0.CO;2-I

Wisselink, H.W., Toirkens, M.J., Wu, Q., Pronk, J.T., van Maris, A.J.A., 2009. Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered *Saccharomyces cerevisiae* strains. *Appl. Environ. Microbiol.* 75, 907–914. doi:10.1128/AEM.02268-08

Wittmann, C., Heinzle, E., 2002. Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing corynebacteria. *Appl. Environ. Microbiol.* 68, 5843–5859.

Wittmann, Heinzle, 1999. Mass spectrometry for metabolic flux analysis. *Biotechnol. Bioeng.* 62, 739–750.

Chapter 2

Methods for Metabolite Extraction and LC-MS/MS Analysis

2.1 Introduction

¹³C-Metabolic Flux Analysis (MFA) requires accurate and precise metabolite labeling data for flux estimation. The earliest ¹³C-MFA studies employed NMR to quantify ¹³C labeling of amino acids obtained from protein hydrolysates (Szyperski, 1995; Marx et al., 1996). The pathways for biosynthesis of these proteinogenic amino acids are well known, and consequently the amino acid labeling information can be used to infer the labeling patterns of central carbon metabolite precursors (Szyperski, 1995; Maaheimo et al., 2001).

The principal advantage of using proteinogenic amino acids to obtain labeling data is that protein is very abundant so that relatively little biomass is required to obtain sufficient signal with either NMR or Mass Spectrometry (MS). However, there are several drawbacks associated with using proteinogenic amino acids as proxies for their central carbon metabolite precursors. The most obvious is that labeling information can only be obtained for metabolites which are substrates for amino acid biosynthesis. Data for other metabolites in the metabolic network are not available with this approach. The second major disadvantage is that this approach requires that the flux analysis be conducted while the cells are in a metabolic steady state and are actively synthesizing protein (i.e. during exponential growth in a batch culture or in a chemostat). For the protein labeling patterns to accurately reflect those of the central carbon metabolite precursors, the steady state must be maintained until nearly all the protein in the system has been synthesized in

the metabolic steady state being investigated. At this point the system approaches an “isotopic steady state” and biomass can be harvested.

An alternative to obtaining labeling data from proteinogenic amino acids is to measure the labeling patterns of the primary central carbon metabolites directly (van Winden et al., 2005). This approach allows collection of labeling data for a large range of metabolites whose labeling patterns are not accessible from proteinogenic amino acid data. It also makes possible the estimation of fluxes in situations where the cell is not actively synthesizing new amino acids and/or proteins. For the work described in this thesis labeling data from primary metabolite measurements were highly desirable. The investigation of xylose metabolism in *Saccharomyces cerevisiae* (Chapter 5) was focused on estimation of fluxes through glycolysis and the Pentose Phosphate Pathway (PPP). However proteinogenic amino acid labeling data yield little information about the metabolites of these pathways; the labeling patterns of metabolites in these pathways must be analyzed directly. For the study of lipid overproduction in *Yarrowia lipolytica* (Chapters 6 and 7), metabolic fluxes were quantified after nitrogen had been depleted from the growth medium and lipid accumulation had commenced. At this point in the fermentation, the yeast are not synthesizing amino acids (as there is a negligible amount of nitrogen in the medium), and so the amino acid labeling patterns will not accurately reflect the fluxes during the lipid accumulation phase. Consequently relevant labeling data could only be obtained through direct measurement of primary metabolites.

However the measurement of labeling patterns of primary metabolites is experimentally challenging. Central carbon metabolite pools are small and turn over quickly—on the order of

seconds to minutes (Canelas et al., 2008). Consequently, to ensure that metabolite pool size and labeling measurements are representative of the *in vivo* metabolic state, enzymatic reactions that consume and produce these metabolites must be quenched during sample preparation. Moreover, sensitive measurement techniques are needed to obtain sufficient signal for accurate and precise quantification of labeling patterns. For this reason, MS is generally preferable to NMR, which is far less sensitive (Wittmann and Heinzle, 1999; Christensen et al., 2002). While amino acid labeling patterns can be determined by Gas Chromatography-Mass Spectrometry (GC-MS), phosphorylated metabolites such as those in glycolysis and the PPP are typically too heavy to be analyzed by GC-MS; consequently a Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) system is required for their analysis.

To facilitate measurement of primary metabolite labeling patterns, we implemented a method for LC-MS/MS analysis of the phosphorylated metabolites of glycolysis and the PPP. The LC-MS/MS method also allows quantification of select cofactors that carry phosphate groups and are consequently inaccessible by GC-MS, including ATP, ADP, AMP, NAD(P)⁺, NAD(P)H, and acetyl-CoA. The quantification of these cofactors yielded important information in the study of xylose metabolism in *S. cerevisiae* (Chapter 5).

We also investigated several protocols for extraction of intracellular metabolites for quantification of pool sizes and ¹³C labeling patterns. In a landmark paper, de Koning and van Dam presented a method in which cell culture samples were added to 60% methanol at -40°C in order to quench metabolism prior to metabolite extraction (de Koning and van Dam, 1992). However, subsequent work suggested that contact with the methanol solution compromises the

cell membrane, resulting in leakage and incomplete recovery of intracellular metabolites (Villas-Bôas et al., 2005; Canelas et al., 2008). It has been proposed that addition of a salt to the quenching solution may reduce leakage of intracellular metabolites by decreasing the osmotic shock experienced by cells (Faijes et al., 2007; Villas-Bôas and Bruheim, 2007; Canelas et al., 2008; Spura et al., 2009). Ideally the salt should be volatile so that it can be removed from samples by evaporation during sample preparation (Faijes et al., 2007). Another study has suggested that pure methanol quenching solution results in less metabolite leakage than 60% methanol solution (Canelas et al., 2008). We compared three quenching solutions to determine which resulted in the highest metabolite yields (and presumably the least metabolite leakage) in metabolite extraction protocols. The quenching solutions were 60% (v/v) methanol in water, 10 mM ammonium acetate in 60% methanol (Ewald et al., 2009), and pure methanol.

In metabolite extraction protocols it is often desirable to have a washing step which removes medium components from the cell pellet prior to extraction. The washing step reduces the amounts of extracellular metabolites and salts in the samples to be analyzed by MS. High concentrations of extracellular metabolites and salts can produce large peaks which interfere with quantification of intracellular metabolite pool sizes and ^{13}C labeling patterns and negatively affect MS performance by suppressing ionization and dirtying the instrument, resulting in increased baseline noise. However, in the event that metabolites leak during sample preparation due to contact with the quenching solution, introduction of a washing step might be expected to increase metabolite leakage by increasing the amount of time cells are in contact with the quenching solution. We therefore compared three methods of washing cell pellets to determine which method resulted in the highest metabolite yields while also satisfactorily removing

extracellular medium components from samples: (1) pouring fresh washing solution (identical to the quenching solution) onto the cell pellet and then immediately decanting the solution; (2) adding fresh washing solution and resuspending the cell pellet by vigorous pipetting followed by centrifugation and removal of the washing solution; and (3) adding fresh washing solution and resuspending the cell pellet by vortexing followed by centrifugation and removal of the washing solution.

Numerous methods for extraction of intracellular metabolites have been presented in the literature (Villas-Bôas et al., 2005; Canelas et al., 2009). An effective metabolite extraction protocol should disrupt the cell membrane and cell wall, resulting in the release of intracellular metabolites into the extraction solution, and also avoid the degradation and interconversion of metabolites to be analyzed. We compared three metabolite extraction protocols to determine which method resulted in the highest metabolite yields: a hot ethanol (HE) extraction (Gonzalez et al., 1997; Canelas et al., 2009), a Chloroform-Methanol-Water (CMW) extraction similar to those used by (de Koning and van Dam, 1992; Villas-Bôas et al., 2005; Canelas et al., 2009), and an Acidic Acetonitrile-Methanol (AAM) extraction which has been reported to give high recovery of nucleotide triphosphates from *Escherichia coli* (Rabinowitz and Kimball, 2007).

Finally, the LC-MS/MS method and an optimized metabolite extraction protocol in which cells were quenched in pure methanol and washed by resuspending the cell pellet by vigorous pipetting prior to a hot ethanol extraction were validated by measuring the Mass Isotopomer Distributions (MIDs) of metabolites labeled to natural abundance and the glucose-6-

phosphate/fructose-6-phosphate (G6P/F6P) ratio and adenylate energy charge (AEC) in *S. cerevisiae*.

2.2 Materials and Methods

2.2.1 Optimization of Compound Parameters

Five compound parameters were optimized for each metabolite to be analyzed in the LC-MS/MS method. For each analyte a roughly 100 μ M standard solution was prepared and directly injected into an AB Sciex API 2000 Tandem Mass Spectrometer at a flow rate of 10 μ l/min using a syringe pump built into the MS. Each parameter to be optimized was then varied over the entire parameter range in order to determine the value which resulted in the greatest signal intensity for the metabolite being investigated. The dwell time was 0.61 s. The source/gas parameters were fixed at the following values:

Curtain Gas (CUR)	20
IonSpray Voltage (IS)	-4500
Collision Gas (CAD)	5
Temperature (TEM)	500
Ion Source Gas 1 (GS1)	65
Ion Source Gas 2 (GS2)	60

2.2.2 Optimization of Source/Gas Parameters

Source/gas parameters were optimized by repeatedly injecting a mixture of Fructose-1,6-bisphosphate (FBP), Glucose-6-phosphate (G6P), Glyceraldehyde-3-phosphate (GAP), and 3-

Phosphoglycerate (3PG) into the LC-MS/MS with different parameter values and determining which values gave the maximum signal intensity. The concentration of each metabolite was roughly 10 μ M and the injection volume was 5 μ l. The Agilent 1100 Series HPLC system was operated without a chromatographic column. The flow rate was maintained at 300 μ l/min and the mobile phase was a 50:50 mixture of solvent A (10 mM tributylamine with 15 mM acetic acid in water) and solvent B (methanol). The compound parameter values were chosen to be the optimal values determined by Optimization of Compound Parameters (section 2.2.1). During the optimization of each source/gas parameter the values of the other source/gas parameters were fixed at the values specified in the Optimization of Compound Parameters section.

2.2.3 Liquid Chromatography-Tandem Mass Spectrometry

Wherever relevant, dried metabolite extracts were resuspended in 80 μ l Millipore water. LC-MS/MS analysis was performed using two different methods. The first method used a HPLC flow rate (200 μ l/min) and solvent gradient identical to those employed by (Luo et al., 2007). The other used an increased flow rate of 300 μ l/min and a fast gradient similar to that employed by (Shastri, 2008; Young et al., 2011) to increase sample throughput. The solvent gradients for the slow and fast methods are presented in Tables 2-1 and 2-2, respectively. The HPLC column oven was held at 25 $^{\circ}$ C for both methods. The injection volume was 20 μ l. The majority of the data in this thesis were collected using the fast LC-MS/MS method, including the data for Comparison of Hot Ethanol and Chloroform-Methanol-Water Extraction Protocols with LC-MS/MS (2.2.7) and Measurement of Mass Isotopomer Distributions (MIDs) of Metabolites Labeled to Natural Abundance (2.2.8).

Each LC-MS/MS method was broken up into two “periods” in the Analyst software to allow a longer dwell time for each metabolite being analyzed. For the slow method, period 1 spanned 0-35 min and period 2 35-80 min. For the fast method period 1 and period 2 spanned 0–21.5 min and 21.5-42 min, respectively. The system was generally washed after each sample by injecting 20 µl water, and then allowed to equilibrate before injection of the subsequent sample. For the slow method, after injection of water the system was run at 10% A for 7.5 min, switched to 100% A over 0.1 min, and equilibrated at 100% A for 17.4 min (a total of 25 min elapsed between samples.) For the fast method, after injection of water the system was run at 0% A for 5 min, switched to 100% A over 1 min, and equilibrated at 100% A for 15 min (total of 21 min elapsed between samples.)

Table 2-1 Solvent gradient for slow LC-MS/MS method

Time (min)	% A
0	100
15	100
25	80
55	80
60	65
65	65
70	40
75	40
75.1	10
80	10

Table 2-2 Solvent gradient for fast LC-MS/MS method

Time (min)	% A
0	100
8	100
18	77.5
28	60
32	40
34	10
36	10
37	0
42	0

2.2.4 Strain and Culture Conditions

All experiments were conducted with a xylose-consuming *S. cerevisiae* strain similar to H131-A3-AL^{CS} (Zhou et al., 2012) cultivated at 30 °C. 5 ml aerobic starter cultures grown in aerobic culture tubes (BD Falcon, #352059) were inoculated from 15% glycerol -80 °C freezer stocks. The starter cultures were used to inoculate 50 ml shake flask cultures. For the Comparison of the Hot Ethanol and Chloroform-Methanol-Water Extraction Protocols with LC-MS/MS (2.2.7), the strain was cultured in minimal medium with 20 g/l xylose as sole carbon source and 6.7 g/l Yeast Nitrogen Base (YNB) without amino acids (Difco) as a source of salts, vitamins, and trace elements (YNBX medium). For Measurement of MIDs of Metabolites Labeled to Natural Abundance (2.2.8), the strain was cultured in minimal medium with 20 g/l glucose as sole carbon

source and 6.7 g/l YNB (YNBG medium). For all other experiments, starter cultures were grown in Yeast Extract-Peptone-Dextrose (YPD) medium and shake flask cultures in YNBG medium.

2.2.5 Comparison of Quenching Methods

Cultures were harvested at $OD_{600} \sim 1-1.4$. 5 ml culture was quenched in 25 ml quenching solution held at low temperature ($\sim -40\text{ }^{\circ}\text{C}$) in a cold ethanol bath. Samples were centrifuged for 5 min at 1540g and $-10\text{ }^{\circ}\text{C}$. The supernatants were decanted and cell pellets were washed by one of three methods with 25 ml cold ($\sim -40\text{ }^{\circ}\text{C}$) quenching solution.

A hot ethanol extraction (Gonzalez et al., 1997; Canelas et al., 2009) was used for extraction of intracellular metabolites. 5 ml hot ($80\text{ }^{\circ}\text{C}$) 75% (v/v) ethanol solution was added to each cell pellet and samples were vortexed for 30 s, incubated in an $80\text{ }^{\circ}\text{C}$ water bath for 3 min, and vortexed a second time for 30 s. Samples were briefly cooled in the cold ethanol bath and centrifuged for 5 min at 3270g and $-10\text{ }^{\circ}\text{C}$ to remove cell debris. Two 2 ml volumes of each metabolite extract were stored at $-80\text{ }^{\circ}\text{C}$ and subsequently evaporated under airflow to complete dryness using a Pierce Reacti-Therm III Heating/Stirring Module prior to GC-MS analysis. For each sample, one 2 ml volume was used for GC-MS analysis of TMS derivatives and the other for GC-MS analysis of TBDMS derivatives (see sections 2.2.10 and 2.2.11 below).

A total of nine shake flask cultures were prepared for metabolite extractions. Three samples were taken from each flask, resulting in a total of 27 samples. From each flask, one sample was quenched in each of the three quenching solutions being investigated: (1) pure methanol; (2) 60% methanol (v/v); and (3) 60% methanol containing 10 mM ammonium acetate. Three

different methods of washing cell pellets were investigated. Each method was used for three of the nine shake flasks. Thus in total, there were nine combinations of quenching solutions and cell pellet washing protocols. Three samples were prepared with each of the nine combinations. The three methods for washing cell pellets were: (1) add 25 ml quenching solution to the cell pellet and immediately decant quenching solution; (2) add 25 ml quenching solution to the cell pellet, resuspend cells by pipetting quenching solution up and down, centrifuge for 5 min at 1540g and -10 °C, and decant supernatant; (3) add 25 ml quenching solution to the cell pellet, resuspend cells by vortexing, centrifuge for 5 min at 1540g and -10 °C, and decant supernatant.

2.2.6 Comparison of Metabolite Extraction Protocols with GC-MS

Cultures were harvested at $OD_{600} \sim 1-1.4$. 5 ml culture was quenched in 25 ml pure methanol (Canelas et al., 2008) held at low temperature (~ -40 °C) in a cold ethanol bath. Samples were centrifuged for 5 min at 1540g and -10 °C and the supernatants decanted. Cell pellets were resuspended in 25 ml cold (~ -40 °C) methanol by pipetting. Samples were then centrifuged and supernatants decanted as before. Metabolites were extracted from cell pellets by one of three methods:

2.2.6.1 Hot Ethanol Extraction

The Hot Ethanol Extraction was performed as described above in the Comparison of Quenching Methods section (2.2.5).

2.2.6.2 Chloroform-Methanol-Water Extraction

A mixture containing 2.5 ml methanol, 2 ml Millipore water, and 5 ml chloroform was

added to each cell pellet and samples were vortexed 45 min at -20 °C. Samples were then centrifuged for 20 min at 770g and -10 °C to separate the polar and chloroform phases. The aqueous layer was collected and a solution containing 2 ml methanol and 2 ml Millipore water was added to the chloroform phase and samples were vortexed 30 s and centrifuged for 5 min at 3270g and -10 °C. The aqueous layer was collected and added to the first aqueous layer. The aqueous metabolite extract was split into two fractions for GC-MS analysis of TMS and TBDMS derivatives, respectively.

2.2.6.3 Acidic Acetonitrile Methanol Extraction

1 ml extraction solution (0.1 M formic acid in a 40:40:20 (by volume) mixture of acetonitrile:methanol:Millipore water) was added to each cell pellet. Samples were vortexed to resuspend cell pellets, incubated at -20 °C for 15 min, and centrifuged for 5 min at 3270g and -10 °C. The supernatant was collected and metabolites were extracted from the cell pellet two additional times by addition of 1 ml extraction solution, vortexing 20 s, incubation at -20 °C for 15 min, and centrifugation. Supernatants from all three extractions were pooled. Metabolite extracts were split into two equal fractions for GC-MS analysis of TMS and TBDMS derivatives.

Each extraction was performed with two biological replicate shake flask cultures with two technical replicates from each flask. Thus there was a total of four samples for each extraction protocol.

2.2.7 Comparison of Hot Ethanol and Chloroform-Methanol-Water Extraction Protocols with LC-MS/MS

Cultures were harvested at $OD_{600} \sim 1.1$. 2 ml culture was quenched in 10 ml pure methanol (Canelas et al., 2008) held at low temperature (~ -40 °C) in a cold ethanol bath. Samples were centrifuged for 5 min at 3270g and -10 °C and supernatants were discarded. Cell pellets were resuspended in 10 ml cold (~ -40 °C) methanol by pipetting. Samples were then centrifuged and the supernatants discarded as before. The hot ethanol extraction was performed as described above (section 2.2.5). The chloroform-methanol-water extraction was performed as described above (2.2.6.2) with a few modifications. The 45 min vortex step was conducted at 4 °C rather than -20 °C and samples were subsequently centrifuged at 3270g rather than 770g. Metabolite extracts were stored at -80 °C and subsequently dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module prior to LC-MS/MS analysis.

2.2.8 Measurement of Mass Isotopomer Distributions (MIDs) of Metabolites Labeled to Natural Abundance

The shake flask culture was harvested at $OD \sim 1$. 5 ml culture was quenched in 25 ml pure methanol held at low temperature (~ -40 °C) in a cold ethanol bath. Samples were centrifuged for 5 min at 3270g and -10 °C and supernatants were discarded. Cell pellets were resuspended in 25 ml cold (~ -40 °C) methanol by pipetting. Samples were then centrifuged and the supernatants discarded as before. Metabolites were extracted using the hot ethanol extraction protocol as described as above (section 2.2.5). The metabolite extracts were stored at -80 °C and subsequently dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module prior to LC-MS/MS analysis.

2.2.9 Validation of Protocols by Measurement of Metabolite Ratios

The quenching, hot ethanol extraction, and LC-MS/MS protocols were validated using data from the investigation of xylose metabolism in *S. cerevisiae* (Chapter 5). See Chapter 5 for relevant Materials and Methods information.

2.2.10 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of TMS Derivatives

Dried metabolite extracts were resuspended in 25 μ l 2% methoxyamine-hydrogen chloride in pyridine (Methoxamine (MOX) Reagent, Thermo Scientific #TS-45950) and incubated at 37 $^{\circ}$ C for 1.5 h. Following the methoximation reaction, 25 μ l *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) (Thermo Scientific #TS-48915) was added to each sample and samples were incubated at 37 $^{\circ}$ C for an additional 1.5 h. The resulting trimethylsilyl (TMS) derivatives were analyzed by GC/MS using an Agilent 6890N Network GC System coupled to an Agilent 5975B Inert XL MSD. Samples were injected with injection volume 1 μ l, split ratio 2:1, and inlet temperature 270 $^{\circ}$ C. Metabolites were separated on an Agilent J&W DB-35ms column (#122-3832, 30.0 m x 250 μ m x 0.25 μ m) with helium carrier gas at a flow rate of 1 ml/min. The GC column oven temperature was initially held at 80 $^{\circ}$ C for 6 min, ramped to 300 $^{\circ}$ C at 6 $^{\circ}$ C/min, held at 300 $^{\circ}$ C for 10 min, ramped to 320 $^{\circ}$ C at 10 $^{\circ}$ C/min, and held at 320 $^{\circ}$ C for 2 min, resulting in a total run time of 56.67 min. The MS was operated in electron ionization mode with an electron energy of 69.9 eV and source and quadrupole temperatures of 230 $^{\circ}$ C and 150 $^{\circ}$ C, respectively. Mass spectra were acquired by scanning the mass range 100-700 m/z.

2.2.11 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of TBDMS Derivatives

Dried metabolite extracts were resuspended in 20 μ l 2% methoxyamine-hydrogen chloride in pyridine and incubated at 37 °C for 2.5 h. 30 μ l *N*-*tert*-Butyldimethylsilyl-*N*-methyltrifluoroacetamide (MTBSTFA) with 1% *tert*-Butyldimethylchlorosilane (TBDMCS) (Thermo Scientific #TS-48927) was added to each sample and samples were incubated at 55 °C for 1 h. The resulting *tert*-butyldimethylsilyl (TBDMS) derivatives were analyzed by GC-MS. The GC-MS conditions were the same as for analysis of TMS derivatives, with the following exceptions: The GC column oven was initially held at 100 °C for 1 min, ramped to 105 °C at 2.5 °C/min, held at 105 °C for 2 min, and ramped to 300 °C at 3.5 °C/min and then to 320 °C at 5 °C/min, resulting in a run time of 64.71 min. The mass range scanned was 100-600 m/z.

2.3 Results and Discussion

2.3.1 Optimization of Compound Parameters

In order to maximize the sensitivity of the LC-MS/MS method, a set of compound parameters was optimized for each metabolite to be analyzed. The optimal values of the compound parameters are independent of the HPLC conditions and consequently can be determined for each analyte by direct injection of a standard solution into the MS. The compound parameters are then varied in order to find the set of parameters which results in maximum signal intensity. The MS was operated in Multiple Reaction Monitoring (MRM) mode such that each analyte must pass through two mass filters in order to reach the detector. The ions which enter the MS

are first filtered to select for parent ions of a desired mass-to-charge ratio (Q1 mass). The ions that pass through this initial filter are then fragmented in a collision cell and the fragments are filtered a second time to select for product ions of a desired mass-to-charge ratio (Q3 mass). The application of two mass filters in series results in greater selectivity and reduces background noise. When operating in MRM mode, the API 2000 requires optimization of five parameters for each metabolite to be detected:

Declustering Potential (DP)

Focusing Potential (FP)

Entrance Potential (EP)

Collision Energy (CE)

Collision Cell Exit Potential (CXP)

The DP, FP, and EP parameters are voltages which influence the filtering of the parent ions (Q1). The CE parameter determines how much energy is added to parent ions to facilitate fragmentation in the collision cell. The CXP parameter is a voltage which influences the filtering of the product ions (Q3). The optimized parameter values for each metabolite are listed in Table 2-3.

Table 2-3 Optimal Metabolite Compound Parameters

Metabolite	Q1 mass	Q3 mass	U- ¹³ C-labeled Q1 mass	U- ¹³ C-labeled Q3 mass	DP	FP	EP	CE	CXP
G6P	259	97	265	97	-13.2	-375	-4	-20	-6
F6P	259	97	265	97	-13.2	-375	-4	-20	-6
R5P	229	97	234	97	-10	-315	-5.5	-16	-6
Ru5P	229	97	234	97	-10	-315	-5.5	-16	-6
X5P	229	97	234	97	-10	-315	-5.5	-16	-6
DHAP	169	97	172	97	-15	-324	-3	-13	-6
GAP	169	97	172	97	-15	-324	-3	-13	-6
E4P	199	97	203	97	-19	-340	-3	-14	-6
S7P	289	97	296	97	-20	-380	-4	-25	-5
6PG	275	79	281	79	-12	-270	-5.5	-20	-5.5
FBP	339	97	345	97	-20	-340	-6	-26	-6
PEP	167	79	170	79	-14	-300	-3.5	-17	-12
3PG	185	79	188	79	-17	-290	-4	-39	-5
NAD ⁺	662.1	79	683.2	79	-4	-380	-6	-120	-12
NADP ⁺	742.1	79	763.1	79	-8	-365	-7.5	-124	-12
NADH	664.1	79	685.2	79	-75	-360	-9	-115	-11
NADPH	744.1	79	765.2	79	-70	-300	-9.5	-120	-11
ATP	506	79	516	79	-50	-370	-7	-110	-11
ADP	426	79	436.1	79	-28	-370	-8	-55	-12
AMP	346.1	79	356.1	79	-26	-365	-7	-42	-10
AcCoA	808.1	79	831.2	79	-75	-200	-10	-115	-4
Pyr	87	43	90	45	-16	-300	-3	-12	-8
Cit	191	87	197	90	-17	-320	-5	-22	-5.5
Suc	117	73	121	76	-15	-300	-4.5	-15.5	-6
Fum	115	71	119	74	-17	-305	-6	-10	-4
Mal	133	115	137	119	-15	-305	-4.5	-15.5	-8
AKG	145	101	150	105	-15	-280	-3	-13	-5.5
OAA	131	87	135	90	-23	-265	-7	-11	-6
Glx	73	45	75	46	-20	-315	-4	-11	-7

“U-¹³C-labeled Q1 mass” and “U-¹³C-labeled Q3 mass” indicate the Q1 and Q3 masses for metabolites in which all carbon atoms are ¹³C. Metabolite abbreviations: G6P, Glucose 6-phosphate; F6P, Fructose 6-phosphate; R5P, Ribose 5-phosphate; Ru5P, Ribulose 5-phosphate; X5P, Xylulose 5-phosphate; DHAP, Dihydroxyacetone phosphate; GAP, Glyceraldehyde 3-phosphate; E4P, Erythrose-4-phosphate; S7P, Sedoheptulose-7-phosphate; 6PG, 6-

Phosphogluconate; FBP, Fructose 1,6-bisphosphate; PEP, Phosphoenolpyruvate; 3PG, 3-Phosphoglycerate; NAD⁺, Nicotinamide adenine dinucleotide (oxidized); NADP⁺, Nicotinamide adenine dinucleotide phosphate (oxidized); NADH, Nicotinamide adenine dinucleotide (reduced); NADPH, Nicotinamide adenine dinucleotide phosphate (reduced); ATP, Adenosine triphosphate; ADP, Adenosine diphosphate; AMP, Adenosine monophosphate; AcCoA, Acetyl-CoA; Pyr, Pyruvate; Cit, Citrate; Suc, Succinate; Fum, Fumarate; Mal, Malate; AKG, α -Ketoglutarate; OAA, Oxaloacetate; Glx, Glyoxylate

2.3.2 Optimization of Source/Gas Parameters

The optimal values of the source/gas parameters are dependent on the HPLC conditions but independent of the analyte being measured. In order to determine the optimal values of the source/gas parameters, a mixture of Fructose-1,6-bisphosphate (FBP), Glucose-6-phosphate (G6P), glyceraldehyde-3-phosphate (GAP), and 3-phosphoglycerate (3PG) was analyzed with different values of each source/gas parameter. The five parameters to be optimized were:

Curtain Gas (CUR)

IonSpray Voltage (IS)

Collision Gas (CAD)

Ion Source Gas 1 (GS1)

Ion Source Gas 2 (GS2)

The Temperature (TEM) was chosen to be 500 and was not varied.

The API 2000 manufacturer AB Sciex recommends that the curtain gas parameter CUR be set to the highest value which does not sacrifice sensitivity in order to keep the system as clean as possible. The other source/gas parameters are chosen so as to maximize the signal-to-noise ratio. The dependences of the analyte peak areas on the source/gas parameters are shown in Figures 2-1, 2-2, 2-3, 2-4, and 2-5. The optimal parameter values are listed in Table 2-4.

The CUR parameter was chosen to be 30, as this was the highest value which did not result in a decrease in signal intensity. The optimal values of IS and GS2 were found to be -1500 and 80, respectively. All values of the CAD parameter greater than 2 that were tested resulted in similar signal intensities; the value of the CAD parameter was chosen to be 4. The signal intensity was relatively insensitive to the GS1 parameter, which was chosen to be 50.

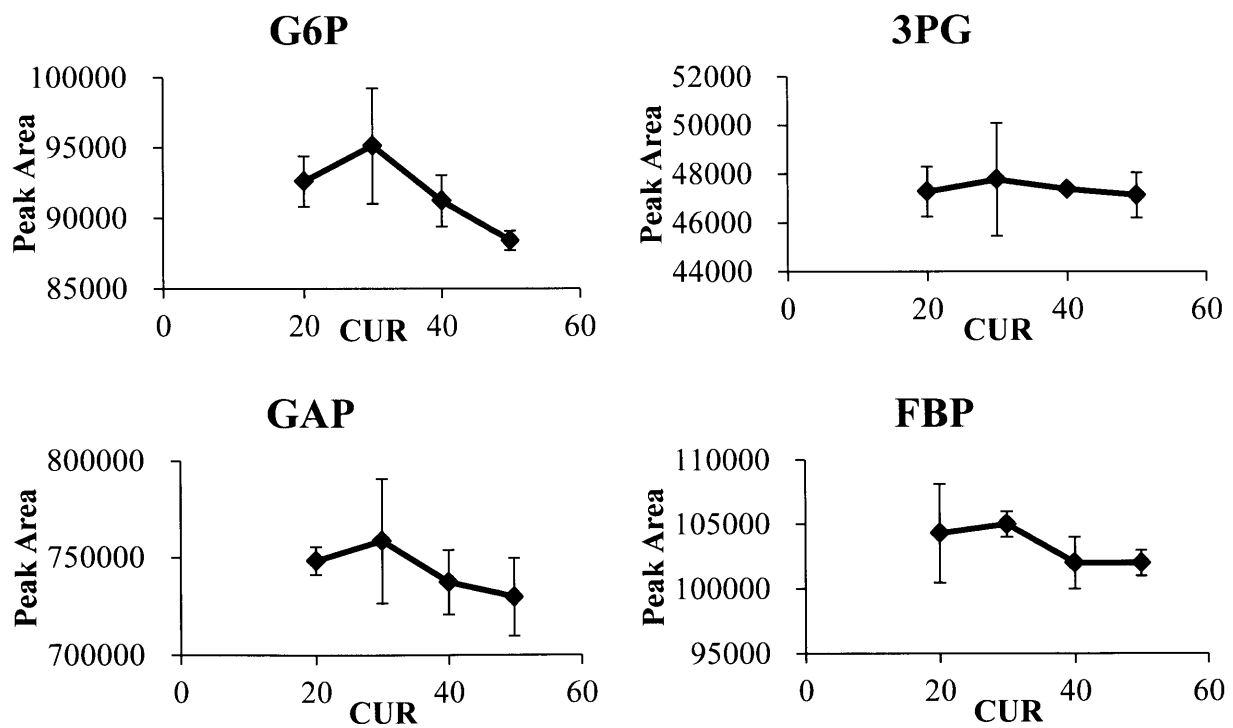


Figure 2-1 Curtain Gas (CUR) optimization

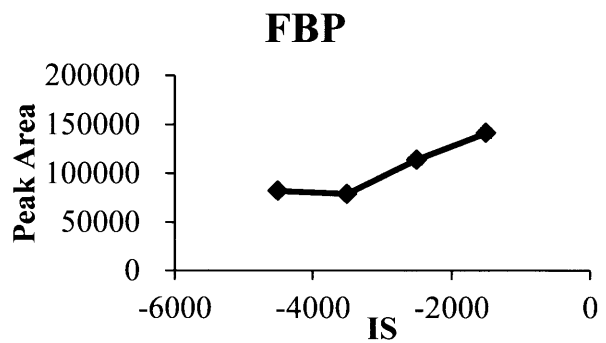
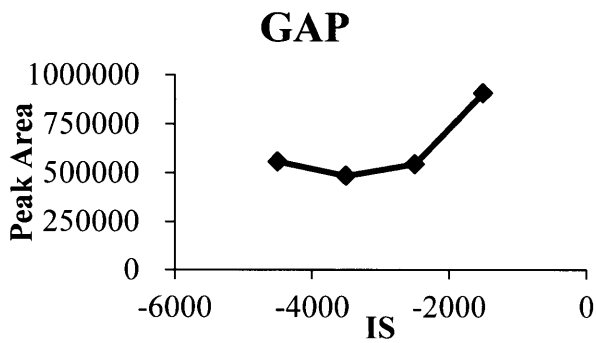
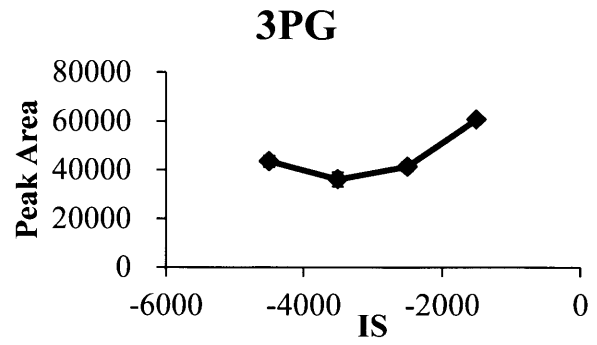
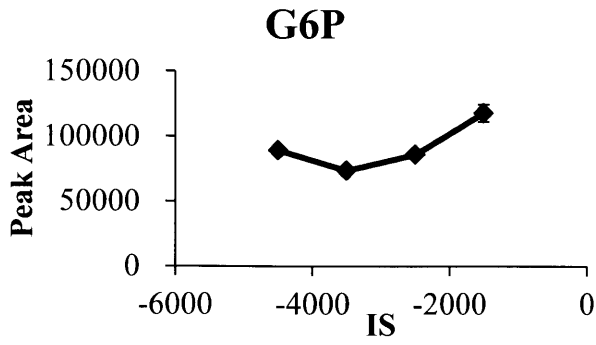


Figure 2-2 IonSpray Voltage (IS) optimization

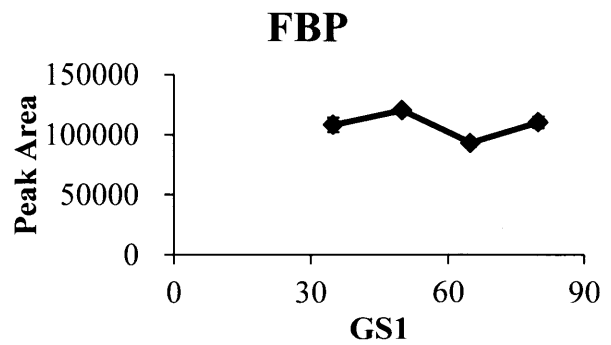
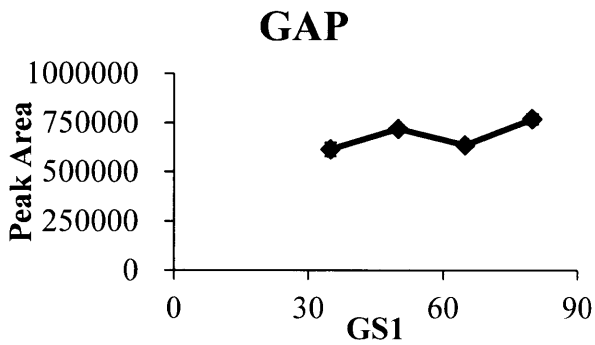
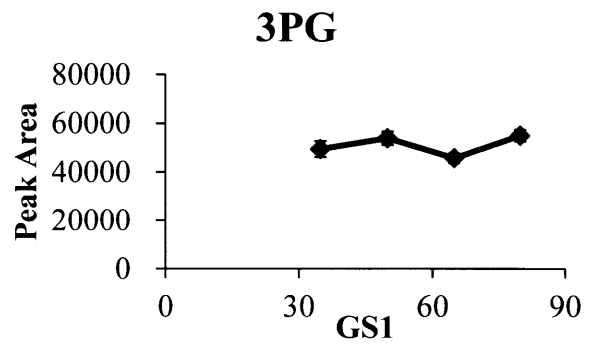
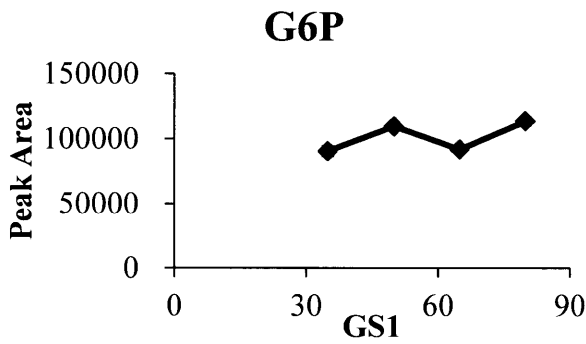


Figure 2-3 Ion Source Gas 1 (GS1) optimization

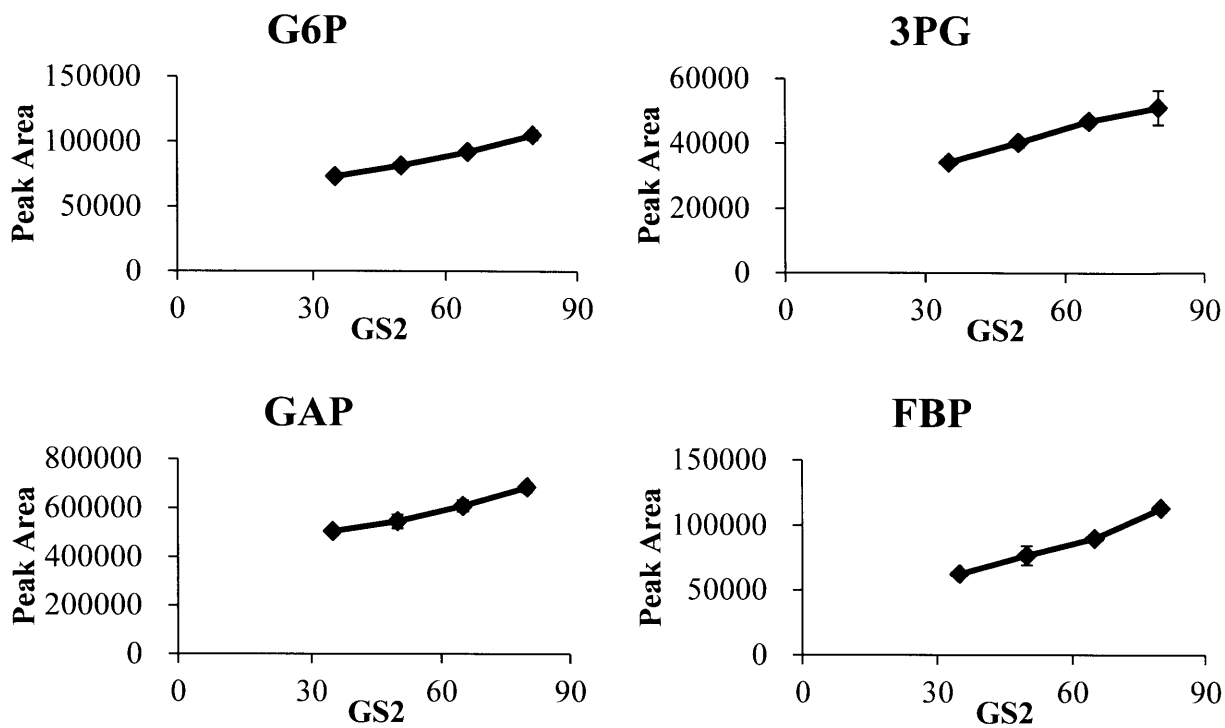


Figure 2-4 Ion Source Gas 2 (GS2) optimization

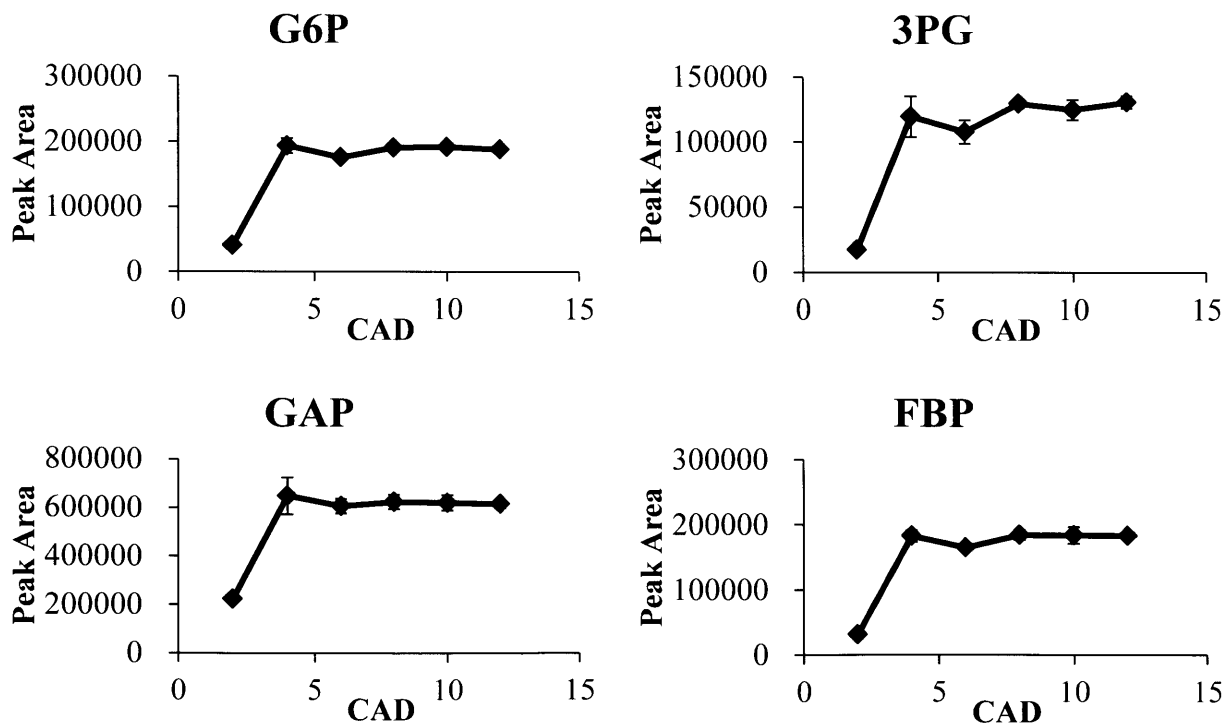


Figure 2-5 Collision Gas (CAD) optimization

Table 2-4 Optimal Source/Gas parameter values

Parameter	Optimal Value
CAD	4
CUR	30
GS1	50
GS2	80
IS	-1500

2.3.3 Liquid Chromatography-Tandem Mass Spectrometry

Two different HPLC methods were tested for LC-MS/MS analysis, as described in the Materials and Methods section. The first method uses a flow rate of 200 $\mu\text{l}/\text{min}$ and has a run time of 80 min for each sample (not including equilibration of the column). The second method uses a flow rate of 300 $\mu\text{l}/\text{min}$ and has a run time of 42 min per sample. Similar metabolite resolution was achieved with both methods, and so the fast method was used for the majority of the LC-MS/MS analysis conducted for this thesis. Approximate metabolite retention times for each method are listed in Table 2-5.

Table 2-5 Typical metabolite retention times (in minutes) for fast and slow HPLC methods

Metabolite	RT (fast)	RT (slow)
G6P	8.8-11.5	10.9-14.1
F6P	10.5-13.7	12.9-16.2
R5P	10.0-13.0	12.1-15.6
Ru5P	13.1-15.8	16.1-20.2
X5P	12.6-15.5	16.1-19.4
DHAP	14.8-18.1	22.7-25
GAP	10.0-12.0	13-16
E4P	N/A	N/A
S7P	10.2-13.3	12.5-15.9
6PG	26.2-29.0	41.2-49.3
FBP	26.2-29.5	52.3-59
PEP	28.6-30.8	58.2-65.8
3PG	26.2-29	49.2-53
NAD ⁺	19.4-20.6	30.3
NADP ⁺	28.6-30.5	62
NADH	30.4-32.4	68.7
NADPH	34.2-36.0	75.3
ATP	34.2-35.1	75.6
ADP	29.9-31.8	67.4
AMP	22.6-23.7	34.8
AcCoA	35.6-36.4	76.7
Pyr	15.9-18	22.1-25.3
Cit	29.7-30.2	N/A
Suc	23-24.1	34.5-36.2
Fum	24.5-26.2	36.6-40.2
Mal	25.5-26.2	38.6-41.7
AKG	26.0-27.5	41-42.3
OAA	N/A	N/A
Glx	N/A	N/A

2.3.4 Comparison of Quenching Methods

Three different quenching solutions (60% (v/v) methanol in water (MW), 10 mM ammonium acetate in 60% methanol (AAMW), and pure methanol (PM)) and three different washing

methods ((1) pouring fresh washing solution (identical to the quenching solution) onto the cell pellet and then immediately decanting the solution (“Pour”); (2) adding fresh washing solution and resuspending the cell pellet by vigorous pipetting followed by centrifugation and removal of the washing solution (“Pip”); and (3) adding fresh washing solution and resuspending the cell pellet by vortexing followed by centrifugation and removal of the washing solution (“Vor”)) were investigated to determine the quenching solution and washing method which gave the highest metabolite yields. Hot ethanol metabolite extractions were performed with all nine possible combinations of the three quenching solutions and three washing methods.

Pure methanol was found to be the best quenching solution, in agreement with (Canelas et al., 2008). When cells were washed by resuspending the cell pellet by pipetting, pure methanol resulted in the highest metabolite yield for all 22 TBDMS metabolite derivatives that were detected in all samples. Results for four representative metabolites are shown in Figure 2-6. When cells were washed with the vortexing method, pure methanol gave the highest yield for 19/20 metabolites, and when cells were washed by the immediate decanting method pure methanol gave the highest yield for 5/8 metabolites. Thus pure methanol was the best quenching solution independent of the method for washing cell pellets.

The pipetting method was chosen as the best method for washing cell pellets. With pure methanol as the quenching solution, the pipetting method gave the highest yield for 18 of 21 metabolites detected in all samples. (The vortexing method gave the highest yield for the other three metabolites.) However the differences were generally small and not statistically significant (Figure 2-7). Vortexing is the harsher of the two methods for resuspension of cell pellets and

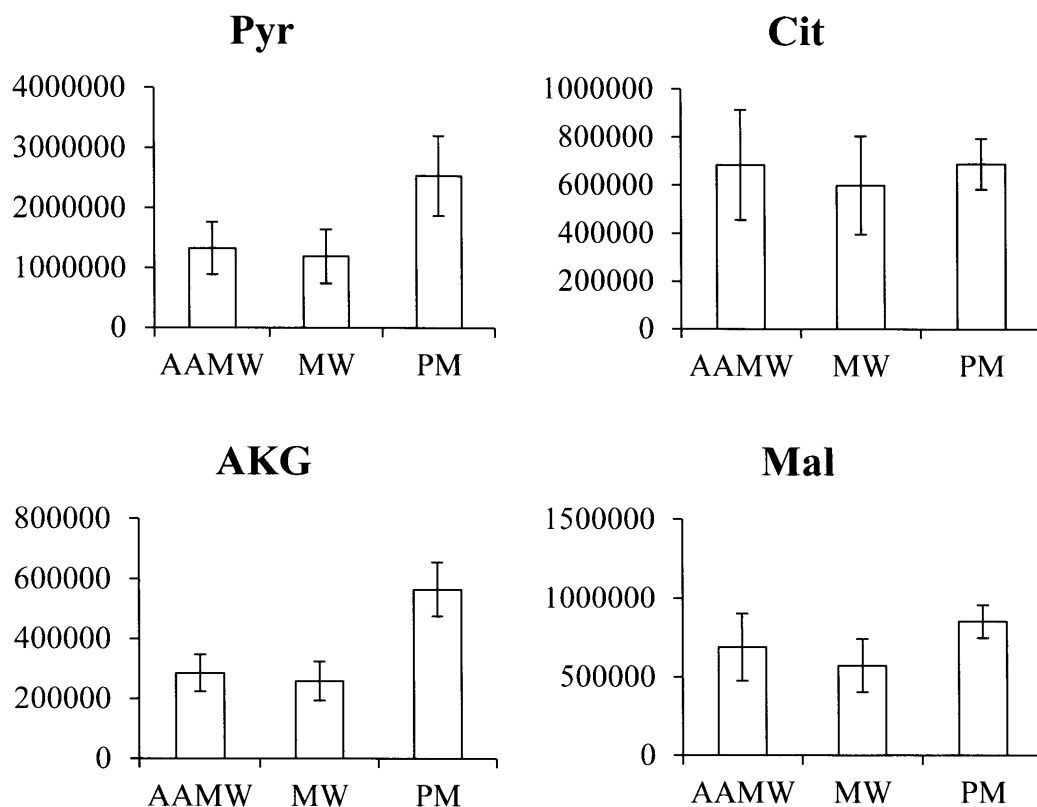


Figure 2-6 Comparison of quenching solutions using the pipetting method for washing pellets. The y-axis is in units of peak area/OD. Abbreviations: AAMW, 10 mM Ammonium Acetate in 60% Methanol; MW, 60% (v/v) Methanol in Water; PM, Pure Methanol.

might therefore result in greater disruption of the cell membrane and cell wall. The immediate decanting method resulted in the lowest yield for the majority of metabolites. This may be because this method did not employ a second centrifugation step. Consequently the pouring of washing solution onto the cell pellet may have disturbed the pellet and resuspended a small fraction of cells, which would then have been discarded along with the washing solution. The immediate decanting method would also be expected to be least effective in removing

extracellular metabolites and salts; since the cell pellet is not resuspended in this method, medium trapped inside the cell pellet will not be removed.

Quenching in pure methanol followed by washing the cells with the pipetting method was selected as the optimal quenching protocol. This quenching protocol was used for the comparison of metabolite extraction protocols.

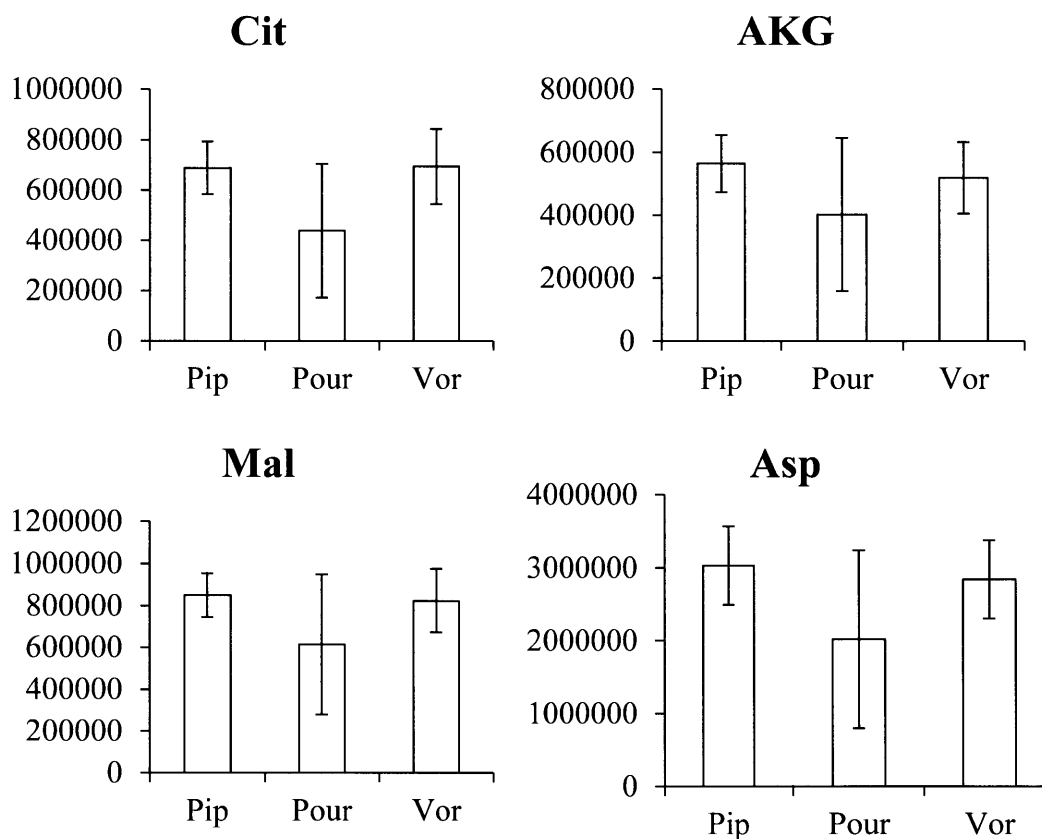


Figure 2-7 Comparison of methods for washing cell pellets with pure methanol quenching solution. The y-axis is in units of peak area/OD. Abbreviations: Pip, resuspension by pipetting; Pour, immediate decanting without resuspension; Vor, resuspension by vortexing.

2.3.5 Comparison of Metabolite Extraction Protocols with GC-MS

Three different metabolite extraction protocols were compared. The hot ethanol extraction (Gonzalez et al., 1997; Canelas et al., 2009) is relatively fast and simple to execute and does not require toxic organic solvents. However, it utilizes high temperatures which may reduce recovery of labile metabolites. The chloroform-methanol-water protocol (de Koning and van Dam, 1992; Villas-Bôas et al., 2005; Canelas et al., 2009) allows samples to be kept cold throughout and has an additional advantage in that it allows recovery of lipid species from the chloroform phase. However, this protocol is relatively time-consuming and generates a biphasic waste stream which contains the hazardous organic solvents chloroform and methanol. The acidic acetonitrile-methanol extraction (Rabinowitz and Kimball, 2007) had been previously shown to result in high recovery of nucleotide triphosphates, suggesting it might perform well in extraction of sugar-phosphates from Glycolysis and the PPP. However, this protocol had been developed for *E. coli* cells, which are generally easier to lyse than yeast, and consequently needed to be tested in *S. cerevisiae*.

For TBDMS metabolite derivatives (generally organic acids and amino acids) the hot ethanol and acidic acetonitrile-methanol methods performed well. The hot ethanol extraction resulted in the highest metabolite yields for 11 of 21 TBDMS derivatives detected in all samples, while the acidic acetonitrile-methanol method gave the highest yields for the other 10. Results for a few organic acids from the Tricarboxylic Acid (TCA) Cycle are shown in Figure 2-8.

The results were different for TMS derivatives of sugar-phosphates from glycolysis and the PPP (Figure 2-9). For some of these metabolites the hot ethanol extraction gave the highest yields,

while for others the chloroform-methanol-water protocol was best. The acidic acetonitrile-methanol method generally gave the poorest results for these metabolites, despite the fact that it had been reported to give high yields for nucleotide triphosphates in *E. coli*. As the metabolites of glycolysis and the PPP were of particular interest for the investigation of xylose metabolism, the acidic acetonitrile-methanol method was eliminated from consideration. A second experiment was then conducted to compare the hot ethanol and chloroform-methanol-water extraction protocols using LC-MS/MS for metabolite analysis.

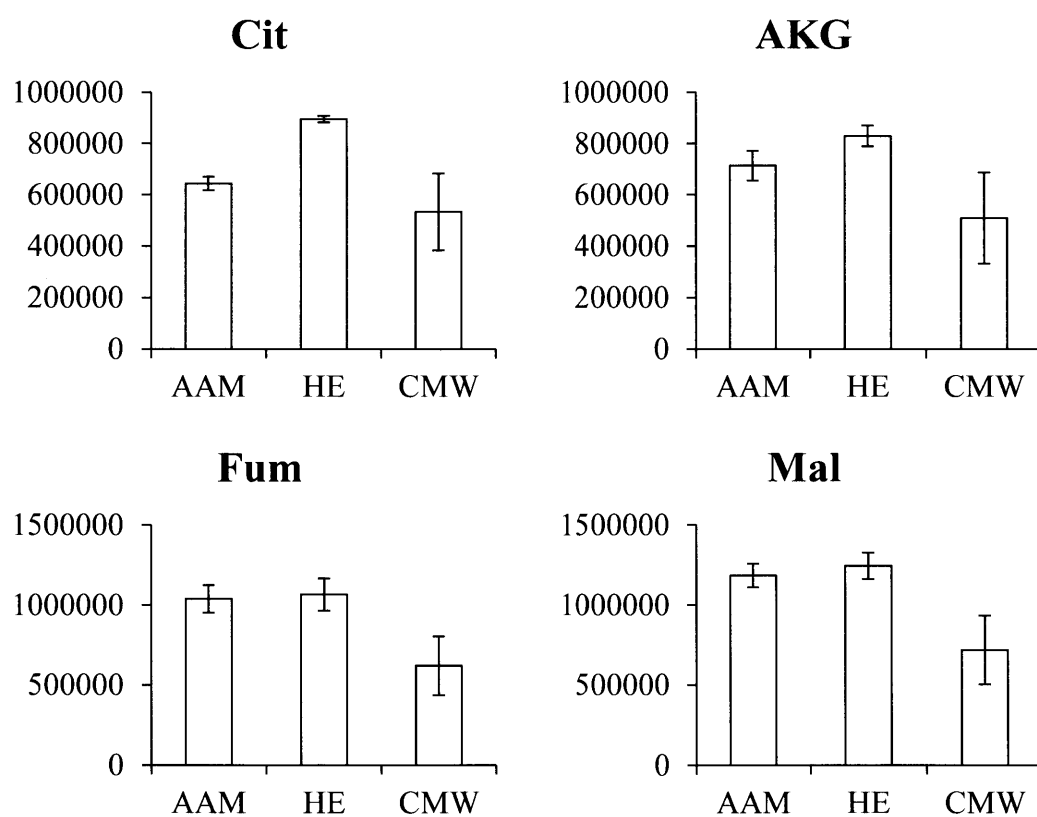


Figure 2-8 Comparison of extraction protocols by GC-MS with TBDMS derivatives. Abbreviations: AAM, Acidic Acetonitrile-Methanol; HE, Hot Ethanol; CMW, Chloroform-Methanol-Water.

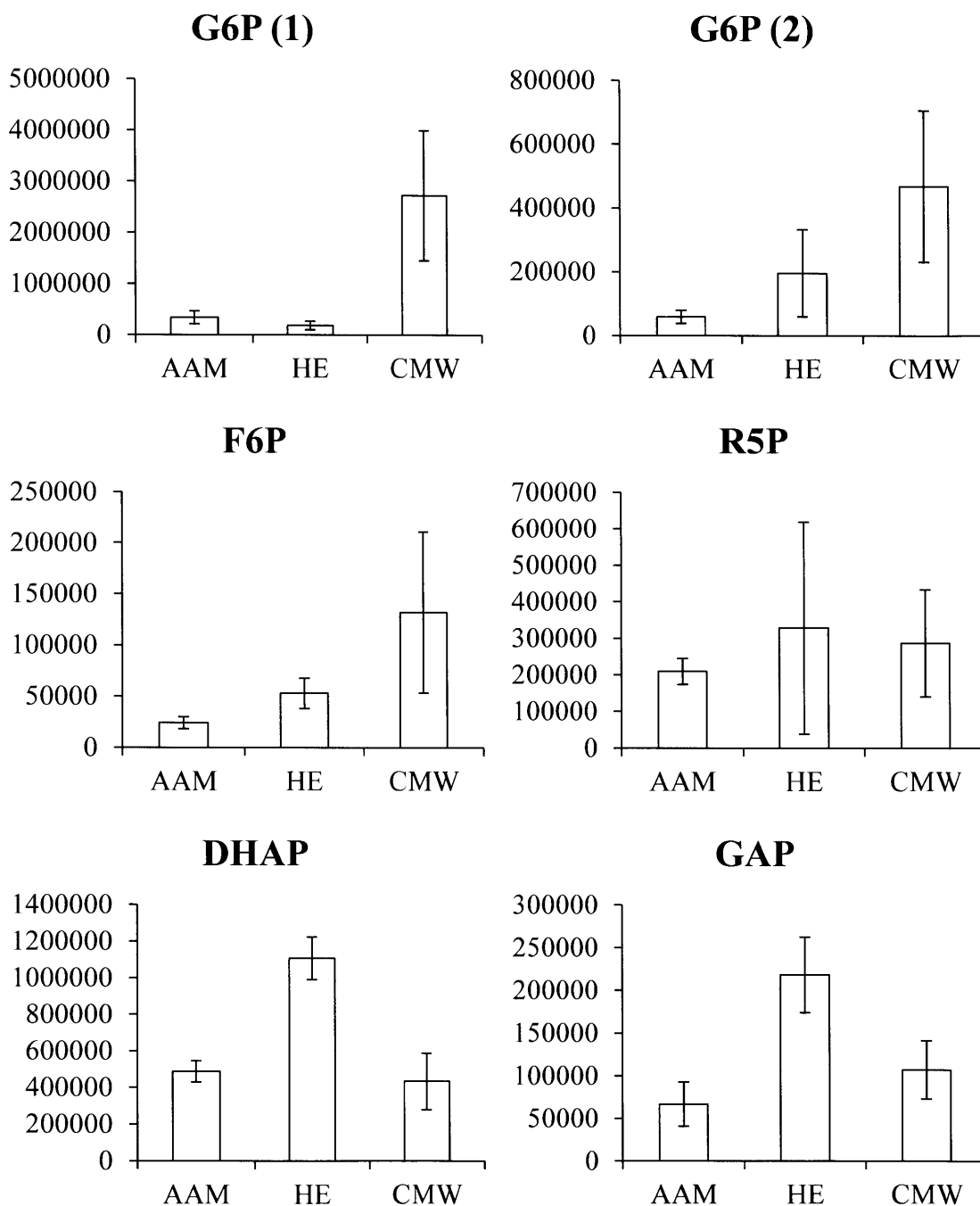


Figure 2-9 Comparison of extraction protocols by GC-MS with TMS derivatives. Abbreviations: AAM, Acidic Acetonitrile-Methanol; HE, Hot Ethanol; CMW, Chloroform-Methanol-Water.

2.3.6 Comparison of Hot Ethanol and Chloroform-Methanol-Water Extraction Protocols with LC-MS/MS

GC-MS analysis of TMS derivatives allowed detection of only a small subset of the phosphorylated metabolites from glycolysis and the PPP, and the signals for these metabolites were low and noisy. Therefore, a second experiment was conducted using LC-MS/MS to compare the metabolite yields of the hot ethanol and chloroform-methanol-water protocols. The results are shown in Figure 2-10. It can be seen that the hot ethanol extraction gave higher metabolite yields for the unphosphorylated organic acids of the TCA Cycle while the chloroform-methanol-water extraction gave higher yields for the majority of sugar phosphates. However the differences were generally small and not statistically significant. It is possible that the high temperatures in the hot ethanol protocol result in slight degradation of some of the phosphorylated metabolites, resulting in lower yields.

Nonetheless, the hot ethanol protocol was selected for metabolite extractions for this thesis. The hot ethanol extraction protocol is simpler and takes less time (45 min as opposed to almost two hours), making it more amenable to time courses and experiments in which multiple cultures must be harvested in rapid succession. It also has fewer safety concerns and avoids the use of chloroform. Although there may be minor degradation of some phosphorylated metabolites, this will not affect metabolite ^{13}C -labeling measurements as long as all isotopomers of a metabolite degrade at the same rate, which would seem to be a good assumption. Metabolite pool sizes can still be quantified accurately if uniformly ^{13}C -labeled metabolite extract is used as internal standard (Wu et al., 2005), as in the study of xylose metabolism in *S. cerevisiae* (Chapter 5), again assuming that different isotopomers of a metabolite degrade at the same rate.

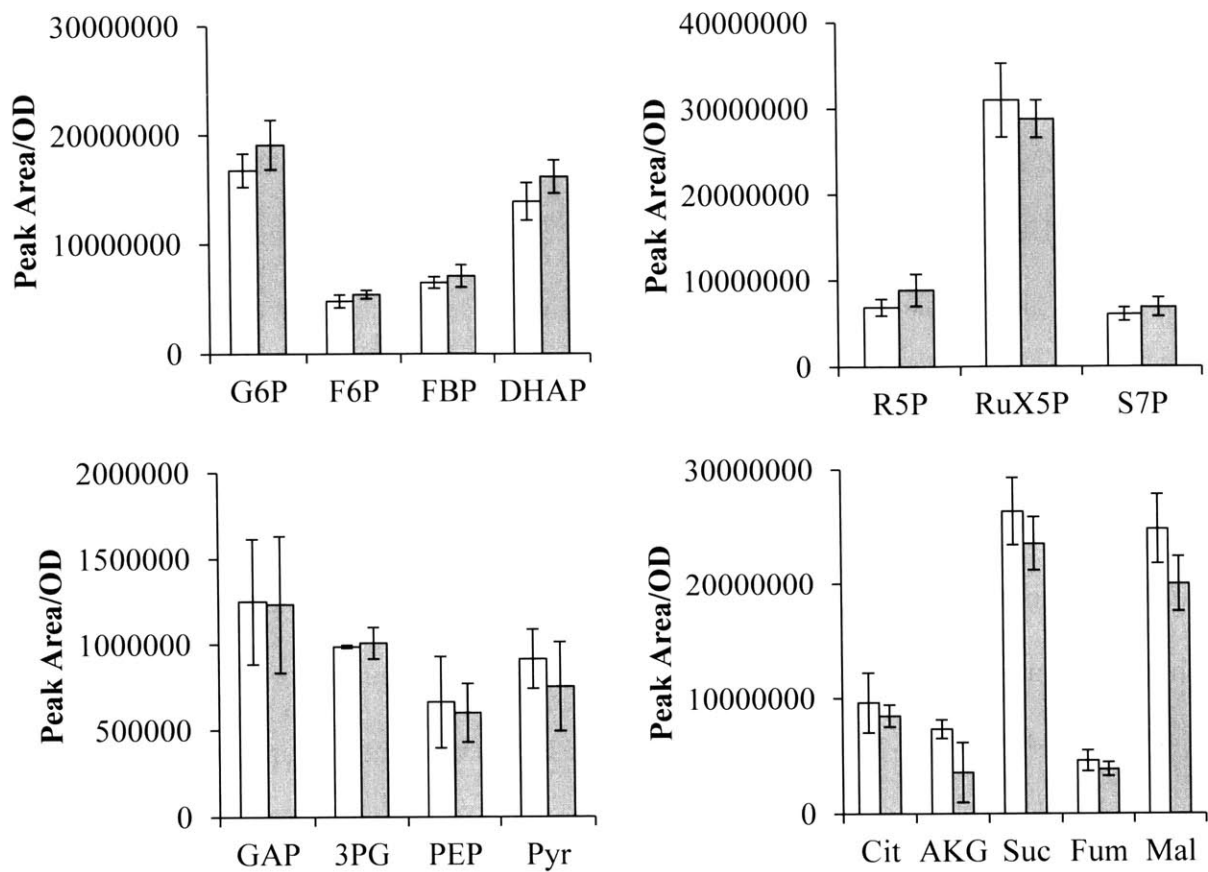


Figure 2-10 Comparison of hot ethanol and chloroform-methanol-water extraction protocols by LC-MS/MS. White bars, Hot Ethanol; grey bars, Chloroform-Methanol-Water.

2.3.7 Measurement of Mass Isotopomer Distributions (MIDs) of Metabolites Labeled to Natural Abundance

The LC-MS/MS method for quantification of metabolite ^{13}C labeling patterns was validated by measuring MIDs of metabolites labeled to natural abundance. A *S. cerevisiae* strain was grown with natural glucose as the sole carbon source. The MID of each metabolite in this culture can be predicted from the metabolite's molecular formula and the natural abundances of isotopes of carbon, hydrogen, oxygen, and phosphorous (e.g. the ^{13}C isotope natural abundance is 1.1%).

Metabolite samples were obtained using the optimized hot ethanol extraction protocol described above. Each metabolite MID was quantified by LC-MS/MS scanning for Q1 masses $M+0$, $M+1$, $M+2$, ..., $M+n$, where M is the molecular mass when all atoms are the most commonly occurring isotope and n is the number of carbons in the metabolite backbone. The Q3 fragment was generally a phosphate fragment ($\text{H}_2\text{PO}_4^- = 97$ or $\text{PO}_3^- = 79$). Since we are interested in ^{13}C labeling, the Q3 mass was held constant and the atoms in the Q3 fragment were omitted from the molecular formula used for calculating theoretical MIDs.

The theoretical and measured MIDs for a few representative metabolites are shown in Figure 2-11. It can be seen that good agreement between theory and experiment was obtained. Thus phosphorylated metabolite MIDs can be measured accurately using the methods for metabolite extraction and LC-MS/MS analysis described above.

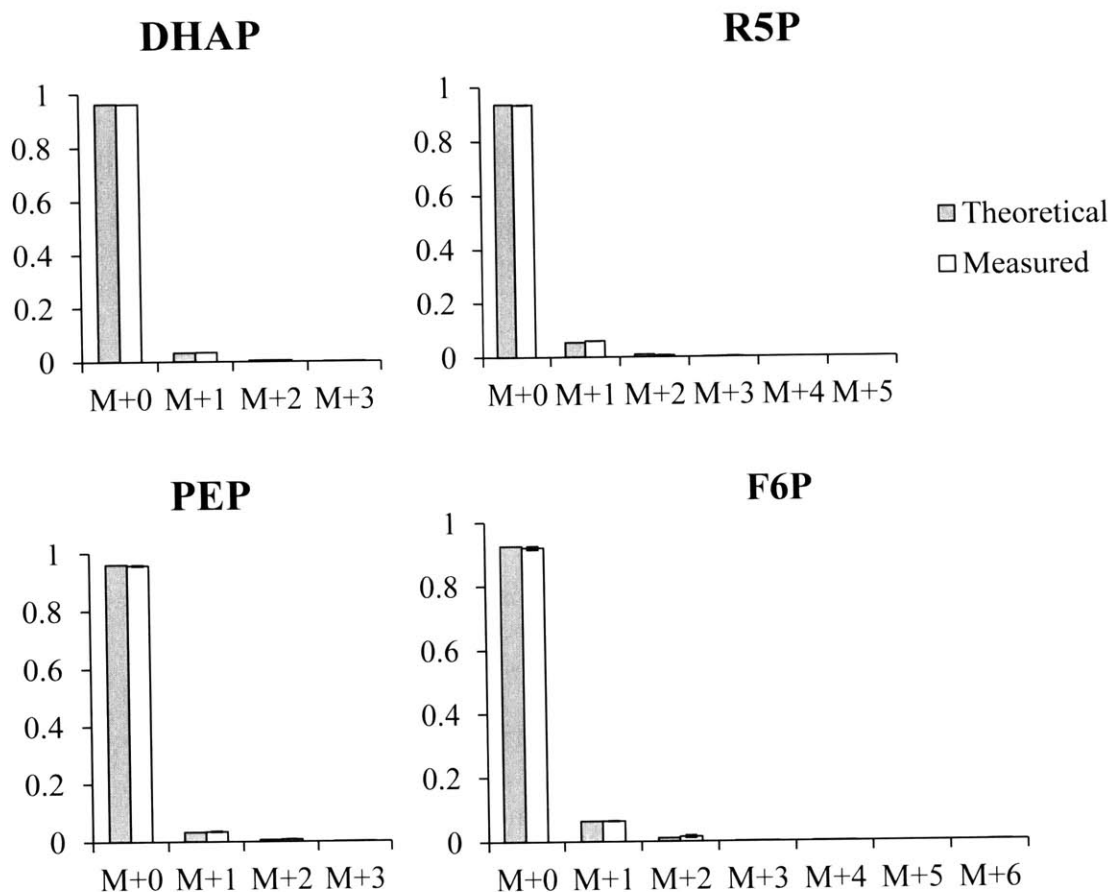


Figure 2-11 Validation of LC-MS/MS method for measuring metabolite labeling patterns. Measured mass isotopomer relative abundances are compared to mass isotopomer distributions predicted from theory.

