

**EXTREMELY THERMOPHILIC *GEOBACILLUS LC300*:
GENETIC IDENTIFICATION,
GROWTH CHARACTERIZATION,
METABOLIC MODELING,
AND SCALE-UP**

by

Lauren Taylor Cordova

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Bachelor in Chemical Engineering with Distinction

Spring 2015

© 2015 Lauren Cordova
All Rights Reserved

**EXTREMELY THERMOPHILIC *GEOBACILLUS LC300*:
GENETIC IDENTIFICATION,
GROWTH CHARACTERIZATION,
METABOLIC MODELING,
AND SCALE-UP**

by

Lauren Taylor Cordova

Approved: _____
Maciek Antoniewicz, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Wilfred Chen, Ph.D.
Committee member from the Department of Chemical Engineering

Approved: _____
Susan Groh, Ph.D.
Committee member from the Board of Senior Thesis Readers

Approved: _____
Michael Arnold, Ph.D.
Directory, University Honors Program

ACKNOWLEDGMENTS

This thesis has been possible due the influence of many people during my time at the University of Delaware. I offer my sincerest gratitude to everyone who has touched my life in the last four years.

First and foremost, I would like to thank my research advisor, Dr. Maciek Antoniewicz. Thank you for providing me the opportunity to work on this dynamic project. You guided me through the research process working through each portion of my project. You worked hands-on with me in the beginning and then let me be more independent over time. Also, I thank you for the opportunity to teach and train another undergraduate student as this was personally rewarding. I can honestly say I would not be the researcher and student I am today without you, thank you.

Second, I would like to thank my academic advisor, Dr. Wilfred Chen for his guidance during the entire course of my undergraduate career. I greatly appreciated your input during the graduate school selection process as well.

I would also like to thank my thesis committee, Dr. Antoniewicz, Dr. Chen, and Dr. Susan Groh for assistance and support through the senior thesis process.

Next, I would like to thank my group members, Jen, Chris, Niko, Jen, Camil, and Brian for insightful conversation, laboratory assistance and of course as a source of frozen snacks. Thanks to all of you, I have fond memories of the Antoniewicz research group at UD.

I also would like to thank all of my friends at the University of Delaware for listening to me discuss my research and its progress.

Finally, I would like to take this opportunity to thank my parents for their unwavering support of my education. They have been a constant source of encouragement, inspiration and motivation for me. I am indebted to them more than they know. Thank you for all you have done for me.

Lauren Cordova

May 2015

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES.....	vi
ABSTRACT.....	viii
1 INTRODUCTION	1
1.1 Motivation	1
1.2 <i>Geobacillus</i> and <i>Geobacillus</i> LC300	1
1.3 Aim of thesis.....	2
2 MATERIALS AND METHODS	3
2.1 Materials	3
2.2 Strain and growth conditions	3
2.3 Tracer experiments.....	5
2.4 Analytical methods	5
2.4.1 Determination of xylose concentration using standards.....	6
2.4.2 Derivatization of amino acids	6
2.4.3 Determination of biomass composition	7
2.5 Spectroscopy analysis	7
2.5.1 GC-MS for amino acids and sugar detection	7
2.5.2 On-line GC-MS for CO ₂ off gas analysis	8
2.6 Genome sequencing, assembly and analysis	8
2.7 Metabolic Flux determination and statistical analysis	8
2.8 Scale-up and fed-batch fermentation	9
3 RESULTS AND DISCUSSION.....	12
3.1 Genome features	12
3.2 General metabolism	13
3.3 Growth characterization	16
3.3.1 Biomass Composition.....	17
3.4 Metabolic network reconstruction	18
3.5 Off-gas and biomass labeling analysis	19
3.6 Biomass labeling	21
3.7 Metabolic Flux Analysis	24

3.7.1	Metabolic model development.....	24
3.7.1.1	Metabolism of <i>Geobacillus</i> LC300.....	25
3.8	Scale-up analysis.....	29
3.8.1	Batch Experiments	29
3.8.2	Fed-Batch Experiments	32
4	CONCLUSIONS	36
4.1	Conclusions	36
4.2	Recommendations for future work	38
	REFERENCES	39
A	Amino Acid Biosynthesis Pathways.....	41
B	<i>Geobacillus</i> LC300 Metabolic Model	42
C	CO ₂ Off-gas Data from Parallel Labeling Experiments	49
D	Benchtop Bioreactor Experimental Data	52

LIST OF TABLES

Table 2.9: Scale-up experimental design.....	11
Table 3.1: Species with high sequence similarity to <i>Geobacillus LC300</i>	12
Table 3.2: General genome features of <i>Geobacillus LC300</i>	13
Table 3.5: Average percent CO ₂ labeling for parallel glucose and parallel xylose tracer experiments.....	20
Table C.1: Average CO ₂ labeling of off-gas produced from growth on ¹³ C-labeled glucose or xylose	49
Table D.1: Complete data set from Experiment 8: High density growth on glucose with nitrogen, vitamins, minerals, and yeast extract in feed solution.	55
Table D.2: Complete data set from Experiment 8: High density growth on xylose with nitrogen, vitamins, minerals, and yeast extract in feed solution.	56

LIST OF FIGURES

Figure 2.1: Bioreactor configuration with real-time off-gas (CO ₂) monitoring by ProLine mass spectrometer.	4
Figure 3.1: Functional distribution of annotated genes	13
Figure 3.2: Central carbon metabolism labeled by EC number and location within annotation. The second reaction (3.1.1.31) in the oxidative pentose phosphate pathway was not identified by RAST.....	15
Figure 3.3: Growth rate determination during exponential growth as determined by OD ₆₀₀	16
Figure 3.3.2: Growth rate as a function of initial xylose concentration.	17
Figure 3.3: Biomass composition of <i>Geobacillus</i> LC300 compared to other model organisms. Protein, RNA, lipids, and glycogen were experimentally measured for <i>Geobacillus</i> LC300 only	18
Figure 3.5.1: Production and labeling of CO ₂ in off-gas from [2- ¹³ C] xylose	19
Figure 3.6.1: Average carbon labeling of amino acid fragments from glucose parallel labeling experiment	22
Figure 3.6.2 Average carbon labeling of amino acid fragments from xylose parallel labeling experiment.....	23
Figure 3.7.1: ¹³ CO ₂ labeling to elucidate between the ribulose-monophosphate pathway and the oxidative pentose phosphate pathway in upper glycolysis for metabolic modeling.....	25
Figure 3.7.2: Flux map of <i>Geobacillus</i> LC300 from COMPLETE-MFA for growth on xylose	26
Figure 3.7.3: Flux map of <i>Geobacillus</i> LC300 from COMPLETE-MFA for growth on xylose	28
Figure 3.9.1: <i>Geobacillus</i> LC300 Cell Density (OD ₆₀₀) during growth in batch with various control schemes	30
Figure 3.9.2 Measured pH of <i>Geobacillus</i> LC300 under batch culture various control schemes	31

Figure 3.9.3: <i>Geobacillus</i> LC300 Cell Density (OD ₆₀₀) during fed-batch growth with various control schemes	33
Figure C.2: Average CO ₂ labeling (%) in CO ₂ produced in off-gas during growth on labeled glucose.....	49
Figure C.3: CO ₂ Production and CO ₂ Labeling for <i>Geobacillus</i> LC300 growth on labeled xylose	50
Figure C.4: CO ₂ Production and CO ₂ Labeling for <i>Geobacillus</i> LC300 growth on labeled glucose	51
Figure D.1: Glucose concentration profile for batch experiments with 2L scale bioreactor under various controls	52
Figure D.2: Glucose or Xylose concentration profile for fed-batch experiments with 2L bioreactor under various controls	53
Figure E. 3: Dissolved oxygen and agitator speed profiles for 2L bioreactor experiments.....	54

ABSTRACT

A new thermophilic microorganism, *Geobacillus* LC300 was discovered with interesting characteristics. This organism grows optimally at 72 °C and has a fast growth rate (1.5 hr⁻¹) utilizing xylose as the sole carbon source. Thermophilic organisms and more specifically, *Geobacillus* LC300, are of interest in biotechnology for the production of low-boiling fuels and thermostable enzymes. In order to effectively produce these compounds, metabolism must be understood and subsequently manipulated. In this work, a thermophilic strain, identified as *Geobacillus* LC300, was examined to determine the genomic content and growth characteristics for metabolic model reconstruction and metabolic flux analysis.

First, *Geobacillus* LC300 was isolated through plating on agar plates and the genome sequenced using PacBio. The organism was cultured in 10 mL custom mini-bioreactors designed for thermophiles at 72 °C. The maximum growth rate of 1.5 hr⁻¹ was observed with xylose as the sole carbon source.

From the genome sequence, central carbon metabolism and amino acid biosynthesis pathways were annotated. Major metabolic pathways such as glycolysis, the pentose phosphate pathway and the tricarboxylic acid (TCA) cycle were complete. These complete reactions were used to reconstruct a metabolic model which included 73 reactions with a lumped biomass reaction. The biomass composition was experimentally determined for accurate metabolic modeling. The reconstructed *Geobacillus* LC300 model was then validated with ¹³C-labeling experiments and analyzed through ¹³C-metabolic flux analysis for growth on xylose. A ¹³C-labeling experiment was also completed with parallel glucose tracers to detect any major differences between glucose and xylose metabolism on central carbon metabolism. As

expected, upper glycolysis is more active under glucose consumption compared to xylose. Likewise, the pentose phosphate pathway has higher flux during growth on xylose.

Finally, scale-up analysis enabled high cell densities of *Geobacillus* LC300 which is industrially relevant. Increasing reactor control and feed components (glucose, nitrogen, vitamins, minerals, and yeast extract) led to higher cell densities. A fed-batch experiment with glucose reached a maximum OD₆₀₀ (cell density) of 22, while growth on xylose reached a maximum of 25.

This extensive analysis provides a genomic, metabolic and physical understanding of *Geobacillus* LC300 to create a strong foundation for future metabolic engineering and use in biotechnology applications.

Chapter 1

INTRODUCTION

1.1 Motivation

Microorganisms have been widely used to produce valuable chemicals for several decades. Organisms such as *E. coli* and *S. cerevisiae* have been extensively studied and genetically engineered for the production of chemical compounds. Thermophilic organisms are of increasing interest due to several characteristics at high temperatures. Growth at elevated temperatures enables high reaction rates, lower cooling costs of substrates and facilitates the removal of products such as low-boiling fuels. Several thermophilic bacteria also have the ability to consume pentose and hexose sugars leading to higher conversion from cellulosic biomass¹.

1.2 *Geobacillus* and *Geobacillus* LC300

The genus *Geobacillus* is a wide class of thermophilic microbes which have a variety of applications for industrial biotechnological processes. For example, *Geobacilli* have the potential for biofuels production, bioremediation, and a source of thermostable enzymes². *Geobacillus* species typically have optimal growth temperatures at 55-60°C but are capable of growth at up to 70-80 °C³.

¹ Lynd 1989

² Cripps, 2009; Feng, 2007; Verma, 2013; Ezeji, 2005

³ Coorevitis, 2012

Geobacillus LC300 was isolated and thoroughly studied in this thesis. It has desirable growth characteristics such as rapid growth on the pentose sugar, xylose. This organism has the potential for applications across the biotechnology industry from efficient chemical to thermostable enzyme production.

1.3 Aim of thesis

The aim of this study is a complete metabolic analysis of *Geobacillus* LC300. This microorganism grows rapidly in an extremely thermophilic environment (75°C) with glucose or xylose as the sole carbon source. This thesis seeks to understand the metabolic reaction network which enables this behavior. Genomic analysis provided the basis for model development which was subsequently verified using ¹³C metabolic flux analysis. This work is the first extensive genetic, metabolic and scale-up analysis of an extremely thermophilic microorganism.

Chapter 2

MATERIALS AND METHODS

2.1 Materials

Media and chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Wolfe's vitamins and minerals were purchased from ATCC (Manassas, VA). Xylose tracers ([1-¹³C], [2-¹³C], [3-¹³C], [4-¹³C], [5-¹³C], [U-¹³C]) and glucose tracers ([1-¹³C], [2-¹³C], [3-¹³C], [4-¹³C], [5-¹³C], [6-¹³C], [U-¹³C]) were purchased from Cambridge Isotope Laboratories (Andover, MA). Xylose stock solutions (20 wt %), glucose stock solutions (20%), and yeast extract stock solution (1%) were prepared with distilled water and filter sterilized.

For *Geobacillus* LC300 growth, minimal defined media was used with the following composition (per L of medium): 0.50 g K₂HPO₄, 0.30 g KH₂PO₄, 0.50 g NH₄Cl, 0.50 g NaCl, 0.20 g MgCl₂·6H₂O, 0.04 g CaSO₄·2H₂O, 40 mL of 1 M Tris, 5 mL of Wolfe's minerals, 5 mL of Wolfe's vitamins, and 0.05 g/L of yeast extract. Xylose and/or glucose was added as indicated.

2.2 Strain and growth conditions

Geobacillus strain LC300 was isolated from a contaminated culture of *Thermus thermophilus* that was grown aerobically at 72°C. The extremely fast growth rate and change in physical appearance suggested the presence of a contaminant which obtained the ability to grow under the same thermophilic conditions. To isolate the contaminant, a diluted sample of the mixed culture was streaked on a xylose-agar plate and incubated overnight at 60°C. A single colony was selected and cultured in defined growth media with 2 g/L xylose. During exponential phase, cells were collected and

re-suspended in medium containing 15 wt% glycerol and frozen at -80°C . The frozen stock was termed LC300 and used for all further experiments.

Most experiments were performed in custom 10mL bioreactors designed for growing thermophiles. Figure 2.1 below shows the reactor setup.

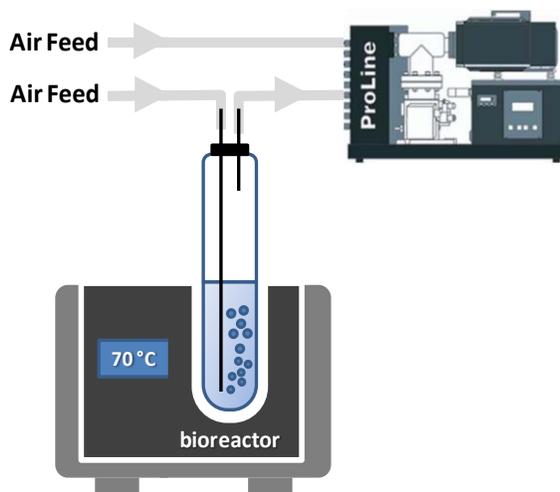


Figure 2.1: Bioreactor configuration with real-time off-gas (CO_2) monitoring by ProLine mass spectrometer.

The glass culture vessels were 15 mL anaerobic Hungate tubes (BellcoGlass Cat. No.2047-16125). The reactor was sealed with a rubber septum pierced with needles for: supply of filtered air, sampling of cell culture and venting of off-gas for spectroscopy analysis. See section 2.5.2 for details of off-gas analysis. These reactors were autoclaved before inoculation. The air flow rate was controlled by a high-precision multi-channel peristaltic pump (Watson Marlow, Wilmington, MA). Mixing was achieved through the rise of air bubbles from the bottom to the top of the reactor.

