AN ABSTRACT OF THE THESIS OF

Tyler M. Kirkendall for the degree of Master of Science in Environmental Engineering presented on August, 24, 2015.

Title: Genome-Scale Metabolic Reconstruction and Analysis of a Pure Culture Anaerobic Digester.

Abstract approved:

_____

Tyler S. Radniecki

Anaerobic digestion is a biological process in which organic matter is decomposed by a community of microbes in the absence of oxygen. The end product of anaerobic digestion is biogas, composed of methane and carbon dioxide, which is often recovered and used to generate energy. Commonly, biogas is not produced in sufficient quantities suitable for economic recovery, resulting in the flaring of biogas to the atmosphere. To optimize anaerobic digestion performance, *in silico* models have been created. Current models examine the critical components of anaerobic digestion, approaching the system at a macroscopic level. To address the limits of current macroscopic *in silico* models, a genome-scale model (GEM) of a pure culture anaerobic digester was constructed to evaluate both the individual organism's metabolic activity and the community level fitness. An *in silico* pure culture anaerobic digester GEM was defined to include the acidogenic bacteria *Clostridium acetobutylicum* (iCac802), the syntrophic short chain fatty acid oxidizer *Syntrophomonas wolfei* (iTK530), and the methanogenic archaea *Methanosarcina barkeri* (iMG746). While GEMs for both *C. acetobutylicum* and *M. barkeri* have been previously published, a novel GEM for *S. wolfei* (iTK530) was curated. Flux balance analysis was performed on the *S. wolfei* GEM to determine reaction fluxes

through the system for various experimental conditions. The pure culture anaerobic digester was analyzed through the application of OptCom and descriptive-OptCom, which utilized multi-level and multi-objective optimization techniques. The construction of a pure culture anaerobic digester GEM presents a high resolution platform for *in silico* analysis and continued investigation of anaerobic digestion at the genome level.

Genome-Scale Metabolic Reconstruction and Analysis of a Pure Culture Anaerobic Digester

by
Tyler M. Kirkendall

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented August 24, 2015
Commencement June 2016

Master of Science thesis of Tyler Kirkendall presented on August 24, 2015.

APPROVED:

Major Professor, representing Environmental Engineering

Head of the School of Chemical, Biological, and Environmental Engineering

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Tyler M. Kirkendall, Author

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Tyler Radniecki for serving as my advisor and mentor throughout my entire engineering education. From working as an undergraduate in his lab at San Diego State University to working with him as a graduate student at Oregon State University, Dr. Radniecki has continued to foster my academic, personal, and professional growth. I greatly appreciate the unrelenting support, dedication, and enthusiasm Dr. Radniecki has provided me with and for always looking towards the bright side, even when our research isn't going as smoothly as expected.

I would also like to express my sincerest thanks to Dr. Frank Chaplen, who not only served as a spark of inspiration for this project, but who has continued to serve as an advisor and as a source of invaluable expertise on the subject. I would like to thank Dr. Chaplen for providing the patience, guidance, support, and modeling knowledge that directly enabled the success of this project. I would also like to Inara Scott for serving as my Minor Professor and for providing me with the opportunity to connect my engineering knowledge to real world and business situations. The guidance and expertise that Inara provided has played a critical role in my growth as an engineer. I would like to thank Dr. Mark Dolan for serving as a committee member and I would also like to thank Dr. Ingrid Arocho for serving as my Graduate Council Representative.

I would like to express my sincerest gratitude to all the CBEE graduate students, faculty, and graduate students in Owen 406 who have continued to provide support, advice, and a much needed break from research these past two years. Most importantly, I would like to thank my parents, Mike and Pam, and my brother, David, for continuing to provide me the support and stability to allow me to accomplish my goals.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

Genome-Scale Metabolic Reconstruction and Analysis of a Pure Culture Anaerobic Digester

CHAPTER 1
INTRODUCTION

Anaerobic digestion is a natural process in which microorganisms break down organic material in the absence of oxygen, producing biogas, a mixture of methane and carbon dioxide. (Gujer & Zehnder, 1983) Naturally occurring in numerous environments, anaerobic activity has been found in wetlands, benthic deposits, hot springs, and the intestinal tract of various mammals including cattle and humans. (Gerardi, 2003) Anaerobic digestion is commonly used as the terminal step in the wastewater treatment process, due to its ability to stabilize wastewater sludge, reduce sludge volume, and produce a sustainable source of energy. (Chen et al, 2008)

Anaerobic digestion has also gained traction for its ability to reduce the quantity and relative strength of greenhouse gasses, due to the capture and combustion of the methane produced. Methane released to the atmosphere has a 21 times greater global warming potential than carbon dioxide. (IPCC, 2001) Exacerbated by methane's increased global warming potential, methane emissions have been estimated to contribute 20.7% of total anthropogenic greenhouse gas emissions. (US EPA, 2012) The global scientific community has recognized the importance of anaerobic digestion for its capacity to reduce greenhouse gas emissions and produce a renewable energy source, and has sought out methods to optimize digester performance. (Kythreotou, 2014)

Due to the complexity of anaerobic digestion, mathematical models have been developed to increase digester efficiency and understand digester operating conditions. Mathematical models of anaerobic digestion have been evolving since the early 1960's, and have been developed to understand a wide range of conditions though a variety of approaches. (Andrews, 1969; Kythreotou, 2014) The majority of

mathematical models produced for anaerobic digestion utilize reaction kinetics, chemical reactions, or a combination of both kinetics and chemical reactions to estimate digester performance. (Lyberatos & Skiadas, 1999)

Advances in genomic sequencing techniques have allowed for the construction of genome-scale models (GEMs), which provide a high-resolution platform for *in silico* analysis. (Milne et al, 2009) While simple kinetic based solutions may capture the rate of a desired metabolite produced, these models often neglect the remaining metabolism of the microorganism, which provide valuable insight into the function and structure of the microorganism's metabolism. Through the development of GEMs, inter-species interactions and the metabolic network of each species can be modeled and analyzed. (Oberhardt et al, 2009) The application of GEMs to the anaerobic digestion process will allow researchers to better elucidate the fundamental metabolic processes and microbial population dynamics that occur in an anaerobic digester. The construction of an anaerobic digester GEM will assist in the optimization of anaerobic digestion, allowing for enhanced substrate utilization and methane production. Additionally, an anaerobic digester GEM will serve as a discovery platform for the further investigation of a microorganism's metabolic processes. Through the investigation of the metabolism of a microorganism, commercially valuable metabolic byproducts can be uncovered and the critical rate limiting steps of the microorganism's metabolism can be identified and analyzed.