2.3.8 Validation of Quenching and Extraction Protocols by Measurement of Metabolite Ratios

The quenching, extraction, and LC-MS/MS protocols can be validated by measurement of metabolite ratios whose values can be predicted under certain conditions. Two such ratios are the G6P/F6P ratio and the AEC, which is defined as:

$$AEC = \frac{[ATP] + 0.5[ADP]}{[ATP] + [ADP] + [AMP]} \quad (2.1)$$

The G6P/F6P ratio is expected to be maintained at a value greater than one in cells that are rapidly metabolizing glucose due to the directionality of the flux through phosphoglucose isomerase. The AEC is expected to be maintained at a value close to one.

For cells grown on glucose, once the cells are pelleted the glucose availability and glucose uptake rate will decrease significantly. If the phosphoglucose isomerase reaction is not quenched, G6P will continue to be metabolized but will not be replenished. Consequently if the quenching of metabolism during sample preparation is not adequate the intracellular G6P/F6P ratio will drop rapidly. Similarly, once cells are harvested they will likely be unable to continue producing ATP at an appreciable rate. Thus if the ATP-consuming reactions in the cell are not satisfactorily quenched, the AEC will decrease over time.

In the study of xylose metabolism described in Chapter 5, the G6P/F6P ratio and AEC were measured in a xylose-consuming *S. cerevisiae* strain during aerobic growth on glucose as sole carbon source. The measured values were similar to those obtained in a previously published study (Ewald et al., 2009) (Table 2-6). This suggests that metabolism is satisfactorily quenched during sample preparation and that the effects of metabolite degradation and interconversion are negligible.

Table 2-6 Measurement of key metabolite ratios for validation of quenching, extraction, and LC-MS/MS methods

	This work	Ewald, <i>et al.</i>
G6P/F6P Ratio	4.3	4.2
AEC	0.90	>0.8

2.4 Conclusions

Methods for extraction and quantification of yeast central carbon metabolites have been implemented. In the optimized metabolite extraction protocol, cells are quenched in pure methanol and washed once by resuspending the cell pellet in fresh methanol solution by pipetting. A hot ethanol extraction was selected as the best of three protocols investigated for extraction of metabolites from washed cell pellets. A LC-MS/MS method from the literature has been implemented for quantification of metabolite mass isotopomers, and MS parameters have been optimized to maximize signal-to-noise ratio. The extraction and LC-MS/MS methods have been validated by measuring MIDs of metabolites labeled to natural abundance and by measuring two key metabolite ratios in a xylose-consuming *S. cerevisiae* strain. Together, the protocols described here allow for measurement of central carbon metabolite pool sizes and ¹³C labeling patterns for studies of yeast metabolism.

2.5 References

Canelas, A.B., Ras, C., Pierick, A., Dam, J.C., Heijnen, J.J., Gulik, W.M., 2008. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics* 4, 226–239.

doi:10.1007/s11306-008-0116-4

Canelas, A.B., ten Pierick, A., Ras, C., Seifar, R.M., van Dam, J.C., van Gulik, W.M., Heijnen, J.J., 2009. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. *Anal. Chem.* 81, 7379–7389. doi:10.1021/ac900999t

Christensen, B., Gombert, A.K., Nielsen, J., 2002. Analysis of flux estimates based on (13)C-labelling experiments. *Eur. J. Biochem. FEBS* 269, 2795–2800.

De Koning, W., van Dam, K., 1992. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Anal. Biochem.* 204, 118–123.

Ewald, J.C., Heux, S., Zamboni, N., 2009. High-throughput quantitative metabolomics: workflow for cultivation, quenching, and analysis of yeast in a multiwell format. *Anal. Chem.* 81, 3623–3629. doi:10.1021/ac900002u

Faijes, M., Mars, A.E., Smid, E.J., 2007. Comparison of quenching and extraction methodologies for metabolome analysis of *Lactobacillus plantarum*. *Microb. Cell Factories* 6, 27.

doi:10.1186/1475-2859-6-27

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

Luo, B., Groenke, K., Takors, R., Wandrey, C., Oldiges, M., 2007. Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry. *J. Chromatogr. A* 1147, 153–164. doi:10.1016/j.chroma.2007.02.034

Maaheimo, H., Fiaux, J., Cakar, Z.P., Bailey, J.E., Sauer, U., Szyperski, T., 2001. Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional (¹³C) labeling of common amino acids. *Eur. J. Biochem. FEBS* 268, 2464–2479.

Marx, A., de Graaf, A.A., Wiechert, W., Eggeling, L., Sahm, H., 1996. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 49, 111–129. doi:10.1002/(SICI)1097-0290(19960120)49:2<111::AID-BIT1>3.0.CO;2-T

Rabinowitz, J.D., Kimball, E., 2007. Acidic acetonitrile for cellular metabolome extraction from *Escherichia coli*. *Anal. Chem.* 79, 6167–6173. doi:10.1021/ac070470c

Shastri, A.A., 2008. Metabolic flux analysis of photosynthetic systems. Purdue University.

Spura, J., Reimer, L.C., Wieloch, P., Schreiber, K., Buchinger, S., Schomburg, D., 2009. A method for enzyme quenching in microbial metabolome analysis successfully applied to gram-positive and gram-negative bacteria and yeast. *Anal. Biochem.* 394, 192–201. doi:10.1016/j.ab.2009.07.016

Szyperski, T., 1995. Biosynthetically directed fractional ¹³C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *Eur. J. Biochem.* FEBS 232, 433–448.

Van Winden, W.A., van Dam, J.C., Ras, C., Kleijn, R.J., Vinke, J.L., van Gulik, W.M., Heijnen, J.J., 2005. Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of (¹³C)-labeled primary metabolites. *FEMS Yeast Res.* 5, 559–568. doi:10.1016/j.femsyr.2004.10.007

Villas-Bôas, S.G., Bruheim, P., 2007. Cold glycerol-saline: the promising quenching solution for accurate intracellular metabolite analysis of microbial cells. *Anal. Biochem.* 370, 87–97. doi:10.1016/j.ab.2007.06.028

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

Wittmann, Heinzle, 1999. Mass spectrometry for metabolic flux analysis. *Biotechnol. Bioeng.* 62, 739–750.

Wu, L., Mashego, M.R., van Dam, J.C., Proell, A.M., Vinke, J.L., Ras, C., van Winden, W.A., van Gulik, W.M., Heijnen, J.J., 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ¹³C-labeled cell extracts as internal standards. *Anal. Biochem.* 336, 164–171. doi:10.1016/j.ab.2004.09.001

Young, J.D., Shastri, A.A., Stephanopoulos, G., Morgan, J.A., 2011. Mapping photoautotrophic metabolism with isotopically nonstationary (13)C flux analysis. *Metab. Eng.* 13, 656–665.

doi:10.1016/j.ymben.2011.08.002

Zhou, H., Cheng, J.-S., Wang, B.L., Fink, G.R., Stephanopoulos, G., 2012. Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*.

Metab. Eng. 14, 611–622. doi:10.1016/j.ymben.2012.07.011

Chapter 3

Modeling of Propagation of Error in Mass Spectrometry

Mass Isotopomer Distribution Data

3.1 Introduction

Intracellular metabolic flux estimation as described in Chapter 1 is essentially a standard non-linear parameter estimation problem. The most likely metabolic flux distribution is determined by finding the set of fluxes \vec{v} that minimizes the weighted sum of squared residuals (WSSR) (see equation (1.5), which is reproduced here as equation (3.1)):

$$\begin{aligned} \min_{\vec{v}} \quad & \sum_i \frac{(y_i^{meas} - y_i(\vec{v}))^2}{\sigma_i^2} + \sum_{k \in EF} \frac{(v_k^{meas} - v_k)^2}{\sigma_k^2} \\ \text{s. t.} \quad & \mathbf{S} \cdot \vec{v} = \vec{\mathbf{0}} \end{aligned} \tag{3.1}$$

In equation (3.1), each squared residual is weighted by the variance of the associated measurement; this ensures that measurements with small errors are given greater weight, and conversely measurements with large errors have less influence on the estimation of fluxes.

If all squared residuals are properly weighted a chi-square test for goodness-of-fit can be used to determine whether the metabolic model adequately describes the labeling data. Such a test was recently used to identify an error in a published metabolic model of *Escherichia coli* (Leighty and Antoniewicz, 2012). Without such a test for goodness-of-fit, one cannot determine whether

the system being investigated has been modeled satisfactorily, and the reliability of the estimated metabolic flux distribution cannot be ascertained. In addition, the proper weighting of squared residuals allows one to determine the precision with which fluxes are known; only by computing confidence intervals for estimated fluxes can one determine whether differences between flux distributions estimated from different strains or under different culture conditions are statistically significant. Confidence intervals for estimated metabolic fluxes have been computed using linearization theory (Wiechert et al., 1997; Araúzo-Bravo and Shimizu, 2003) and Monte Carlo methods (Malloy et al., 1990; Schmidt et al., 1999), the former method giving less exact confidence intervals and the latter being computationally expensive. Recently a new method has been proposed in which the confidence interval for each flux is computed by determining the change in the WSSR as the flux value is varied (Antoniewicz et al., 2006). However, all three of these methods for computing confidence intervals rely on accurate assessment of the σ_i and σ_k , which allows proper weighting of the squared residuals. Thus accurate assessment of measurement errors is essential for interpretation of the results of flux estimation.

However, the measurement errors in mass spectrometry (MS) MID data are still relatively unknown. In 1999 Möllney, *et al.* wrote that “[a]n important problem for the future will be the assignment of reasonable measurement error bars for MS experiments. To this end, more theoretical and experimental studies are required. Even if it turns out that [MID] data really have an outstanding quality it must be pointed out that the assumption of very small error bars will almost surely produce significant inconsistency problems for fitting the large data sets. These inconsistencies stem from the simplifying assumptions that are still present in the network formulation and will undoubtedly make data evaluation more difficult” (Möllney et al., 1999).

In 2001 Wiechert, *et al.* noted that in most studies the assumed measurement variances used to weight the squared residuals of the labeling data in the objective function are “rather crude values” (Wiechert et al., 2001). Unfortunately more than a decade later this is still the case. In many instances researchers simply assume a measurement variance with little justification.

Assessment of the errors in MS data is difficult because while mass spectrometers are incredibly precise instruments, this extraordinary precision may lead to overly optimistic estimation of the accuracy of labeling data. In a study of GC-MS MID data for 29 reliable amino acid fragments, Antoniewicz, *et al.* observed that for many of the fragments the precision of the computed isotopomer mole fractions was superior to the accuracy. In fact the accuracy threshold for a reliable fragment was chosen to be <0.5 mole %, while the precision threshold was only 0.2 mole % (Antoniewicz et al., 2007). Consequently, the measurement errors in GC-MS MID data cannot be estimated from standard deviations calculated from many replicate samples—the run-to-run variation in the mass isotopomer mole fractions is less than the true error in the mole fractions.

Determination of the σ_i is also complicated by the fact that the mole fractions y_i for a given metabolite are not truly measured quantities. When ^{13}C labeling data are obtained by MS, each y_i is computed from the set of MS ion counts for the metabolite x_j :

$$y_i \equiv \frac{x_i}{\sum_{j=0}^N x_j} \quad (3.2)$$

where N is the total number of carbons in the metabolite (or, more precisely, the number of labeled atoms in the highest mass isotopomer of the MID). Thus the ion counts x_j are the true

measured quantities. Although this chapter will focus on MS ^{13}C labeling data, similar arguments could be made for labeling data obtained through NMR.

In this chapter, we sought to use the theory of error propagation to model the error in MID mole fractions \bar{y} obtained from MS ion counts \bar{x} . In particular we examined whether the error in a mole fraction y_i could be modeled as a function of y_i in order to determine which mole fraction measurements should receive the greatest weight in the WSSR.

3.2 Error Model

As noted above, mass isotopomer mole fractions are not measured quantities *per se*; the mole fractions are computed from mass spectrometer measurements of mass isotopomer ion counts. In order to model the error in y_i , Gaussian noise should be added to the ion counts x_i (as noted in (Srouf et al., 2011)):

$$y_i^{meas} = \frac{x_i + \varepsilon_i}{\sum_{j=0}^N (x_j + \varepsilon_j)} \quad (3.3)$$

and the propagation of error from the ion counts to the mole fractions should be investigated. In equation (3.3), ε_j represents the error in the measurement of the j^{th} ion count. To investigate the propagation of error, we examined the partial derivatives of y_i with respect to the i^{th} ion count x_i and with respect to the other ion counts $x_{j,j \neq i}$:

$$\frac{\partial y_i}{\partial x_i} = \frac{\sum_{j=0}^N x_j - x_i}{(\sum_{j=0}^N x_j)^2} = \frac{1}{\sum_{j=0}^N x_j} (1 - y_i) = \frac{1 - y_i}{TIC} \quad (3.4)$$

$$\frac{\partial y_i}{\partial x_{j,j \neq i}} = \frac{-x_i}{(\sum_{j=0}^N x_j)^2} = \frac{-y_i}{TIC} \quad (3.5)$$

where the definition of y_i in equation (3.2) has been used to simplify both expressions and we note that:

$$\sum_{j=0}^N x_j = TIC \quad (3.6)$$

where TIC is the total ion count for the ion cluster.

The total error in y_i arising from the measurement errors in all ion counts x_j can be approximated by the total differential of y_i :

$$\begin{aligned} \delta y_i &= \sum_{j=0}^N \frac{\partial y_i}{\partial x_j} \delta x_j = \frac{\partial y_i}{\partial x_i} \delta x_i + \sum_{j=0, i \neq j}^N \frac{\partial y_i}{\partial x_j} \delta x_j \quad (3.7) \\ &= \frac{1}{TIC} \left[(1 - y_i) \delta x_i - \sum_{j=0, i \neq j}^N y_i \delta x_j \right] \end{aligned}$$

In order to estimate the deviation in the mass isotopomer mole fraction δy_i , we must first model the deviations in the measured ion counts, the δx_j . Dauner, *et al.* have proposed an error model for NMR data which will be adopted here (Dauner et al., 2001). The error model for the measured quantities x_j assumes that the error has a proportional and a constant term:

$$\delta x_j = a_1 + a_2 x_j \quad (3.8)$$

where a_1 and a_2 are unknown constants. A similar model was proposed for GC-MS data (Dauner, 2000).

It is informative to split the total differential of y_i into two terms, the first being the deviation that results from measurement error in x_i and the second being the deviation that results from errors in the $x_{j,j \neq i}$.

$$\delta y_i = \delta y_i^1 + \delta y_i^2 \quad (3.9)$$

$$\delta y_i^1 = \frac{1 - y_i}{TIC} \delta x_i \quad (3.10)$$

$$\delta y_i^2 = \frac{-1}{TIC} \sum_{j=0, i \neq j}^N y_j \delta x_j \quad (3.11)$$

We look at two limiting cases, the first case being proportional error ($a_1 = 0$) and the second being constant error ($a_2 = 0$). Dauner found that the errors in GC-MS data were proportional as long as the signal-to-noise ratio was sufficiently high (Dauner, 2000), but there are reasons to believe that some errors in GC-MS data may be constant (discussed below).

In the case of proportional error, the error δx_j is proportional to x_j which in turn is proportional to y_j . Consequently the two terms in the total differential of y_i can be simplified to:

$$\delta y_i^1 \propto \frac{(1 - y_i)y_i}{TIC} \quad (3.12)$$

$$\delta y_i^2 \propto \frac{-y_i}{TIC} \sum_{j=0, i \neq j}^N y_j \quad (3.13)$$

However, using the fact that

$$\sum_{j=0}^N y_j = 1 \quad (3.14)$$

and therefore

$$\sum_{j=0, i \neq j}^N y_j = 1 - y_i \quad (3.15)$$

we can simplify the second term to

$$\delta y_i^2 \propto \frac{-(1 - y_i)y_i}{TIC} \quad (3.16)$$

yielding the desired result. In the case of proportional error, both terms δy_i^1 and δy_i^2 are proportional to $(1 - y_i)y_i$, and thus ***in the case of proportional error the standard deviation in the mole fraction of a mass isotopomer is expected to depend quadratically on the value of the mole fraction.*** The standard deviations in the mole fractions of mass isotopomers are expected to scale as $(1 - y_i)y_i$.

In the case of constant error, the result is less simple. Assuming the δx_j are equal to a constant δx independent of the ion count we have

$$\delta y_i^1 = \frac{(1 - y_i)}{TIC} \delta x \quad (3.17)$$

$$\delta y_i^2 = \frac{-N y_i \delta x}{TIC} \quad (3.18)$$

As $y_i \rightarrow 1$, we expect the term δy_i^1 to be small and the term δy_i^2 to dominate so that the standard deviation in y_i will increase linearly with y_i . As $y_i \rightarrow 0$ the opposite is true and the standard deviation in y_i should decrease with y_i as $(1 - y_i)$. Thus ***for constant error the standard deviation in a mole fraction of a mass isotopomer distribution is again expected to depend on the value of the mole fraction. For small mole fractions the standard deviation is expected to decrease with y_i and scale as $(1 - y_i)$ whereas for large mole fractions the standard deviation is expected to scale with y_i .***

The applicability of the error model was investigated using both theory and experiments.

3.3 Materials and Methods

3.3.1 Simulation of Noisy MS Labeling Data

Random MIDs were generated in Matlab. Each MID consisted of four mass isotopomer mole fractions (i.e ranging from $M + 0$ to $M + 3$). For each MID, the $M + 0$ mole fraction y_0 was fixed to one of 21 different values: 0, 0.05, 0.1, 0.15, ..., 0.95, 1. 100 different MIDs were generated with each y_0 value to ensure complete coverage of the mole fraction space, resulting in a total of 2100 MIDs. For each MID the other three mole fractions y_1 , y_2 , and y_3 were assigned by generating three random numbers (using the Matlab function `rand`) and then normalizing such that the sum of the three random numbers was equal to $1 - y_0$. For each MID, the ion counts x_i were then computed by multiplying the y_i by the total ion count, which was assumed to be 10,000 for all MIDs.

Random error was then added to the ion counts associated with each MID to determine how errors in the ion counts propagated to the mole fractions. Proportional error was added to ion count x_i by using the Matlab function `normrnd` to generate a random number δx_i drawn from a Gaussian distribution with a mean of zero and a standard deviation equal to 1% of x_i . The “noisy” ion count x_i^{noise} is then equal to:

$$x_i^{noise} = x_i + \delta x_i \quad (3.19)$$

In the case of constant error each δx_i was drawn from a Gaussian distribution with a mean of zero and a standard deviation equal to 0.3% of the total ion count (i.e. 30), independent of the

value of x_i . The “noisy” MID mole fractions y_i^{noise} can then be computed from equation (3.20), which is analogous to equation (3.2):

$$y_i^{noise} = \frac{x_i^{noise}}{\sum_{j=0}^3 x_j^{noise}} \quad (3.20)$$

For each case considered (proportional error, constant error, and both proportional and constant errors), noise was added to each MID 1000 times and the standard deviation δy_i of each MID mole fraction was computed by calculating the standard deviation of the 1000 y_i^{noise} values from the 1000 “noisy” MIDs.

3.3.2 Strain and Culture Conditions

Protein hydrolysate samples were obtained from *E. coli* BW 25113. The strain was cultured at 37 °C in MOPS minimal medium (Teknova, M2106) with 5 g/l glucose as the sole carbon source. A 5 ml starter culture grown in an aerobic culture tube (BD Falcon, #352059) was inoculated from 15% glycerol -80 °C freezer stock. After 24 h, a 50 ml shake flask culture was inoculated with 500 μ l (1% by volume) starter culture. All cultures were shaken at 250 rpm.

3.3.3 Protein Hydrolysis

The shake flask culture was harvested at $OD_{600} \sim 1$. Ten 1 ml culture samples were centrifuged 5 min at maximum speed (18,000g) in pre-chilled Eppendorf tubes and the supernatants removed. Cell pellets were washed twice by resuspension in 1 ml 9 g/l sodium chloride solution, centrifugation 5 min at maximum speed, and removal of supernatant. Samples were kept on ice throughout harvesting and washing. 700 μ l 6 M hydrochloric acid was added to each cell pellet. Samples were vortexed, transferred to vacuum hydrolysis tubes, and incubated 24 h at 105 °C.

Hydrolysates were transferred to clean Eppendorf tubes and dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module. Dried samples were stored at -80 °C prior to GC-MS analysis.

3.3.4 Gas Chromatography/Mass Spectrometry (GC-MS)

Dried protein hydrolysates were resuspended in 50 μ l pyridine. 70 μ l *N-tert*-Butyldimethylsilyl-*N*-methyltrifluoroacetamide with 1% *tert*-Butyldimethylchlorosilane (Sigma-Aldrich, #375934) was added to each sample and samples were incubated 30 min at 60 °C. Samples were centrifuged 5 min at maximum speed and supernatants analyzed by GC-MS using an Agilent 6890N Network GC System coupled to an Agilent 5975B Inert XL MSD. The GC/MS conditions have been described in (Wasylenko and Stephanopoulos, 2015) (see also Chapter 5). All samples were injected with a split ratio of 4:1.

3.4 Results

3.4.1 Simulation of Noisy MS Labeling Data

Random MIDs were generated as described in the Methods section. Random noise was added to the ion counts and the standard deviations in the MID mole fractions δy_i were computed. The mole fraction standard deviations are plotted as a function of the mole fraction y_i for the cases of proportional errors (Figure 3-1), constant errors (Figure 3-2), and both proportional and constant errors in the ion counts (Figure 3-3). The results agree with the predictions of the error model. In the case of proportional error in the ion counts, the standard deviation δy_i depends

quadratically on the mole fraction y_i , reaching a maximum at $y_i = 0.5$. In the case of constant error, the standard deviation decreases with y_i for small values of y_i before reaching a minimum and then increasing approximately linearly with y_i . When both proportional and constant errors are included the relationship between δy_i and y_i is more complex and depends on the relative sizes of the proportional and constant error terms. In the example shown in Figure 3-3, as $y_i \rightarrow 0$ and $y_i \rightarrow 1$ the contribution of the proportional error is small, but for intermediate values of y_i the proportional error term contributes significantly.

When the level of noise was doubled (by doubling the a_1 and a_2 parameters) the plots of δy_i as a function of y_i were qualitatively similar, with δy_i increasing approximately two-fold at each value of y_i (Figures 3-4, 3-5, and 3-6). This result suggests δy_i is roughly proportional to the noise in the ion counts, as expected, and shows that the error model predicts the qualitative dependence of δy_i on y_i at different levels of noise. The error model also predicts that with constant error in the ion counts the error δy_i increases with N (equation (3.18)); thus larger errors are expected for MIDs containing larger numbers of mass isotopomer mole fractions in the case of constant errors, but not in the case of proportional errors. Simulations were performed with MIDs of different sizes. Representative results are shown for the case of $N = 6$ in Figures 3-7 and 3-8. Comparing to Figures 3-1 and 3-2 respectively, it can be seen that in the case of constant errors in the ion counts δy_i increases with N , particularly for larger values of y_i where the δy_i^2 term is significant. In the limit as $y_i \rightarrow 0$ (where the δy_i^1 term dominates), δy_i is largely independent of N , even in the case of constant errors. The magnitude of δy_i did not increase with N in the case of proportional error (Figures 3-1 and 3-7), which is consistent with the error model.

The results of the Monte Carlo simulations show that when random errors are added to MS ion counts *in silico* the errors propagate to the MID mole fractions in a complex manner, leading to an unexpected dependence of the error on the magnitude of the mole fraction. The error model described above successfully predicted the relationship between δy_i and y_i in the cases of both proportional and constant errors in mass isotopomer ion counts.

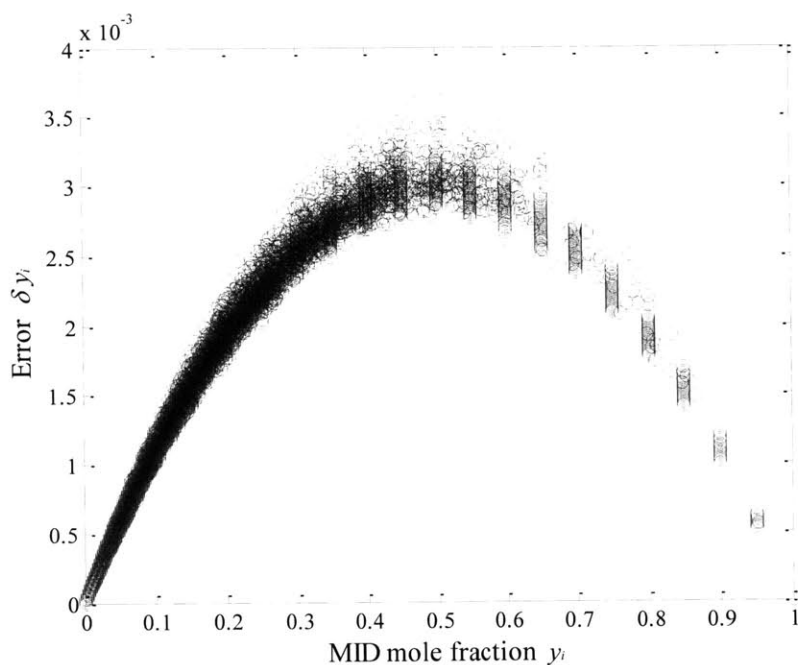


Figure 3-1 Results of Monte Carlo simulations used to validate MS data error model in the case of proportional error only in ion counts ($a_1 = 0$, $a_2 = 0.01$, $N = 3$)

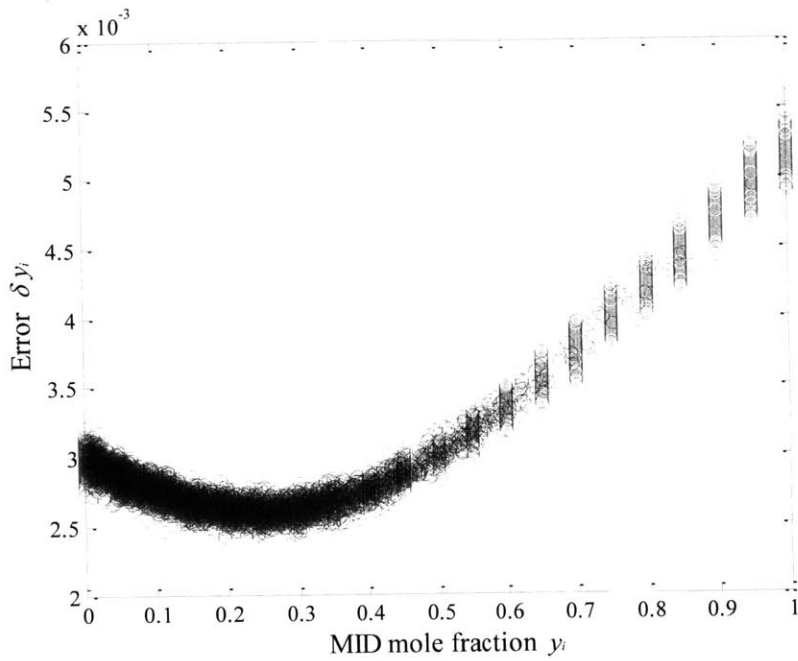


Figure 3-2 Results of Monte Carlo simulations used to validate MS data error model in the case of constant error only in ion counts ($a_1 = 0.003$, $a_2 = 0$, $N = 3$)

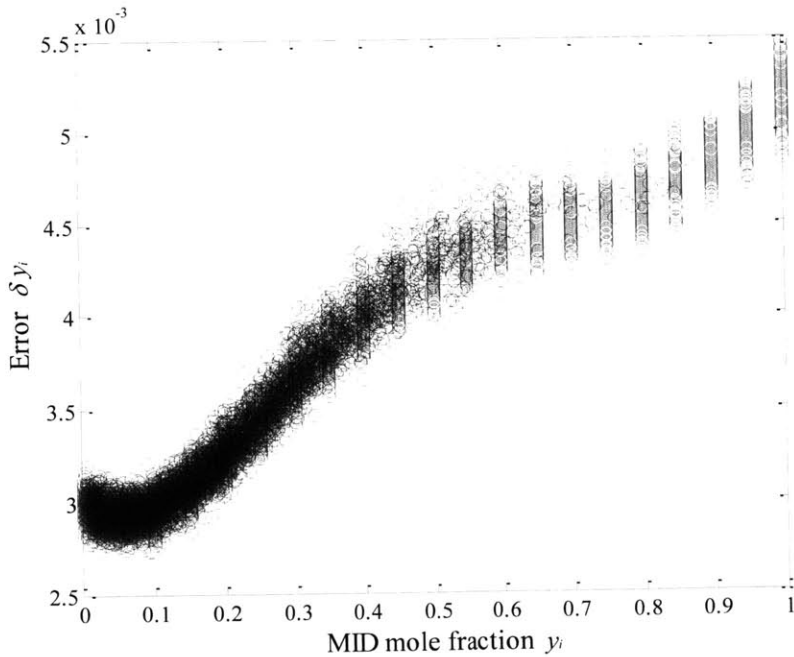


Figure 3-3 Results of Monte Carlo simulations used to validate MS data error model in the case of both proportional and constant errors in ion counts ($a_1 = 0.003$, $a_2 = 0.01$, $N = 3$)

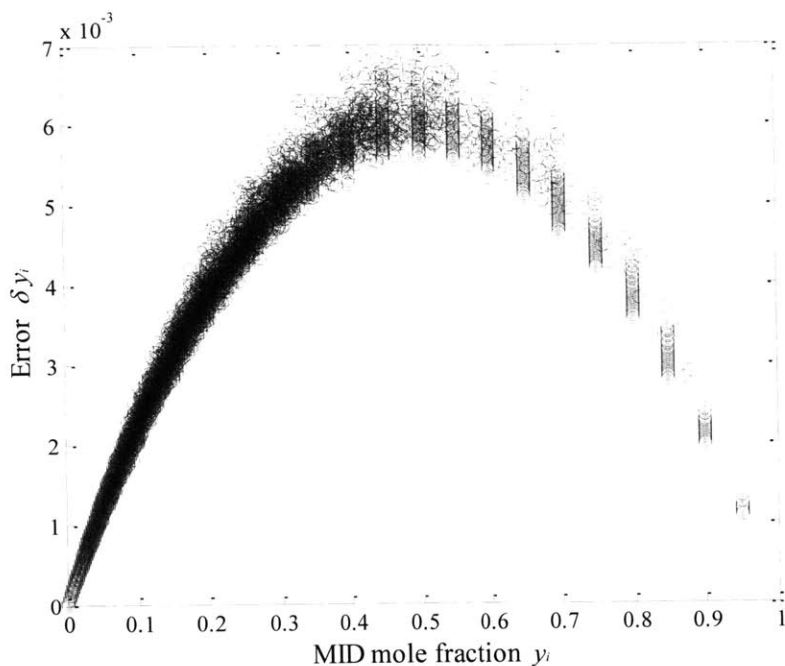


Figure 3-4 Results of Monte Carlo simulations used to validate MS data error model with double proportional error only in ion counts ($a_1 = 0$, $a_2 = 0.02$, $N = 3$)

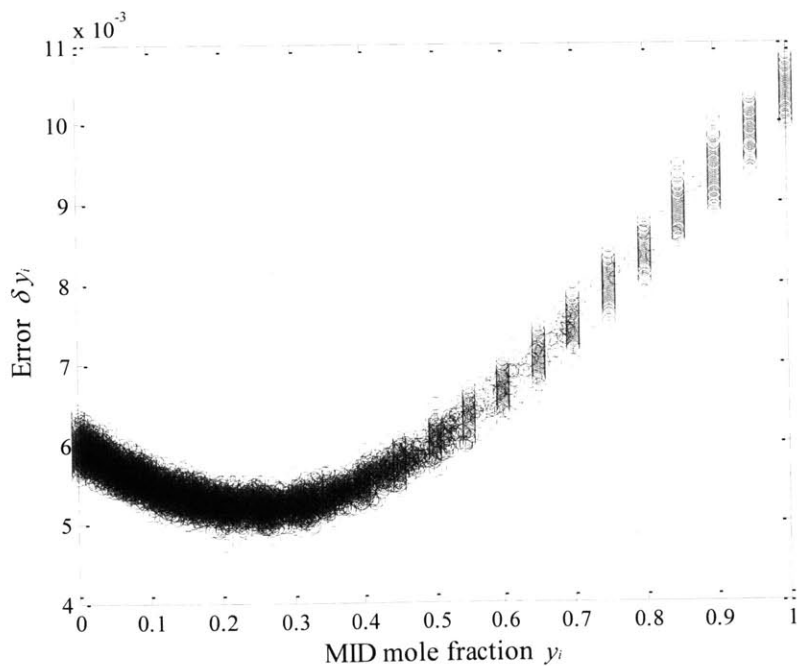


Figure 3-5 Results of Monte Carlo simulations used to validate MS data error model with double constant error only in ion counts ($a_1 = 0.006$, $a_2 = 0$, $N = 3$)

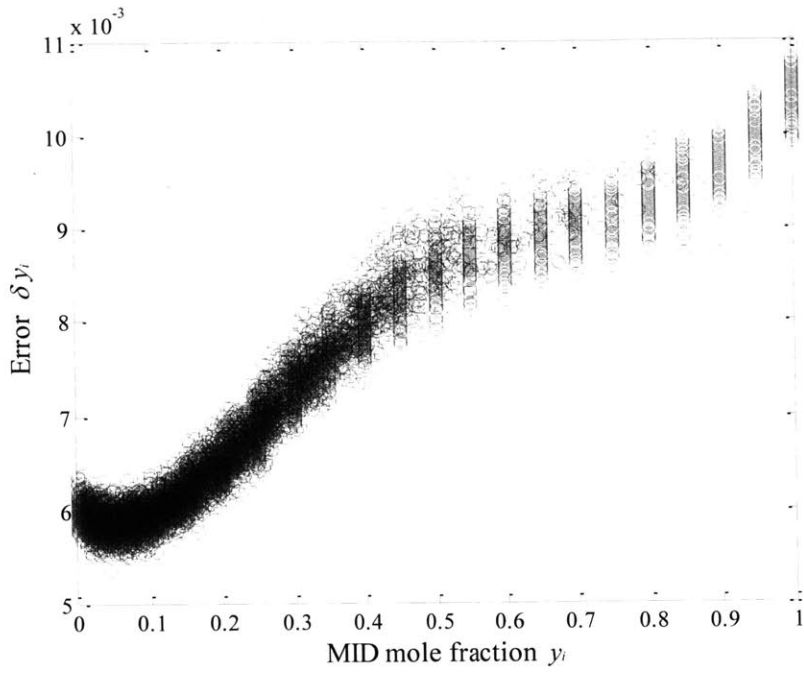


Figure 3-6 Results of Monte Carlo simulations used to validate MS data error model with double proportional and constant errors in ion counts ($a_1 = 0.006$, $a_2 = 0.02$, $N = 3$)

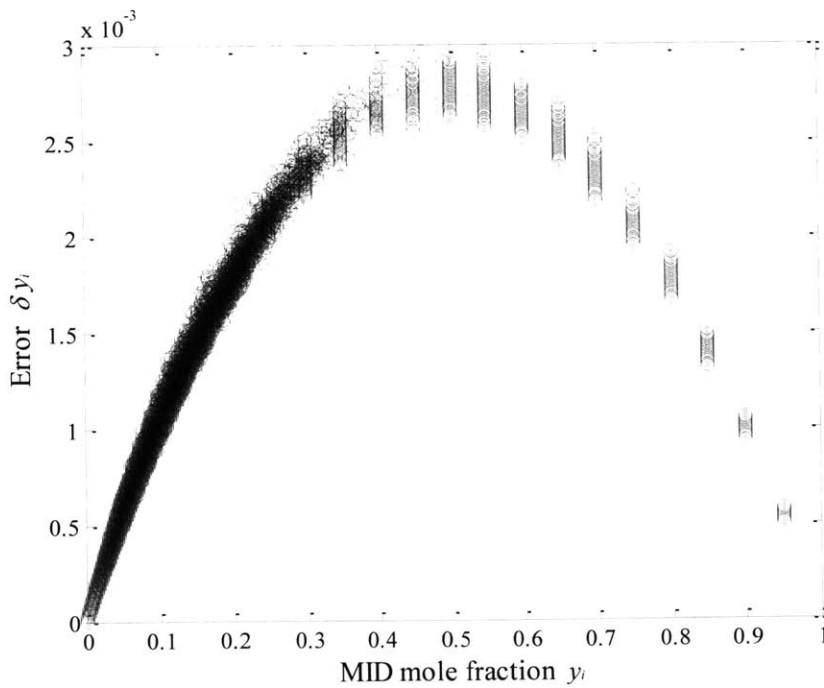


Figure 3-7 Results of Monte Carlo simulations used to validate MS data error model in the case of proportional error only in ion counts ($a_1 = 0$, $a_2 = 0.01$, $N = 6$)

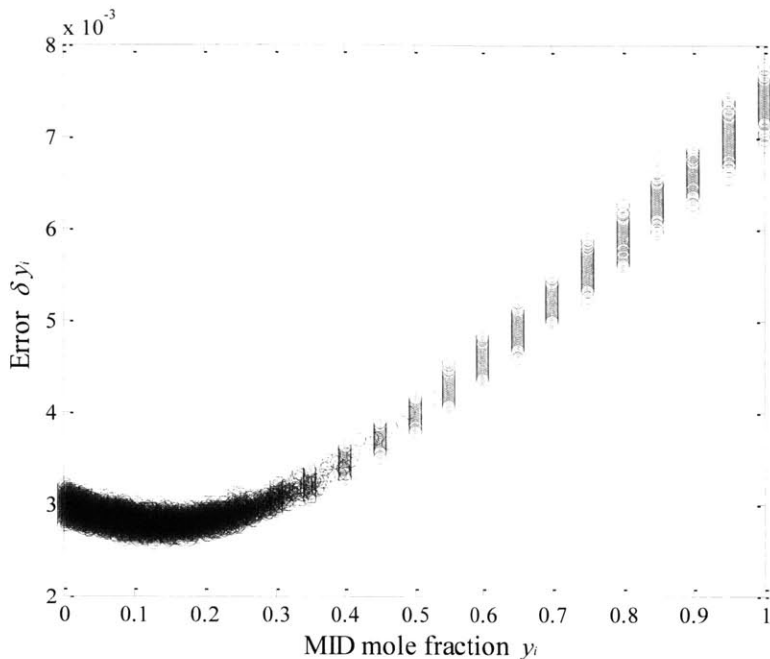


Figure 3-8 Results of Monte Carlo simulations used to validate MS data error model in the case of constant error only in ion counts ($a_1 = 0.003$, $a_2 = 0$, $N = 6$)

3.4.2 Analysis of *E. coli* Protein Hydrolysates by GC-MS

After *in silico* validation of the model for error propagation, a large GC-MS MID dataset was generated in order to determine whether the error model could be used to identify the appropriate variances σ_i^2 for the objective function in equation (3.1). More specifically, we sought to determine whether the weight given to a residual should be a function of the MID mole fraction y_i . Protein hydrolysates labeled to natural abundance were obtained from an *E. coli* culture and analyzed by GC-MS. The MID mole fraction errors δy_i were then calculated from the differences between the MID mole fractions measured in the experiment y_i^{meas} and the MID mole fractions predicted from theory y_i and plotted as a function of y_i . All MIDs with an error $\delta y_i > 0.01$ were omitted from the analysis. The results are shown in Figure 3-9. The MID mole

fractions formed two clusters, with the M+0 mole fractions typically falling between 0.6 and 0.9 and the higher isotopomer mole fractions typically <0.25 . A quantitative correlation between δy_i and y_i could not be obtained. Larger MID mole fractions tended to have larger errors, suggesting there is a significant “constant” component in the ion count errors. However due to the large spread in the data and the lack of data points between mole fractions 0.25 and 0.6, it was not possible to use the GC-MS dataset to determine how the σ_i should depend on y_i^{meas} .

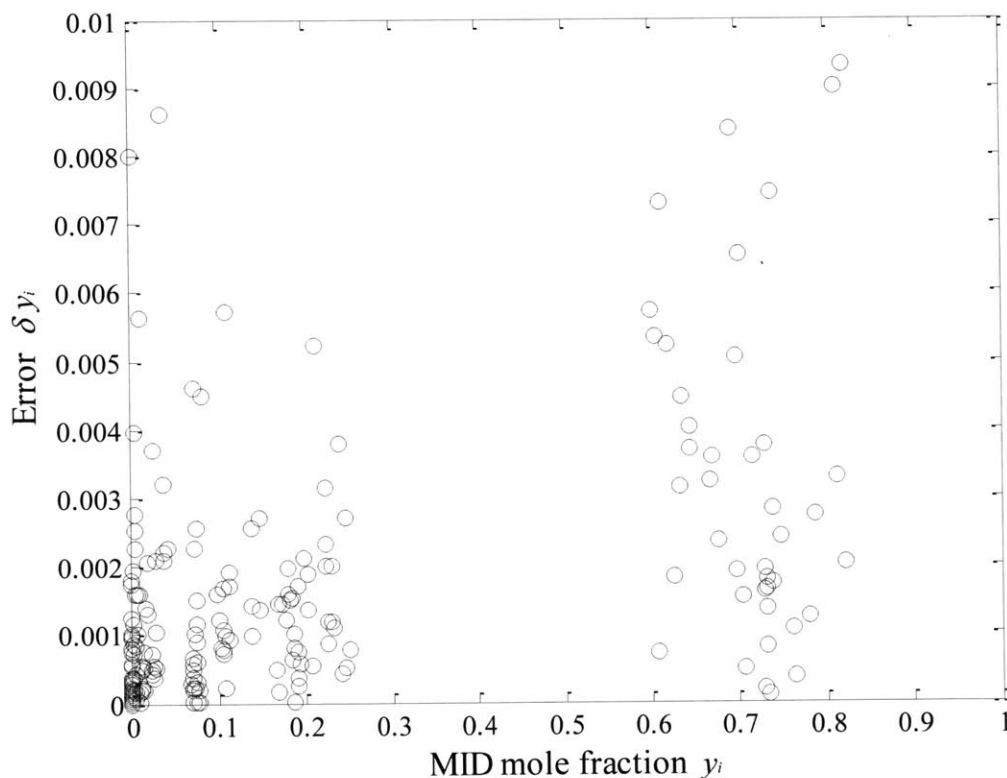


Figure 3-9 GC-MS measurement errors in MID mole fractions of proteinogenic amino acids labeled to natural abundance

3.5 Discussion

Although the model for propagation of error from ion counts to MID mole fractions was validated *in silico*, we were not able to use the error model to identify the appropriate weights σ_i for the objective function in equation (3.1). It is possible that the dataset considered here was not large enough to allow determination of the functional form of the dependence of δy_i on y_i . Additionally, because the protein hydrolysates were labeled to natural abundance the MID mole fractions formed two clusters (the M+0 mole fractions, which were typically in the range 0.6-0.9, and the higher isotopomer mole fractions, which were typically <0.25). No MID mole fractions between 0.3 and 0.6 were obtained. This “gap” in the MID mole fraction space may obscure correlations between MID mole fraction and error that are present in the data. However, the errors in GC-MS data likely arise from a number of sources and it is also possible that these errors cannot be modeled as a simple function of y_i .

Sources of error in MS labeling data include the integration of raw chromatograms to compute mass isotopomer peak areas (Antoniewicz et al., 2007) and non-linearities in detector response (Dauner and Sauer, 2000). Saturation of the detector (which will result in underestimation of ion counts and mole fractions for the most abundant mass isotopomers) can be avoided by decreasing the amount of sample injected into the MS. However, as the amount of sample is decreased the signal-to-noise ratio for less abundant ions will decrease. In biological samples different analytes will often be present at very different concentrations; even the concentrations of different mass isotopomers of the same analyte may differ by more than an order of magnitude. Thus it is difficult if not impossible to achieve an optimal concentration of all

analytes in a sample in a single injection, and care must be taken to ensure that all signals are sufficiently high to limit the influence of noise but sufficiently low that the detector is not overloaded.

Concentration dependence of computed MIDs has been reported by (Fagerquist et al., 1999) and (Antoniewicz et al., 2007). In these studies, the computed mole fractions of the measured MIDs depended on the amount of sample injected into the GC-MS, with the largest mole fractions decreasing relative to smaller mole fractions as sample size increased. This effect has been attributed to chemistry occurring in the gas phase (Fagerquist et al., 2001) and is mostly likely instrument-dependent (Antoniewicz et al., 2007).

In some cases, undesired metabolites or undesired fragments of a desired metabolite may have retention times and masses that are similar to those of a fragment for which MS data are desired. In this case, the undesired fragments can contribute to the ion counts of the desired fragment, leading to a skew in the resulting MID. The effects of overlapping ion fragments are similar to the effects of measurement noise, but generally cannot be reduced by increasing the signal strength as both the desired and undesired signals will increase similarly. In fact increasing sample concentration might worsen the temporal resolution of two analytes, exacerbating the problem. The errors associated with interference by overlapping ion clusters and measurement noise are expected to be “constant,” independent of the ion count x_i .

The final source of error in GC-MS data that will be discussed has been proposed by Christensen and Nielsen (Christensen and Nielsen, 1999). The ion count x_i of a mass isotopomer is a

function of the concentration of the measured mass isotopomer in the sample c_i and the “ionization efficiency” η_i , which is the fraction of molecules that are ionized and detected by the mass spectrometer:

$$x_i = \eta_i c_i \quad (3.21)$$

In mass isotopomer analysis it is implicitly assumed that, for a given ion cluster, the η_i are equal for all i . However, this is not necessarily the case. Because they contain an extra neutron, ^{13}C atoms form bonds with slightly different bond strengths than those formed by ^{12}C atoms (Christensen and Nielsen, 1999). Because ionization frequently involves the breaking of these bonds, a bias toward lighter mass isotopomers might be expected, since these isotopomers have more ^{12}C atoms and will generally form slightly weaker bonds and could ionize more easily. In fact different isotopomers with the same number of labeled atoms (and which therefore contribute to the same mass isotopomer) may have different ionization efficiencies due to the different positions of the ^{12}C and ^{13}C atoms, and so the ionization efficiency η_i of a mass isotopomer could in fact depend on the composition of the mass isotopomer in terms of its constituent positional isotopomers.

These various sources of error could contribute to different extents for different analytes. Consequently, it is perhaps not surprising that we were unable to find a simple functional form for the dependence of δy_i on y_i . Moreover, in determining the appropriate σ_i for ^{13}C -MFA one must consider not only the measurements errors but also potential sources of modeling error, one of which will be discussed in the next Chapter.

In sum, due to the complex nature of the errors in MS ^{13}C labeling data, we were unable to determine how the objective function weights σ_i should depend on the mass isotopomer mole fractions y_i . Consequently we have assumed a constant σ_i independent of y_i for each of the ^{13}C -MFA studies in this thesis.

3.6 References

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2006. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* 8, 324–337. doi:10.1016/j.ymben.2006.01.004

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Anal. Chem.* 79, 7554–7559. doi:10.1021/ac0708893

Araújo-Bravo, M.J., Shimizu, K., 2003. An improved method for statistical analysis of metabolic flux analysis using isotopomer mapping matrices with analytical expressions. *J. Biotechnol.* 105, 117–133.

Christensen, B., Nielsen, J., 1999. Isotopomer analysis using GC-MS. *Metab. Eng.* 1, 282–290. doi:10.1006/mben.1999.0117

Dauner, M., 2000. Intracellular carbon flux analysis by ^{13}C -tracer experiments (PhD). ETH Zurich.

- Dauner, M., Bailey, J.E., Sauer, U., 2001. Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*. *Biotechnol. Bioeng.* 76, 144–156.
- Dauner, M., Sauer, U., 2000. GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing. *Biotechnol. Prog.* 16, 642–649. doi:10.1021/bp000058h
- Fagerquist, C.K., Hellerstein, M.K., Faubert, D., Bertrand, M.J., 2001. Elimination of the concentration dependence in mass isotopomer abundance mass spectrometry of methyl palmitate using metastable atom bombardment. *J. Am. Soc. Mass Spectrom.* 12, 754–761. doi:10.1016/S1044-0305(01)00227-6
- Fagerquist, C.K., Neese, R.A., Hellerstein, M.K., 1999. Molecular ion fragmentation and its effects on mass isotopomer abundances of fatty acid methyl esters ionized by electron impact. *J. Am. Soc. Mass Spectrom.* 10, 430–439. doi:10.1016/S1044-0305(99)00003-3
- Leighty, R.W., Antoniewicz, M.R., 2012. Parallel labeling experiments with [U-¹³C]glucose validate *E. coli* metabolic network model for ¹³C metabolic flux analysis. *Metab. Eng.* 14, 533–541. doi:10.1016/j.ymben.2012.06.003
- Malloy, C.R., Sherry, A.D., Jeffrey, F.M., 1990. Analysis of tricarboxylic acid cycle of the heart using ¹³C isotope isomers. *Am. J. Physiol.* 259, H987–995.
- Möllney, M., Wiechert, W., Kownatzki, D., de Graaf, A.A., 1999. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. *Biotechnol. Bioeng.* 66, 86–103.
- Schmidt, K., Nørregaard, L.C., Pedersen, B., Meissner, A., Duus, J.O., Nielsen, J.O., Villadsen, J., 1999. Quantification of intracellular metabolic fluxes from fractional enrichment and ¹³C-

¹³C coupling constraints on the isotopomer distribution in labeled biomass components. *Metab. Eng.* 1, 166–179. doi:10.1006/mben.1999.0114

Srour, O., Young, J.D., Eldar, Y.C., 2011. Fluxomers: a new approach for ¹³C metabolic flux analysis. *BMC Syst. Biol.* 5, 129. doi:10.1186/1752-0509-5-129

Wasylenko, T.M., Stephanopoulos, G., 2015. Metabolomic and (¹³) C-metabolic flux analysis of a xylose-consuming *Saccharomyces cerevisiae* strain expressing xylose isomerase. *Biotechnol. Bioeng.* 112, 470-483. doi:10.1002/bit.25447

Wiechert, W., Möllney, M., Petersen, S., de Graaf, A.A., 2001. A universal framework for ¹³C metabolic flux analysis. *Metab. Eng.* 3, 265–283. doi:10.1006/mben.2001.0188

Wiechert, W., Siefke, C., de Graaf, A.A., Marx, A., 1997. Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis. *Biotechnol. Bioeng.* 55, 118–135. doi:10.1002/(SICI)1097-0290(19970705)55:1<118::AID-BIT13>3.0.CO;2-I

Chapter 4

Kinetic Isotope Effects Significantly Influence Intracellular Metabolite ^{13}C Labeling Patterns and Flux Determination

This chapter is adapted from (Wasylenko and Stephanopoulos, 2013).

4.1 Introduction

As described in Chapter 1, an important step in ^{13}C -MFA is the prediction of the ^{13}C metabolite labeling patterns expected to result from a putative flux distribution by solution of a set of isotopomer balance equations. (The isotopomer balance equations are typically recast in terms of cumomers, EMUs, or something similar to facilitate the solution (Antoniewicz et al., 2007a; Wiechert et al., 1999)). However, these isotopomer balance equations implicitly assume the absence of isotope effects on the rates of the enzyme-catalyzed reactions of central carbon metabolism. That is, these equations inherently assume that the enzymes involved in central carbon metabolism will turn over all isotopomers of their substrate metabolites at the same rate. There is in fact a large body of literature showing that isotope effects do occur in central carbon metabolism. For instance it is well known that plants assimilating inorganic carbon preferentially assimilate ^{12}C over ^{13}C , and that the extent of the discrimination against ^{13}C depends on whether carbon is assimilated through Rubisco or PEP carboxylase (Whelan et al., 1970, 1973). More generally, ^{13}C atoms form stronger bonds than ^{12}C atoms so that the presence

of ^{13}C atoms in a metabolite is expected to slow the rate of its enzymatic conversion (Christensen and Nielsen, 1999; Heinzle et al., 2008). Nonetheless in ^{13}C -MFA it has commonly been assumed, either explicitly (Christensen and Nielsen, 1999; van Winden et al., 2001; Wiechert and de Graaf, 1997) or implicitly, that such carbon isotope effects are negligible in the prediction of the ^{13}C labeling states of intracellular metabolites. To our knowledge, the validity of this assumption has never been investigated quantitatively. Although Christensen and Nielsen concluded that isotope effects were unlikely to significantly affect ^{13}C labeling patterns of intracellular metabolites (Christensen and Nielsen, 1999), using Gas Chromatography-Combustion-Isotope Ratio Mass Spectrometry (GC-C-IRMS) Heinzle, *et al.* have shown that isotope effects do significantly influence ^{13}C labeling data in specialized ^{13}C -MFA experiments at low ^{13}C enrichments, and that correction for these isotope effects is necessary for flux estimation in these systems (Heinzle et al., 2008; Yuan et al., 2010). We sought to quantify the modeling error that could result from neglecting carbon isotope effects in traditional ^{13}C -MFA experiments, with standard levels of ^{13}C enrichment and conventional MS measurements used to obtain ^{13}C labeling data for flux estimation. As a case study, we investigated the potential for kinetic isotope effects to cause isotopic fractionation at the pyruvate node.

4.2 Modeling Isotopic Fractionation at the Pyruvate Node

It has been known for more than 50 years that lipids and the carboxyl group of leucine, both of which are derived from acetyl-CoA, are generally depleted in the ^{13}C isotope relative to other biomass constituents (Abelson and Hoering, 1961; Park and Epstein, 1961). Monson and Hayes used indirect evidence from measurement of positional ^{13}C enrichment in lipids to conclude that

this depletion of ^{13}C was a result of an isotope effect on the reaction catalyzed by pyruvate dehydrogenase (PDH) (Monson and Hayes, 1982), and the isotope effects associated with PDH enzymes from *Escherichia coli* and *Saccharomyces cerevisiae* have been characterized by Melzer and Schmidt (Melzer and Schmidt, 1987). However, the influence that these isotope effects will have on metabolite isotopomer distributions is difficult to predict, as isotopic fractionation due to isotope effects will depend on the metabolic flux distribution and the kinetic state of the system (DeNiro and Epstein, 1977; Schmidt, 2003). We performed simulations to investigate the potential consequences of isotope effects at the pyruvate node in ^{13}C -MFA labeling experiments.

To model the influence of isotope effects at the pyruvate branch point on intracellular metabolite isotopomer distributions, we constructed a simplified model of the pyruvate node (Figure 4-1). Pyruvate is produced from phosphoenolpyruvate in the reaction catalyzed by Pyruvate Kinase (PK). We assume that this reaction is effectively irreversible, and the flux through PK is denoted v^{PK} . We further assume that all reactions other than the PDH reaction that consume pyruvate can be lumped into a single reaction forming a metabolite X , as in (Monson and Hayes, 1982). These reactions that form metabolite X could include anaplerotic reactions, amino acid biosynthesis (e.g. synthesis of alanine), or production of byproducts such as lactate. We further assumed that the reaction converting pyruvate to X is irreversible and has no significant isotope effects. The latter assumption is justified by the fact that reactions in which carbon-carbon bonds are broken (such as the PDH reaction) are generally expected to have much larger ^{13}C isotope effects than reactions in which all carbon-carbon bonds are left intact (such as production of lactate or alanine from pyruvate). The flux through the reaction converting pyruvate to

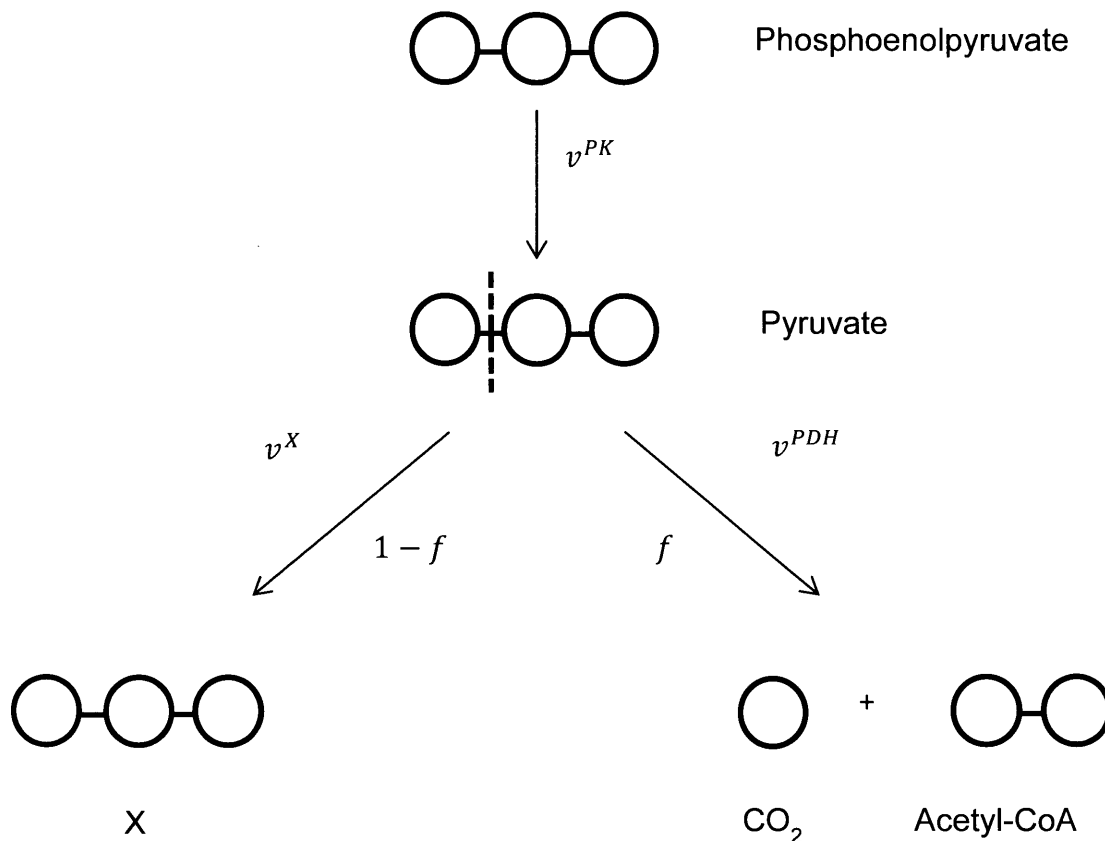


Figure 4-1 The pyruvate node model system. We assume that pyruvate is produced through Pyruvate Kinase (PK) and consumed through one of two reactions. In the Pyruvate Dehydrogenase (PDH) reaction, the bond between C_1 and C_2 of pyruvate is broken (represented by the dashed line), yielding one molecule of CO_2 and one acetyl-CoA two-carbon unit. This reaction is subject to isotope effects, as measured by Melzer and Schmidt (see text). In the other reaction, pyruvate is converted to a metabolite X. In this reaction all carbon-carbon bonds in pyruvate remain intact, so we assume the isotope effects on this reaction are small relative to those on the PDH reaction. The ratio of the flux through PDH to the flux through PK is equal to f , while the ratio of the flux to metabolite X to the flux through PK is equal to $1 - f$. Circles represent the carbon atoms of each metabolite.

metabolite X is denoted v^X . Finally, for simplicity we assumed that acetyl-CoA is synthesized exclusively from pyruvate through PDH. This reaction is assumed to be irreversible, and the flux from pyruvate to acetyl-CoA is denoted v^{PDH} .

Here we define some nomenclature that will be used in the following discussion. The positional isotopomers of pyruvate will be denoted ijk , where the binary variables i , j , and k represent the labeling states at C_1 , C_2 , and C_3 of pyruvate, respectively. These variables take on values $i, j, k = 0$ if the carbon at the respective position in the pyruvate molecule is ^{12}C and $i, j, k = 1$ if the carbon is ^{13}C . So for instance the isotopomer 001 has ^{12}C atoms at C_1 and C_2 and a ^{13}C atom at C_3 . The ^{13}C labeling state of pyruvate can be described by a set of eight isotopomer mole fractions y_{ijk}^P , where y_{ijk}^P is equal to the mole fraction of the total pyruvate pool with the labeling pattern ijk . The labeling state of the two-carbon unit of acetyl-CoA will similarly be represented by four isotopomer mole fractions y_{ij}^A .

At steady state, a mass balance on pyruvate dictates that the rate of pyruvate production must equal the rate of pyruvate consumption:

$$v^{PK} = v^X + v^{PDH} \quad (4.1)$$

Similarly, a mass balance can be written for each of the isotopomers of pyruvate:

$$v_{ijk}^{PK} = v_{ijk}^X + v_{ijk}^{PDH} \quad (4.2)$$

where v_{ijk}^{PK} is the rate at which pyruvate isotopomer ijk is produced through PK and v_{ijk}^X and v_{ijk}^{PDH} are the rates at which pyruvate isotopomer ijk is metabolized to metabolite X and acetyl-CoA, respectively. We define f to be the ratio of the flux through PDH to the flux through PK (Figure 4-1):

$$f \equiv \frac{v^{PDH}}{v^{PK}} \quad (4.3)$$

and the variables f_{ijk} are defined analogously for each of the isotopomers of pyruvate:

$$f_{ijk} \equiv \frac{v_{ijk}^{PDH}}{v_{ijk}^{PK}} \quad (4.4)$$

The total flux of pyruvate through PDH must equal the sum of the fluxes of each of the pyruvate isotopomers:

$$v^{PDH} = \sum v_{ijk}^{PDH} \quad (4.5)$$

Using the definition of f_{ijk} , this can be rewritten as:

$$v^{PDH} = \sum v_{ijk}^{PK} f_{ijk} \quad (4.6)$$

which is equivalent to:

$$f = \sum \tilde{v}_{ijk}^{PK} f_{ijk} \quad (4.7)$$

where we have introduced the “isotopomer flux fraction” \tilde{v}_{ijk}^{PK} :

$$\tilde{v}_{ijk}^{PK} \equiv \frac{v_{ijk}^{PK}}{v^{PK}} \quad (4.8)$$

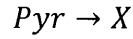
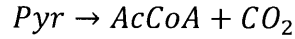
The isotopomer flux fraction \tilde{v}_{ijk}^{PK} is equal to the fraction of the total flux through PK that produces the pyruvate isotopomer ijk . We note that although under the typical assumption that isotope effects are negligible the isotopomer flux fractions are equal to the pyruvate isotopomer mole fractions y_{ijk}^P , in the case where isotope effects are present this is not the case; the y_{ijk}^P will deviate from the \tilde{v}_{ijk}^{PK} and equation (4.7) must be written as a flux fraction-weighted average rather than as a mole fraction-weighted average.

In order to quantify the effects of isotopic fractionation at the pyruvate branch point, we sought to determine the f_{ijk} . In the idealized case (without isotope effects),

$$f_{ijk}^{ideal} = f, \quad \forall i, j, k \quad (4.9)$$

i.e. the f_{ijk} are independent of i, j , and k . However, when isotope effects are present this will not be the case.

The split ratio at any branch point can be viewed as the ratio of two effective rate constants. In the model system described above, there are two competing reactions which consume pyruvate:



We assume the fluxes of these reactions can be approximated by equations (4.10) and (4.11), respectively:

$$v^{PDH} = k^{PDH}[Pyr] \quad (4.10)$$

$$v^X = k^X[Pyr] \quad (4.11)$$

where k^{PDH} is the effective rate constant for the PDH reaction, k^X is the effective rate constant for the reaction forming metabolite X , and $[Pyr]$ is the intracellular concentration of pyruvate. Substituting equation (4.1) into equation (4.3), f can be expressed as:

$$f = \frac{v^{PDH}}{v^{PDH} + v^X} \quad (4.12)$$

Substituting the expressions from equations (4.10) and (4.11) into (4.12), we have:

$$f = \frac{k^{PDH}}{k^{PDH} + k^X} \quad (4.13)$$

so that f is expressed as a ratio of rate constants. The f_{ijk} can similarly be expressed as ratios of rate constants:

$$f_{ijk} = \frac{k_{ijk}^{PDH}}{k_{ijk}^{PDH} + k^X} \quad (4.14)$$

where k_{ijk}^{PDH} is the rate constant for the PDH reaction with the pyruvate isotopomer ijk as the substrate. k^X is constant for all isotopomers of pyruvate since we have assumed the isotope effects on this reaction are negligible. The k_{ijk}^{PDH} can be expressed relative to k_{000}^{PDH} (which can be fixed to equal one). We define a set of β_{ijk} which are equal to the ratios $k_{ijk}^{PDH}/k_{000}^{PDH}$ so that the rate constants can be expressed as:

$$k_{ijk}^{PDH} = \beta_{ijk} k_{000}^{PDH} \quad (4.15)$$

The β_{ijk} for singly-labeled isotopomers of pyruvate can be determined directly from the data of Melzer and Schmidt, who have measured the isotope effects on C₁, C₂, and C₃ of pyruvate. These are denoted α_1 , α_2 , and α_3 , respectively:

$$\begin{aligned}
\alpha_1 &= \frac{k_{000}^A}{k_{100}^A} \\
\alpha_2 &= \frac{k_{000}^A}{k_{010}^A} \\
\alpha_3 &= \frac{k_{000}^A}{k_{001}^A}
\end{aligned} \tag{4.16}$$

Melzer and Schmidt have determined these parameters to be $\alpha_1 = 1.0093$ for *E. coli* and 1.0238 for *S. cerevisiae*, $\alpha_2 = 1.0213$ for *E. coli* and 1.0254 for *S. cerevisiae*, and $\alpha_3 = 1.0031$ for the enzymes of both organisms (Melzer and Schmidt, 1987).

The rates of reaction for isotopomers with multiple ^{13}C atoms were not measured by Melzer and Schmidt. Thus the rate constants for these isotopomers must be approximated from the rate constants for the singly-labeled isotopomers. We investigated two different cases in which the β_{ijk} are defined by the following two equations:

$$\beta_{ijk} = \left(i \times \frac{1}{\alpha_1} + (1 - i) \right) \times \left(j \times \frac{1}{\alpha_2} + (1 - j) \right) \times \left(k \times \frac{1}{\alpha_3} + (1 - k) \right) \tag{4.17}$$

$$\beta_{ijk} = \min \left[\left(i \times \frac{1}{\alpha_1} + (1 - i) \right), \left(j \times \frac{1}{\alpha_2} + (1 - j) \right), \left(k \times \frac{1}{\alpha_3} + (1 - k) \right) \right] \tag{4.18}$$

Equation (4.17) assumes that the isotope effects are roughly additive. Each β_{ijk} is equal to the product of the isotope effect contributions of the three individual carbon atoms of pyruvate, where the isotope effect contribution of each individual carbon atom is equal to one (if the carbon is ^{12}C) or the reciprocal (multiplicative inverse) of its respective α value (if the carbon is ^{13}C). The result is that multiply-labeled isotopomers will have rate constants that reflect the product of the isotope effects on each of their labeled atoms. Equation (4.18) assumes that the

isotope effects are not additive and that the ^{13}C atom with the largest isotope effect will dominate the kinetics. In this case each β_{ijk} is equal to the minimum of the individual carbon atom contributions, where the individual carbon atom contributions are the same as above. Thus multiply-labeled isotopomers will not have rate constants that are smaller than those of the singly-labeled isotopomers. This latter case will represent a lower bound on the error that can be expected to result from neglecting isotope effects in the enzymatic reactions of central carbon metabolism.

Under the assumptions listed above, the fractionation of isotopes at the pyruvate node (and therefore the error associated with neglecting isotope effects) will depend solely on f and the isotopomer flux fractions \tilde{v}_{ijk}^{PK} . To determine realistic values for the \tilde{v}_{ijk}^{PK} , we performed simulations based on a flux distribution similar to the one estimated in (Antoniewicz et al., 2007c) using five commonly used glucose tracer mixtures and glucose labeled to natural abundance as the substrates. For simplicity, we assumed that isotope effects in the metabolism of glucose to pyruvate were negligible. Under this assumption, the \tilde{v}_{ijk}^{PK} are equal to the y_{ijk}^P obtained from simulation of the normal isotopomer balance equations. The selected tracers were 1- ^{13}C -glucose (commonly used to estimate flux through the oxidative pentose phosphate pathway (Blank et al., 2005; Fendt and Sauer, 2010; Fischer and Sauer, 2003)), 1,2- $^{13}\text{C}_2$ -glucose (determined to be the optimal tracer for a mammalian cell network in (Metallo et al., 2009)), 3,4- $^{13}\text{C}_2$ -glucose (determined to be optimal for resolution of pyruvate carboxylase flux in (Crown et al., 2012)), 20% U- $^{13}\text{C}_6$ -glucose (Blank et al., 2005; Fendt and Sauer, 2010; Fischer and Sauer, 2003), and 75% 1- ^{13}C -glucose + 25% U- $^{13}\text{C}_6$ -glucose (the tracer mixture used in the original study (Antoniewicz et al., 2007c)). The isotopomer flux fractions resulting from application of

these tracers and natural abundance glucose given the flux distribution in (Antoniewicz et al., 2007c) and neglecting isotope effects in the metabolism of glucose to pyruvate are summarized in Table 4-1.

Table 4-1 Predicted Isotopomer Flux Fractions

Isotopomer Flux Fractions	Tracers					
	NA	1	1,2	3,4	20%U	75%1 + 25%U
\tilde{v}_{000}^{PK}	0.968	0.706	0.606	0.132	0.754	0.518
\tilde{v}_{001}^{PK}	0.011	0.243	0.069	0.005	0.019	0.187
\tilde{v}_{010}^{PK}	0.011	0.021	0.015	0.043	0.013	0.016
\tilde{v}_{011}^{PK}	1.2E-04	0.006	0.228	0.001	0.018	0.027
\tilde{v}_{100}^{PK}	0.011	0.017	0.017	0.755	0.029	0.031
\tilde{v}_{101}^{PK}	1.2E-04	0.005	0.039	0.009	0.002	0.008
\tilde{v}_{110}^{PK}	1.2E-04	0.002	0.012	0.052	0.008	0.011
\tilde{v}_{111}^{PK}	1.3E-06	5.0E-04	0.013	0.002	0.157	0.202

The isotopomer flux fractions \tilde{v}_{ijk}^{PK} are predicted for each of the five glucose tracer mixtures listed in the text and for glucose labeled to natural abundance. Abbreviations: NA = natural abundance glucose; 1 = 1-¹³C-glucose; 1,2 = 1,2-¹³C₂-glucose; 3,4 = 3,4-¹³C₂-glucose; 20%U = 20% U-¹³C₆-glucose; 75%1 + 25%U = 75% 1-¹³C-glucose + 25% U-¹³C₆-glucose.