Parallel cultures with a working volume of 10 mL were set up with 10 mL of base medium, 50 μ L of Wolfe's vitamins, 50 μ L of Wolfe's minerals, and yeast extract, glucose and/or xylose as specified by the experiment.

2.3 Tracer experiments

For parallel labeling experiments, cells from a -80°C frozen stock were pre-grown in medium containing 250 μ L yeast extract (1%) and 1 mL unlabeled xylose (20%). Once the pre-culture reached an OD_{600} of about 1.5, 50 μ L of the pre-culture was used to inoculate the culture tubes containing 10 mL medium with either a glucose or xylose tracer. For the xylose tracer experiment, 100 μ L [1- ^{13}C], [2- ^{13}C], [3- ^{13}C], [4- ^{13}C], and [5- ^{13}C] xylose were added to the 10 mL of culture media. For the glucose tracer experiment, 200 μ L [1- ^{13}C], [2- ^{13}C], [3- ^{13}C], [4- ^{13}C], [5- ^{13}C], and [6- ^{13}C] glucose were added to the culture. Cell density (OD_{600}) was measured periodically throughout the experiment which lasted 8 to 12 hours. Cell biomass samples were collected at an OD_{600} of approximately 0.5 and 1.0 for analysis of biomass amino acids. The supernatant liquid from these samples was also frozen at -4°C for subsequent analysis.

2.4 Analytical methods

Biomass concentration was measured using optical density at 600 nm (OD_{600}) with a spectrophotometer (Eppendorf BioPhotometer). The OD_{600} values were converted to cell dry weight concentration using a determined OD_{600} -dry cell weight relationship for *Geobacillus* LC300: $1.0 \text{ OD}_{600} = 0.27 \text{ g}_{\text{DW}}/\text{L}$. This numerical relationship was determined using a 50 mL sample of approximately OD_{600} 1.0.

2.4.1 Determination of xylose concentration using standards

Xylose concentrations were determined by isotope ratio analysis using [U-¹³C] xylose as a standard. Using techniques developed in the Antoniewicz lab, 50 µL of supernatant was combined with 50 µL of a 3.5 g/L [U-¹³C] xylose solution. This mixture was then evaporated and derivatized using the aldonitrile propionate derivatization with 2wt% hydroxylamine hydrochloride in pyridine solution and propionic anhydride. The last two carbons of xylose were detected via GC-MS analysis of ion fragment (*m/z* 173). This derivatization procedure (aldonitrile propionate) was also used to measure labeling of sugars, glucose or xylose without the addition of standards.

2.4.2 Derivatization of amino acids

Tertbutyldimethylsilyl (TBDMS) derivatization of biomass amino acids was performed prior to GC/MS analysis. 1 mL samples were collected at OD₆₀₀ approximately 0.5 and 1.0 (the end of the experiment). The samples were centrifuged for 5 minutes at 14,000 rpm and the supernatant separated from the cell pellet. The pellet was hydrolyzed with 500 µL of 6N HCl overnight at 110°C. The samples were then centrifuged at 14,000 rpm for 5 minutes to remove cell debris. The contents were transferred to an Eppendorf tube and evaporated to dryness under airflow at 65°C. The dried sample was dissolved in 35 µL of pyridine and 50 µL of N-(ter-butyl dimethylsilyl)-N-methyl-trifluoroacetamide (MTBSTFA) + 1% TBDMCS and reacted for 30 minutes at 60°C. The sample was centrifuged again to remove cell debris and transferred to a GC injection vial.

2.4.3 Determination of biomass composition

Biomass composition was experimentally determined for accurate flux analysis. This included direct measurement of weight percentages of protein, RNA, glycogen and lipids (wt% of dry biomass), using methods recently developed in the Antoniewicz lab⁴. The weight percentage of DNA and ions and metabolites were assumed to be constant (3% each). The remaining fraction was assumed to be cell wall components, peptidoglycan and teichoic acids. To accurately determine the weight percentages, the dry cell weight was determined using 50 mL of liquid culture with a measured optical density of approximately 1.0. These cells were centrifuged, supernatant liquid removed and allowed to dry. Cell mass was measured for several subsequent days to ensure complete drying.

2.5 Spectroscopy analysis

Several forms of spectroscopy analysis were used to collect useful data. Gas chromatography- mass spectroscopy techniques were used to accurately measure labeling of sugars, amino acids and CO₂. The methods to quantify each are outlined in the subsections below.

2.5.1 GC-MS for amino acids and sugar detection

GC-MS analysis was used to quantify the labeling of amino acids and sugars. 1 μ L of derivatized sample was injected with split ratios of 1:10, 1:40 and/or 1:100. An Agilent 7890A GC system was connected to a Waters Quattro Micro Tandem Mass Spectrometer (GC-MS/MS). Mass isotopomer distributions were obtained by integration of selected ion chromatograms and corrected for natural abundances.

⁴ Long, 2014

2.5.2 On-line GC-MS for CO₂ off gas analysis

The inlet air and the off-gas from the mini-bioreactors were monitored in real time by an on-line mass spectrometer (Ametek Proline). Molar percentages of nitrogen, oxygen, carbon dioxide and ¹³C-labeled carbon dioxide were measured to monitor growth. The total CO₂ produced by the cells and the fractional ¹³C-labeling of CO₂ were calculated using the equations below.

$$\text{Net CO}_2 \text{ produced} = [^{12}\text{CO}_2 (\text{off-gas}) + ^{13}\text{CO}_2 (\text{off-gas}) - [^{12}\text{CO}_2 (\text{inlet}) + ^{13}\text{CO}_2 (\text{inlet})]$$

$$^{13}\text{C} - \text{labeling of CO}_2 = [^{13}\text{CO}_2 (\text{off-gas}) - ^{13}\text{CO}_2 (\text{inlet})] / [\text{Net CO}_2 \text{ produced}]$$

2.6 Genome sequencing, assembly and analysis

Genomic DNA of *Geobacillus* LC300 was isolated in the Papoutsakis laboratory using Qiagen's Genomic tip protocol, followed by restriction enzyme digestion. The isolated DNA was sequenced using PacBio's RSII third generation DNA sequencing system known as SMRT (Single Molecule, Real Time) sequencing at the University of Delaware Sequencing and Genotyping Center. A central genome and a plasmid were assembled and annotated using the RAST pipeline⁵.

2.7 Metabolic Flux determination and statistical analysis

The mass isotopomer distribution obtained in the GC/MS analysis was further analyzed using a metabolic modeling software called Metran. This software performs ¹³C metabolic flux analysis (¹³C-MFA) using the elementary metabolite framework (EMU)⁶. A network model for *Geobacillus* LC300 was developed from genomic

⁵ Aziz, 2008

⁶ Antoniewicz, 2007

annotation. The model includes all major metabolic pathways of central carbon metabolism, including glycolysis, pentose phosphate pathway, tricarboxylic acid cycle, one-carbon metabolism and amino acid biosynthesis pathways. The model also accounts for CO₂ exchange which can dilute intracellular carbon labeling.

The metabolic fluxes were estimated using non-least square regression by minimizing the variance-weighted sum of squared residuals between experimentally determined and simulated measurements. For advanced flux analysis of parallel labeling experiments, all data were fitted simultaneously to a single flux model. Flux estimation was repeated at least 10 times starting with random initial values to find a global solution. Confidence intervals (95%) were computed with converged flux results.

2.8 Scale-up and fed-batch fermentation

To understand how *Geobacillus* LC300 grows at a larger scale, a 2 L benchtop bioreactor constructed by New Brunswick (BioFlow/CellGro 115) was used. The benchtop bioreactor had a working volume of 1.5 L. The media for these experiments contained 1.5 L of the minimal media described above including the appropriate 5 mL of Wolfe's Vitamins, Wolfe's Minerals, and Yeast Extract, except where otherwise specified. Glucose or xylose was initially added at 5 g/L.

Several levels of increasing control were implemented in an effort to obtain high cell densities. pH control was maintained at 6.60 via base addition (1N NaOH) while the dissolved oxygen set point was 40% controlled by the agitator speed (rpm). The minimum agitator speed was 400 rpm with a maximum allowable speed of 1000 rpm. The bioreactor unit controlled these set parameters using a PI control scheme. At the limit of pH and DO control, additional glucose and ammonium chloride (nitrogen

source) were added in stoichiometric amounts to sustain additional growth. For these experiments, 500 mg NH_4Cl were added to the reactor in addition to the amount provided from the base media. For the feed solutions, 0.761 moles of N were added per mole of carbon added; this translated to 0.135 g per g glucose.

Extra Wolfe's vitamins and minerals were also added at various points during the fermentation, either as a 5 mL bolus or within the feed solution. To sustain long term, high density cell cultures, the reactor was continuously fed with a nutrient solution for those experiments indicated fed-batch. This solution contained some or all of the following components as dictated by the experiment: 7.5 g glucose, 2.5 g nitrogen, 8 mL Wolfe's vitamins, 8 mL Wolfe's minerals and yeast extract in a 5:1 g/g substrate ratio. Fed-batch reactor cultures used a pump set point of 15% but this was increased or decreased mid-experiment to maintain a constant glucose concentration. The level of culture control was systematically increased with the goal of reaching the highest OD_{600} . For the experiment with xylose, the xylose concentration was not able to be determined rapidly and consequently the pump set point remained at 15%. Table 2.9 outlines the high density cell culture experiments and the control conditions for each.

Table 2.9: Scale-up experimental design

Exp #	Batch/ Fed-batch	pH Control	Oxygen Control	Notes
1	Batch	N	N	
2	Batch	N	Y	
3	Batch	Y	N	
4	Batch	Y	Y	
5	Batch	Y	Y	Addition of 500 mg NH ₄ Cl
6	Fed-Batch	Y	Y	Feed glucose and NH ₄ Cl solution
7	Fed-Batch	Y	Y	Single addition of Vitamins & Minerals, Feed glucose and NH ₄ Cl solution
8	Fed-Batch	Y	Y	Feed glucose, NH ₄ Cl, Vitamins, Minerals, & Yeast Extract
9	Fed-Batch	Y	Y	Feed xylose, NH ₄ Cl, Vitamins, Minerals & Yeast Extract

Chapter 3

RESULTS AND DISCUSSION

3.1 Genome features

The genome of strain LC300 was sequenced using Pacific Biosciences RSII third generation DNA sequencing system which is referred to as SMRT (Single-Molecule, Real-time) sequencing. Analysis of the 16S rDNA identified the strain LC300 as a member of the genus *Geobacillus*. BLAST analysis of the 16S rDNA gene sequence identified several strains of high similarity listed in the table below. *G. stearothermophilus*, the species most closely related to LC300 is the model strain for the *Geobacillus* genus⁷.

Table 3.1: Species with high sequence similarity to *Geobacillus LC300*.

Species	Similarity (% match)
<i>G. stearothermophilus</i>	99.7
<i>G. thermocatenuatus</i>	98.9
<i>G. kaustophilus</i>	98.7
<i>G. thermoleovorans</i>	98.5
<i>G. thermodenitrificans</i>	98.5
<i>G. thermoglucosidasius</i>	96.7

The genome was then annotated using the RAST (Rapid Annotations using Subsystems Technology) servers to map genes for metabolic reconstruction⁸. The genome of *Geobacillus LC300* is composed of a 3,494,216 base-pair main

⁷ Coorevits 2012

⁸ Aziz, 2008

chromosome with a plasmid of 38,364-bp. The chromosome had 3,988 identified protein coding sequences as well as 88 tRNA genes for each of the 20 amino acids. The other genomic features are outlined in Table 3.2 below. From genomic annotation, genes with known EC functions were sorted into categories of similar substrates (Figure 3.1).

Table 3.2: General genome features of *Geobacillus* LC300

	Chromosome	Plasmid
Length (bp)	3,494,216	38,364
GC content (%)	52.2	46.8
Coding Sequences (#)	3988	51
Known (EC) function	950	
RNAs	111	0

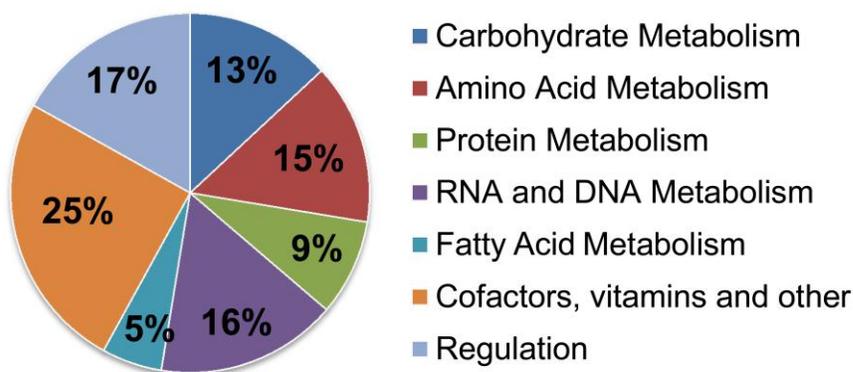


Figure 3.1: Functional distribution of annotated genes

3.2 General metabolism

From genomic annotation, *Geobacillus* LC300 contains most major metabolic pathways such as glycolysis, pentose phosphate pathway, tricarboxylic acid (TCA)

cycle, the glyoxylate shunt, and the ribulose monophosphate (RuMP) pathway but not the Entner-Doudoroff pathway. Although one of the enzymes (EC 3.1.1.31) in the oxidative branch of the pentose phosphate pathway was not identified, this reaction can proceed spontaneously and the pathway was deemed complete⁹. Figure 3.2 below shows the EC numbers and genome location (by peg in annotation) of the enzymes required for central carbon metabolism. A similar annotation map for amino acid biosynthesis is located in Appendix A.

⁹ Kupor, 1969; Thomason, 2004; Zimenkov, 2005

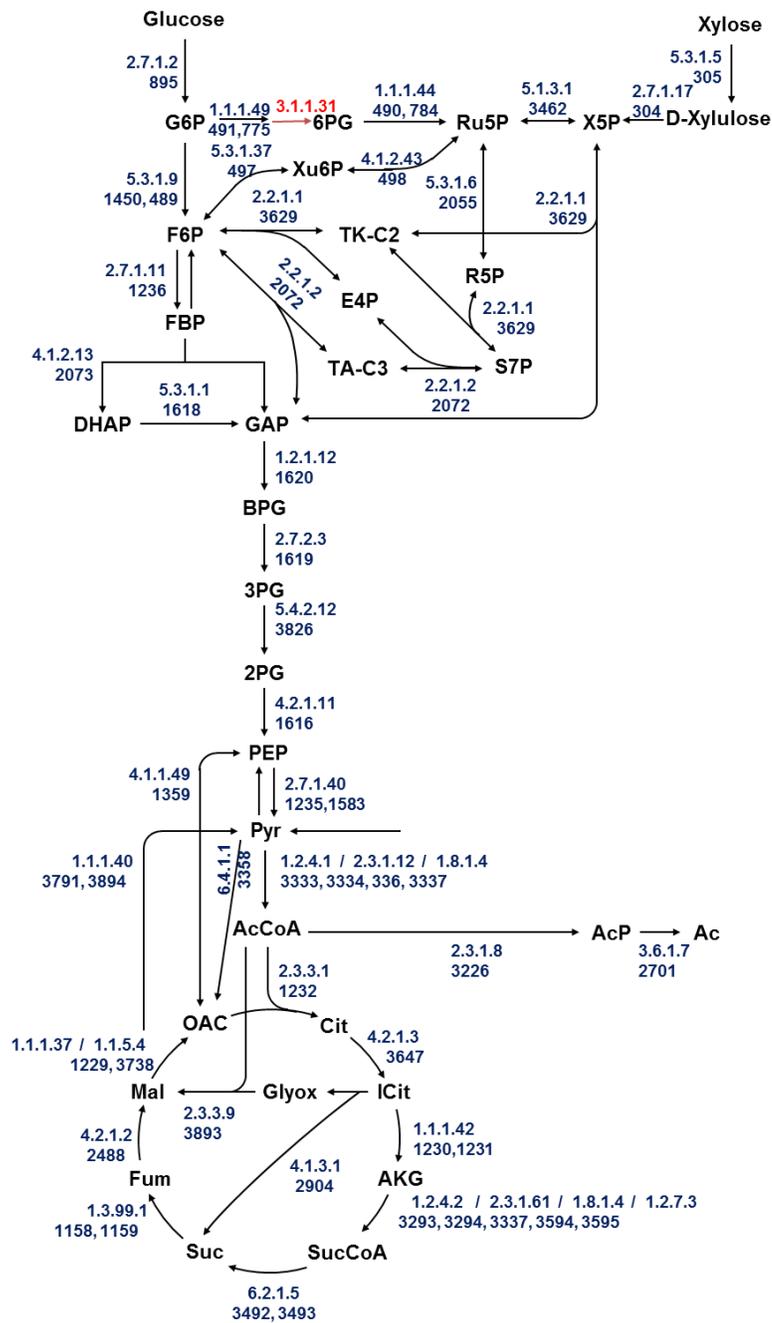


Figure 3.2: Central carbon metabolism labeled by EC number and location within annotation. The second reaction (3.1.1.31) in the oxidative pentose phosphate pathway was not identified by RAST.