CHAPTER 2
LITERATURE REVIEW

2.1 History of Anaerobic Digestion

The process of anaerobic digestion was first uncovered in the late 1700s, where Van Helmont first observed that decaying organic matter produced a flammable gas. (Abbasi et al, 2011) In 1776, Volta then determined that the quantity of flammable gas was dependent on the quantity of decaying organic material. This flammable gas was later determined to be methane during the early 1800's by John Dalton and Humphrey Davy, who conducted experiments independently from 1804-1808. (Tietjen, 1975)

Anaerobic digestion was first reported to be a microbial mediated process in 1868 by Bechamp. Microorganisms involved in anaerobic digestion were first isolated by Omelianski in the early 1890s. Omelianski determined that the isolated microbes were responsible for the production of hydrogen, acetic acid, and butyric acid during the anaerobic degradation of cellulose.

In addition to the isolation of microbes that were critical to the anaerobic degradation of organic material, Omelianski proposed that methane was generated in a reaction between hydrogen and carbon dioxide, mediated by a microorganism. Omelianski's hypothesis was later confirmed by Sohngen in 1910. Sohngen also proposed that the fermentation of complex macromolecules occurs through oxidation-reduction reactions, producing hydrogen, carbon dioxide, and acetic acid. Sohngen was also the first to propose that methane is generated from the decarboxylation of acetic acid. (McCarty et al, 1982)

Anaerobic digestion was first applied in the treatment of wastewater in 1881 by Mouras using an "automatic scavenger" design. (Moigno, 1881) This design was later refined and termed a "septic tank" by Cameron in England. The first municipal

use of anaerobic digestion was in 1897 when the local government of Exeter, England used septic tanks to treat the city's wastewater. (McCarty et al, 1982) The septic tanks in Exeter were also one of the first implementations of heating and lighting systems designed to use biogas produced from anaerobic digestion. (Chawla, 1986)

Anaerobic digestion has continued to be used in the treatment of wastewater and the stabilization of putrescible solids. Receiving very little dedicated research in the early 1900s, the fundamental technology of the anaerobic digestion of wastewater sludge has remained largely the same. Reinvigorated by the oil crisis in the early 1970's and developing pollution restrictions, scientific interest returned to anaerobic digestion, as indicated by Figure 2.1. (Abbasi et al, 2011) The main focus of this renewed research was to enhance biogas production from organic substrates. At its current rate, anaerobic digesters typically do not produce enough gas to make capture and purification for energy generation economical. (Kapdi et al, 2005) By increasing the volume of biogas produced, energy or clean natural gas can be produced at an economical rate.

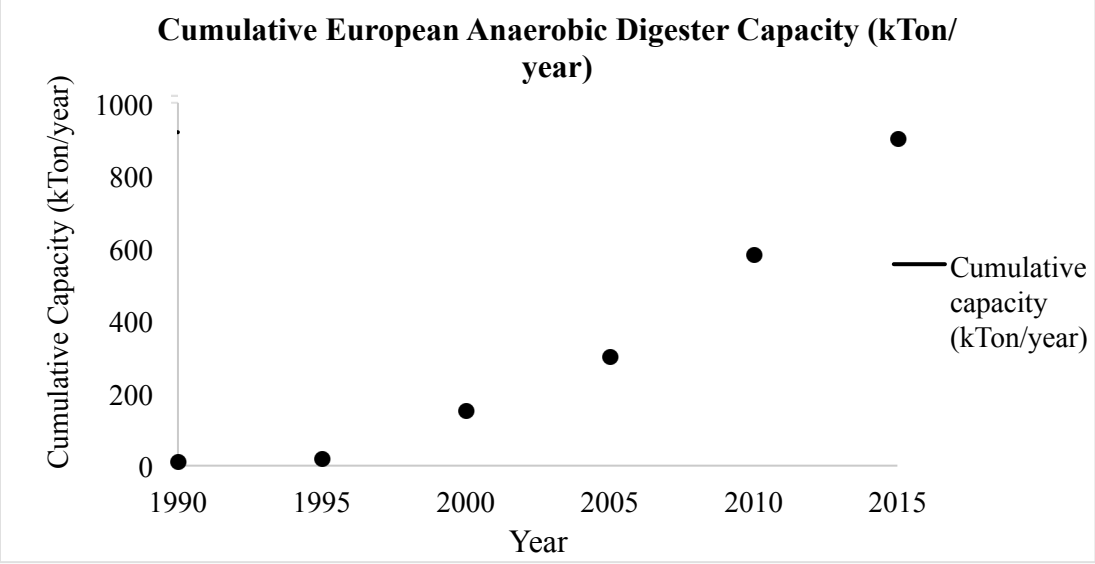

Figure 2.1: Cumulative Anaerobic Digestion Capacity in Europe (adopted from European Bioplastics, 2015)

2.2 Introduction to Microbiology of Anaerobic Digestion

Anaerobic digestion is performed by a diverse community of microorganisms who each play a critical role in this process. Anaerobic microorganisms can be divided into two distinct groups determined by their reaction to free molecular oxygen. Oxygen-tolerant anaerobes will exhibit significant metabolic inhibition, but will survive in an environment with free molecular oxygen. Common oxygen tolerant anaerobes found in an anaerobic digester include acidogenic, which ferment organic monomers to a wide arrange of volatile fatty acids (VFAs), and acetogenic bacteria, which transform VFAs to acetic acid. (Speece, 1983; Gerardi, 2003)

Strict anaerobic microorganisms will die in the presence of free molecular oxygen. Common strict anaerobes include the methanogenic community. Anaerobes are most active when the oxidation-reduction potential (ORP) of the system is between -200 and -400 millivolts (mV). Common anaerobic digestion substrates, including primary sludge and waste activated sludge, feature a low ORP, ranging from -100 to -300 mV. Increasing the ORP has been shown to inhibit anaerobic activity including hydrolysis, acetogenesis, and methanogenesis. (Gerardi, 2003)

2.3 Primary Processes of Anaerobic Digestion

Anaerobic digestion is composed of four primary processes including hydrolysis, acidogenesis, acetogenesis, and methanogenesis. Insoluble organic material and organic macromolecules are hydrolyzed by numerous hydrolytic enzymes, breaking down the organic material into smaller soluble monomers. The small monomers produced are then fermented by a group of acidogenic bacteria, producing a wide range of VFAs. VFAs are then transformed to acetate by a group of acetogenic bacteria. The final process in anaerobic digestion is methanogenesis, where acetate, carbon dioxide, and hydrogen are transformed to methane by a group

of methanogenic Archaea. (Parkin & Owen, 1986) A figure illustrating the anaerobic digestion process can be found below in Figure 2.2. (Rapport et al, 2008)