For a given value of f , the isotopomer distribution of the two-carbon unit in acetyl-CoA, y_{ij}^A , can be predicted for a specified tracer and the associated isotopomer flux fractions \tilde{v}_{ijk}^{PK} in the following way: the choice of the organism determines the values of α_1 , α_2 , and α_3 . The β_{ijk} are then computed using either equation (4.17) or (4.18), depending on whether isotope effects are assumed to be additive or not. The rate constants k_{ijk}^{PDH} are computed using equation (4.15) (with k_{000}^{PDH} fixed equal to one—the results are independent of the value of k_{000}^{PDH}) and substituted into equation (4.14). Equations (4.14) and (4.7) then yield nine equations with nine unknowns (k^X and f_{ijk}), which can be solved using a non-linear equation solver (we used the Matlab function `fsolve.m`). Assuming there are no isotope effects on the reactions downstream of acetyl-CoA that significantly affect its labeling pattern, the isotopomer distribution of acetyl-CoA will then be given by:

$$\begin{aligned}
 y_{00}^{A \text{ isotope}} &= (f_{000}\tilde{v}_{000}^{PK} + f_{100}\tilde{v}_{100}^{PK})/f \\
 y_{01}^{A \text{ isotope}} &= (f_{001}\tilde{v}_{001}^{PK} + f_{101}\tilde{v}_{101}^{PK})/f \\
 y_{10}^{A \text{ isotope}} &= (f_{010}\tilde{v}_{010}^{PK} + f_{110}\tilde{v}_{110}^{PK})/f \\
 y_{11}^{A \text{ isotope}} &= (f_{011}\tilde{v}_{011}^{PK} + f_{111}\tilde{v}_{111}^{PK})/f
 \end{aligned} \tag{4.19}$$

In the idealized case (neglecting isotope effects), using equation (4.9) this can be simplified to:

$$\begin{aligned}
y_{00}^A{}^{ideal} &= \tilde{v}_{000}^{PK} + \tilde{v}_{100}^{PK} \\
y_{01}^A{}^{ideal} &= \tilde{v}_{001}^{PK} + \tilde{v}_{101}^{PK} \\
y_{10}^A{}^{ideal} &= \tilde{v}_{010}^{PK} + \tilde{v}_{110}^{PK} \\
y_{11}^A{}^{ideal} &= \tilde{v}_{011}^{PK} + \tilde{v}_{111}^{PK}
\end{aligned} \tag{4.20}$$

As mentioned in Chapter 1, ^{13}C labeling data for ^{13}C -MFA are often obtained from MS, which resolves isotopomers by molecular mass. Consequently, isotopomers with the same number of ^{13}C atoms cannot be differentiated and are lumped into a single “mass isotopomer,” which is simply a set of all the isotopomers of a metabolite with a particular mass (or equivalently, a particular number of ^{13}C atoms). For a metabolite with n carbons and a mass M when all of the carbon atoms are ^{12}C , MS will yield a set of $n + 1$ mass isotopomer mole fractions, where the i^{th} mole fraction is equal to the mole fraction of the total metabolite pool with mass $M + i$ ($i = 0, 1, 2, \dots, n$). This set of mole fractions is referred to as the “Mass Isotopomer Distribution” (MID) of the metabolite. The mole fraction of each mass isotopomer is equal to sum of the mole fractions of its constituent isotopomers; for instance the $M + 2$ mole fraction of the pyruvate MID will be equal to the sum $y_{011}^P + y_{101}^P + y_{110}^P$. The set of MIDs obtained from MS (for as many metabolites as can be detected and accurately quantitated) makes up the ^{13}C labeling dataset used in the flux estimation algorithms described above. In order to estimate the error in MS ^{13}C labeling data that could result from neglecting isotope effects, we converted both the $y_{ij}^A{}^{ideal}$ and $y_{ij}^A{}^{isotope}$ isotopomer distributions into MIDs:

$$\begin{aligned}
y_0^A &= y_{00}^A \\
y_1^A &= y_{01}^A + y_{10}^A \\
y_2^A &= y_{11}^A
\end{aligned} \tag{4.21}$$

where y_i^A is equal to the mole fraction of the $M + i$ mass isotopomer in the MID of the acetyl-CoA two-carbon unit. These equations hold for both the idealized case and the case with isotope effects included. The errors associated with the assumption that isotope effects are negligible can be estimated to be the difference between y_i^{Aideal} (the MID mole fractions expected to be obtained from a simulation which neglects isotope effects) and $y_i^{Aisotope}$ (the MID mole fractions expected to be obtained from a ^{13}C labeling experiment):

$$\begin{aligned}
 y_0^{Aerror} &= y_0^{Aideal} - y_0^{Aisotope} \\
 y_1^{Aerror} &= y_1^{Aideal} - y_1^{Aisotope} \\
 y_2^{Aerror} &= y_2^{Aideal} - y_2^{Aisotope}
 \end{aligned}
 \tag{4.22}$$

4.3 Results

We computed the y_i^{Aerror} as functions of f for each of the five glucose tracers listed above and for glucose labeled to natural abundance using the α_i values measured by Melzer and Schmidt for both the *E. coli* and *S. cerevisiae* PDH enzymes (Melzer and Schmidt, 1987). Two typical results for the *E. coli* PDH enzyme assuming additivity of isotopes effects are shown in Figure 4-2.

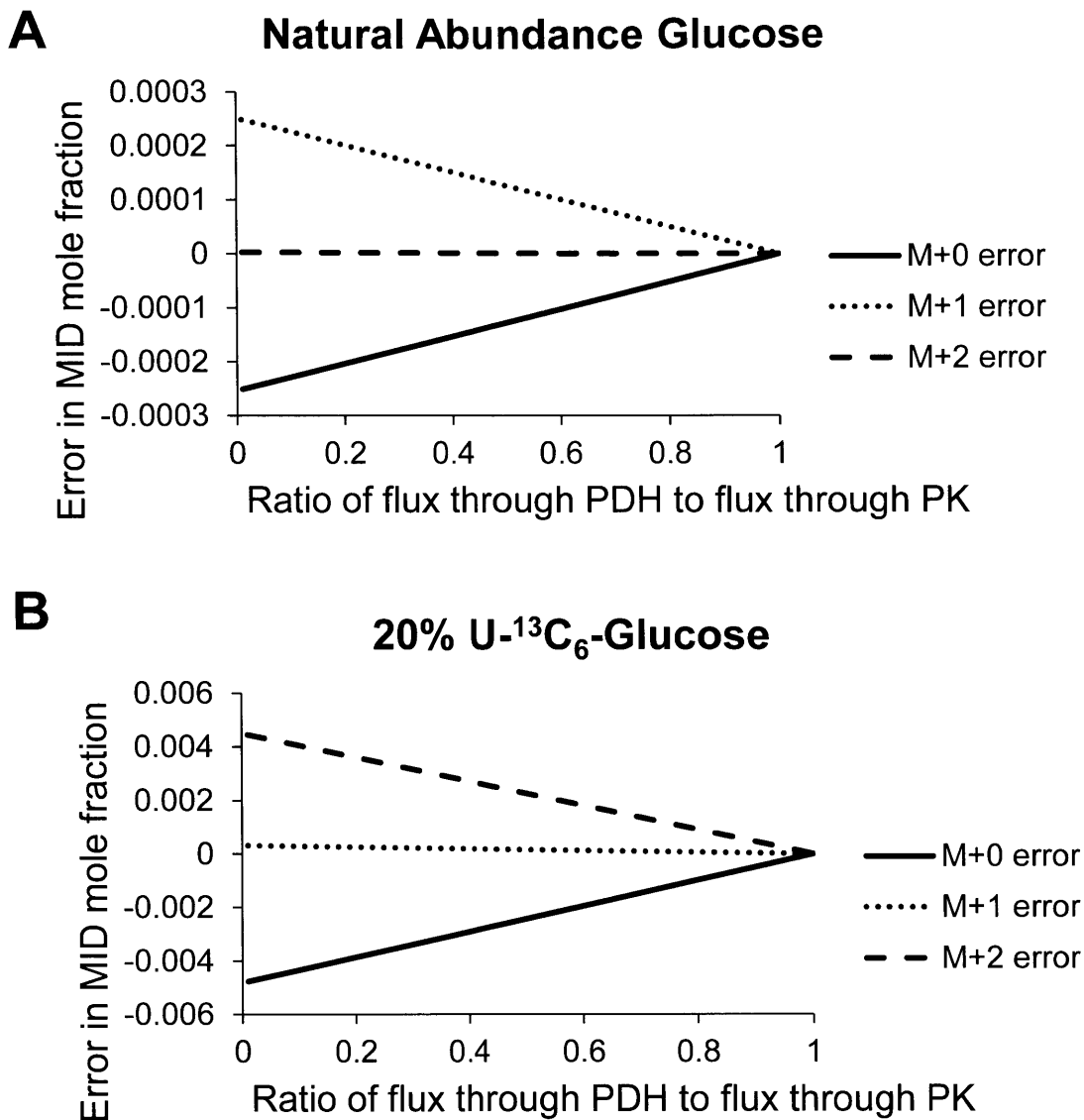


Figure 4-2 Typical results for errors associated with neglecting isotope effects as a function of f . The predicted modeling errors associated with neglecting isotope effects on the *E. coli* PDH enzyme (assuming additivity of isotope effects for multiply-labeled isotopomers of pyruvate) with natural abundance glucose (A) and 20% U-¹³C₆-glucose (B) substrates are plotted as a function of the ratio of flux through PDH to flux through PK, f . The errors for each of the three mass isotopomer mole fractions of the acetyl-CoA two-carbon unit MID are shown.

It can be seen that the magnitude of the error is maximum when f is small and decreases linearly as f increases, approaching zero as f approaches unity. This is a well-known result (DeNiro and Epstein, 1977; Monson and Hayes, 1982). There is zero error associated with isotope effects when $f = 1$ because isotopic fractionation can only occur at a branch point in the metabolic network (Schmidt, 2003). When metabolism is at steady state, if one enzyme discriminates against an isotopomer of its substrate metabolite then there must be another enzyme in the cell that turns over a disproportionately high fraction of that isotopomer due to conservation of mass. For instance, in the model system investigated above PDH discriminates against the heavy isotopomers of pyruvate and consequently the enzyme forming metabolite X will metabolize a disproportionately high fraction of these heavy isotopomers. If the conversion of the PDH reaction is 100%, the PDH enzyme must turn over all isotopomers of pyruvate completely and so there can be no discrimination against any of the pyruvate isotopomers at steady state.

Under the assumption of additive isotope effects, the error associated with the *E. coli* PDH enzyme with glucose labeled to natural abundance as the substrate approaches 0.025 mol % as f approaches zero (Figure 4-2). An identical result (to the nearest thousandth of a mole percent) is obtained if isotope effects are assumed to be non-additive (data not shown); because metabolism of glucose labeled to natural abundance produces very few multiply-labeled isotopomers, the predicted errors with this substrate are relatively insensitive to the assumption on additivity of isotope effects. The predicted maximum error of 0.025 mol % is in good agreement with the results of Monson and Hayes, who predicted acetyl-CoA would be depleted in ^{13}C relative to pyruvate by approximately 23 per mille for small f (see Figure 4 of (Monson and Hayes, 1982)). When the ^{13}C enrichment in glucose is increased by introduction of ^{13}C -labeled tracers, the errors

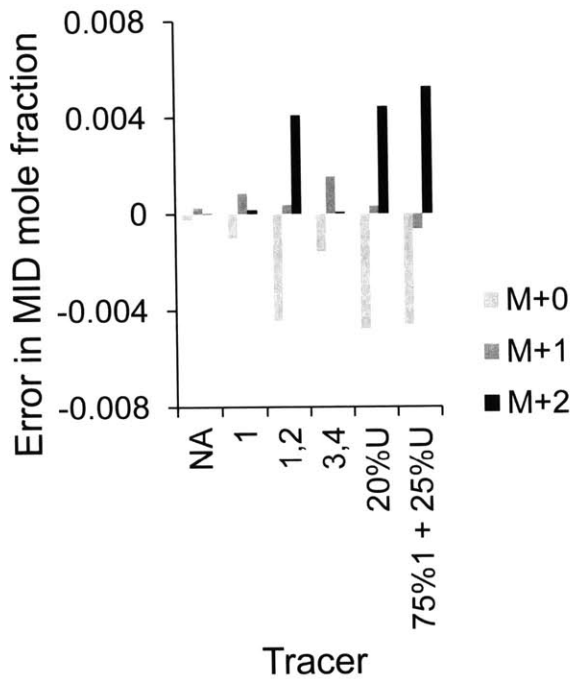
in MID measurements associated with neglecting isotope effects increase. The results with 20% U- $^{13}\text{C}_6$ -glucose as the tracer indicate that these errors can approach 0.5 mol % for the *E. coli* PDH enzyme for small f .

In order to investigate the effect of the organism (*E. coli* vs. *S. cerevisiae*), the glucose tracer, and the assumption on the additivity of isotope effects for multiply-labeled isotopomers of pyruvate (equations (4.17) and (4.18)), we visualized the error in each case with f fixed to a value of 1% (Figure 4-3). This represents something of a “worst case scenario” since the errors are maximum for small values of f . The errors associated with isotope effects are predicted to be larger in *S. cerevisiae* than in *E. coli*, which is expected since the α_1 and α_2 values measured by Melzer and Schmidt were greater for the *S. cerevisiae* PDH. The expected errors with a 1- ^{13}C -glucose tracer are relatively small because metabolism of this tracer yields pyruvate primarily labeled at C_3 , which does not participate in the bond that is broken in the PDH reaction and consequently has a relatively small isotope effect. The expected errors with a 3,4- $^{13}\text{C}_2$ -glucose tracer are also relatively minor. At first this may seem surprising since metabolism of this tracer yields pyruvate primarily labeled at C_1 , and there is a significant isotope effect on this carbon. However, the set of isotopomer flux fractions that result from application of this tracer is dominated by a single isotopomer (the isotopomer with labeling pattern 100). In fact, in this case 85% of the flux to pyruvate produces the isotopomers with the labeling patterns 100, 010, or 110, and the unlabeled isotopomer 000 accounts for only approximately 13% of the total flux to pyruvate. The small flux producing the unlabeled isotopomer, which is the isotopomer turned over at the highest rate by PDH, mitigates the effects of isotopic discrimination so that the errors associated with neglecting isotope effects are smaller than might be expected. The applications

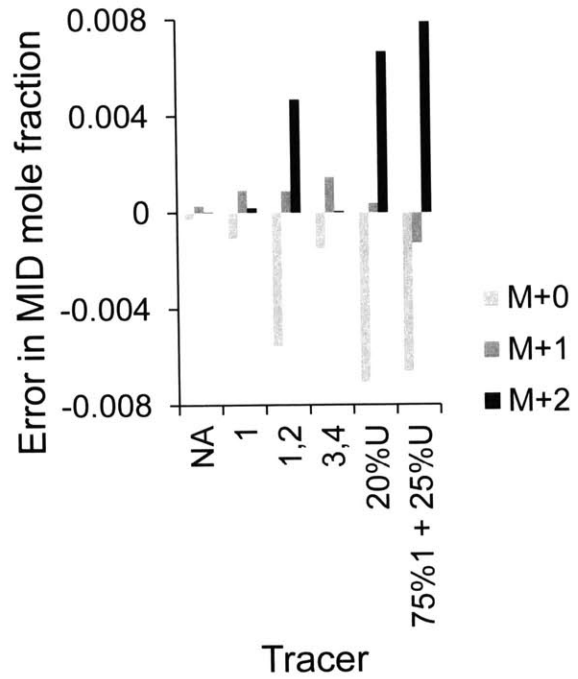
of the tracers 1,2-¹³C₂-glucose, 20% U-¹³C₆-glucose, and 75% 1-¹³C-glucose + 25% U-¹³C₆-glucose result in the highest expected errors. Metabolism of each of these tracers results in an unlabeled (000) isotopomer flux fraction greater than 50% and significant flux fractions for isotopomers labeled at one or both of the carbons with large isotope effects. Consequently, with these tracers neglecting isotope effects is expected to result in relatively large errors.

It can be seen from Figure 4-3 that, under the assumption that isotope effects are additive, the errors associated with isotope effects can exceed 0.5 mol % and can even approach 0.8 mol %. Antoniewicz, *et al.* have shown that MIDs of amino acids labeled to natural abundance can be measured to an accuracy of 0.4 mol % and with a precision of 0.2 mol % using Gas Chromatography (GC) coupled to MS (Antoniewicz et al., 2007b). Thus, in some cases the errors associated with isotope effects may exceed the measurement errors associated with the GC-MS instrument and could in fact be the dominant source of error in the modeling of GC-MS MID data for flux estimation. Even in the case where isotope effects are assumed to be completely non-additive, the errors associated with isotopic discrimination can approach the 0.4 mol % upper bound for error in GC-MS data which was given in (Antoniewicz et al., 2007b).

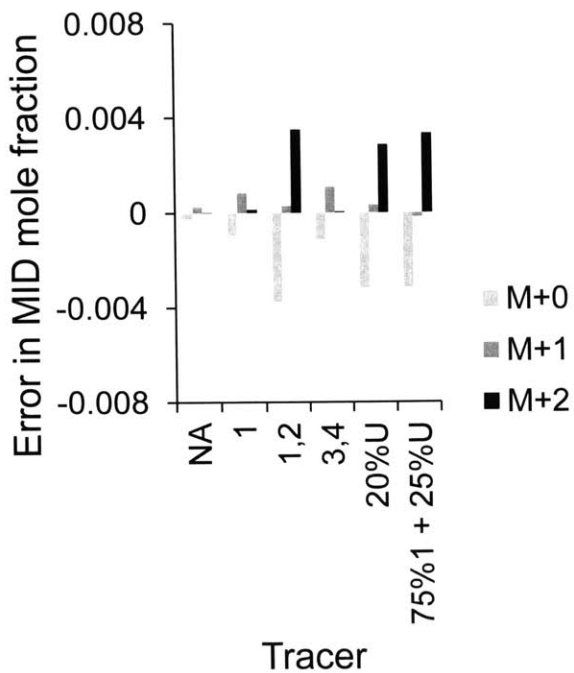
A *E. coli*, Additive Isotope Effects



B *S. cerevisiae*, Additive Isotope Effects



C *E. coli*, Non-additive Isotope Effects



D *S. cerevisiae*, Non-Additive Isotope Effects

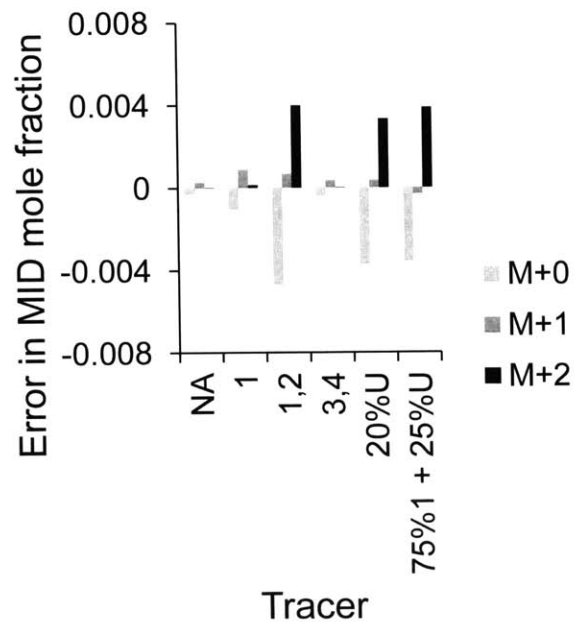


Figure 4-3 Expected errors associated with isotope effects at low values of f . The predicted modeling errors associated with neglecting isotope effects at a fixed value of $f = 0.01$ for different combinations of organism, tracer, and assumption on the additivity of isotope effects for multiply-labeled isotopomers of pyruvate. Predicted errors for each glucose substrate are plotted for the cases of *E. coli* assuming additive isotope effects (A), *S. cerevisiae* assuming additive isotope effects (B), *E. coli* assuming non-additive isotope effects (C), and *S. cerevisiae* assuming non-additive isotope effects (D). Abbreviations: NA = natural abundance glucose; 1 = 1-¹³C-glucose; 1,2 = 1,2-¹³C₂-glucose; 3,4 = 3,4-¹³C₂-glucose; 20%U = 20% U-¹³C₆-glucose; 75%1 + 25%U = 75% 1-¹³C-glucose + 25% U-¹³C₆-glucose.

4.4 Discussion

It was noted above that the errors in Figure 4-3 represent something of a “worst-case scenario.” Because the magnitudes of the errors associated with neglecting isotope effects decrease as f increases, in systems in which the majority of pyruvate is metabolized through PDH the actual errors will be significantly smaller. For these systems, the ratio of the actual error to the error shown in Figure 4-3 will be approximately equal to $1 - f$, since the magnitude of the error decreases linearly as f increases and is equal to zero when f equals unity. However in many cases the errors associated with isotope effects will be significant. For instance, a recent ¹³C-MFA study in the A549 cancer cell line revealed that in these cells the majority of pyruvate was metabolized to lactate so that in this system f was only approximately 10% (Metallo et al., 2009). Thus the errors associated with the acetyl-CoA two-carbon unit MID in ¹³C-MFA studies in cancer cells may approach the errors shown in Figure 4-3.

Moreover, the dependence of reaction kinetics on ^{13}C -labeling is not unique to the PDH enzyme. In fact, for enzymes catalyzing carbon-carbon bond-breaking reactions, significant isotope effects appear to be the norm rather than the exception. Gleixner, *et al.* observed an isotope effect of 1.6% on C_3 of fructose-1,6-bisphosphate (FBP) for FBP aldolase from rabbit muscle (although the isotope effect on C_4 was negligible) (Gleixner and Schmidt, 1997), and Hermes, *et al.* determined the isotope effect on C_4 of malate to be approximately 3% for malic enzyme from chicken liver (Hermes *et al.*, 1982). Significant isotope effects have also been observed for the enzymes of the oxidative pentose phosphate pathway. Hermes, *et al.* determined the isotope effect on C_1 of glucose-6-phosphate (G6P) to be 1.65% for the G6P dehydrogenase from *Leuconostoc mesenteroides* (Hermes *et al.*, 1982), while the isotope effect on C_1 of 6-phosphogluconate (6PG) has been estimated to be anywhere from 0.59% to 2.25% for 6PG dehydrogenases from various sources (Hwang *et al.*, 1998; Hwang and Cook, 1998; Rendina *et al.*, 1984). The results for 6PG dehydrogenase (and the results for the PDH enzymes of *E. coli* and *S. cerevisiae* discussed above) reveal that the isotope effect on a single chemical reaction can vary significantly across different enzymes and different organisms. In fact, the isotope effect associated with a single enzyme can in some cases be highly condition-dependent—Grissom and Cleland measured the isotope effect on C_6 of isocitrate for NADP-dependent isocitrate dehydrogenase over a range of pH values and found that the isotope effect varied from 0.3% (pH 7.0) to 2.76% (pH 4.1) (Grissom and Cleland, 1988). The presence of significant isotope effects in the kinetics of such a large number of the enzymes of central carbon metabolism and the variation in the magnitudes of these effects across different organisms and under different conditions makes it difficult to predict the impact that isotopic discrimination will have on the ^{13}C labeling patterns of intracellular metabolites. However, given the preceding analysis (in

which we neglected isotope effects in all reactions upstream and downstream of PDH for simplicity) and the ubiquity of enzymes with kinetics subject to isotope effects in the various pathways of central metabolism, it seems likely that isotope effects will exert significant influence on the distribution of ^{13}C label in metabolic systems.

Rigorous interpretation of ^{13}C -MFA results requires an accurate assessment of the measurement errors in the ^{13}C labeling data. Knowledge of the measurement errors is necessary to conduct goodness-of-fit tests, which quantify how well a set of labeling data fit a metabolic model, and to compute confidence intervals for the estimated fluxes, which is necessary to determine if differences in estimated fluxes in different strains or under different conditions are statistically significant (Antoniewicz et al., 2006; Möllney et al., 1999; van Winden et al., 2001). Failure to account for any significant source of error in a ^{13}C -MFA can potentially lead to underestimation of the errors associated with labeling data. This can result in a biased estimated metabolic flux distribution, failure of a chi-square test for goodness-of-fit even with a complete model metabolic network, and overly optimistic estimation of the confidence intervals for the computed fluxes. To-date, the most common methods for determining the magnitude of the error in MS data have been measurement of MIDs of metabolites labeled to natural abundance (for which the MIDs are theoretically independent of the flux distribution and can be computed from the well-known natural abundances of heavy isotopes) (Antoniewicz et al., 2007b; Dauner and Sauer, 2000; van Winden et al., 2005), calculation of standard deviations from replicate samples (which is really a measure of precision, not accuracy) (van Winden et al., 2005; Wittmann and Heinzle, 2002), and comparison of labeling data for metabolites which should in principle have identical labeling patterns (i.e. comparison of redundant data) (Schmidt et al., 1999). None of these

methods are capable of quantifying the errors associated with neglecting isotope effects in central carbon metabolism. These errors are structural in nature and can only be quantified by modeling the impact of isotope effects on the distribution of ^{13}C atoms in intracellular metabolites, which at present seems to be a challenging task. We further note that as long as the isotopomer balance equations used for prediction of the labeling state associated with a flux distribution neglect isotope effects, the errors associated with isotope effects place a bound on the accuracy that can be achieved in modeling ^{13}C labeling data. No matter how accurate the measurement technique, the errors will never be smaller than those associated with the assumption that isotope effects are negligible.

In this contribution, we have modeled isotopic fractionation at the pyruvate node to quantitatively assess the potential modeling errors associated with kinetic isotope effects. We have also proposed a framework that can be used to estimate modeling errors associated with isotope effects at other branch points in metabolism. Our results show that under some conditions the modeling errors associated with isotope effects are significant and will be comparable in size to the measurements errors associated with GC-MS. The quantification of these errors will facilitate accurate interpretation of the results of future ^{13}C -MFA studies.

4.5 References

Abelson, P.H., Hoering, T.C., 1961. Carbon isotope fractionation in formation of amino acids by photosynthetic organisms. *Proc. Natl. Acad. Sci. U.S.A.* 47, 623–632.

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2006. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* 8, 324–337. doi:10.1016/j.ymben.2006.01.004

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007a. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng.* 9, 68–86. doi:10.1016/j.ymben.2006.09.001

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007b. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Anal. Chem.* 79, 7554–7559. doi:10.1021/ac0708893

Antoniewicz, M.R., Kraynie, D.F., Laffend, L.A., González-Lergier, J., Kelleher, J.K., Stephanopoulos, G., 2007c. Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol. *Metab. Eng.* 9, 277–292. doi:10.1016/j.ymben.2007.01.003

Blank, L.M., Kuepfer, L., Sauer, U., 2005. Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* 6, R49. doi:10.1186/gb-2005-6-6-r49

Christensen, B., Nielsen, J., 1999. Isotopomer analysis using GC-MS. *Metab. Eng.* 1, 282–290. doi:10.1006/mben.1999.0117

Crown, S.B., Ahn, W.S., Antoniewicz, M.R., 2012. Rational design of ¹³C-labeling experiments for metabolic flux analysis in mammalian cells. *BMC Syst Biol* 6, 43. doi:10.1186/1752-0509-6-43

Dauner, M., Sauer, U., 2000. GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing. *Biotechnol. Prog.* 16, 642–649. doi:10.1021/bp000058h

DeNiro, M.J., Epstein, S., 1977. Mechanism of carbon isotope fractionation associated with lipid synthesis. *Science* 197, 261–263.

Fendt, S.-M., Sauer, U., 2010. Transcriptional regulation of respiration in yeast metabolizing differently repressive carbon substrates. *BMC Syst Biol* 4, 12. doi:10.1186/1752-0509-4-12

Fischer, E., Sauer, U., 2003. Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *Eur. J. Biochem.* 270, 880–891.

Gleixner, G., Schmidt, H.L., 1997. Carbon isotope effects on the fructose-1,6-bisphosphate aldolase reaction, origin for non-statistical ^{13}C distributions in carbohydrates. *J. Biol. Chem.* 272, 5382–5387.

Grissom, C.B., Cleland, W.W., 1988. Isotope effect studies of the chemical mechanism of pig heart NADP isocitrate dehydrogenase. *Biochemistry* 27, 2934–2943.

Heinzle, E., Yuan, Y., Kumar, S., Wittmann, C., Gehre, M., Richnow, H.-H., Wehrung, P., Adam, P., Albrecht, P., 2008. Analysis of ^{13}C labeling enrichment in microbial culture applying metabolic tracer experiments using gas chromatography-combustion-isotope ratio mass spectrometry. *Anal. Biochem.* 380, 202–210. doi:10.1016/j.ab.2008.05.039

Hermes, J.D., Roeske, C.A., O'Leary, M.H., Cleland, W.W., 1982. Use of multiple isotope effects to determine enzyme mechanisms and intrinsic isotope effects. Malic enzyme and glucose-6-phosphate dehydrogenase. *Biochemistry* 21, 5106–5114.

Hwang, C.C., Berdis, A.J., Karsten, W.E., Cleland, W.W., Cook, P.F., 1998. Oxidative decarboxylation of 6-phosphogluconate by 6-phosphogluconate dehydrogenase proceeds by a stepwise mechanism with NADP and APADP as oxidants. *Biochemistry* 37, 12596–12602.
doi:10.1021/bi980611s

Hwang, C.C., Cook, P.F., 1998. Multiple isotope effects as a probe of proton and hydride transfer in the 6-phosphogluconate dehydrogenase reaction. *Biochemistry* 37, 15698–15702.
doi:10.1021/bi981288w

Melzer, E., Schmidt, H.L., 1987. Carbon isotope effects on the pyruvate dehydrogenase reaction and their importance for relative carbon-13 depletion in lipids. *J. Biol. Chem.* 262, 8159–8164.

Metallo, C.M., Walther, J.L., Stephanopoulos, G., 2009. Evaluation of ¹³C isotopic tracers for metabolic flux analysis in mammalian cells. *J. Biotechnol.* 144, 167–174.
doi:10.1016/j.jbiotec.2009.07.010

Möllney, M., Wiechert, W., Kownatzki, D., de Graaf, A.A., 1999. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. *Biotechnol. Bioeng.* 66, 86–103.

Monson, K.D., Hayes, J., 1982. Carbon isotopic fractionation in the biosynthesis of bacterial fatty acids. Ozonolysis of unsaturated fatty acids as a means of determining the intramolecular distribution of carbon isotopes. *Geochimica et Cosmochimica Acta* 46, 139–149.
doi:10.1016/0016-7037(82)90241-1

Park, R., Epstein, S., 1961. Metabolic fractionation of C¹³ & C¹² in plants. *Plant Physiol.* 36, 133–138.

Rendina, A.R., Hermes, J.D., Cleland, W.W., 1984. Use of multiple isotope effects to study the mechanism of 6-phosphogluconate dehydrogenase. *Biochemistry* 23, 6257–6262.

Schmidt, H.-L., 2003. Fundamentals and systematics of the non-statistical distributions of isotopes in natural compounds. *Naturwissenschaften* 90, 537–552. doi:10.1007/s00114-003-0485-5

Schmidt, K., Nørregaard, L.C., Pedersen, B., Meissner, A., Duus, J.O., Nielsen, J.O., Villadsen, J., 1999. Quantification of intracellular metabolic fluxes from fractional enrichment and ^{13}C - ^{13}C coupling constraints on the isotopomer distribution in labeled biomass components. *Metab. Eng.* 1, 166–179. doi:10.1006/mben.1999.0114

Van Winden, W.A., van Dam, J.C., Ras, C., Kleijn, R.J., Vinke, J.L., van Gulik, W.M., Heijnen, J.J., 2005. Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of (^{13}C)-labeled primary metabolites. *FEMS Yeast Res.* 5, 559–568. doi:10.1016/j.femsyr.2004.10.007

Van Winden, W., Verheijen, P., Heijnen, S., 2001. Possible pitfalls of flux calculations based on (^{13}C)-labeling. *Metab. Eng.* 3, 151–162. doi:10.1006/mben.2000.0174

Wasylenko, T.M., Stephanopoulos, G., 2013. Kinetic isotope effects significantly influence intracellular metabolite (^{13}C) labeling patterns and flux determination. *Biotechnol J* 8, 1080–1089. doi:10.1002/biot.201200276

Whelan, T., Sackett, W.M., Benedict, C.R., 1970. Carbon isotope discrimination in a plant possessing the C_4 dicarboxylic acid pathway. *Biochem. Biophys. Res. Commun.* 41, 1205–1210.

Whelan, T., Sackett, W.M., Benedict, C.R., 1973. Enzymatic fractionation of carbon isotopes by phosphoenolpyruvate carboxylase from c(4) plants. *Plant Physiol.* 51, 1051–1054.

Wiechert, W., de Graaf, A.A., 1997. Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments. *Biotechnol. Bioeng.* 55, 101–117. doi:10.1002/(SICI)1097-0290(19970705)55:1<101::AID-BIT12>3.0.CO;2-P

Wiechert, W., Möllney, M., Isermann, N., Wurzel, M., de Graaf, A.A., 1999. Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. *Biotechnol. Bioeng.* 66, 69–85.

Wittmann, C., Heinzle, E., 2002. Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing corynebacteria. *Appl. Environ. Microbiol.* 68, 5843–5859.

Yuan, Y., Yang, T.H., Heinzle, E., 2010. ¹³C metabolic flux analysis for larger scale cultivation using gas chromatography-combustion-isotope ratio mass spectrometry. *Metab. Eng.* 12, 392–400. doi:10.1016/j.ymben.2010.02.001

Chapter 5

Metabolomic and ^{13}C -Metabolic Flux Analysis of a Xylose-Consuming *Saccharomyces cerevisiae* Strain Expressing Xylose Isomerase

This chapter is adapted from (Wasylenko and Stephanopoulos, 2015).

5.1 Introduction

In recent years, increasing energy demand and concerns about climate change and the sustainability of heavy reliance on fossil fuels have motivated the development of technologies for production of liquid fuels from plant biomass. Ethanol, which serves as a fuel additive, can be readily produced by fermentation of hexose sugars derived from cornstarch and sucrose. However, production of feed stocks such as corn and sugarcane requires large amounts of arable land and may compete with the food supply. Lignocellulosic feed stocks represent an attractive alternative, but technologies for production of liquid fuels from lignocellulosic material are relatively immature. The Baker's yeast *Saccharomyces cerevisiae* is a promising biocatalyst for production of liquid fuels from lignocellulosic biomass because its high rates of ethanol production under anaerobic conditions and high ethanol tolerance allow ethanol to be produced at high yield, productivity, and final titer. *S. cerevisiae* also exhibits relatively high tolerance to

inhibitors such as furan derivatives, weak acids, and phenolics present in lignocellulosic hydrolysates (Almeida et al., 2007; Lau et al., 2010), and the insusceptibility of yeast to bacteriophage and its ability to grow at low pH minimize the risk of contamination, allowing the avoidance of costs associated with reactor sterilization in industrial processes. However *S. cerevisiae* cannot natively metabolize the pentose sugars xylose and arabinose, which make up more than one-third of the carbohydrate biomass in some agricultural residues such as corn stover, wheat straw, and bagasse, with xylose being by far the more abundant of the two (van Maris et al., 2006). For production of biofuels from lignocellulosic feed stocks to be cost-effective, it will be necessary to effect the conversion of all sugars present in hydrolysates to liquid fuels (Stephanopoulos, 2007; Carroll and Somerville, 2009).

The ability to consume xylose can be conferred on *S. cerevisiae* strains by introduction of a heterologous pathway for conversion of xylose to its isomer xylulose. While many bacteria use a xylose isomerase (XI) enzyme to catalyze this conversion directly without the use of pyridine nucleotide cofactors, xylose-consuming eukaryotes generally effect the isomerization through a two-step redox pathway in which xylose reductase (XR) first catalyzes the reduction of xylose to xylitol, which is then oxidized via xylitol dehydrogenase (XDH) to form xylulose. Initial attempts to express heterologous *xylA* (encoding XI) genes in *S. cerevisiae* were unsuccessful; in several cases putative *xylA* transcripts were detected in Northern blots but putative XI protein products were insoluble and inactive (Sarthy et al., 1987; Amore et al., 1989; Gárdonyi and Hahn-Hägerdal, 2003). Consequently, the majority of xylose-consuming strains have been constructed using the XR-XDH pathway. However, while XR uses NADPH as its preferred cofactor substrate, XDH is strictly NAD⁺-dependent. This mismatch in cofactor specificities

results in a “cofactor imbalance” whereby NADP⁺ and NADH accumulate (and NADPH and NAD⁺ are depleted). The accumulation of NADH is especially problematic under industrially relevant anaerobic conditions. Without oxygen as a terminal electron acceptor, NADH cannot be efficiently re-oxidized to NAD⁺, severely inhibiting xylose metabolism (Bruinenberg et al., 1983; Bruinenberg et al., 1984). In early xylose-consuming *S. cerevisiae* strains the low availability of NAD⁺ for the XDH reaction also resulted in secretion of large amounts of the by-product xylitol (Kötter and Ciriacy, 1993; Tantirungkij et al., 1993), compromising ethanol yield.

A major breakthrough occurred with the discovery that the anaerobic fungus *Piromyces* sp. strain E2 metabolizes xylose using the XI pathway (Harhangi et al., 2003). The XI from this organism was functionally expressed in *S. cerevisiae* (Kuyper et al., 2003), and both evolutionary and rational metabolic engineering were used to construct efficient xylose-consuming strains capable of anaerobic growth on xylose (Kuyper et al., 2004; Kuyper et al., 2005a; Kuyper et al., 2005b). Our lab has recently used a similar approach along with a multi-stage evolutionary strategy to engineer the *S. cerevisiae* strain H131-A3-AL^{CS}, the fastest xylose-consuming strain reported to date (Zhou et al., 2012). However, the rates of growth and ethanol production on xylose are still significantly lower than those on glucose for reasons that are not completely understood.

Many hypotheses for bottlenecks in xylose metabolism have been presented. These include xylose transport, which may be especially limiting at low extracellular xylose concentrations (Gárdonyi et al., 2003; Runquist et al., 2009a; Runquist et al., 2010; Young et al., 2012); conversion of xylose to xylulose (Lönn et al., 2003; Jeppsson et al., 2003; Karhumaa et al., 2005; Kim et al., 2012); phosphorylation of xylulose (Toivari et al., 2001; Jin et al., 2003); and

conversion of xylulose-5-phosphate (X5P) to the glycolytic intermediates fructose-6-phosphate (F6P) and glyceraldehyde-3-phosphate (GAP) via the reactions of the non-oxidative Pentose Phosphate Pathway (PPP) (Walfridsson et al., 1995; Kuyper et al., 2005a). Once xylose is metabolized to F6P and GAP xylose metabolism is in principle identical to glucose metabolism. However while glucose induces a strong carbon catabolite repression (CCR) response in *S. cerevisiae* (Gancedo, 2008), there is evidence that xylose is not recognized as a fermentable carbon source and fails to fully activate the CCR program (Jin et al., 2004; Salusjärvi et al., 2008). Consequently, bottlenecks downstream of GAP could result from altered gene expression during xylose utilization.

In this study, we sought to characterize the metabolism of strain H131-A3-AL^{CS} by quantifying central carbon metabolite pool sizes and using ¹³C-Metabolic Flux Analysis (MFA) to estimate the fluxes through central metabolism in order to identify rate-limiting steps in xylose utilization. Although previous metabolomic and MFA studies have been conducted on xylose-consuming *S. cerevisiae* strains, many of these studies investigated strains that utilized the XR-XDH pathway to effect xylose isomerization (Wahlbom et al., 2001; Pitkänen et al., 2003; Sonderegger et al., 2004; Grotkjaer et al., 2005; Klimacek et al., 2010; Feng and Zhao, 2013a; Feng and Zhao, 2013b; Matsushika et al., 2013). Consequently, in these strains the redox cofactor imbalance is expected to exert a large influence on metabolism. Moreover, many of the strains employed exhibited low xylose consumption rates and growth rates, and in several cases analysis of metabolite pool sizes revealed signs of carbon starvation (Klimacek et al., 2010; Bergdahl et al., 2012; Matsushika et al., 2013). In this study, we investigated a strain capable of rapid xylose utilization via the xylose isomerase pathway in order to observe the differences between glucose

and xylose metabolism under aerobic and anaerobic conditions with high rates of xylose consumption and in the absence of the redox cofactor imbalance associated with action of the XR-XDH pathway. We present indirect evidence that in this strain there is an apparent bottleneck in xylose metabolism downstream of GAP, in the lower glycolysis pathway.

5.2 Materials and Methods

5.2.1 Strain and Culture Conditions

All experiments were performed with a previously engineered xylose-consuming *S. cerevisiae* strain similar to H131-A3-AL^{CS} (Zhou et al., 2012). The strain was cultivated at 30 °C in minimal medium containing 6.7 g/l Yeast Nitrogen Base (YNB) without amino acids (Difco) as a source of salts, vitamins, and trace elements and 20 g/l of either glucose (YNBG) or xylose (YNBX) as the sole carbon source. Media were supplemented with 0.42 g/l Tween 80 and 0.01 g/l ergosterol to facilitate growth under anaerobic conditions (Andreasen and Stier, 1953; Andreasen and Stier, 1954).

The yeast strain was cultivated in both YNBG and YNBX medium under both aerobic and anaerobic conditions, resulting in a total of four culture conditions: Glucose Aerobic (GA), Glucose Anaerobic (GN), Xylose Aerobic (XA), and Xylose Anaerobic (XN). For each culture condition, 5 ml starter cultures were inoculated from 15% glycerol -80 °C freezer stocks. 50 ml main cultures grown in 250 ml bottles (VWR, 89000-236) were inoculated with 50 µl (0.1% by volume) of the starter culture with the same carbon source and aeration condition. All cultures were shaken at 250 rpm to facilitate mixing.

Aerobic starter cultures were grown in aerobic culture tubes (BD Falcon, #352059). The 250 ml aerobic main culture bottle caps were left loose, allowing exchange between the culture headspace and the (aerobic) ambient atmosphere. Anaerobic starter cultures were grown in airtight Hungate tubes (ChemGlass, #CLS-4208-01). Anaerobic starter culture media were sparged with Ultra High Purity (UHP) Nitrogen (Airgas) for 3 min and left in an anaerobic chamber (Coy Labs) under an atmosphere of nitrogen and hydrogen for approximately 12 h prior to inoculation. Anaerobic main cultures were grown with 250 ml bottle caps completely tightened, preventing aeration from the ambient atmosphere. Anaerobic main culture media were sparged with UHP Nitrogen for 15 min and left in the anaerobic chamber for 12 h prior to inoculation. Inoculation and all sampling were conducted inside the anaerobic chamber.

For ^{13}C -MFA, the YNBG and YNBX carbon sources were 20 g/l 1,2- $^{13}\text{C}_2$ -glucose and 20 g/l 1,2- $^{13}\text{C}_2$ -xylose (Cambridge Isotope Laboratories), respectively. The ^{13}C -MFA cultures were otherwise identical.

5.2.2 Analytical Methods

Cell density was monitored by measuring the absorbance (“optical density”) of yeast cultures at 600 nm (OD_{600}) using an Ultrospec 2100 *pro* UV/Visible Spectrophotometer (Amersham Biosciences). Cell density was calculated using pre-determined correlations between Dry Cell Weight (DCW) and OD_{600} (see section 5.2.3 below). Glucose, xylose, ethanol, glycerol, and acetate concentrations were determined by High-Performance Liquid Chromatography (HPLC). Culture samples were filtered through 0.20 μm Nylon syringe filters (Microliter Analytical #F13-2020-1GF) and supernatants were analyzed on an Agilent 1200 series HPLC system equipped

with a refractive index detector. Analytes were separated on an Aminex HPX-87H column (Bio-Rad) with 5 mM sulfuric acid mobile phase at a flow rate of 0.6 ml/min and a temperature of 55 °C.

5.2.3 Determination of Optical Density-Dry Cell Weight Correlations

Triplicate cultures were prepared for each culture condition: GA, GN, XA, and XN. OD₆₀₀ was measured for each culture. Dry cell weight was subsequently determined by vacuum-filtering 40 ml culture with a pre-weighed cellulose nitrate filter (0.2 µm, Whatman #10401312). Each filter was washed with two volumes Millipore water and dried at 75-80 °C. Filters were weighed intermittently; it was observed that the mass of each filter decreased over time, reached a minimum, and then increased, presumably due to oxidation of biomass. The minimum mass was used for determination of dry cell weight. Control samples were prepared in triplicate by filtering 40 ml medium without cells to correct dry cell weight measurements for insolubles in the medium and changes in filter mass during drying.

For each culture condition j , DCW was plotted against OD₆₀₀ and the slope of the best-fit line k_j^{DCW} and the standard error of the slope σ_j^{DCW} were calculated using the Microsoft Excel function LINEST. k_j^{DCW} is equal to the biomass concentration that is equivalent to one OD₆₀₀ unit for culture condition j .

5.2.4 Estimation of Extracellular Fluxes

Specific growth rates, specific sugar consumption rates, and specific ethanol, glycerol, and acetate production rates were estimated from OD₆₀₀ and HPLC data from five biological

replicate cultures for each culture condition. For each biological replicate culture i the specific growth rate μ_i was determined by fitting the OD₆₀₀ data to a function of the form:

$$OD(t) = OD_0 \exp(\mu_i t)$$

where $OD(t)$ is the measured value of OD₆₀₀ at time t and OD_0 is the fitted value of OD₆₀₀ at time zero. Parameter estimation was achieved using the Matlab function `lsqnonlin`.

To determine specific metabolite consumption and production rates for biological replicate culture i , the concentration of each metabolite k was plotted against OD₆₀₀ and the slope of the best-fit line m_{ik} was computed. For each culture condition j the average specific growth rate μ_j and the average slopes m_{jk} were computed by averaging the μ_i and m_{ik} , respectively. Uncertainties were assumed to be equal to the standard deviations of the μ_i and m_{ik} . The specific consumption or production rate of metabolite k in culture condition j is then equal to:

$$q_{jk} = \frac{m_{jk} \mu_j}{k_j^{DCW}}$$

where k_j^{DCW} is the biomass concentration that is equivalent to one OD₆₀₀ unit (see above). The error in q_{jk} , denoted σ_{jk}^q , was estimated by assuming the relative error in q_{jk} is equal to the sum of the relative errors in μ_j , m_{jk} , and k_j^{DCW} :

$$\frac{\sigma_{jk}^q}{q_{jk}} = \frac{\sigma_j^\mu}{\mu_j} + \frac{\sigma_{jk}^m}{m_{jk}} + \frac{\sigma_j^{DCW}}{k_j^{DCW}}$$

5.2.5 Preparation of Uniformly ^{13}C -Labeled Cell Extract for Internal Standard

Cells were grown in medium containing 10 g/l U- $^{13}\text{C}_6$ -glucose as the sole carbon source and 6.7 g/l YNB without amino acids. 5 ml aerobic starter cultures were inoculated from freezer stocks (as above). 40 ml shake flask cultures were inoculated from starter cultures and harvested at $\text{OD}_{600} \sim 1$. Uniformly ^{13}C -labeled metabolite extracts were prepared using a protocol similar to the one described for preparation of samples for pool size measurements and ^{13}C -MFA (see below). Extracts were stored at $-80\text{ }^\circ\text{C}$ and subsequently dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module and resuspended in 100 μl Millipore water. A total of 48 extracts (four extracts from each of 12 shake flasks) were pooled and the pooled cell extract was aliquoted. Aliquots were stored at $-80\text{ }^\circ\text{C}$ for future use as internal standard.

5.2.6 Metabolite Extractions for Pool Size Measurements and ^{13}C -Metabolic Flux Analysis

Cultures were harvested in mid-exponential phase ($\text{OD}_{600} \sim 0.5\text{-}0.6$). 7.5 ml culture was quenched in 37.5 ml pure methanol (Canelas et al., 2008) held at low temperature ($< -70\text{ }^\circ\text{C}$) in a cold ethanol bath. Samples were centrifuged for 5 min at 3270g and $-10\text{ }^\circ\text{C}$. Centrifuge adapters for holding sample tubes were pre-cooled in a $-80\text{ }^\circ\text{C}$ freezer to keep the sample temperature as low as possible during centrifugation. The supernatants were discarded and cell pellets were washed by resuspension in 40 ml cold ($< -70\text{ }^\circ\text{C}$) methanol. Samples were centrifuged a second time and supernatants discarded. 50 μl ^{13}C -labeled cell extract was added to cell pellets for pool size measurement samples (which are labeled to natural abundance) only for use as internal standard (Wu et al., 2005). Internal standard solution was not added to ^{13}C -labeled samples for

MFA.

A hot ethanol extraction (Gonzalez et al., 1997; Canelas et al., 2009) was used for extraction of intracellular metabolites. 5 ml hot (80 °C) 75% (v/v) ethanol solution was added to each cell pellet and samples were vortexed for 30 s, incubated in an 80 °C water bath for 3 min, and vortexed a second time for 30 s. Samples were briefly cooled in the cold (<-70 °C) ethanol bath and centrifuged to remove cell debris. 3.5 ml supernatant was set aside for LC-MS/MS analysis and the remaining supernatant (~ 1.5 ml) was used for GC/MS analysis. Metabolite extracts were stored at -80 °C and subsequently dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module prior to GC/MS and LC-MS/MS analysis.

5.2.7 Gas Chromatography/Mass Spectrometry (GC-MS)

All metabolites without phosphate groups (organic acids and amino acids) were analyzed by GC-MS. Dried metabolite extracts were resuspended in 20 µl 2% methoxyamine-hydrogen chloride in pyridine (Methoxamine (MOX) Reagent, Thermo Scientific #TS-45950) and incubated at 37 °C for 1.5 h. Following the methoximation reaction, 30 µl *N-tert*-Butyldimethylsilyl-*N*-methyltrifluoroacetamide with 1% *tert*-Butyldimethylchlorosilane (Sigma-Aldrich, #375934) was added to each sample and samples were incubated at 55 °C for 1 h. The resulting *tert*-butyldimethylsilyl derivatives were analyzed by GC-MS using an Agilent 6890N Network GC System coupled to an Agilent 5975B Inert XL MSD. 1 µl sample was injected in splitless mode with an inlet temperature of 270 °C. Metabolite separation was achieved using an Agilent J&W DB-35ms column (#122-3832, 30.0 m x 250 µm x 0.25 µm) with helium carrier gas at a flow rate of 1 ml/min. The GC column oven temperature was initially held at 100 °C for 1 min,

ramped to 105 °C at 2.5 °C/min, held at 105 °C for 2 min, ramped to 250 °C at 3.5 °C/min and then to 320 °C at 20 °C/min, for a total run time of approximately 50 min. The MS was operated in electron ionization mode with an electron energy of 69.9 eV and source and quadrupole temperatures of 230 °C and 150 °C, respectively. Mass spectra were measured by scanning the range 100-650 m/z.

5.2.8 Liquid Chromatography/Tandem Mass Spectrometry (LC-MS/MS)

All phosphorylated metabolites (sugar phosphates from glycolysis and the PPP, acetyl-CoA, and all cofactors depicted in Figure 5-4) were analyzed by LC-MS/MS. Dried metabolite extracts were resuspended in 80 µl Millipore water and analyzed using an ion pair chromatography method adapted from (Luo et al., 2007). Metabolite separation was achieved with an Agilent 1100 Series HPLC system using a Waters XBridge C18 Column (2.1 mm x 150 mm, 130 Å, 3.5 µm; #186003023). Metabolite isotopomers were quantified with an API 2000 MS/MS (AB Sciex) operating in multiple reaction monitoring (MRM) mode. The injection volume was 20 µl. The majority of samples were run at an increased flow rate of 300 µl/min with a fast gradient similar to that employed by (Young et al., 2011) in order to increase sample throughput. Details can be found in Chapter 2.

5.2.9 Quantification of Metabolite Pool Sizes

Pool size measurements were conducted in biological triplicate with technical duplicates for each of the three biological replicates. Accurate quantification was achieved using uniformly ¹³C-labeled metabolite extract as internal standard. Internal standard was added to calibration standards as well as samples so that calibration curves could be constructed in terms of the ratio

of the M+0 (from sample) and M+N (from internal standard) isotopomer peak areas (where N is the number of carbons in the metabolite backbone) (Wu et al., 2005).

5.2.10 Metabolic Flux Estimation

A model metabolic network was constructed for estimation of intracellular metabolic fluxes. The model network consisted of the reactions of Glycolysis, the PPP, the Tricarboxylic Acid (TCA) Cycle, One-Carbon Metabolism, pathways for xylose assimilation and ethanol, glycerol, and acetate production, and biosynthetic pathways for synthesis of biomass constituents. The reactions of the non-oxidative PPP were modeled using half-reactions as in (Kleijn et al., 2005). The composition of *S. cerevisiae* biomass was taken from (Förster et al., 2003). For anaerobic cultures it was assumed unsaturated fatty acids and sterols were taken up from the medium rather than synthesized from the sugar carbon source (Andreasen and Stier, 1953; Andreasen and Stier, 1954) and the biomass equation was adjusted accordingly. The modeling of compartmentalization was similar to that in (Maaheimo et al., 2001); both cytosolic and mitochondrial pools of pyruvate, acetyl-CoA, and oxaloacetate were included in the model, the anaplerotic pyruvate carboxylase reaction was assumed to occur in the cytosol, and the malic enzyme reaction was assumed to occur in the mitochondria. As in (Maaheimo et al., 2001), pyruvate transport was assumed to be unidirectional, from cytosol to mitochondria. However, oxaloacetate transport was assumed to occur in both directions in order to represent the possible impact of the malate-aspartate (or malate-oxaloacetate) shuttle on metabolite labeling patterns. Both cytosolic and mitochondrial reactions were included for alanine aminotransferase. The reactions of the metabolic model and relevant atom transitions are listed in Table A-1 (Appendix A).

For flux estimation, an in-house elementary metabolite unit (EMU)-based software was used to simulate the extracellular flux and mass spectrometry metabolite labeling data predicted to result from a given flux distribution (Antoniewicz et al., 2007a). The shift in mass isotopomer distribution (MID) mole fractions expected to result from natural abundance of heavy isotopes was accounted for by including naturally occurring heavy isotopes in simulations (Wittmann and Heinzle, 1999; van Winden et al., 2002). Reversible reactions were modeled in terms of net and exchange fluxes rather than forward and backward fluxes (Wiechert and de Graaf, 1997). For each culture condition, the best estimate for the true metabolic flux distribution was computed by minimizing the lack-of-fit between experimentally measured and simulated datasets. Only metabolite MIDs which could be measured with high accuracy were included in the MFA. In the GC-MS and LC-MS/MS analysis, control samples labeled to natural abundance were analyzed side-by-side with ^{13}C -labeled samples. The MIDs of the metabolites in the natural abundance samples can be predicted from theory, and MIDs which could not be measured to sufficient accuracy in these samples were excluded from the flux estimation models. In some cases, a metabolite pool size could be quantified by measuring the M+0 isotopomer in samples labeled to natural abundance, but the metabolite MID could not be measured to sufficient accuracy for MFA because of interference by co-eluting molecules that overlapped with higher isotopomers or because the MS signal was not sufficiently high to accurately quantify less abundant isotopomers.

Each flux estimation was performed 500 times with random initial guesses and the flux distribution with the lowest weighted sum-of-squared residuals was assumed to be the global optimum. A chi-square test for goodness-of-fit was used to evaluate whether each flux model

adequately described the data, and nonlinear 68% and 95% confidence intervals were computed for each estimated flux by computing the sensitivity of the weighted sum-of-squared residuals to variations in the flux value (Antoniewicz et al., 2006). For the chi-square test and confidence interval computation, an error of 0.325 mol % was assumed for the mole fractions in the MID data, which is reasonable given the expected measurement (Antoniewicz et al., 2007b) and modeling (Wasylenko and Stephanopoulos, 2013) errors. The assumed errors in extracellular flux data σ_{jk}^q were discussed above.

5.3 Results and Discussion

5.3.1 Determination of Optical Density-Dry Cell Weight Correlations

The correlation between OD₆₀₀ and DCW was measured to allow monitoring of cell density during pool size measurement and flux estimation experiments. Figure 5-1 shows DCW plotted against OD₆₀₀ for each of four culture conditions: Glucose Aerobic (GA), Glucose Anaerobic (GN), Xylose Aerobic (XA), and Xylose Anaerobic (XN). The calculated conversion factors for converting OD₆₀₀ to DCW, k_j^{DCW} , are listed in Table 5-1.

Table 5-1 Conversion factors for converting OD₆₀₀ to DCW, k_j^{DCW}

Culture Condition	1 OD ₆₀₀ unit =
Glucose Aerobic	0.394 ± 0.008 g/l
Xylose Aerobic	0.393 ± 0.011 g/l
Glucose Anaerobic	0.390 ± 0.017 g/l
Xylose Anaerobic	0.375 ± 0.006 g/l

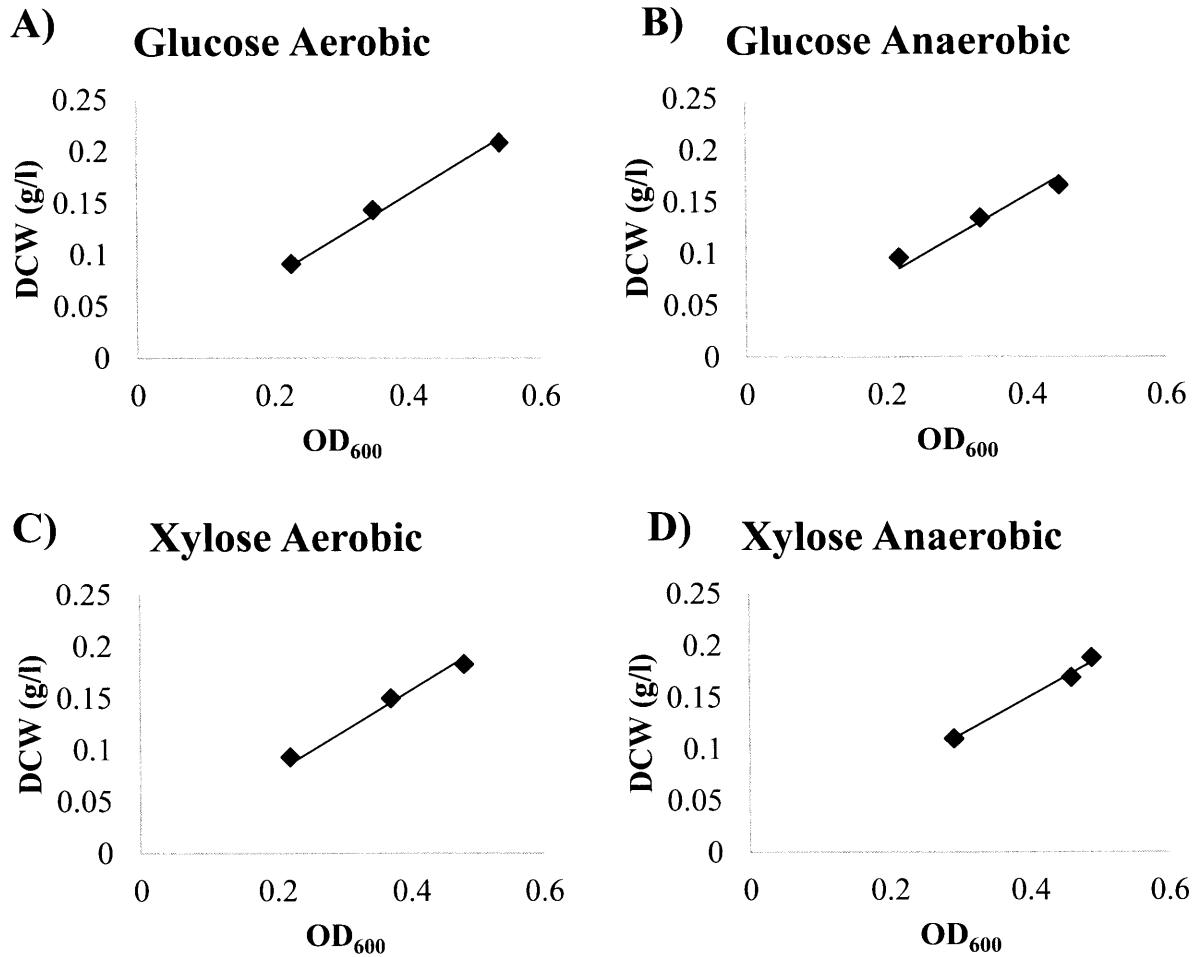


Figure 5-1 DCW vs. OD₆₀₀ for each set of culture conditions: A) Glucose Aerobic, B) Glucose Anaerobic, C) Xylose Aerobic, D) Xylose Anaerobic

5.3.2 Estimation of Extracellular Fluxes

The metabolism of the xylose-consuming strain was studied under GA, GN, XA, and XN culture conditions in order to investigate the effects of carbon source and aeration on metabolism. The specific growth rates, specific metabolite consumption and production rates, and their uncertainties were computed from OD₆₀₀ and HPLC data from five biological replicate experiments for each culture condition. The values used for the flux estimation algorithm are

listed in Table 5-2. It can be seen that glucose carbon source and aerobic conditions allowed faster growth than xylose and anaerobic conditions, respectively. However the growth rates in the GA, XA, and GN conditions were relatively similar while the growth rate in the XN condition was significantly lower than the rest. Specific sugar consumption and ethanol production rates were higher with glucose as carbon source (relative to xylose) and under anaerobic conditions (relative to aerobic conditions). Glycerol production was significantly increased under anaerobic conditions on both carbon sources. This was expected as glycerol serves as an electron sink, and its production allows for regeneration of NAD^+ when oxygen is not available as a terminal electron acceptor. Acetate production was low under all conditions tested.

Table 5-2 Specific growth rates and specific metabolite consumption and production rates under different conditions

Rate	GA	GN	XA	XN
μ (h^{-1})	0.243 ± 0.007	0.228 ± 0.003	0.231 ± 0.010	0.178 ± 0.006
$q_{\text{glucose/xylose}}$ (mmol/g/h)	-11.6 ± 1.9	-13.8 ± 1.3	-10.0 ± 1.4	-12.3 ± 3.1
q_{ethanol} (mmol/g/h)	16.0 ± 1.1	20.6 ± 2.0	12.6 ± 1.3	14.8 ± 1.1
q_{glycerol} (mmol/g/h)	1.28 ± 0.08	3.04 ± 0.32	0.75 ± 0.16	2.15 ± 0.18
q_{acetate} (mmol/g/h)	0.21 ± 0.05	0.33 ± 0.05	0.32 ± 0.07	0.23 ± 0.05

GA = Glucose Aerobic, GN = Glucose Anaerobic, XA = Xylose Aerobic, XN = Xylose Anaerobic

5.3.3 Metabolite Pool Size Measurements and Metabolic Flux Estimation

Central carbon metabolite pool sizes were measured under each of the four culture conditions. The results for metabolites of Glycolysis and the PPP, metabolites of the TCA Cycle, and select nucleotide cofactors are shown in Figures 5-2, 5-3, and 5-4, respectively. The high ratio of glucose-6-phosphate (G6P) to F6P during growth on glucose (4.3 aerobic, 6.4 anaerobic) and the high adenylate energy charge under all conditions (>0.90) suggest metabolism was satisfactorily quenched during sample preparation (Ewald et al., 2009).

The intracellular metabolic flux distribution for each culture condition was estimated by fitting OD_{600} and HPLC data and metabolite labeling data obtained from GC-MS and LC-MS/MS to a metabolic model. The best-fit metabolic flux distributions are shown in Figure 5-5. The Weighted Sums of Squared Residuals (WSSRs) from the parameter estimations are shown in Figure 5-6. It can be seen that for each of the four culture conditions the WSSR falls in the range that would be expected from the assumption that the WSSR is drawn from a chi-square distribution, suggesting the model adequately describes the data. WSSRs for model variants are also shown and are discussed below. The measured and simulated ^{13}C -labeling data for the best-fit flux models can be compared in Table A-2 (Appendix A). Non-linear 68% and 95% flux confidence intervals were computed and are shown for select fluxes in Figure 5-7. The complete set of best-fit fluxes and confidence intervals is presented in Table A-3 (Appendix A). The pool size and flux estimation results are discussed below, with the results organized by metabolic pathway.