3.3 Growth characterization

One of the interesting characteristics of *Geobacillus* LC300 is its rapid growth rate on xylose. Under aerobic conditions, a temperature of 72°C, and an initial xylose concentration of 30 g/L, the average specific growth rate was 1.52 hr⁻¹ as seen in the figure below.

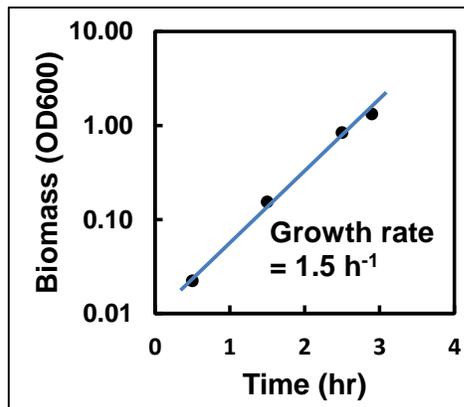


Figure 3.3: Growth rate determination during exponential growth as determined by OD₆₀₀.

To ensure optimal growth, the growth rate was measured as a function of initial xylose concentration. Figure 3.4 below shows *Geobacillus* LC300 follows Monod growth kinetics with a maximum growth rate with 1000 µL of 20 wt% xylose solution. For all future growth experiments, 1000 µL was used per 10 mL of base media (unless otherwise specified).

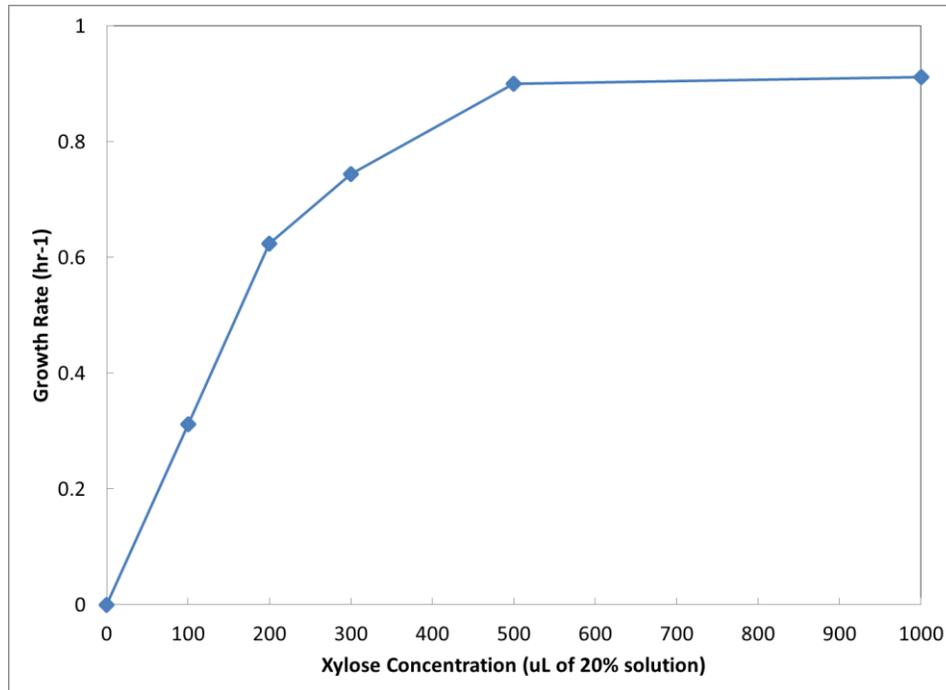


Figure 3.3.2: Growth rate as a function of initial xylose concentration.

3.3.1 Biomass Composition

In order to complete the metabolic reconstruction of *Geobacillus* LC300 for subsequent flux analysis, the distribution of biomass components must be known. This enables the most accurate stoichiometry for the lumped biomass equation in the model. Figure 3.3.1 below shows the determined biomass composition for *Geobacillus* LC300 in addition to that of *B. subtilis* and *E. coli* for comparison. *Geobacillus* LC300 had a relatively higher fraction of RNA which is characteristics of fast growing organisms¹⁰.

¹⁰ Dauner, 2001; Neidhart, 1987

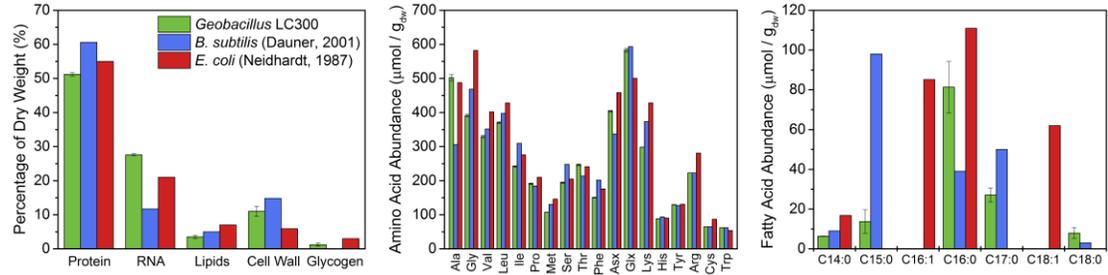


Figure 3.3: Biomass composition of *Geobacillus* LC300 compared to other model organisms. Protein, RNA, lipids, and glycogen were experimentally measured for *Geobacillus* LC300 only

The fraction of proteins in *Geobacillus* LC300 was slightly lower compared to both *B. subtilis* and *E. coli* which is most likely due to the higher mass of RNA. The amino acid distribution of *Geobacillus* LC300 follows similar trends compared to the other organisms. The predominant fatty acids in *Geobacillus* LC300 are saturated C16:0 and C17:0 with smaller amounts of C14:0, C15:0, and C18:0.

3.4 Metabolic network reconstruction

To enable metabolic flux analysis, a detailed metabolic network model for ¹³C metabolic flux analysis (¹³C-MFA) was developed from the genomic annotation. This model is contained in Appendix C. This model includes the identified pathways in central carbon metabolism: glycolysis, pentose phosphate pathway and the TCA cycle. The model also includes amino acid biosynthesis pathways and a lumped reaction for cellular biomass. Exchange reactions account for the exchange of intracellular and atmospheric CO₂. These reactions dilute labeling of metabolites and inclusion of these

reactions increases model fits significantly¹¹. Overall, this model contains 73 reactions, 63 intracellular metabolites, and 9 external metabolites.

3.5 Off-gas and biomass labeling analysis

Geobacillus LC300 was cultured aerobically in either five or six parallel batch cultures. The singly labeled tracers used were [1-¹³C], [2-¹³C], [3-¹³C], [4-¹³C], and [5-¹³C] xylose. This experiment was repeated using the complete set of glucose tracers: [1-¹³C], [2-¹³C], [3-¹³C], [4-¹³C], [5-¹³C], and [6-¹³C] glucose. Figure 3.5.1 shows the off-gas analysis for the [2-¹³C] xylose tracer. ¹³C-Labeling of CO₂ was constant at about 18.3 % ± 0.8%. Figure 3.5.2 shows the off-gas labeling for all five xylose tracers. The off-gas labeling data for the glucose tracers is contained in Appendix D with the CO₂ production profiles for both growth on both substrates.

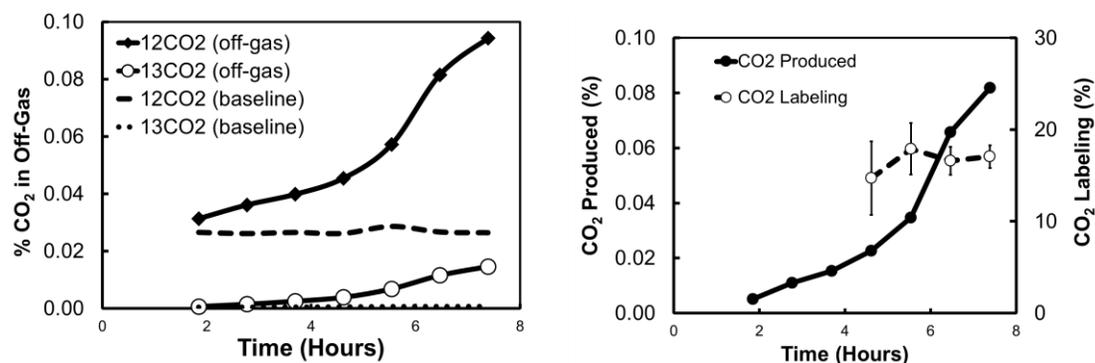
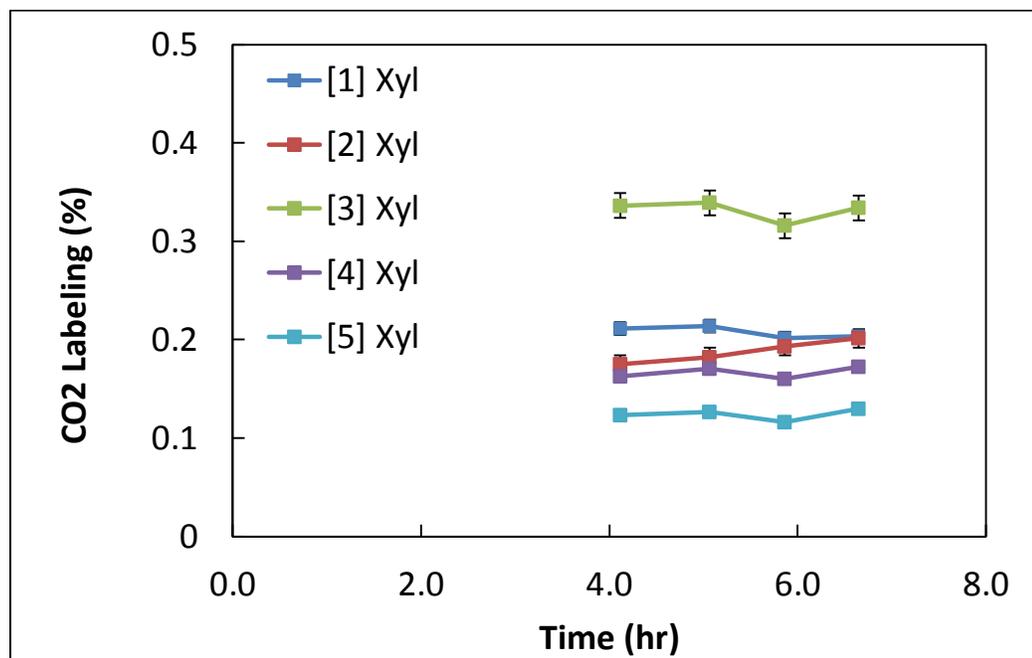


Figure 3.5.1: Production and labeling of CO₂ in off-gas from [2-¹³C] xylose

¹¹ Au, 2014; Leighty, 2012



3.5.2: Average CO₂ labeling for parallel ¹³C xylose tracers as sole substrate for *Geobacillus* LC300

The ¹³C- Labeling of CO₂ was constant for all tracers, suggesting that the cells were at metabolic and isotopic steady state. Table 3.5 shows the average CO₂ labeling for the glucose and xylose tracers.

Table 3.5: Average percent CO₂ labeling for parallel glucose and parallel xylose tracer experiments

Tracer	Average % CO ₂ labeling	SD (± %)	Tracer	Average % CO ₂ labeling	SD (± %)
[1- ¹³ C]glucose	16.0%	3.4%	[1- ¹³ C]xylose	20.9%	0.6%
[2- ¹³ C]glucose	7.9%	1.4%	[2- ¹³ C]xylose	18.3%	0.9%
[3- ¹³ C]glucose	27.3%	1.7%	[3- ¹³ C]xylose	33.0%	1.3%
[4- ¹³ C]glucose	35.6%	2.9%	[4- ¹³ C]xylose	16.4%	0.6%
[5- ¹³ C]glucose	5.9%	0.2%	[5- ¹³ C]xylose	12.2%	0.5%
[6- ¹³ C]glucose	4.4%	0.5%			

The highest average labeling by glucose was with [4-¹³C] glucose, while [3-¹³C] xylose produced the highest labeling by xylose. High labeling of [3-¹³C] xylose, [3-¹³C] glucose, and [4-¹³C] glucose indicates a relatively high contribution of glycolysis (via pyruvate dehydrogenase) to CO₂ production. Additionally, the high labeling of [2-¹³C] and [5-¹³C] xylose indicates a high TCA cycle contribution while the low labeling of these tracers of glucose suggests low TCA cycle contribution to CO₂ production¹². The off-gas analysis of these two experiments identifies variability, as expected, in central carbon metabolism between glucose (hexose substrate) and xylose (pentose substrate).

3.6 Biomass labeling

Incorporation of the substrate tracer leads to labelling of intracellular compounds including amino acids. The ¹³C-labeling of these proteinogenic amino acids was measured by GC-MS for both parallel labeling experiments. To quantitate the carbon flow from xylose to amino acids, the average carbon labeling was calculated for each amino acid fragment following the equation below.

$$^{13}\text{C labeling} = \frac{1}{n} \sum_{i=1}^n i (M+i)$$

In this equation, n is the number of carbon atoms in a given amino acid fragment and M+i is the relative mass isotopomer abundances after natural isotope abundance

¹² Leighty, 2013

correction. Figures 3.6.1 and 3.6.2 show the ^{13}C -labeling of all amino acid fragments for xylose and glucose experiments, respectively.