The degradation of complex macromolecules to methane requires a high level of syntrophic relationships between organisms and relies on the flow of metabolites through the entire system. Due to the sensitive nature of the microorganisms involved, an excess of any of the metabolic products produced during anaerobic digestion will likely cause inhibition of the system. (Gerardi, 2003)

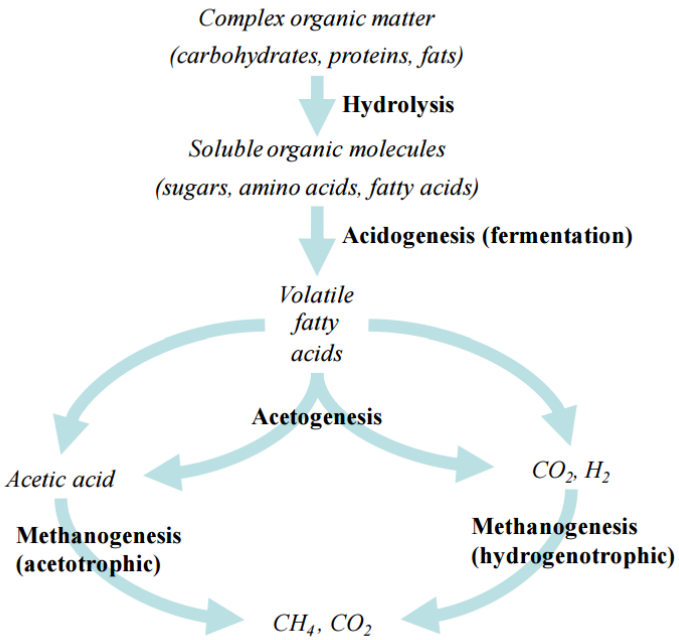

Figure 2.2: Anaerobic digestion process (adopted from Rapport et al, 2008)

2.3.1 Hydrolysis

The first step of anaerobic digestion is hydrolysis, in which complex organic macromolecules are cleaved and the hydroxide and hydrogen ions from water are attached to the separate products at the cleavage point. An example of hydrolysis can be seen in Figure 2.3. During hydrolysis, complex organic polymers are hydrolyzed by extracellular hydrolytic enzymes, which catalyze the degradation of long chain

polymers into simple monomers. The primary constituents that are hydrolyzed include large compounds such as polysaccharides, proteins, and lipids, which are hydrolyzed by extracellular enzymes secreted by a wide range of fermentative bacteria. Extracellular enzymes produced by a single bacterium only degrade a unique or small set of compounds. Examples of the wide range of bacteria required to hydrolyze common substrates and their resulting products can be in Table 2.1. Each species can only degrade a subset of the total variety of substrates, therefore anaerobic digestion requires a diverse community of bacteria for the hydrolysis of complex organic substrates. (Gerardi, 2003)

Table 2.1: Examples of common hydrolytic products and bacteria (adopted from Gerardi, 2003)

Substrate to be Degraded	**Exoenzyme Needed**	**Example**	**Bacterium**	**Product**
Polysaccharides	Saccharolytic	Cellulase	*Cellulomonas*	Simple sugars
Proteins	Proteolytic	Protease	*Bacillus*	Amino acids
Lipids	Lipolytic	Lipase	*Mycobacterium*	Fatty acids

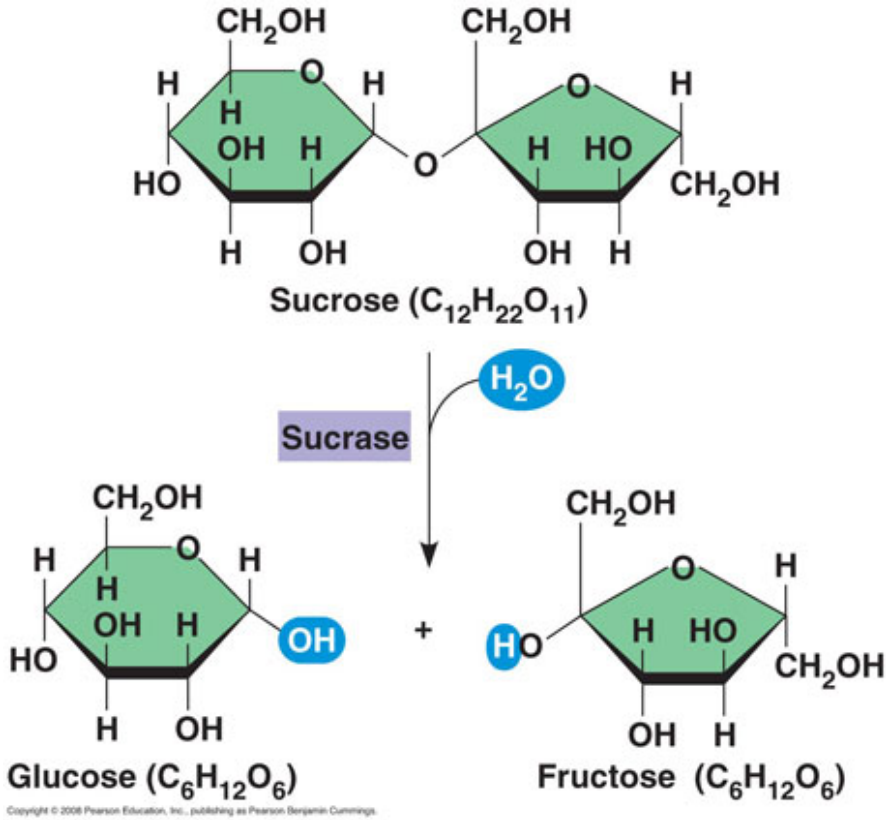

Figure 2.3: Hydrolysis example: Sucrose + Water → Glucose + Fructose (adopted from Reece et al, 2011)

The hydrolysis of complex organic substrates into simple soluble monomers has been identified as one the key rate limiting steps of anaerobic digestion due to the slow kinetics of hydrolytic enzymes and the bottleneck imposed by the lack of suitable substrates for anaerobic digestion. (Tomei et al, 2009) Previous studies have found that slow hydrolytic conversion of biological sludge leads to the slowdown of the entire anaerobic digestion process, significant lag times during startup, and digester failure. (Eastman & Ferguson, 1981; Gujer & Zehnder, 1983; Li & Noike, 1987; Tomei et al, 2009)

2.3.2 Acidogenesis

Following hydrolysis, the soluble monomers that were produced are transformed into VFAs, alcohols, organic acids, organic nitrogen compounds, organic sulfur compounds, carbon dioxide, and hydrogen through a process called acidogenesis. (Gerardi, 2003) Acidogenic bacteria typically produce a large variety of potential products through fermentation. This can be clearly seen by the fermentative bacteria and obligate anaerobe *Clostridium acetobutylicum,* which is capable of producing acetic acid, butyric acid, acetone, butanol, ethanol, hydrogen, and carbon dioxide, as detailed in Figure 2.4.