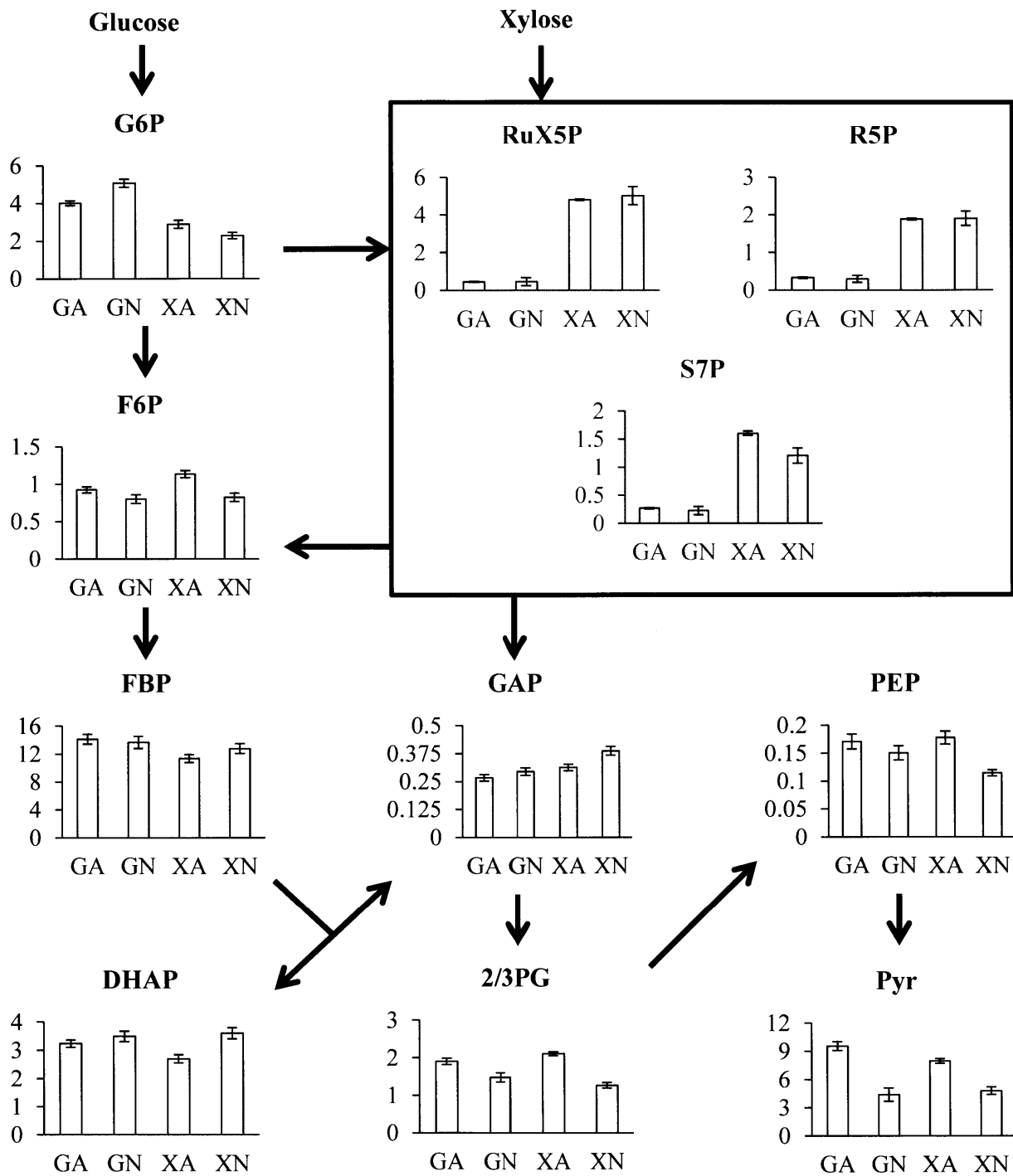


Figure 5-2 Metabolite pool sizes for metabolites in Glycolysis and the PPP. All pool sizes are in units of $\mu\text{mol/g DCW}$. Error bars represent uncertainties from standard deviations of three biological replicates as well as uncertainties in OD_{600} -DCW correlations. GA = Glucose Aerobic, GN = Glucose Anaerobic, XA = Xylose Aerobic, XN = Xylose Anaerobic

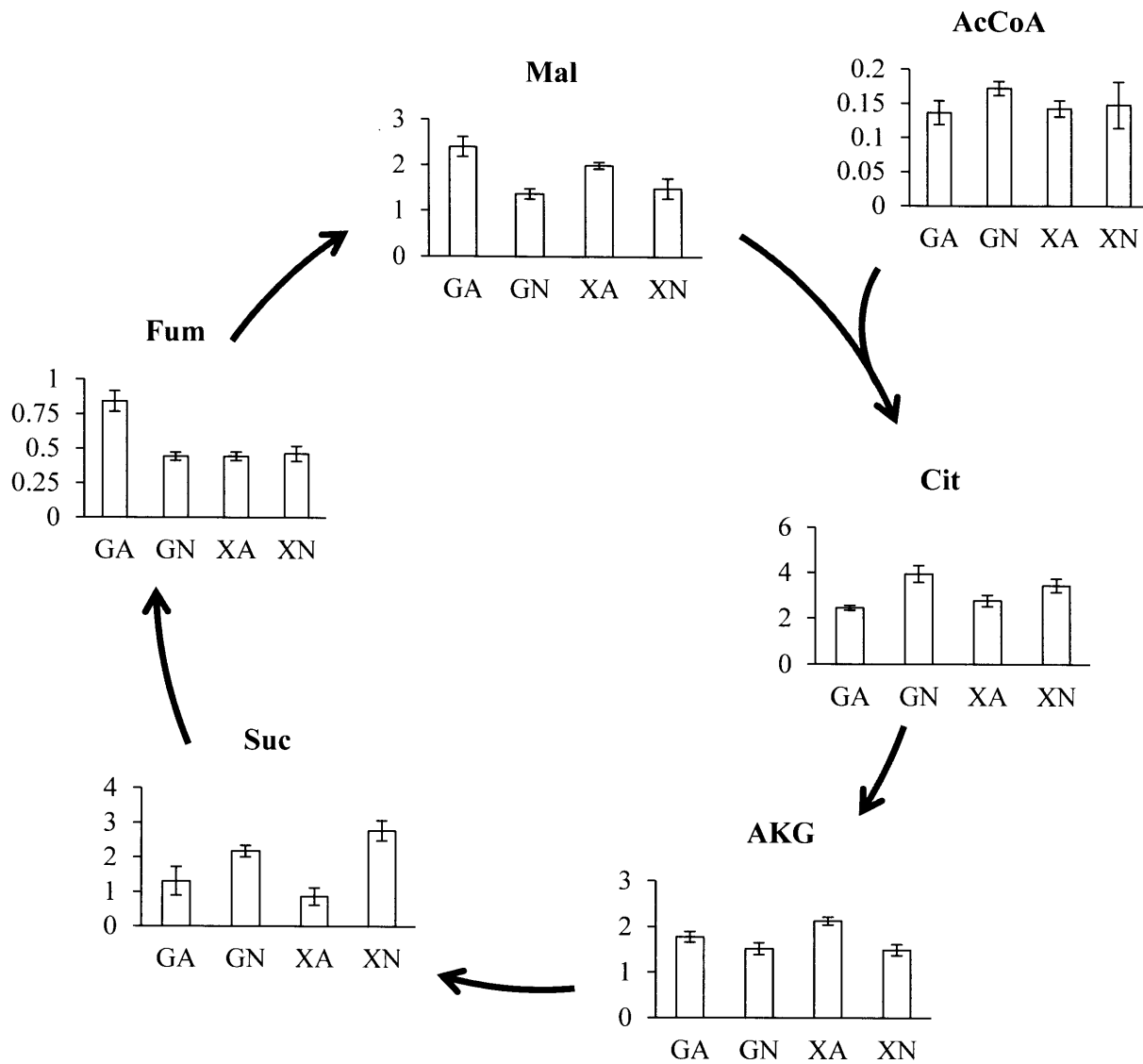


Figure 5-3 Metabolite pool sizes for metabolites in the TCA Cycle. All pool sizes are in units of $\mu\text{mol/g DCW}$. Error bars represent uncertainties from standard deviations of three biological replicates as well as uncertainties in OD_{600} -DCW correlations. GA = Glucose Aerobic, GN = Glucose Anaerobic, XA = Xylose Aerobic, XN = Xylose Anaerobic

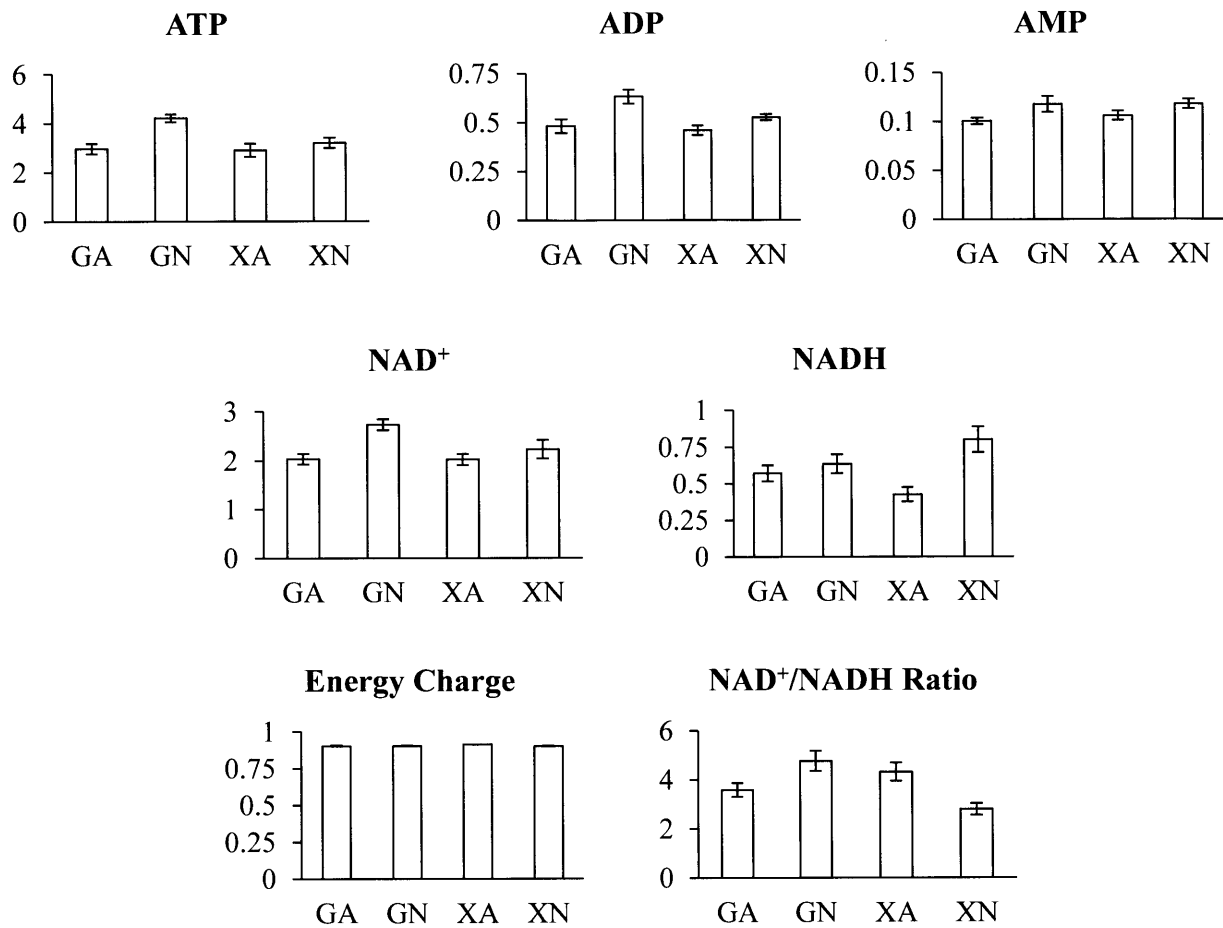


Figure 5-4 Metabolite pool sizes and ratios for select cofactors. All pool sizes are in units of $\mu\text{mol/g DCW}$. Pool size error bars represent uncertainties from standard deviations of three biological replicates as well as uncertainties in OD_{600} -DCW correlations. Metabolite ratio error bars represent uncertainties from standard deviations of three biological replicates only. GA = Glucose Aerobic, GN = Glucose Anaerobic, XA = Xylose Aerobic, XN = Xylose Anaerobic

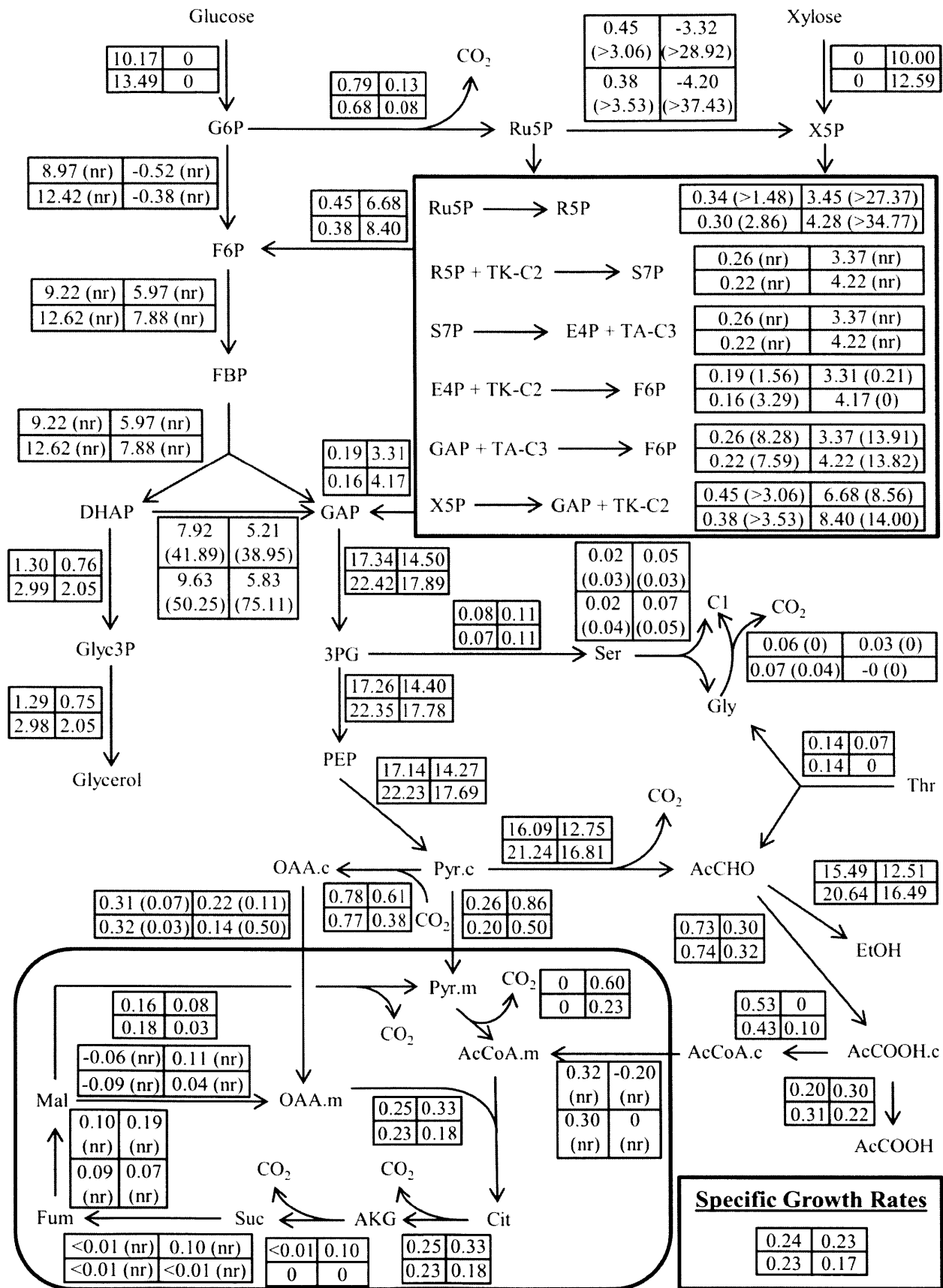


Figure 5-5 Best-fit metabolic flux distributions in xylose-consuming *S. cerevisiae* strain under four sets of conditions. All fluxes are given in units of mmol/g/h. For each reversible reaction the net flux is given with the exchange flux indicated inside parentheses; “nr” indicates the exchange flux could not be resolved to within one order of magnitude. Specific growth rates (shown in the bottom right of the Figure) are in units of h^{-1} . Four values are listed for each flux. These are the best-fit values for the Glucose Aerobic (top left), Glucose Anaerobic (bottom left), Xylose Aerobic (top right), and Xylose Anaerobic (bottom right) cultures

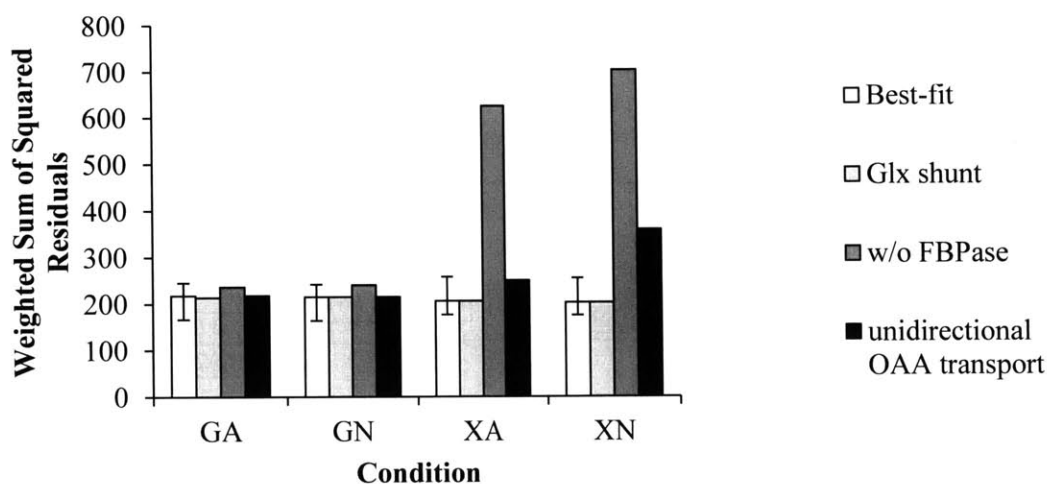


Figure 5-6 Weighted Sums of Squared Residuals from flux estimations with different models. “Best-fit” = the final metabolic model used for flux estimation, “Glx shunt” = the final model with addition of the glyoxylate shunt pathway, “w/o FBPase” = the final model with removal of the gluconeogenic enzyme Fructose-1,6-Bisphosphatase, “unidirectional OAA transport” = the final model with removal of OAA transport from mitochondria to cytosol (i.e. via Malate-Aspartate or Malate-Oxaloacetate Shuttle), so that OAA can only be transported from cytosol to mitochondria. Error bars for each “Best-fit” model show the expected range for the weighted sum of squared residuals in the chi-square goodness-of-fit test. GA = Glucose Aerobic, GN = Glucose Anaerobic, XA = Xylose Aerobic, XN = Xylose Anaerobic

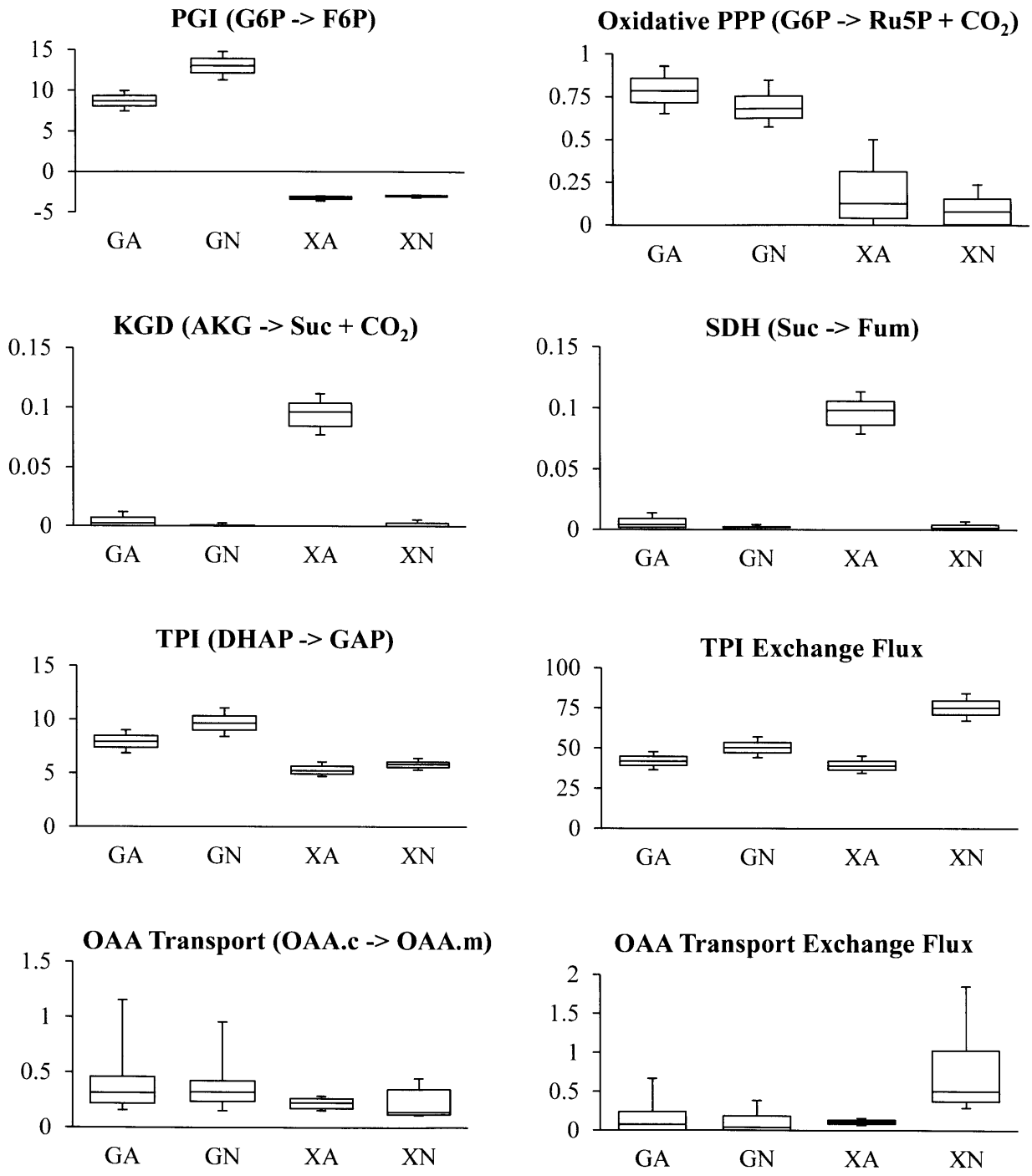


Figure 5-7 Confidence intervals for select fluxes. All fluxes are in units of mmol/g/h. The line inside each box represents the best-fit value of the flux. Each box represents the flux 68% confidence interval and the error bars (“whiskers”) represent the flux 95% confidence interval.

GA = Glucose Aerobic, GN = Glucose Anaerobic, XA = Xylose Aerobic, XN = Xylose Anaerobic. Enzyme abbreviations: PGI, Phosphoglucose isomerase; KGD, α -Ketoglutarate dehydrogenase; SDH, Succinate dehydrogenase; TPI, Triosephosphate isomerase

5.3.4 Pentose Phosphate Pathway

It can be seen that when xylose is the sole carbon source the flux through the non-oxidative PPP is more than one order of magnitude higher than when glucose is the sole carbon source (Figure 5-5). This is a direct result of the metabolic network topology. All carbon assimilated from xylose enters central metabolism at X5P and consequently must pass through the reactions of the non-oxidative PPP. In contrast, during growth on glucose the majority of carbon flux is channeled through glycolysis with only a fraction being diverted through the PPP to allow production of NADPH for biosynthesis and formation of ribose-5-phosphate (R5P) and erythrose-4-phosphate (E4P) carbon backbones which serve as precursors for amino acids and nucleotides. It has been shown that PPP flux in yeasts metabolizing glucose correlates with biomass yield, likely because high biomass yields increase demand for NADPH produced in the oxidative PPP (Blank et al., 2005). Under fermentative conditions in which the majority of carbon is converted to ethanol the PPP flux on glucose is expected to be low (Gombert et al., 2001; Fiaux et al., 2003). The increased flux through the non-oxidative PPP on xylose results in a significant increase in the pool sizes of the metabolites of this pathway (Figure 5-2), consistent with previous observations in *S. cerevisiae* during pentose utilization (Kötter and Ciriacy, 1993; Wisselink et al., 2010).

In contrast, the flux through the oxidative PPP is low under all conditions (Figures 5-5 and 5-7).

Carbon from xylose need not pass through this phase of the PPP to be assimilated into central metabolism. Previous work with strains employing the XR-XDH pathway for xylose assimilation has suggested that the flux through the oxidative PPP is high during xylose utilization, likely due to the redox cofactor imbalance which increases demand for NADPH (Pitkänen et al., 2003; Sonderegger et al., 2004; Grotkjaer et al., 2005; Runquist et al., 2009b; Feng and Zhao, 2013a). The high flux through this pathway decreases the ethanol yield that can be achieved with these strains because one mole of carbon is lost as CO₂ for every two moles of NADPH produced (Pitkänen et al., 2003; Grotkjaer et al., 2005). Because XI is used to effect the conversion of xylose to xylulose in the strain investigated here, the redox balance in the cell is not perturbed by xylose assimilation and only minimal flux through the oxidative PPP is required to satisfy NADPH demand; consequently higher ethanol yields can be obtained. Low oxidative PPP flux has also been found in an arabinose-consuming strain of *S. cerevisiae* utilizing an isomerase pathway for arabinose assimilation (Wisselink et al., 2010).

5.3.5 Glycolysis

Carbon derived from xylose via the non-oxidative PPP enters the glycolytic pathway at F6P and GAP, with the PPP stoichiometry dictating that roughly two moles of F6P are produced for every mole of GAP. (The ratio deviates slightly from 2:1 due to withdrawal of R5P and E4P for biosynthesis). Because carbon enters the pathway below the G6P node, the flux through phosphoglucose isomerase (PGI) is reversed during xylose utilization, with a net flux from F6P to G6P to allow synthesis of carbohydrates and small flux through the oxidative PPP (Figures 5-5 and 5-7). This reversal of flux results in a depletion of the G6P pool during growth on xylose, which is especially remarkable under anaerobic conditions (Figure 5-2). The majority of F6P

derived from xylose is directed to GAP through the upper glycolytic pathway, although this flux is somewhat attenuated on xylose because some carbon from the non-oxidative PPP is converted to GAP directly. Once GAP is formed, the pathways for glucose and xylose utilization are identical.

5.3.6 Tricarboxylic Acid Cycle

The fluxes through the TCA Cycle are observed to be low in all conditions, as the majority of carbon from pyruvate is converted to ethanol rather than being transported to the mitochondria and oxidized to form CO₂. Under anaerobic conditions the TCA Cycle is essentially inactive, and the reactions of the pathway serve only to provide carbon building blocks for biosynthesis. The flux through α -ketoglutarate dehydrogenase (KGD) is estimated to be zero and the flux through succinate dehydrogenase (SDH) is minimal (Figures 5-5, 5-7). (The small flux through SDH is a result of formation of succinate as a by-product of cysteine biosynthesis.) The lack of flux through the TCA Cycle is expected under anaerobic conditions since oxygen is not available as a terminal electron acceptor for re-oxidation of mitochondrial NADH, and it can be seen that the pool sizes of metabolites in this pathway depend more on aeration than carbon source (Figure 5-3). In the GA condition there is a small flux through KGD, but inspection of the flux confidence intervals reveals that this flux is not significantly different from zero (Figure 5-7). This minimal flux through the TCA Cycle under the GA condition is likely a result of glucose-repression of this pathway (Polakis and Bartley, 1965), and is consistent with previous work in *S. cerevisiae* under fermentative conditions (Gombert et al., 2001). Only in the XA case do the labeling data suggest a significant (albeit still small) flux through KGD (Figure 5-7).

5.3.7 Incomplete Activation of Fermentation Program

The small but significant flux through the TCA Cycle in the XA condition suggests an incomplete induction of the CCR response, which has previously been observed during xylose assimilation (Jin et al., 2004; Salusjärvi et al., 2008). Other glucose-repressed metabolic pathways include gluconeogenesis, the glyoxylate shunt, and oxidative phosphorylation. Previous transcriptomic studies of H131-A3-AL^{CS} (Zhou et al., 2012) and other xylose-consuming strains (Wahlbom et al., 2003; Jin et al., 2004; Salusjärvi et al., 2008; Runquist et al., 2009b; Scalcinati et al., 2012) have revealed upregulation of genes associated with these pathways during growth on xylose. Inclusion of the glyoxylate shunt in the metabolic network did not significantly affect the results of flux estimation. The WSSRs for the models with and without this pathway were virtually identical in all cases (Figure 5-6), and when the glyoxylate shunt was included the estimated flux through the pathway was negligible (data not shown). Thus, the labeling data do not provide any evidence for the activity of this pathway.

The labeling data do suggest that the gluconeogenic enzyme fructose-1,6-bisphosphatase (FBPase1) is active during growth on xylose. FBPase1 essentially catalyzes the reverse of the reaction catalyzed by phosphofructokinase-1 (Pfk1), which is coupled to ATP hydrolysis and consequently practically irreversible. The activation of FBPase1 is modeled by making the Pfk1 reaction for conversion of F6P to fructose-1,6-bisphosphate (FBP) reversible. Although the exchange flux (equivalent to the backward flux for a reversible reaction) for the Pfk1 reaction could not be resolved, Figure 5-6 shows that when FBPase1 was excluded from the model (i.e. when the Pfk1 reaction was specified as irreversible) a dramatic increase in the WSSRs for the xylose flux estimation models was observed. This implies that the model without FBPase1 does

not describe the xylose labeling datasets as well as the model that includes FBPase1. The large increases in the WSSRs resulted from an inability to fit the F6P labeling data without the FBPase1 reaction in the model (data not shown). However, it can be seen that when FBPase1 was omitted from the glucose flux estimation models the effect was relatively small, and the WSSRs still fell in the range expected from the chi-square distribution. Thus the data provide evidence for activity of FBPase1 on xylose but not on glucose.

In *S. cerevisiae*, FBPase1 flux is repressed in the presence of glucose through several redundant mechanisms, as futile cycling between Pfk1 and FBPase1 is detrimental to the cell (Navas and Gancedo, 1996). The transcription of the associated gene FBP1 is repressed and the degradation of FBP1 mRNA accelerated in the presence of even low concentrations of glucose. This effect can be triggered by the extracellular glucose sensor Snf3 or the Ras-cAMP pathway (Yin et al., 2000). It has also been proposed that accumulation of G6P (which was depleted on xylose; see above) prevents localization of Snf1 to the nucleus, contributing to the repression of gluconeogenic genes (Vincent et al., 2001), and previous work suggests that accumulation of G6P and F6P is both necessary and sufficient for complete repression of FBPase1 activity (Gancedo and Gancedo, 1979). The FBPase1 protein is also degraded in the presence of glucose (Belinchón and Gancedo, 2007), and FBPase1 activity is inhibited by accumulation of fructose-2,6-bisphosphate (F-2,6-BP), which is an allosteric inhibitor and also stimulates the phosphorylation and inactivation of the enzyme by protein kinase A (PKA) (Gancedo et al., 1983). The accumulation of F-2,6-BP in the presence of glucose is itself a result of PKA-dependent phosphorylation and activation of phosphofructokinase-2 (Pfk2) (Vaseghi et al., 2001; Dihazi et al., 2003). PKA can be activated by the extracellular glucose sensor Gpr1 or the Ras-

cAMP pathway (Rolland et al., 2000; Gancedo, 2008). Activation of the Ras-cAMP pathway has been shown to be dependent on glucose phosphorylation (Rolland et al., 2000) and may be incomplete on xylose due to the depletion of G6P. Thus the increased FBPase1 flux on xylose likely results from some combination of absence of stimulation of the extracellular glucose sensors Gpr1, Rgt2, and Snf3, incomplete activation of the Ras-cAMP pathway, and depletion of G6P which possibly results in nuclear localization of Snf1.

Although expression of FBPase1 on xylose would be expected to result in futile cycling between Pfk1 and FBPase1 and unnecessary ATP expenditure, the high energy charges in XA and XN conditions (Figure 5-4) suggest that this has little impact on xylose metabolism. The gluconeogenic enzyme PEP carboxykinase (PEPCK) was also shown to be upregulated in H131-A3-AL^{CS} (Zhou et al., 2012), but was not included in the metabolic model because in the glucose flux models the flux through this enzyme could not be resolved; the flux through the cycle formed by pyruvate kinase, pyruvate carboxylase, and PEPCK could be arbitrarily large without exerting a significant influence on metabolite labeling patterns (data not shown). When PEPCK was included in either xylose flux model the estimated flux was zero, indicating that if the enzyme is active it may not carry significant flux.

The data also indicate the exchange flux for transport of oxaloacetate (OAA) across the mitochondrial membrane is increased on xylose. While the transport of OAA from mitochondria to cytosol is not significant during growth on glucose, it is on xylose (Figures 5-5 and 5-7), particularly under anaerobic conditions. Figure 5-6 shows that the WSSRs were increased in the xylose flux estimation models when OAA transport from cytosol to mitochondria was modeled

as an irreversible reaction. The effect was again more pronounced under anaerobic conditions, where the WSSR of the irreversible OAA transport model fell outside the range expected from the chi-square distribution, suggesting reversible OAA transport is necessary to satisfactorily describe the XN labeling data.

The biological significance of the increased OAA transport exchange flux on xylose is not obvious. OAA transport into the mitochondria is driven by the proton gradient and is generally unidirectional (Palmieri et al., 1999; Maaheimo et al., 2001). The increased exchange flux on xylose could represent export of four carbon units from the TCA Cycle to the cytosol due to incomplete repression of gluconeogenic pathways. The labeling data also could be explained by synthesis of aspartate by mitochondrial aspartate aminotransferase (in the metabolic model aspartate was assumed to be synthesized exclusively in the cytosol, as in (Maaheimo et al., 2001)). However, an attractive explanation is that the high OAA transport exchange flux is a result of shuttling of NADH equivalents across the mitochondrial membrane. Several redox shuttles that effectively transport NADH equivalents across the mitochondrial membrane exist in *S. cerevisiae*. These include the malate-aspartate shuttle, the malate-oxaloacetate shuttle, and the ethanol-acetaldehyde shuttle. In eukaryotes, the malate-aspartate shuttle generally transports NADH equivalents generated in the cytosol (e.g. by glycolysis) into the mitochondria for oxidative phosphorylation, and may be able to operate against a concentration gradient. The latter two shuttles can in principle transport reducing equivalents in either direction but likely not against a concentration gradient (Bakker et al., 2001).

Operation of the malate-aspartate (or malate-oxaloacetate) shuttle would have an effect on

metabolite labeling distributions similar to that of reversible OAA transport. Although the malate-aspartate shuttle exists in *S. cerevisiae* (Cavero et al., 2003), it is associated with respiratory metabolism and is not expected to carry significant flux during glucose fermentation, when NADH is re-oxidized primarily by ethanol formation rather than oxidative phosphorylation. Previous work has shown the cytosolic malate dehydrogenase Mdh2, a key component of the pathway and also of the malate-oxaloacetate shuttle, is inactive during growth on glucose (Minard and McAlister-Henn, 1992; Minard and McAlister-Henn, 1994). On xylose, activation of Mdh2 due to incomplete induction of CCR could result in increased redox shuttling activity. This possibility is discussed below.

5.3.8 Lower Glycolysis as a Potential Bottleneck in Xylose Metabolism

The *S. cerevisiae* strain employed in this study exhibited a specific growth rate in the XN condition that was significantly lower than those observed in the other three conditions, indicating the presence of bottlenecks in anaerobic xylose metabolism. It can be seen that there were significant increases in both the GAP pool size (Figure 5-2) and the exchange flux for the triosephosphate isomerase (TPI) reaction (Figures 5-5 and 5-7) in the XN condition. Taken together, these data suggest the presence of a bottleneck downstream of GAP, which causes GAP to accumulate and increases the rate of the back reaction to DHAP. The 2-phosphoglycerate/3-phosphateglycerate (2/3PG) and phosphoenolpyruvate (PEP) pools were depleted in the XN condition, suggesting the bottleneck occurs in lower glycolysis.

A bottleneck in lower glycolysis could result from incomplete activation of the fermentation program on xylose. Although the signaling mechanisms for activation of CCR are complex and

incompletely understood, intracellular metabolite concentrations likely play a prominent role (Belinchón and Gancedo, 2003; Gancedo, 2008). In particular, accumulation of G6P (as well as other glycolytic metabolites) has been shown to be required for complete activation of enzymes associated with lower glycolysis and ethanol formation (Boles et al., 1993; Boles and Zimmermann, 1993; Müller et al., 1995; Boles et al., 1996). It was noted above that the G6P pool was depleted in the XN condition, so it is plausible that the activities of lower glycolytic enzymes are reduced during xylose utilization. Previous transcriptomic studies have revealed decreased expression of several lower glycolytic genes on xylose relative to on glucose (Wahlbom et al., 2003; Jin et al., 2004), and a transcriptomic analysis of H131-A3-AL^{CS} resulted in similar findings, although the changes in expression were not significant (i.e. less than two-fold) (Zhou et al., 2012).

Under fermentative conditions, cytosolic NADH produced in the GAP dehydrogenase (GAPDH) reaction is re-oxidized largely through ethanol formation. Consequently, a bottleneck in lower glycolysis or ethanogenesis might lead to inefficient regeneration of cytosolic NAD⁺. Under aerobic conditions a bottleneck in cytosolic NADH re-oxidation could be alleviated by shuttling NADH into the mitochondria, where it could be re-oxidized in the electron transport chain with oxygen as the terminal electron acceptor. However under anaerobic conditions inefficient ethanol formation would likely result in accumulation of NADH and low NAD⁺ availability. Figure 5-4 shows that the NADH pool size was indeed increased (and the NAD⁺/NADH ratio decreased) in the XN condition relative to all other conditions. The low availability of NAD⁺ would then limit the rate of the GAPDH reaction, consistent with the observed accumulation of GAP, increased TPI exchange flux, and depletion of 2/3PG and PEP pools.

Interestingly, on xylose Matsushika, *et al.* observed even more dramatic accumulation of metabolites upstream of GAP and depletion of metabolites downstream of GAP than those reported here (Matsushika et al., 2013). The study employed a strain that utilized the XR-XDH pathway to effect xylose isomerization. Consequently, in this strain the NAD^+ availability would be expected to be very low due to the redox cofactor imbalance associated with that pathway. The authors concluded that low glycolytic flux limited the rate of xylose utilization. Although they did not measure the pool sizes of NAD^+ or NADH, our data suggest that in their study also the glycolytic flux may have been limited by the availability of NAD^+ for the GAPDH reaction.

Inefficient re-oxidation and accumulation of NADH could also explain the apparent increased flux of NADH equivalents across the mitochondrial membrane discussed above. Under aerobic conditions the best-fit OAA transport exchange flux was slightly increased on xylose relative to glucose (Figure 5-5), and the flux confidence intervals reveal that this flux is significantly different from zero on xylose but not on glucose (Figure 5-7). Under XA conditions, NADH equivalents could be shuttled into the mitochondria via the malate-aspartate (or malate-oxaloacetate) shuttle and oxidized in the electron transport chain. The increased NADH shuttling activity would likely result from activation of Mdh2 due to incomplete CCR, inefficient oxidation of NADH in the cytosol due to low activities of enzymes in lower glycolysis and ethanologensis, or a combination of the two.

Surprisingly the increase in OAA transport exchange flux is more pronounced in the XN condition, where oxygen is not available as a terminal electron acceptor and NADH equivalents shuttled into the mitochondria cannot be efficiently re-oxidized in the electron transport chain.

However, we note that the flux of NADH equivalents into the mitochondria is relatively low. In the XN condition the best-fit OAA transport exchange flux is only 0.50 mmol/g/h while the flux through GAPDH is 17.89 mmol/g/h. Thus, only a few percent of the NADH equivalents generated in glycolysis would be shuttled into the mitochondria. The NADH shuttling flux may be higher under anaerobic conditions because the glycolytic flux (and the rate of NADH production) is higher, increasing the burden on the lower glycolytic and ethanologenic pathways for NAD⁺ regeneration. It is unclear how the NADH shuttled into the mitochondria would be re-oxidized, but several redox shuttling systems exist in *S. cerevisiae*, as noted above. One interesting possibility is that the mitochondrial alcohol dehydrogenase Adh3 may catalyze the conversion of acetaldehyde to ethanol in the mitochondria. It has previously been proposed that under anaerobic conditions Adh3 oxidizes NADH generated in the mitochondria by amino acid biosynthesis (Nissen et al., 1997).

Taken together then, the data suggest that in the strain employed here slow metabolism of GAP to pyruvate and ultimately to ethanol and inefficient re-oxidation of NADH produced in glycolysis may limit xylose consumption under anaerobic conditions, where oxygen is unavailable to serve as an electron acceptor for regeneration of NAD⁺. The presence of the bottleneck in lower glycolysis hypothesized here would most likely arise due to the extensive engineering of strain H131-A3-AL^{CS}. Upon introduction of the xylose assimilation pathway, bottlenecks are far more likely to occur in xylose transport or conversion to the glycolytic metabolites F6P and GAP. However, overexpression of *xy1A*, *XYL3*, and the genes of the non-oxidative PPP along with multi-stage evolutionary engineering which resulted in a multi-copy integration of *xy1A* into the chromosome (Zhou et al., 2012) appear to have removed all

bottlenecks in the assimilation of xylose to central carbon metabolism. Consequently, the changes in gene expression that result from incomplete CCR on xylose may have become the dominant factor in limiting the rate of xylose metabolism.

Because CCR is a complex phenomenon, it is not obvious how one would engineer a yeast strain with a fully active fermentation program during metabolism of xylose as sole carbon source. If glycolytic metabolite pool sizes do play a role in regulation, it would be difficult to recapitulate the metabolite profile from glucose fermentation on xylose due to differences in reaction network topology. It may be possible to realize faster xylose consumption by engineering the promoters of genes associated with lower glycolysis and ethanol formation to achieve higher activities during xylose utilization. For biofuels applications, another strategy may be to engineer strains that consume the glucose and xylose in biomass hydrolysates simultaneously. Such an approach would allow metabolism of xylose with glucose present to support flux through upper glycolysis, allowing for larger pools of metabolites in that pathway, and activate the fermentation program. Although glucose represses xylose uptake, an interesting strategy for co-fermentation of cellobiose and xylose has recently been presented (Ha et al., 2011). Such approaches hold much promise for production of lignocellulosic biofuels.

5.4 References

Almeida, J.R., Modig, T., Petersson, A., Hähn-Hägerdal, B., Lidén, G., Gorwa-Grauslund, M.F., 2007. Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biotechnol.* 82, 340–349. doi:10.1002/jctb.1676

Amore, R., Wilhelm, M., Hollenberg, C.P., 1989. The fermentation of xylose--an analysis of the expression of *Bacillus* and *Actinoplanes* xylose isomerase genes in yeast. *Appl. Microbiol. Biotechnol.* 30, 351–357. doi:10.1007/BF00296623

Andreasen, A.A., Stier, T.J., 1954. Anaerobic nutrition of *Saccharomyces cerevisiae*. II. Unsaturated fatty acid requirement for growth in a defined medium. *J. Cell. Physiol.* 43, 271–281.

Andreasen, A.A., Stier, T.J.B., 1953. Anaerobic nutrition of *Saccharomyces cerevisiae*. I. Ergosterol requirement for growth in a defined medium. *J. Cell. Physiol.* 41, 23–36.

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2006. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* 8, 324–337. doi:10.1016/j.ymben.2006.01.004

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007a. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng.* 9, 68–86. doi:10.1016/j.ymben.2006.09.001

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007b. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Anal. Chem.* 79, 7554–7559. doi:10.1021/ac0708893

Bakker, B.M., Overkamp, K.M., van Maris AJ, null, Kötter, P., Luttik, M.A., van Dijken JP, null, Pronk, J.T., 2001. Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 25, 15–37.

Belinchón, M.M., Gancedo, J.M., 2003. Xylose and some non-sugar carbon sources cause catabolite repression in *Saccharomyces cerevisiae*. *Arch. Microbiol.* 180, 293–297.

doi:10.1007/s00203-003-0593-9

Belinchón, M.M., Gancedo, J.M., 2007. Glucose controls multiple processes in *Saccharomyces cerevisiae* through diverse combinations of signaling pathways. *FEMS Yeast Res.* 7, 808–818.

doi:10.1111/j.1567-1364.2007.00236.x

Bergdahl, B., Heer, D., Sauer, U., Hahn-Hägerdal, B., van Niel, E.W., 2012. Dynamic metabolomics differentiates between carbon and energy starvation in recombinant *Saccharomyces cerevisiae* fermenting xylose. *Biotechnol. Biofuels* 5, 34. doi:10.1186/1754-

6834-5-34

Blank, L.M., Lehmbeck, F., Sauer, U., 2005. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Res.* 5, 545–558. doi:10.1016/j.femsyr.2004.09.008

Boles, E., Heinisch, J., Zimmermann, F.K., 1993. Different signals control the activation of glycolysis in the yeast *Saccharomyces cerevisiae*. *Yeast Chichester Engl.* 9, 761–770.

doi:10.1002/yea.320090710

Boles, E., Müller, S., Zimmermann, F.K., 1996. A multi-layered sensory system controls yeast glycolytic gene expression. *Mol. Microbiol.* 19, 641–642.

Boles, E., Zimmermann, F.K., 1993. Induction of pyruvate decarboxylase in glycolysis mutants of *Saccharomyces cerevisiae* correlates with the concentrations of three-carbon glycolytic metabolites. *Arch. Microbiol.* 160, 324–328.

Bruinenberg, P., Bot, P.M., Dijken, J., Scheffers, W.A., 1984. NADH-linked aldose reductase: the key to anaerobic alcoholic fermentation of xylose by yeasts. *Appl. Microbiol. Biotechnol.* 19. doi:10.1007/BF00251847

Bruinenberg, P.M., Bot, P.H.M., Dijken, J.P., Scheffers, W.A., 1983. The role of redox balances in the anaerobic fermentation of xylose by yeasts. *Eur. J. Appl. Microbiol. Biotechnol.* 18, 287–292. doi:10.1007/BF00500493

Canelas, A.B., Ras, C., Pierick, A., Dam, J.C., Heijnen, J.J., Gulik, W.M., 2008. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics* 4, 226–239. doi:10.1007/s11306-008-0116-4

Canelas, A.B., ten Pierick, A., Ras, C., Seifar, R.M., van Dam, J.C., van Gulik, W.M., Heijnen, J.J., 2009. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. *Anal. Chem.* 81, 7379–7389. doi:10.1021/ac900999t

Carroll, A., Somerville, C., 2009. Cellulosic biofuels. *Annu. Rev. Plant Biol.* 60, 165–182. doi:10.1146/annurev.arplant.043008.092125

Cavero, S., Vozza, A., del Arco, A., Palmieri, L., Villa, A., Blanco, E., Runswick, M.J., Walker, J.E., Cerdán, S., Palmieri, F., Satrústegui, J., 2003. Identification and metabolic role of the mitochondrial aspartate-glutamate transporter in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 50, 1257–1269.

Dihazi, H., Kessler, R., Eschrich, K., 2003. Glucose-induced stimulation of the Ras-cAMP pathway in yeast leads to multiple phosphorylations and activation of 6-phosphofructo-2-kinase. *Biochemistry (Mosc.)* 42, 6275–6282. doi:10.1021/bi034167r

- Ewald, J.C., Heux, S., Zamboni, N., 2009. High-throughput quantitative metabolomics: workflow for cultivation, quenching, and analysis of yeast in a multiwell format. *Anal. Chem.* 81, 3623–3629. doi:10.1021/ac900002u
- Feng, X., Zhao, H., 2013a. Investigating xylose metabolism in recombinant *Saccharomyces cerevisiae* via ^{13}C metabolic flux analysis. *Microb. Cell Factories* 12, 114. doi:10.1186/1475-2859-12-114
- Feng, X., Zhao, H., 2013b. Investigating glucose and xylose metabolism in *Saccharomyces cerevisiae* and *Scheffersomyces stipitis* via ^{13}C metabolic flux analysis. *AIChE J.* 59, 3195–3202. doi:10.1002/aic.14182
- Fiaux, J., Cakar, Z.P., Sonderegger, M., Wüthrich, K., Szyperski, T., Sauer, U., 2003. Metabolic-flux profiling of the yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*. *Eukaryot. Cell* 2, 170–180.
- Förster, J., Famili, I., Fu, P., Palsson, B.Ø., Nielsen, J., 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* 13, 244–253. doi:10.1101/gr.234503
- Gancedo, J.M., 2008. The early steps of glucose signalling in yeast. *FEMS Microbiol. Rev.* 32, 673–704. doi:10.1111/j.1574-6976.2008.00117.x
- Gancedo, J.M., Gancedo, C., 1979. Inactivation of gluconeogenic enzymes in glycolytic mutants of *Saccharomyces cerevisiae*. *Eur. J. Biochem. FEBS* 101, 455–460.
- Gancedo, J.M., Mazón, M.J., Gancedo, C., 1983. Fructose 2,6-bisphosphate activates the cAMP-dependent phosphorylation of yeast fructose-1,6-bisphosphatase in vitro. *J. Biol. Chem.* 258, 5998–5999.

Gárdonyi, M., Hahn-Hägerdal, B., 2003. The *Streptomyces rubiginosus* xylose isomerase is misfolded when expressed in *Saccharomyces cerevisiae*. *Enzyme Microb. Technol.* 32, 252–259. doi:10.1016/S0141-0229(02)00285-5

Gárdonyi, M., Jeppsson, M., Lidén, G., Gorwa-Grauslund, M.F., Hahn-Hägerdal, B., 2003. Control of xylose consumption by xylose transport in recombinant *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 82, 818–824. doi:10.1002/bit.10631

Gombert, A.K., Moreira dos Santos, M., Christensen, B., Nielsen, J., 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* 183, 1441–1451. doi:10.1128/JB.183.4.1441-1451.2001

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

Grotkjaer, T., Christakopoulos, P., Nielsen, J., Olsson, L., 2005. Comparative metabolic network analysis of two xylose fermenting recombinant *Saccharomyces cerevisiae* strains. *Metab. Eng.* 7, 437–444. doi:10.1016/j.ymben.2005.07.003

Harhangi, H.R., Akhmanova, A.S., Emmens, R., van der Drift, C., de Laat, W.T.A.M., van Dijken, J.P., Jetten, M.S.M., Pronk, J.T., Op den Camp, H.J.M., 2003. Xylose metabolism in the anaerobic fungus *Piromyces* sp. strain E2 follows the bacterial pathway. *Arch. Microbiol.* 180, 134–141. doi:10.1007/s00203-003-0565-0

Ha, S.-J., Galazka, J.M., Kim, S.R., Choi, J.-H., Yang, X., Seo, J.-H., Glass, N.L., Cate, J.H.D., Jin, Y.-S., 2011. Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 504–509.

doi:10.1073/pnas.1010456108

Jeppsson, M., Träff, K., Johansson, B., Hahn-Hägerdal, B., Gorwa-Grauslund, M.F., 2003. Effect of enhanced xylose reductase activity on xylose consumption and product distribution in xylose-fermenting recombinant *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 3, 167–175.

Jin, Y.-S., Laplaza, J.M., Jeffries, T.W., 2004. *Saccharomyces cerevisiae* engineered for xylose metabolism exhibits a respiratory response. *Appl. Environ. Microbiol.* 70, 6816–6825.

doi:10.1128/AEM.70.11.6816-6825.2004

Jin, Y.-S., Ni, H., Laplaza, J.M., Jeffries, T.W., 2003. Optimal growth and ethanol production from xylose by recombinant *Saccharomyces cerevisiae* require moderate D-xylose kinase activity. *Appl. Environ. Microbiol.* 69, 495–503.

Karhumaa, K., Hahn-Hägerdal, B., Gorwa-Grauslund, M.-F., 2005. Investigation of limiting metabolic steps in the utilization of xylose by recombinant *Saccharomyces cerevisiae* using metabolic engineering. *Yeast Chichester Engl.* 22, 359–368. doi:10.1002/yea.1216

Kim, S.R., Ha, S.-J., Kong, I.I., Jin, Y.-S., 2012. High expression of XYL2 coding for xylitol dehydrogenase is necessary for efficient xylose fermentation by engineered *Saccharomyces cerevisiae*. *Metab. Eng.* 14, 336–343. doi:10.1016/j.ymben.2012.04.001

Kleijn, R.J., van Winden, W.A., van Gulik, W.M., Heijnen, J.J., 2005. Revisiting the ¹³C-label distribution of the non-oxidative branch of the pentose phosphate pathway based upon kinetic and genetic evidence. *FEBS J.* 272, 4970–4982. doi:10.1111/j.1742-4658.2005.04907.x

Klimacek, M., Krahulec, S., Sauer, U., Nidetzky, B., 2010. Limitations in xylose-fermenting *Saccharomyces cerevisiae*, made evident through comprehensive metabolite profiling and thermodynamic analysis. *Appl. Environ. Microbiol.* 76, 7566–7574. doi:10.1128/AEM.01787-10

Kötter, P., Ciriacy, M., 1993. Xylose fermentation by *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 38, 776–783. doi:10.1007/BF00167144

Kuyper, M., Harhangi, H.R., Stave, A.K., Winkler, A.A., Jetten, M.S.M., de Laat, W.T.A.M., den Ridder, J.J.J., Op den Camp, H.J.M., van Dijken, J.P., Pronk, J.T., 2003. High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res.* 4, 69–78.

Kuyper, M., Hartog, M.M.P., Toirkens, M.J., Almering, M.J.H., Winkler, A.A., van Dijken, J.P., Pronk, J.T., 2005a. Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res.* 5, 399–409. doi:10.1016/j.femsyr.2004.09.010

Kuyper, M., Toirkens, M.J., Diderich, J.A., Winkler, A.A., van Dijken, J.P., Pronk, J.T., 2005b. Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. *FEMS Yeast Res.* 5, 925–934. doi:10.1016/j.femsyr.2005.04.004

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

FEMS Yeast Res. 4, 655–664. doi:10.1016/j.femsyr.2004.01.003

Lau, M.W., Gunawan, C., Balan, V., Dale, B.E., 2010. Comparing the fermentation performance of *Escherichia coli* KO11, *Saccharomyces cerevisiae* 424A(LNH-ST) and *Zymomonas mobilis*

AX101 for cellulosic ethanol production. Biotechnol. Biofuels 3, 11. doi:10.1186/1754-6834-3-11

Lönn, A., Träff-Bjerre, K., Cordero Otero, R., van Zyl, W., Hahn-Hägerdal, B., 2003. Xylose isomerase activity influences xylose fermentation with recombinant *Saccharomyces cerevisiae*

strains expressing mutated *xylA* from *Thermus thermophilus*. Enzyme Microb. Technol. 32, 567–573. doi:10.1016/S0141-0229(03)00024-3

Luo, B., Groenke, K., Takors, R., Wandrey, C., Oldiges, M., 2007. Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic

acid cycle by liquid chromatography-mass spectrometry. J. Chromatogr. A 1147, 153–164. doi:10.1016/j.chroma.2007.02.034

Maaheimo, H., Fiaux, J., Cakar, Z.P., Bailey, J.E., Sauer, U., Szyperski, T., 2001. Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional (¹³C) labeling of

common amino acids. Eur. J. Biochem. FEBS 268, 2464–2479.

Matsushika, A., Nagashima, A., Goshima, T., Hoshino, T., 2013. Fermentation of xylose causes inefficient metabolic state due to carbon/energy starvation and reduced glycolytic flux in

recombinant industrial *Saccharomyces cerevisiae*. PloS One 8, e69005.

doi:10.1371/journal.pone.0069005

Minard, K.I., McAlister-Henn, L., 1992. Glucose-induced degradation of the MDH2 isozyme of malate dehydrogenase in yeast. *J. Biol. Chem.* 267, 17458–17464.

Minard, K.I., McAlister-Henn, L., 1994. Glucose-induced phosphorylation of the MDH2 isozyme of malate dehydrogenase in *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.* 315, 302–309.

Müller, S., Boles, E., May, M., Zimmermann, F.K., 1995. Different internal metabolites trigger the induction of glycolytic gene expression in *Saccharomyces cerevisiae*. *J. Bacteriol.* 177, 4517–4519.

Navas, M.A., Gancedo, J.M., 1996. The regulatory characteristics of yeast fructose-1,6-bisphosphatase confer only a small selective advantage. *J. Bacteriol.* 178, 1809–1812.

Nissen, T.L., Schulze, U., Nielsen, J., Villadsen, J., 1997. Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*. *Microbiol. Read. Engl.* 143 (Pt 1), 203–218.

Palmieri, L., Vozza, A., Agrimi, G., De Marco, V., Runswick, M.J., Palmieri, F., Walker, J.E., 1999. Identification of the yeast mitochondrial transporter for oxaloacetate and sulfate. *J. Biol. Chem.* 274, 22184–22190.

Pitkänen, J.-P., Aristidou, A., Salusjärvi, L., Ruohonen, L., Penttilä, M., 2003. Metabolic flux analysis of xylose metabolism in recombinant *Saccharomyces cerevisiae* using continuous culture. *Metab. Eng.* 5, 16–31.

Polakis, E.S., Bartley, W., 1965. Changes in the enzyme activities of *Saccharomyces cerevisiae* during aerobic growth on different carbon sources. *Biochem. J.* 97, 284–297.

- Rolland, F., De Winde, J.H., Lemaire, K., Boles, E., Thevelein, J.M., Winderickx, J., 2000. Glucose-induced cAMP signalling in yeast requires both a G-protein coupled receptor system for extracellular glucose detection and a separable hexose kinase-dependent sensing process. *Mol. Microbiol.* 38, 348–358.
- Runquist, D., Fonseca, C., Rådström, P., Spencer-Martins, I., Hahn-Hägerdal, B., 2009. Expression of the Gxf1 transporter from *Candida intermedia* improves fermentation performance in recombinant xylose-utilizing *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 82, 123–130. doi:10.1007/s00253-008-1773-y
- Runquist, D., Hahn-Hägerdal, B., Bettiga, M., 2009. Increased expression of the oxidative pentose phosphate pathway and gluconeogenesis in anaerobically growing xylose-utilizing *Saccharomyces cerevisiae*. *Microb. Cell Factories* 8, 49. doi:10.1186/1475-2859-8-49
- Runquist, D., Hahn-Hägerdal, B., Rådström, P., 2010. Comparison of heterologous xylose transporters in recombinant *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 3, 5. doi:10.1186/1754-6834-3-5
- Salusjärvi, L., Kankainen, M., Soliymani, R., Pitkänen, J.-P., Penttilä, M., Ruohonen, L., 2008. Regulation of xylose metabolism in recombinant *Saccharomyces cerevisiae*. *Microb. Cell Factories* 7, 18. doi:10.1186/1475-2859-7-18
- Sarthy, A.V., McConaughy, B.L., Lobo, Z., Sundstrom, J.A., Furlong, C.E., Hall, B.D., 1987. Expression of the *Escherichia coli* xylose isomerase gene in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 53, 1996–2000.

- Scalcinati, G., Otero, J.M., Van Vleet, J.R.H., Jeffries, T.W., Olsson, L., Nielsen, J., 2012. Evolutionary engineering of *Saccharomyces cerevisiae* for efficient aerobic xylose consumption. *FEMS Yeast Res.* 12, 582–597. doi:10.1111/j.1567-1364.2012.00808.x
- Sonderegger, M., Jeppsson, M., Hahn-Hägerdal, B., Sauer, U., 2004. Molecular basis for anaerobic growth of *Saccharomyces cerevisiae* on xylose, investigated by global gene expression and metabolic flux analysis. *Appl. Environ. Microbiol.* 70, 2307–2317.
- Stephanopoulos, G., 2007. Challenges in engineering microbes for biofuels production. *Science* 315, 801–804. doi:10.1126/science.1139612
- Tantirungkij, M., Nakashima, N., Seki, T., Yoshida, T., 1993. Construction of xylose-assimilating *Saccharomyces cerevisiae*. *J. Ferment. Bioeng.* 75, 83–88. doi:10.1016/0922-338X(93)90214-S
- Toivari, M.H., Aristidou, A., Ruohonen, L., Penttilä, M., 2001. Conversion of xylose to ethanol by recombinant *Saccharomyces cerevisiae*: importance of xylulokinase (XKS1) and oxygen availability. *Metab. Eng.* 3, 236–249. doi:10.1006/mben.2000.0191
- Van Maris, A.J.A., Abbott, D.A., Bellissimi, E., van den Brink, J., Kuyper, M., Luttik, M.A.H., Wisselink, H.W., Scheffers, W.A., van Dijken, J.P., Pronk, J.T., 2006. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie Van Leeuwenhoek* 90, 391–418. doi:10.1007/s10482-006-9085-7
- Van Winden, W.A., Wittmann, C., Heinzle, E., Heijnen, J.J., 2002. Correcting mass isotopomer distributions for naturally occurring isotopes. *Biotechnol. Bioeng.* 80, 477–479. doi:10.1002/bit.10393

Vaseghi, S., Macherhammer, F., Zibek, S., Reuss, M., 2001. Signal transduction dynamics of the protein kinase-A/phosphofruktokinase-2 system in *Saccharomyces cerevisiae*. *Metab. Eng.* 3, 163–172. doi:10.1006/mben.2000.0179

Vincent, O., Townley, R., Kuchin, S., Carlson, M., 2001. Subcellular localization of the Snf1 kinase is regulated by specific beta subunits and a novel glucose signaling mechanism. *Genes Dev.* 15, 1104–1114. doi:10.1101/gad.879301

Wahlbom, C.F., Cordero Otero, R.R., van Zyl, W.H., Hahn-Hägerdal, B., Jönsson, L.J., 2003. Molecular analysis of a *Saccharomyces cerevisiae* mutant with improved ability to utilize xylose shows enhanced expression of proteins involved in transport, initial xylose metabolism, and the pentose phosphate pathway. *Appl. Environ. Microbiol.* 69, 740–746.

Wahlbom, C.F., Eliasson, A., Hahn-Hägerdal, B., 2001. Intracellular fluxes in a recombinant xylose-utilizing *Saccharomyces cerevisiae* cultivated anaerobically at different dilution rates and feed concentrations. *Biotechnol. Bioeng.* 72, 289–296.

Walfridsson, M., Hallborn, J., Penttilä, M., Keränen, S., Hahn-Hägerdal, B., 1995. Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Appl. Environ. Microbiol.* 61, 4184–4190.

Wasylenko, T.M., Stephanopoulos, G., 2013. Kinetic isotope effects significantly influence intracellular metabolite (¹³C) labeling patterns and flux determination. *Biotechnol. J.* 8, 1080–1089. doi:10.1002/biot.201200276

- Wasylenko, T.M., Stephanopoulos, G., 2015. Metabolomic and (13) C-metabolic flux analysis of a xylose-consuming *Saccharomyces cerevisiae* strain expressing xylose isomerase. *Biotechnol. Bioeng.* 112, 470-483. doi:10.1002/bit.25447
- Wiechert, W., de Graaf, A.A., 1997. Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments. *Biotechnol. Bioeng.* 55, 101–117. doi:10.1002/(SICI)1097-0290(19970705)55:1<101::AID-BIT12>3.0.CO;2-P
- Wisselink, H.W., Cipollina, C., Oud, B., Crimi, B., Heijnen, J.J., Pronk, J.T., van Maris, A.J.A., 2010. Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. *Metab. Eng.* 12, 537–551. doi:10.1016/j.ymben.2010.08.003
- Wittmann, Heinzle, 1999. Mass spectrometry for metabolic flux analysis. *Biotechnol. Bioeng.* 62, 739–750.
- Wu, L., Mashego, M.R., van Dam, J.C., Proell, A.M., Vinke, J.L., Ras, C., van Winden, W.A., van Gulik, W.M., Heijnen, J.J., 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ¹³C-labeled cell extracts as internal standards. *Anal. Biochem.* 336, 164–171. doi:10.1016/j.ab.2004.09.001
- Yin, Z., Hatton, L., Brown, A.J., 2000. Differential post-transcriptional regulation of yeast mRNAs in response to high and low glucose concentrations. *Mol. Microbiol.* 35, 553–565.
- Young, E.M., Comer, A.D., Huang, H., Alper, H.S., 2012. A molecular transporter engineering approach to improving xylose catabolism in *Saccharomyces cerevisiae*. *Metab. Eng.* 14, 401–411. doi:10.1016/j.ymben.2012.03.004

Young, J.D., Shastri, A.A., Stephanopoulos, G., Morgan, J.A., 2011. Mapping photoautotrophic metabolism with isotopically nonstationary (13)C flux analysis. *Metab. Eng.* 13, 656–665.

doi:10.1016/j.ymben.2011.08.002

Zhou, H., Cheng, J.-S., Wang, B.L., Fink, G.R., Stephanopoulos, G., 2012. Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*.

Metab. Eng. 14, 611–622. doi:10.1016/j.ymben.2012.07.011

Chapter 6

The Oxidative Pentose Phosphate Pathway is the Primary Source of NADPH for Lipid Overproduction from Glucose in *Yarrowia lipolytica*

This chapter is adapted from a manuscript which has been submitted to Metabolic Engineering.

6.1 Introduction

In recent years, concerns about the sustainability and environmental impact of fossil fuels have given impetus to research into alternative sources of liquid fuels for the transportation industry. Biodiesel has emerged as a promising biofuel and is currently produced predominantly by transesterification of vegetable oils derived from rapeseed, palm, and soybean, with a small fraction being produced from waste oils (Gui et al., 2008). However, given the finite supply of arable land and increasing demand for food, it is unlikely that food-based biofuels will contribute significantly to meeting future transportation energy needs (Hill et al., 2006). It has been proposed that biodiesel could be produced utilizing oleaginous microbes capable of accumulating lipids produced from a wide range of feedstocks (Li et al., 2008; Meng et al., 2009; Ageitos et al., 2011). Such a process for production of “single-cell oil” would have several advantages over current biodiesel production technologies. The process would be independent of climate and variations in weather (e.g. drought), and production facilities could be located on

non-arable land so as to not compete directly with the food supply (Beopoulos et al., 2011). Most importantly, land productivity in the case of carbohydrate production and conversion to oil is much higher than can be achieved by direct production of oil crops. Economical production of biodiesel using oleaginous microbes will require optimization of the biocatalysts, which can be achieved through metabolic engineering (Bailey, 1991). The oleaginous yeast *Yarrowia lipolytica* has emerged as both a model oleaginous organism and a promising host for single-cell oil production because of the availability of genome sequence data (Dujon et al., 2004) and well-developed tools for genetic engineering (Beopoulos et al., 2009). Although much work has been done to elucidate the mechanism of lipid accumulation in oleaginous microbes (Ratledge, 2002), to guide the engineering of *Y. lipolytica* strains which can accumulate lipids at high yield, titer, and productivity a more comprehensive understanding will be required.

High lipid accumulation in oleaginous microbes is typically achieved with nitrogen-limiting culture conditions. After nitrogen is depleted from the growth medium, cells can no longer synthesize nucleotides and proteins. Consequently, oleaginous microbes begin to convert the excess carbon source remaining in the medium to lipids, which are generally stored as triacylglycerols (TAGs). It is believed that in many oleaginous organisms lipid accumulation is initiated by a decrease in the intracellular adenosine monophosphate (AMP) pool (Boulton and Ratledge, 1983), which may be mediated by AMP deaminase (which converts AMP to inosine monophosphate and ammonium in order to scavenge for nitrogen) (Evans and Ratledge, 1985; Wynn et al., 2001) or may simply result from a decreased demand for ATP upon cessation of nucleotide and protein synthesis (Marchal et al., 1977). AMP is an allosteric regulator of mitochondrial isocitrate dehydrogenase (IDH) (Mitsushima et al., 1978; Botham and Ratledge,

1979), and the resultant decrease in IDH activity causes an accumulation of isocitrate and citrate (which may be rapidly interconverted by aconitase) in the mitochondria. The accumulated citrate is transported out to the cytosol (usually the mechanism is citrate-malate antiport) (Evans et al., 1983a, 1983b), where it is acted upon by ATP citrate lyase, generating a molecule of oxaloacetate and a molecule of acetyl-CoA (AcCoA) (Boulton and Ratledge, 1981). The AcCoA is the substrate for fatty acid biosynthesis. The oxaloacetate can be converted to malate by cytosolic malate dehydrogenase (MDH). As one molecule of malate is transported into the mitochondria in exchange for citrate, the malate can be converted back to oxaloacetate by mitochondrial MDH and condense with a mitochondrial AcCoA two-carbon unit (catalyzed by citrate synthase) to regenerate mitochondrial citrate. Thus, the net effect of this process is the shuttling of one molecule of AcCoA from the mitochondria to the cytosol (and one NADH reducing equivalent from the cytosol into the mitochondria), fueling lipid overproduction (Evans et al., 1983a). The cycle for shuttling AcCoA from the mitochondria to the cytosol for fatty acid synthesis is depicted in Figure 6-1.

Because lipids are highly reduced species, TAG overproduction also requires large quantities of NADPH, with two NADPH being oxidized to NADP^+ for each fatty acid elongation step and one NADPH being consumed for each fatty acid desaturation reaction. It is unclear how oleaginous microbes meet this high NADPH demand during lipid accumulation. There is evidence that malic enzyme plays a major role in providing lipogenic NADPH in the oleaginous fungi *Mucor circinelloides* and *Mortierella alpina* (Wynn et al., 1997, 1999), and overexpression of malic enzyme in *M. circinelloides* resulted in a 2.5-fold increase in lipid content and an increase in the degree of unsaturation of accumulated fatty acids (Zhang et al., 2007). Overexpression of the *M.*

circinelloides malic enzyme in *Rhodotorula glutinis* also increased lipid content more than two-fold (Z. Li et al., 2013). A recent study found that malic enzyme activity increased significantly in *Y. lipolytica* after the onset of lipid accumulation (Ochoa-Estopier and Guillouet, 2014). However, the single malic enzyme detected in the *Y. lipolytica* genome is predicted to localize to the mitochondria, while lipid accumulation requires generation of NADPH in the cytosol, where lipid biosynthesis occurs (Beopoulos et al., 2011). Moreover, an *in vitro* study has shown that the *Y. lipolytica* malic enzyme prefers NAD⁺ to NADP⁺ (Zhang et al., 2013). Taken together, these results suggest malic enzyme does not play a meaningful role in providing reducing power for lipid accumulation in *Y. lipolytica*. Indeed, a transcriptomic study showed that malic enzyme expression did not change significantly upon depletion of nitrogen from the growth medium (Morin et al., 2011), and overexpression of neither the endogenous malic enzyme (Beopoulos et al., 2011) nor the NADP⁺-dependent malic enzyme from *M. alpina* (Zhang et al., 2013) had a significant effect on lipid accumulation in this oleaginous yeast. These results have led to speculation that the oxidative Pentose Phosphate Pathway (PPP) is the primary source of lipogenic NADPH in *Y. lipolytica* (Zhang et al., 2013; Tai and Stephanopoulos, 2013).

In order to investigate the impact of lipid overproduction on central carbon metabolism, we have conducted ¹³C-Metabolic Flux Analysis (MFA) (Wiechert, 2001) on a previously engineered *Y. lipolytica* strain overexpressing acetyl-CoA carboxylase and diacylglycerol acyltransferase, which is capable of producing lipids from glucose at high yield and to high final lipid content (Tai and Stephanopoulos, 2013). Precise flux estimates were obtained by fitting data from parallel labeling experiments with two complementary ¹³C-glucose tracers (Crown and

Antoniewicz, 2013). The results indicate the oxidative PPP is the primary source of lipogenic NADPH in *Y. lipolytica* overproducing TAGs from glucose.

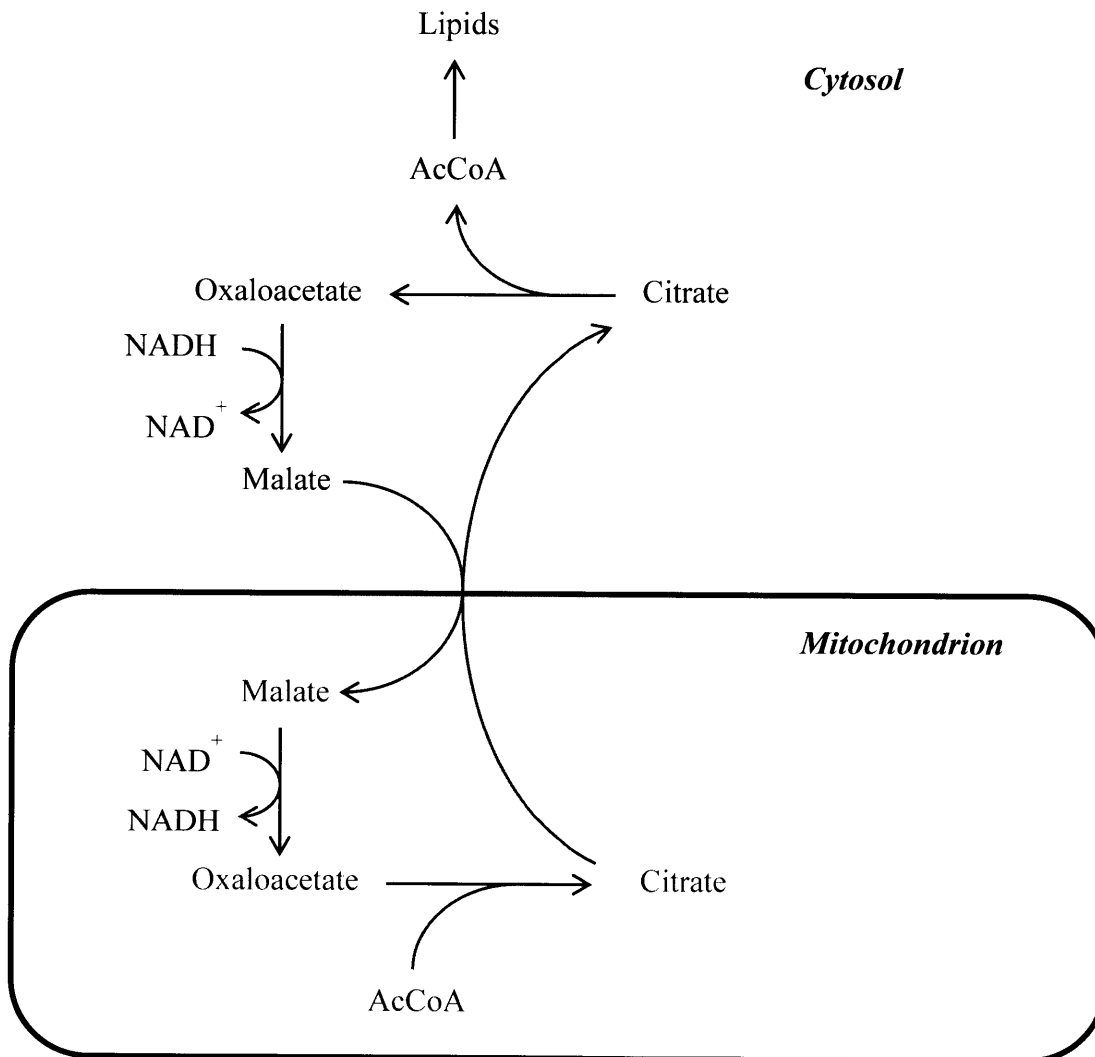


Figure 6-1 Cycle for shuttling acetyl-CoA from the mitochondria to the cytosol for fatty acid synthesis. Citrate acts as the carrier for acetyl-CoA two-carbon units. The cycle also results in transport of one NADH equivalent from the cytosol to the mitochondria and consumes one ATP in the ATP citrate lyase reaction.

6.2 Materials and Methods

6.2.1 Strains and Culture Conditions

Experiments were conducted with two previously described *Y. lipolytica* strains: a control strain MTYL037 and a lipid-overproducing strain MTYL065 which overexpresses *ACCI* (encoding acetyl-CoA carboxylase) and *DGAI* (encoding diacylglycerol acyltransferase) (Tai and Stephanopoulos, 2013). Strains were maintained at 4 °C on Yeast Extract-Peptone-Dextrose (YPD) plates and cultured at 30 °C in minimal media containing 20 g/l glucose as the sole carbon source, ammonium sulfate as the sole nitrogen source, and 1.7 g/l Yeast Nitrogen Base without amino acids and ammonium sulfate (Amresco, Solon, OH) as a source of salts, vitamins, and trace elements. All cultures were shaken at 250 rpm.

Initially a single 50 ml shake flask starter culture of each strain was inoculated from the corresponding YPD plate. The starter culture medium contained 5 g/l ammonium sulfate, resulting in an initial carbon/nitrogen (C/N) ratio of 8.8. After approximately 24 h, each starter culture was used to inoculate three shake flask cultures. The media for these cultures contained 0.44 g/l ammonium sulfate, resulting in an initial C/N ratio of 100. For each strain, each of the three shake flask cultures contained a different glucose substrate: (1) glucose labeled to natural abundance; (2) 1,2-¹³C₂-glucose (Omicron Biochemicals, South Bend, IN); and (3) 20% U-¹³C₆-glucose (Cambridge Isotope Laboratories, Tewksbury, MA). The three cultures were otherwise identical.

Cells were washed prior to inoculation in order to remove residual starter culture medium and prevent carryover of unlabeled glucose from the starter cultures to the cultures containing ^{13}C -labeled glucose substrates. 1 ml starter culture was centrifuged 5 min at maximum speed (18,000g) and the supernatant discarded. The cell pellet was resuspended in 1 ml of the medium to be used in the labeling experiment (e.g. the cell pellet to be used for inoculation of a 1,2- $^{13}\text{C}_2$ -glucose culture would have been resuspended in 1,2- $^{13}\text{C}_2$ -glucose medium), centrifuged a second time, and the supernatant discarded. Finally, cells were resuspended in 1 ml of the medium to be used in the labeling experiment and transferred to the 50 ml shake flask culture.

6.2.2 Extracellular Metabolite Quantification

Extracellular glucose, citrate, mannitol, and erythritol concentrations were measured by High-Performance Liquid Chromatography (HPLC). Cells were removed from culture samples by filtration through 0.20 μm Nylon syringe filters (Microliter Analytical #F13-2020-1GF). HPLC was performed with a Waters 2695 Separation Module coupled to a Waters 410 Differential Refractometer. Metabolite separation was achieved on an Aminex HPX-87H column (Bio-Rad, Hercules, CA) with an isocratic flow of 14 mM sulfuric acid mobile phase at 0.7 ml/min and a temperature of 50 °C. The injection volume was 10 μl .