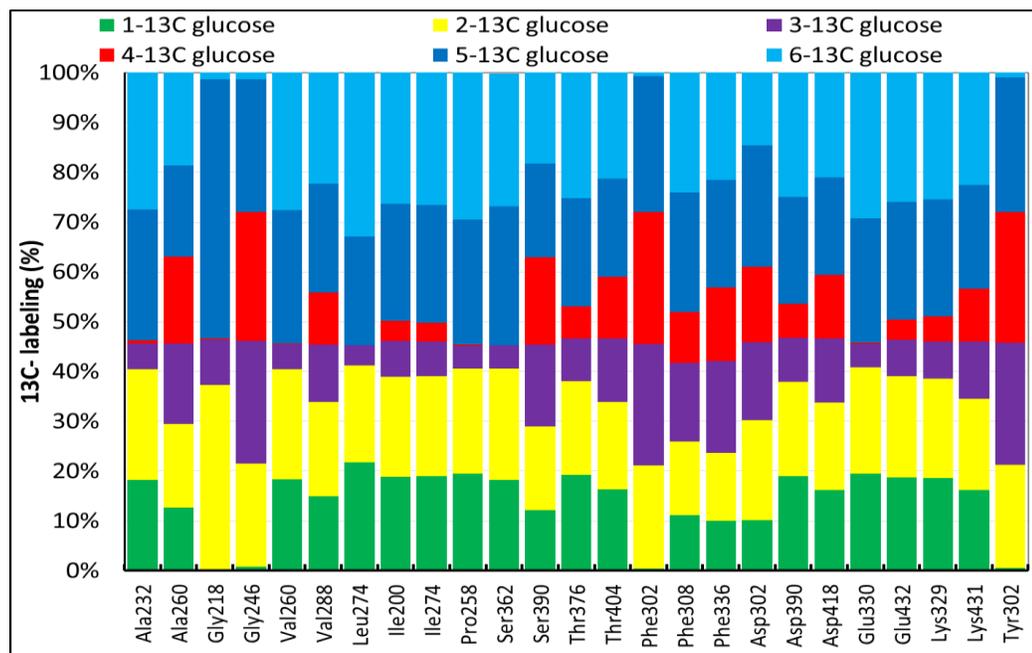


Figure 3.6.1: Average carbon labeling of amino acid fragments from glucose parallel labeling experiment

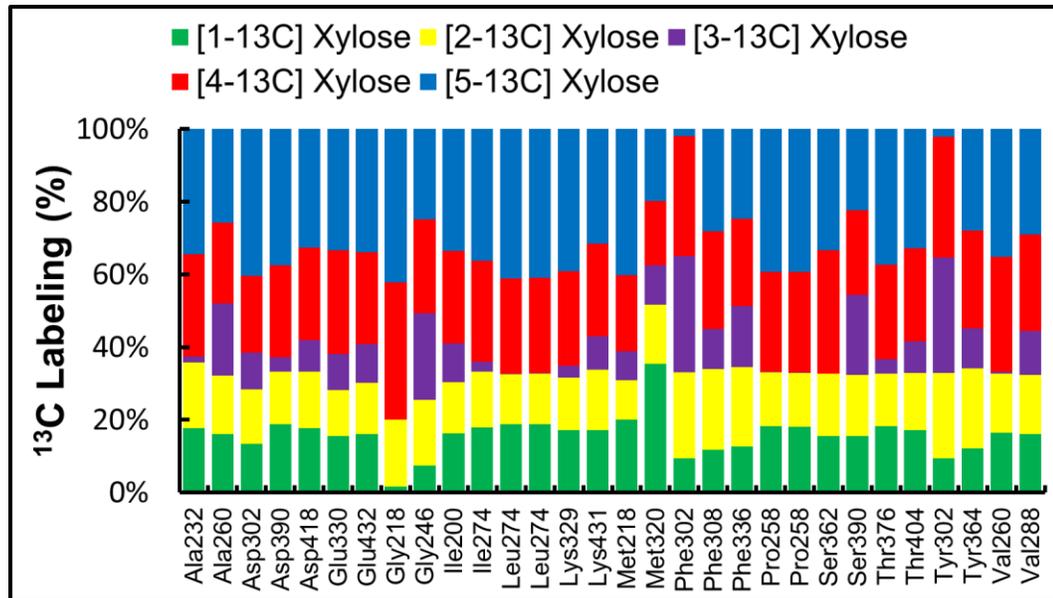


Figure 3.6.2 Average carbon labeling of amino acid fragments from xylose parallel labeling experiment

The expected summed labeling is approximately 97% for each carbon source due to the tracer purity (99% ^{13}C labeled) and a 2% unlabeled inoculum. The summed ^{13}C labeling of glucose ranged from 76% to 95%. For xylose, the summed ^{13}C labeling ranged from 91% to 98%. These two summed labeling ranges suggest that unlabeled carbon atoms were incorporated into amino acids. Additionally, unlabeled dilution was more significant during growth on glucose due to a higher inoculum size. Previous work with aerobic *E. coli* observed a similar carbon dilution for amino acids¹³. The biomass amino acid labeling results confirm that incorporation of unlabeled atmospheric CO_2 must be included in the metabolic model for ^{13}C -flux analysis.

¹³ Leighty, 2012; Leighty, 2013

3.7 Metabolic Flux Analysis

Metabolic fluxes were determined for each tracer experiment (five for xylose, six for glucose) using traditional ^{13}C -metabolic flux analysis (MFA). The combined data set, either glucose or xylose tracers, was analyzed using COMPLETE-MFA where all data sets are fitted simultaneously to a flux model. Using the methods outlined above, the sum of the squared residuals was minimized through iteration. Statistically accepted fits were obtained for all individual experiments, i.e. with [1- ^{13}C], [2- ^{13}C], [3- ^{13}C], [4- ^{13}C], and [5- ^{13}C] xylose tracers and [1- ^{13}C], [2- ^{13}C], [3- ^{13}C], [4- ^{13}C], [5- ^{13}C], [6- ^{13}C] glucose tracers. The COMPLETE-MFA model results were also well within the 95% confidence interval with several model fits deemed ‘over-fitted’.

3.7.1 Metabolic model development

Analysis of biomass amino acid labeling suggested loss of carbon in upper metabolism. This issue was resolved using from the xylose parallel labeling tracer experiment. Following the genome annotation, there are two pathways which enable *Geobacillus* LC300 to lose carbon this early in metabolism: the oxidative pentose phosphate pathway (oxPPP) and the ribulose-monophosphate pathway (RuMP). Due to the carbon arrangement of these two pathways, ^{13}C -MFA based on amino acids cannot identify which pathway is active. In this case, off-gas data analysis was used to confirm the activity of the oxidative pentose phosphate pathway. The oxPPP leads to different ^{13}C -labeling of CO_2 than the RuMP. Inclusion of each pathway individually in COMPLETE-MFA led to a predicted $^{13}\text{CO}_2$ labeling as seen in Figure 3.7.1 below. The oxPPP accurately predicting this labeling while the RuMP underestimated labeling from [1- ^{13}C] xylose and overestimated the other tracers. Additionally, the

measured ^{13}C -labeling of CO_2 was added to the data set for COMPLETE-MFA and yielded an active oxPPP.

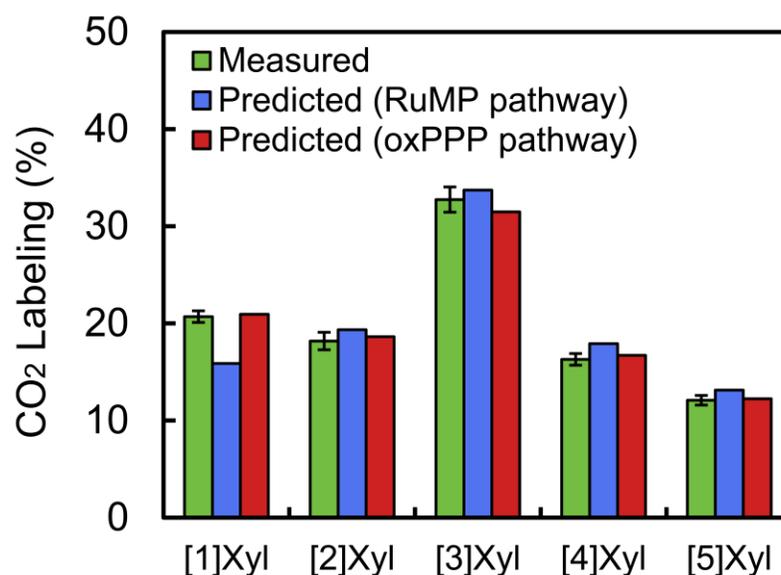


Figure 3.7.1: $^{13}\text{CO}_2$ labeling to elucidate between the ribulose-monophosphate pathway and the oxidative pentose phosphate pathway in upper glycolysis for metabolic modeling

3.7.1.1 Metabolism of *Geobacillus* LC300

Metabolic flux analysis was performed with the base model including the oxidative pentose phosphate pathway for *Geobacillus* LC300 with glucose or xylose as the primary carbon source. The results of COMPLETE-MFA with xylose as the carbon source are shown in Figure 3.7.2 below. The fluxes are normalized to a xylose uptake rate of 100.

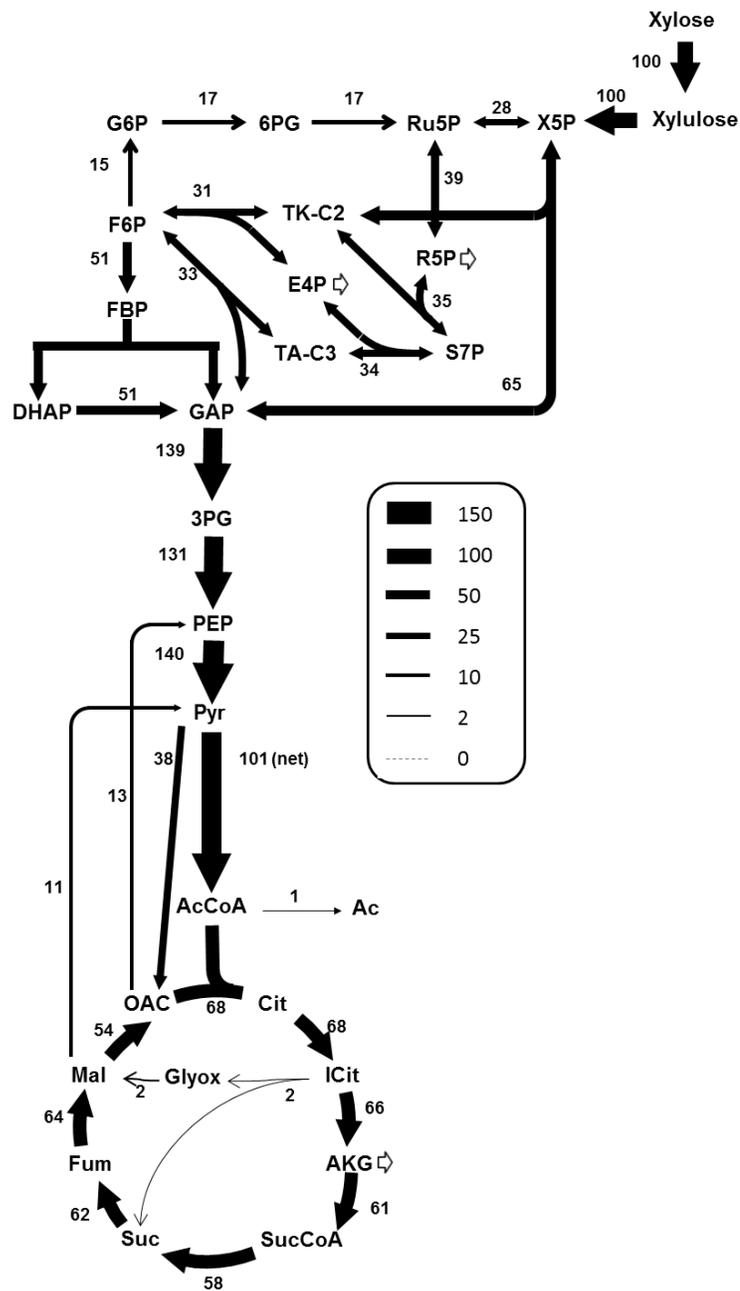


Figure 3.7.2: Flux map of *Geobacillus* LC300 from COMPLETE-MFA for growth on xylose

Xylose metabolism is characterized by high glycolytic and pentose phosphate pathway flux. Upper glycolysis has minimal flux as expected by the entrance site of xylose into central carbon metabolism. The tricarboxylic acid (TCA) cycle was also highly active during xylose consumption.

For comparison, the xylose parallel labeling experiment which generated the data above was repeated with glucose tracers to determine the major differences between hexose and pentose metabolism. Figure 3.7.3 below shows the flux distribution from COMPLETE-MFA of glucose tracers. Again, the fluxes are normalized by glucose uptake, set at 100.

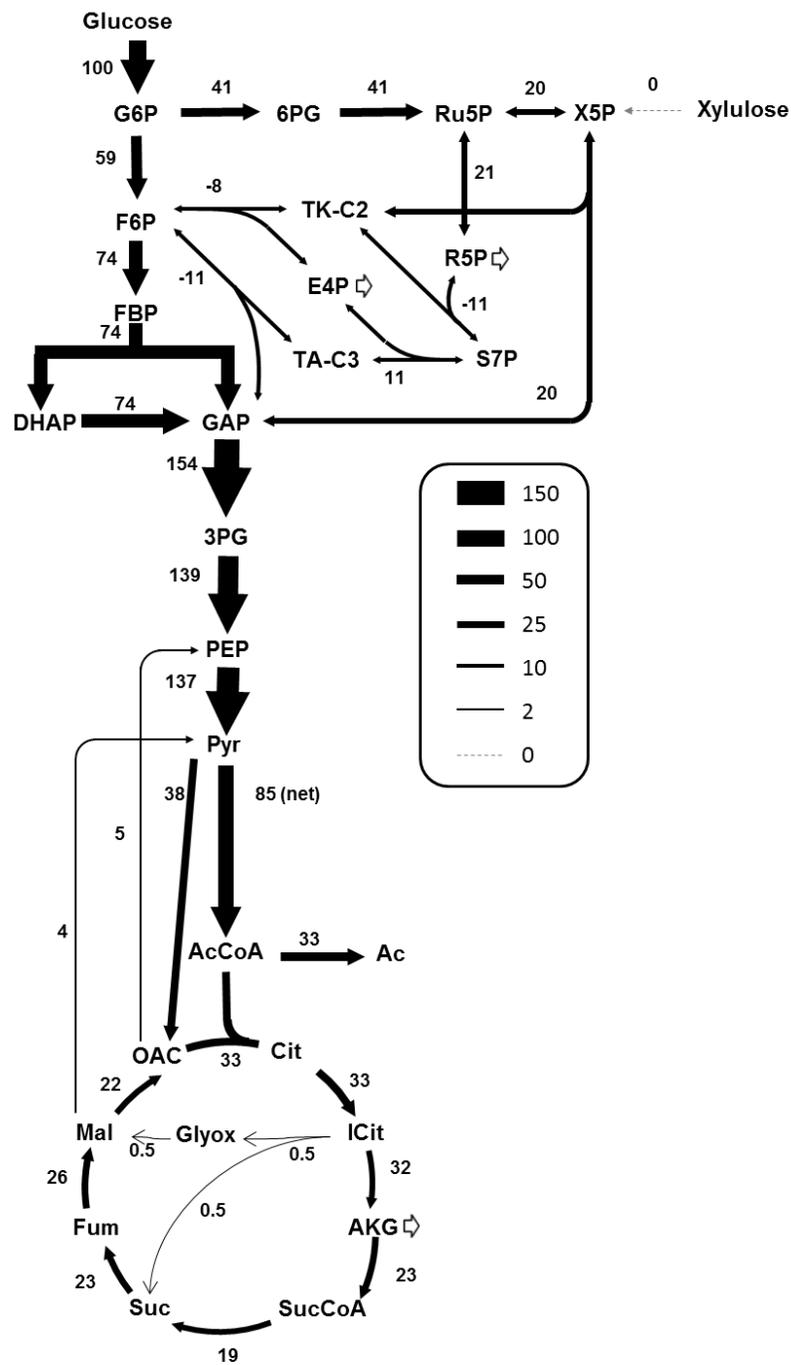


Figure 3.7.3: Flux map of *Geobacillus* LC300 from COMPLETE-MFA for growth on xylose

Glucose metabolism retains high glycolytic flux as seen in xylose metabolism. It does have significantly lower flux through the TCA cycle and the pentose phosphate pathway. Lower pentose phosphate pathway activity is expected as this pathway only serves to generate nucleotides and other complex molecules during growth on glucose. The lessened TCA cycle activity aligns with the off-gas CO₂ labeling data from the parallel labeling experiment.