Acidogenesis is performed by a wide variety of fast growing anaerobic bacteria through an array of fermentative pathways, and is typically the most rapid step in the anaerobic digestion process. Due to the rapid fermentation of a substrate, sudden drops in pH may result from the increased production of acidogenic products. As a result of the increase in the VFAs concentrations, the acetogenic and methanogenic populations may become inhibited. (Beaty & McInerney, 1989) The inhibition of the acetogenic and methanogenic populations creates cyclic affect, as the decrease in the removal of organic acids leads to a continued drop in pH, and ultimately digester failure. (Gerardi, 2003; Henze, 2008)

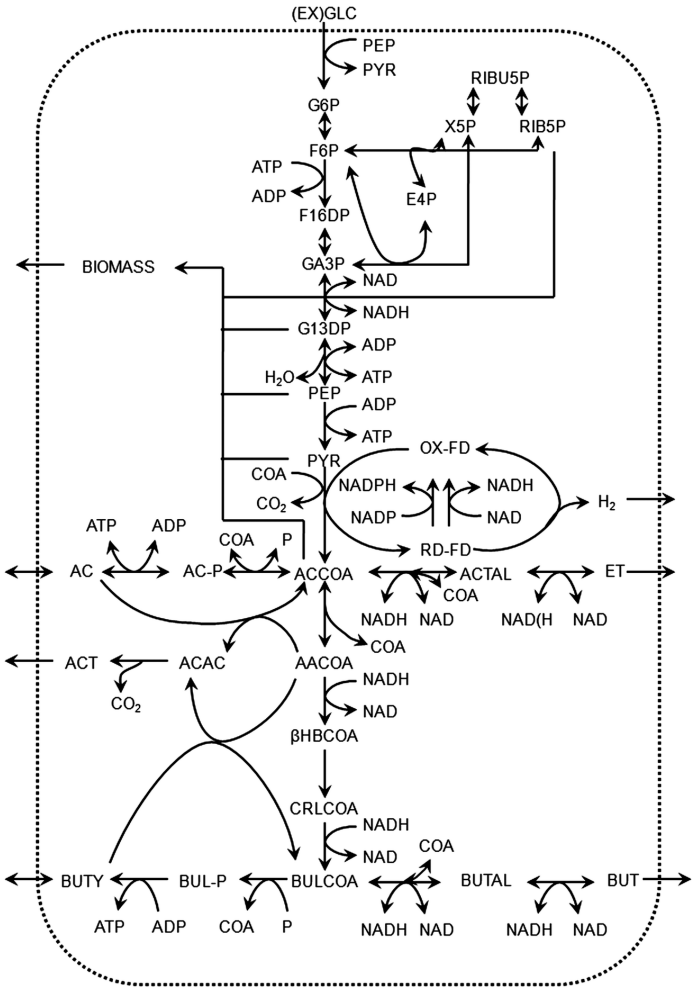

Figure 2.4: Fermentation of glucose by *C. acetobutylicum*

Fermentation products produced by acidogenic bacteria are often dependent on environmental and operating conditions, with system pH cited as a critical condition. (Bahl et al, 1982) The role pH plays on product production can be seen by the fermentative bacteria and obligate anaerobe *Clostridium acetobutylicum.* *C. acetobutylicum* primarily produces acetate and butyrate when the pH is greater than 5.1. Under operating conditions at a pH lower than 5.1, acetate and butyrate are utilized to produce acetone, butanol, and ethanol. These alternative metabolic products have been shown to be toxic to most bacteria, resulting in the further inhibition of the anaerobic digester. (Huang et al, 1986; Dash et al, 2014) Common

acidogenic reactions from sucrose as a substrate are listed in Table 2.2, adopted from Henze, 2008.

Table 2.2: Acidogenic reactions with sucrose as the substrate (adopted from Henze, 2008)

Common Acidogenic Reactions
$C_{12}H_{22}O_{11} + 9H_2O \rightarrow 4CH_3COO^- + 4HCO_3^- + 8H^+ + 8H_2$
$C_{12}H_{22}O_{11} + 5H_2O \rightarrow 2CH_3CH_2CH_2COO^- + 4HCO_2^- + 6H^+ + 4H_2$
$C_{12}H_{22}O_{11} + 3H_2O \rightarrow 2CH_3COO^- + 2CH_3CH_2COO^- + 2HCO_3^- + 6H^+ + 2H_2$

2.3.3 Acetogenesis

While acetate is also produced during acidogenesis, a wide array of VFAs, alcohols, organic acids, organic nitrogen compounds, and organic sulfur compounds are produced as well. The diverse end products produced through acidogenesis are transformed to acetate through a process called acetogenesis. (Gerardi, 2003) Examples of common acetogenic reactions can be seen in Table 2.3. Acetogenesis is performed by a group of obligate anaerobic bacteria known as acetogens. Acetogens can be further separated into two categories including homoacetogens, which includes carbon dioxide reducing acetogens, and syntrophic acetogens. (Drake, 2012) Homoacetogens produce acetate as its sole fermentation product, while carbon dioxide reducing acetogens reduce carbon dioxide and hydrogen to acetate. (Ragsdale & Pierce, 2008) Syntrophic acetogens are capable of producing acetate from a variety of substrates through a syntrophic relationship with a microbial partner. (Beaty & McInerney, 1989)

Table 2.3 Common acetogenic transformation of propionate, glucose, and ethanol

Acetogenic Reactions
$CH_3CH_2COO^- + 3H_2O \leftrightarrow CH_3COO^- + H^+ + HCO_3^- + 3H_2$
$C_6H_{12}O_6 + 2H_2O \leftrightarrow 2CH_3COOH + 2CO_2 + 4H_2$
$CH_3CH_2OH + 2H_2O \leftrightarrow CH_3COO^- + 2H_2 + H^+$

2.3.3.1 Homoacetogens

Homoacetogens produce acetate as its sole fermentation product. Acetogenic microbes area found dispersed throughout many phyla including *Chloroflexi, Firmicutes, and Spirochaetes*, demonstrating that acetogenesis is a metabolic trait, and not a phylogenetic trait. While acetogens are found throughout many phyla, homoacetogens present a rarer subset of acetogens. (Ragsdale & Pierce, 2008) In the homoacetogenic fermentation of glucose to acetate, glucose is first metabolized to two acetates, carbon dioxide, and hydrogen though the Embden-Meyerhof-Parnas pathway, a glycolytic pathway to metabolize glucose to pyruvate coupled with the generation of ATP and NADH. The carbon dioxide produced is then reduced via the Wood-Ljungdahl pathway, a highly elucidated metabolic pathway responsible for the production of acetate from carbon dioxide. (Leang et al, 2013) Homoacetogenic fermentation of glucose to acetate is detailed in Table 2.4 and Figure 2.5. (Diekert & Wohlfarth, 1994)