Extracellular ammonium was quantified using a commercially available Ammonia Assay Kit (Sigma-Aldrich, #AA0100).

6.2.3 Dry Cell Weight Measurements

Cell densities were monitored by measuring dry cell weight (DCW). 1 ml culture was vacuum-filtered on a pre-weighed cellulose nitrate filter (0.2 μm , Whatman #10401312). Filters were washed with two volumes Millipore water, dried at 60 °C, and weighed intermittently until the mass no longer changed over time. At each time point two control filters were prepared by filtering 1 ml cell-free natural abundance glucose medium. The control filters were used to correct for the losses in filter mass that occur during sample preparation.

6.2.4 Lipid Analysis

The five principal fatty acid species synthesized by *Y. lipolytica* (palmitate, C16:0; palmitoleate, C16:1; stearate, C18:0; oleate, C18:1; linoleate, C18:2) were quantified using a Gas Chromatography-Flame Ionization Detector (GC-FID) setup. No other fatty acid species were detected in significant quantities. Depending on cell density, 0.7-1 ml cell culture was centrifuged 10 min at 18,000g and the supernatant removed. The volume of cell culture was chosen so that each sample would contain roughly 2-3 mg biomass (as measured by DCW). Cell pellets were stored at -20 °C until derivatization. For lipid analysis, 100 μl internal standard solution (2 mg/ml methyl tridecanoate (Sigma-Aldrich #T0627) + 2 mg/ml glyceryl triheptadecanoate (Sigma-Aldrich #T2151) in hexane) was added to each cell pellet. Methyl tridecanoate was used to correct for volume losses during sample preparation and glyceryl triheptadecanoate was used to correct for transesterification efficiency. Lipids were transesterified to fatty acid methyl esters (FAMES) by addition of 500 μl 0.5 N sodium methoxide (20 g/l sodium hydroxide in methanol) followed by 60 min vortexing at 1200 rpm. Samples were then neutralized by addition of 40 μl sulfuric acid. FAMES were extracted by

addition of 500 μ l hexane followed by another 30 min vortexing at 1200 rpm. Samples were centrifuged 1 min at 8000 rpm (roughly 6000g) and the hexane layer was subsequently analyzed on a Bruker 450-GC Gas Chromatograph equipped with a FID. Samples were injected with an injection volume of 1 μ l, split ratio of 10, and injector temperature of 260 °C. FAME species were separated on an Agilent J&W HP-INNOWax capillary column (30 m x 0.25 mm x 0.15 μ m, #19091N-033) with helium carrier gas at a flow rate of 1.5 ml/min. The column oven was held at a constant temperature of 200 °C for the duration of the method. The FID was operated at a temperature of 260 °C with a helium make up gas flow of 25 ml/min, hydrogen flow of 30 ml/min, and air flow of 300 ml/min.

6.2.5 Estimation of Extracellular Fluxes

The GC-FID lipid data were used to calculate the total moles of AcCoA and glycerol-3-phosphate (Glyc3P) consumed for lipid biosynthesis by assuming that one mole AcCoA was consumed for every two moles of carbon incorporated into fatty acids and that all lipids were in the form of TAGs and therefore one mole Glyc3P was required for every three moles of fatty acids produced. The total NADPH consumption for lipid biosynthesis was also calculated by assuming that two NADPH were required for each fatty acid elongation step and one NADPH was required for each fatty acid desaturation step. However, no mass balance constraint on NADPH was included in the metabolic model, as inclusion of such redox cofactor balances may introduce significant errors due to assumptions about enzyme cofactor preferences and possible omission of important sources or sinks from the metabolic model (Marx et al., 1996; Schmidt et al., 1998). Consequently, the NADPH consumption calculations were not used in the metabolic

flux estimation. The AcCoA, Glyc3P, and NADPH requirements for synthesis of each of the five principal fatty acids produced by *Y. lipolytica* are summarized in Table 6-1.

Table 6-1 Acetyl-CoA, Glycerol-3-Phosphate, and NADPH requirements for biosynthesis of each of the five principal fatty acids in *Y. lipolytica*

Fatty Acid	AcCoA	Glyc3P	NADPH
Palmitate (C16:0)	8	0.333	14
Palmitoleate (C16:1)	8	0.333	15
Stearate (C18:0)	9	0.333	16
Oleate (C18:1)	9	0.333	17
Linoleate (C18:2)	9	0.333	18

All metabolic fluxes were normalized to a glucose uptake rate of 100. For each culture, the molar concentrations of citrate, mannitol, and erythritol and the molar consumption of lipogenic AcCoA, Glyc3P, and NADPH were plotted against the molar glucose concentration and the slope of each best-fit line was computed using the Microsoft Excel function LINEST. These slopes were then multiplied by 100 to determine the moles of citrate, mannitol, and erythritol produced and the moles of lipogenic AcCoA, Glyc3P, and NADPH consumed per 100 moles of glucose taken up. For each strain, the results from each of the three shake flask cultures were averaged and the standard deviations were computed as estimates of the uncertainties in the computed extracellular flux values.

6.2.6 Metabolite Extractions

7.5 ml culture was added to 37.5 ml pure methanol quenching solution (Canelas et al., 2008) held at low temperature (< -70 °C) in a cold ethanol bath. Cells were pelleted by centrifugation for 5 min at 3270g and -10 °C and the supernatant removed by aspiration. Cells were washed by

resuspension in 40 ml cold (< -70 °C) methanol followed by centrifugation and aspiration of the supernatant (as before). Metabolites were extracted from quenched cell pellets in 5 ml hot (80 °C) 75% (v/v) ethanol extraction solvent (Gonzalez et al., 1997; Canelas et al., 2009). Samples were vortexed for 30 s, held at 80 °C in a water bath for 3 min, vortexed again for 30 s, cooled briefly in the cold ethanol bath, and centrifuged to clear cell debris. Each metabolite extract was split into two equal fractions (each roughly 2.5 ml) for GC-MS and LC-MS/MS analysis, respectively, and stored at -80 °C.

6.2.7 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of Metabolite Extracts

Metabolite extracts were dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module and resuspended in 20 μ l 2% methoxyamine-hydrogen chloride in pyridine (Methoxamine (MOX) Reagent, Thermo Scientific #TS-45950). The methoximation reaction was allowed to proceed for 90 min at 37 °C. Subsequently a silylation reaction was performed by addition of 25 μ l N-tert-Butyldimethylsilyl-N-methyltrifluoroacetamide with 1% tert-Butyldimethylchlorosilane (Sigma-Aldrich, #375934) to each sample and incubation at 60 °C for 30 min. Samples were centrifuged 2 min to remove cell debris and the supernatants analyzed on an Agilent 6890N Network GC System coupled to an Agilent 5975B Inert XL MSD. The injection volume was 3 μ l. The GC-MS conditions have been described elsewhere (Wasylenko and Stephanopoulos, 2015) (see Chapter 5). Mass spectra were acquired in Selective Ion Monitoring (SIM) mode (Ahn and Antoniewicz, 2013).

6.2.8 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)

Metabolite extracts were dried under airflow using a Pierce Reacti-Therm III Heating/Stirring Module, resuspended in 80 μ l Millipore water, and analyzed on an Agilent 1100 Series HPLC system coupled to an API 2000 MS/MS (AB Sciex, Framingham, MA). The injection volume was 20 μ l. Metabolites were separated on a Waters XBridge C18 Column (2.1 mm x 150 mm, 130 Å, 3.5 μ m; #186003023) using an ion pair chromatography method adapted from (Luo et al., 2007; Young et al., 2011). The HPLC flow rate was 300 μ l/min and the solvent gradient was: 0% B (0 min), 0% B (8 min), 22.5% B (18 min), 40% B (28 min), 60% B (32 min), 90% B (34 min), 90% B (36 min), 100% B (37 min), 100% B (42 min), where A is 10 mM tributylamine + 15 mM acetic acid in water and B is methanol. Mass spectra were acquired in multiple reaction monitoring (MRM) mode (see Chapter 2).

6.2.9 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis of Glucose

The isotopic purities of 1,2-¹³C₂- and U-¹³C₆-glucose tracers were determined by GC-MS analysis of aldonitrile pentapropionate glucose derivatives (Antoniewicz et al., 2011). 5 μ l fresh glucose medium was dried using a Pierce Reacti-Therm III Heating/Stirring Module. Dry residues were dissolved in 50 μ l pyridine solution containing 20 mg/ml hydroxylamine hydrochloride and incubated at 90 °C for 60 min. 100 μ l propionic anhydride was added to each sample, followed by incubation at 60 °C for an additional 30 min. Samples were dried, resuspended in 100 μ l ethyl acetate, and centrifuged to remove insolubles. The supernatants were then analyzed by GC-MS. The injection volume was 1 μ l with split ratio 9:1, and the inlet temperature was 250 °C. The helium carrier gas flow rate was 0.9 ml/min. The GC column oven temperature gradient was: 80 °C (0 min), 80 °C (1 min), 280 °C (11 min), 280 °C (15 min).

Mass spectra were acquired by scanning the range 150-450 m/z. All other details of the GC-MS method were identical to those for analysis of metabolite extracts. Glucose tracer isotopic purities were estimated from ^{13}C -labeling of fragments at 259 m/z ($\text{U-}^{13}\text{C}_6\text{-glucose}$ only) and 284 and 370 m/z (both tracers) (Antoniewicz et al., 2011).

6.2.10 Metabolic Flux Estimation

The metabolic fluxes through a compartmentalized model reaction network were computed by nonlinear parameter estimation. The model network comprised the reactions of the Glycolysis, PPP, and Tricarboxylic Acid (TCA) Cycle pathways as well as pathways for production of extracellular citrate, mannitol, and erythritol and incorporation of AcCoA and Glyc3P into TAGs. Cytosolic pyruvate carboxylase (Otto et al., 2012) and mitochondrial malic enzyme (Beopoulos et al., 2011) reactions were also included. The non-oxidative PPP was modeled using the half-reaction method of (Kleijn et al., 2005). The model contained separate cytosolic and mitochondrial pools of AcCoA, citrate, malate, oxaloacetate, and pyruvate. For malate, oxaloacetate, and pyruvate, the ^{13}C -labeling patterns of the cytosolic and mitochondrial pools are not necessarily the same. For each of these “compartmentalized” metabolites, both the cytosolic and mitochondrial pools were allowed to contribute to the ^{13}C -labeling data, with the contribution of each pool left as a free parameter in the model. Pyruvate was assumed to be transported unidirectionally from the cytosol to the mitochondria, as has been assumed for *Saccharomyces cerevisiae* (Maaheimo et al., 2001). Citrate was assumed to be transported across the mitochondrial membrane in exchange with malate (Evans et al., 1983a, 1983b). Anaplerosis was assumed to be mediated by irreversible transport of malate from the cytosol to the mitochondria via the dicarboxylate carrier (Palmieri et al., 1996; Luévano-Martínez et al.,

2010), likely in exchange for phosphate and indirectly driven by the proton gradient across the mitochondrial membrane (Palmieri et al., 2000). Inclusion of either an oxaloacetate carrier reaction (Palmieri et al., 1999; Luévano-Martínez et al., 2010) or a reversible malate transport reaction did not have any significant effect on the flux estimation, and these reactions were therefore excluded. Of the gluconeogenic enzymes, PEP carboxykinase was omitted from the model while fructose-1,6-bisphosphatase was included, as the former is glucose-repressed (albeit incompletely) in *Y. lipolytica* while the latter is not (Jardón et al., 2008). The glyoxylate shunt was also omitted from the reaction network. Although this pathway is only incompletely repressed on glucose in *Y. lipolytica* (Flores and Gancedo, 2005; Jardón et al., 2008), Blank, *et al.* found no evidence of flux through the pathway in *Y. lipolytica* during growth on glucose (Blank et al., 2005). Consistent with these results, we found that the estimated flux through the glyoxylate shunt was negligible when the pathway was included in the model (data not shown). The complete model reaction network with atom transitions can be found in Table B-1 (Appendix B).

Metabolic flux estimation for each strain was achieved by generating a random initial flux distribution, using the model network to simulate the expected extracellular flux and metabolite labeling data corresponding to this flux distribution, and then iteratively refining the flux distribution to minimize the lack-of-fit (as measured by the weighted sum of squared residuals) between the simulated data and the data obtained in the carbon labeling experiments. The flux estimation for each strain was repeated 500 times with different initial guesses and the smallest weighted sum of squared residuals obtained taken to be the global optimum. The best-fit flux distributions were subjected to a chi-square test for goodness-of-fit, and 68% and 95%

confidence intervals were calculated for each flux using a parameter continuation technique (Antoniewicz et al., 2006). The uncertainties in the ^{13}C -labeling data mass isotopomer mole fractions were assumed to be 0.4 mol % for the goodness-of-fit test and confidence interval determination (Antoniewicz et al., 2007a; Wasylenko and Stephanopoulos, 2013). All computations were performed with an in-house software that formulates the simulation problem in terms of elementary metabolite units (EMUs) (Antoniewicz et al., 2007b). The expected effects of naturally occurring heavy isotopes were added to simulated data, and the corrected simulated data were compared directly to experiment data (Wittmann and Heinzle, 1999; van Winden et al., 2002). Reversible reactions were modeled as two fluxes, a net flux and an exchange flux (Wiechert and de Graaf, 1997).

6.3 Results

6.3.1 Fermentation Profiles and Establishment of Metabolic Steady State

We compared the performances of the engineered lipid-overproducing strain MTYL065 and the control strain MTYL037 in a low-nitrogen medium that promotes lipid accumulation (initial C/N ratio = 100). Each strain was cultured in triplicate, and for each strain one replicate was cultured in 1,2- $^{13}\text{C}_2$ -glucose medium and another in 20% U- $^{13}\text{C}_6$ -glucose medium to facilitate ^{13}C -MFA. The third replicate culture for each strain contained glucose labeled to natural abundance. The fermentation profiles for the six shake flask cultures are shown in Figures 6-2 and 6-3. The engineered strain MTYL065 consumed glucose and produced lipids more rapidly than the control strain MTYL037. The three MTYL065 cultures achieved final lipid contents of 57-59% compared to only 22-26% for MTYL037. The final lipid profiles of the two strains were similar

(Figure 6-4), with oleate and palmitate being the most abundant fatty acid species and accounting for roughly 50% and 25% of the total lipids, respectively. The only by-products detected by HPLC were citrate, mannitol, and erythritol. The two strains produced mannitol at similar rates and titers, although, since MTYL065 consumed glucose more rapidly, its mannitol yield was significantly lower than that of MTYL037. The engineered strain produced small amounts of citrate during the initial phase of the fermentation, but in the latter stages (after the ammonium supply was exhausted) citrate production was negligible. The control strain continued to produce citrate even after the ammonium in the medium was depleted. Both strains produced only small amounts of erythritol.

In conventional ^{13}C -MFA, it is assumed that intracellular metabolite labeling patterns are at steady state. To satisfy this assumption, cultures must be maintained in a “metabolic steady state” where the metabolic fluxes are invariant in time until the intracellular metabolite ^{13}C -labeling patterns reach “isotopic steady state,” at which point metabolites are harvested (Wiechert and de Graaf, 1997; Wiechert, 2001). Labeling data for this study were obtained from primary central carbon metabolites directly rather than from biomass macromolecule hydrolysates. Due to the fast turnover times of these central carbon metabolites, if a culture can be maintained in metabolic steady state the metabolite ^{13}C -labeling patterns are expected to reach isotopic steady state relatively quickly (van Winden et al., 2005; Canelas et al., 2008).

Metabolic steady state conditions can be achieved in chemostat cultures, and in batch cultures it is generally assumed that exponential growth phase approximates a metabolic steady state. In a *Y. lipolytica* culture with a high initial C/N ratio, the cells are expected to grow roughly

exponentially until the ammonium nitrogen source is depleted from the medium. At this point, the cells can no longer produce nucleotide and protein biomass and lipid accumulation commences. This lipid accumulation phase is more relevant than exponential growth phase for lipid overproduction; however it was not obvious that the cultures would be in a metabolic steady state during this phase.

If the cultures were to reach a steady state during the lipid accumulation phase, the number of cells would be expected to be roughly constant due to the absence of nitrogen in the growth medium. A constant number of cells with constant metabolic fluxes would be expected to generate metabolite profiles which vary linearly in time. We observed that by 40 h the ammonium supply had been exhausted in all cultures (Figure 6-2) and between 40 and 64 h all cultures exhibited roughly linear glucose consumption and lipid, citrate, mannitol, and erythritol production profiles (Figures 6-2 and 6-3). Thus we concluded that between 40 and 64 h the cultures were in lipid accumulation phase and could be assumed to approximate a metabolic steady state. The metabolic flux estimation was performed using data acquired during this 24 h timeframe, which is sufficiently long that the central carbon metabolites could be expected to be at isotopic steady state at the end of the experiment.

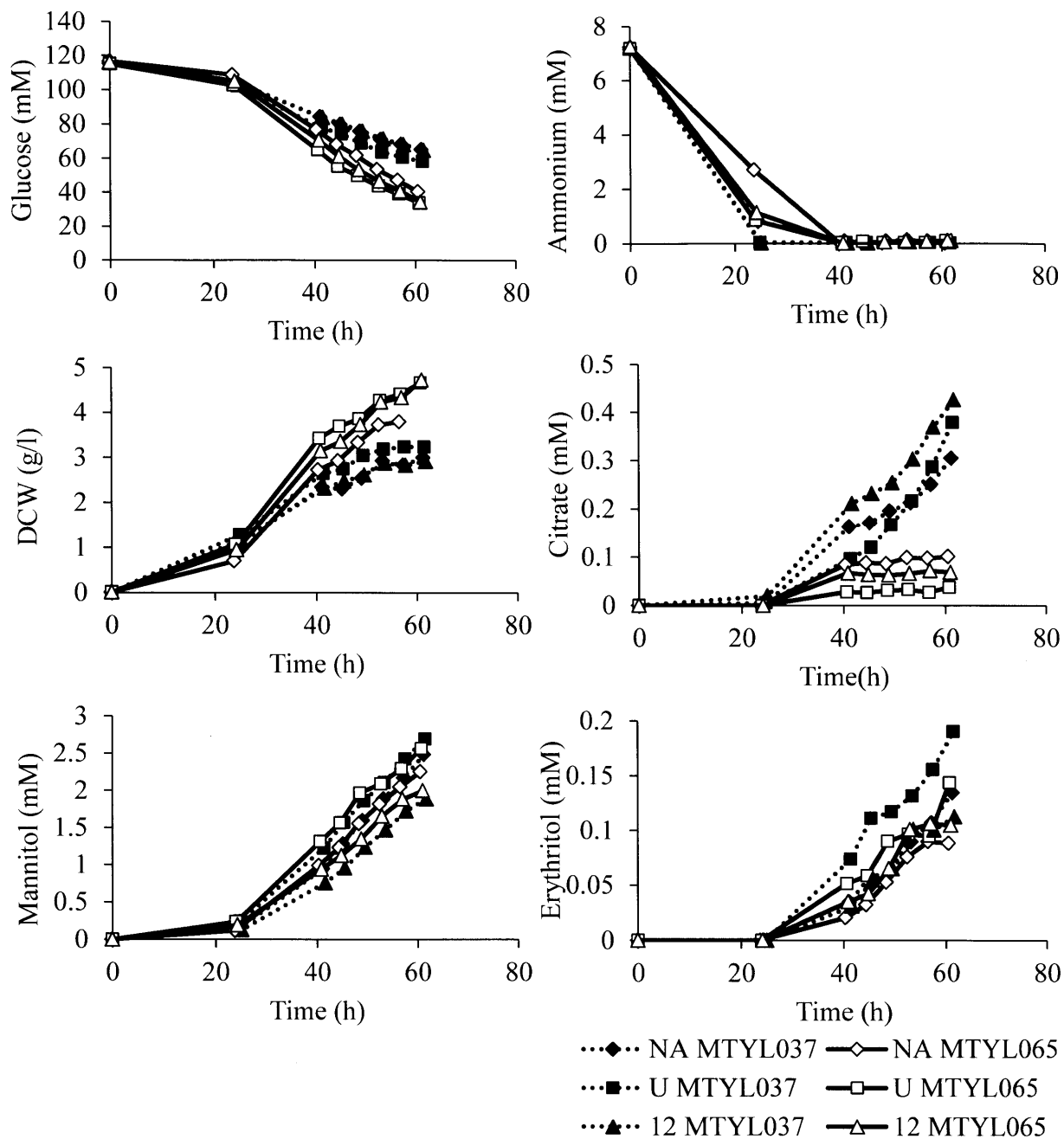


Figure 6-2 Extracellular metabolite fermentation profiles. Time courses for glucose and ammonium consumption and dry cell weight, citrate, mannitol, and erythritol accumulation in shake flask cultures of strains MTYL037 and MTYL065. Abbreviations denoting glucose substrate for each culture: NA, natural abundance glucose; U, 20% U-¹³C₆-glucose; 12, 1,2-¹³C₂-glucose.

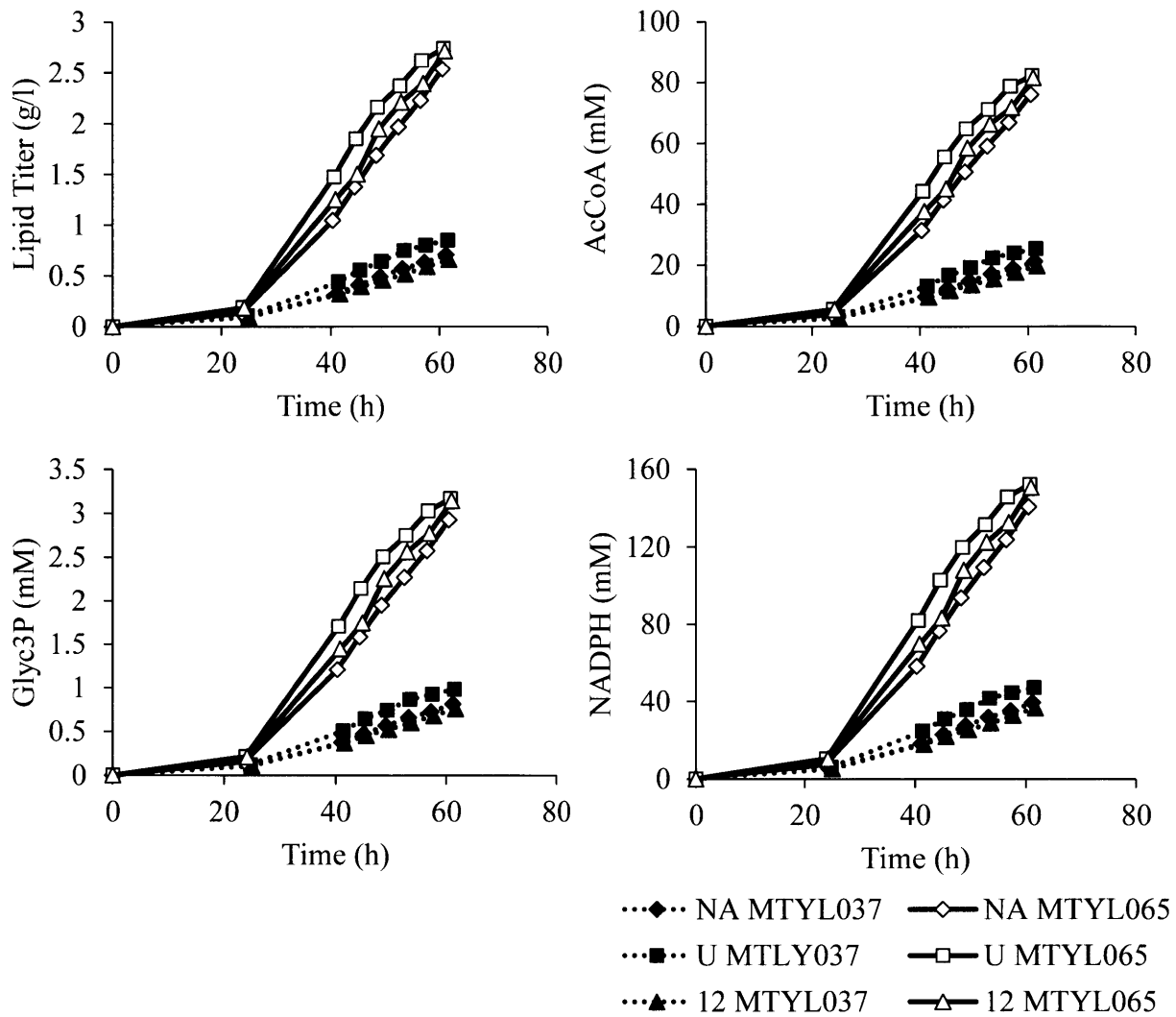


Figure 6-3 Fatty acid fermentation profiles. Time courses for total lipid titer and total lipogenic AcCoA, Glyc3P, and NADPH consumed for TAG biosynthesis in shake flask cultures of strains MTYL037 and MTYL065. Abbreviations denoting glucose substrate for each culture: NA, natural abundance glucose; U, 20% U-¹³C₆-glucose; 12, 1,2-¹³C₂-glucose.

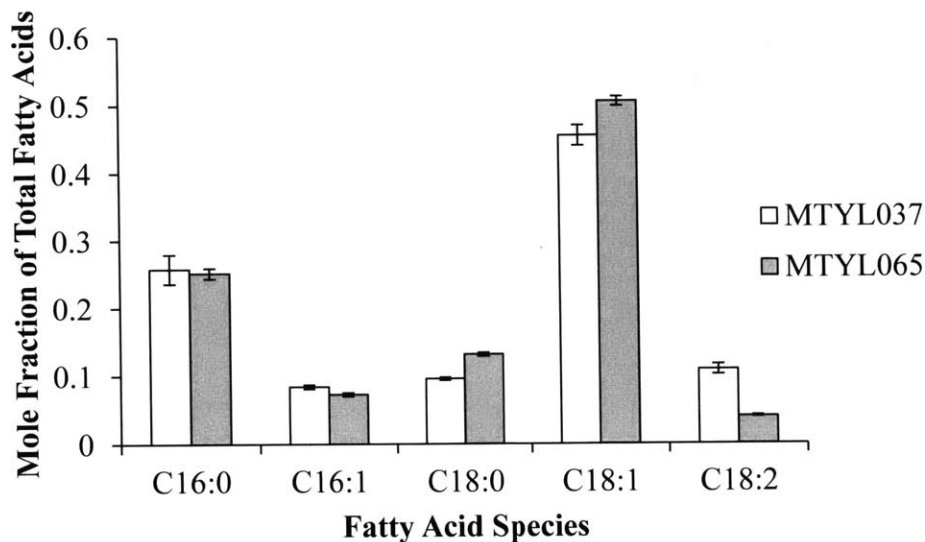


Figure 6-4 Final fatty acid distribution profiles. The relative abundances of the five principal fatty acid species in strains MTYL037 and MTYL065 at the time cells were harvested for ^{13}C -MFA.

6.3.2 Estimation of Extracellular Fluxes

The glucose uptake rate was fixed to an arbitrary value of 100 for all cultures to facilitate comparison of the intracellular metabolic flux distributions in the two strains. The other extracellular fluxes were then determined from the yields of citrate, mannitol, erythritol, and lipogenic AcCoA, Glyc3P, and NADPH per 100 mol glucose, which were calculated by modeling the HPLC and GC-FID data acquired between 40 and 64 h through linear regression (see Materials and Methods). The data used for the linear regressions and the resulting best-fit lines are shown in Figures B-1-B-6 (Appendix B). The extracellular flux values and their uncertainties used in metabolic flux estimation are listed in Table 6-2. It can be seen that the lipid yield achieved by the engineered strain was more than two-fold higher than that of the

control. The control strain produced all three by-products citrate, mannitol, and erythritol at higher yield than the engineered strain.

Table 6-2 Extracellular flux values used for ¹³C-MFA

	MTYL037	MTYL065
Glucose	100	100
Citrate	2.1 ± 0.7	0
Mannitol	6.9 ± 0.9	3.6 ± 0.5
Erythritol	0.47 ± 0.03	0.24 ± 0.03
AcCoA	56 ± 4	124 ± 3
Glyc3P	2.2 ± 0.2	4.75 ± 0.09
NADPH	104 ± 8	228 ± 5

6.3.3 Metabolic Flux Estimation

¹³C-MFA was performed on strains MTYL037 and MTYL065 using the estimated extracellular flux values and GC-MS and LC-MS/MS metabolite labeling data as inputs. GC-MS analysis of the glucose tracers confirmed that each tracer had an isotopic purity of approximately 99%, consistent with manufacturer specifications. The best-fit metabolic flux distributions are shown in Figure 6-5, and the confidence intervals for key fluxes are shown in Figure 6-6. The weighted sums of squared residuals for the MTYL037 and MTYL065 flux estimation models were 198.9 and 165.9, respectively. These values fell within the range that would be expected given that the models adequately describe the data, which was 145.6 to 220.2. The measured and simulated ¹³C-labeling data from the flux estimation algorithm are presented in Table B-2 (Appendix B), and the full set of best-fit flux values and flux confidence intervals are tabulated in Table B-3 (Appendix B).

The computed metabolic flux distributions reveal that the flux of AcCoA to lipids in the engineered strain MTYL065 is increased more than twofold over the control strain, and there is a dramatic rearrangement of the metabolic flux distribution to support lipid overproduction. The majority of mitochondrial citrate produced by citrate synthase is exported to the cytosol as a shuttle for lipogenic AcCoA. Consequently, the flux through isocitrate dehydrogenase is decreased by roughly 70%, although there is still significant flux through the TCA Cycle. The fluxes through ATP citrate lyase and cytosolic malate dehydrogenase are also increased to support the shuttling of AcCoA across the mitochondrial membrane. Malate is transported from the cytosol to the mitochondria in exchange for citrate (and also by the dicarboxylate carrier). However, the most striking feature of the metabolic flux distributions is the large increase in flux through the oxidative PPP in the engineered strain.

In the control strain MTYL037 the flux through the oxidative PPP is estimated to be 52.0 (Figure 6-5). As 2 moles of NADPH are produced for each mole of glucose-6-phosphate (G6P) metabolized through this pathway, the total NADPH production by the oxidative PPP is estimated to be 104 moles per 100 moles glucose consumed. This number is in perfect agreement with the estimate for NADPH consumed in fatty acid biosynthesis (Table 6-2). (Note that an NADPH mass balance was not included in the metabolic network, so the matching of the rates of NADPH consumption and production is non-trivial.) Thus the flux through the oxidative PPP appears to be sufficient to supply all of the NADPH required for lipid production in MTYL037.

In the engineered strain MTYL065, the flux of AcCoA to TAG biosynthesis is increased roughly two-fold (from 59.8 to 123.3). The flux through the oxidative PPP is also increased by roughly a factor of two, from 52.0 to 102.2. In fact, the flux through the oxidative PPP exceeds the glucose uptake flux; the flux through phosphoglucose isomerase is reversed with a small net flux from fructose-6-phosphate supplying the G6P substrate for the oxidative PPP. In this strain, the predicted NADPH production rate of 204.4 is in good agreement with the estimated NADPH requirement for fatty acid biosynthesis, which is 228 (Table 6-2), although there is a roughly 10% gap between the rates of NADPH production and consumption. This gap could result from experimental error or could indicate that an alternative source of NADPH not included in the model contributes to NADPH production in MTYL065. Regardless, the oxidative PPP appears to supply the majority of the reducing power required for fatty acid biosynthesis in this strain.

The flux confidence intervals in Figure 6-6 show that the flux through the oxidative PPP was estimated with good resolution in both strains. The correlation between the flux of AcCoA to lipids and the oxidative PPP flux is clear. However, the flux through malic enzyme was not significantly different in the two strains, suggesting that malic enzyme does not play a significant role in lipid production in *Y. lipolytica*.

The increase in flux through the oxidative PPP in MTYL065 is evident from the ^{13}C -labeling data. The mass isotopomer distributions (MIDs) of several key metabolites from the 1,2- $^{13}\text{C}_2$ -glucose labeling experiment are shown in Figure 6-7. With this tracer, the majority of the carbon atoms lost as CO_2 in the 6-phosphogluconate dehydrogenase reaction will be ^{13}C . Consequently, increased flux through the oxidative PPP will result in increased abundance of light mass

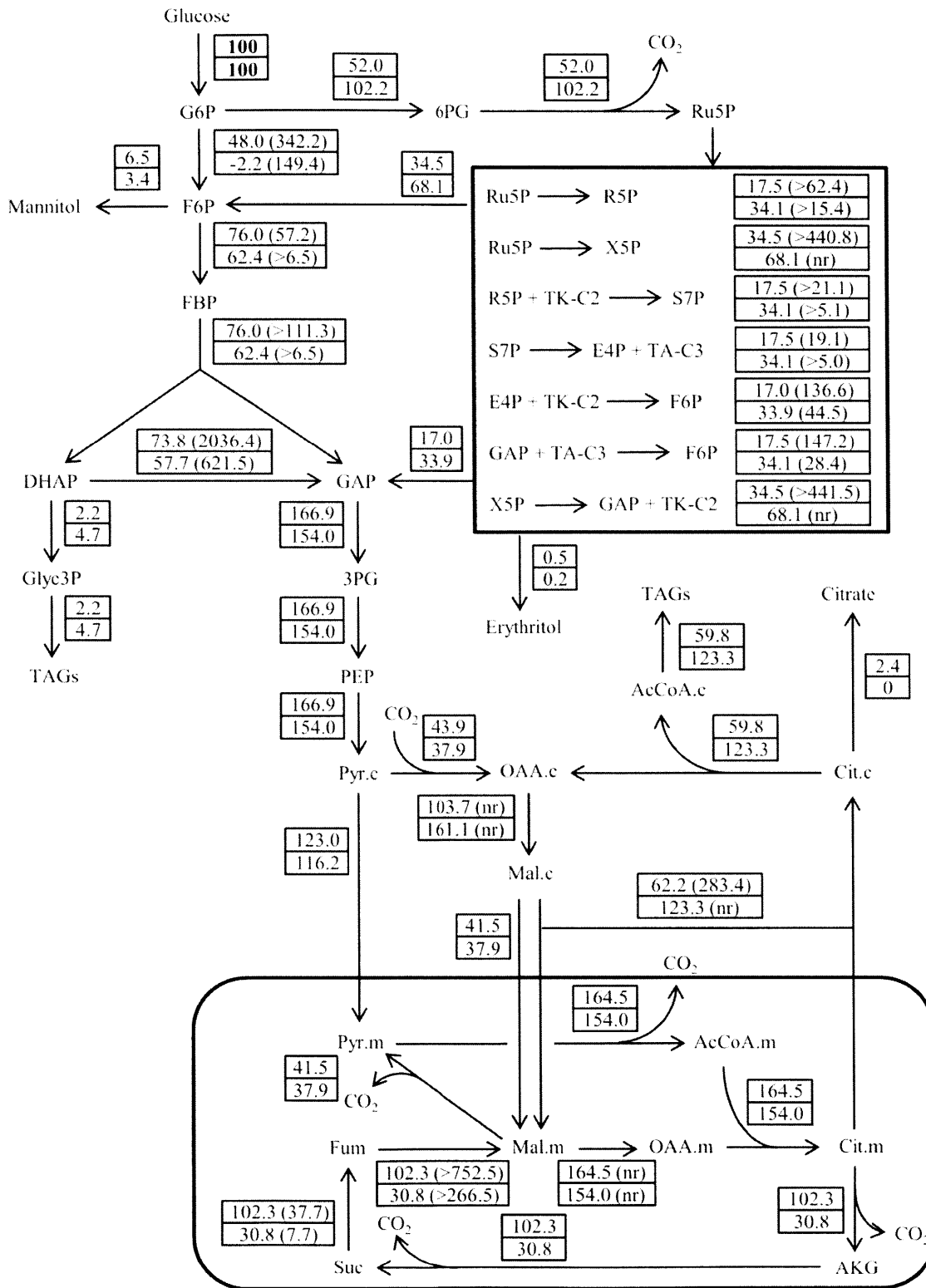


Figure 6-5 Estimated metabolic flux distributions. Two flux values are listed for each reaction; these are the fluxes in the control strain MTYL037 (top) and the engineered strain MTYL065 (bottom). Each reversible reaction is described by the net flux with the exchange flux listed inside parentheses. Exchange fluxes denoted “nr” could not be resolved within one order of magnitude.

isotopomers and decreased abundance of heavy mass isotopomers in the MIDs. G6P and 6-phosphogluconate (6PG) must in theory have identical ^{13}C -labeling patterns, since 6PG is derived from G6P via a linear pathway. The MIDs of G6P and 6PG are in good agreement for both strains. It can be seen that the relative abundances of the lighter M+1 and M+2 mass isotopomers are higher in MTYL065 than in MTYL037 while the relative abundances of the heavier mass isotopomers (M+3 and higher) are lower, reflecting the increased oxidative PPP flux in the engineered strain. Dihydroxyacetone phosphate (DHAP) and 3-phosphoglycerate (3PG) are also expected to have similar ^{13}C -labeling patterns due to the high exchange flux of the triosephosphate isomerase reaction (Figure 6-5), although the DHAP MID is skewed by the presence of derivatization groups added to facilitate GC-MS analysis. Nonetheless, it can be seen that in both the DHAP and 3PG MIDs the relative abundances of the M+0 and M+1 mass isotopomers are higher in MTYL065 than in MTYL037 while the relative abundances of the heavier mass isotopomers (M+2 and higher) are lower, providing further confirmation that the oxidative PPP flux is elevated in the engineered strain.

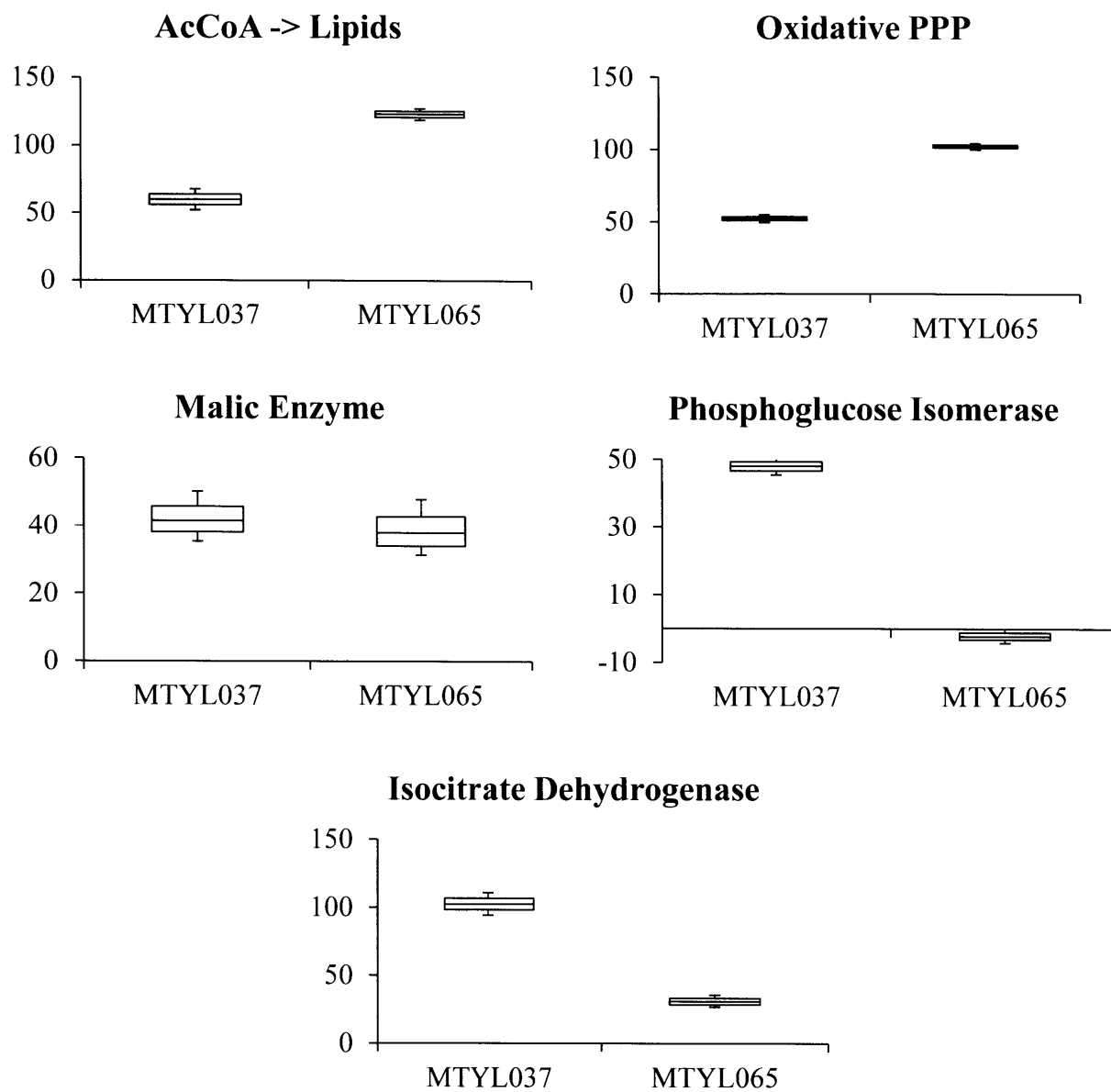


Figure 6-6 Select flux confidence intervals. Confidence intervals are shown for fluxes which illustrate the key differences between the flux distributions in strains MTYL037 and MTYL065. 68% confidence intervals are indicated by boxes and 95% confidence intervals by error bars (“whiskers”). The line inside each box indicates the best-fit flux value. All fluxes are normalized to a glucose uptake rate of 100.

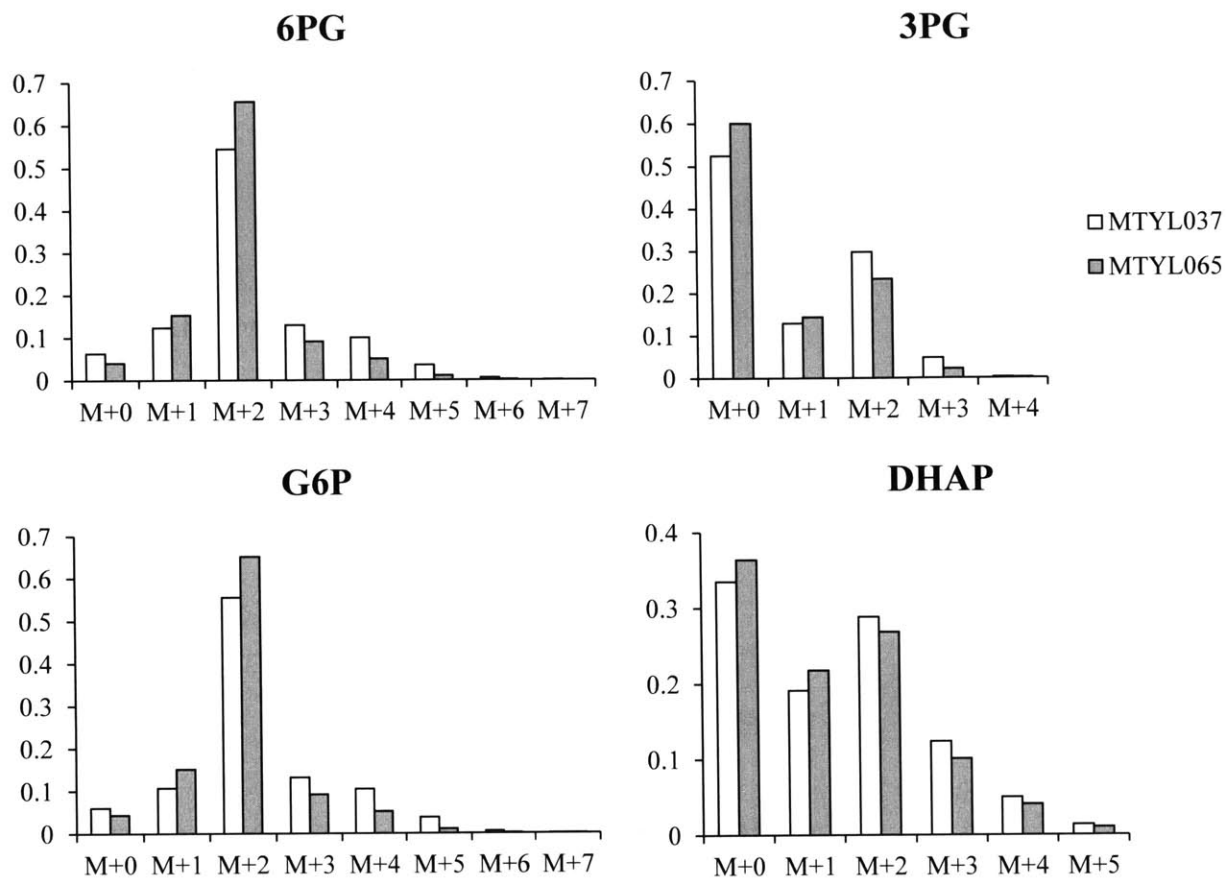


Figure 6-7 Mass isotopomer distributions (MIDs) of illustrative metabolites from the 1,2-¹³C₂-glucose labeling experiment. The relative abundances (i.e. mole fractions) of the various mass isotopomers of 6-phosphogluconate (6PG), glucose-6-phosphate (G6P), 3-phosphoglycerate (3PG), and dihydroxyacetone phosphate (DHAP) in strains MTYL037 and MTYL065 are shown. 6PG, G6P, and 3PG were measured by LC-MS/MS, DHAP by GC-MS.

6.4 Discussion

We have conducted ^{13}C -MFA of two *Y. lipolytica* strains during lipid accumulation on glucose. High flux resolution was achieved by combining ^{13}C -labeling data from parallel labeling experiments utilizing two different ^{13}C -glucose tracers, 1,2- $^{13}\text{C}_2$ -glucose and 20% U- $^{13}\text{C}_6$ glucose (Crown and Antoniewicz, 2013). The ^{13}C -labeling data imply that there is a dramatic increase in flux through the oxidative PPP in the lipid-overproducing strain MTYL065 relative to the control strain MTYL037. Indeed, quantitative metabolic flux estimation revealed that the NADPH-generating flux through the oxidative PPP is comparable to the NADPH-consuming flux of AcCoA to fatty acid biosynthesis in both strains. The flux through malic enzyme was similar in both strains and did not correlate with lipid production. Taken together, these results indicate that the oxidative PPP is likely the primary source of lipogenic NADPH during lipid overproduction from glucose in *Y. lipolytica*.

As noted above, an NADPH mass balance constraint was not included in the metabolic model; the model contained no information about production or consumption of NADPH equivalents. Therefore the result that the estimated flux through the NADPH-producing reactions of the oxidative PPP increases roughly proportionally to the NADPH-consuming flux of AcCoA to lipids in MTYL065 is non-trivial. With the modeling approach used, it would have been mathematically possible for the flux of AcCoA to lipids to increase without any increase in flux through NADPH-producing reactions. The flux of AcCoA to lipids in each strain is determined largely by the HPLC and GC-FID data (see Table 6-2). Conversely, the increase in flux through the oxidative PPP in MTYL065 reflects differences in the metabolite ^{13}C -labeling patterns of the

two strains (see Figure 6-7). Thus the conclusion that the oxidative PPP and fatty acid biosynthesis fluxes are correlated in the two *Y. lipolytica* strains investigated here comes directly from the HPLC, GC-FID, and ¹³C-labeling data and is not dependent on any modeling assumptions about NADPH production and consumption.

The utilization of the oxidative PPP to supply NADPH for fatty acid biosynthesis may not be unique to *Y. lipolytica* among oleaginous yeasts; upregulation of 6-phosphogluconate dehydrogenase during lipid accumulation has previously been observed in proteomic studies of *Rhodospiridium toruloides* and *Lipomyces starkeyi* (Liu et al., 2009, 2011). The heavy usage of the oxidative PPP for NADPH generation is suboptimal for production of lipids from glucose at high yields, as one mole of carbon is lost as carbon dioxide for every two moles of NADPH produced. A recent analysis showed that higher lipid yields could be achieved if NADPH were produced from cytosolic NADH generated in glycolysis (e.g. by a transhydrogenase cycle mediated by pyruvate carboxylase, malate dehydrogenase, and cytosolic NADP-dependent malic enzyme) or from mitochondrial NADH (e.g. by a transhydrogenase cycle consisting of mitochondrial NAD⁺-dependent IDH and cytosolic NADP⁺-dependent IDH operating in opposite directions) without net generation of carbon dioxide (Ratledge, 2014). In *S. cerevisiae*, cytosolic NADP⁺-dependent IDH and a citrate/ α -ketoglutarate antiporter apparently play an important role in supplying cytosolic NADPH (Castegna et al., 2010). However, *Y. lipolytica* possesses only a single NADP⁺-dependent IDH, which likely localizes to the mitochondria (X. Li et al., 2013). The malic enzyme transhydrogenase cycle is apparently intimately related to lipid overproduction in other oleaginous organisms (Wynn et al., 1997, 1999; Zhang et al., 2007), but not *Y. lipolytica* (Beopoulos et al., 2011; Zhang et al., 2013). The observed production of the

sugar alcohols mannitol and erythritol, which are formed by reduction of mannose and erythrose respectively, suggests that under the conditions investigated here the cell does indeed have an excess supply of cytosolic NADH which must be re-oxidized to NAD^+ and could therefore be used to generate NADPH equivalents. Engineering of pathways for production of NADPH without loss of carbon represents a promising avenue for future research.

In sum, the overexpression of *ACC1* and *DGAI* in MTYL065 drastically increases the rate of lipid synthesis in this strain, roughly doubling the demand for NADPH. To satisfy this increased demand, fluxes are redistributed at the G6P node such as to increase flux through the oxidative PPP. There does not appear to be any regulation preventing this redistribution; thus in *Y. lipolytica* G6P is likely a flexible branch point (Stephanopoulos and Vallino, 1991) at which the partitioning of flux readily changes to meet requirements for NADPH production. An increased rate of lipid production may feedback to the G6P node by increasing the rate of NADPH oxidation, resulting in a larger fraction of G6P being metabolized through the oxidative PPP to meet the increased demand for lipogenic NADPH. However, this mechanism is suboptimal from a yield perspective, and lipid yields could potentially benefit from increasing the supply of NADPH through mechanisms which do not involve net loss of carbon through decarboxylation reactions.

6.5 References

- Ageitos, J.M., Vallejo, J.A., Veiga-Crespo, P., Villa, T.G., 2011. Oily yeasts as oleaginous cell factories. *Appl. Microbiol. Biotechnol.* 90, 1219–1227. doi:10.1007/s00253-011-3200-z
- Ahn, W.S., Antoniewicz, M.R., 2013. Parallel labeling experiments with [1,2-(13)C]glucose and [U-(13)C]glutamine provide new insights into CHO cell metabolism. *Metab. Eng.* 15, 34–47. doi:10.1016/j.ymben.2012.10.001
- Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2006. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* 8, 324–337. doi:10.1016/j.ymben.2006.01.004
- Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007a. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Anal. Chem.* 79, 7554–7559. doi:10.1021/ac0708893
- Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2007b. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng.* 9, 68–86. doi:10.1016/j.ymben.2006.09.001
- Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2011. Measuring deuterium enrichment of glucose hydrogen atoms by gas chromatography/mass spectrometry. *Anal. Chem.* 83, 3211–3216. doi:10.1021/ac200012p
- Bailey, J.E., 1991. Toward a science of metabolic engineering. *Science* 252, 1668–1675.

Beopoulos, A., Cescut, J., Haddouche, R., Uribe Larrea, J.-L., Molina-Jouve, C., Nicaud, J.-M., 2009. *Yarrowia lipolytica* as a model for bio-oil production. *Prog. Lipid Res.* 48, 375–387. doi:10.1016/j.plipres.2009.08.005

Beopoulos, A., Nicaud, J.-M., Gaillardin, C., 2011. An overview of lipid metabolism in yeasts and its impact on biotechnological processes. *Appl. Microbiol. Biotechnol.* 90, 1193–1206. doi:10.1007/s00253-011-3212-8

Blank, L.M., Lehmbeck, F., Sauer, U., 2005. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Res.* 5, 545–558. doi:10.1016/j.femsyr.2004.09.008

Botham, P.A., Ratledge, C., 1979. A biochemical explanation for lipid accumulation in *Candida* 107 and other oleaginous micro-organisms. *J. Gen. Microbiol.* 114, 361–375.

Boulton, C.A., Ratledge, C., 1981. Correlation of Lipid Accumulation in Yeasts with Possession of ATP: Citrate Lyase. *Microbiology* 127, 169–176. doi:10.1099/00221287-127-1-169

Boulton, C.A., Ratledge, C., 1983. Use of Transition Studies in Continuous Cultures of *Lipomyces starkeyi*, an oleaginous yeast, to Investigate the Physiology of Lipid Accumulation. *Microbiology* 129, 2871–2876. doi:10.1099/00221287-129-9-2871

Canelas, A.B., Ras, C., Pierick, A., Dam, J.C., Heijnen, J.J., Gulik, W.M., 2008. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics* 4, 226–239. doi:10.1007/s11306-008-0116-4

Canelas, A.B., ten Pierick, A., Ras, C., Seifar, R.M., van Dam, J.C., van Gulik, W.M., Heijnen, J.J., 2009. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. *Anal. Chem.* 81, 7379–7389. doi:10.1021/ac900999t

Castegna, A., Scarcia, P., Agrimi, G., Palmieri, L., Rottensteiner, H., Spera, I., Germinario, L., Palmieri, F., 2010. Identification and functional characterization of a novel mitochondrial carrier for citrate and oxoglutarate in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285, 17359–17370.

doi:10.1074/jbc.M109.097188

Crown, S.B., Antoniewicz, M.R., 2013. Parallel labeling experiments and metabolic flux analysis: Past, present and future methodologies. *Metab. Eng.* 16, 21–32.

doi:10.1016/j.ymben.2012.11.010

Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., De Montigny, J., Marck, C., Neuvéglise, C., Talla, E., Goffard, N., Frangeul, L., Aigle, M., Anthouard, V., Babour, A., Barbe, V., Barnay, S., Blanchin, S., Beckerich, J.-M., Beyne, E., Bleykasten, C., Boisramé, A., Boyer, J., Cattolico, L., Confanioleri, F., De Daruvar, A., Despons, L., Fabre, E., Fairhead, C., Ferry-Dumazet, H., Groppi, A., Hantraye, F., Hennequin, C., Jauniaux, N., Joyet, P., Kachouri, R., Kerrest, A., Koszul, R., Lemaire, M., Lesur, I., Ma, L., Muller, H., Nicaud, J.-M., Nikolski, M., Oztas, S., Ozier-Kalogeropoulos, O., Pellenz, S., Potier, S., Richard, G.-F., Straub, M.-L., Suleau, A., Swennen, D., Tekaia, F., Wésolowski-Louvel, M., Westhof, E., Wirth, B., Zeniou-Meyer, M., Zivanovic, I., Bolotin-Fukuhara, M., Thierry, A., Bouchier, C., Caudron, B., Scarpelli, C., Gaillardin, C., Weissenbach, J., Wincker, P., Souciet, J.-L., 2004. Genome evolution in yeasts. *Nature* 430, 35–44. doi:10.1038/nature02579

Evans, C.T., Ratledge, C., 1985. Possible regulatory roles of ATP:citrate lyase, malic enzyme, and AMP deaminase in lipid accumulation by *Rhodospiridium toruloides* CBS 14. *Can. J. Microbiol.* 31, 1000–1005. doi:10.1139/m85-189

- Evans, C.T., Scragg, A.H., Ratledge, C., 1983a. A comparative study of citrate efflux from mitochondria of oleaginous and non-oleaginous yeasts. *Eur. J. Biochem. FEBS* 130, 195–204.
- Evans, C.T., Scragg, A.H., Ratledge, C., 1983b. Regulation of citrate efflux from mitochondria of oleaginous and non-oleaginous yeasts by adenine nucleotides. *Eur. J. Biochem. FEBS* 132, 609–615.
- Flores, C.-L., Gancedo, C., 2005. *Yarrowia lipolytica* mutants devoid of pyruvate carboxylase activity show an unusual growth phenotype. *Eukaryot. Cell* 4, 356–364. doi:10.1128/EC.4.2.356-364.2005
- Gonzalez, B., François, J., Renaud, M., 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast Chichester Engl.* 13, 1347–1355. doi:10.1002/(SICI)1097-0061(199711)13:14<1347::AID-YEA176>3.0.CO;2-O
- Gui, M.M., Lee, K.T., Bhatia, S., 2008. Feasibility of edible oil vs. non-edible oil vs. waste edible oil as biodiesel feedstock. *Energy* 33, 1646–1653. doi:10.1016/j.energy.2008.06.002
- Hill, J., Nelson, E., Tilman, D., Polasky, S., Tiffany, D., 2006. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc. Natl. Acad. Sci. U. S. A.* 103, 11206–11210. doi:10.1073/pnas.0604600103
- Jardón, R., Gancedo, C., Flores, C.-L., 2008. The gluconeogenic enzyme fructose-1,6-bisphosphatase is dispensable for growth of the yeast *Yarrowia lipolytica* in gluconeogenic substrates. *Eukaryot. Cell* 7, 1742–1749. doi:10.1128/EC.00169-08

- Kleijn, R.J., van Winden, W.A., van Gulik, W.M., Heijnen, J.J., 2005. Revisiting the ¹³C-label distribution of the non-oxidative branch of the pentose phosphate pathway based upon kinetic and genetic evidence. *FEBS J.* 272, 4970–4982. doi:10.1111/j.1742-4658.2005.04907.x
- Li, Q., Du, W., Liu, D., 2008. Perspectives of microbial oils for biodiesel production. *Appl. Microbiol. Biotechnol.* 80, 749–756. doi:10.1007/s00253-008-1625-9
- Li, X., Wang, P., Ge, Y., Wang, W., Abbas, A., Zhu, G., 2013. NADP(+)-specific isocitrate dehydrogenase from oleaginous yeast *Yarrowia lipolytica* CLIB122: biochemical characterization and coenzyme sites evaluation. *Appl. Biochem. Biotechnol.* 171, 403–416. doi:10.1007/s12010-013-0373-1
- Li, Z., Sun, H., Mo, X., Li, X., Xu, B., Tian, P., 2013. Overexpression of malic enzyme (ME) of *Mucor circinelloides* improved lipid accumulation in engineered *Rhodotorula glutinis*. *Appl. Microbiol. Biotechnol.* 97, 4927–4936. doi:10.1007/s00253-012-4571-5
- Liu, H., Zhao, X., Wang, F., Jiang, X., Zhang, S., Ye, M., Zhao, Z.K., Zou, H., 2011. The proteome analysis of oleaginous yeast *Lipomyces starkeyi*. *FEMS Yeast Res.* 11, 42–51. doi:10.1111/j.1567-1364.2010.00687.x
- Liu, H., Zhao, X., Wang, F., Li, Y., Jiang, X., Ye, M., Zhao, Z.K., Zou, H., 2009. Comparative proteomic analysis of *Rhodospiridium toruloides* during lipid accumulation. *Yeast Chichester Engl.* 26, 553–566. doi:10.1002/yea.1706
- Luévano-Martínez, L.A., Moyano, E., de Lacoba, M.G., Rial, E., Uribe-Carvajal, S., 2010. Identification of the mitochondrial carrier that provides *Yarrowia lipolytica* with a fatty acid-

induced and nucleotide-sensitive uncoupling protein-like activity. *Biochim. Biophys. Acta* 1797, 81–88. doi:10.1016/j.bbabi.2009.09.003

Luo, B., Groenke, K., Takors, R., Wandrey, C., Oldiges, M., 2007. Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry. *J. Chromatogr. A* 1147, 153–164. doi:10.1016/j.chroma.2007.02.034

Maaheimo, H., Fiaux, J., Cakar, Z.P., Bailey, J.E., Sauer, U., Szyperski, T., 2001. Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional (¹³C) labeling of common amino acids. *Eur. J. Biochem. FEBS* 268, 2464–2479.

Marchal, R., Vandecasteele, J.P., Metche, M., 1977. Regulation of the central metabolism in relation to citric acid production in *Saccharomyces lipolytica*. *Arch. Microbiol.* 113, 99–104.

Marx, A., de Graaf, A.A., Wiechert, W., Eggeling, L., Sahm, H., 1996. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 49, 111–129. doi:10.1002/(SICI)1097-0290(19960120)49:2<111::AID-BIT1>3.0.CO;2-T

Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., Xian, M., 2009. Biodiesel production from oleaginous microorganisms. *Renew. Energy* 34, 1–5. doi:10.1016/j.renene.2008.04.014

Mitsushima, K., Shinmyo, A., Enatsu, T., 1978. Control of citrate and 2-oxoglutarate formation in *Candida lipolytica* mitochondria by adenine nucleotides. *Biochim. Biophys. Acta* 538, 481–492.

Morin, N., Cescut, J., Beopoulos, A., Lelandais, G., Le Berre, V., Uribelarrea, J.-L., Molina-Jouve, C., Nicaud, J.-M., 2011. Transcriptomic analyses during the transition from biomass production to lipid accumulation in the oleaginous yeast *Yarrowia lipolytica*. *PLoS One* 6, e27966. doi:10.1371/journal.pone.0027966

Ochoa-Estopier, A., Guillouet, S.E., 2014. D-stat culture for studying the metabolic shifts from oxidative metabolism to lipid accumulation and citric acid production in *Yarrowia lipolytica*. *J. Biotechnol.* 170, 35–41. doi:10.1016/j.jbiotec.2013.11.008

Otto, C., Yovkova, V., Aurich, A., Mauersberger, S., Barth, G., 2012. Variation of the by-product spectrum during α -ketoglutaric acid production from raw glycerol by overexpression of fumarase and pyruvate carboxylase genes in *Yarrowia lipolytica*. *Appl. Microbiol. Biotechnol.* 95, 905–917. doi:10.1007/s00253-012-4085-1

Palmieri, L., Palmieri, F., Runswick, M.J., Walker, J.E., 1996. Identification by bacterial expression and functional reconstitution of the yeast genomic sequence encoding the mitochondrial dicarboxylate carrier protein. *FEBS Lett.* 399, 299–302.

Palmieri, L., Runswick, M.J., Fiermonte, G., Walker, J.E., Palmieri, F., 2000. Yeast mitochondrial carriers: bacterial expression, biochemical identification and metabolic significance. *J. Bioenerg. Biomembr.* 32, 67–77.

Palmieri, L., Voza, A., Agrimi, G., De Marco, V., Runswick, M.J., Palmieri, F., Walker, J.E., 1999. Identification of the yeast mitochondrial transporter for oxaloacetate and sulfate. *J. Biol. Chem.* 274, 22184–22190.

Ratledge, C., 2002. Regulation of lipid accumulation in oleaginous micro-organisms. *Biochem. Soc. Trans.* 30, 1047–1050. doi:10.1042/

Ratledge, C., 2014. The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: a reappraisal and unsolved problems. *Biotechnol. Lett.* doi:10.1007/s10529-014-1532-3

Schmidt, Marx, de Graaf AA, Wiechert, Sahm, Nielsen, Villadsen, 1998. ^{13}C tracer experiments and metabolite balancing for metabolic flux analysis: comparing two approaches. *Biotechnol. Bioeng.* 58, 254–257.

Stephanopoulos, G., Vallino, J.J., 1991. Network rigidity and metabolic engineering in metabolite overproduction. *Science* 252, 1675–1681.

Tai, M., Stephanopoulos, G., 2013. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metab. Eng.* 15, 1–9. doi:10.1016/j.ymben.2012.08.007

Van Winden, W.A., van Dam, J.C., Ras, C., Kleijn, R.J., Vinke, J.L., van Gulik, W.M., Heijnen, J.J., 2005. Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of (^{13}C) -labeled primary metabolites. *FEMS Yeast Res.* 5, 559–568. doi:10.1016/j.femsyr.2004.10.007

Van Winden, W.A., Wittmann, C., Heinzle, E., Heijnen, J.J., 2002. Correcting mass isotopomer distributions for naturally occurring isotopes. *Biotechnol. Bioeng.* 80, 477–479. doi:10.1002/bit.10393

- Wasylenko, T.M., Stephanopoulos, G., 2013. Kinetic isotope effects significantly influence intracellular metabolite (¹³C) labeling patterns and flux determination. *Biotechnol. J.* 8, 1080–1089. doi:10.1002/biot.201200276
- Wasylenko, T.M., Stephanopoulos, G., 2015. Metabolomic and (¹³C)-metabolic flux analysis of a xylose-consuming *Saccharomyces cerevisiae* strain expressing xylose isomerase. *Biotechnol. Bioeng.* 112, 470–483. doi:10.1002/bit.25447
- Wiechert, W., 2001. ¹³C Metabolic Flux Analysis. *Metab. Eng.* 3, 195–206. doi:10.1006/mben.2001.0187
- Wiechert, W., de Graaf, A.A., 1997. Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments. *Biotechnol. Bioeng.* 55, 101–117. doi:10.1002/(SICI)1097-0290(19970705)55:1<101::AID-BIT12>3.0.CO;2-P
- Wittmann, Heinzle, 1999. Mass spectrometry for metabolic flux analysis. *Biotechnol. Bioeng.* 62, 739–750.
- Wynn, J.P., bin Abdul Hamid, A., Ratledge, C., 1999. The role of malic enzyme in the regulation of lipid accumulation in filamentous fungi. *Microbiol. Read. Engl.* 145 (Pt 8), 1911–1917.
- Wynn, J.P., Hamid, A.A., Li, Y., Ratledge, C., 2001. Biochemical events leading to the diversion of carbon into storage lipids in the oleaginous fungi *Mucor circinelloides* and *Mortierella alpina*. *Microbiol. Read. Engl.* 147, 2857–2864.
- Wynn, J.P., Kendrick, A., Ratledge, C., 1997. Sesamol as an inhibitor of growth and lipid metabolism in *Mucor circinelloides* via its action on malic enzyme. *Lipids* 32, 605–610.

Young, J.D., Shastri, A.A., Stephanopoulos, G., Morgan, J.A., 2011. Mapping photoautotrophic metabolism with isotopically nonstationary (13)C flux analysis. *Metab. Eng.* 13, 656–665.

doi:10.1016/j.ymben.2011.08.002

Zhang, H., Zhang, L., Chen, H., Chen, Y.Q., Ratledge, C., Song, Y., Chen, W., 2013. Regulatory properties of malic enzyme in the oleaginous yeast, *Yarrowia lipolytica*, and its non-involvement in lipid accumulation. *Biotechnol. Lett.* 35, 2091–2098. doi:10.1007/s10529-013-1302-7

Zhang, Y., Adams, I.P., Ratledge, C., 2007. Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiol. Read. Engl.* 153, 2013–2025. doi:10.1099/mic.0.2006/002683-

0

Chapter 7

¹³C-Metabolic Flux Analysis and Flux Balance Analysis of Lipid Production from Acetate in *Yarrowia lipolytica*

7.1 Introduction

The last chapter focused on overproduction of lipids from glucose. However, a major advantage of *Yarrowia lipolytica* as a biocatalyst is its ability to utilize numerous diverse substrates (Fickers et al., 2005; Rodrigues and Pais, 2000), many of which are far less costly than glucose. Specifically, *Y. lipolytica* can use a wide range of weak carboxylic acids as sole carbon and energy source (Rodrigues and Pais, 2000). Lipid accumulation in *Y. lipolytica* from the volatile fatty acid (VFA) carbon sources acetic, propionic, and butyric acids has been demonstrated, although the best results were obtained when biomass was grown up on glucose or glycerol prior to conversion of VFAs to lipids (Fontanille et al., 2012). VFA mixtures and acetic acid have also been investigated as carbon sources for production of lipids in the oleaginous yeasts *Cryptococcus albidus* (Fei et al., 2011a, 2011b) and *Cryptococcus curvatus* (Christophe et al., 2012).

The conversion of VFAs to lipids is of interest for biodiesel applications because VFAs can be produced by fermentation of the organic fractions of municipal solid waste and sewage sludge (Sans et al., 1995a, 1995b; Morgan-Sagastume et al., 2011), and acetate can be produced

efficiently by syngas fermentation (Hu et al., 2013). These feed stocks are cheap (in some cases possibly even negative cost) relative to glucose, and consequently engineering of *Y. lipolytica* strains capable of efficiently producing lipids from VFAs in general and from acetate in particular could make possible the cost-effective production of biodiesel. In this chapter we focus on acetate as a model VFA, as its metabolism is the simplest and it is also likely to be available in the largest quantities, making it the most industrially relevant.

While acetate is more cheaply available than glucose, it is also a more challenging substrate for production of lipids at high yield, titer, and productivity. Every mole of acetate metabolized must first be activated to acetyl-CoA (AcCoA), which requires two ATP equivalents. (The enzyme acetyl-CoA synthetase which catalyzes this reaction hydrolyzes ATP to AMP and pyrophosphate.) Moreover, while metabolism of glucose through the glycolytic pathway generates ATP and NADH, production of triacylglycerols (TAGs) from acetate requires gluconeogenic flux in order to generate the glucose-6-phosphate (G6P) substrate for the NADPH-producing reactions of the oxidative Pentose Phosphate Pathway (PPP) as well as glycerol-3-phosphate (Glyc3P) for production of TAG backbones. This gluconeogenic flux consumes ATP and NADH, which must be regenerated primarily through the Tricarboxylic Acid (TCA) Cycle and Oxidative Phosphorylation.

We have performed ^{13}C -Metabolic Flux Analysis (MFA) to characterize the metabolism of the lipid-overproducing *Y. lipolytica* strain MTYL065 (Tai and Stephanopoulos, 2013) during production of lipids from acetate. We have also performed Flux Balance Analysis (FBA) to investigate the maximum theoretical yield of lipids from acetate given different assumptions

about the source of NADPH for fatty acid biosynthesis. The differences between the experimentally measured and theoretically optimal flux distributions are discussed.

7.2 Materials and Methods

7.2.1 Strain and Culture Conditions

The ^{13}C labeling experiment was conducted with the lipid-overproducing *Y. lipolytica* strain MTYL065 (Tai and Stephanopoulos, 2013). The strain was maintained at 4 °C on a Yeast Extract-Peptide-Acetate (YPA) plate, which was used to inoculate a 50 ml shake flask starter culture. The starter culture medium contained 27.5 g/l sodium acetate as sole carbon source, 5 g/l ammonium sulfate as sole nitrogen source, and 1.7 g/l Yeast Nitrogen Base (YNB) without amino acids and ammonium sulfate (Amresco, Solon, OH) as a source of salts, vitamins, and trace elements. The culture was incubated at 30 °C and shaken at 250 rpm.

The shake flask starter culture was used to inoculate a batch bioreactor culture with 200 ml working volume for ^{13}C -MFA. The bioreactor medium contained 1- ^{13}C -sodium acetate (Cambridge Isotope Laboratories, Tewksbury, MA) as sole carbon source, 0.44 g/l ammonium sulfate as sole nitrogen source, and 1.7 g/l YNB. The initial sodium acetate concentration was nominally 27.83 g/l (which is the same molar concentration as 27.5 g/l unlabeled sodium acetate); however HPLC analysis revealed that the true initial sodium acetate concentration was somewhat lower. The nominal initial carbon/nitrogen (C/N) ratio was 100.