3.8 Scale-up analysis

3.8.1 Batch Experiments

Geobacillus LC300 was cultured in a 2L bioreactor with a working volume of 1.5 L. The goal of these experiments was to reach the highest cell density possible. This quantity is directly related to the number and weight of cells in culture by an experimentally determined conversion factor.

A series of batch experiments determined several limiting factors of cell growth. Under no control schemes, the maximum culture density was 1.73 with consumption of 2 mM glucose. After 4 hours of growth, both the dissolved oxygen and the pH of this uncontrolled culture starting decreasing. To increase the culture density, pH and dissolved oxygen (DO) were placed under bioreactor control both individually and collectively. Figures 3.9.1, and 3.9.2 show the measured culture density and the pH of these batch cultures, respectively.

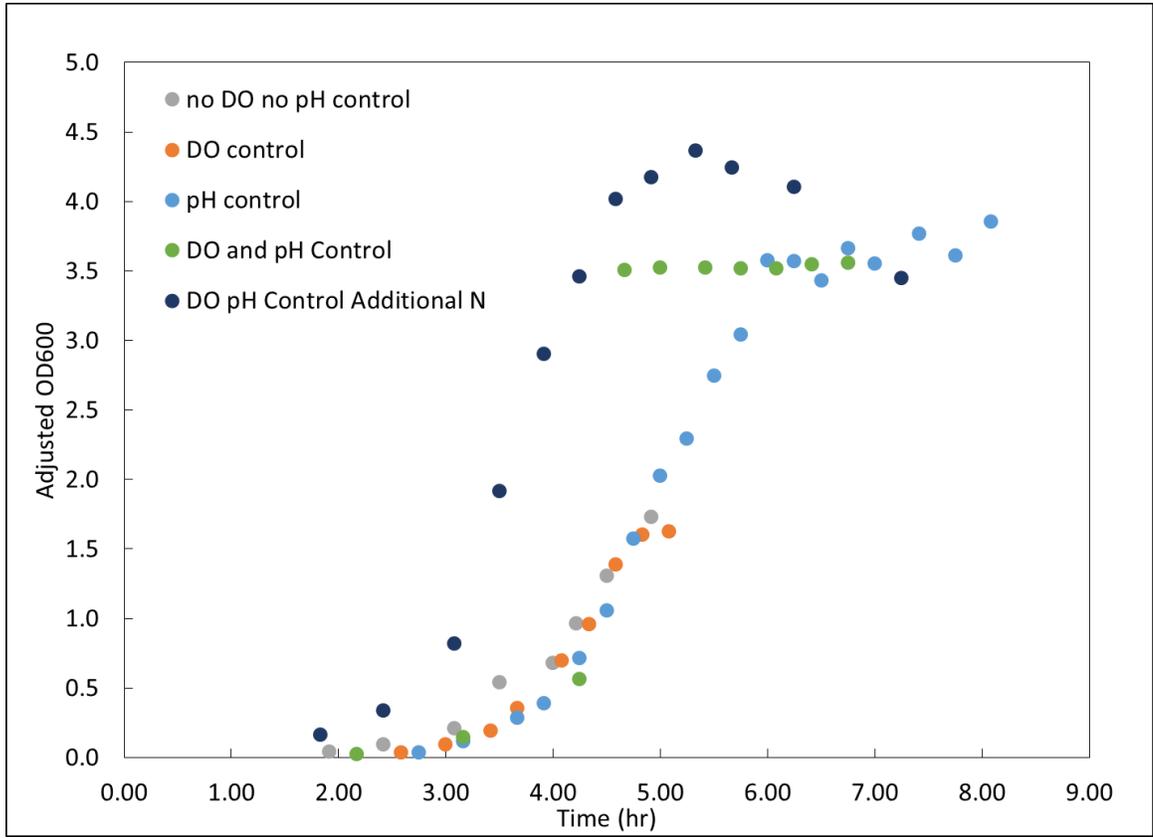


Figure 3.9.1: *Geobacillus* LC300 Cell Density (OD_{600}) during growth in batch with various control schemes

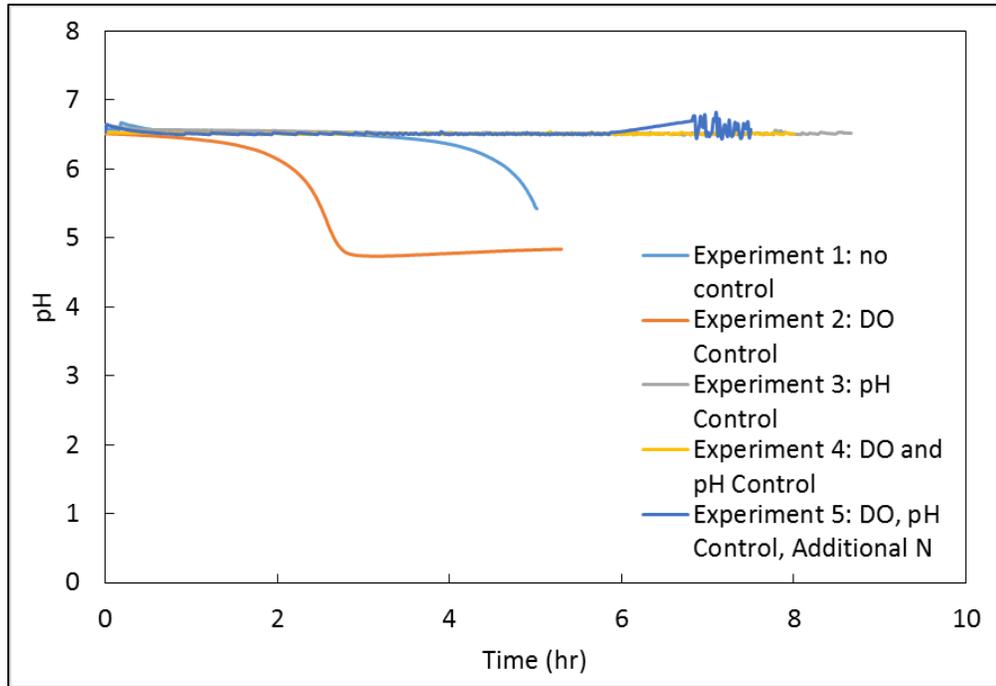


Figure 3.9.2 Measured pH of *Geobacillus* LC300 under batch culture various control schemes

Dissolved oxygen and pH were the first two limiting factors to cell growth. Culture pH is a key limitation to reaching higher cell densities, as this level of control led to significant increase in culture density from an original maximum of approximately 1.73 to a new maximum of 3.85. Successful control of culture pH confirmed *Geobacillus* LC300 produces acidic compounds during growth. The accumulation of these compounds and the subsequent decrease in pH are cause of cell death for uncontrolled cultures, such as the 10mL mini-bioreactors.

Control of dissolved oxygen led to a more rapid pH drop at an earlier time in the culture. This highlighted the need to control both of these reactor parameters to achieve high density culture. While dual control led to faster growth, it did not have an

additive effect on culture density as seen in Figure 3.9.1. At this level of growth, it was suspected that *Geobacillus* LC300 had consumed the available supply of nitrogen. To confirm, an extra bolus of ammonium chloride (500 mg) was added during inoculation. As Figure 3.9.1 above shows, the culture grew faster and to a higher density. During all of these experiments, the glucose concentration was monitored. The data are contained within Appendix E. For these controlled cultures, the initial glucose was consumed or nearly consumed before the end of cell growth.

3.8.2 Fed-Batch Experiments

To prevent the depreciation of substrate, the remaining experiments were conducted under fed-batch conditions. The glucose feed attempted to maintain a constant glucose concentration of 20mM within the reactor. The initial feed solution contained only glucose and nitrogen in stoichiometric amounts. The substrate feed enabled the density to rise to a maximum density of 4.23 as seen in Figure 3.9.3. Addition of single bolus of Wolfe's Vitamins and Minerals led to a further increase to a density of 7.35.

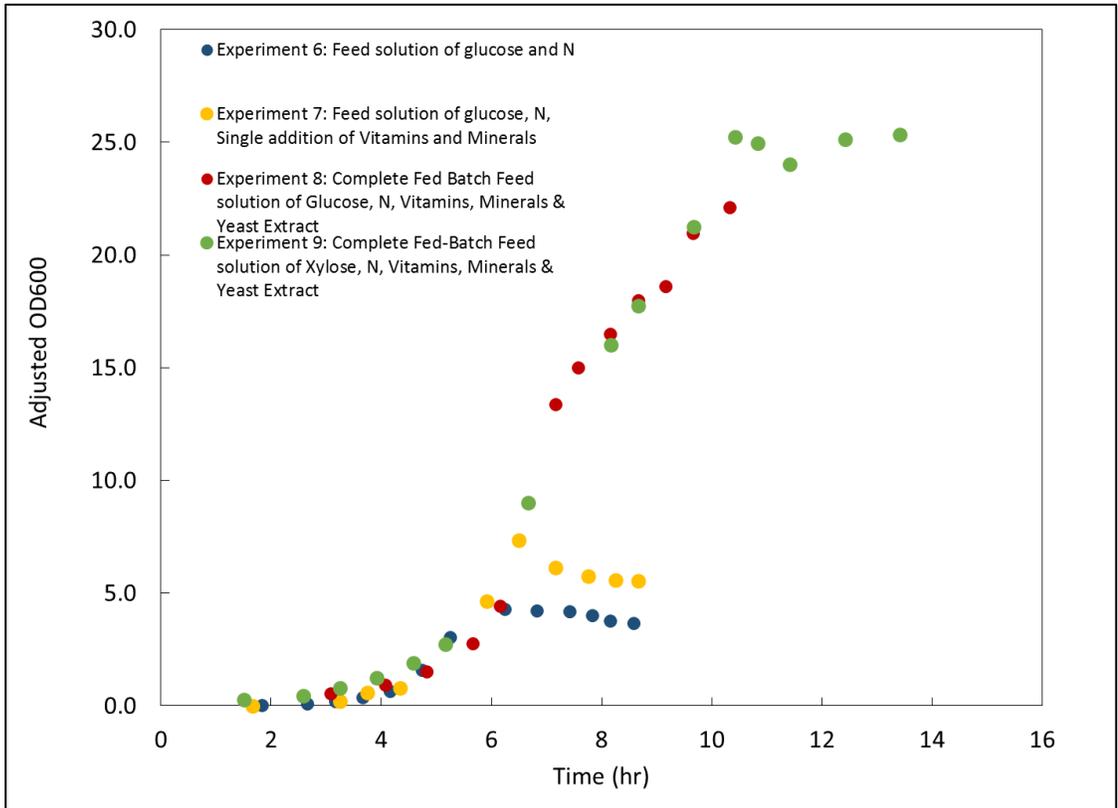


Figure 3.9.3: *Geobacillus* LC300 Cell Density (OD_{600}) during fed-batch growth with various control schemes

A feed solution containing glucose, ammonium chloride (Nitrogen source), Wolfe's Vitamins, Minerals and yeast extract achieved the highest cell densities. The feed solution contained stoichiometric amounts of the listed components with a 5:1 g/g ratio for glucose to yeast extract. Using this high-nutrient feed with glucose, a maximum cell density of 22.12 was achieved. At such high density, by-product accumulation is likely the next limiting factor. As this cannot be prevented or circumvented with existing laboratory technology, an OD_{600} of 22 was deemed the maximum density for growth on glucose.

This final, high nutrient fed-batch experiment was repeated using xylose since *Geobacillus* LC300 is typically cultured with xylose as the sole substrate. Under these growth conditions, the cell density exceeded that of glucose growth with a maximum OD₆₀₀ of 25.35.

In both high nutrient, fed-batch experiments, *Geobacillus* LC300 reached a maximum density but continued to be metabolically active. Over the course of the experiment, *Geobacillus* LC300 produced acetate and lactate. The full table of time course data for each of these experiments is contained within Appendix D. During growth on glucose, *Geobacillus* LC300 produced acetate at high levels (over 100 mM) while significantly less acetate was produced during growth on xylose (maximum of 32 mM). Lactate was also produced at minimal levels (0.470 mM) during both experiments. Additionally, the dissolved oxygen in both of these cultures was not able to maintain 40% during the last few hours due to the agitator reaching its maximum speed of 1000 rpm. At this point in the experiment, the environment was microaerobic rather than aerobic as designed.

Successful culture of *Geobacillus* LC300 to densities above 20 suggests that this organism is feasible for medium scale industrial production. The lack of high levels of acetate production during growth on xylose indicate either the presence of an additional fermentation product or excessive glycolysis. Xylose was continuously added during the course of this experiment and the concentration remained low until the experiment was terminated. This suggests *Geobacillus* LC300 continued to be metabolically active at high cell densities and low dissolved oxygen. The continual consume xylose while not growing indicates either high turnover of glycolysis to maintain cell function, or the production of an additional fermentation product.

Additional research and experiments are necessary to determine which of these two options is physically occurring under high density, microaerobic conditions.

Chapter 4

CONCLUSIONS

4.1 Conclusions

In this work, a *Geobacillus* strain LC300 was extensively studied to understand the genomic, growth, and metabolic features. *Geobacillus* LC300 was also cultured at the benchtop bioreactor scale (2L) to obtain high density cultures which is industrially relevant.

Geobacillus LC300 was successfully isolated and the genome sequenced and annotated. High sequence similarity led to identification as a new *Geobacillus* strain termed LC300. The 3.49 Mbp chromosome contained 3988 coding sequences and 11 tRNA coding regions. From annotation using the RAST server, the major metabolic pathways are deemed complete in *Geobacillus* LC300. Glycolysis, the pentose phosphate pathway (oxidative and nonoxidative), the tricarboxylic acid (TCA) cycle, the glyoxylate shunt and the ribulose monophosphate (RuMP) pathway are present in the genome while the Entner-Doudoroff pathway is not. One of the enzymes responsible for the oxidative pentose phosphate pathway was not identified but the reaction can proceed spontaneously completing the pathway. All amino acid biosynthesis pathways were also complete with the exception of an enzyme for serine biosynthesis. This enzyme, phosphoserine phosphatase, has not been identified in any sequenced *Geobacilli* genome.