Table 2.4: Homoacetogenic transformation of glucose and carbon dioxide (adopted from Diekert & Wohlfarth, 1994)

	$C_6H_{12}O_6 + 2H_2O \rightarrow 2CH_3COO^- + 2CO_2 + 8H^+$
	$8H^+ + 2CO_2 \rightarrow CH_3COO^- + H^+ + 2H_2O$
Σ	$C_6H_{12}O_6 \rightarrow 3CH_3COO^-$

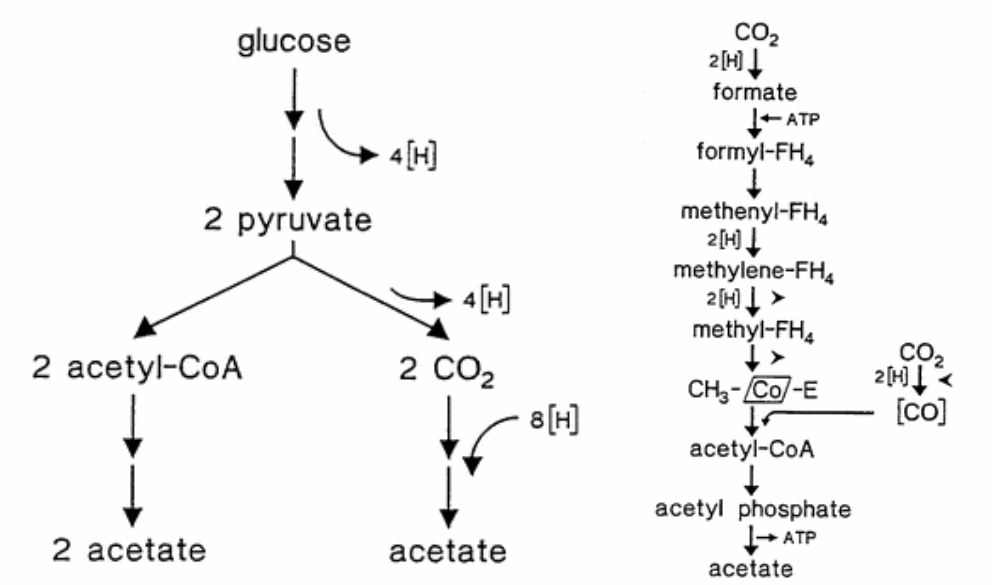

Figure 2.5: Homoacetogenic transformation of glucose and carbon dioxide (adopted from Diekert & Wohlfarth, 1994)

2.3.3.2 Syntrophic Acetogens

Syntrophic acetogens are capable of utilizing various VFAs, such as butyrate, to produce acetate. In the production of acetate, H_2 is produced which must be removed by a syntrophic partner, typically a methanogen or sulfate reducing bacteria. (Sieber et al, 2010; McInerney et al, 1979) The degradation of VFAs to acetate is a thermodynamically unfavorable reaction, and only remains favorable through the constant removal of the H_2 produced.

A model syntrophic acetogen is the gram negative anaerobic bacterium *Syntrophomonas wolfei,* which beta-oxidizes butyrate to acetate through a syntrophic relationship with a methanogenic partner. (Sieber et al, 2010) As detailed in Table 2.5, the degradation of butyrate to acetate has a Gibbs free energy change of +48.6 KJ/mol, which is thermodynamically impractical. (Thauer et al, 1977) When the H_2 partial pressure is kept at 1 Pa by a syntrophic partner, the Gibbs free energy change for the degradation of butyrate to acetate becomes -39.2 KJ/mol. As evident

by the low available energy, syntrophic microorganisms typically feature slow growth rates and low growth yields. (Schöcke & Schink, 1997; Scholten & Conrad, 2000; Jackson & McInerney, 2002; Sieber et al, 2010) In a stable anaerobic digester, hydrogenotrophic methanogens utilize free H_2 rapidly so that the H_2 partial pressure remains below 10^{-4} atm. This zone of allowable H_2 partial pressures that maintains thermodynamic favorability can be seen in Figure 2.6. (Henze, 2008)

Table 2.5: Thermodynamics of the degradation of butyrate to acetate, (adopted from Thauer et al, 1977; Sieber et al, 2010)

Acetogenic Reaction	**Without Hydrogen Removal ΔG°' (KJ/mol)**	**With Hydrogen Removal ΔG°' (KJ/mol)**
$Butyrate^- + 2H_2O \dashrightarrow 2\ Acetate^- + H^+ + 2H_2$	+48.6	-39.2

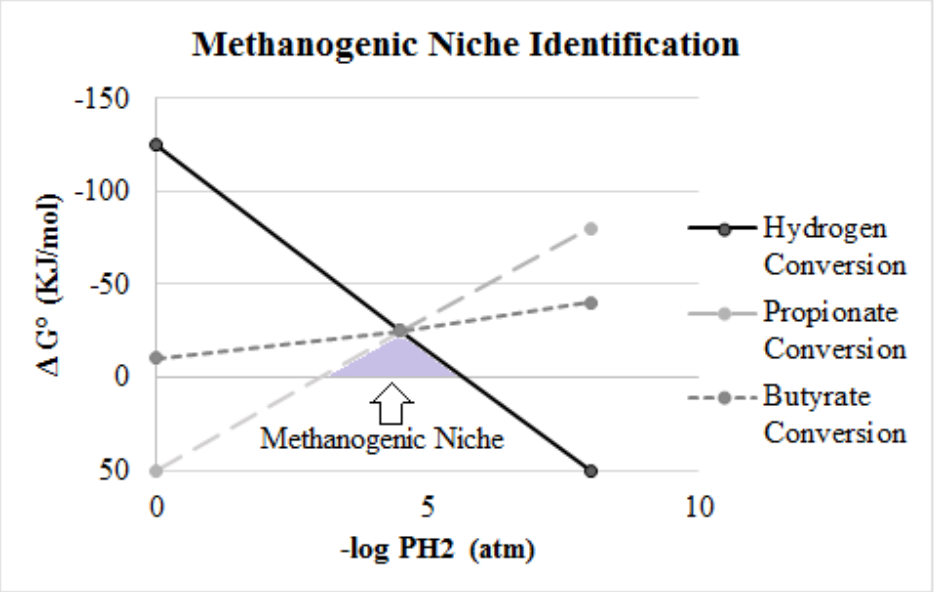

Figure 2.6: Free energy change versus H_2 partial pressure (adopted from Henze, 2008)

2.3.4 Methanogenesis

The terminal step in anaerobic digestion is the conversion of single carbon compounds and hydrogen, produced during acidogenesis and acetogenesis, to methane. Methanogenesis is performed by a unique group of highly specialized obligate anaerobic Archaea called methanogens. Methanogens are a morphologically diverse and have diverse growth patterns, shapes, and sizes. (Vogels et al, 1988; Boone et al, 1993a) Methanogen cells are typically 0.1–15 um in diameter, with filaments ranging up to 200 um in length. (Gerardi, 2003) Methanogens can be divided into two main categories, acetoclastic or acetotrophic methanogens and hydrogenotrophic methanogens. Acetoclastic methanogens are responsible for converting acetate to methane and carbon dioxide while hydrogenotrophic methanogens are responsible for converting hydrogen and carbon dioxide to methane.