For inoculation, 4 ml starter culture was washed with 10x YNB solution. The washed cell pellet was resuspended in 10x YNB and added to the bioreactor. The bioreactor culture was maintained at 30 °C with 0.75 vvm aeration and an agitator stirring speed of 850 rpm. The pH was maintained at 7.0 by addition of 6 M hydrochloric acid. 200 µl 50% (v/v) Antifoam B solution was added to the reactor every 24 h to prevent foaming.

7.2.2 Extracellular Metabolite Quantification

Extracellular acetate and citrate were quantified by High-Performance Liquid Chromatography (HPLC) using an Agilent 1260 Infinity HPLC system. Sample preparation, HPLC conditions, and ammonium quantification are described in Chapter 6.

7.2.3 Dry Cell Weight Measurements

Dry cell weight measurements were performed as in Chapter 6 with the exceptions that cells were filtered through 0.45 µm nitrocellulose filters (Millipore) and samples were dried at 70 °C for 12 h.

7.2.4 Analytical Methods

Lipid quantification, metabolite extractions, and GC-MS and LC-MS/MS measurement of metabolite labeling patterns were performed as in Chapter 6.

7.2.5 Estimation of Extracellular Fluxes

The estimation of extracellular fluxes was performed as in Chapter 6. All metabolic fluxes were normalized to an acetate uptake rate of 100. Fluxes for production of citrate and lipogenic

AcCoA, Glyc3P, and NADPH were calculated from yields computed using the Microsoft Excel function LINEST. Mannitol and erythritol were not detected in significant quantities. Because there was only a single culture, uncertainties were computed from the standard errors of the slopes calculated with the LINEST function, rather than from standard deviations of replicate cultures.

7.2.6 Metabolic Flux Estimation

Metabolic flux estimation was performed as in Chapter 6. The model metabolic network was modified slightly to reflect the differences between acetate and glucose metabolism. The gluconeogenic enzyme PEP carboxykinase and the glyoxylate shunt pathway were added to the model, as these reactions are necessary for synthesis of TAGs from acetate as sole carbon source. The addition of the glyoxylate shunt pathway necessitated the introduction of separate cytosolic and mitochondrial pools of succinate, and both pools were allowed to contribute to the ^{13}C -labeling data for succinate. ATP citrate lyase was omitted from the metabolic network since on acetate the cytosolic AcCoA for fatty acid biosynthesis is generated by acetyl-CoA synthetase directly. This cytosolic AcCoA was also assumed to be transported into the mitochondria (where it could serve as the substrate for citrate synthase) by the carnitine shuttle (Palmieri et al., 1999a), which was assumed to be reversible (as in Chapter 5).

Reactions for unidirectional transport of malate, pyruvate, and succinate from the cytosol to the mitochondria and of citrate from the mitochondria to the cytosol were included. Citrate and succinate transport are required for activity of the glyoxylate shunt. The pyruvate and malate transport reactions were identical to those in Chapter 6. Malate and succinate transport could be

mediated by the dicarboxylate carrier (Palmieri et al., 1996; Luévano-Martínez et al., 2010), which in *Saccharomyces cerevisiae* is essential for growth on acetate (Palmieri et al., 1999c). The dicarboxylate carrier would likely transport both malate and succinate in exchange for phosphate, with the proton gradient across the mitochondrial membrane as the driving force (Palmieri et al., 2000). The oxaloacetate carrier reaction (Palmieri et al., 1999b; Luévano-Martínez et al., 2010) was omitted from the model, as addition of this reaction had no significant effect on the results of the flux estimation (data not shown). Similarly, inclusion of a reversible reaction for transport of citrate across the mitochondrial membrane had no significant effect, and so the unidirectional transport reaction was deemed to be sufficient.

The complete model reaction network with atom transitions can be found in Table C-1 (Appendix C).

7.2.7 Flux Balance Analysis

Flux Balance Analysis (Orth et al., 2010) was performed to compute the optimal flux distributions for TAG overproduction with different pathways for NADPH production. The problem formulation was:

$$\begin{aligned}
 & \min_{\vec{v}} \quad -v_2 \\
 \text{s. t.} \quad & \mathbf{S} \cdot \vec{v} = \vec{\mathbf{0}} \\
 & v_1 = 100 \\
 & v_k \geq 0 \quad k \in IR
 \end{aligned}$$

where \bar{v} is the metabolic flux distribution, S is the stoichiometric matrix, v_1 is the acetate uptake flux, v_2 is the TAG production flux, and IR is the set of fluxes assumed to be irreversible. For simplicity, it was assumed that all fatty acids were oleate and that the stoichiometry of TAG biosynthesis was therefore:



where it has been assumed that 1 ATP (for acetyl-CoA carboxylase) and 2 NADPH are consumed in each fatty acid elongation step and 1 NADPH is consumed in each fatty acid desaturation step. The LP problem was solved using the Matlab function linprog. The list of reactions included in FBA model can be found in Table C-4 (Appendix C).

7.3 Results and Discussion

7.3.1 Fermentation Profiles and Metabolic Steady State

The lipid-overproducing strain MTYL065 was cultured in a bioreactor with pH control, as in shake flask cultures acetic acid consumption resulted in a dramatic rise in pH which inhibited growth and lipid accumulation (data not shown). The time courses for acetate and ammonium consumption and citrate, dry cell weight, and lipid accumulation are shown in Figures 7-1 and 7-2. The strain produced significant quantities of citrate, which was the only by-product detected. Fatty acid production was relatively poor, with the final lipid content reaching only 18% by the time cells were harvested for ^{13}C -MFA. The final fatty acid distribution (Figure 7-3) was similar

to that observed on glucose (Chapter 6), with oleate and palmitate being the dominant fatty acid species.

The ammonium supply had been exhausted by the 48 h time point. As was the case on glucose (see Chapter 6), the culture reached an apparent metabolic steady state with linear acetate consumption and citrate production profiles. The lipid production profile was also roughly linear, although there was significant variation in the data from the final four time points. ^{13}C -MFA was performed with data from samples collected between 84 and 108 h.

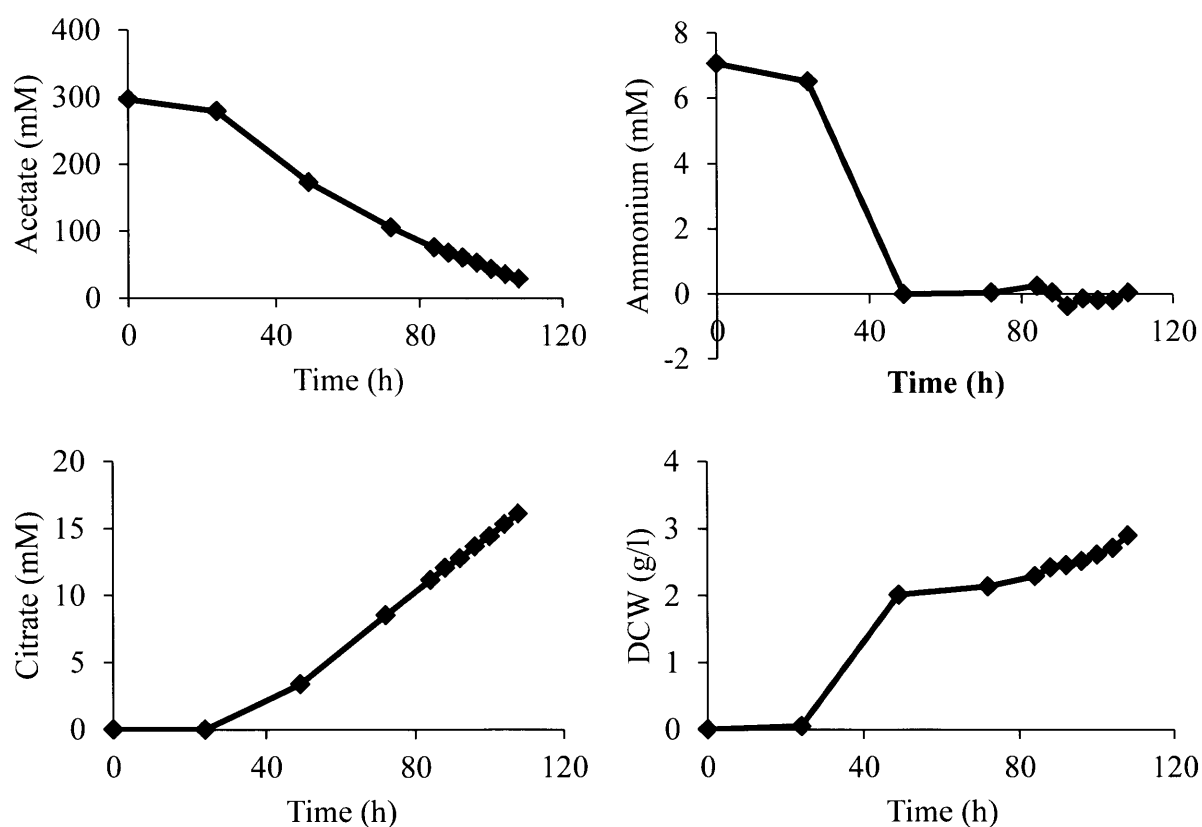


Figure 7-1 Extracellular metabolite fermentation profiles. Time courses for acetate and ammonium consumption and citrate and dry cell weight accumulation in ^{13}C labeling experiment.

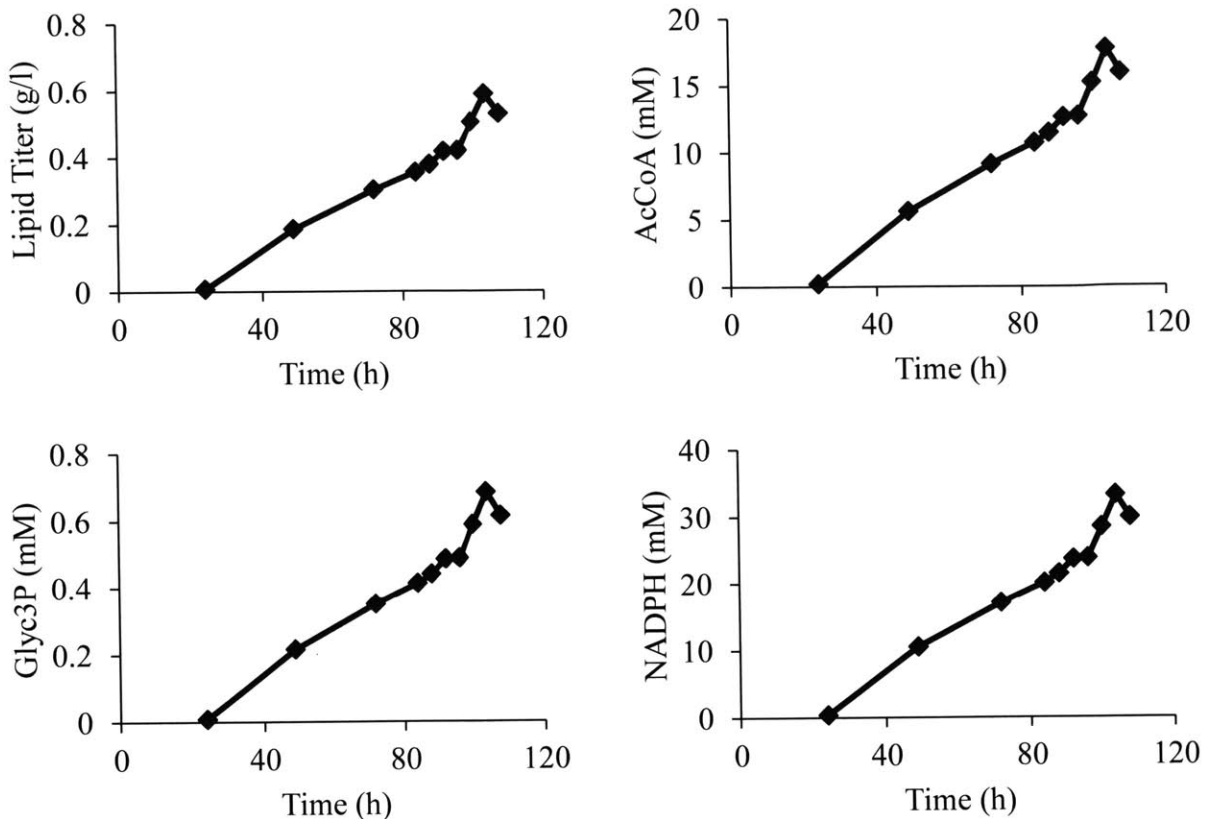


Figure 7-2 Fatty acid fermentation profiles. Time courses for total lipid titer and total lipogenic AcCoA, Glyc3P, and NADPH consumed for TAG biosynthesis in ^{13}C labeling experiment.

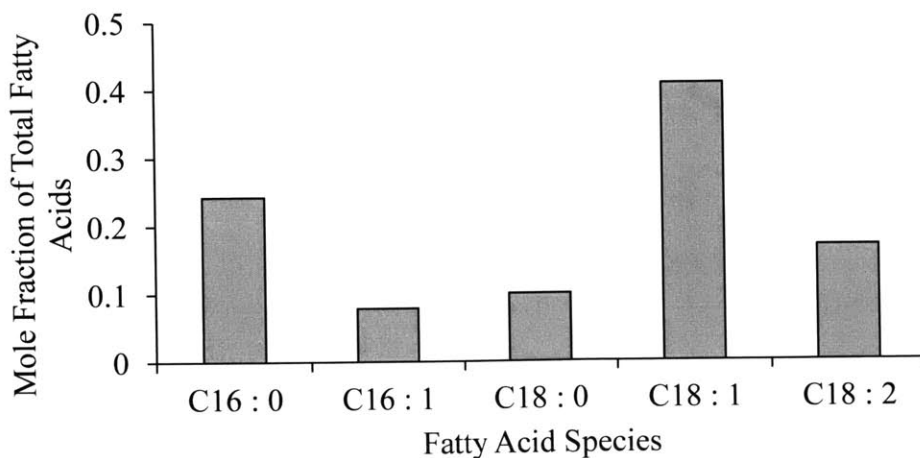


Figure 7-3 Final fatty acid distribution profile. The relative abundances of the five principal fatty acid species at the time cells were harvested for ^{13}C -MFA.

7.3.2 Estimation of Extracellular Fluxes

The acetate consumption flux was fixed to 100 and the other extracellular fluxes were calculated by using linear regression to compute the yields of citrate and lipogenic AcCoA, Glyc3P, and NADPH per 100 mol acetate, as in Chapter 6. The data used for the linear regressions and the resulting best-fit lines are shown in Figure C-1 (Appendix C). The extracellular flux values used for ¹³C-MFA are presented in Table 7-1.

Table 7-1 Extracellular flux values used for ¹³C-MFA

Acetate	100
Citrate	10.5 ± 0.2
AcCoA	14 ± 2
Glyc3P	0.55 ± 0.09
NADPH	26 ± 4

7.3.3 Metabolic Flux Estimation

¹³C-MFA was performed on strain MTYL065 during the lipid production phase on acetate. The estimated metabolic flux distribution is presented in Figure 7-4. The metabolic model described the labeling experiment data satisfactorily, with the weighted sum of squared residuals of 90.4 falling within the expected range of 63.9 to 115.8. The measured and simulated ¹³C labeling data from the flux estimation algorithm are presented in Table C-2 (Appendix C), and the full set of best-fit flux values and flux confidence intervals are tabulated in Table C-3 (Appendix C).

The lipid production in the labeling experiment was relatively poor—the lipid yield during the lipid production phase was only 0.08 g/g acetate. This is only about one-third of the maximum theoretical yield (see section on Flux Balance Analysis below). More than 30% of the carbon

from acetate was secreted as citrate. Only roughly 16% was incorporated into fatty acids and glycerol backbones for TAG production. The rest of the carbon (slightly more than 50%) was oxidized to CO₂, largely in reactions producing mitochondrial NADH (TCA Cycle, malic enzyme, pyruvate dehydrogenase). This mitochondrial NADH is expected to be used mostly for production of ATP through oxidative phosphorylation.

In order to determine whether a decrease in the amount of carbon oxidized to CO₂ in the mitochondria could result in improved lipid yield, an ATP balance was performed (Table 7-2). It was assumed that NADH and FADH₂ could be oxidized to generate 2.5 and 1.5 mol ATP, respectively, through oxidative phosphorylation. It can be seen that the rate of ATP production exceeds the rate of ATP consumption by approximately 243 mol per 100 mol acetate consumed. However this number is difficult to interpret, since the amount of ATP required for maintenance (e.g. macromolecule turnover, maintenance of ion gradients) is unknown. The largest ATP sink is the acetyl-CoA synthetase reaction, which converts acetate to AcCoA. The next largest sink is a futile cycle formed by pyruvate kinase, pyruvate carboxylase, and PEP carboxykinase. The ATP expenditure for fatty acid biosynthesis is relatively small.

Interestingly, satisfactory flux fits could not be obtained without significant flux through the futile cycle formed by pyruvate kinase, pyruvate carboxylase, and PEP carboxykinase. When pyruvate carboxylase was omitted from the metabolic model to “break” this cycle, the weighted sum of squared residuals increased to 165.4, which is significantly greater than the maximum value expected given that the model accurately describes the data (117.0). The flux estimation model predicts that roughly 49 mol ATP are consumed in this futile cycle. If the activity of this cycle is indeed so high, then the rates of mitochondrial NADH production and oxidative

phosphorylation may exceed those required for efficient production of lipids, and the excess ATP may be recycled to ADP through futile cycling. A flux distribution which generates more ATP than is necessary for fatty acid biosynthesis will be inefficient from the standpoint of lipid overproduction, as some carbon will be unnecessarily burned to CO₂, compromising lipid yield.

Table 7-2 ATP balance

Reaction	Flux	ATP produced	NADH produced	FADH ₂ produced
Acetyl-CoA synthetase	100	-200	0	0
AcCoA.c -> Lipids	15.58	-13.78	0	0
PEP carboxykinase	53.94	-53.94	0	0
Pyruvate kinase	49.27	49.27	0	0
Phosphoglycerate kinase	4.67	-4.67	0	0
GAP dehydrogenase	4.67	0	-4.67	0
Glycerol dehydrogenase	0.55	0	-0.55	0
Malic enzyme	6.45	0	6.45	0
Pyruvate carboxylase	49.27	-49.27	0	0
Malate dehydrogenase (cytosol)	4.67	0	4.67	0
Pyruvate dehydrogenase	6.45	0	6.45	0
Isocitrate dehydrogenase	37.16	0	37.16	0
AKG dehydrogenase/Succinyl-CoA synthetase	37.16	37.16	37.16	0
Succinate dehydrogenase	58.77	0	0	58.77
Malate dehydrogenase (mitochondria)	69.26	0	69.26	0
NADH -> NAD + 2.5 ATP (OXPHOS)	155.93	389.82	-155.93	0
FADH ₂ -> FAD + 1.5 ATP (OXPHOS)	58.77	88.15	0	-58.77
Sum		242.74		

As was the case on glucose (Chapter 6), the estimated rate of NADPH consumption in fatty acid biosynthesis is in good agreement with the estimated rate of NADPH production in the oxidative PPP. The rate of NADPH consumption is estimated to be 26 mol/100 mol acetate consumed (Table 7-1) while the rate of NADPH production is estimated to be 24.6 mol/100 mol acetate (Figure 7-4), and the difference between these two numbers is not statistically significant. Thus

the results suggest that the oxidative PPP may be the primary source of NADPH for lipid production on acetate, as was the case on glucose. However, as both the oxidative PPP and fatty acid biosynthesis fluxes were relatively small in the experiment described here, this finding should be confirmed with ^{13}C -MFA experiments under conditions which result in higher lipid yields.

7.3.4 Flux Balance Analysis

Because the lipid yield in the ^{13}C -labeling experiment was low, Flux Balance Analysis (FBA) was performed to determine the metabolic flux distribution which maximizes the yield of fatty acids on acetate. For simplicity, it was assumed that the all synthesized fatty acids would be oleate (C18:1), which was the most abundant fatty acid species in the labeling experiment (Figure 7-3). The optimal metabolic flux distribution is presented in Figure 7-5. The maximum lipid yield is determined to be 0.244 g oleate/g acetate, roughly threefold higher than the yield achieved in the labeling experiment. The flux through the oxidative PPP is increased to provide the NADPH for lipid biosynthesis. The TCA Cycle flux is decreased by about half compared to that estimated by ^{13}C -MFA, resulting in lower rates of production of mitochondrial NADH and ATP, and there is no futile cycling as the flux through pyruvate carboxylase is zero. The glyoxylate shunt functions primarily to provide PEP for gluconeogenesis, which produces glycerol backbones for TAG biosynthesis and G6P to be oxidized in the PPP for NADPH production. The citrate secretion flux is zero.

Figure 7-4 Estimated metabolic flux distribution in strain MTYL065 on acetate. Each reversible reaction is described by the net flux with the exchange flux listed inside parentheses. Exchange fluxes denoted “nr” could not be resolved within one order of magnitude.

In the original FBA model the oxidative PPP was the only source of cytosolic NADPH, as *Y. lipolytica* likely does not possess cytosolic, NADP-dependent malic enzyme or isocitrate dehydrogenase (Beopoulos et al., 2011; Zhang et al., 2013; Li et al., 2013). It has been shown that expression of either of these enzymes will increase the maximum theoretical lipid yield on glucose (Ratledge, 2014). Therefore the FBA was repeated with these reactions included in the metabolic model to observe the effect on the theoretical lipid yield on acetate.

When malic enzyme was added to the metabolic model the optimal metabolic flux distribution was identical to that in Figure 7-5, indicating that malic enzyme does not improve the maximum theoretical lipid yield. To investigate the flux distribution with malic enzyme as the source of lipogenic NADPH, the oxidative PPP flux was fixed to zero and the FBA repeated. The results are shown in Figure 7-6. The malate substrate for malic enzyme is produced from acetate through the glyoxylate shunt. The high glyoxylate shunt activity results in production of more ATP than is required for fatty acid biosynthesis. Consequently, the excess ATP is recycled to ADP in the pyruvate kinase-pyruvate carboxylase-PEP carboxykinase futile cycle discussed above. The maximum lipid yield is decreased to 0.177 g oleate/g acetate, suggesting that malic enzyme is a less efficient means of producing NADPH than the oxidative PPP on acetate.

In Figure 7-6, an excess of mitochondrial NADH is converted to ATP, which is then recycled to ADP through futile cycling. Lipid production with malic enzyme as the sole source of NADPH can be more efficient if excess mitochondrial NADH is used to generate NADPH through a transhydrogenase cycle. However, this requires shuttling of reducing equivalents from the mitochondria to the cytosol. To test the effect of such a transhydrogenase cycle, the transport reactions in the metabolic model were altered to allow oxaloacetate to be transported into the mitochondria and malate to be transported from the mitochondria to the cytosol. This creates a potential transhydrogenase cycle in which oxaloacetate is reduced to malate using mitochondrial NADH and the malate product transported to the cytosol, where it is oxidized by malic enzyme, generating cytosolic NADPH, pyruvate, and CO₂. The pyruvate can be carboxylated to form oxaloacetate, which can then be transported into the mitochondria, closing the cycle. The pyruvate carboxylase reaction consumes one mole of ATP. Inclusion of this transhydrogenase cycle resulted in the flux distribution shown in Figure 7-7. The maximum theoretical lipid yield increased to 0.232 g oleate/g acetate, which is still lower than that obtained with the base model. Thus even with the transhydrogenase cycle, malic enzyme appears to be a less efficient means of generating NADPH than the oxidative PPP.

Next the effect of including NADP-dependent cytosolic isocitrate dehydrogenase in the model was investigated. A citrate/ α -ketoglutarate (AKG) antiporter (Castegna et al., 2010) was also included to allow transport of the AKG product into the mitochondria. The results of the FBA are shown in Figure 7-8. The maximum theoretical yield increased to 0.258 g oleate/g acetate, an increase of roughly 6% over that obtained with the base model with the oxidative PPP as the sole source of NADPH. In the optimal flux distribution NADPH for fatty acid biosynthesis is

produced exclusively by NADP-dependent isocitrate dehydrogenase, with about half being produced in a transhydrogenase cycle consisting of mitochondrial NAD-dependent isocitrate dehydrogenase and cytosolic NADP-dependent isocitrate dehydrogenase operating in opposite directions. This cycle requires reductive flux through the mitochondrial enzyme. When this flux was constrained to be positive (in the oxidative direction), lipogenic NADPH was produced by a combination of the oxidative PPP and NADP-dependent isocitrate dehydrogenase, with the majority being produced by PPP (Figure 7-9). The maximum theoretical yield decreased to 0.248 g oleate/g acetate, which is only about 2% greater than the yield obtained with the base model.

The maximum theoretical yields obtained with the five different FBA models are summarized in Figure 7-10. The results show that expression of NADP-dependent isocitrate dehydrogenase could increase the maximum theoretical lipid yield on acetate if a transhydrogenase cycle with the mitochondrial isozyme can be established. However, malic enzyme is a less efficient means of NADPH production than the oxidative PPP, and consequently production of NADPH via malic enzyme is expected to decrease the lipid yield on acetate.

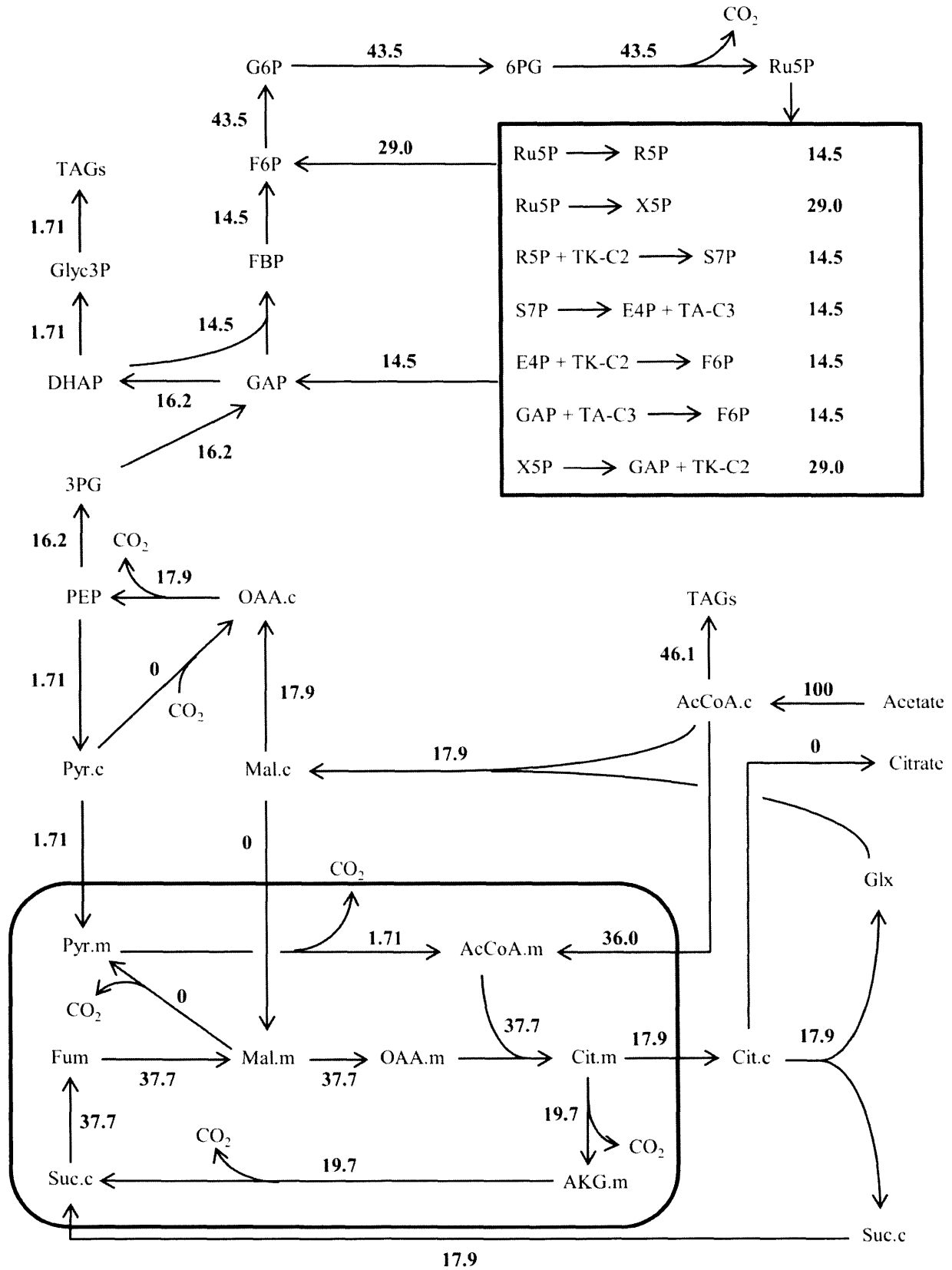


Figure 7-5 Results of Flux Balance Analysis with base model

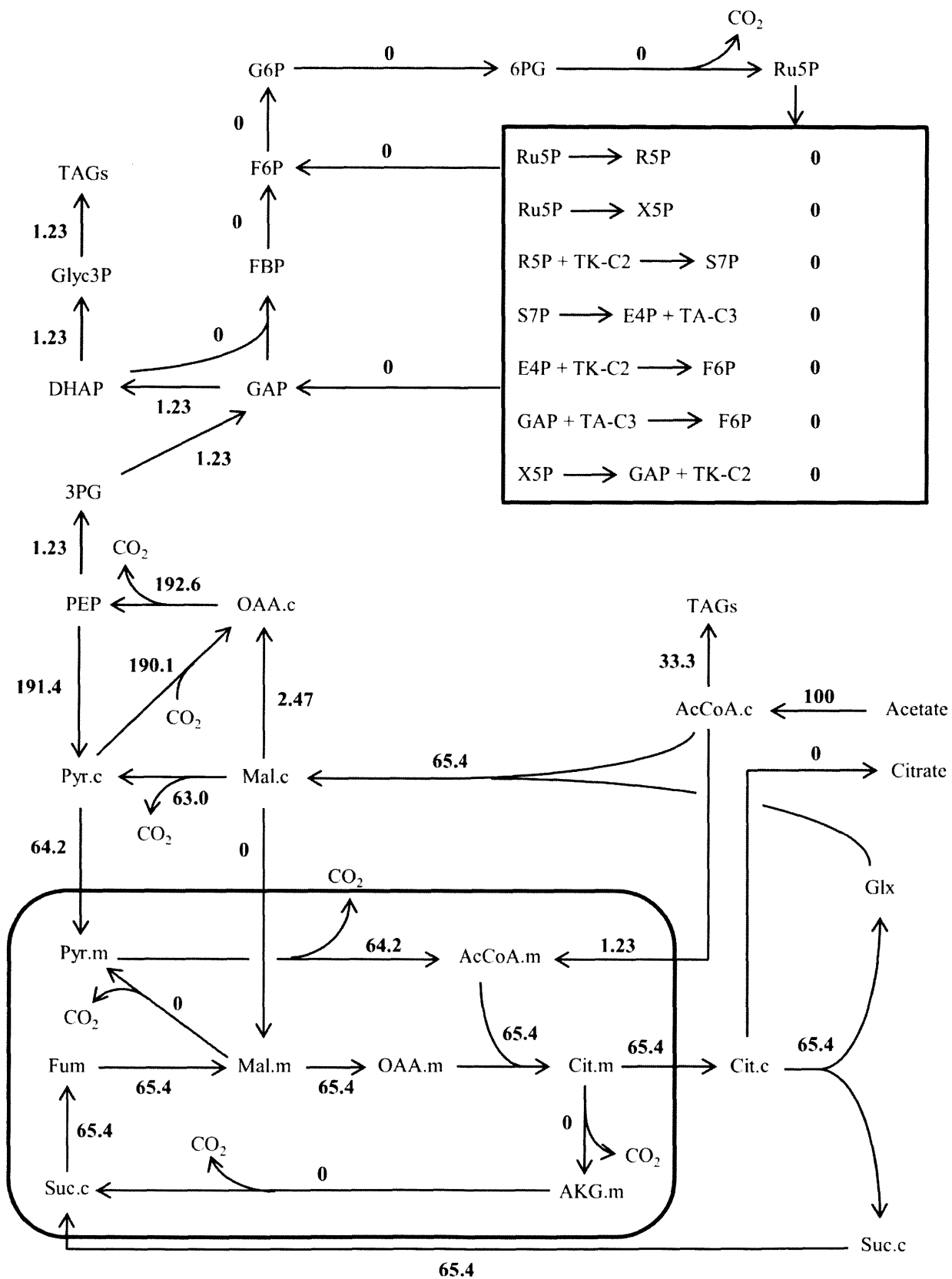


Figure 7-6 Results of Flux Balance Analysis with malic enzyme as only source of NADPH

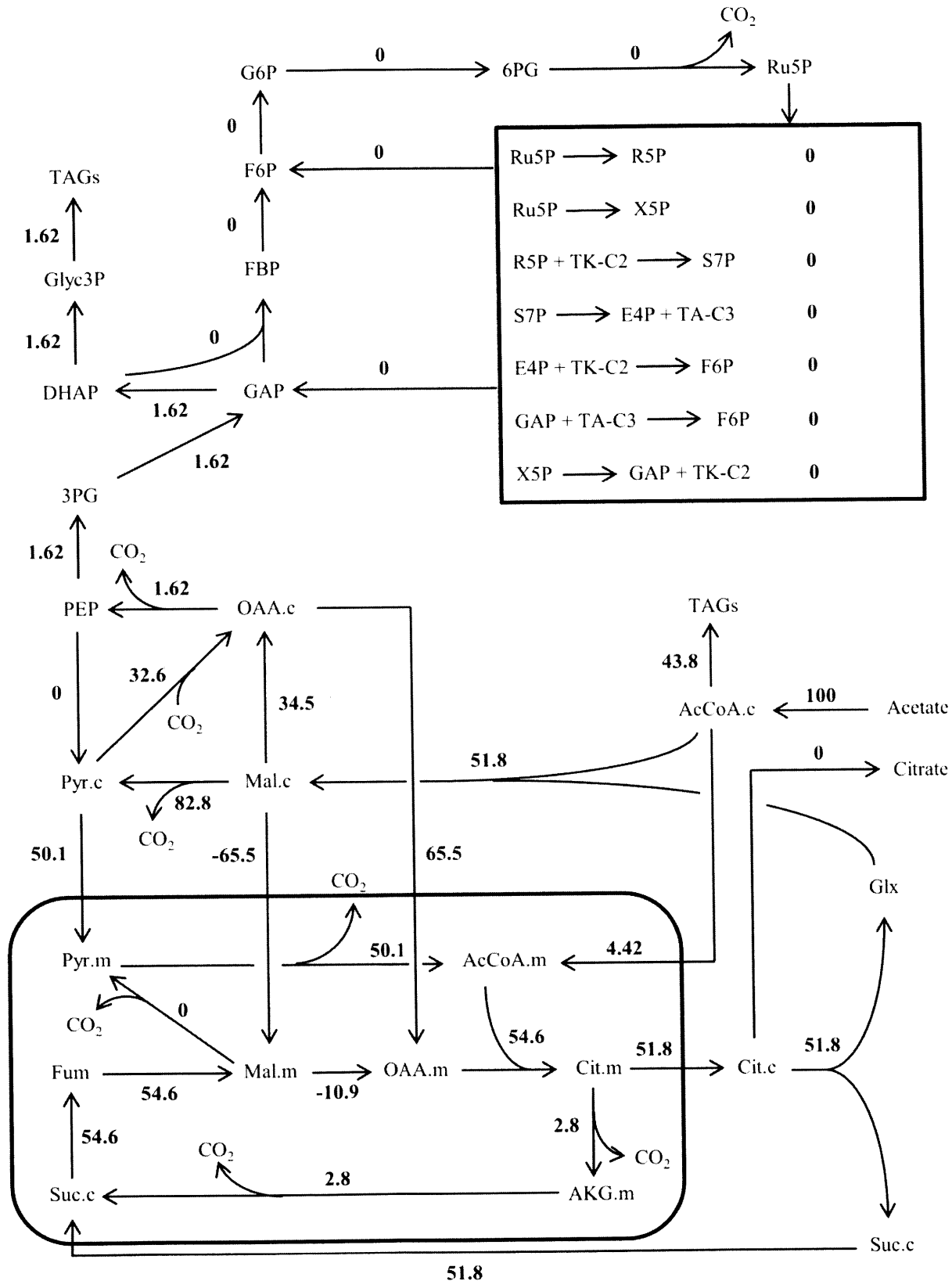


Figure 7-7 Results of Flux Balance Analysis with malic enzyme as only source of NADPH and malate/oxaloacetate shuttling

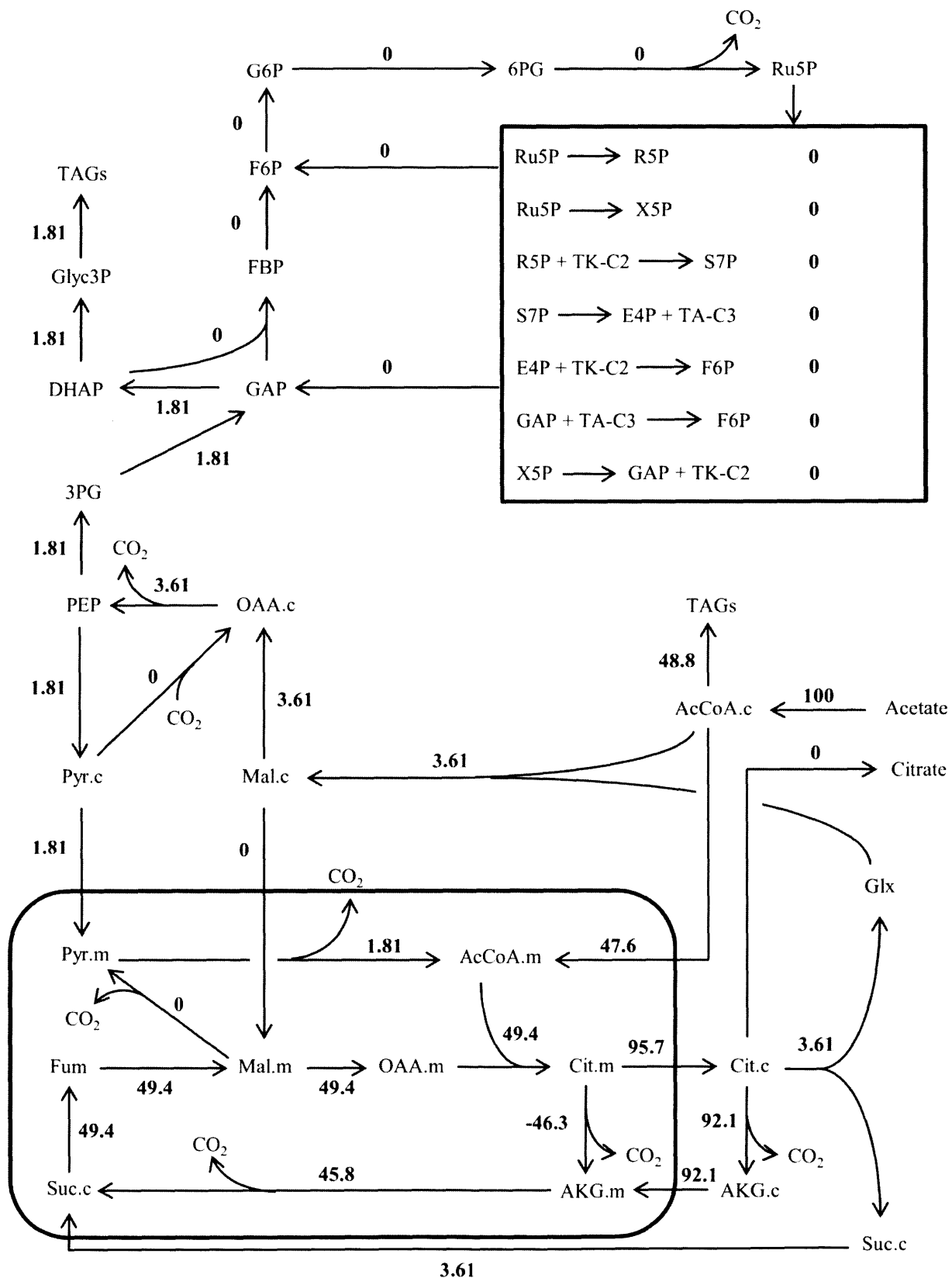


Figure 7-8 Results of Flux Balance Analysis with addition of cytosolic NADP-dependent isocitrate dehydrogenase

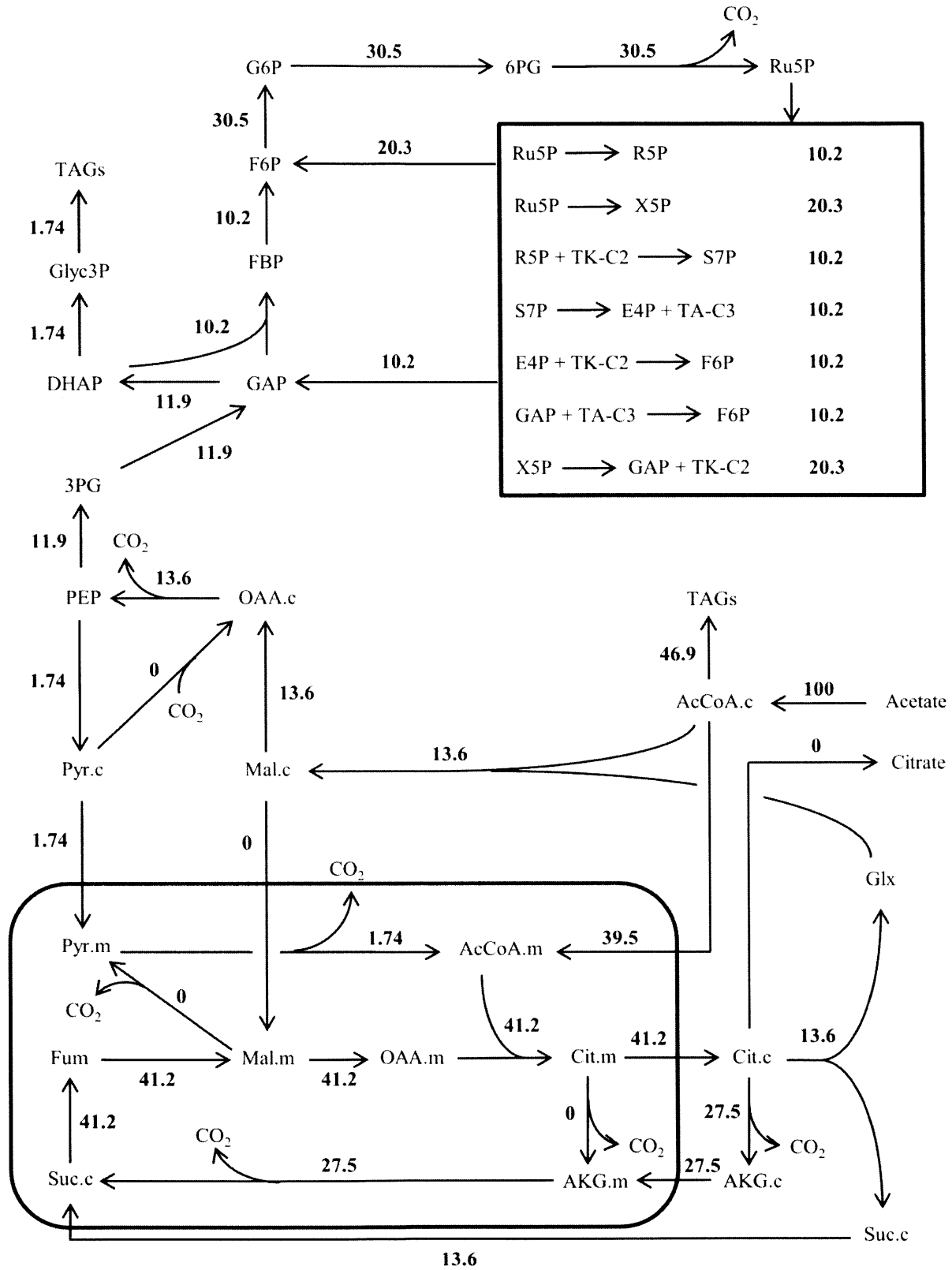


Figure 7-9 Results of Flux Balance Analysis with cytosolic NADP-dependent isocitrate dehydrogenase and irreversible mitochondrial NAD-dependent isocitrate dehydrogenase

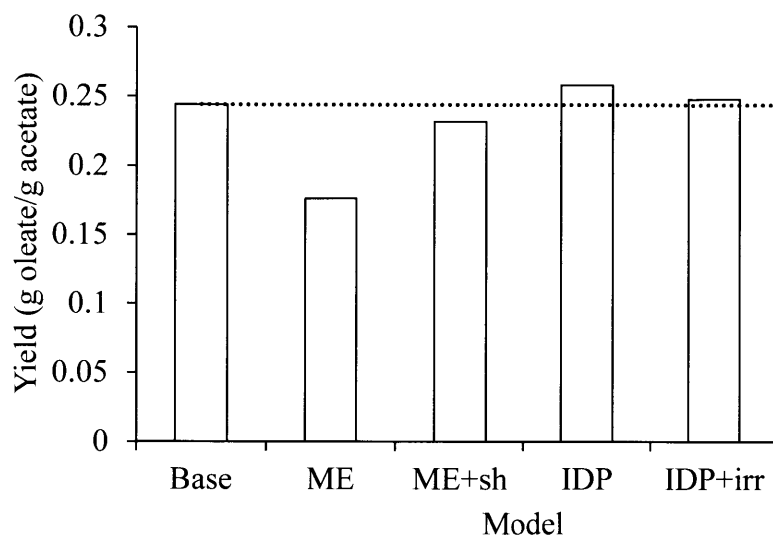


Figure 7-10 Maximum theoretical oleate yields as calculated by Flux Balance Analysis under different assumptions about NADPH sources. Abbreviations: Base, base model (PPP is sole source of NADPH); ME, model with NADP-dependent malic enzyme as sole source of NADPH; ME+sh, model with NADP-dependent malic enzyme as sole source of NADPH and malate/oxaloacetate shuttling; IDP, model including cytosolic NADP-dependent isocitrate dehydrogenase; IDP+irr, model including cytosolic NADP-dependent isocitrate dehydrogenase and irreversible mitochondrial NAD-dependent isocitrate dehydrogenase.

7.3.5 Stoichiometric Analysis of NADPH Production Pathways

The FBA predicts that NADP-dependent isocitrate dehydrogenase is the most efficient source of NADPH. The oxidative PPP is the next most efficient and cytosolic NADP-dependent malic enzyme is the least efficient source of NADPH considered here. These results can be confirmed by relatively simple calculations which are presented below.

When NADPH is formed through the oxidative PPP, two moles of AcCoA are first converted to a mole of oxaloacetate through the glyoxylate cycle. This pathway is initiated by the condensation of a mole of oxaloacetate with a mole of AcCoA, forming citrate (citrate synthase). Aconitase and isocitrate lyase catalyze the conversion of citrate to one mole of succinate and one mole of glyoxylate. The succinate can be oxidized in the TCA Cycle to regenerate the oxaloacetate needed to initiate the cycle. This results in production of one mole FADH₂ (succinate dehydrogenase) and one mole NADH (malate dehydrogenase). The glyoxylate can then condense with a second mole of AcCoA to form malate (malate synthase). The malate is converted to oxaloacetate by malate dehydrogenase, forming a second mole of NADH. Thus the net effect of the glyoxylate cycle is conversion of two moles AcCoA to oxaloacetate with the concomitant production of 2 moles NADH and 1 mole FADH₂.

To generate NADPH, the oxaloacetate is first converted to glyceraldehyde-3-phosphate through the reactions of gluconeogenesis. This results in consumption of two moles ATP (PEP carboxykinase and phosphoglycerate kinase) and one mole NADH (GAP dehydrogenase). One carbon from oxaloacetate is lost as CO₂ in the PEP carboxykinase reaction. The three carbons in glyceraldehyde-3-phosphate can then be oxidized to 3 CO₂ through gluconeogenesis and the PPP, generating 6 moles NADPH. Summing the effects of the two phases of the process (glyoxylate cycle and gluconeogenesis/PPP), we find that two moles AcCoA are converted to CO₂ resulting in consumption of two moles ATP and production of one mole NADH, one mole FADH₂, and six moles NADPH. If we assume oxidation of NADH and FADH₂ can be coupled to production of 2.5 and 1.5 moles ATP, respectively, through oxidative phosphorylation, then the net effect of the entire process is:

2 AcCoA converted to 4 CO₂, 2 ATP and 6 NADPH produced.

We now consider production of NADPH through NADP-dependent isocitrate dehydrogenase and cytosolic NADP-dependent malic enzyme. The most efficient means of production of NADPH through NADP-dependent isocitrate dehydrogenase is with a transhydrogenase cycle consisting of NAD- and NADP-dependent isozymes operating in opposite directions. In such a cycle, production of six moles NADPH necessitates consumption of six moles NADH. The six NADH can be regenerated by burning AcCoA in the TCA Cycle. Consumption of two moles AcCoA can be coupled to production of the six moles NADH (isocitrate dehydrogenase, α -ketoglutarate dehydrogenase, and malate dehydrogenase), two moles FADH₂ (succinate dehydrogenase), and two moles ATP (succinyl-CoA synthetase). If we assume the two moles FADH₂ are oxidized to form 3 moles ATP, then the net effect of this process is:

2 AcCoA converted to 4 CO₂, 5 ATP and 6 NADPH produced.

In the case of malic enzyme, NADPH is produced from NADH through a transhydrogenase cycle consisting of pyruvate carboxylase, malate dehydrogenase, and NADP-dependent malic enzyme. This cycle consumes 1 mole ATP per mole NADPH produced (in the pyruvate carboxylase reaction). Thus production of six moles NADPH requires consumption of six moles NADH and six moles ATP. As in the case of NADP-dependent isocitrate dehydrogenase, NADH can be regenerated by oxidation of two moles AcCoA in the TCA Cycle, which generates six moles NADH and five moles ATP (including 3 moles produced by oxidation of FADH₂). Thus when NADPH is generated by malic enzyme the net effect of the process is:

2 AcCoA converted to 4 CO₂, 6 NADPH produced, 1 ATP consumed.

These calculations make the reasons for the relative efficiencies of the three potential NADPH-producing pathways clear. In each case, 2 moles AcCoA are oxidized to 4 CO₂ and 6 NADPH are produced. The NADP-dependent isocitrate dehydrogenase process concomitantly produces 5 ATP while the oxidative PPP pathway process produces only 2 ATP. Thus production of NADPH through NADP-dependent isocitrate dehydrogenase results in higher theoretical lipid yield.

As discussed above, the transhydrogenase cycle including malic enzyme generally requires production of cytosolic NADPH from mitochondrial NADH. This requires some sort of electron shuttling mechanism (e.g. reduction of oxaloacetate to malate in the mitochondria and shuttling of malate to the cytosol, where it is the substrate for malic enzyme, see above). When no electron shuttling mechanism is present, the FBA model predicts an alternative process for production of NADPH via malic enzyme (Figure 7-6).

Two moles of AcCoA are converted to one mole of malate through the glyoxylate cycle. This generates one mole NADH (malate dehydrogenase) and one mole FADH₂ (succinate dehydrogenase). The malate is oxidized to pyruvate and CO₂ by malic enzyme, producing one mole NADPH. The pyruvate is then oxidized by pyruvate dehydrogenase, resulting in the loss of a second mole of carbon as CO₂ as well as formation of one mole AcCoA and one mole NADH. This regenerates one of the two moles of AcCoA which were used to synthesize malate in the

glyoxylate cycle. The two moles NADH and one mole FADH₂ produced can be oxidized to produce 6.5 moles ATP. Thus the net effect of this process is:

1 AcCoA converted to 2 CO₂, 6.5 ATP and 1 NADPH produced

Although this process generates 6.5 moles ATP, only 1 mole NADPH is produced for every 2 moles of carbon converted to CO₂, compared to 3 moles NADPH for every 2 moles carbon converted to CO₂ in the processes described above. Consequently, many moles of AcCoA must be burned to meet the NADPH demand for fatty acid biosynthesis, and the lipid yield is relatively low. The ATP generated through NADPH production is more than is needed for fatty acid biosynthesis, and the FBA model predicts a large flux through a futile cycle consisting of PEP carboxykinase, pyruvate kinase, and pyruvate carboxylase to consume the excess ATP.

7.4 Conclusions

We have conducted ¹³C-MFA of the lipid-overproducing *Y. lipolytica* strain MTYL065 during the lipid accumulation phase on acetate. Under the conditions investigated here, the lipid yield was relatively low. The rate of NADPH production through the oxidative PPP was in good agreement with the rate of NADPH consumption for fatty acid biosynthesis. Taken together with literature data suggesting that *Y. lipolytica* does not possess cytosolic NADP-dependent versions of either isocitrate dehydrogenase (Li et al., 2013) or malic enzyme (Beopoulos et al., 2011; Zhang et al., 2013), it seems likely that *Y. lipolytica* natively produces the majority of lipogenic NADPH through the oxidative PPP, as was the case on glucose (Chapter 6). However,

this result should be confirmed by performing ^{13}C -MFA under conditions which allow for higher lipid yield. The results also suggest that the rate of ATP production may exceed what is required for efficient lipid production, resulting in unnecessary loss of carbon as CO_2 . The ATP that is not used for lipid production or maintenance may be recycled to ADP through futile cycling.

FBA was conducted to explore potential strategies for improving the lipid yield by metabolic engineering, including by introduction of additional sources of NADPH. The results presented here suggest that expression of NADP-dependent isocitrate dehydrogenase (possibly in conjunction with a citrate/ α -ketoglutarate antiporter (Castegna et al., 2010)) could improve lipid yield. When NADPH is generated in the oxidative PPP, carbon in AcCoA must be metabolized to oxaloacetate via the glyoxylate shunt and then travel up the gluconeogenic pathway to produce G6P, which is oxidized to generate NADPH. Gluconeogenesis consumes NADH and ATP, which then become scarce on acetate since on glucose these cofactors are produced through glycolysis. Utilization of NADP-dependent isocitrate dehydrogenase for production of lipogenic NADPH would allow NADPH production without gluconeogenic flux, conserving NADH and ATP and improving yields. It is also possible that expression of NADP-dependent isocitrate dehydrogenase might increase the rates of NADPH and lipid production on acetate by alleviating any bottlenecks associated with the gluconeogenesis pathway. However, such predictions about rates cannot be tested using FBA and would have to be confirmed experimentally. Expression of NADP-dependent malic enzyme is not predicted to improve the theoretical lipid yield on acetate (in contrast to the situation on glucose (Ratledge, 2014)), and does not appear to be a promising strategy for fatty acid overproduction from acetate.

7.5 References

Beopoulos, A., Nicaud, J.-M., Gaillardin, C., 2011. An overview of lipid metabolism in yeasts and its impact on biotechnological processes. *Appl. Microbiol. Biotechnol.* 90, 1193–1206. doi:10.1007/s00253-011-3212-8

Castegna, A., Scarcia, P., Agrimi, G., Palmieri, L., Rottensteiner, H., Spera, I., Germinario, L., Palmieri, F., 2010. Identification and functional characterization of a novel mitochondrial carrier for citrate and oxoglutarate in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285, 17359–17370. doi:10.1074/jbc.M109.097188

Christophe, G., Deo, J.L., Kumar, V., Nouaille, R., Fontanille, P., Larroche, C., 2012. Production of Oils from Acetic Acid by the Oleaginous Yeast *Cryptococcus curvatus*. *Appl. Biochem. Biotechnol.* 167, 1270–1279. doi:10.1007/s12010-011-9507-5

Fei, Q., Chang, H.N., Shang, L., Choi, J., 2011a. Exploring low-cost carbon sources for microbial lipids production by fed-batch cultivation of *Cryptococcus albidus*. *Biotechnol. Bioprocess Eng.* 16, 482–487. doi:10.1007/s12257-010-0370-y

Fei, Q., Chang, H.N., Shang, L., Choi, J., Kim, N., Kang, J., 2011b. The effect of volatile fatty acids as a sole carbon source on lipid accumulation by *Cryptococcus albidus* for biodiesel production. *Bioresour. Technol.* 102, 2695–2701. doi:10.1016/j.biortech.2010.10.141

Fickers, P., Benetti, P.-H., Waché, Y., Marty, A., Mauersberger, S., Smit, M.S., Nicaud, J.-M., 2005. Hydrophobic substrate utilisation by the yeast *Yarrowia lipolytica*, and its potential applications. *FEMS Yeast Res.* 5, 527–543. doi:10.1016/j.femsyr.2004.09.004

Fontanille, P., Kumar, V., Christophe, G., Nouaille, R., Larroche, C., 2012. Bioconversion of volatile fatty acids into lipids by the oleaginous yeast *Yarrowia lipolytica*. *Bioresour. Technol.* 114, 443–449. doi:10.1016/j.biortech.2012.02.091

Hu, P., Rismani-Yazdi, H., Stephanopoulos, G., 2013. Anaerobic CO₂ fixation by the acetogenic bacterium *Moorella thermoacetica*. *AIChE J.* 59, 3176–3183. doi:10.1002/aic.14127

Li, X., Wang, P., Ge, Y., Wang, W., Abbas, A., Zhu, G., 2013. NADP(+)-specific isocitrate dehydrogenase from oleaginous yeast *Yarrowia lipolytica* CLIB122: biochemical characterization and coenzyme sites evaluation. *Appl. Biochem. Biotechnol.* 171, 403–416. doi:10.1007/s12010-013-0373-1

Luévano-Martínez, L.A., Moyano, E., de Lacoba, M.G., Rial, E., Uribe-Carvajal, S., 2010. Identification of the mitochondrial carrier that provides *Yarrowia lipolytica* with a fatty acid-induced and nucleotide-sensitive uncoupling protein-like activity. *Biochim. Biophys. Acta* 1797, 81–88. doi:10.1016/j.bbabi.2009.09.003

Morgan-Sagastume, F., Pratt, S., Karlsson, A., Cirne, D., Lant, P., Werker, A., 2011. Production of volatile fatty acids by fermentation of waste activated sludge pre-treated in full-scale thermal hydrolysis plants. *Bioresour. Technol.* 102, 3089–3097. doi:10.1016/j.biortech.2010.10.054

Orth, J.D., Thiele, I., Palsson, B.Ø., 2010. What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi:10.1038/nbt.1614

Palmieri, L., Lasorsa, F.M., Iacobazzi, V., Runswick, M.J., Palmieri, F., Walker, J.E., 1999a. Identification of the mitochondrial carnitine carrier in *Saccharomyces cerevisiae*. *FEBS Lett.* 462, 472–476.

Palmieri, L., Palmieri, F., Runswick, M.J., Walker, J.E., 1996. Identification by bacterial expression and functional reconstitution of the yeast genomic sequence encoding the mitochondrial dicarboxylate carrier protein. *FEBS Lett.* 399, 299–302.

Palmieri, L., Runswick, M.J., Fiermonte, G., Walker, J.E., Palmieri, F., 2000. Yeast mitochondrial carriers: bacterial expression, biochemical identification and metabolic significance. *J. Bioenerg. Biomembr.* 32, 67–77.

Palmieri, L., Vozza, A., Agrimi, G., De Marco, V., Runswick, M.J., Palmieri, F., Walker, J.E., 1999b. Identification of the yeast mitochondrial transporter for oxaloacetate and sulfate. *J. Biol. Chem.* 274, 22184–22190.

Palmieri, L., Vozza, A., Hönlinger, A., Dietmeier, K., Palmisano, A., Zara, V., Palmieri, F., 1999c. The mitochondrial dicarboxylate carrier is essential for the growth of *Saccharomyces cerevisiae* on ethanol or acetate as the sole carbon source. *Mol. Microbiol.* 31, 569–577.

Ratledge, C., 2014. The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: a reappraisal and unsolved problems. *Biotechnol. Lett.* doi:10.1007/s10529-014-1532-3

Rodrigues, G., Pais, C., 2000. The influence of acetic and other weak carboxylic acids on growth and cellular death of the yeast *Yarrowia lipolytica*. *Food Technol. Biotechnol.* 38, 27–32.

Sans, C., Mata-Alvarez, J., Cecchi, F., Pavan, P., Bassetti, A., 1995a. Volatile fatty acids production by mesophilic fermentation of mechanically-sorted urban organic wastes in a plug-flow reactor. *Bioresour. Technol.* 51, 89–96. doi:10.1016/0960-8524(95)95866-Z

Sans, C., Mata-Alvarez, J., Cecchi, F., Pavan, P., Bassetti, A., 1995b. Acidogenic fermentation of organic urban wastes in a plug-flow reactor under thermophilic conditions. *Bioresour. Technol.* 54, 105–110. doi:10.1016/0960-8524(95)00098-4

Tai, M., Stephanopoulos, G., 2013. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metab. Eng.* 15, 1–9. doi:10.1016/j.ymben.2012.08.007

Zhang, H., Zhang, L., Chen, H., Chen, Y.Q., Ratledge, C., Song, Y., Chen, W., 2013. Regulatory properties of malic enzyme in the oleaginous yeast, *Yarrowia lipolytica*, and its non-involvement in lipid accumulation. *Biotechnol. Lett.* 35, 2091–2098. doi:10.1007/s10529-013-1302-7

Chapter 8

Conclusions and Recommendations for Future Work

8.1 Conclusions

Protocols for metabolite extraction and mass isotopomer quantification have been optimized and applied for ^{13}C -Metabolic Flux Analysis (MFA) of two industrially relevant recombinant yeast strains. The optimized metabolite extraction protocol comprises three principal steps: quenching of cell metabolism in cold methanol solution, a washing step, and a hot ethanol metabolite extraction. The protocol results in rapid quenching of cell metabolism, as evidenced by measurements of the glucose-6-phosphate/fructose-6-phosphate ratio and adenylate energy charge, and high recovery of the majority of central carbon metabolites. A Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) method has been implemented for detection and quantification of metabolite mass isotopomers, allowing measurement of central carbon metabolite ^{13}C labeling patterns and pool sizes. The LC-MS/MS method was validated by measurement of Mass Isotopomer Distributions (MIDs) of metabolites labeled to natural abundance, for which MIDs can be predicted from theory.

An investigation of the errors associated with modeling the distribution of ^{13}C atoms through a metabolic network was also conducted, as these errors must be accurately known for rigorous statistical analysis of the results of metabolic flux estimation. A study of how MID mole fraction measurement errors depend on the values of the MID mole fractions was inconclusive, but simulations revealed that modeling errors which result from failure to include kinetic isotope

effects in ^{13}C -MFA models are expected to be similar in size to measurement errors associated with mass spectrometry.

Finally, ^{13}C -MFA was applied to two industrially relevant systems, xylose fermentation in *Saccharomyces cerevisiae* and lipid overproduction in *Yarrowia lipolytica*. These systems are important because they could lead to biofuel production processes superior to production of bioethanol from cornstarch or sucrose by allowing for use of less expensive, more widely available feed stocks (xylose fermentation, Chapter 5), production of superior fuel molecules (overproduction of fatty acids which may be transesterified to biodiesel, Chapter 6), or both (overproduction of fatty acids from volatile fatty acids such as acetate, Chapter 7).

The flux estimations were performed with compartmentalized metabolic models which contain separate cytosolic and mitochondrial pools of key metabolites which are involved in metabolism in both compartments. Such detailed models are essential for accurate modeling of metabolism in eukaryotic systems (Niklas et al., 2010). The ^{13}C -MFA results highlight the importance of redox cofactor balancing in metabolism. The results presented in Chapter 5 suggest anaerobic xylose metabolism is limited by re-oxidation of NADH formed in glycolysis, while the work with *Y. lipolytica* suggests that overproduction of lipids requires substantial rearrangements to the central carbon metabolic flux distribution for production of NADPH, which provides reducing power for lipid biosynthesis. Depending on the pathway utilized, production of NADPH may result in loss of carbon that could otherwise be used for lipid production.

8.2 Recommendations for Future Work

The investigation of the dependence of mass spectrometry measurement errors on mass isotopomer mole fraction was inconclusive. Although a theoretical model for propagation of error from ion counts to mass isotopomer mole fractions was validated *in silico*, the model was not successfully fit to experiment data. One problem was a large “gap” in the Gas Chromatography/Mass Spectrometry (GC-MS) MID data which resulted from the fact that MIDs were obtained from metabolites labeled to natural abundance, for which the M+0 mass isotopomer is dominant. Metabolites labeled to natural abundance were used because, to fit the error model to experiment data, it is necessary that the measured MIDs can be predicted from theory. Metabolites with known MIDs can also be obtained by culturing cells on one-carbon carbon sources, which lead to MIDs that can be predicted from the binomial distribution (Shen et al., 2009). These MIDs would more completely cover the mass isotopomer mole fraction space. However in this case errors associated with isotopic impurities in ^{13}C -labeled substrates and kinetic isotope effects (Chapter 4) may be comparable to or even larger than those associated with the mass spectrometry measurements, which would prevent fitting the measurement errors to the error model.

The results of ^{13}C -MFA of xylose-fermenting *S. cerevisiae* suggested that re-oxidation of cytosolic NADH may limit the rate of anaerobic xylose fermentation. Thus overexpression of enzymes in lower glycolysis or alcohol dehydrogenase might allow for xylose fermentation at rates comparable to those observed in glucose fermentation. Further insight into xylose metabolism in *S. cerevisiae* might also be gained by analysis of strains from intermediate points in the evolutionary engineering process. These strains currently exist in the Stephanopoulos Lab.

Such experiments would reveal the progression of the strain adaptation, from the parent strain incapable of xylose utilization to the final strain, which is the fast xylose-consuming strain published to date (Zhou et al., 2012).

^{13}C -MFA of lipid-overproducing *Y. lipolytica* revealed that the oxidative Pentose Phosphate Pathway is the primary source of lipogenic NADPH in this oleaginous yeast. However, high flux through this pathway results in loss of carbon as carbon dioxide, which is suboptimal from the standpoint of lipid yield. During lipid production from glucose, the rates of production of NADH and ATP in glycolysis greatly exceed those required for fatty acid biosynthesis, so expression of metabolic enzymes which create transhydrogenase cycles which use this extra ATP and NADH to produce NADPH without loss of carbon could improve lipid yields (Ratledge, 2014). Work along these lines has already been initiated in the Stephanopoulos Lab.

Lipid production was relatively poor in the *Y. lipolytica* ^{13}C -MFA experiment on acetate. This experiment could be repeated under conditions which allow better lipid production. The rate of aeration may be a key variable, as high aeration seems to be associated with high rates of citrate secretion, which compromise lipid yield. However, this effect needs to be investigated more rigorously.

Flux Balance Analysis showed that on acetate production of NADPH through a transhydrogenase cycle involving malic enzyme would be less efficient and result in lower theoretical yield than production of NADPH through the oxidative Pentose Phosphate Pathway, which is likely the pathway utilized by current lipid-overproducing strains. This is in contrast to

the situation on glucose; whereas on glucose ATP and NADH are necessarily produced in excess through glycolysis and can therefore be used to drive a transhydrogenase cycle for NADPH production, on acetate in the absence of glycolytic flux these cofactors are more scarce. However production of NADPH through isocitrate dehydrogenase would consume NADH but not ATP, and consequently expression of cytosolic NADP-dependent isocitrate dehydrogenase could increase the theoretical lipid yield on acetate. Expression of this pathway would also allow production of NADPH without flux through gluconeogenesis, possibly increasing the rates of NADPH and lipid production. Effective implementation of such a strategy may require expression of a citrate/ α -ketoglutarate antiporter (Castegna et al., 2010) or another transporter which allows transport of α -ketoglutarate from the cytosol to the mitochondria.

8.3 References

- Castegna, A., Scarcia, P., Agrimi, G., Palmieri, L., Rottensteiner, H., Spera, I., Germinario, L., Palmieri, F., 2010. Identification and functional characterization of a novel mitochondrial carrier for citrate and oxoglutarate in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285, 17359–17370. doi:10.1074/jbc.M109.097188
- Niklas, J., Schneider, K., Heinzle, E., 2010. Metabolic flux analysis in eukaryotes. *Curr. Opin. Biotechnol.* 21, 63–69. doi:10.1016/j.copbio.2010.01.011
- Ratledge, C., 2014. The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: a reappraisal and unsolved problems. *Biotechnol. Lett.* doi:10.1007/s10529-014-1532-3

Shen, T., Shen, W., Xiong, Y., Liu, H., Zheng, H., Zhou, H., Rui, B., Liu, J., Wu, J., Shi, Y.,
2009. Increasing the accuracy of mass isotopomer analysis through calibration curves
constructed using biologically synthesized compounds. *J. Mass Spectrom.* JMS 44, 1066–1080.
doi:10.1002/jms.1583

Zhou, H., Cheng, J.-S., Wang, B.L., Fink, G.R., Stephanopoulos, G., 2012. Xylose isomerase
overexpression along with engineering of the pentose phosphate pathway and evolutionary
engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*.
Metab. Eng. 14, 611–622. doi:10.1016/j.ymben.2012.07.011

Appendix A

Saccharomyces cerevisiae Xylose Flux Estimation Details

This Appendix contains the Model Metabolic Reaction Network used for ^{13}C -MFA of the xylose-consuming *S. cerevisiae* strain (Table A-1), measured and simulated ^{13}C labeling data from the flux estimations (Table A-2), the best-fit flux values and 68% and 95% confidence intervals for the estimated fluxes under each condition (Table A-3).