The growth of *Geobacillus* LC300 was well characterized to fully understand its growth behavior. Optimal conditions for *Geobacillus* are aerobic, a temperature of 72°C, and an initial xylose concentration of 30 g/L. Under these conditions, the average specific growth rate was 1.52 hr⁻¹ which is 3 times higher than that of *E. coli*

under the same conditions. *Geobacillus* LC300 also follows Monod growth kinetics with the maximum growth rate with 30 g/L xylose. For an accurate metabolic model, the biomass composition was determined with *Geobacillus* LC300 having a higher relative fraction of RNA.

A metabolic model was reconstructed for *Geobacillus* LC300 for metabolic flux analysis. This model included all identified central carbon metabolism pathways, amino acid biosynthesis and a lumped reaction for biomass accumulation. Off-gas labeling indicated ^{13}C - labeling experiments were at metabolic and isotopic steady state. These labeling experiments were completed using singly-labeled glucose and xylose tracers to elucidate the differences in metabolism. The CO_2 average labeling indicated high glycolysis activity during growth on glucose and xylose. The TCA cycle has high activity when consuming xylose and lower activity when metabolizing glucose. Biomass labeling of amino acids indicated the necessity of a CO_2 exchange flux to include the incorporation of unlabeled atmospheric CO_2 .

Metabolic flux analysis (^{13}C -MFA) was completed for individually for each tracer and then in combination using COMPLETE-MFA for each data set, either glucose or xylose tracers. Use of off-gas CO_2 labeling data in conjunction with flux analysis confirmed the activity of the oxidative pentose phosphate pathway. As expected, xylose metabolism is characterized by high glycolytic and pentose phosphate pathway flux, while glucose metabolism has a high glycolytic and low pentose phosphate pathway flux. As predicted by the CO_2 off-gas analysis, the TCA cycle is more active during xylose metabolism than during glucose metabolism.

The final portion of this work successfully scaled *Geobacillus* L300 from 10 mL mini-bioreactors to a 1.5 L working volume reactor. A series of batch and fed-

batch experiments identified the necessity of dissolved oxygen and pH control as well as continuous addition of nutrients to reach high cell densities. The maximum achieved cell density was 22 and 25 for growth on glucose and xylose, respectively. Acetate and lactate production were detected during both of these high cell density experiments. Lower production of acetate during growth on xylose indicates differences in cell metabolism based on the substrate. Additionally, xylose consumption with minimal increases in biomass suggests an additional unknown fermentation product. This analysis identified that *Geobacillus* LC300 can be grown to high densities in larger volume reactors.

4.2 Recommendations for future work

From the analysis presented in this thesis, *Geobacillus* LC300 is well characterized for future experiments. Under the xylose growth conditions outlined above, the only significant by-product is acetate at moderate levels. The experimental results suggest the production of an additional fermentation product. Future studies are necessary to determine if these results are reproducible with consistent growth behavior. Additionally, off-gas analysis and high performance liquid chromatography (HPLC) analysis would determine the identity of the unknown fermentation product. If this product is of scientific interest, further experiments could optimize the system for high production.

In addition to the experiments presented above, *Geobacillus* LC300 contains highly efficient xylose transporters as evident by the specific uptake rate. Over-expression, characterization and use of these enzymes in other microbial systems would increase understanding and efficiency of substrate uptake.

REFERENCES

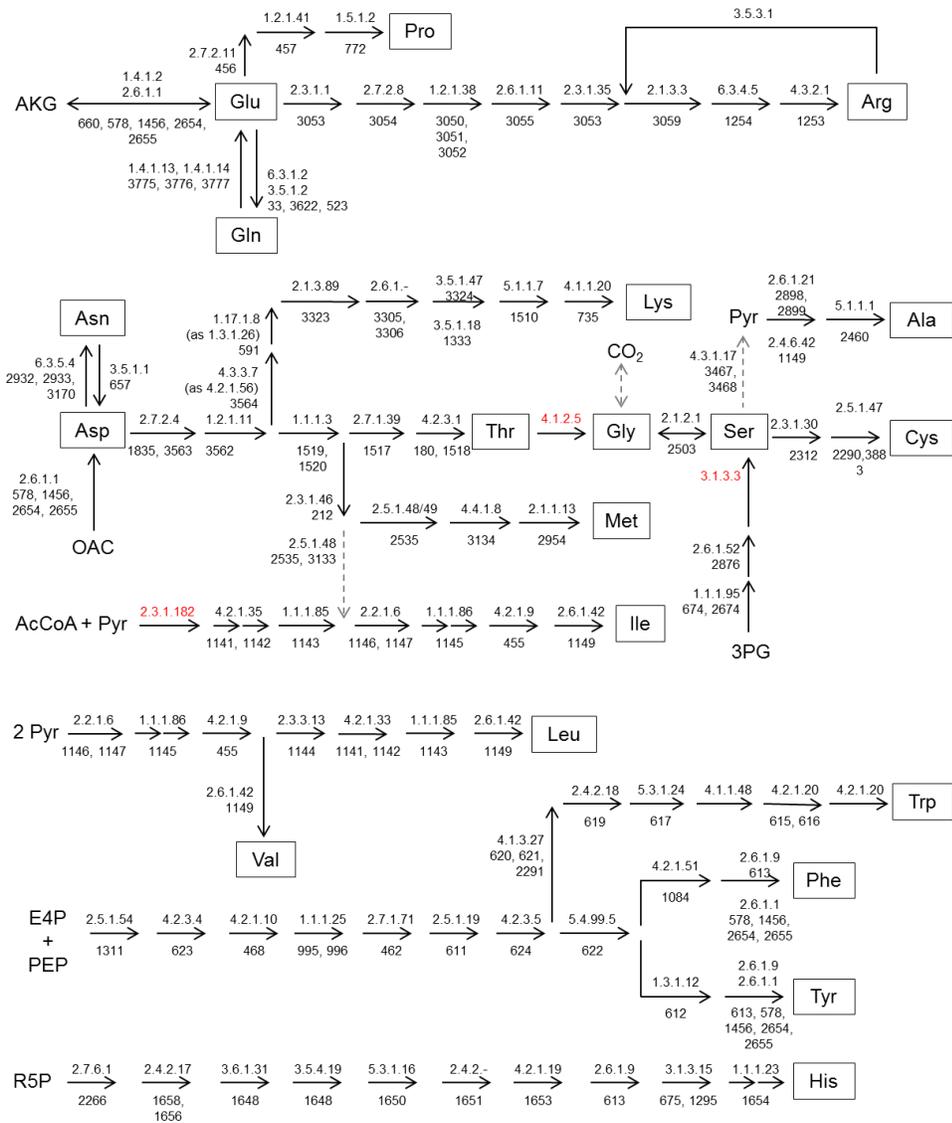
1. Lynd, L.R. (1987) Production of ethanol from lignocellulosic materials using thermophilic bacteria. United States: Hanover, NH (US); Dartmouth College.
2. Cripps, R. E., Eley, K., Leak, D. J., Rudd, B., Taylor, M., Todd, M., Boakes, S., Martin, S., Atkinson, T (2009) Metabolic engineering of *Geobacillus thermoglucosidasius* for high ethanol production. *Metabolic Engineering*, 11, 398-408.
3. Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, C., Liu, X., Han, W., Peng, X., Liu, R., Wang, L. (2007) Genome and proteome of long-chain degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. *Proceedings of the National Academy of Sciences USA*. 104, 5602-5607.
4. Verma, D., Anand, A., Satyanarayana T. (2013) Thermostable and Alkalistable Endoxylanase of the Extremely Thermophilic Bacterium *Geobacillus thermodenitrificans* TSAA1: Cloning, Expression, Characteristics and Its Applicability in Generating Xylooligosaccharides and Fermentable Sugars. *Applied Biochemistry and Biotechnology*. 170, 119-130.
6. Ezeji, T. C., Wolf, A., Bahl, H., (2005) Isolation, characterization and identification of *Geobacillus thermodenitrificans* HRO10, an α -amylase and α -glucosidase producing thermophile. *Canadian Journal of Microbiology*. 51, 685-693.
7. Coorevits, A., Dinsdale, A. E., Halket, G., Lebbe, L., De Vos, P., Van Landschoot, A., Logan, N. A., (2012) Taxonomic revision of the genus *Geobacillus*: emendation of *Geobacillus*, *G. stearothermophilus*, *G. jurassicus*, *G. toebii*, *G. thermodenitrificans* and *G. thermoglucosidans* (nom. corrig., formerly 'thermoglucosidasius'); transfer of *Bacillus thermantarcticus* to the genus as *G. thermantarcticus* comb. nov.; proposal of *Caldibacillus debilis* gen. nov., comb. nov.; transfer of *G. tepidamans* to *Anoxybacillus* as *A. tepidamans* comb. nov.; and proposal of *Anoxybacillus caldiproteolyticus* sp. nov. *International journal of systematic and evolutionary microbiology*. 62, 1470-85.
8. Long, C. P., Antoniewicz, M. R., (2014) Quantifying biomass composition by gas chromatography/mass spectrometry. *Analytical Chemistry*. 86, 9423-7.

9. Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formisano, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. 9, 75.
10. Antoniewicz, M. R., Kelleher, J. K., Stephanopoulos, G., (2007) Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Analytical Chemistry*. 79, 7554-9.
11. Kupor, S. R., Fraenkel, D. G., (1969) 6-phosphogluconolactonase mutants of *Escherichia coli* and a maltose blue gene. *Journal of Bacteriology*. 100, 1296-301.
12. Thomason, L. C., Court, D. L., Datta, A. R., Khanna, R., Rosner, J. L., (2004) Identification of the *Escherichia coli* K-12 *ybhE* gene as *pgl*, encoding 6-phosphogluconolactonase. *Journal of Bacteriology*. 186, 8248-53.
13. Zimenkov, D., Gulevich, A., Skorokhodova, A., Biriukova, I., Kozlov, Y., Mashko, S., (2005) *Escherichia coli* ORF *ybhE* is *pgl* gene encoding 6-phosphogluconolactonase (EC 3.1.1.31) that has no homology with known 6PGLs from other organisms. *FEMS Microbiology Letters*. 244, 275-80.
14. Dauner, M., Sauer, U., (2001) Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. *Biotechnology and Bioengineering*. 76, 132-43.
15. Neidhardt, F. C., (1987) *Escherichia Coli* and *Salmonella Typhimurium*. ASM Press.
16. Au, J., Choi, J., Jones, S. W., Venkataramanan, K. P., Antoniewicz, M. R., (2014) Parallel labeling experiments validate *Clostridium acetobutylicum* metabolic network model for C metabolic flux analysis. *Metabolic Engineering*. 26, 23-33.
17. 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. *Metabolic Engineering*. 14, 533-41.
18. Leighty, R. W., Antoniewicz, M. R., (2013) COMPLETE-MFA: complementary parallel labeling experiments technique for metabolic flux analysis. *Metabolic Engineering*. 20, 49-55.

Appendix A

Amino Acid Biosynthesis Pathways

Amino Acid Biosynthesis Pathways – *Geobacillus* LC300



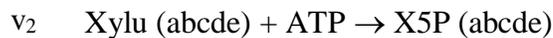
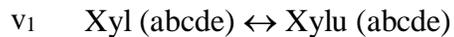
EC numbers highlighted in red were not identified from *Geobacillus* LC300 genome annotation.

Appendix B

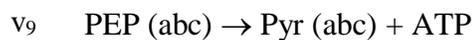
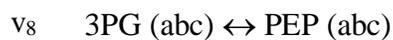
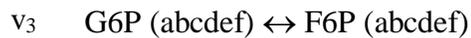
Geobacillus LC300 Metabolic Model

The *Geobacillus* LC300 metabolic model was reconstructed from genome annotation data. 73 reactions were included within this model and are shown below by metabolism classification. The lowercase letters (abcdef) correspond to the carbon transitions, which must be specified. The lumped biomass formation equation was determined experimentally as previously specified.

Xylose Metabolism



Glycolysis



Pentose Phosphate Pathway

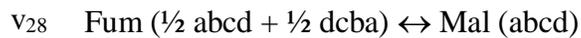
- v10 G6P (abcdef) \rightarrow 6PG (abcdef) + NADPH
- v11 6PG (abcdef) \rightarrow Ru5P (bcdef) + CO₂ (a) + NADPH
- v12 Ru5P (abcde) \leftrightarrow X5P (abcde)
- v13 Ru5P (abcde) \leftrightarrow R5P (abcde)
- v14 X5P (abcde) \leftrightarrow TK-C2 (ab) + GAP (cde)
- v15 F6P (abcdef) \leftrightarrow TK-C2 (ab) + E4P (cdef)
- v16 S7P (abcdefg) \leftrightarrow TK-C2 (ab) + R5P (cdefg)
- v17 F6P (abcdef) \leftrightarrow TA-C3 (abc) + GAP (def)
- v18 S7P (abcdefg) \leftrightarrow TA-C3 (abc) + E4P (defg)

Ribulose Monophosphate Pathway

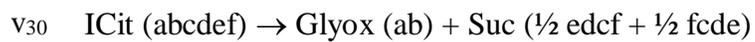
- v19 F6P (abcdef) \leftrightarrow Hu6P (abcdef)
- v20 Hu6P (abcdef) \leftrightarrow Ru5P (bcdef) + Formald (a)

TCA Cycle

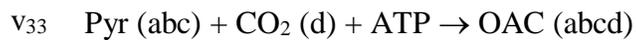
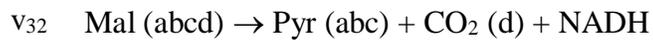
- v21 Pyr (abc) \rightarrow AcCoA (bc) + CO₂ (a) + NADH
- v22 OAC (abcd) + AcCoA (ef) \rightarrow Cit (dcbfea)
- v23 Cit (abcdef) \leftrightarrow ICit (abcdef)
- v24 ICit (abcdef) \leftrightarrow AKG (abcde) + CO₂ (f) + NADPH



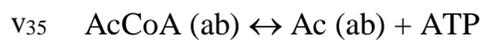
Glyoxylate Shunt



Amphibolic Reactions



Acid Formation



Amino Acid Biosynthesis

- v37 AKG (abcde) + NADPH + NH₃ → Glu (abcde)
- v38 Glu (abcde) + ATP + NH₃ → Gln (abcde)
- v39 Glu (abcde) + ATP + 2 NADPH → Pro (abcde)
- v40 Glu (abcde) + CO₂ (f) + Gln (ghijk) + Asp (lmno) + AcCoA (pq) + 5 ATP + NADPH
→ Arg (abcdef) + AKG (ghijk) + Fum (lmno) + Ac (pq)
- v41 OAC (abcd) + Glu (efghi) → Asp (abcd) + AKG (efghi)
- v42 Asp (abcd) + 2 ATP + NH₃ → Asn (abcd)
- v43 Pyr (abc) + Glu (defgh) → Ala (abc) + AKG (defgh)
- v44 3PG (abc) + Glu (defgh) → Ser (abc) + AKG (defgh) + NADH
- v45 Ser (abc) ↔ Gly (ab) + MEETHF (c)
- v46 Gly (ab) ↔ CO₂ (a) + MEETHF (b) + NADH + NH₃
- v47 Ser (abc) + AcCoA (de) + 3 ATP + 4 NADPH + SO₄ → Cys (abc) + Ac (de)
- v48 Asp (abcd) + Pyr (efg) + Glu (hijkl) + SucCoA (mnop) + ATP + 2 NADPH →
LL-DAP (½ abcdgfe + ½ efgdcba) + AKG (hijkl) + Suc (½ mnop + ½ ponm)
- v49 LL-DAP (½ abcdefg + ½ gfedcba) → Lys (abcdef) + CO₂ (g)
- v50 Asp (abcd) + 2 ATP + 2 NADPH → Thr (abcd)
- v51 Asp (abcd) + METHF (e) + Cys (fgh) + SucCoA (ijkl) + ATP + 2 NADPH →
Met (abcde) + Pyr (fgh) + Suc (½ ijkl + ½ lkji) + NH₃
- v52 Pyr (abc) + Pyr (def) + Glu (ghijk) + NADPH → Val (abcef) + CO₂ (d) + AKG
(ghijk)