Some species of methanogens are capable of utilizing a wide variety of single carbon compounds in addition to the acetotrophic and hydrogenotrophic pathways used to produce methane. One such model methanogen, *Methanosarcina barkeri*, is capable of fermenting hydrogen, acetate, methanol, and methylamines to methane. (Balch et al, 1979) It is estimated that two-thirds of the methane produced in nature is derived from the methyl group of acetate, while the remaining third is derived from the reduction of carbon dioxide with electrons derived from the oxidation of hydrogen or formate. (Demirel & Scherer, 2008) Both classes of methanogens are critical in anaerobic digestion, as accumulation of organic acids and hydrogen have been shown to inhibit both acetogenesis and methanogenesis. (Chen et al, 2008)

2.3.4.1 Acetoclastic Methanogens

Acetoclastic, or acetotrophic, methanogens are a group of methanogens responsible for converting acetate to methane. (Ferry, 1992) The fermentation of acetate by acetoclastic methanogens has only recently been recently elucidated, and

proceeds by reducing the cleaved methyl group of acetate to methane with electrons derived from the oxidation of the carbonyl group of acetate to carbon dioxide. (Vogels et al, 1988; Ferry, 1992) The net reaction for this process can be seen in Equation 1, with the metabolic process detailed in Figure 2.7.

Equation 1: Fermentation of acetate to methane

$$CH_3COO^- + H^+ \rightarrow CH_4 + CO_2$$

($\Delta G^{\circ\prime} = -36$ kJ/mol CH_4)

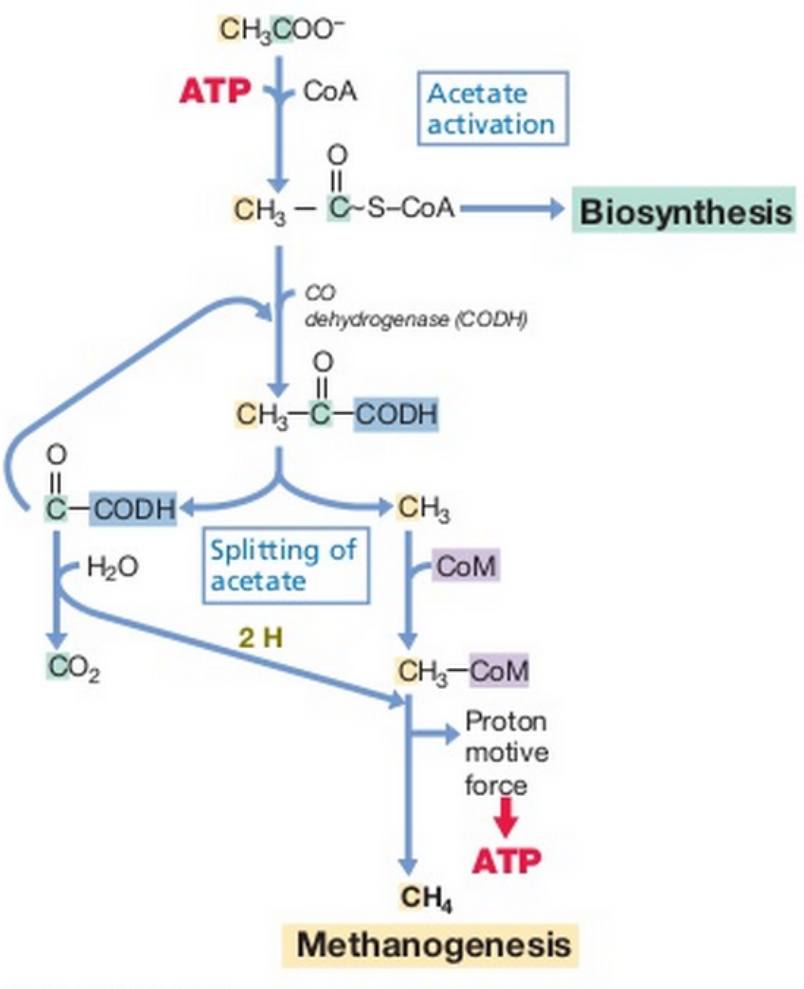

Figure 2.7: Methane production from acetotrophic methanogenesis (adopted from Madigan et al, 2010)

Reactor design, operating parameters, and substrate type play a critical role in the quantity and species of acetoclastic methanogens found in an anaerobic digester. (Yu et al, 2005) Summaries of studies on digester startup and operation have found that *Methanosaeta* spp. was the dominant acetoclastic methanogen in anaerobic digesters with low acetate loadings. (Griffin et al, 2000; Zheng & Raskin, 2000) In digesters with elevated acetate concentrations or unstable conditions, *Methanosarcina* spp. was found to be the dominant acetoclastic methanogen. (Demirel & Scherer, 2008) Elevated acetate concentrations and unstable digester conditions commonly occur in the anaerobic co-digestion of high strength organic wastes. Subsequently, *Methanosarcina* spp. was found to be the dominant acetoclastic methanogen in co-digestion reactors, which co-digest municipal solid wastes along with other high strength wastes such as fats, oils, and greases (FOG). (Stroot et al, 2001; McMahon et al, 2001; Demirel & Scherer, 2008) Acetoclastic methanogens play a critical role in digester activity, as high loading of acetate has been shown to inhibit both acetogens and methanogens (Demirel & Scherer, 2008)

2.3.4.2 Hydrogenotrophic Methanogens

Hydrogenotrophic methanogens play a critical role in digester stability, and are responsible for controlling the hydrogen partial pressure in an anaerobic digester. Through the conversion of hydrogen and carbon dioxide to methane, hydrogenotrophic methanogens maintain a low hydrogen partial pressure, which allows the fermentative pathways that occur in acidogenesis and acetogenesis to remain energetically favorable. (Hedderich & Whitman, 2006; Sarmiento et al, 2011)

Hydrogenotrophic methanogens utilize carbon dioxide as a carbon source and hydrogen as the reducing agent to produce methane. In the reduction of carbon dioxide to methane, carbon dioxide is first bound to the coenzyme methanofuran (MFR) and is then reduced to the formyl level. The formyl group of formylmethanofuran is then transferred to the coenzyme H4MPT and subsequently

dehydrated to methenyl-H4MPT. The carbon group of methenyl-H4MPT is then reduced down to methyl-H4MPT where the methyl group is transferred to coenzyme M (CoM). Ultimately, The methyl-CoM is then reduced to methane using coenzyme B (CoB) as the final electron donor. (Liu & Whitman, 2008; Sarmiento et al, 2011) The net reaction of this process can be seen in Equation 2, with the complete metabolic process detailed Figure 2.8, adopted from Goldman et al, 2009.