Table A-1: Model Metabolic Reaction Network for ^{13}C -MFA

Extracellular Fluxes

```
v1:  Glc.ext (abcdef) -> G6P (abcdef) OR
      Xylo.ext (abcde) -> X5P (abcde)
v2:  EtOH (ab) -> EtOH.ext (ab)
v3:  AcCOOH (ab) -> AcCOOH.ext (ab)
v4:  Glyc (abc) -> Glyc.ext (abc)
v5:  0.459 Ala + 0.161 Arg + 0.102 Asn + 0.297 Asp + 0.007 Cys +
      0.105 Gln + 0.302 Glu + 0.29 Gly + 0.066 His + 0.193 Ile +
      0.296 Leu + 0.286 Lys + 0.051 Met + 0.134 Phe + 0.165 Pro +
      0.185 Ser + 0.191 Thr + 0.028 Trp + 0.102 Tyr + 0.265 Val +
      0.519 Glycogen + 0.023 Trehalose + 0.808 Mannan + 1.135
      Glucan + 0.051 AMP + 0.05 CMP + 0.051 GMP + 0.067 UMP +
      0.0036 DAMP + 0.0024 DCMP + 0.0024 DGMP + 0.0036 DTMP +
      0.0066 TAG + 0.0007 ergosterol + 0.0015 zymosterol + 0.0006
      PA + 0.0062 PC + 0.0046 PE + 0.0055 PI + 0.0017 PS + 0.4623
      AcCoA.c -> Biomass
```

Pyruvate compartmentalization

```
v6:  0 Pyr.c (abc) -> Pyr.mnt (abc)
v7:  0 Pyr.m (abc) -> Pyr.mnt (abc)
v8:  Pyr.mnt (abc) -> Pyr.fix (abc)
```

Glycolysis

```
v9:  G6P (abcdef) <-> F6P (abcdef)
v10: F6P (abcdef) <-> FBP (abcdef)
v11: FBP (abcdef) <-> DHAP (cba) + G3P (def)
```

v12: DHAP (abc) <-> G3P (abc)
v13: G3P (abc) -> PG3 (abc)
v14: PG3 (abc) -> PEP (abc)
v15: PEP (abc) -> Pyr.c (abc)
v16: DHAP (abc) -> Glyc3P (abc)
v17: Glyc3P (abc) -> Glyc (abc)

Pentose Phosphate Pathway

v18: G6P (abcdef) -> CO2 (a) + Ru5P (bcdef)
v19: Ru5P (abcde) <-> R5P (abcde)
v20: Ru5P (abcde) <-> X5P (abcde)
v21: X5P (abcde) <-> G3P (cde) + TKC2 (ab)
v22: F6P (abcdef) <-> E4P (cdef) + TKC2 (ab)
v23: S7P (abcdefg) <-> R5P (cdefg) + TKC2 (ab)
v24: F6P (abcdef) <-> G3P (def) + TAC3 (abc)
v25: S7P (abcdefg) <-> E4P (defg) + TAC3 (abc)

Pyruvate metabolism

v26: Mal (abcd) -> Pyr.m (abc) + CO2 (d)
v27: Pyr.c (abc) + CO2 (d) -> OAA.c (abcd)
v28: Pyr.c (abc) -> AcCHO (bc) + CO2 (a)
v29: AcCHO (ab) -> EtOH (ab)
v30: AcCHO (ab) -> AcCOOH (ab)
v31: AcCHO (ab) -> AcCoA.c (ab)

Citric Acid Cycle

v32: Pyr.m (abc) -> AcCoA.m (bc) + CO2 (a)
v33: OAA.m (abcd) + AcCoA.m (ef) -> Cit (dcbfea)
v34: Cit (abcdef) -> ICit (abcdef)
v35: ICit (abcdef) -> AKG (abcde) + CO2 (f)
v36: AKG (abcde) -> Suc (bcde) + CO2 (a)
v37: Suc (abcd) <-> Fum (abcd)
v38: Fum (abcd) <-> Mal (abcd)
v39: Mal (abcd) <-> OAA.m (abcd)

One-carbon metabolism

v40: PG3 (abc) + Glu (defgh) -> Ser (abc) + AKG (defgh)
v41: Ser (abc) <-> Gly (ab) + C1 (c)
v42: Gly (ab) <-> CO2 (a) + C1 (b)

Transport across mitochondrial membrane

v43: Pyr.c (abc) -> Pyr.m (abc)
v44: OAA.c (abcd) <-> OAA.m (abcd)
v45: AcCoA.c (ab) <-> AcCoA.m (ab)

Amino Acid Biosynthesis

v46: AKG (abcde) -> Glu (abcde)

v47: Glu (abcde) -> Gln (abcde)
v48: Glu (abcde) -> Pro (abcde)
v49: Glu (abcde) + Glu (fghij) + CO2 (k) + Gln (lmnop) +
Asp (qrst) -> AKG (abcde) + Arg (fghijk) + Glu (lmnop) +
Fum (qrst)
v50: OAA.c (abcd) + Glu (efghi) -> Asp (abcd) + AKG (efghi)
v51: Asp (abcd) + Gln (efghi) -> Asn (abcd) + Glu (efghi)
v52: Pyr.c (abc) + Glu (defgh) -> Ala (abc) + AKG (defgh)
v53: Pyr.m (abc) + Glu (defgh) -> Ala (abc) + AKG (defgh)
v54: Asp (abcd) -> Thr (abcd)
v55: Thr (abcd) -> Gly (ab) + AcCHO (cd)
v56: Ser (abc) + AcCoA.c (de) + Asp (fghi) + CO2(j) -> Cys (abc)
+ AcCOOH (de) + Suc (jihj) + CO2 (f)
v57: Asp (abcd) + Cl (e) + AcCoA.c (fg) -> Met (abcde) +
AcCOOH (fg)
v58: AKG (abcde) + AcCoA.c (fg) + Glu (hijkl) + Glu (mnopq) ->
Lys (fgbcde) + CO2 (a) + AKG (hijkl) + AKG (mnopq)
v59: AcCoA.m (ab) + Pyr.m (cde) + Pyr.m (fgh) + Glu (ijklm) ->
Leu (abdghe) + CO2 (c) + CO2 (f) + AKG (ijklm)
v60: Thr (abcd) + Pyr.m (efg) + Glu (hijkl) -> Ile (abfcdg) +
CO2 (e) + AKG (hijkl)
v61: Pyr.m (abc) + Pyr.m (def) + Glu (ghijk) -> Val (abecf) +
CO2 (d) + AKG (ghijk)
v62: PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) ->
Phe (abceefghij) + CO2 (d) + AKG (klmno)
v63: PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) ->
Tyr (abceefghij) + CO2 (d) + AKG (klmno)
v64: Ser (abc) + R5P (defgh) + PEP (ijk) + E4P (lmno) +
PEP (pqr) + Gln (stuvw) -> Trp (abcdklmnoj) + CO2 (i) +
G3P (fgh) + Pyr.c (pqr) + Glu (stuvw)
v65: R5P (abcde) + Cl (f) + Gln (ghijk) + Asp (lmno) ->
His (edcbaf) + AKG (ghijk) + Fum (lmno)

Nucleotide Biosynthesis

v66: CO2 (b) + Asp (nfed) + R5P (vwxyz) + Gln (ijklm) ->
UMP (bdefvwxyz) + CO2 (n) + Glu (ijklm)
v67: UMP (bdefvwxyz) + Gln (ijklm) -> CMP (bdefvwxyz) +
Glu (ijklm)
v68: R5P (vwxyz) + Gln (ijklm) + Gln (nopqr) + Gly (de) +
Asp (stua) + Cl (b) + Cl (h) + CO2 (f) -> IMP (bdefhvwxyz)
+ Glu (ijklm) + Glu (nopqr) + Fum (stua)
v69: IMP (bdefhvwxyz) + Asp (ijkl) -> AMP (bdefhvwxyz) +
Fum (ijkl)
v70: IMP (bdefhvwxyz) + Gln (ijklm) -> GMP (bdefhvwxyz) +
Glu (ijklm)
v71: AMP (abcdefghij) -> DAMP (abcdefghij)
v72: CMP (abcdefghi) -> DCMP (abcdefghi)

v73: GMP (abcdefghij) -> DGMP (abcdefghij)
v74: UMP (abcdefghi) -> DUMP (abcdefghi)
v75: DUMP (abcdefghi) + C1 (j) -> DTMP (abcdjefghi)

Lipid Synthesis

v76: Glyc3P (abc) -> PA (abc)
v77: Glyc3P (def) + Ser (abc) -> PS (abcdef)
v78: Glyc3P (def) + Ser (abc) -> PE (cbdef) + CO2 (a)
v79: Glyc3P (def) + Ser (abc) + C1 (g) + C1 (h) + C1 (i) ->
PC (cbghidef) + CO2 (a)
v80: Glyc3P (ghi) + G6P (abcdef) -> PI (fedcbaghi)
v81: Glyc3P (abc) -> TAG (abc)
v82: AcCoA.c (ab) + AcCoA.c (cd) -> AcAcCoA (abcd)
v83: AcAcCoA (abcd) + AcCoA.c (ef) -> HMGCoA (abcfed)
v84: HMGCoA (abcdef) -> MVL (edcbaf)
v85: MVL (abcdef) -> IPPP (edcbf) + CO2 (a)
v86: IPPP (abcde) -> DMAPP (abcde)
v87: DMAPP (abcde) + IPPP (fghij) + IPPP + DMAPP + IPPP + IPPP
-> CO2 (d) + CO2 (e) + CO2 (j) + zymosterol
v88: zymosterol + C1 -> ergosterol

Carbohydrate Biosynthesis

v89: G6P (abcdef) + G6P (ghijkl) -> Trehalose (abcdefghijkl)
v90: G6P (abcdef) -> Glycogen (abcdef)
v91: G6P (abcdef) -> Glucan (abcdef)
v92: F6P (abcdef) -> Mannan (abcdef)

CO2 Transport

v93: CO2.ext (a) -> CO2 (a)
v94: CO2 (a) -> CO2.out (a)

Table A-2: Measured and Simulated Mass Spectrometry Data

Notation:

Fragment: the ion measured by the mass spectrometer. Each ion is denoted by a metabolite abbreviation followed by an underscore and the mass of the unlabeled M+0 isotopomer.

Mass Isotopomer: each mass isotopomer is denoted by its mass relative to the mass of the unlabeled parent ion M.

Measured: mass isotopomer mole fractions measured by mass spectrometry in ¹³C labeling experiments

Simulated: mass isotopomer mole fractions obtained by simulation of best-fit flux models

Glucose MS data:

GA = Glucose Aerobic

GN = Glucose Anaerobic

Fragment	Mass Isotopomer	GA		GN	
		Measured	Simulated	Measured	Simulated
Pyr_174	M+0	0.435700	0.440489	0.437800	0.439665
	M+1	0.077500	0.078069	0.076250	0.075720
	M+2	0.410800	0.404553	0.411650	0.409423
	M+3	0.055800	0.057006	0.054400	0.055359
	M+4	0.018600	0.018149	0.018550	0.018178
	M+5	0.001600	0.001589	0.001450	0.001513
	M+6				
	M+7				
Ala_260	M+0	0.379500	0.383341	0.380450	0.385124
	M+1	0.110150	0.111040	0.103250	0.106333
	M+2	0.378600	0.373718	0.385800	0.378687
	M+3	0.090950	0.091400	0.089950	0.089590
	M+4	0.034050	0.033955	0.034250	0.033908
	M+5	0.005700	0.005527	0.005500	0.005361
	M+6	0.000900	0.000915	0.000850	0.000897
	M+7	0.000150	0.000096		
Gly_218	M+0	0.420950	0.420638	0.412800	0.413401
	M+1	0.435050	0.434183	0.438850	0.439995
	M+2	0.104550	0.105125	0.105000	0.105932

	M+3	0.035050	0.034607	0.037500	0.035146
	M+4	0.004400	0.004653	0.005300	0.004718
	M+5			0.000550	0.000747
Gly_246	M+0	0.407800	0.406490	0.402500	0.401722
	M+1	0.424150	0.426663	0.431000	0.431817
	M+2	0.121700	0.121115	0.120100	0.120574
	M+3	0.038750	0.038335	0.039150	0.038552
	M+4	0.006600	0.006298	0.006350	0.006242
	M+5	0.000950	0.000990	0.000850	0.000987
	M+6				
Ser_288	M+0	0.325350	0.330941	0.316700	0.319234
	M+1	0.204500	0.205629	0.218250	0.219281
	M+2	0.348700	0.343426	0.344550	0.341527
	M+3	0.085100	0.085107	0.085150	0.085341
	M+4	0.030600	0.029475	0.029650	0.029239
	M+5	0.004750	0.004611	0.004950	0.004577
	M+6	0.000950	0.000734	0.000800	0.000726
Ser_302	M+0	0.387250	0.389402	0.381800	0.386523
	M+1	0.418800	0.425212	0.420150	0.428784
	M+2	0.134450	0.134167	0.134600	0.133530
	M+3	0.042550	0.042089	0.044050	0.042122
	M+4	0.012550	0.007733	0.014250	0.007656
	M+5	0.003100	0.001242	0.003550	0.001231
	M+6	0.001250	0.000143	0.001600	0.000140
Ser_362	M+0	0.293750	0.297626	0.282800	0.287098
	M+1	0.206000	0.206966	0.217100	0.218464
	M+2	0.339750	0.333523	0.338650	0.332336
	M+3	0.106750	0.106506	0.107350	0.107033
	M+4	0.044000	0.043761	0.044000	0.043512
	M+5	0.008100	0.009236	0.008400	0.009197
	M+6	0.001650	0.002040	0.001650	0.002023
Ser_390	M+0	0.289350	0.291735	0.279450	0.281922
	M+1	0.200850	0.201905	0.214250	0.215367
	M+2	0.331300	0.327720	0.330300	0.326125
	M+3	0.117800	0.117062	0.116500	0.116161
	M+4	0.047250	0.047484	0.046350	0.046690
	M+5	0.010850	0.011136	0.010700	0.010864
	M+6	0.002350	0.002489	0.002200	0.002419
	M+7	0.000350	0.000405	0.000300	0.000390
	M+ 8				
Ser_432	M+0	0.278400	0.282107	0.269850	0.272618
	M+1	0.205700	0.204760	0.219350	0.217457
	M+2	0.330600	0.323601	0.325400	0.322495

	M+3	0.122900	0.123967	0.124650	0.123049
	M+4	0.049550	0.049860	0.049700	0.049062
	M+5	0.011300	0.012362	0.010600	0.012073
	M+6	0.001200	0.002788	0.000450	0.002712
	M+7	0.000300	0.000477		
	M+ 8				
Asp_302	M+0	0.384600	0.388425	0.383200	0.387183
	M+1	0.425750	0.425582	0.431100	0.428415
	M+2	0.135200	0.134588	0.133100	0.133336
	M+3	0.043550	0.042222	0.042500	0.042049
	M+4	0.008800	0.007776	0.008150	0.007637
	M+5	0.001800	0.001249	0.001600	0.001227
	M+6	0.000300	0.000144	0.000300	0.000140
	M+7	0.000050	0.000012		
Asp_376	M+0	0.351600	0.349323	0.348550	0.348206
	M+1	0.405500	0.408605	0.410350	0.411070
	M+2	0.162400	0.162262	0.161400	0.161283
	M+3	0.061950	0.061376	0.060900	0.061230
	M+4	0.015150	0.014653	0.014300	0.014478
	M+5	0.003350	0.003184	0.003850	0.003149
	M+6			0.000600	0.000506
Asp_390	M+0	0.311850	0.313463	0.319100	0.320994
	M+1	0.146850	0.149146	0.137900	0.138997
	M+2	0.345050	0.340553	0.352900	0.349258
	M+3	0.127000	0.127196	0.122250	0.122660
	M+4	0.053150	0.053037	0.052400	0.052373
	M+5	0.012900	0.013071	0.012200	0.012367
	M+6	0.002950	0.002952	0.002850	0.002815
	M+7	0.000300	0.000501	0.000400	0.000462
Asp_418	M+0	0.307350	0.307571	0.313700	0.315893
	M+1	0.145900	0.146297	0.137950	0.138964
	M+2	0.333250	0.331588	0.343000	0.339213
	M+3	0.136050	0.137294	0.131250	0.131643
	M+4	0.057700	0.057526	0.055900	0.055930
	M+5	0.015350	0.015363	0.014100	0.014325
	M+6	0.003550	0.003588	0.003250	0.003338
	M+7	0.000750	0.000657	0.000650	0.000593
	M+ 8	0.000150	0.000101	0.000200	0.000090
	M+ 9				
Asp_460	M+0	0.294000	0.297420	0.301750	0.305467
	M+1	0.155700	0.151504	0.146150	0.144684
	M+2	0.328550	0.325534	0.336350	0.332670
	M+3	0.142750	0.143637	0.139000	0.138418

	M+4	0.060500	0.060232	0.059300	0.058507
	M+5	0.015800	0.016785	0.014900	0.015727
	M+6	0.002750	0.003993	0.002550	0.003716
	M+7				
	M+ 8				
Glu_330	M+0	0.202050	0.190937	0.203350	0.191326
	M+1	0.086700	0.080705	0.075450	0.072958
	M+2	0.361550	0.361985	0.369550	0.367318
	M+3	0.114950	0.117689	0.107750	0.112328
	M+4	0.176700	0.186582	0.183700	0.192224
	M+5	0.040800	0.043660	0.042100	0.044738
	M+6	0.014550	0.015552	0.015200	0.016113
	M+7	0.002300	0.002439	0.002450	0.002526
	M+ 8	0.000400	0.000406	0.000400	0.000422
Glu_404	M+0	0.180400	0.171716	0.181300	0.172066
	M+1	0.090650	0.085295	0.080800	0.078354
	M+2	0.338850	0.337252	0.344650	0.341545
	M+3	0.130300	0.132783	0.123800	0.128060
	M+4	0.179500	0.187723	0.186150	0.192611
	M+5	0.052750	0.055903	0.054350	0.057074
	M+6	0.021900	0.023206	0.023100	0.023965
	M+7	0.004500	0.004842	0.004600	0.004999
	M+ 8	0.001050	0.001093	0.001100	0.001134
	M+ 9	0.000150	0.000162	0.000100	0.000168
Glu_432	M+0	0.171300	0.162705	0.174750	0.166264
	M+1	0.094100	0.087686	0.085000	0.080274
	M+2	0.320300	0.322905	0.330350	0.331058
	M+3	0.143600	0.143235	0.134550	0.135552
	M+4	0.176200	0.185023	0.182650	0.189924
	M+5	0.062350	0.064762	0.060700	0.063415
	M+6	0.024550	0.025811	0.024750	0.025877
	M+7	0.005900	0.006198	0.005700	0.006008
	M+ 8	0.001300	0.001398	0.001300	0.001367
	M+ 9	0.000300	0.000236	0.000250	0.000225
	M+10	0.000050	0.000034		
Glu_474	M+0	0.162800	0.157336	0.166950	0.160776
	M+1	0.097000	0.090100	0.089500	0.083049
	M+2	0.313650	0.315170	0.320750	0.322813
	M+3	0.150000	0.149076	0.142200	0.141910
	M+4	0.175600	0.183712	0.181650	0.188203
	M+5	0.066800	0.068716	0.065100	0.067570
	M+6	0.026000	0.027142	0.026100	0.027164
	M+7	0.006550	0.006860	0.006300	0.006678

	M+ 8	0.001500	0.001564	0.001450	0.001528
	M+ 9				
Gln_431	M+0	0.170050	0.162477	0.173250	0.166031
	M+1	0.093350	0.088123	0.084850	0.080734
	M+2	0.321750	0.322422	0.331800	0.330531
	M+3	0.143200	0.143967	0.134450	0.136338
	M+4	0.178050	0.184598	0.183850	0.189449
	M+5	0.063000	0.065015	0.061350	0.063702
	M+6	0.024550	0.025620	0.024600	0.025672
	M+7	0.006050	0.006146	0.005750	0.005959
Cit_459	M+0	0.166100	0.156311	0.170750	0.160873
	M+1	0.091450	0.087924	0.084450	0.081636
	M+2	0.311450	0.312186	0.321700	0.321488
	M+3	0.145600	0.149317	0.138400	0.141648
	M+4	0.176450	0.183919	0.181800	0.188012
	M+5	0.069050	0.071580	0.065650	0.069047
	M+6	0.029000	0.028927	0.027650	0.028149
	M+7	0.008100	0.007623	0.007200	0.007119
	M+ 8	0.002550	0.001812	0.002150	0.001673
	M+ 9	0.000300	0.000338	0.000250	0.000301
	M+10			0.000050	0.000047
Val_260	M+0	0.191550	0.200264	0.189750	0.200556
	M+1	0.075400	0.073921	0.067050	0.066232
	M+2	0.375150	0.376584	0.381100	0.382222
	M+3	0.105650	0.105005	0.098750	0.099445
	M+4	0.197950	0.190860	0.205900	0.196760
	M+5	0.037300	0.037034	0.038650	0.037869
	M+6	0.014600	0.014192	0.016550	0.014702
	M+7	0.002400	0.001811	0.002250	0.001873
	M+ 8				
Val_288	M+0	0.190950	0.194894	0.192000	0.195744
	M+1	0.074600	0.073690	0.068050	0.067793
	M+2	0.366600	0.365960	0.373100	0.371694
	M+3	0.111500	0.113055	0.105900	0.107714
	M+4	0.191950	0.188103	0.196450	0.192970
	M+5	0.045950	0.045345	0.046200	0.044947
	M+6	0.016100	0.015909	0.016100	0.016136
	M+7	0.002300	0.002576	0.002300	0.002538
	M+ 8				
Ile_200	M+0	0.198550	0.206839	0.202000	0.211963
	M+1	0.074700	0.076299	0.064050	0.065001
	M+2	0.380850	0.381482	0.394850	0.395074
	M+3	0.106950	0.106905	0.093750	0.094160

	M+4	0.192600	0.183506	0.203350	0.191970
	M+5	0.036850	0.035871	0.034050	0.032960
	M+6	0.008150	0.008129	0.007950	0.008029
	M+7	0.001100	0.000923		
	M+ 8	0.000250	0.000046		
Ile_274	M+0	0.178750	0.186017	0.183450	0.190625
	M+1	0.085050	0.082391	0.078800	0.072572
	M+2	0.354500	0.355021	0.363650	0.366661
	M+3	0.124900	0.124250	0.115950	0.113323
	M+4	0.190350	0.184882	0.198650	0.192088
	M+5	0.050100	0.048351	0.043200	0.045885
	M+6	0.016300	0.015901	0.016300	0.015889
Phe_302	M+0	0.395800	0.389676	0.392200	0.387766
	M+1	0.421800	0.425108	0.425200	0.428045
	M+2	0.132150	0.134048	0.130900	0.133197
	M+3	0.040650	0.042052	0.041700	0.041991
	M+4	0.007950	0.007721	0.007800	0.007624
	M+5	0.001650	0.001240	0.002250	0.001225
His_440	M+0	0.018750	0.017808	0.020400	0.019778
	M+1	0.114050	0.115839	0.087300	0.087768
	M+2	0.227450	0.226308	0.212750	0.212005
	M+3	0.221500	0.221263	0.218600	0.218492
	M+4	0.198000	0.202392	0.210150	0.209395
	M+5	0.144850	0.141747	0.164750	0.164835
	M+6	0.052050	0.052031	0.059150	0.060478
	M+7	0.017900	0.017579	0.020800	0.021165
	M+ 8	0.004050	0.004051	0.004800	0.004881
	M+ 9	0.001000	0.000833	0.001150	0.001019
	M+10	0.000250	0.000131	0.000200	0.000160
F6P_259/97	M+0	0.017184	0.015792	0.015610	0.017133
	M+1	0.027998	0.030425	0.026568	0.027014
	M+2	0.666088	0.665661	0.708304	0.710075
	M+3	0.073012	0.073225	0.068064	0.065475
	M+4	0.185640	0.185444	0.153299	0.155508
	M+5	0.027912	0.026418	0.025713	0.022286
	M+6	0.002165	0.002749	0.002443	0.002269
	M+7	0.000000	0.000268	0.000000	0.000226
DHAP_169/97	M+0	0.412677	0.410636	0.402386	0.400864
	M+1	0.037217	0.041627	0.034593	0.036412
	M+2	0.521095	0.519125	0.537695	0.535757
	M+3	0.026430	0.026350	0.022989	0.024645
	M+4	0.002582	0.002154	0.002336	0.002219
3PG_185/79	M+0	0.503373	0.497511	0.504488	0.498198

	M+1	0.036384	0.040542	0.036554	0.036140
	M+2	0.431780	0.435539	0.435294	0.441047
	M+3	0.024245	0.022618	0.020981	0.020800
	M+4	0.004218	0.003595	0.002683	0.003636
PEP_167/79	M+0	0.496763	0.498874	0.499144	0.499562
	M+1	0.037030	0.040308	0.035158	0.035893
	M+2	0.439736	0.435684	0.442787	0.441208
	M+3	0.022541	0.022296	0.019230	0.020478
	M+4	0.003930	0.002698	0.003681	0.002729

Xylose MS data:

XA = Xylose Aerobic

XN = Xylose Anaerobic

MS Fragment	Mass Isotopomer	XA		XN	
		Measured	Simulated	Measured	Simulated
Pyr_174	M+0	0.523100	0.528144	0.516150	0.522169
	M+1	0.066050	0.068745	0.067050	0.067844
	M+2	0.029650	0.032346	0.030250	0.031942
	M+3	0.335900	0.326563	0.340550	0.332981
	M+4	0.030400	0.029785	0.030900	0.030361
	M+5	0.014200	0.013710	0.014600	0.013982
	M+6	0.000600	0.000621	0.000500	0.000633
Ala_260	M+0	0.452150	0.458573	0.444100	0.451401
	M+1	0.109950	0.111849	0.110300	0.112018
	M+2	0.054450	0.054234	0.055700	0.055827
	M+3	0.296250	0.289242	0.301150	0.293353
	M+4	0.057950	0.057458	0.059450	0.058355
	M+5	0.025350	0.024774	0.025550	0.025120
	M+6	0.003150	0.003229	0.003300	0.003276
Gly_218	M+0	0.481800	0.476704	0.479500	0.471062
	M+1	0.383150	0.389162	0.387250	0.393692
	M+2	0.097750	0.098868	0.098650	0.099497
	M+3	0.032250	0.030434	0.030500	0.030854
	M+4	0.004800	0.004146	0.004150	0.004197
Gly_246	M+0	0.465900	0.463448	0.452050	0.454040

	M+1	0.114050	0.115258	0.121350	0.121181
	M+2	0.327850	0.328108	0.333900	0.330448
	M+3	0.062150	0.063188	0.061900	0.064083
	M+4	0.026600	0.025992	0.027450	0.026204
	M+5	0.003250	0.003352	0.003350	0.003385
	M+6	0.000300	0.000601	0.000050	0.000606
Ser_288	M+0	0.409750	0.410570	0.391700	0.391371
	M+1	0.193400	0.191695	0.215950	0.215040
	M+2	0.296050	0.296546	0.290550	0.292707
	M+3	0.071400	0.071898	0.072600	0.072141
	M+4	0.024850	0.024762	0.024150	0.024297
	M+5	0.003750	0.003853	0.004300	0.003786
	M+6	0.000850	0.000612	0.000750	0.000597
Ser_302	M+0	0.443350	0.445460	0.433350	0.438911
	M+1	0.125300	0.126100	0.124050	0.127057
	M+2	0.318950	0.320469	0.320100	0.324454
	M+3	0.074600	0.074773	0.077550	0.075890
	M+4	0.028550	0.027976	0.030600	0.028387
	M+5	0.006900	0.004396	0.009850	0.004463
	M+6	0.002400	0.000744	0.004450	0.000755
Ser_362	M+0	0.368450	0.369239	0.353250	0.351972
	M+1	0.202400	0.199737	0.221250	0.219453
	M+2	0.295100	0.293076	0.291950	0.290541
	M+3	0.091200	0.091110	0.091750	0.091831
	M+4	0.036900	0.037082	0.036500	0.036571
	M+5	0.005250	0.007762	0.004600	0.007677
	M+6	0.000700	0.001709	0.000750	0.001676
Ser_390	M+0	0.361100	0.361412	0.344800	0.343743
	M+1	0.163550	0.165281	0.171050	0.172680
	M+2	0.111700	0.113110	0.126500	0.127597
	M+3	0.244550	0.240602	0.240400	0.237754
	M+4	0.076800	0.076644	0.076100	0.076262
	M+5	0.033750	0.033780	0.032800	0.033011
	M+6	0.006850	0.007216	0.006900	0.007055
	M+7	0.001500	0.001666	0.001400	0.001618
	M+ 8	0.000100	0.000249		
Ser_432	M+0	0.350900	0.349484	0.330800	0.332398
	M+1	0.171350	0.171618	0.175300	0.178196
	M+2	0.113450	0.114906	0.131550	0.129149
	M+3	0.238700	0.236414	0.237800	0.234135
	M+4	0.081100	0.082007	0.080150	0.081550
	M+5	0.035150	0.035257	0.034750	0.034499
	M+6	0.007650	0.008109	0.006700	0.007929

	M+7	0.001750	0.001859	0.002650	0.001808
	M+ 8			0.000150	0.000289
Asp_302	M+0	0.439550	0.440160	0.428650	0.429271
	M+1	0.129800	0.135462	0.139900	0.144042
	M+2	0.319000	0.317188	0.321550	0.318519
	M+3	0.077000	0.074430	0.075700	0.075271
	M+4	0.028400	0.027607	0.028200	0.027720
	M+5	0.004900	0.004340	0.004750	0.004362
	M+6	0.001050	0.000732	0.001100	0.000734
	M+7	0.000200	0.000075	0.000150	0.000075
Asp_376	M+0	0.390650	0.395850	0.381050	0.386057
	M+1	0.155450	0.151135	0.162500	0.158126
	M+2	0.309750	0.308877	0.311400	0.310284
	M+3	0.093300	0.092922	0.094700	0.094042
	M+4	0.041050	0.040448	0.039750	0.040656
	M+5	0.009000	0.008487	0.009150	0.008544
	M+6	0.000800	0.001948	0.001450	0.001955
Asp_390	M+0	0.259150	0.252143	0.256000	0.252907
	M+1	0.239450	0.239010	0.232450	0.233610
	M+2	0.234500	0.239574	0.230500	0.233813
	M+3	0.179050	0.180127	0.189000	0.186968
	M+4	0.060500	0.061722	0.063350	0.063762
	M+5	0.021650	0.021635	0.022600	0.022823
	M+6	0.004800	0.004647	0.004900	0.004902
	M+7	0.000900	0.000976	0.001100	0.001040
Asp_418	M+0	0.243800	0.242017	0.239250	0.237656
	M+1	0.235500	0.236290	0.236650	0.238363
	M+2	0.105350	0.105849	0.095700	0.098315
	M+3	0.185000	0.185839	0.189050	0.188989
	M+4	0.150850	0.150267	0.156500	0.154372
	M+5	0.054550	0.054567	0.056350	0.056273
	M+6	0.019650	0.019714	0.020450	0.020382
	M+7	0.004050	0.004354	0.004700	0.004507
	M+ 8	0.001100	0.000939	0.001200	0.000973
	M+ 9	0.000150	0.000142	0.000100	0.000147
Asp_460	M+0	0.233450	0.530870	0.232200	0.229812
	M+1	0.238100	0.389604	0.236450	0.238249
	M+2	0.110500	0.436034	0.101550	0.102937
	M+3	0.181750	0.325387	0.187100	0.186050
	M+4	0.153350	0.213582	0.157800	0.155481
	M+5	0.057900	0.074685	0.059750	0.059524
	M+6	0.021200	0.025193	0.020900	0.021603
	M+7	0.003750	0.003750	0.004100	0.005044

M+8	0.001096	0.000100	0.000100	0.001096
M+0	0.267217	0.274100	0.274100	0.267217
M+1	0.081248	0.084350	0.084350	0.081248
M+2	0.366377	0.364100	0.364100	0.366377
M+3	0.097748	0.095500	0.095500	0.097748
M+4	0.141923	0.138100	0.138100	0.141923
M+5	0.032035	0.030850	0.030850	0.032035
M+6	0.011360	0.010900	0.010900	0.011360
M+7	0.001765	0.001700	0.001700	0.001765
M+8	0.000294	0.000350	0.000350	0.000294
M+0	0.240317	0.245500	0.245500	0.240317
M+1	0.090862	0.093050	0.093050	0.090862
M+2	0.343768	0.343900	0.343900	0.343768
M+3	0.115224	0.112600	0.112600	0.115224
M+4	0.146384	0.143700	0.143700	0.146384
M+5	0.041817	0.040800	0.040800	0.041817
M+6	0.017164	0.016800	0.016800	0.017164
M+7	0.003533	0.003000	0.003000	0.003533
M+8	0.000795	0.000600	0.000600	0.000795
M+9	0.156695	0.160500	0.160500	0.156695
M+1	0.142085	0.147300	0.147300	0.142085
M+2	0.238176	0.236600	0.236600	0.238176
M+3	0.205886	0.205000	0.205000	0.205886
M+4	0.128017	0.125650	0.125650	0.128017
M+5	0.088278	0.085700	0.085700	0.088278
M+6	0.028325	0.027250	0.027250	0.028325
M+7	0.009908	0.009500	0.009500	0.009908
M+8	0.002104	0.002000	0.002000	0.002104
M+9	0.000449	0.000500	0.000500	0.000449
M+10	0.000067	0.000100	0.000100	0.000067
M+0	0.151524	0.153100	0.153100	0.151524
M+1	0.142508	0.148100	0.148100	0.142508
M+2	0.235010	0.232700	0.232700	0.235010
M+3	0.206915	0.207250	0.207250	0.206915
M+4	0.130599	0.128700	0.128700	0.130599
M+5	0.089619	0.087800	0.087800	0.089619
M+6	0.030319	0.029650	0.029650	0.030319
M+7	0.010538	0.010250	0.010250	0.010538
M+8	0.002369	0.002350	0.002350	0.002369
M+9	0.000507	0.000150	0.000150	0.000507
M+0	0.156475	0.159250	0.159250	0.156475
M+1	0.142426	0.147350	0.147350	0.142426

	M+2	0.235650	0.234342	0.238300	0.238012
	M+3	0.203550	0.205192	0.205600	0.206127
	M+4	0.120100	0.121919	0.126550	0.128060
	M+5	0.084550	0.088130	0.086000	0.088174
	M+6	0.026600	0.028033	0.027300	0.028328
	M+7	0.009500	0.009904	0.009600	0.009811
Cit_459	M+0	0.148950	0.144335	0.144000	0.143053
	M+1	0.138850	0.137097	0.149700	0.147356
	M+2	0.178700	0.181167	0.149300	0.153711
	M+3	0.188600	0.190281	0.205150	0.206032
	M+4	0.148050	0.147139	0.136700	0.133625
	M+5	0.098650	0.100550	0.107950	0.108680
	M+6	0.065400	0.067219	0.070850	0.071955
	M+7	0.021750	0.022265	0.023850	0.024612
	M+ 8	0.008550	0.007814	0.009550	0.008603
	M+ 9	0.002150	0.001696	0.002450	0.001886
	M+10	0.000400	0.000371	0.000450	0.000412
Val_260	M+0	0.277200	0.477464	0.273550	0.280270
	M+1	0.081150	0.155071	0.073150	0.072118
	M+2	0.370700	0.747284	0.376850	0.380189
	M+3	0.090700	0.195705	0.085350	0.084840
	M+4	0.140150	0.331010	0.147300	0.143935
	M+5	0.028150	0.065184	0.030200	0.026823
	M+6	0.010050	0.024242	0.010950	0.010292
	M+7	0.001800	0.003611	0.002600	0.001298
	M+ 8	0.000150	0.000150		
Val_288	M+0	0.274400	0.274405	0.268700	0.271755
	M+1	0.080300	0.082663	0.077000	0.075878
	M+2	0.203100	0.205822	0.203000	0.206879
	M+3	0.219750	0.217088	0.223050	0.220281
	M+4	0.061500	0.064315	0.061250	0.061814
	M+5	0.126200	0.121253	0.130750	0.127304
	M+6	0.024400	0.023584	0.025500	0.024606
	M+7	0.009750	0.009415	0.010150	0.009945
	M+ 8	0.000600	0.001220	0.000700	0.001287
Ile_200	M+0	0.191550	0.197606	0.192100	0.197421
	M+1	0.157450	0.158239	0.148550	0.150213
	M+2	0.267600	0.268210	0.263950	0.265726
	M+3	0.197250	0.195467	0.202150	0.199388
	M+4	0.108850	0.105095	0.108750	0.104861
	M+5	0.066150	0.063838	0.072850	0.069873
	M+6	0.009250	0.009319	0.008600	0.010042
	M+7	0.001650	0.002091	0.002050	0.002325

Ile_274	M+ 8	0.000250	0.000131	0.000950	0.000146
	M+0	0.174550	0.177714	0.169600	0.177547
	M+1	0.156550	0.155468	0.152300	0.148238
	M+2	0.259500	0.258301	0.256550	0.255526
	M+3	0.199800	0.199065	0.205300	0.202159
	M+4	0.119900	0.116574	0.119750	0.116535
	M+5	0.072950	0.071130	0.077450	0.076669
Phe_302	M+6	0.016850	0.016300	0.018950	0.017348
	M+0	0.454750	0.446168	0.448550	0.441118
	M+1	0.118200	0.124849	0.116000	0.123168
	M+2	0.322950	0.320907	0.329150	0.325814
	M+3	0.072450	0.074819	0.074050	0.076031
	M+4	0.027250	0.028025	0.027550	0.028540
His_440	M+5	0.004350	0.004403	0.004700	0.004486
	M+0	0.001350	0.001118	0.002050	0.001201
	M+1	0.000000	0.004507	0.000000	0.004346
	M+2	0.312650	0.324015	0.315800	0.310327
	M+3	0.326950	0.313874	0.303200	0.305719
	M+4	0.127450	0.127384	0.120650	0.124295
	M+5	0.114000	0.115784	0.126400	0.124818
	M+6	0.078650	0.075355	0.087150	0.085664
	M+7	0.026900	0.026487	0.030700	0.030326
	M+ 8	0.009600	0.009040	0.010500	0.010473
F6P_259/97	M+ 9	0.002000	0.001964	0.002550	0.002280
	M+10	0.000500	0.000405	0.000900	0.000472
	M+0	0.018305	0.023830	0.025969	0.032792
	M+1	0.014058	0.008402	0.014569	0.008507
	M+2	0.021189	0.014723	0.020532	0.014130
	M+3	0.587049	0.592394	0.577954	0.584256
	M+4	0.033627	0.028790	0.034992	0.028056
	M+5	0.024477	0.021712	0.023963	0.021155
	M+6	0.300494	0.305756	0.300969	0.306704
	M+7	0.000802	0.001243	0.001052	0.001241
DHAP_169/97	M+0	0.514827	0.514018	0.533812	0.531604
	M+1	0.020536	0.022686	0.020385	0.022783
	M+2	0.015796	0.013292	0.010651	0.012597
	M+3	0.448195	0.447502	0.434528	0.430609
	M+4	0.000645	0.000668	0.000623	0.000641
3PG_185/79	M+0	0.602569	0.592986	0.598751	0.586277
	M+1	0.021955	0.025066	0.021521	0.024644
	M+2	0.011550	0.014044	0.008299	0.013855
	M+3	0.363234	0.363927	0.370569	0.371171
	M+4	0.000692	0.000985	0.000860	0.001001

PEP_167/79	M+0	0.593093	0.594610	0.589040	0.587883
	M+1	0.021008	0.024723	0.020385	0.024304
	M+2	0.010162	0.012849	0.007776	0.012673
	M+3	0.375498	0.364863	0.382283	0.372128
	M+4	0.000239	0.000708	0.000516	0.000720

Metabolite Abbreviations: Pyr, Pyruvate; Ala, Alanine; Gly, Glycine; Ser, Serine; Asp, Aspartate; Glu, Glutamate; Gln, Glutamine; Cit, Citrate; Val, Valine; Ile, Isoleucine; Phe, Phenylalanine; His, Histidine; F6P, Fructose 6-phosphate; DHAP, Dihydroxyacetone phosphate; 3PG, 3-Phosphoglycerate; PEP, phosphoenolpyruvate.

Table A-3 Best-Fit Fluxes and Flux Confidence Intervals

Notation:

95 lb = lower bound of flux 95% confidence interval

68 lb = lower bound of flux 68% confidence interval

best = best-fit flux value

68 ub = upper bound of flux 68% confidence interval

95 ub = upper bound of flux 95% confidence interval

“net” indicates net flux for a reversible reaction.

“exch” indicates exchange flux for a reversible reaction.

Glucose Aerobic

Flux	95 lb	68 lb	best	68 ub	95 ub
v1	9.0445	9.6010	10.1730	10.7430	11.2930
v2	13.3170	14.3880	15.4940	16.5970	17.6580
v3	0.1179	0.1643	0.2120	0.2597	0.3061
v4	1.1307	1.2095	1.2908	1.3720	1.4505
v5	0.2308	0.2378	0.2449	0.2520	0.2589
v6	0.0000	0.0000	0.4796	0.8733	1.0000
v7	0.0000	0.1267	0.5204	1.0000	1.0000
v8	1.0000	1.0000	1.0000	1.0000	1.0000
v9 net	7.9601	8.4585	8.9690	9.4775	9.9668
v9 exch	0.0000	0.0000	0.6229	5.2831	16.6900
v10 net	8.1328	8.6696	9.2208	9.7702	10.2990
v10 exch	0.3278	0.4614	12.6530	Inf	Inf
v11 net	8.1328	8.6696	9.2208	9.7702	10.2990
v11 exch	0.3282	0.4619	1.1085	Inf	Inf
v12 net	6.8413	7.3736	7.9239	8.4724	9.0003
v12 exch	36.4530	39.1200	41.8910	44.7750	47.6590
v13	15.1390	16.2250	17.3440	18.4600	19.5340
v14	15.0880	16.1580	17.2640	18.3660	19.4270
v15	14.9590	16.0290	17.1350	18.2370	19.2980
v16	1.1368	1.2157	1.2969	1.3782	1.4566
v17	1.1307	1.2095	1.2908	1.3720	1.4505
v18	0.6536	0.7168	0.7865	0.8588	0.9308
v19 net	0.2924	0.3135	0.3368	0.3609	0.3849
v19 exch	1.4761	3.2555	9.5409	Inf	Inf
v20 net	0.3609	0.4031	0.4497	0.4981	0.5462
v20 exch	3.0584	3.9448	11.6820	Inf	Inf

v21 net	0.3609	0.4031	0.4497	0.4981	0.5462
v21 exch	3.0586	3.9446	9.0449	Inf	Inf
v22 net	-0.2410	-0.2168	-0.1925	-0.1691	-0.1479
v22 exch	1.2490	1.3795	1.5631	1.8546	2.2810
v23 net	-0.3053	-0.2813	-0.2572	-0.2340	-0.2129
v23 exch	0.0801	0.1203	0.1797	Inf	Inf
v24 net	-0.3053	-0.2813	-0.2572	-0.2340	-0.2129
v24 exch	7.2239	7.7086	8.2824	8.9462	9.6131
v25 net	0.2129	0.2340	0.2572	0.2813	0.3053
v25 exch	0.0801	0.1203	4.8582	Inf	Inf
v26	0.0059	0.0653	0.1617	0.3073	0.9963
v27	0.5047	0.6209	0.7838	0.9700	1.1398
v28	13.8450	14.9490	16.0850	17.2170	18.3060
v29	13.3170	14.3880	15.4940	16.5970	17.6580
v30	0.1037	0.1501	0.1978	0.2455	0.2919
v31	0.0000	0.4756	0.5315	0.5478	0.5636
v32	0.0000	0.0000	0.0000	0.0540	0.5624
v33	0.2356	0.2432	0.2519	0.2607	0.2695
v34	0.2356	0.2432	0.2519	0.2607	0.2695
v35	0.2356	0.2432	0.2519	0.2607	0.2695
v36	0.0000	0.0000	0.0023	0.0071	0.0118
v37 net	0.0016	0.0017	0.0041	0.0088	0.0136
v37 exch	0.0000	0.0000	0.4878	Inf	Inf
v38 net	0.0917	0.0947	0.0995	0.1051	0.1108
v38 exch	0.0980	0.3906	2.4247	Inf	Inf
v39 net	-0.8906	-0.2076	-0.0622	0.0320	0.0903
v39 exch	0.1462	0.3822	2.9216	Inf	Inf
v40	0.0388	0.0402	0.0801	0.1275	0.1562
v41 net	-0.0172	-0.0165	0.0231	0.0707	0.0966
v41 exch	0.0151	0.0158	0.0314	0.0496	0.0626
v42 net	-0.0054	0.0162	0.0640	0.1053	0.1089
v42 exch	0.0000	0.0000	0.0000	0.0221	0.0904
v43	0.0000	0.0868	0.2565	0.3689	0.8901
v44 net	0.1579	0.2179	0.3141	0.4601	1.1527
v44 exch	0.0000	0.0000	0.0726	0.2374	0.6631
v45 net	-0.2168	0.2694	0.3244	0.3350	0.3455
v45 exch	0.0000	0.0000	2.1382	Inf	Inf
v46	1.1410	1.2159	1.2798	1.3452	1.3910
v47	0.2093	0.2156	0.2221	0.2284	0.2347
v48	0.0381	0.0392	0.0404	0.0416	0.0427
v49	0.0372	0.0383	0.0394	0.0406	0.0417
v50	0.3124	0.3722	0.4698	0.5592	0.5782
v51	0.0235	0.0243	0.0250	0.0257	0.0264
v52	0.0000	0.0000	0.0163	0.0727	0.1154
v53	0.0000	0.0398	0.0961	0.1156	0.1188
v54	0.0882	0.1362	0.2324	0.3171	0.3278

v55	0.0000	0.0426	0.1383	0.2208	0.2286
v56	0.0016	0.0017	0.0017	0.0018	0.0018
v57	0.0118	0.0121	0.0125	0.0129	0.0132
v58	0.0660	0.0680	0.0700	0.0720	0.0740
v59	0.0683	0.0704	0.0725	0.0746	0.0766
v60	0.0446	0.0459	0.0473	0.0486	0.0500
v61	0.0612	0.0630	0.0649	0.0668	0.0686
v62	0.0309	0.0319	0.0328	0.0337	0.0347
v63	0.0235	0.0243	0.0250	0.0257	0.0264
v64	0.0065	0.0067	0.0069	0.0071	0.0072
v65	0.0152	0.0157	0.0162	0.0166	0.0171
v66	0.0284	0.0292	0.0301	0.0310	0.0318
v67	0.0121	0.0125	0.0128	0.0132	0.0136
v68	0.0249	0.0257	0.0264	0.0272	0.0280
v69	0.0126	0.0130	0.0134	0.0138	0.0141
v70	0.0123	0.0127	0.0131	0.0135	0.0138
v71	0.0008	0.0009	0.0009	0.0009	0.0009
v72	0.0006	0.0006	0.0006	0.0006	0.0006
v73	0.0006	0.0006	0.0006	0.0006	0.0006
v74	0.0008	0.0009	0.0009	0.0009	0.0009
v75	0.0008	0.0009	0.0009	0.0009	0.0009
v76	0.0001	0.0001	0.0001	0.0002	0.0002
v77	0.0004	0.0004	0.0004	0.0004	0.0004
v78	0.0011	0.0011	0.0011	0.0012	0.0012
v79	0.0014	0.0015	0.0015	0.0016	0.0016
v80	0.0013	0.0013	0.0013	0.0014	0.0014
v81	0.0015	0.0016	0.0016	0.0017	0.0017
v82	0.0030	0.0031	0.0032	0.0033	0.0034
v83	0.0030	0.0031	0.0032	0.0033	0.0034
v84	0.0030	0.0031	0.0032	0.0033	0.0034
v85	0.0030	0.0031	0.0032	0.0033	0.0034
v86	0.0010	0.0010	0.0011	0.0011	0.0011
v87	0.0005	0.0005	0.0005	0.0006	0.0006
v88	0.0002	0.0002	0.0002	0.0002	0.0002
v89	0.0053	0.0055	0.0056	0.0058	0.0060
v90	0.1198	0.1234	0.1271	0.1308	0.1344
v91	0.2620	0.2699	0.2779	0.2859	0.2938
v92	0.1865	0.1921	0.1979	0.2034	0.2092
v93	0.0000	0.7690	2.2080	3.7460	5.2701
v94	15.0940	16.9610	19.1090	21.4760	23.8170

Glucose Anaerobic

	95 lb	68 lb	best fit	68 ub	95 ub
v1	12.0290	12.7480	13.4920	14.2430	15.0020
v2	17.9810	19.3010	20.6360	22.0390	23.5110

v3	0.2324	0.2797	0.3280	0.3768	0.4241
v4	2.4184	2.7123	2.9849	3.2346	3.4804
v5	0.2222	0.2254	0.2287	0.2320	0.2351
v6	0.0000	0.0000	0.0000	0.4193	0.9177
v7	0.0823	0.5807	1.0000	1.0000	1.0000
v8	1.0000	1.0000	1.0000	1.0000	1.0000
v9 net	11.0450	11.7230	12.4230	13.1280	13.8250
v9 exch	0.0000	0.0000	0.0000	9.5416	215.1700
v10 net	11.1870	11.8920	12.6210	13.3570	14.0920
v10 exch	0.5699	0.7552	10.1750	Inf	Inf
v11 net	11.1870	11.8920	12.6210	13.3570	14.0920
v11 exch	0.5702	0.7557	2.3569	Inf	Inf
v12 net	8.3893	8.9935	9.6306	10.3200	11.0640
v12 exch	44.1370	47.1230	50.2530	53.5090	56.8950
v13	19.7520	21.1000	22.4200	23.8300	25.3040
v14	19.6960	20.9950	22.3510	23.7550	25.2270
v15	19.5760	20.8750	22.2310	23.6340	25.1060
v16	2.4228	2.7183	2.9907	3.2401	3.4862
v17	2.4184	2.7123	2.9849	3.2346	3.4804
v18	0.5734	0.6238	0.6792	0.7527	0.8457
v19 net	0.2607	0.2776	0.2961	0.3206	0.3516
v19 exch	0.7255	1.5537	2.8569	6.1448	32.2360
v20 net	0.3126	0.3462	0.3832	0.4322	0.4942
v20 exch	3.5289	4.8592	10.9080	Inf	Inf
v21 net	0.3126	0.3462	0.3832	0.4322	0.4942
v21 exch	3.5288	4.8593	14.2100	Inf	Inf
v22 net	-0.2169	-0.1859	-0.1614	-0.1429	-0.1260
v22 exch	2.3483	2.7446	3.2864	4.2765	7.0157
v23 net	-0.2772	-0.2463	-0.2218	-0.2033	-0.1865
v23 exch	0.0579	0.0975	0.9969	Inf	Inf
v24 net	-0.2772	-0.2463	-0.2218	-0.2033	-0.1865
v24 exch	6.6373	7.1001	7.5901	8.1816	8.8821
v25 net	0.1865	0.2033	0.2218	0.2463	0.2772
v25 exch	0.0579	0.0975	0.1871	Inf	Inf
v26	0.0110	0.0953	0.1797	0.2803	0.8106
v27	0.5027	0.6349	0.7732	0.9232	1.0980
v28	18.5530	19.8730	21.2380	22.6540	24.2580
v29	17.9810	19.3010	20.6360	22.0390	23.5110
v30	0.2192	0.2665	0.3148	0.3635	0.4108
v31	0.0000	0.3338	0.4290	0.4351	0.4412
v32	0.0000	0.0000	0.0000	0.0951	0.4392
v33	0.2264	0.2297	0.2330	0.2362	0.2396
v34	0.2264	0.2297	0.2330	0.2362	0.2396
v35	0.2264	0.2297	0.2330	0.2362	0.2396
v36	0.0000	0.0000	0.0000	0.0007	0.0026
v37 net	0.0016	0.0016	0.0016	0.0023	0.0041

v37 exch	0.0000	0.0000	1.6906	Inf	Inf
v38 net	0.0881	0.0894	0.0907	0.0920	0.0939
v38 exch	0.7185	3.3195	1.1319E+05	Inf	Inf
v39 net	-0.7201	-0.1904	-0.0890	-0.0044	0.0796
v39 exch	0.6720	3.2258	1.0000E+07	Inf	Inf
v40	0.0339	0.0345	0.0684	0.1151	0.1417
v41 net	-0.0190	-0.0185	0.0153	0.0619	0.0876
v41 exch	0.0176	0.0180	0.0357	0.0599	0.0747
v42 net	-0.0050	0.0193	0.0659	0.1006	0.1022
v42 exch	0.0000	0.0031	0.0433	0.1618	0.4626
v43	0.0000	0.0741	0.1955	0.3117	0.7444
v44 net	0.1533	0.2376	0.3220	0.4236	0.9535
v44 exch	0.0000	0.0000	0.0350	0.1826	0.3842
v45 net	-0.1314	0.2056	0.3007	0.3050	0.3092
v45 exch	0.0000	0.0000	4.3476	Inf	Inf
v46	1.1097	1.1512	1.2013	1.2484	1.2681
v47	0.2015	0.2044	0.2074	0.2103	0.2132
v48	0.0367	0.0372	0.0377	0.0383	0.0388
v49	0.0358	0.0363	0.0368	0.0373	0.0379
v50	0.3038	0.3576	0.4512	0.5240	0.5321
v51	0.0227	0.0230	0.0233	0.0237	0.0240
v52	0.0000	0.0000	0.0306	0.0984	0.1074
v53	0.0000	0.0066	0.0744	0.1063	0.1078
v54	0.0862	0.1361	0.2295	0.2999	0.3046
v55	0.0000	0.0484	0.1417	0.2111	0.2145
v56	0.0016	0.0016	0.0016	0.0016	0.0016
v57	0.0113	0.0115	0.0117	0.0118	0.0120
v58	0.0636	0.0645	0.0654	0.0663	0.0672
v59	0.0658	0.0667	0.0677	0.0686	0.0696
v60	0.0429	0.0435	0.0441	0.0448	0.0454
v61	0.0589	0.0597	0.0606	0.0615	0.0623
v62	0.0298	0.0302	0.0306	0.0311	0.0315
v63	0.0227	0.0230	0.0233	0.0237	0.0240
v64	0.0062	0.0063	0.0064	0.0065	0.0066
v65	0.0147	0.0149	0.0151	0.0153	0.0155
v66	0.0273	0.0277	0.0281	0.0285	0.0289
v67	0.0116	0.0118	0.0120	0.0122	0.0123
v68	0.0240	0.0243	0.0247	0.0250	0.0254
v69	0.0121	0.0123	0.0125	0.0127	0.0128
v70	0.0119	0.0120	0.0122	0.0124	0.0126
v71	0.0008	0.0008	0.0008	0.0008	0.0008
v72	0.0005	0.0005	0.0005	0.0006	0.0006
v73	0.0005	0.0005	0.0005	0.0006	0.0006
v74	0.0008	0.0008	0.0008	0.0008	0.0008
v75	0.0008	0.0008	0.0008	0.0008	0.0008
v76	0.0001	0.0001	0.0001	0.0001	0.0001

v77	0.0004	0.0004	0.0004	0.0004	0.0004
v78	0.0010	0.0010	0.0011	0.0011	0.0011
v79	0.0014	0.0014	0.0014	0.0014	0.0015
v80	0.0012	0.0012	0.0013	0.0013	0.0013
v81	0.0015	0.0015	0.0015	0.0015	0.0016
v82	N/A	N/A	N/A	N/A	N/A
v83	N/A	N/A	N/A	N/A	N/A
v84	N/A	N/A	N/A	N/A	N/A
v85	N/A	N/A	N/A	N/A	N/A
v86	N/A	N/A	N/A	N/A	N/A
v87	N/A	N/A	N/A	N/A	N/A
v88	N/A	N/A	N/A	N/A	N/A
v89	0.0051	0.0052	0.0053	0.0053	0.0054
v90	0.1153	0.1170	0.1187	0.1204	0.1220
v91	0.2522	0.2558	0.2595	0.2632	0.2668
v92	0.1795	0.1821	0.1848	0.1874	0.1900
v93	0.8934	2.9314	5.3805	8.2626	11.6010
v94	21.2400	24.0450	27.3100	31.0100	35.1070

Xylose Aerobic

	95 lb	68 lb	best fit	68 ub	95 ub
v1	9.0276	9.4766	10.0010	10.7110	11.3940
v2	11.1580	11.7680	12.5080	13.6320	14.7110
v3	0.1827	0.2493	0.3177	0.3864	0.4530
v4	0.4391	0.5933	0.7530	0.9086	1.0561
v5	0.2125	0.2219	0.2317	0.2412	0.2501
v6	0.7886	0.9370	1.0000	1.0000	1.0000
v7	0.0000	0.0000	0.0000	0.0630	0.2114
v8	1.0000	1.0000	1.0000	1.0000	1.0000
v9 net	-0.8843	-0.7006	-0.5216	-0.4360	-0.3810
v9 exch	0.0000	0.0000	0.1182	Inf	Inf
v10 net	5.3685	5.6486	5.9724	6.3962	6.8029
v10 exch	0.6217	0.6913	3.0417	Inf	Inf
v11 net	5.3685	5.6486	5.9724	6.3962	6.8029
v11 exch	0.6216	0.6911	3.0538	Inf	Inf
v12 net	4.7027	4.9325	5.2136	5.6452	6.0587
v12 exch	34.4100	36.5080	38.9510	42.0520	45.1160
v13	13.0870	13.7310	14.5020	15.6280	16.7100
v14	12.9860	13.6240	14.3950	15.5250	16.5990
v15	12.8690	13.5060	14.2720	15.3970	16.4780
v16	0.4448	0.5991	0.7588	0.9144	1.0619
v17	0.4391	0.5933	0.7530	0.9086	1.0561
v18	0.0000	0.0407	0.1264	0.3131	0.4988
v19 net	3.0948	3.2544	3.4465	3.7353	4.0046
v19 exch	27.3740	116.9400	6.9037E+04	Inf	Inf
v20 net	-3.6343	-3.4796	-3.3201	-3.1681	-3.0263

v20	exch	28.9170	121.8100	9.9998E+06	Inf	Inf
v21	net	5.9857	6.3004	6.6812	7.2523	7.8134
v21	exch	7.4321	7.9486	8.5616	9.2711	9.9569
v22	net	-3.8737	-3.5958	-3.3100	-3.1201	-2.9634
v22	exch	0.0000	0.0000	0.2070	0.9975	1.7541
v23	net	-3.9370	-3.6573	-3.3712	-3.1803	-3.0223
v23	exch	1.2486	2.1117	11.4180	Inf	Inf
v24	net	-3.9370	-3.6573	-3.3712	-3.1803	-3.0223
v24	exch	10.2730	11.7200	13.9140	16.5720	19.3410
v25	net	3.0223	3.1803	3.3712	3.6573	3.9370
v25	exch	1.2487	2.1133	6.0605	Inf	Inf
v26		0.0150	0.0310	0.0808	0.1165	0.1376
v27		0.4757	0.5170	0.6052	0.7222	0.8123
v28		11.3910	12.0020	12.7460	14.0710	15.1540
v29		11.1580	11.7680	12.5080	13.6320	14.7110
v30		0.1694	0.2359	0.3042	0.3730	0.4395
v31		0.0000	0.0000	0.0000	0.4257	0.6208
v32		0.0000	0.1631	0.5972	0.6224	0.6465
v33		0.2983	0.3125	0.3326	0.3476	0.3622
v34		0.2983	0.3125	0.3326	0.3476	0.3622
v35		0.2983	0.3125	0.3326	0.3476	0.3622
v36		0.0773	0.0846	0.0965	0.1040	0.1117
v37	net	0.0789	0.0861	0.0981	0.1056	0.1134
v37	exch	0.0000	0.0000	0.4029	Inf	Inf
v38	net	0.1647	0.1737	0.1884	0.1984	0.2084
v38	exch	0.1990	0.2478	0.5073	Inf	Inf
v39	net	0.0558	0.0742	0.1076	0.1455	0.1687
v39	exch	0.1988	0.2476	1.3747	Inf	Inf
v40		0.0385	0.0603	0.1079	0.1347	0.1465
v41	net	-0.0136	0.0056	0.0540	0.0807	0.0892
v41	exch	0.0117	0.0182	0.0333	0.0418	0.0462
v42	net	-0.0025	0.0019	0.0284	0.0771	0.1014
v42	exch	0.0000	0.0000	0.0000	0.0010	0.0025
v43		0.2748	0.4565	0.8600	0.9276	0.9793
v44	net	0.1540	0.1742	0.2250	0.2628	0.2860
v44	exch	0.0666	0.0840	0.1083	0.1338	0.1577
v45	net	-0.2116	-0.2040	-0.1960	0.2294	0.4140
v45	exch	0.0000	0.0000	0.0000	Inf	Inf
v46		1.0688	1.1236	1.1789	1.2438	1.3274
v47		0.1927	0.2012	0.2101	0.2187	0.2268
v48		0.0351	0.0366	0.0382	0.0398	0.0413
v49		0.0342	0.0357	0.0373	0.0388	0.0403
v50		0.2989	0.3259	0.3803	0.4788	0.5473
v51		0.0217	0.0226	0.0236	0.0246	0.0255
v52		0.0059	0.0395	0.0675	0.0923	0.1116
v53		0.0000	0.0144	0.0388	0.0671	0.1009

v54	0.0895	0.1024	0.1556	0.2513	0.3080
v55	0.0047	0.0135	0.0666	0.1629	0.2132
v56	0.0015	0.0016	0.0016	0.0017	0.0018
v57	0.0108	0.0113	0.0118	0.0123	0.0128
v58	0.0608	0.0635	0.0663	0.0690	0.0715
v59	0.0629	0.0657	0.0686	0.0714	0.0740
v60	0.0410	0.0428	0.0447	0.0465	0.0483
v61	0.0563	0.0588	0.0614	0.0639	0.0663
v62	0.0285	0.0297	0.0311	0.0323	0.0335
v63	0.0217	0.0226	0.0236	0.0246	0.0255
v64	0.0059	0.0062	0.0065	0.0068	0.0070
v65	0.0140	0.0146	0.0153	0.0159	0.0165
v66	0.0261	0.0273	0.0285	0.0297	0.0308
v67	0.0111	0.0116	0.0121	0.0126	0.0131
v68	0.0229	0.0240	0.0250	0.0260	0.0270
v69	0.0116	0.0121	0.0127	0.0132	0.0137
v70	0.0113	0.0119	0.0124	0.0129	0.0134
v71	0.0008	0.0008	0.0008	0.0009	0.0009
v72	0.0005	0.0005	0.0006	0.0006	0.0006
v73	0.0005	0.0005	0.0006	0.0006	0.0006
v74	0.0008	0.0008	0.0008	0.0009	0.0009
v75	0.0008	0.0008	0.0008	0.0009	0.0009
v76	0.0001	0.0001	0.0001	0.0001	0.0002
v77	0.0004	0.0004	0.0004	0.0004	0.0004
v78	0.0010	0.0010	0.0011	0.0011	0.0012
v79	0.0013	0.0014	0.0014	0.0015	0.0016
v80	0.0012	0.0012	0.0013	0.0013	0.0014
v81	0.0014	0.0015	0.0015	0.0016	0.0017
v82	0.0028	0.0029	0.0031	0.0032	0.0033
v83	0.0028	0.0029	0.0031	0.0032	0.0033
v84	0.0028	0.0029	0.0031	0.0032	0.0033
v85	0.0028	0.0029	0.0031	0.0032	0.0033
v86	0.0009	0.0010	0.0010	0.0011	0.0011
v87	0.0005	0.0005	0.0005	0.0005	0.0006
v88	0.0001	0.0002	0.0002	0.0002	0.0002
v89	0.0049	0.0051	0.0053	0.0055	0.0058
v90	0.1103	0.1152	0.1203	0.1252	0.1298
v91	0.2412	0.2519	0.2630	0.2737	0.2839
v92	0.1717	0.1793	0.1872	0.1949	0.2021
v93	0.0000	0.0000	0.0311	0.3088	0.5826
v94	12.2330	12.8950	13.7490	15.2780	16.7480

Xylose Anaerobic

	95 lb	68 lb	best fit	68 ub	95 ub
v1	11.6190	12.0940	12.5940	13.1040	13.6100

v2	15.0630	15.7540	16.4870	17.2580	18.0050
v3	0.1261	0.1776	0.2318	0.2855	0.3377
v4	1.7101	1.8737	2.0458	2.2182	2.3853
v5	0.1641	0.1690	0.1746	0.1801	0.1854
v6	0.4266	0.8017	1.0000	1.0000	1.0000
v7	0.0000	0.0000	0.0000	0.1983	0.5734
v8	1.0000	1.0000	1.0000	1.0000	1.0000
v9 net	-0.5348	-0.4368	-0.3767	-0.3027	-0.2832
v9 exch	0.0000	0.0000	1.1836	Inf	Inf
v10 net	7.2538	7.5576	7.8776	8.1965	8.5298
v10 exch	1.0405	1.1311	3.0554	Inf	Inf
v11 net	7.2538	7.5576	7.8776	8.1965	8.5298
v11 exch	1.0405	1.1325	7.3811	Inf	Inf
v12 net	5.2949	5.5514	5.8274	6.0589	6.3982
v12 exch	67.1100	70.9560	75.1090	79.5700	84.2000
v13	16.4380	17.1420	17.8850	18.5190	19.4140
v14	16.3340	17.0370	17.7790	18.1980	19.3080
v15	16.2440	16.9450	17.6860	18.4500	19.2110
v16	1.7112	1.9047	2.0502	2.2163	2.3901
v17	1.7101	1.8737	2.0458	2.2182	2.3853
v18	0.0000	0.0045	0.0790	0.1533	0.2380
v19 net	3.9312	4.0996	4.2775	4.4585	4.6416
v19 exch	34.7700	143.9700	1.0000E+07	Inf	Inf
v20 net	-4.5112	-4.3559	-4.1985	-4.0435	-3.8954
v20 exch	37.4260	152.9800	1.0000E+07	Inf	Inf
v21 net	7.7045	8.0411	8.3954	8.7591	9.1211
v21 exch	12.4610	13.1980	13.9990	14.8510	15.7140
v22 net	-4.5371	-4.3563	-4.1746	-3.9977	-3.8304
v22 exch	0.0000	0.0000	0.0000	0.1910	0.5892
v23 net	-4.5839	-4.4028	-4.2207	-4.0434	-3.8756
v23 exch	2.1365	2.7527	13.1480	Inf	Inf
v24 net	-4.5839	-4.4028	-4.2207	-4.0434	-3.8756
v24 exch	10.4500	12.1070	13.8170	15.9540	18.4010
v25 net	3.8756	4.0434	4.2207	4.4028	4.5839
v25 exch	2.1383	2.7526	5.7819	Inf	Inf
v26	0.0061	0.0114	0.0315	0.2377	0.3343
v27	0.3378	0.3529	0.3772	0.5868	0.7176
v28	15.2900	15.9840	16.8060	17.7960	18.5450
v29	15.0630	15.7540	16.4870	17.2580	18.0050
v30	0.1159	0.1680	0.2217	0.2754	0.3275
v31	0.0000	0.0000	0.0980	0.3389	0.3484
v32	0.0000	0.0000	0.2296	0.3381	0.3486
v33	0.1669	0.1722	0.1779	0.1842	0.1903
v34	0.1669	0.1722	0.1779	0.1842	0.1903
v35	0.1669	0.1722	0.1779	0.1842	0.1903
v36	0.0000	0.0000	0.0000	0.0027	0.0054

v37 net	0.0011	0.0012	0.0012	0.0040	0.0066
v37 exch	0.0000	0.0000	0.2755	Inf	Inf
v38 net	0.0654	0.0671	0.0692	0.0728	0.0763
v38 exch	0.0119	0.0342	1.3611	Inf	Inf
v39 net	-0.2642	-0.1678	0.0377	0.0581	0.0638
v39 exch	0.0119	0.0220	0.0557	Inf	Inf
v40	0.0560	0.0736	0.1060	0.1095	0.1128
v41 net	0.0156	0.0322	0.0654	0.0677	0.0697
v41 exch	0.0259	0.0336	0.0493	0.0514	0.0534
v42 net	-0.0040	-0.0038	-0.0034	0.0296	0.0466
v42 exch	0.0000	0.0000	0.0000	0.0291	0.0622
v43	0.1150	0.1661	0.5027	0.6245	0.6466
v44 net	0.1126	0.1189	0.1402	0.3467	0.4437
v44 exch	0.2912	0.3726	0.5033	1.0255	1.8440
v45 net	-0.1040	-0.1010	0.0000	0.2370	0.2446
v45 exch	0.0000	0.0000	0.0000	Inf	Inf
v46	0.8097	0.8362	0.8634	0.9044	0.9454
v47	0.1487	0.1538	0.1583	0.1633	0.1681
v48	0.0270	0.0279	0.0288	0.0297	0.0306
v49	0.0264	0.0272	0.0281	0.0290	0.0299
v50	0.2219	0.2294	0.2370	0.3046	0.3402
v51	0.0167	0.0172	0.0178	0.0184	0.0189
v52	0.0000	0.0000	0.0050	0.0423	0.0639
v53	0.0164	0.0379	0.0751	0.0826	0.0851
v54	0.0630	0.0651	0.0677	0.1339	0.1684
v55	0.0000	0.0000	0.0007	0.0667	0.1008
v56	0.0012	0.0012	0.0012	0.0013	0.0013
v57	0.0084	0.0086	0.0089	0.0092	0.0095
v58	0.0468	0.0486	0.0499	0.0515	0.0530
v59	0.0485	0.0501	0.0517	0.0533	0.0549
v60	0.0317	0.0326	0.0337	0.0348	0.0358
v61	0.0434	0.0448	0.0463	0.0477	0.0491
v62	0.0219	0.0227	0.0234	0.0241	0.0248
v63	0.0168	0.0172	0.0178	0.0184	0.0189
v64	0.0046	0.0047	0.0049	0.0050	0.0052
v65	0.0109	0.0112	0.0115	0.0119	0.0122
v66	0.0203	0.0208	0.0215	0.0221	0.0228
v67	0.0086	0.0089	0.0091	0.0094	0.0097
v68	0.0178	0.0183	0.0189	0.0194	0.0200
v69	0.0089	0.0092	0.0095	0.0098	0.0101
v70	0.0088	0.0090	0.0093	0.0096	0.0099
v71	0.0006	0.0006	0.0006	0.0006	0.0007
v72	0.0004	0.0004	0.0004	0.0004	0.0004
v73	0.0004	0.0004	0.0004	0.0004	0.0004
v74	0.0006	0.0006	0.0006	0.0006	0.0007
v75	0.0006	0.0006	0.0006	0.0006	0.0007

v76	0.0001	0.0001	0.0001	0.0001	0.0001
v77	0.0003	0.0003	0.0003	0.0003	0.0003
v78	0.0008	0.0008	0.0008	0.0008	0.0009
v79	0.0010	0.0010	0.0011	0.0011	0.0011
v80	0.0009	0.0009	0.0010	0.0010	0.0010
v81	0.0011	0.0011	0.0012	0.0012	0.0012
v82	N/A	N/A	N/A	N/A	N/A
v83	N/A	N/A	N/A	N/A	N/A
v84	N/A	N/A	N/A	N/A	N/A
v85	N/A	N/A	N/A	N/A	N/A
v86	N/A	N/A	N/A	N/A	N/A
v87	N/A	N/A	N/A	N/A	N/A
v88	N/A	N/A	N/A	N/A	N/A
v89	0.0038	0.0039	0.0040	0.0041	0.0043
v90	0.0852	0.0882	0.0906	0.0935	0.0962
v91	0.1861	0.1923	0.1981	0.2044	0.2105
v92	0.1328	0.1374	0.1411	0.1455	0.1498
v93	0.0000	0.0000	0.0000	0.0438	0.1549
v94	15.7020	16.4090	17.1780	17.9520	18.7810

Appendix B

Yarrowia lipolytica Glucose Flux Estimation Details

This Appendix contains details of the ^{13}C -MFA of *Y. lipolytica* on glucose, including figures showing the linear regressions for Estimation of Extracellular Fluxes (Figures B-1-B-6), the Model Metabolic Reaction Network (Table B-1), measured and simulated ^{13}C labeling data from the flux estimations (Table B-2), and the best-fit flux values and 68% and 95% confidence intervals for the estimated fluxes for each strain (Table B-3).

In all Figures B-1-B-6, data from only the final three time points were used for the citrate linear regression. All other linear regressions were conducted with data from the final six time points.

Figure B-1 Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL037, natural abundance glucose culture.