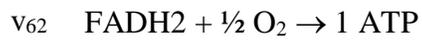
- v53 AcCoA (ab) + Pyr (cde) + Pyr (fgh) + Glu (ijklm) + NADPH →
Leu (abdghe) + CO₂ (c) + CO₂ (f) + AKG (ijklm) + NADH
- v54 Thr (abcd) + Pyr (efg) + Glu (hijkl) + NADPH → Ile (abfcdg) + CO₂ (e) + AKG
(hijkl) + NH₃
- v55 PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) + ATP + NADPH →
Phe (abcefg hij) + CO₂ (d) + AKG (klmno)
- v56 PEP (abc) + PEP (def) + E4P (ghij) + Glu (klmno) + ATP + NADPH →
Tyr (abcefg hij) + CO₂ (d) + AKG (klmno) + NADH
- v57 Ser (abc) + R5P (defgh) + PEP (ijk) + E4P (lmno) + PEP (pqr) + Gln (stuvw) + 3
ATP + NADPH →
Trp (abcdklmnoj) + CO₂ (i) + GAP (fgh) + Pyr (pqr) + Glu (stuvw)
- v58 R5P (abcde) + FTHF (f) + Gln (ghijk) + Asp (lmno) + 5 ATP →
His (edcbaf) + AKG (ghijk) + Fum (lmno) + 2 NADH

One-Carbon Metabolism

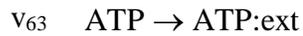
- v59 MEETHF (a) + NADH → METHF (a)
- v60 MEETHF (a) → FTHF (a) + NADPH

Oxidative Phosphorylation

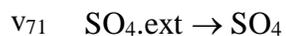
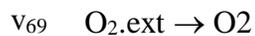
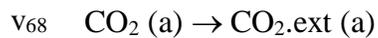
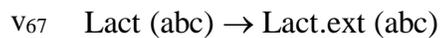
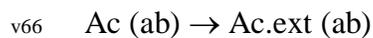
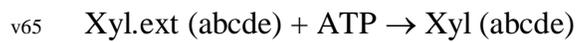
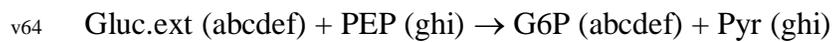
- v61 NADH + ½ O₂ → 2 ATP



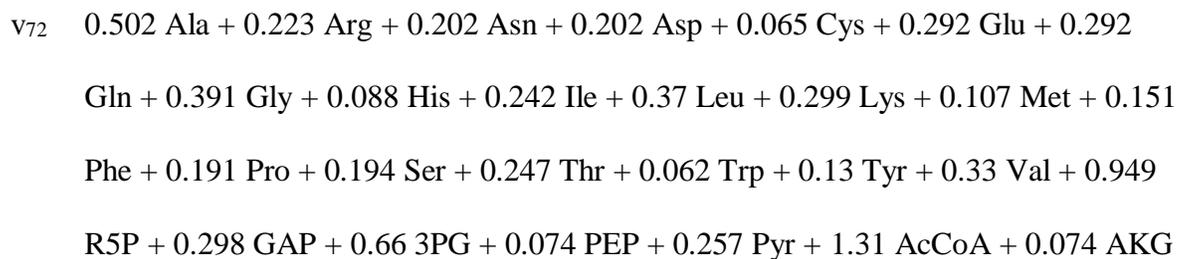
ATP Hydrolysis



Transport



Biomass Formation



+ 0.48 OAC + 0.074 G6P + 0.479 F6P + 35.92 ATP + 3.549 NADPH + 0.571

MEETHF → 39.67 Biomass + 1.525 NADH

CO₂ Exchange

v_{73} CO₂.unlabeled (a) + CO₂ (b) → CO₂ (a) + CO₂.out (b)

The net effect of reaction v_{73} is exchange of intracellular CO₂ for an unlabeled CO₂ without affecting intracellular carbon balances.

Appendix C

CO₂ Off-gas Data from Parallel Labeling Experiments

Table C.1: Average CO₂ labeling of off-gas produced from growth on ¹³C-labeled glucose or xylose

Tracer	Average % CO ₂ labeling	SD (± %)	Tracer	Average % CO ₂ labeling	SD (± %)
[1- ¹³ C]glucose	16.0%	3.4%	[1- ¹³ C]xylose	20.9%	0.6%
[2- ¹³ C]glucose	7.9%	1.4%	[2- ¹³ C]xylose	18.3%	0.9%
[3- ¹³ C]glucose	27.3%	1.7%	[3- ¹³ C]xylose	33.0%	1.3%
[4- ¹³ C]glucose	35.6%	2.9%	[4- ¹³ C]xylose	16.4%	0.6%
[5- ¹³ C]glucose	5.9%	0.2%	[5- ¹³ C]xylose	12.2%	0.5%
[6- ¹³ C]glucose	4.4%	0.5%			

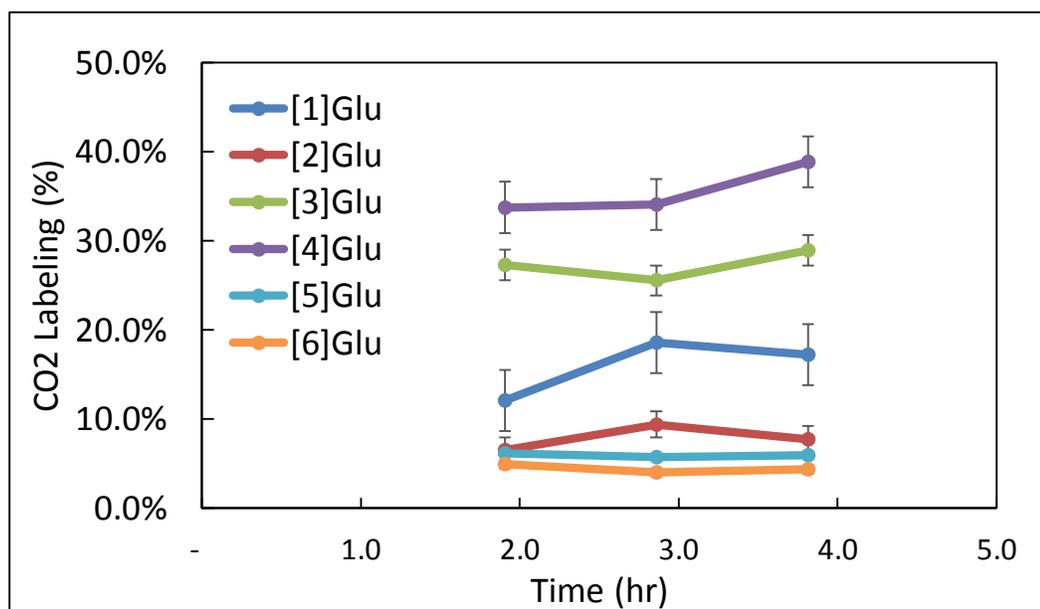


Figure C.2: Average CO₂ labeling (%) in CO₂ produced in off-gas during growth on labeled glucose.

Produced CO₂ and CO₂ labeling for complete glucose and xylose tracer experiments is shown below

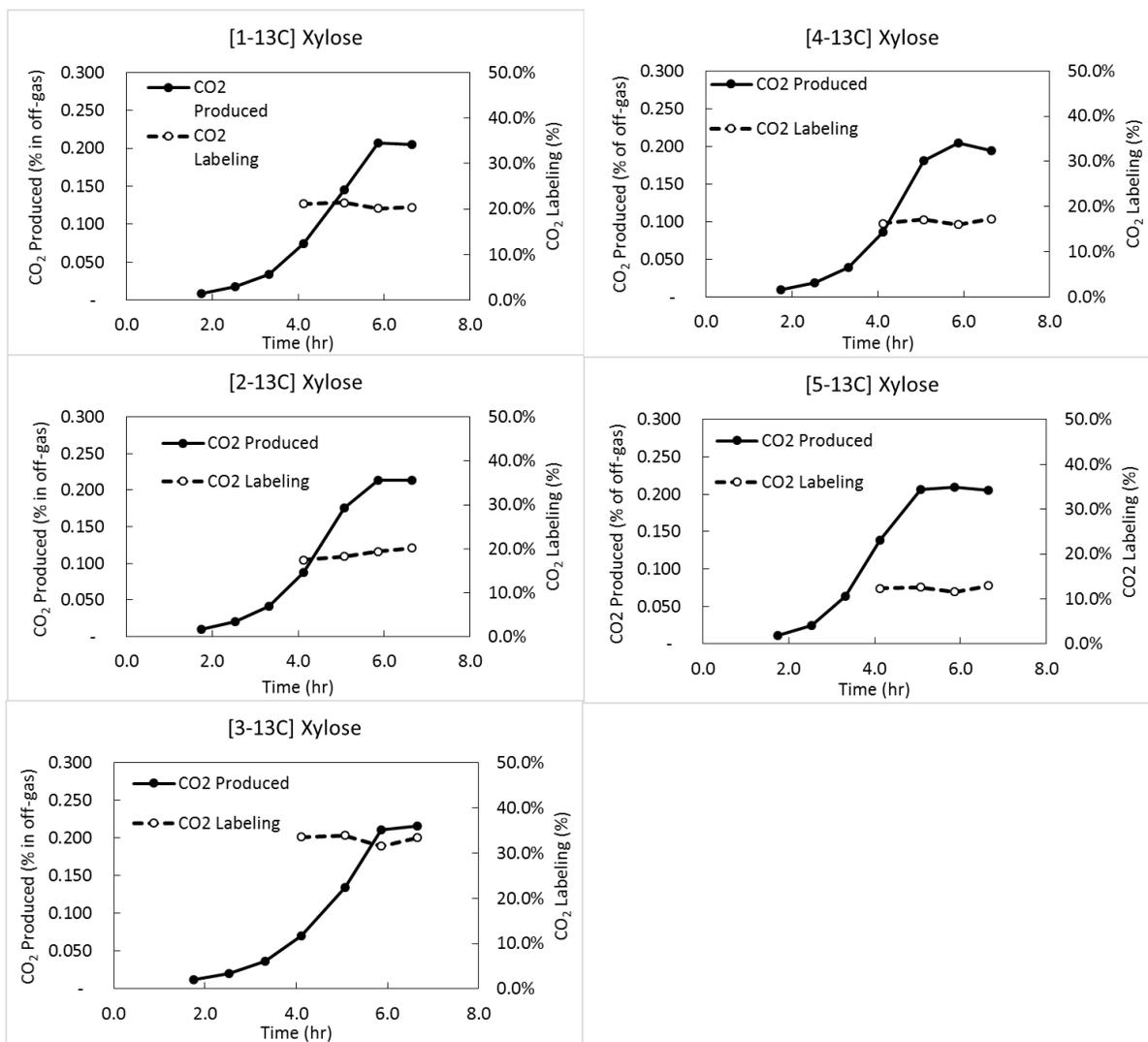


Figure C.3: CO₂ Production and CO₂ Labeling for *Geobacillus* LC300 growth on labeled xylose

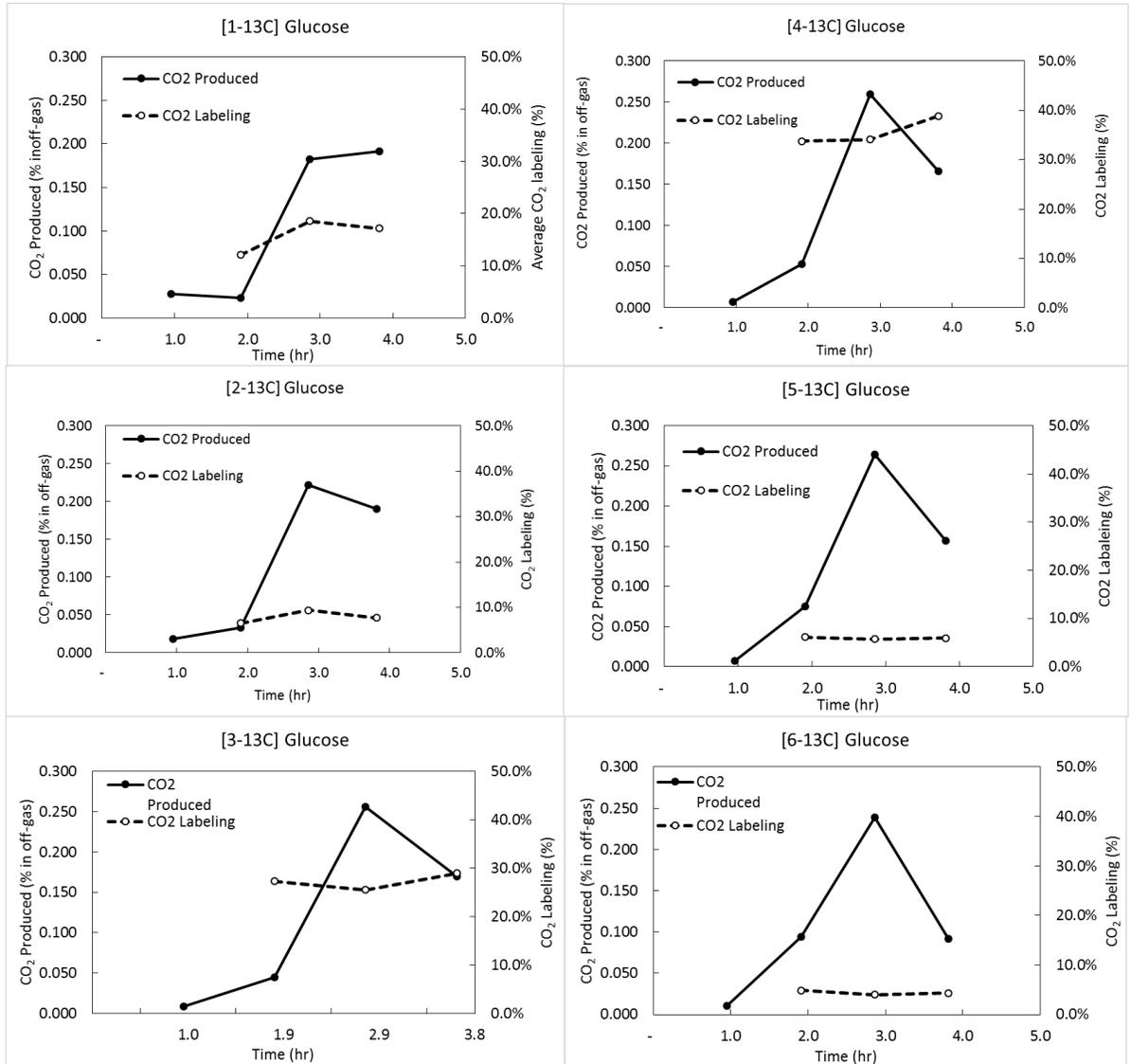


Figure C.4: CO₂ Production and CO₂ Labeling for *Geobacillus* LC300 growth on labeled glucose