Equation 2: Reduction of carbon dioxide to methane

$$H_2 + CO_2 \rightarrow CH_4 + H_2O$$

($\Delta G^{\circ\prime} = -130.4$ kJ/mol CH_4)

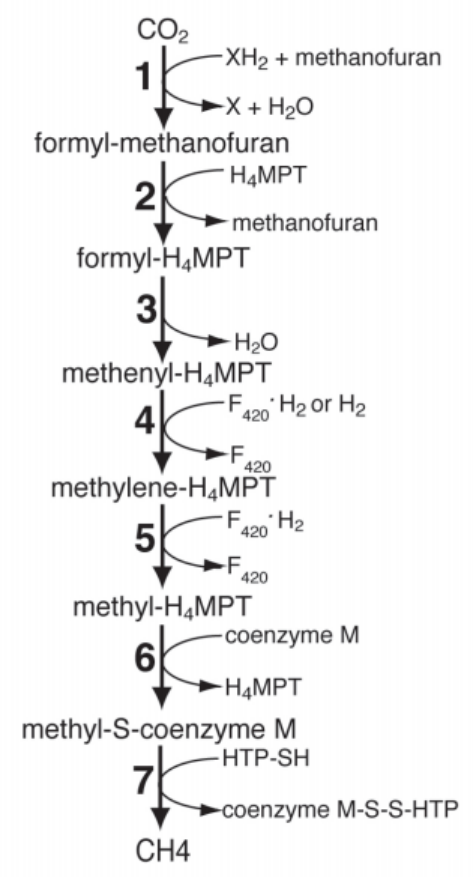

Figure 2.8: Hydrogenotrophic methanogenesis pathway (adopted from Goldman et al, 2009)

Digester operating conditions significantly impact the diversity and activity of the methanogenic community. In a study on thermophilic anaerobic digesters, hydrogenotrophic methanogens were found to be greatly outnumbered by the quantity of acetotrophic methanogens. (Scherer et al, 2000) While hydrogenotrophic methanogens were outnumbered by acetotrophic methanogens, their populations remained steady while the acetotrophic methanogenic population significantly decreased as temperature was increased from 55 to 65 °C. (Ahring et al, 2001)

The methanogenic community of a thermophilic anaerobic digester was monitored from digester start up to steady state by Montero et al, 2008. Hydrogenotrophic methanogens were found to be the dominant methanogen during digester start up, but were quickly displaced by acetotrophic methanogens as the digester reached steady state. The dominance of the hydrogenotrophic methanogens during digester start up may be attributed to the volatile nature of the digester and the extremophilic tolerance of hydrogenotrophic methanogens. Hydrogenotrophic methanogenic activity have been found in environments with a pH as low as 3.8 and as high as 9.9, and are speculated to dominate in acidic environments due to the presence of free acetic acid, which can freely pass through the cell membrane and potentially disrupt the proton motive force. (Kotsyurbenko et al, 2007)

2.4 Anaerobic Digestion Microbial Community

Anaerobic digestion is a sensitive and complex process that requires a rich community of microorganisms. Anaerobic digesters are microbially dense, with greater than 10^{16} cells per milliliter of digestate. (Gerardi, 2003) Feedstock source and practically every operating condition such as temperature, sludge retention time, pH, and concentration of various nutrients and metals have been shown to influence the population of microorganisms found in anaerobic digesters. (Demirel & Scherer, 2008; Ali Shah, 2014) In addition to the variation of microbial populations between digesters, specific phylotypes of common bacteria can be found native to each

digester. (Riviere et al, 2009) To better understand the role each organism plays in this diverse process, research has been conducted to characterize the microbial community of an active anaerobic digester.

There are numerous techniques to identify the species distribution found in an anaerobic digester sample. The most common molecular techniques include fluorescent in-situ hybridization (FISH), microautoradiography-fluorescence in situ hybridization (MAR-FISH), 16S rRNA analysis, and applying polymerase chain reaction along with denaturing gradient gel electrophoresis and pyrosequencing. (Thayanukul et al, 2010; Liu et al, 2010; Sanapareddy et al, 2009; McLellan et al, 2010; Reyes et al, 2015)

Utilizing FISH, the population of a mesophilic anaerobic digester in Valencia, Spain was analyzed. Of the phenotypes sampled, methanogens comprised 30 percent of all bacteria in the digester. The dominant order of methanogens was found to be *Methanosarcinales* (14% of all bacteria), followed by an even distribution of *Methanomicrobiales* and *Methanobacterales* (8% of all bacteria). Sulfate reducing bacteria were found to account for 20 percent of the digester sample while 10 percent were identified as denitrifying bacteria. The remaining 40 percent of the digester were uncharacterized, and assumed to include the acidogenic and acetogenic bacteria found in anaerobic digesters. (Reyes et al, 2015)

A core group of bacteria found in anaerobic digesters were recently defined by Rivere et al, 2009. The study sampled seven anaerobic digesters from France, Chile, and Germany using PCR and 16S rRNA analysis. 16S rRNA sequences were then assigned operational taxonomic units (OTUs). The core group included OTUs affiliated with Betaproteobacteria*, Chloroflexi, Synergistetes, and Bacteroidetes*. OTUs affiliated with these four phylums were found in all seven of the anaerobic digesters sampled. A table of the percentages of *Bacteria* and *Archaea* found in all seven anaerobic digesters is broken down in Table 2.6. (Rivere et al, 2009)

Table 2.6: Percent of phylogenetic groups from digesters sequenced (adopted from Rivere et al, 2009)