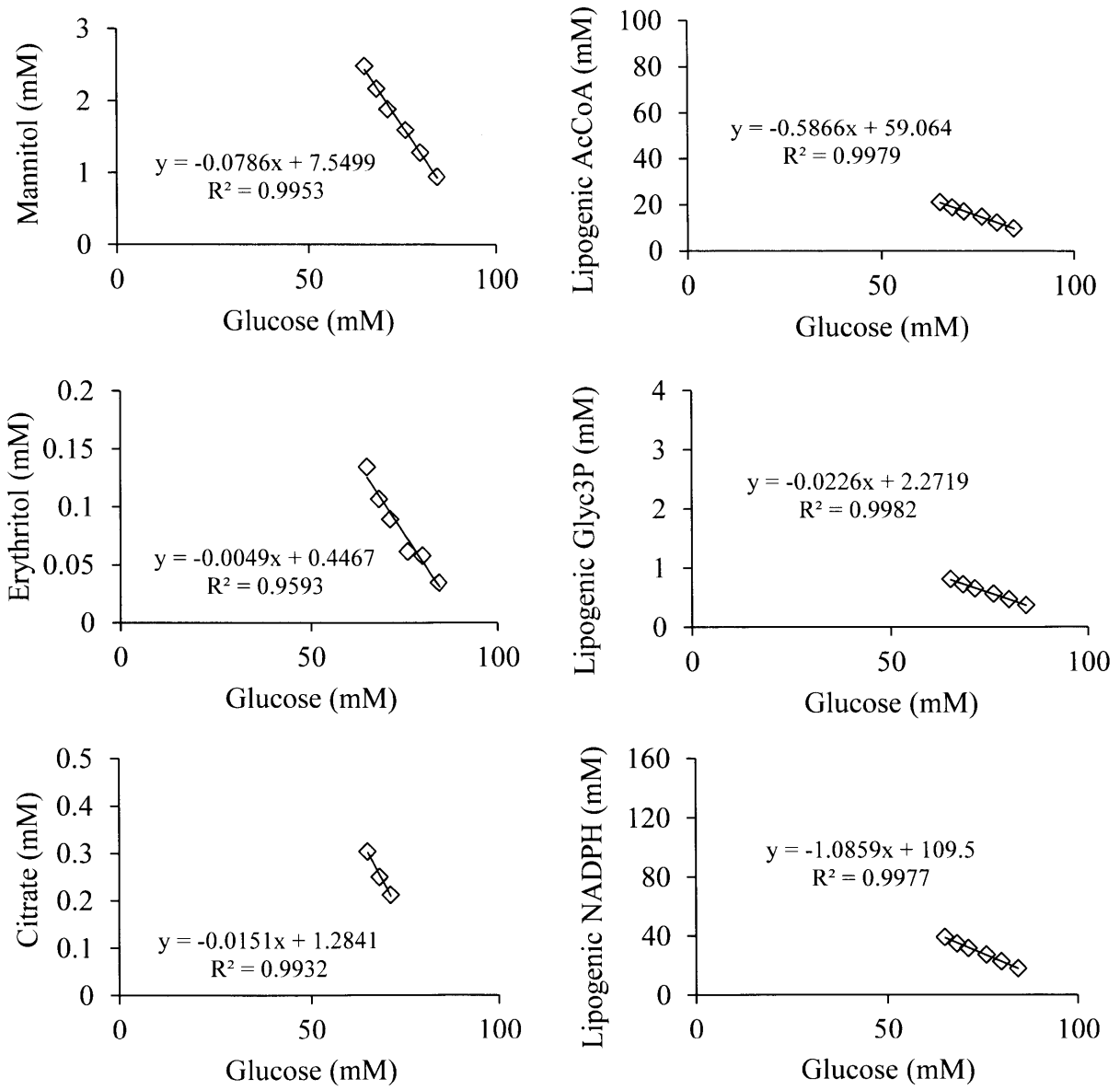


Figure B-2 Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL037, 20% U-¹³C₆-glucose culture.

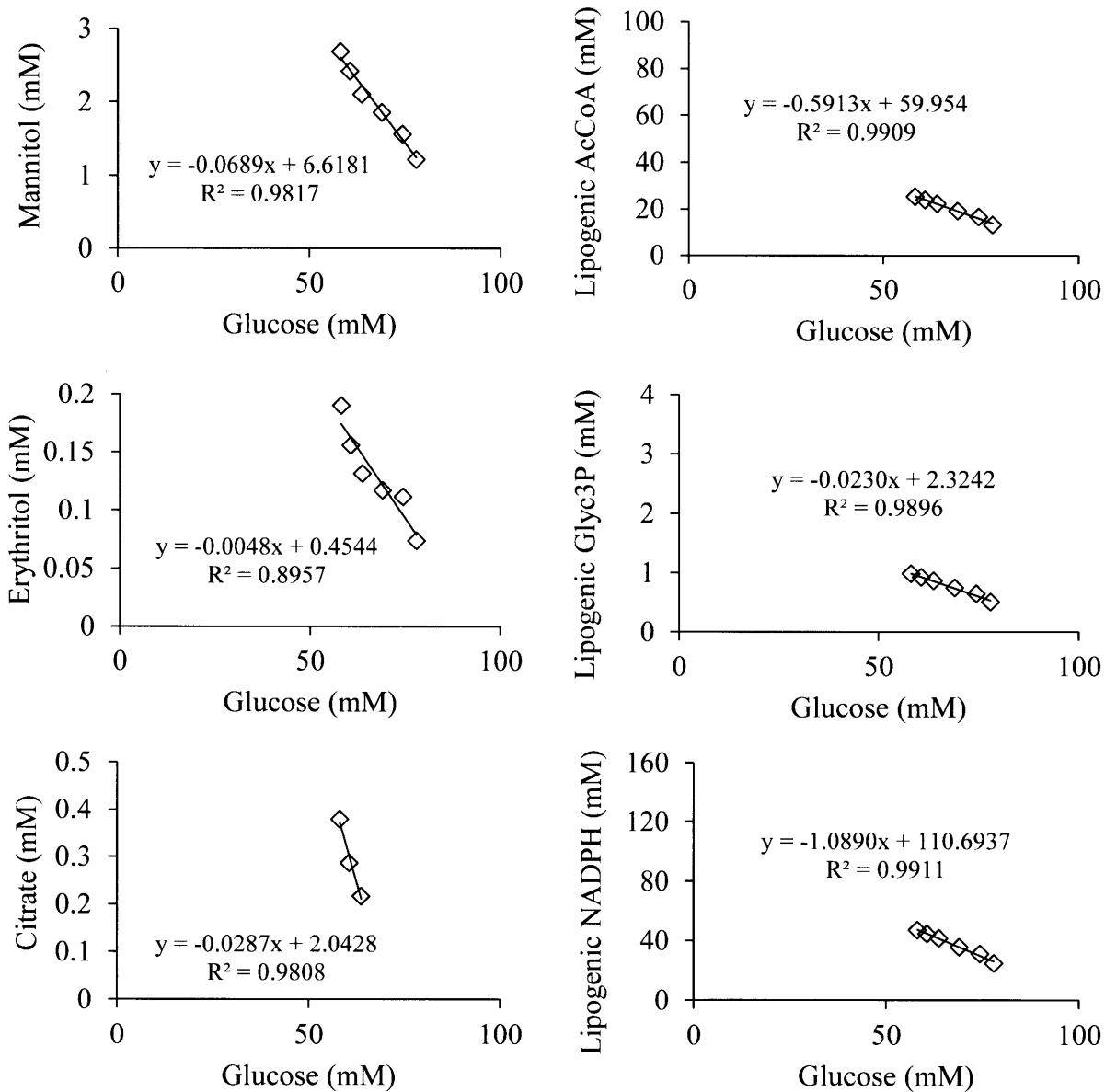


Figure B-3 Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL037, 1,2-¹³C₂-glucose culture.

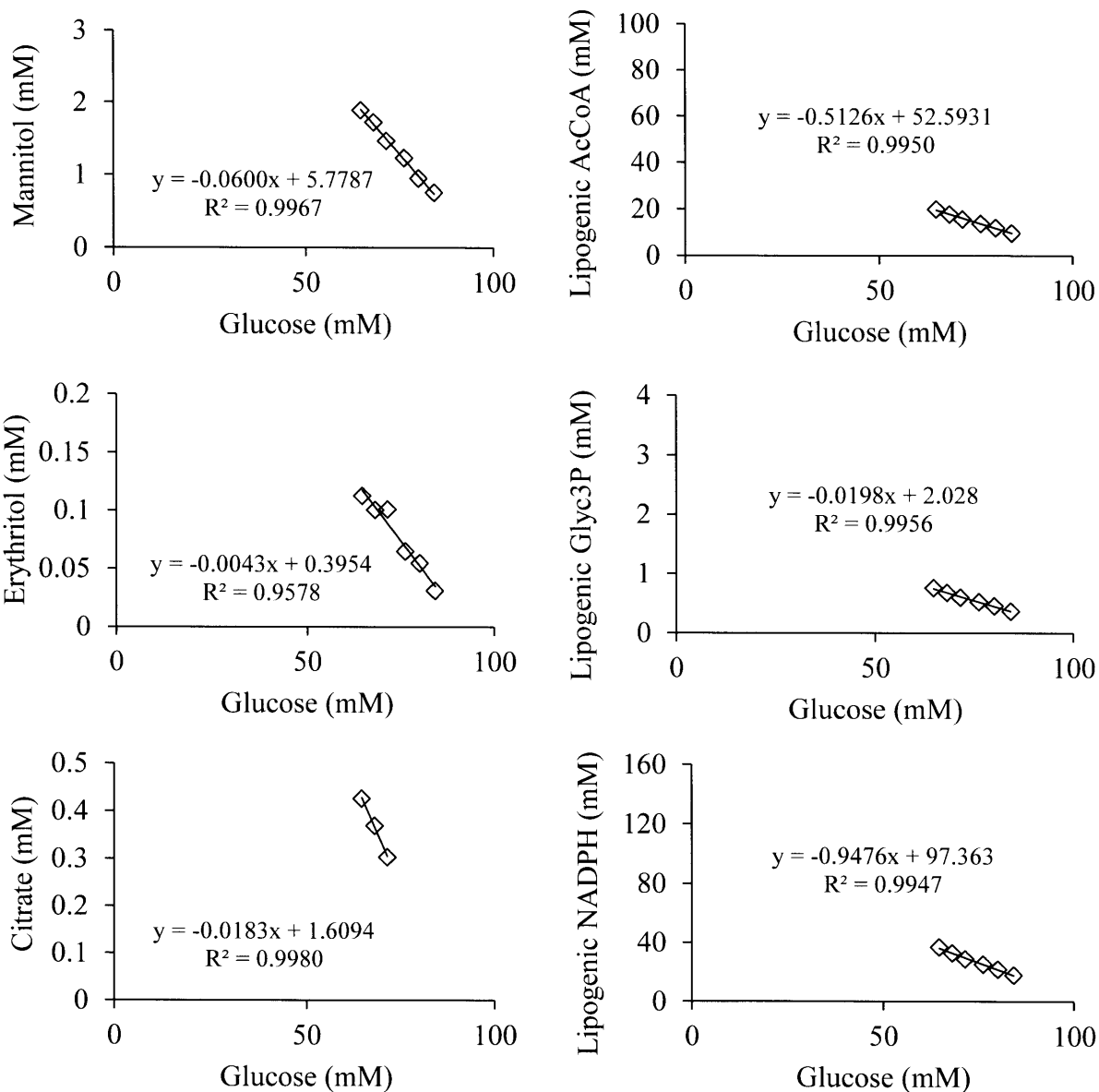


Figure B-4 Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL065, natural abundance glucose culture.

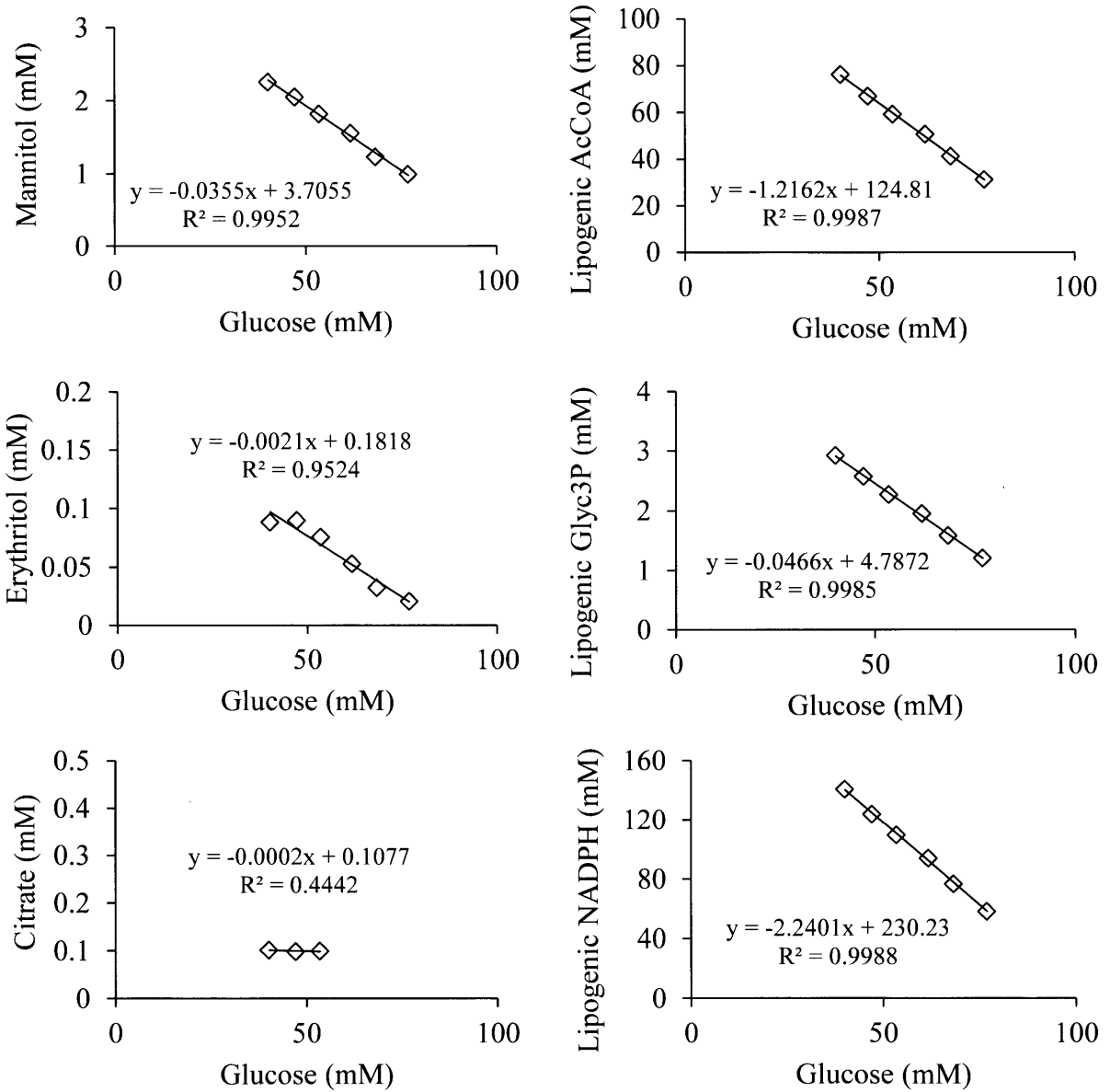


Figure B-5 Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL065, 20% U-¹³C₆-glucose culture.

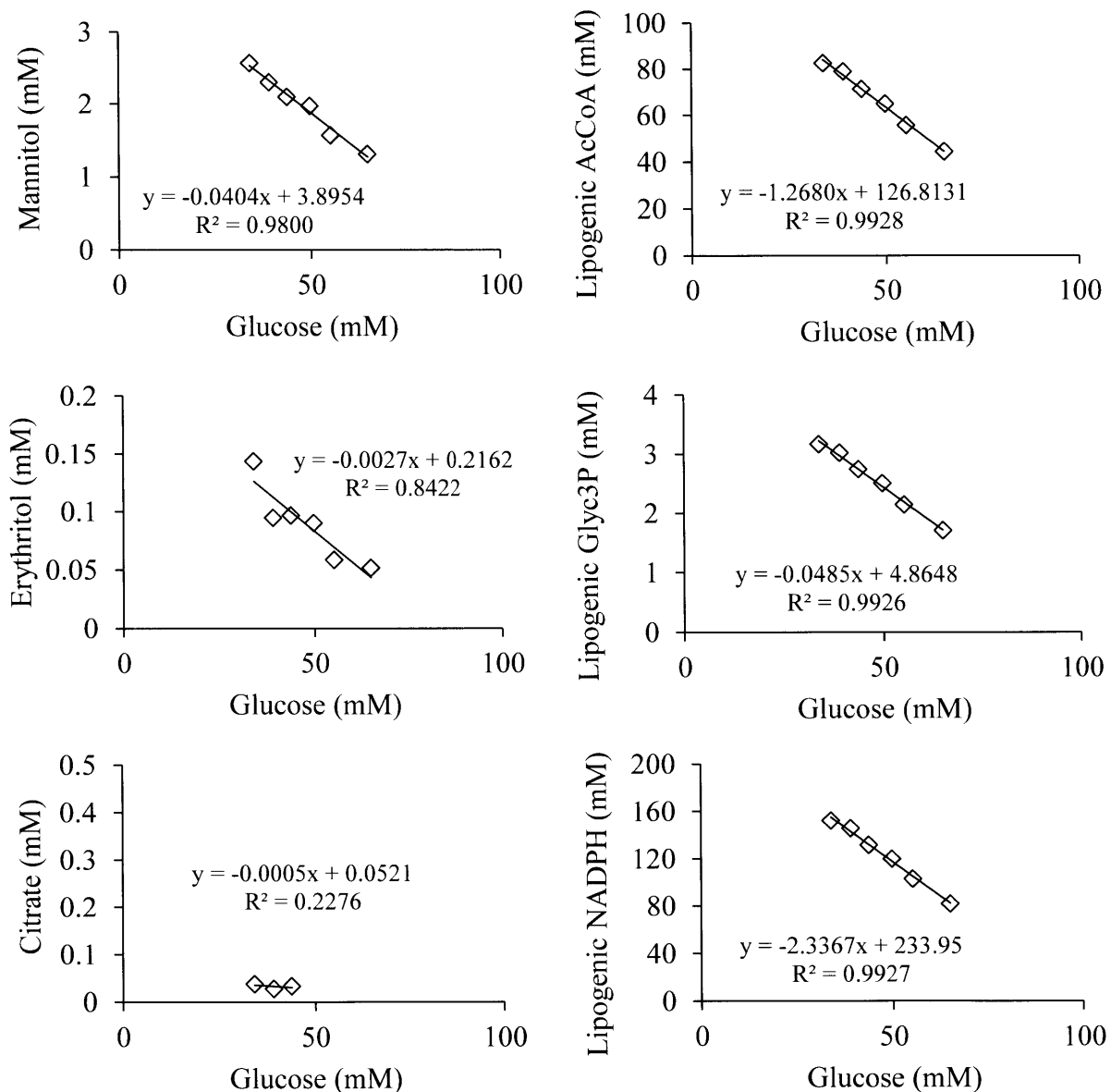


Figure B-6 Linear regressions for Estimation of Extracellular Fluxes: Strain MTYL065, 1,2-¹³C₂-glucose culture.

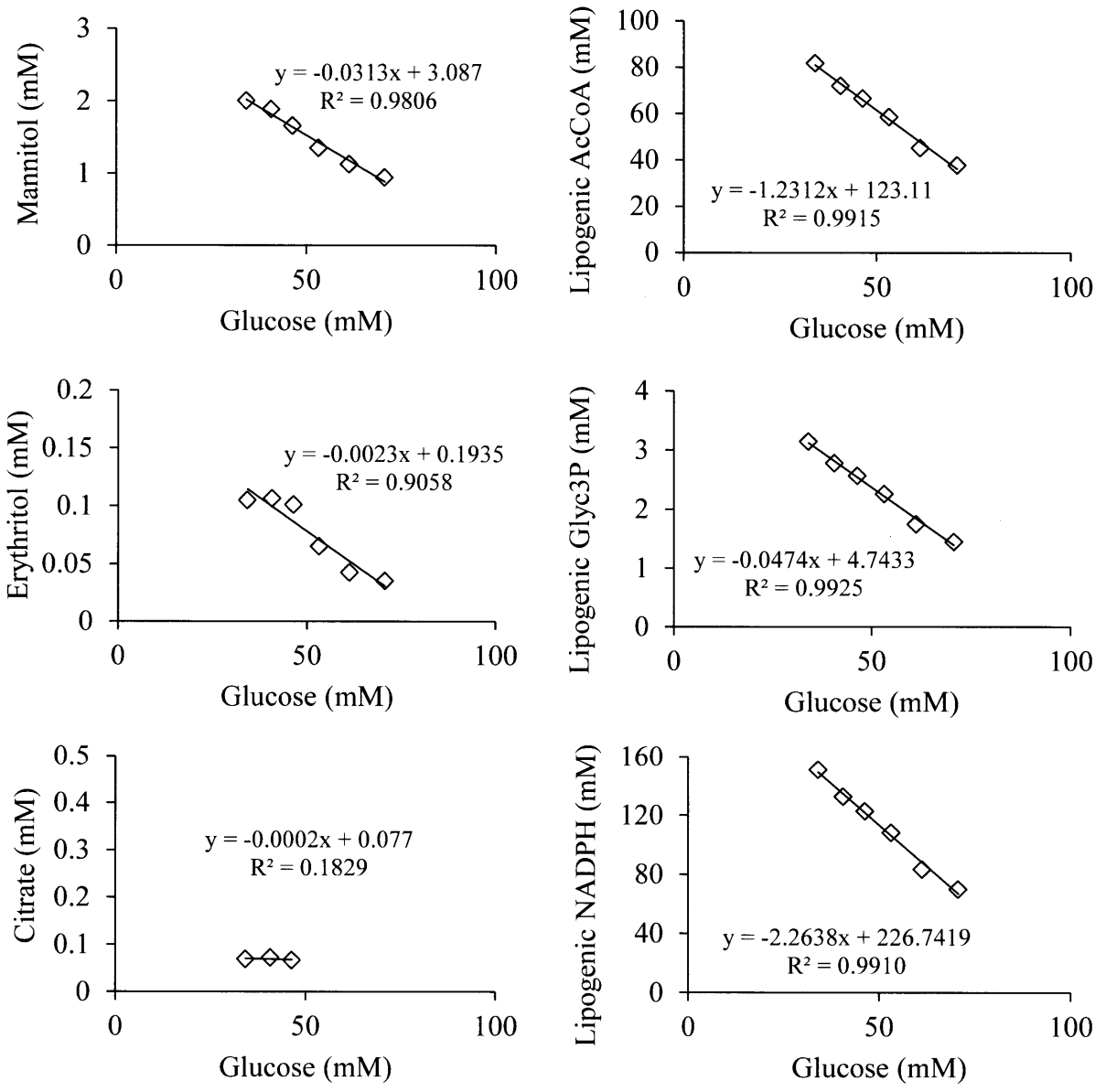


Table B-1: Model Metabolic Reaction Network for ¹³C-MFA

Extracellular fluxes

v1: Glc.ext (abcdef) -> G6P (abcdef)
v2: AcCoA.c (ab) -> Lipids (ab)
v3: Glyc3P (abc) -> TAGs (abc)
v4: F6P (abcdef) -> Mannitol.ext (abcdef)
v5: E4P (abcd) -> Erythritol.ext (abcd)
v6: Cit.c (abcdef) -> Cit.ext (abcdef)

Pyruvate Compartmentalization

v7: 0 Pyr.c (abc) -> Pyr.mnt (abc)
v8: 0 Pyr.m (abc) -> Pyr.mnt (abc)
v9: Pyr.mnt (abc) -> Pyr.fix (abc)

Oxaloacetate Compartmentalization

v10: 0 OAA.c (abcd) -> Asp (abcd)
v11: 0 OAA.m (abcd) -> Asp (abcd)
v12: Asp (abcd) -> Asp.fix (abcd)

Malate Compartmentalization

v13: 0 Mal.c (abcd) -> Mal.mnt (abcd)
v14: 0 Mal.m (abcd) -> Mal.mnt (abcd)
v15: Mal.mnt (abcd) -> Mal.fix (abcd)

Glycolysis

v16: G6P (abcdef) <-> F6P (abcdef)
v17: F6P (abcdef) <-> FBP (abcdef)
v18: FBP (abcdef) <-> DHAP (cba) + G3P (def)
v19: DHAP (abc) <-> G3P (abc)
v20: DHAP (abc) -> Glyc3P (abc)
v21: G3P (abc) -> PG3 (abc)
v22: PG3 (abc) -> PEP (abc)
v23: PEP (abc) -> Pyr.c (abc)

Pentose Phosphate Pathway

v24: G6P (abcdef) -> PG6 (abcdef)
v25: PG6 (abcdef) -> CO2 (a) + Ru5P (bcdef)
v26: Ru5P (abcde) <-> R5P (abcde)
v27: Ru5P (abcde) <-> X5P (abcde)
v28: X5P (abcde) <-> G3P (cde) + TKC2 (ab)
v29: F6P (abcdef) <-> E4P (cdef) + TKC2 (ab)
v30: S7P (abcdefg) <-> R5P (cdefg) + TKC2 (ab)
v31: F6P (abcdef) <-> G3P (def) + TAC3 (abc)
v32: S7P (abcdefg) <-> E4P (defg) + TAC3 (abc)

Pyruvate Metabolism

v33: Mal.m (abcd) -> Pyr.m (abc) + CO2 (d)

v34: Pyr.c (abc) + CO2 (d) -> OAA.c (abcd)

Citric Acid Cycle

v35: Pyr.m (abc) -> AcCoA.m (bc) + CO2 (a)

v36: OAA.m (abcd) + AcCoA.m (ef) -> Cit.m (dcbfea)

v37: Cit.m (abcdef) -> ICit (abcdef)

v38: ICit (abcdef) -> AKG (abcde) + CO2 (f)

v39: AKG (abcde) -> Suc (bcde) + CO2 (a)

v40: Suc (abcd) <-> Fum (abcd)

v41: Fum (abcd) <-> Mal.m (abcd)

v42: Mal.m (abcd) <-> OAA.m (abcd)

Lipid Metabolism

v43: Cit.c (abcdef) -> AcCoA.c (ed) + OAA.c (fcba)

v44: OAA.c (abcd) <-> Mal.c (abcd)

Transport Reactions

v45: Pyr.c (abc) -> Pyr.m (abc)

v46: Cit.m (abcdef) + Mal.c (ghij) <-> Cit.c (abcdef) +
Mal.m (ghij)

v47: Mal.c (abcd) -> Mal.m (abcd)

CO2 Exchange

v48: CO2.ext (a) -> CO2 (a)

v49: CO2 (a) -> CO2.out (a)

Table B-2: Measured and Simulated Mass Spectrometry Data

Notation:

Fragment: the ion measured by the mass spectrometer. Each ion is denoted by a metabolite abbreviation followed by an underscore and the mass of the unlabeled M+0 isotopomer.

Mass Isotopomer: each mass isotopomer is denoted by its mass relative to the mass of the unlabeled parent ion M.

Measured: mass isotopomer mole fractions measured by mass spectrometry in ^{13}C labeling experiments

Simulated: mass isotopomer mole fractions obtained by simulation of best-fit flux models

1,2- $^{13}\text{C}_2$ -Glucose Tracer Data

Fragment	Mass Isotopomer	MTYL037		MTYL065	
		Measured	Simulated	Measured	Simulated
Pyr_174 (GC-MS)	M+0	0.4334135	0.4305397	0.5099935	0.5057099
	M+1	0.193917	0.1873048	0.202911	0.2004867
	M+2	0.273371	0.282922	0.225299	0.2305221
	M+3	0.0800355	0.0807653	0.0484605	0.0506976
	M+4	0.015334	0.0156145	0.0109205	0.0109781
	M+5	0.003929	0.0026747	0.0024155	0.0014965
Suc_289 (GC-MS)	M+0	0.1845895	0.1821514	0.2771105	0.2681379
	M+1	0.24882	0.2502991	0.273488	0.2777015
	M+2	0.278578	0.2826118	0.25676	0.2586127
	M+3	0.172688	0.1735963	0.123098	0.1261701
	M+4	0.0879415	0.0841185	0.0537265	0.0531088
	M+5	0.021378	0.0208051	0.012499	0.012667
	M+6	0.0060055	0.005494	0.003318	0.0030902
Fum_287 (GC-MS)	M+0	0.2089005	0.2066756	0.279667	0.2781334
	M+1	0.2482925	0.2496817	0.2864955	0.2890164
	M+2	0.272158	0.2745772	0.24018	0.2423628
	M+3	0.1673225	0.1684187	0.1284675	0.1291682
	M+4	0.0766595	0.0759346	0.048558	0.0465323

	M+5	0.020807	0.0190791	0.013272	0.0117951
	M+6	0.00586	0.0048191	0.003359	0.0025587
Akg_346	M+0	0.1442655	0.1363935	0.214118	0.2098271
(GC-MS)	M+1	0.2008875	0.2035151	0.2485165	0.2512333
	M+2	0.2707555	0.2713611	0.2652395	0.2643459
	M+3	0.199383	0.2031405	0.1586735	0.1605045
	M+4	0.119583	0.1202095	0.078284	0.0782937
	M+5	0.0491805	0.0489065	0.0268545	0.0271711
	M+6	0.0128075	0.0128046	0.0067575	0.0068525
	M+7	0.0031375	0.0031064	0.0015555	0.0015046
Mal_391	M+0	0.233769	0.230009	0.3045145	0.2986862
(GC-MS)	M+1	0.275696	0.274836	0.2923545	0.2916227
	M+2	0.270664	0.271215	0.2390105	0.2396995
	M+3	0.15068	0.1502973	0.1139	0.1157151
	M+4	0.0529095	0.0529298	0.03912	0.0397814
	M+5	0.016281	0.016355	0.0111005	0.0114912
Mal_419	M+0	0.1875045	0.1852932	0.246841	0.2398272
(GC-MS)	M+1	0.236487	0.2365676	0.27614	0.277905
	M+2	0.267181	0.2667473	0.2477275	0.2467794
	M+3	0.179851	0.1803548	0.1446575	0.1468582
	M+4	0.090387	0.090013	0.0602755	0.0620112
	M+5	0.029914	0.0298898	0.0192505	0.0199521
	M+6	0.008675	0.0088207	0.005108	0.0053155
Asp_390	M+0	0.2262245	0.2259281	0.2946815	0.294979
(GC-MS)	M+1	0.283301	0.2773753	0.301289	0.2939379
	M+2	0.2807085	0.2726401	0.2469775	0.2394359
	M+3	0.1554565	0.1506987	0.117088	0.1171137
	M+4	0.05431	0.0528968	0.0399645	0.0399865
Asp_418	M+0	0.1799635	0.177432	0.238614	0.2402367
(GC-MS)	M+1	0.241259	0.2361397	0.282144	0.2792848
	M+2	0.272953	0.2690345	0.252508	0.245813
	M+3	0.1839985	0.1824301	0.14699	0.1470478
	M+4	0.0917145	0.0927913	0.0605695	0.0614147
	M+5	0.0301125	0.0306996	0.0191745	0.0197182
	M+6	0	0.0091049	0	0.0051836
Glu_330	M+0	0.1681	0.1656708	0.2497845	0.2542147
(GC-MS)	M+1	0.241953	0.247751	0.2740205	0.2753265
	M+2	0.284123	0.2843453	0.264314	0.2632726
	M+3	0.180906	0.1803603	0.1340925	0.1312856
	M+4	0.0936465	0.0905864	0.059151	0.0573497
	M+5	0.024785	0.0239635	0.0149895	0.0144365
	M+6	0.0064865	0.0061931	0.0036475	0.0034937
Glu_432	M+0	0.123247	0.1191538	0.183875	0.1833056

(GC-MS)	M+1	0.1866385	0.1905804	0.235134	0.2391526	
	M+2	0.258221	0.26039	0.2600645	0.2610215	
	M+3	0.206801	0.2094521	0.17257	0.1731453	
	M+4	0.1345955	0.132818	0.0938585	0.0920553	
	M+5	0.063161	0.0607299	0.0377755	0.0364723	
	M+6	0.0211295	0.0198176	0.0130045	0.0112098	
	M+7	0.006207	0.0056152	0.003718	0.002918	
PEP_453	M+0	0.3365535	0.3518211	0.3888835	0.4009206	
(GC-MS)	M+1	0.2024435	0.1930897	0.2266845	0.2194095	
	M+2	0.276073	0.2742369	0.246643	0.2426196	
	M+3	0.121471	0.1180129	0.0918095	0.0907378	
	M+4	0.049563	0.0463725	0.0365105	0.0350247	
	M+5	0.0138965	0.0128627	0.009468	0.008851	
DHAP_484	M+0	0.3346705	0.3333382	0.363185	0.3609924	
	(GC-MS)	M+1	0.190807	0.1939868	0.2175215	0.2213432
		M+2	0.2882765	0.2806326	0.268054	0.2615241
		M+3	0.123375	0.1247534	0.1007425	0.1026753
		M+4	0.04954	0.0493738	0.0403515	0.0401227
M+5		0.013331	0.0139456	0.010146	0.0104266	
G6P_259/97	M+0	0.059691	0.0664414	0.043647	0.0469405	
	(LC-MS/MS)	M+1	0.106857	0.1135323	0.150944	0.1508249
		M+2	0.554517	0.5556232	0.650028	0.6589386
		M+3	0.131526	0.1313863	0.091488	0.0874554
		M+4	0.104586	0.0942167	0.051593	0.0435724
		M+5	0.037298	0.0339825	0.010519	0.0110024
		M+6	0.005085	0.0044259	0.001674	0.0011467
M+7		0.000439	0.0003508	0.000107	0.0001094	
6PG_275/79	M+0	0.063118	0.0660995	0.040422	0.046699	
	(LC-MS/MS)	M+1	0.123482	0.1130193	0.152359	0.1500991
		M+2	0.543208	0.5531563	0.654194	0.6559007
		M+3	0.129187	0.1317661	0.091083	0.0883233
		M+4	0.100011	0.0961339	0.050273	0.0461237
		M+5	0.035529	0.034445	0.010432	0.0113507
		M+6	0.004914	0.0048255	0.001237	0.0013329
M+7		0.000552	0.0004928	0	0.0001553	
3PG_185/79	M+0	0.524606	0.5275214	0.600368	0.6011414	
	(LC-MS/MS)	M+1	0.129118	0.1252467	0.14338	0.1417851
		M+2	0.296649	0.2939157	0.233303	0.2304347
		M+3	0.047092	0.0503985	0.021386	0.0245221
M+4		0.002535	0.0025022	0.001564	0.001916	

20% U-¹³C₆-Glucose Tracer Data

Fragment	Mass Isotopomer	MTYL037		MTYL065	
		Measured	Simulated	Measured	Simulated
Pyr_174 (GC-MS)	M+0	0.605104	0.6007406	0.62017	0.6165729
	M+1	0.170896	0.1719714	0.1544035	0.1536311
	M+2	0.0929065	0.0959275	0.087879	0.0865668
	M+3	0.112493	0.1144188	0.1191685	0.1250508
	M+4	0.01264	0.0122043	0.012666	0.012923
	M+5	0.005961	0.0044963	0.005713	0.0049903
Suc_289 (GC-MS)	M+0	0.390301	0.3869416	0.4263035	0.4118409
	M+1	0.266926	0.2667217	0.224894	0.2279419
	M+2	0.2148225	0.2116514	0.223714	0.2226046
	M+3	0.086049	0.0898011	0.078348	0.0857889
	M+4	0.032775	0.0346892	0.0370535	0.0403062
	M+5	0.0072535	0.0080285	0.00737	0.0088161
	M+6	0.0018735	0.0018621	0.002317	0.0023272
Fum_287 (GC-MS)	M+0	0.3929975	0.3994261	0.4203605	0.4254367
	M+1	0.2685345	0.263682	0.2483915	0.2397058
	M+2	0.186481	0.189698	0.1662115	0.1707052
	M+3	0.0980335	0.0978017	0.104687	0.1047587
	M+4	0.0401865	0.0374862	0.0448935	0.0445665
	M+5	0.0108595	0.0093643	0.0121545	0.0114784
	M+6	0.0029075	0.0021747	0.0033005	0.0028602
Akg_346 (GC-MS)	M+0	0.336044	0.3312469	0.35687	0.3545592
	M+1	0.2499855	0.2519023	0.22371	0.2237706
	M+2	0.2269995	0.2271131	0.2246835	0.2238888
	M+3	0.1158765	0.1174775	0.117842	0.1194501
	M+4	0.048882	0.049404	0.0505835	0.0511864
	M+5	0.017206	0.0175921	0.020253	0.0208325
	M+6	0.004058	0.0041385	0.004724	0.004856
Mal_391 (GC-MS)	M+0	0.394676	0.3936317	0.4142755	0.413107
	M+1	0.2846185	0.2808144	0.2613665	0.2587835
	M+2	0.1926305	0.1932785	0.185476	0.1836873
	M+3	0.090245	0.0913556	0.0969945	0.0989254
	M+4	0.029275	0.0298963	0.0318565	0.032444
	M+5	0.0085545	0.0087648	0.010031	0.0103427
Mal_419 (GC-MS)	M+0	0.3489875	0.3472368	0.369414	0.366721
	M+1	0.269946	0.2671947	0.2563695	0.2506924

	M+2	0.1998945	0.1974455	0.180815	0.1827586
	M+3	0.1125925	0.1154179	0.116725	0.1179662
	M+4	0.0484855	0.0504996	0.053251	0.0560069
	M+5	0.015597	0.0164237	0.0180245	0.0188348
	M+6	0.004497	0.0045918	0.005401	0.005549
Asp_390	M+0	0.388459	0.3911859	0.408101	0.4117988
(GC-MS)	M+1	0.2925415	0.2848803	0.269585	0.2597881
	M+2	0.197284	0.1916219	0.191248	0.185301
	M+3	0.0920025	0.0916735	0.098839	0.0981624
	M+4	0.0297135	0.0297083	0.032228	0.0322308
Asp_418	M+0	0.341404	0.3429088	0.3602035	0.3672203
(GC-MS)	M+1	0.2753855	0.268478	0.26207	0.2526188
	M+2	0.202935	0.2041586	0.184483	0.1788616
	M+3	0.115184	0.1137549	0.1203215	0.118905
	M+4	0.0493575	0.0495126	0.0546	0.0563074
	M+5	0.0157335	0.0157331	0.018322	0.0190473
	M+6	0	0.0043496	0	0.0055727
Glu_330	M+0	0.363428	0.3659083	0.3855725	0.3908759
(GC-MS)	M+1	0.274028	0.2739864	0.2364815	0.2341986
	M+2	0.221283	0.2215103	0.234083	0.2346022
	M+3	0.0940465	0.0923887	0.0899375	0.0876975
	M+4	0.0364935	0.0355095	0.0417585	0.0405461
	M+5	0.008762	0.0084655	0.009733	0.0093272
	M+6	0.001959	0.0019039	0.002435	0.0023446
Glu_432	M+0	0.290332	0.2893784	0.30771	0.3097441
(GC-MS)	M+1	0.2475915	0.251122	0.2269155	0.2287319
	M+2	0.232211	0.2323368	0.2280415	0.227608
	M+3	0.1328165	0.1322726	0.133335	0.1328536
	M+4	0.0628915	0.0616219	0.06438	0.0632198
	M+5	0.0248595	0.0239985	0.028519	0.0270536
	M+6	0.0073195	0.0069765	0.0085515	0.0079658
	M+7	0.001979	0.0018459	0.002547	0.0022626
PEP_453	M+0	0.461926	0.4679881	0.4716	0.4758571
(GC-MS)	M+1	0.221149	0.2131194	0.2134095	0.2069735
	M+2	0.136666	0.1283813	0.1295635	0.1205395
	M+3	0.1234385	0.1287308	0.1274655	0.1327648
	M+4	0.0395695	0.0403887	0.0405255	0.0413947
	M+5	0.0172515	0.0167772	0.0174365	0.0176172
DHAP_484	M+0	0.469809	0.4589038	0.4683365	0.4636258
(GC-MS)	M+1	0.214144	0.2185869	0.2132205	0.215202
	M+2	0.127253	0.1314843	0.1265415	0.1263373
	M+3	0.1293645	0.1278066	0.131949	0.1302708
	M+4	0.041562	0.0414698	0.042067	0.04212

G6P_259/97 (LC-MS/MS)	M+5	0.017867	0.0169377	0.017886	0.0174737
	M+0	0.581274	0.5593024	0.637338	0.6221397
	M+1	0.128596	0.1389598	0.107516	0.1086489
	M+2	0.080728	0.085333	0.058038	0.0608688
	M+3	0.081173	0.0885274	0.044553	0.046115
	M+4	0.033717	0.0316766	0.028074	0.0322892
	M+5	0.016371	0.016946	0.016677	0.0161636
	M+6	0.077708	0.0780155	0.107208	0.112072
6PG_275/79 (LC-MS/MS)	M+7	0.000432	0.0004367	0.000596	0.0005494
	M+0	0.555289	0.556425	0.623947	0.6189391
	M+1	0.139886	0.1388424	0.108274	0.1087546
	M+2	0.091268	0.0873184	0.060494	0.0632034
	M+3	0.086546	0.0887302	0.044649	0.0463867
	M+4	0.036953	0.0319582	0.02865	0.0324229
	M+5	0.017982	0.0172537	0.017588	0.0163032
	M+6	0.071195	0.0777617	0.116136	0.1116445
3PG_185/79 (LC-MS/MS)	M+7	0.000881	0.0005873	0.000262	0.0007323
	M+0	0.700117	0.7017024	0.713632	0.7135012
	M+1	0.104671	0.1010387	0.090999	0.0881493
	M+2	0.058682	0.056913	0.049505	0.0474181
	M+3	0.135695	0.1384362	0.144923	0.1489864
	M+4	0.000834	0.0007742	0.000942	0.0007223

Metabolite Abbreviations: Pyr, Pyruvate; Suc, Succinate; Fum, Fumarate; Akg, α -Ketoglutarate; Mal, Malate; Asp, Aspartate; Glu, Glutamate; PEP, Phosphoenolpyruvate; DHAP, Dihydroxyacetone phosphate; G6P, Glucose-6-phosphate; 6PG, 6-Phosphogluconate; 3PG, 3-Phosphoglycerate.

Table B-3 Best-Fit Fluxes and Flux Confidence Intervals

Notation:

95 lb = lower bound of flux 95% confidence interval

68 lb = lower bound of flux 68% confidence interval

best = best-fit flux value

68 ub = upper bound of flux 68% confidence interval

95 ub = upper bound of flux 95% confidence interval

“net” indicates net flux for a reversible reaction.

“exch” indicates exchange flux for a reversible reaction.

MTYL037

Flux	95 lb	68 lb	best	68 ub	95 ub
v1	100	100	100	100	100
v2	51.995	55.822	59.776	63.607	67.501
v3	1.8468	2.0109	2.18	2.349	2.5131
v4	4.8005	5.6358	6.4956	7.3424	8.1765
v5	0.40152	0.43435	0.46816	0.50197	0.5348
v6	1.0409	1.7056	2.39	3.0609	3.7253
v7	0	0	0.078991	0.17601	0.26041
v8	0.73959	0.82399	0.92101	1	1
v9	1	1	1	1	1
v10	0	0	1.00E-07	0.19752	0.48924
v11	0.51076	0.80248	1	1	1
v12	1	1	1	1	1
v13	0.08908	0.17538	1	1	1
v14	0	0	-2.07E-05	0.82462	0.91092
v15	1	1	1	1	1
v16 net	45.422	46.688	48.007	49.329	50.641
v16 exch	300.52	319.68	342.18	367.74	395.58
v17 net	74.516	75.26	76.017	76.782	77.524
v17 exch	40.645	45.687	57.195	122.39	172.64
v18 net	74.516	75.26	76.017	76.782	77.524
v18 exch	111.32	127.01	246.01	1621.8	6757.6
v19 net	72.302	73.063	73.837	74.62	75.379
v19 exch	1316.6	1597.6	2036.4	2786.5	4284.9

v20	1.8468	2.0109	2.18	2.349	2.5131
v21	163.79	165.32	166.87	168.45	169.98
v22	163.79	165.32	166.87	168.45	169.98
v23	163.79	165.32	166.87	168.45	169.98
v24	49.359	50.671	51.993	53.312	54.578
v25	49.359	50.671	51.993	53.312	54.578
v26 net	16.609	17.046	17.487	17.927	18.349
v26 exch	62.41	159.69	1.00E+07	Inf	Inf
v27 net	32.75	33.624	34.506	35.385	36.23
v27 exch	440.82	1383.1	1.00E+07	Inf	Inf
v28 net	32.75	33.624	34.506	35.385	36.23
v28 exch	441.54	1384	1.00E+07	Inf	Inf
v29 net	-17.882	-17.459	-17.019	-16.578	-16.14
v29 exch	118.72	127.19	136.61	146.65	157.15
v30 net	-18.349	-17.927	-17.487	-17.046	-16.609
v30 exch	21.05	33.587	86.757	2420.8	Inf
v31 net	-18.349	-17.927	-17.487	-17.046	-16.609
v31 exch	115.2	129.73	147.15	167.39	190.33
v32 net	16.609	17.046	17.487	17.927	18.349
v32 exch	9.9426	12.949	19.111	34.354	46.1
v33	35.302	38.033	41.501	45.786	50.192
v34	37.923	40.579	43.891	48.014	52.276
v35	161.16	162.81	164.48	166.19	167.84
v36	161.16	162.81	164.48	166.19	167.84
v37	94.116	98.238	102.32	106.57	110.68
v38	94.116	98.238	102.32	106.57	110.68
v39	94.116	98.238	102.32	106.57	110.68
v40 net	94.116	98.238	102.32	106.57	110.68
v40 exch	13.184	24.059	37.738	53.973	75.021
v41 net	94.116	98.238	102.32	106.57	110.68
v41 exch	752.51	1799.2	1.00E+07	Inf	Inf
v42 net	161.16	162.81	164.48	166.19	167.84
v42 exch	0	0	0.016918	Inf	Inf
v43	51.995	55.822	59.776	63.607	67.501
v44 net	96.391	100.25	103.67	107.67	111.54
v44 exch	0	0	3.74E+06	Inf	Inf
v45	114.85	118.97	122.98	126.24	129.18
v46 net	54.434	58.235	62.166	65.963	69.834
v46 exch	123.09	188.62	283.36	489.9	1054.5
v47	35.302	38.033	41.501	45.786	50.192
v48	0	0.75185	25.281	55.506	92.744

v49	401.78	416.73	444	476.82	515.91
-----	--------	--------	-----	--------	--------

MTYL065

	95 lb	68 lb	best	68 ub	95 ub
v1	100	100	100	100	100
v2	118.66	120.93	123.27	125.48	127.29
v3	4.567	4.6568	4.7492	4.8417	4.9315
v4	2.5402	2.9753	3.413	3.8641	4.3027
v5	0.18057	0.2076	0.23544	0.26329	0.29032
v6	0	0	0	0	0
v7	0	0	0.04613	0.16941	0.27513
v8	0.72487	0.83059	0.95387	1	1
v9	1	1	1	1	1
v10	0	0.17452	0.42994	1	1
v11	0	0	0.57006	0.82548	1
v12	1	1	1	1	1
v13	0	0	0.080834	1	1
v14	0	0	0.91917	1	1
v15	1	1	1	1	1
v16 net	-4.2025	-3.2164	-2.2121	-1.1186	-0.0429
v16 exch	139.35	144.14	149.4	155.7	162.35
v17 net	61.502	61.962	62.438	62.916	63.38
v17 exch	6.4943	7.8703	573.82	Inf	Inf
v18 net	61.502	61.962	62.438	62.916	63.38
v18 exch	6.456	7.9084	10.55	Inf	Inf
v19 net	56.737	57.204	57.689	58.174	58.647
v19 exch	507.2	558.34	621.48	700.48	796.97
v20	4.567	4.6568	4.7492	4.8417	4.9315
v21	152.34	153.18	154.04	154.9	155.74
v22	152.34	153.18	154.04	154.9	155.74
v23	152.34	153.18	154.04	154.9	155.74
v24	100.04	101.12	102.21	103.22	104.2
v25	100.04	101.12	102.21	103.22	104.2
v26 net	33.425	33.784	34.149	34.484	34.813
v26 exch	15.43	56.465	1.00E+07	Inf	Inf
v27 net	66.617	67.334	68.063	68.732	69.39
v27 exch	0	0	1.00E-07	Inf	Inf

v28 net	66.617	67.334	68.063	68.732	69.39
v28 exch	0	0	10.283	Inf	Inf
v29 net	-34.578	-34.249	-33.914	-33.549	-33.19
v29 exch	34.341	39.158	44.521	50.925	57.52
v30 net	-34.813	-34.484	-34.149	-33.784	-33.425
v30 exch	5.0515	8.3114	35.928	Inf	Inf
v31 net	-34.813	-34.484	-34.149	-33.784	-33.425
v31 exch	11.987	19.945	28.416	32.747	37.261
v32 net	33.425	33.784	34.149	34.484	34.813
v32 exch	5.0379	8.3094	35.035	Inf	Inf
v33	31.315	33.985	37.861	42.814	47.894
v34	31.315	33.985	37.861	42.814	47.894
v35	152.34	153.18	154.04	154.9	155.74
v36	152.34	153.18	154.04	154.9	155.74
v37	26.734	28.51	30.77	33.228	35.61
v38	26.734	28.51	30.77	33.228	35.61
v39	26.734	28.51	30.77	33.228	35.61
v40 net	26.734	28.51	30.77	33.228	35.61
v40 exch	3.2751	5.2958	7.6935	10.496	13.709
v41 net	26.734	28.51	30.77	33.228	35.61
v41 exch	266.53	368.44	605.25	1585.2	Inf
v42 net	152.34	153.18	154.04	154.9	155.74
v42 exch	0	0	1.25E-05	Inf	Inf
v43	118.66	120.93	123.27	125.48	127.29
v44 net	156.11	158.27	161.13	164.61	168.27
v44 exch	0	0	6.8681	Inf	Inf
v45	106.51	111.4	116.18	119.98	122.64
v46 net	118.66	120.93	123.27	125.48	127.29
v46 exch	0	0	0.74378	Inf	Inf
v47	31.315	33.985	37.861	42.814	47.894
v48	0	0	1.7312	8.6384	15.848
v49	309.05	313.24	319.52	327.64	335.96

Appendix C

***Yarrowia lipolytica* Acetate Flux Estimation Details**

This Appendix contains details of the ^{13}C -MFA of *Y. lipolytica* strain MTYL065 on acetate, including figures showing the linear regressions for Estimation of Extracellular Fluxes (Figure C-1), the Model Metabolic Reaction Network (Table C-1), measured and simulated ^{13}C labeling data from the flux estimation (Table C-2), and the best-fit flux values and 68% and 95% confidence intervals for the estimated fluxes (Table C-3).

The list of reactions included in the FBA model can be found in Table C-4.

Figure C-1 Linear regressions for Estimation of Extracellular Fluxes

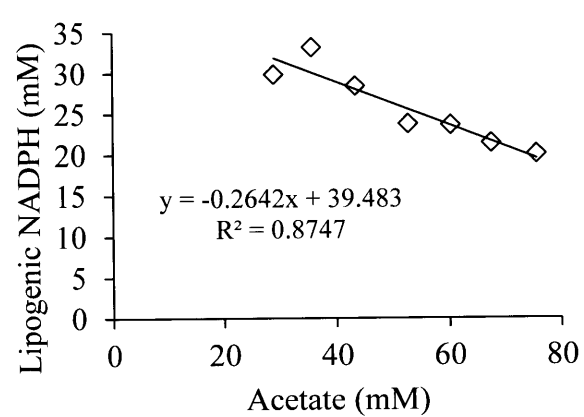
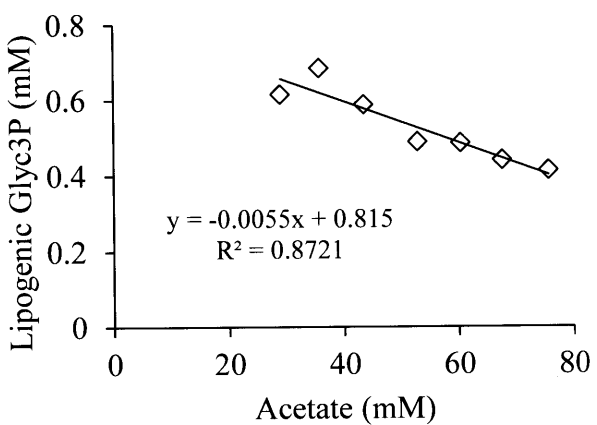
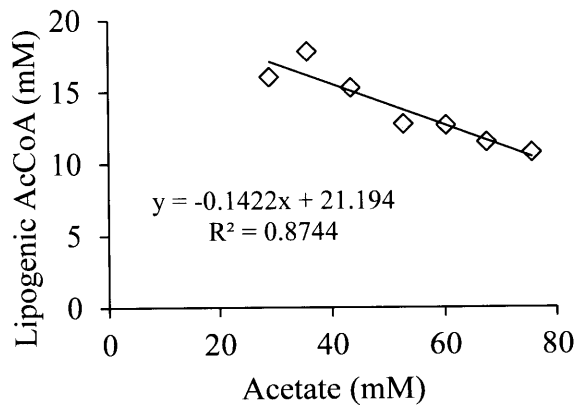
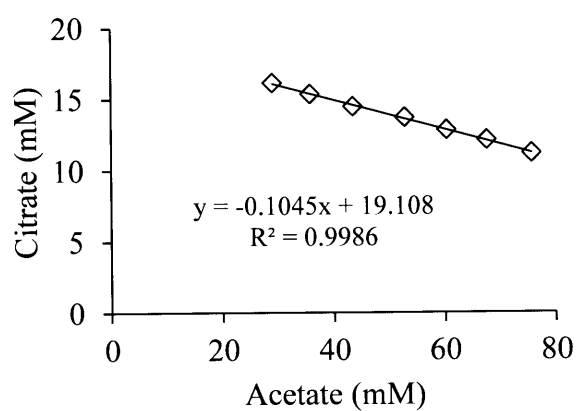


Table C-1: Model Metabolic Reaction Network for ¹³C-MFA

Extracellular fluxes

v1: Acetate.ext (ab) -> AcCoA.c (ab)
v2: AcCoA.c (ab) -> Lipids (ab)
v3: Glyc3P (abc) -> TAGs (abc)
v4: Cit.c (abcdef) -> Cit.ext (abcdef)

Pyruvate compartmentalization

v5: 0 Pyr.c (abc) -> Pyr.mnt (abc)
v6: 0 Pyr.m (abc) -> Pyr.mnt (abc)
v7: Pyr.mnt (abc) -> Pyr.fix (abc)

Malate compartmentalization

v8: 0 Mal.c (abcd) -> Mal.mnt (abcd)
v9: 0 Mal.m (abcd) -> Mal.mnt (abcd)
v10: Mal.mnt (abcd) -> Mal.fix (abcd)

Succinate compartmentalization

v11: 0 Suc.c (abcd) -> Suc.mnt (abcd)
v12: 0 Suc.m (abcd) -> Suc.mnt (abcd)
v13: Suc.mnt (abcd) -> Suc.fix (abcd)

Alanine compartmentalization

v14: 0 Pyr.c (abc) -> Ala (abc)
v15: 0 Pyr.m (abc) -> Ala (abc)
v16: Ala (abc) -> Ala.fix (abc)

Aspartate compartmentalization

v17: 0 OAA.c (abcd) -> Asp (abcd)
v18: 0 OAA.m (abcd) -> Asp (abcd)
v19: Asp (abcd) -> Asp.fix (abcd)

Gluconeogenesis

v20: OAA.c (abcd) -> PEP (abc) + CO2 (d)
v21: PEP (abc) <-> PG3 (abc)
v22: PEP (abc) -> Pyr.c (abc)
v23: PG3 (abc) <-> G3P (abc)
v24: G3P (abc) <-> DHAP (abc)
v25: DHAP (abc) -> Glyc3P (abc)
v26: DHAP (cba) + G3P (def) <-> FBP (abcdef)
v27: FBP (abcdef) <-> F6P (abcdef)
v28: F6P (abcdef) <-> G6P (abcdef)

Pentose Phosphate Pathway

v29: G6P (abcdef) -> CO2 (a) + Ru5P (bcdef)
v30: Ru5P (abcde) <-> R5P (abcde)

v31: Ru5P (abcde) <-> X5P (abcde)
v32: X5P (abcde) <-> G3P (cde) + TKC2 (ab)
v33: E4P (cdef) + TKC2 (ab) <-> F6P (abcdef)
v34: R5P (cdefg) + TKC2 (ab) <-> S7P (abcdefg)
v35: G3P (def) + TAC3 (abc) <-> F6P (abcdef)
v36: S7P (abcdefg) <-> E4P (defg) + TAC3 (abc)

Pyruvate metabolism

v37: Mal.m (abcd) -> Pyr.m (abc) + CO2 (d)
v38: Pyr.c (abc) + CO2 (d) -> OAA.c (abcd)
v39: Mal.c (abcd) <-> OAA.c (abcd)

Citric Acid Cycle

v40: Pyr.m (abc) -> AcCoA.m (bc) + CO2 (a)
v41: OAA.m (abcd) + AcCoA.m (ef) -> Cit.m (dcbfea)
v42: Cit.m (abcdef) <-> ICit.m (abcdef)
v43: ICit.m (abcdef) -> AKG.m (abcde) + CO2 (f)
v44: AKG.m (abcde) -> Suc.m (bcde) + CO2 (a)
v45: Suc.m (abcd) <-> Fum.m (abcd)
v46: Fum.m (abcd) <-> Mal.m (abcd)
v47: Mal.m (abcd) <-> OAA.m (abcd)

Glyoxylate Shunt

v48: Cit.c (abcdef) <-> ICit.c (abcdef)
v49: ICit.c (abcdef) -> Glx (ab) + Suc.c (edcf)
v50: Glx (ab) + AcCoA.c (cd) -> Mal.c (abcd)
v51: Suc.c (abcd) -> Suc.m (abcd)

Transport Reactions

v52: Pyr.c (abc) -> Pyr.m (abc)
v53: Cit.m (abcdef) -> Cit.c (abcdef)
v54: AcCoA.c (ab) <-> AcCoA.m (ab)
v55: Mal.c (abcd) -> Mal.m (abcd)

CO2 Exchange

v56: CO2.ext (a) -> CO2 (a)
v57: CO2 (a) -> CO2.out (a)

Table C-2: Measured and Simulated Mass Spectrometry Data

Notation:

Fragment: the ion measured by the mass spectrometer. Each ion is denoted by a metabolite abbreviation followed by an underscore and the mass of the unlabeled M+0 isotopomer.

Mass Isotopomer: each mass isotopomer is denoted by its mass relative to the mass of the unlabeled parent ion M.

Measured: mass isotopomer mole fractions measured by mass spectrometry in ¹³C labeling experiment

Simulated: mass isotopomer mole fractions obtained by simulation of best-fit flux model

Fragment	Mass Isotopomer	Measured	Simulated
Pyr_174 (GC-MS)	M+0	0.356746	0.359121937
	M+1	0.536299	0.536945112
	M+2	0.078501	0.077737558
	M+3	0.025445	0.024164284
	M+4	0.002277	0.001829254
	M+5	0.000732	0.000190678
Ala_232 (GC-MS)	M+0	0.732153	0.732049316
	M+1	0.180048	0.1809349
	M+2	0.073601	0.072925055
	M+3	0.012094	0.011783452
	M+4	0.002104	0.00207368
Ala_260 (GC-MS)	M+0	0.321659	0.315492235
	M+1	0.49158	0.499100557
	M+2	0.13025	0.130741592
	M+3	0.046644	0.046011087
	M+4	0.007436	0.007283954
Suc_289 (GC-MS)	M+0	0.060341	0.053900555
	M+1	0.510678	0.515385505
	M+2	0.304161	0.307987622
	M+3	0.091499	0.091852108
	M+4	0.026473	0.025938871
	M+5	0.004386	0.004184431
Akg_346	M+6	0.002462	0.000675868
	M+0	0.017483	0.017988431

(GC-MS)	M+1	0.301203	0.31448199
	M+2	0.486375	0.477590633
	M+3	0.13431	0.132525862
	M+4	0.050033	0.047559549
	M+5	0.008689	0.00816595
	M+6	0.001688	0.001493725
	M+7	0.00022	0.000173846
Mal_391	M+0	0.27191	0.267515898
(GC-MS)	M+1	0.466849	0.467541851
	M+2	0.17115	0.17260937
	M+3	0.072898	0.071343744
	M+4	0.017193	0.016510818
Mal_419	M+0	0.057784	0.054341906
(GC-MS)	M+1	0.447149	0.446326647
	M+2	0.313434	0.316039603
	M+3	0.127416	0.127397535
	M+4	0.043799	0.043201733
	M+5	0.010417	0.0101427
Asp_302	M+0	0.309991	0.30395611
(GC-MS)	M+1	0.500981	0.500138854
	M+2	0.138946	0.138852799
	M+3	0.050081	0.047588326
Asp_390	M+0	0.259249	0.265929414
(GC-MS)	M+1	0.474724	0.468025307
	M+2	0.175534	0.173972539
	M+3	0.073248	0.071179687
	M+4	0.017244	0.016494942
Asp_418	M+0	0.064394	0.061706214
(GC-MS)	M+1	0.435015	0.433052578
	M+2	0.318283	0.320752004
	M+3	0.1266	0.127982558
	M+4	0.044439	0.043753254
	M+5	0.01127	0.010208521
	M+6	0	0.002150673
PEP_453	M+0	0.268171	0.270425166
(GC-MS)	M+1	0.461732	0.459225906
	M+2	0.177083	0.174422063
	M+3	0.075373	0.073355274
	M+4	0.01764	0.0175692
	M+5	0	0.004171769
DHAP_484	M+0	0.281864	0.272910403
(GC-MS)	M+1	0.435155	0.434408231
	M+2	0.180017	0.187379823

	M+3	0.08042	0.078346629
	M+4	0.022544	0.020776866
	M+5	0	0.005060319
3PG_585	M+0	0.234663	0.233401667
(GC-MS)	M+1	0.424355	0.419599337
	M+2	0.208804	0.207521558
	M+3	0.099326	0.098258766
	M+4	0.032852	0.030287268
	M+5	0	0.008585666
S7P_289/97	M+0	0.133762	0.138185094
(LC-MS/MS)	M+1	0.34765	0.349053559
	M+2	0.318357	0.31878494
	M+3	0.1478	0.147433517
	M+4	0.043843	0.039153568
	M+5	0.008587	0.006508336
	M+6	0	0.000791639
	M+7	0	8.23448E-05
	M+8	0	6.46754E-06
G6P_259/97	M+0	0.188458	0.185334916
(LC-MS/MS)	M+1	0.410009	0.412977446
	M+2	0.294959	0.293057804
	M+3	0.087263	0.090080833
	M+4	0.017205	0.016436943
	M+5	0.002105	0.001896477
	M+6	0	0.000200443
	M+7	0	1.39846E-05
F6P_259/97	M+0	0.186328	0.185334916
(LC-MS/MS)	M+1	0.411034	0.412977446
	M+2	0.299487	0.293057804
	M+3	0.082186	0.090080833
	M+4	0.017095	0.016436943
	M+5	0.003869	0.001896477
	M+6	0	0.000200443
	M+7	0	1.39846E-05
DHAP_169/97	M+0	0.427231	0.417670452
(LC-MS/MS)	M+1	0.525822	0.52789048
	M+2	0.041046	0.047840661
	M+3	0.0059	0.00638336
	M+4	0	0.000195404
3PG_185/79	M+0	0.408453	0.407071663
(LC-MS/MS)	M+1	0.554943	0.556472076
	M+2	0.030038	0.030010297
	M+3	0.006566	0.00618728

	M+4	0	0.000230659
PEP_167/79	M+0	0.400668	0.406586603
(LC-MS/MS)	M+1	0.571193	0.563555611
	M+2	0.028139	0.025112516
	M+3	0	0.004586513
	M+4	0	0.000144499

Metabolite Abbreviations: Pyr, Pyruvate; Ala, Alanine; Suc, Succinate; Akg, α -Ketoglutarate; Mal, Malate; Asp, Aspartate; PEP, Phosphoenolpyruvate; DHAP, Dihydroxyacetone phosphate; 3PG, 3-Phosphoglycerate; S7P, Sediheptulose-7-phosphate; G6P, Glucose-6-phosphate; F6P, Fructose-6-phosphate.

Table C-3 Best-Fit Fluxes and Flux Confidence Intervals

Notation:

95 lb = lower bound of flux 95% confidence interval

68 lb = lower bound of flux 68% confidence interval

best = best-fit flux value

68 ub = upper bound of flux 68% confidence interval

95 ub = upper bound of flux 95% confidence interval

“net” indicates net flux for a reversible reaction.

“exch” indicates exchange flux for a reversible reaction.

Flux	95 lb	68 lb	best	68 ub	95 ub
v1	100	100	100	100	100
v2	10.78744	13.1144	15.57796	17.0083	19.1185
v3	0.370453	0.461303	0.55488	0.64299	0.733039
v4	10.13982	10.31406	10.49369	10.63821	10.80812
v5	0	0.41126	0.516149	0.887121	1
v6	0	0.112879	0.483851	0.58874	1
v7	1	1	1	1	1
v8	0.013171	0.058087	0.101535	0.11476	0.160303
v9	0.839697	0.88524	0.898465	0.941913	0.986829
v10	1	1	1	1	1
v11	0.004541	0.02654	0.049144	0.056005	0.071157
v12	0.928843	0.943995	0.950856	0.97346	0.995459
v13	1	1	1	1	1
v14	0.790005	0.938766	1.00016	1	1
v15	0	0	-0.00016	0.061234	0.209995
v16	1	1	1	1	1
v17	0.13483	0.17788	0.219404	0.231616	0.259046
v18	0.740954	0.768384	0.780596	0.82212	0.86517
v19	1	1	1	1	1
v20	32.73936	35.75026	53.93669	96.90429	Inf
v21 net	3.047504	4.245244	4.665824	5.34059	6.089388
v21 exch	13.94028	23.23563	34.09543	50.31769	212.3054
v22	27.84679	30.16878	49.27086	91.07298	Inf
v23 net	3.047504	4.245244	4.665824	5.34059	6.089388
v23 exch	5.486453	6.736656	8.965146	9.971524	12.66524
v24 net	3.047504	4.245244	4.665824	5.34059	6.089388
v24 exch	14.68449	39.09093	60.60107	561.2738	Inf

v25	0.370453	0.461303	0.55488	0.64299	0.733039
v26 net	2.479726	3.689325	4.110944	4.791549	5.562884
v26 exch	3.461987	4.10665	5.742752	11.76088	40.85033
v27 net	2.479726	3.689325	4.110944	4.791549	5.562884
v27 exch	6.645616	8.094521	15.56516	85.39839	NaN
v28 net	7.315504	11.06798	12.33283	14.37514	16.68866
v28 exch	0	0	0.00111	Inf	Inf
v29	7.315504	11.06798	12.33283	14.37514	16.68866
v30 net	2.479726	3.689325	4.110944	4.791549	5.562884
v30 exch	37.12914	207.7586	9237615	Inf	Inf
v31 net	4.959452	7.378651	8.221887	9.583098	11.12577
v31 exch	6.636583	9.824783	19.59417	74.60394	Inf
v32 net	4.959452	7.378651	8.221887	9.583098	11.12577
v32 exch	7.053703	12.0293	37.75	779.9478	Inf
v33 net	2.479726	3.689325	4.110944	4.791549	5.562884
v33 exch	0	0	1E-07	5.494653	15.49367
v34 net	2.479726	3.689325	4.110944	4.791549	5.562884
v34 exch	41.95943	151.341	9592770	Inf	Inf
v35 net	2.479726	3.689325	4.110944	4.791549	5.562884
v35 exch	12.72137	20.49333	34263.96	Inf	Inf
v36 net	2.479726	3.689325	4.110944	4.791549	5.562884
v36 exch	15.40139	26.55488	30.49554	37.93997	46.70845
v37	4.286329	5.813916	6.450722	6.944427	7.510069
v38	25.70586	30.76458	49.27086	92.15966	127.4838
v39 net	2.991209	4.245246	4.665824	5.433218	7.449774
v39 exch	336.7234	1176.716	9792466	Inf	Inf
v40	5.397237	5.90753	6.450722	6.604771	7.015965
v41	66.55651	68.2478	69.26253	70.90759	72.64862
v42 net	35.88651	36.7934	37.15859	38.02334	39.69561
v42 exch	0	0	0.000592	Inf	Inf
v43	35.88651	36.7934	37.15859	38.02334	39.69561
v44	35.88651	36.7934	37.15859	38.02334	39.69561
v45 net	56.07843	57.76165	58.76883	60.43098	62.17909
v45 exch	481.1521	1870.484	9999941	Inf	Inf
v46 net	56.07843	57.76165	58.76883	60.43098	62.17909
v46 exch	449.6913	1721.747	9622706	Inf	Inf
v47 net	66.55651	68.2478	69.26253	70.90759	72.64862
v47 exch	0	0	0.000702	Inf	Inf
v48 net	20.13765	21.11767	21.61024	21.85017	22.46773
v48 exch	0	0	1.889242	Inf	Inf
v49	20.13765	21.11767	21.61024	21.85017	22.46773
v50	20.13765	21.11767	21.61024	21.85017	22.46773
v51	20.13765	21.11767	21.61024	21.85017	22.46773

v52	0	0	1E-07	0.738086	2.185848
v53	30.61074	31.59794	32.10393	32.34435	32.96296
v54 net	61.17072	62.33888	62.8118	64.50967	66.20654
v54 exch	124.5033	452.6455	3101654	Inf	Inf
v55	14.64222	16.39134	16.94442	17.43664	18.01255
v56	18.45671	23.97255	42.26565	84.10392	Inf
v57	122.5569	127.4425	146.4829	204.161	Inf

Table C-4 List of Reactions Included in FBA Model

null (Acetate) + 2 ATP -> AcCoA_c
 27 AcCoA_c + 1 Glyc3P + 51 NADPH + 24 ATP -> null (TAGs)
 Cit_c -> null (Cit_ext)
 OAA_c + ATP -> PEP
 PEP -> Pyr_c + ATP
 PEP + ATP + NADH_c <-> GAP + NAD
 GAP <-> DHAP
 DHAP + NADH_c <-> Glyc3P + NAD_c
 GAP + DHAP <-> FBP
 FBP -> F6P
 F6P <-> G6P
 G6P -> Ru5P + CO2 + 2 NADPH
 Ru5P <-> R5P
 Ru5P <-> X5P
 X5P <-> GAP + C2
 E4P + C2 <-> F6P
 R5P + C2 <-> S7P
 GAP + C3 <-> F6P
 S7P <-> E4P + C3
 Mal_m + NAD_m -> Pyr_m + NADH_m
 Pyr_c + ATP -> OAA_c
 Mal_c + NAD_c <-> OAA_c + NADH_c
 Pyr_m + NAD_m -> AcCoA_m + NADH_m
 OAA_m + AcCoA_m -> Cit_m
 Cit_m <-> ICit_m
 ICit_m + NAD_m -> AKG_m + NADH_m
 AKG_m + NAD_m -> Suc_m + NADH_m + ATP
 Suc_m <-> Fum_m + FADH2
 Fum_m <-> Mal_m
 Mal_m + NAD_m <-> OAA_m + NADH_m
 Cit_c <-> ICit_c
 ICit_c -> Glx + Suc_c
 AcCoA_c + Glx -> Mal_c
 Suc_c -> Suc_m
 Pyr_c -> Pyr_m
 Cit_m -> Cit_c
 AcCoA_c <-> AcCoA_m
 Mal_c -> Mal_m (reversible in NADH shuttling case)
 NADH_m -> NAD_m + 2.5 ATP
 FADH2 -> 1.5 ATP
 NADH_c -> NAD_c + 2.5 ATP

Additional reactions included in subsets of special cases

ICit_c -> AKG_c + NADPH
 Cit_m + AKG_c <-> Cit_c + AKG_m

Mal_c -> Pyr_c + NADPH
OAA_c <-> OAA_m