Appendix D

Benchtop Bioreactor Experimental Data

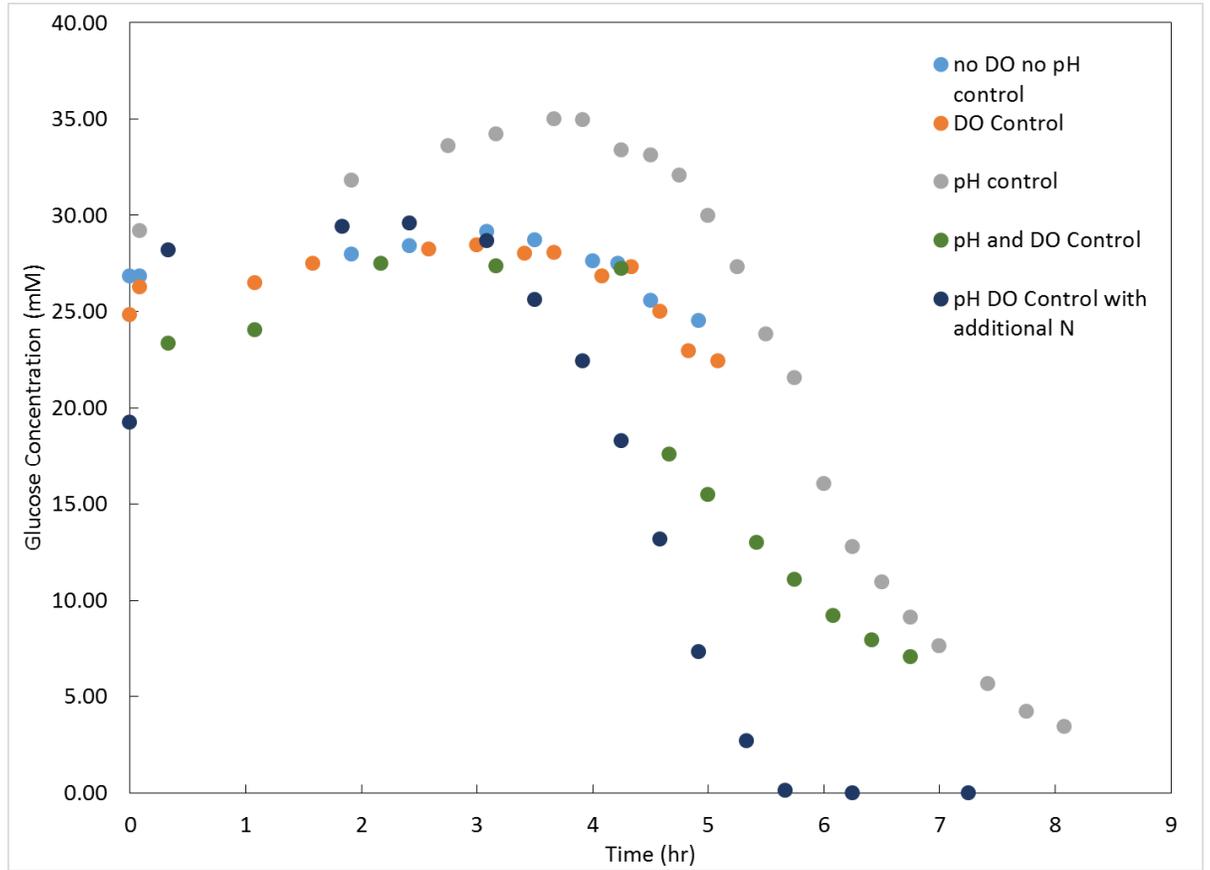


Figure D.1: Glucose concentration profile for batch experiments with 2L scale bioreactor under various controls

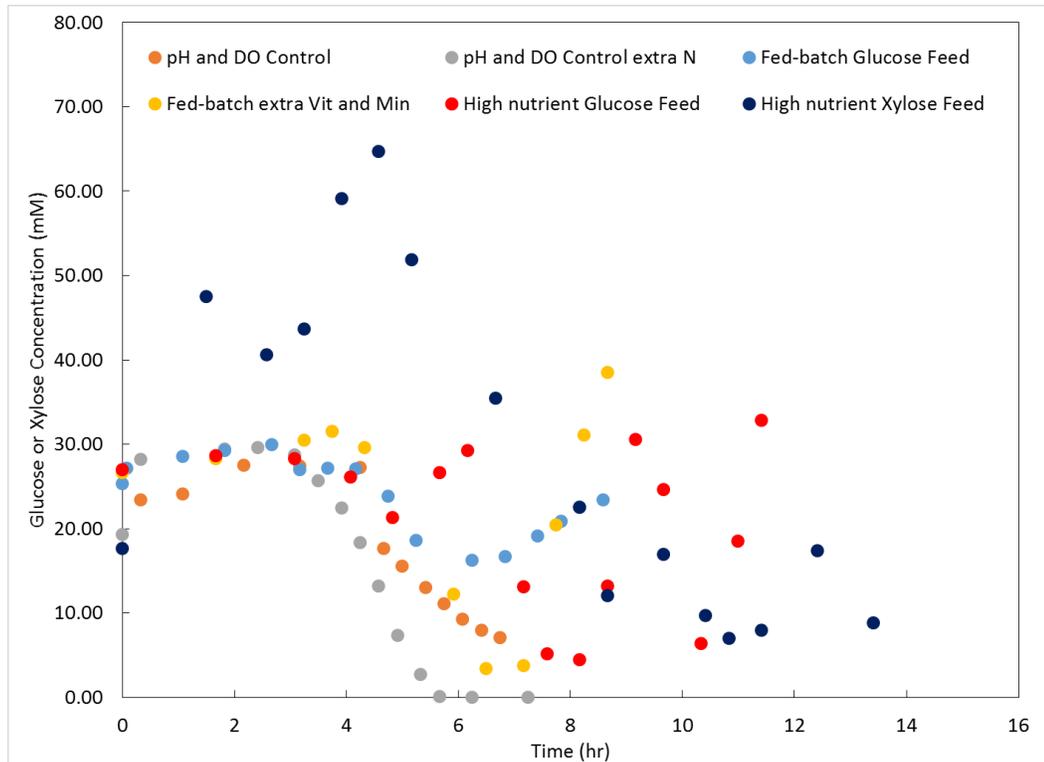


Figure D.2: Glucose or Xylose concentration profile for fed-batch experiments with 2L bioreactor under various controls

To maintain dissolved oxygen, the agitator speed was increased using PI-control. The figure below show the agitator response to low dissolved oxygen (DO) for each experiment.

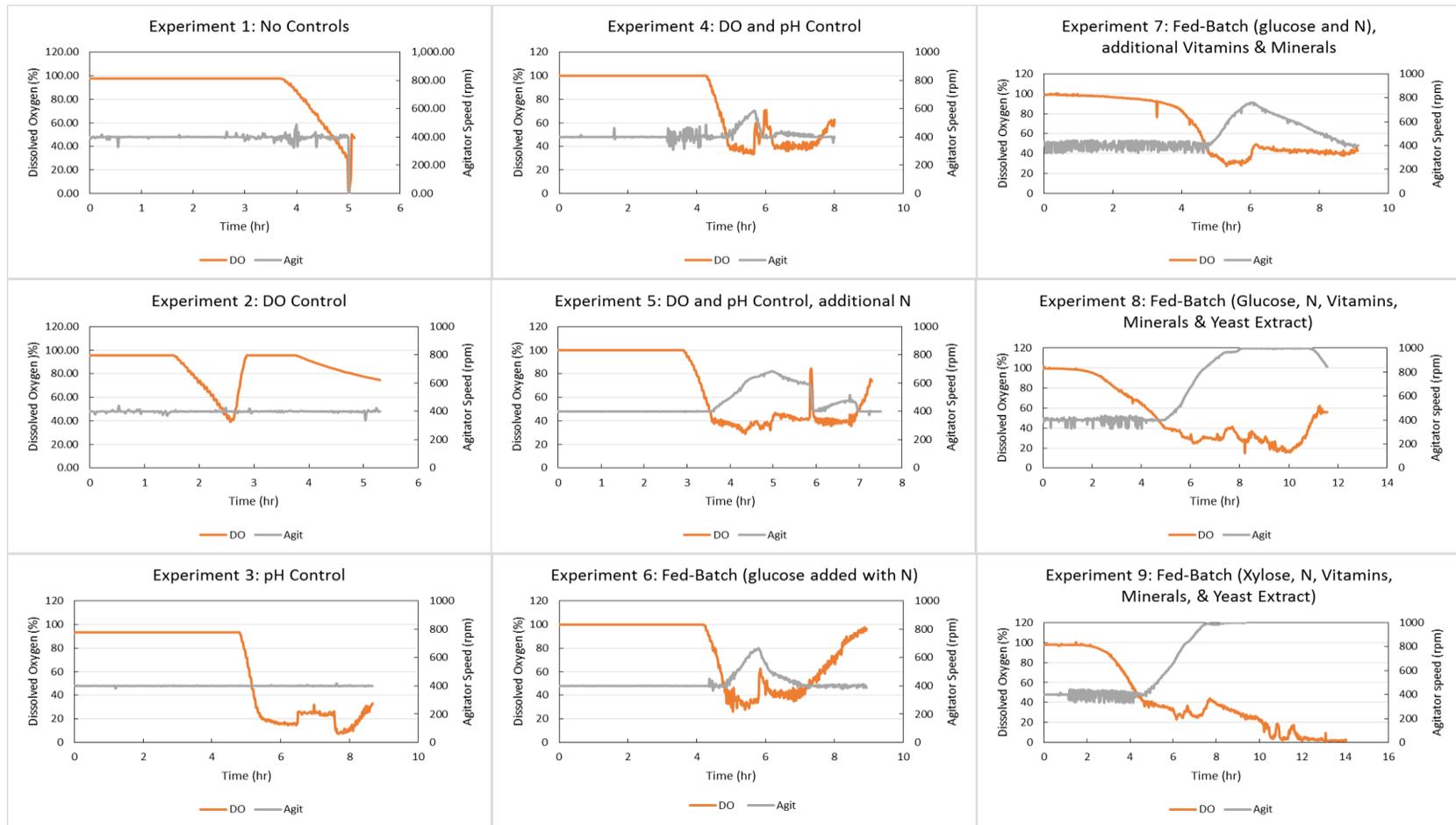


Figure E. 3: Dissolved oxygen and agitator speed profiles for 2L bioreactor experiments

pH was controlled for the fed-batch experiments. The figure below shows the profile of measured pH. The controller works efficiently to prevent significant over or under shooting the target value.

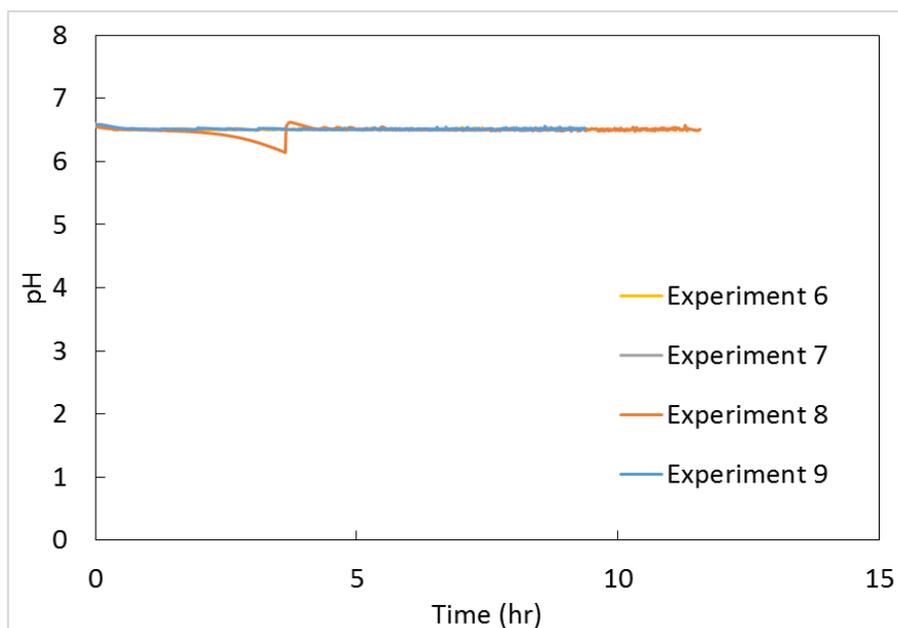


Figure D. 4: Measured pH for fed-batch experiments under various controls (See table 2.9 for experimental set-up).

Table D.1: Complete data set from Experiment 8: High density growth on glucose with nitrogen, vitamins, minerals, and yeast extract in feed solution.

Time (hr)	Average OD	Adjusted OD	Glucose Concentration (mM)	Acetate Concentration (mM)	Lactate Concentration (mM)
0	0	0	26.94	0.86	0.030
1.67	-0.029	-0.005	28.63	0.44	0.093
3.08	0.495	0.550	28.29	2.11	0.067
4.08	0.826	0.928	26.10	4.68	0.061
4.83	1.287	1.505	21.28	5.67	0.074

5.67	2.105	2.745	26.58	6.85	0.098
6.17	2.829	4.405	29.20	16.93	0.060
7.17	2.412	13.350	13.06	52.92	0.094
7.58	2.592	14.982	5.10	53.90	0.091
8.17	2.736	16.496	4.43	74.55	0.230
8.67	2.856	17.985	13.16	110.27	0.050
9.17	2.583	18.614	30.53	73.54	0.106
9.67	2.760	20.963	24.63	106.68	0.174
10.33	2.834	22.108	6.31	162.99	0.297
11.00	2.777	21.215	18.47	101.16	0.224
11.42	2.633	19.230	32.77	106.43	0.474

Table D.2: Complete data set from Experiment 8: High density growth on xylose with nitrogen, vitamins, minerals, and yeast extract in feed solution.

Time (hr)	Average OD	Adjusted OD	Xylose Concentration (mM)	Acetate Concentration (mM)	Lactate Concentration (mM)
0	0	0	17.57	0.56	0.047
1.5	0.282	0.293	47.47	1.46	0.026
2.6	0.416	0.438	40.57	0.53	0.058
3.3	0.739	0.800	43.66	0.99	0.067
3.9	1.123	1.263	59.12	0.91	0.143
4.6	1.608	1.917	64.63	1.40	0.189
5.2	2.121	2.732	51.86	1.61	0.206
6.7	2.883	9.020	35.42	7.80	0.327
8.2	2.715	16.008	22.53	20.36	0.144
8.7	2.862	17.752	12.01	26.75	0.090
9.7	2.802	21.245	16.92	21.39	0.157
10.4	2.346	25.237	9.67	34.58	0.300
10.8	2.330	24.972	7.00	26.11	0.221
11.4	2.270	24.040	7.88	32.49	0.453
12.4	2.339	25.126	17.30	22.44	0.473
13.4	2.354	25.355	8.81	27.94	0.418