Phylogenetic Group	**Average proportion of each group in % ± standard deviation**	***(Min - Max)***
***Bacteria* domain**		
Chloroflexi	32 ± 9	(15 - 45)
Proteobacteria	18 ± 5	(11 - 24)
Bacteroidetes	11 ± 6	(3 - 25)
Firmicutes	9 ± 6	(5 - 25)
Aminanaerobiaa	2 ± 2	(ND - 6)
WWE1	2 ± 4	(ND - 12)
Actinobacteria	2 ± 1	(0.4 - 2)
Synergistetes	4 ± 4	(0.4 - 12)
Coprothermobacteria	1 ± 2	(ND - 7)
Spirochaete	1 ± 1	(ND - 2)
Minor groups	6 ± 3	(2 - 9)
Unclassified	12 ± 4	(6 - 19)
***Archaea* domain**		
Euryarchaeota		
Methanosarcinales	51 ± 30	(12 - 93)
ArcI	36 ± 34	(ND - 77)
Methanomicrobiales	10 ± 2	(2 - 35)
Methanobacteriales	0.2 ± 0.3	(ND - 1)
Crenarchaeota	2 ± 3	(ND - 7)

ND: Not detectable

2.5 Review of Anaerobic Digestion Modeling

Anaerobic digestion is a complex multi-step biological process involving a community of microorganisms. Due to the complexity of the anaerobic digestion system, mathematical models were developed to optimize the process. Through the application of an anaerobic digestion model, design parameters and operating conditions can be efficiently evaluated.

Numerous models have been developed that utilize various parameters and characteristics of the system as the backbone of the model. Mathematical models have been created that utilize various control parameters including substrate inhibited Monod kinetics of methanogens, volatile fatty acid concentration, H_2, pH decrease, and increase of NH_3 concentration in the influent sludge. (Andrews, 1969; Smith et al, 1988; Bryers, 1985; Costello et al, 1991a; Costello et al, 1991b; Siegriest et al, 1993) The wide variety of control parameters modeled is a function of the complexity of the anaerobic digestion system. Early mathematical models were constructed to model the rate-limiting step of anaerobic digestion. These rate-limiting steps are subject to change due to the variable operation conditions, resulting in the wide variety of mathematical models currently available. (Lyberatos & Skiadas, 1999)

The most comprehensive anaerobic digestion model to date, ADM1, was produced by the Anaerobic Digestion Modeling Task Group established by the International Water Association. (Batstone et al, 2002) ADM1 was designed to be as general as possible, and is composed of a large set of reactions and reaction kinetics. ADM1 models every step of the anaerobic digestion process including hydrolysis, acidogenesis, acetogenesis, and methanogenesis. Model outputs by ADM1 include standard output variables such as biogas flow rate, pH, and concentration of volatile organic acids and ammonia. The model is also able to account for various causes of inhibition, including inhibition from pH or decreased biomass growth from nitrogen limitations. (Batstone et al, 2002) A comprehensive overview of variables included in ADM1 can be seen in Figure 2.9.

ADM1 takes a kinetics based approach to modeling anaerobic digestion. All extracellular and hydrolytic reactions are assumed to follow empirically derived first order rate law kinetics. All reactions occurring within the cell follow Monod-type substrate uptake kinetics, and all substrate uptake rates are assumed to remain proportional to the biomass growth rate and biomass concentration. (Batstone et al, 2002; Kythreotou, 2014) While ADM1 has been validated for numerous conditions, it is not well suited for the examination of species dynamics, optimization studies, process control, or detailed studies of the active metabolic pathways of each microorganism, as many of these parameters were not included in ADM1. (Kythreotou, 2014)

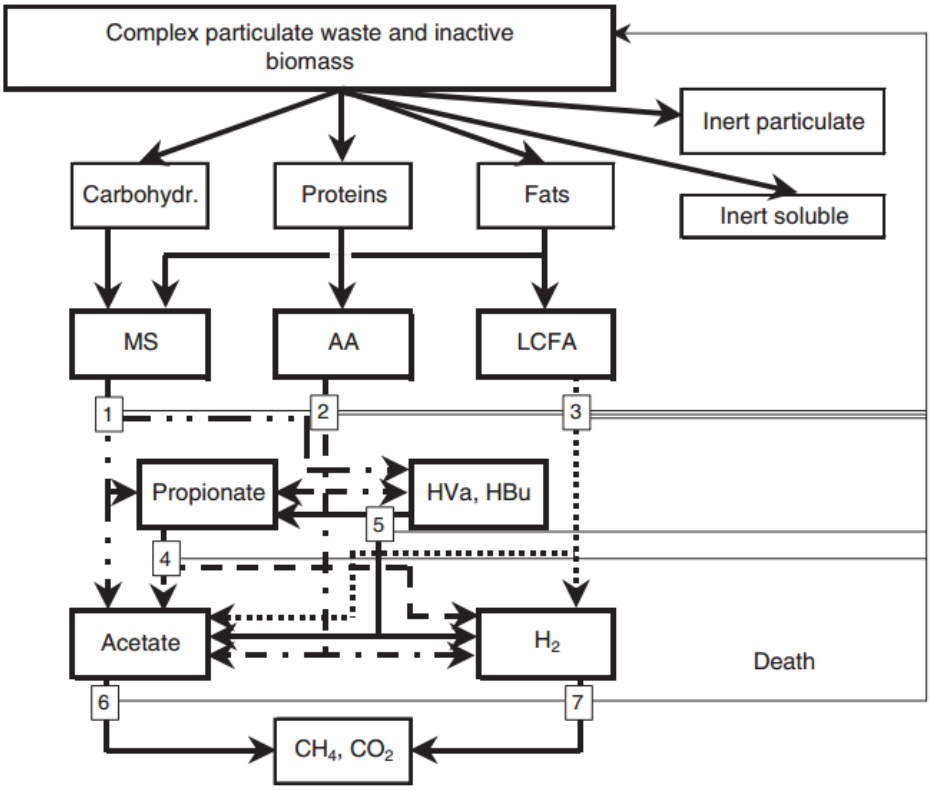

Figure 2.9: ADM1 processes (adopted by Batstone et al, 2002)

2.6 Genome-Scale Modeling

As the number of sequenced genomes and development of high-throughput biological technologies has rapidly increased, an *in silico* platform was needed to examine and visualize the vast array of accumulating biological data. Comprised of a combination of genomic, transcriptomic, proteomic, metabolomics and fluxomic data ("omics" data), GEMS provide a powerful tool to connect "omics" data with microbial physiology. (Milne et al, 2009; Edwards & Palsson, 2000) GEMS are generally comprised of "(a) a list of reactions that occur through each metabolic pathway of an organism, including reaction stoichiometry and reversibility, (b) a group of Gene-Protein-Reaction (GPR) associations that correlate gene activity with metabolic reactions, and (c) a biomass composition reaction that specifies which compounds are necessary for microbial growth." (Devoid et al, 2013)

GEMS have been applied in various ways, most notably for "(a) contextualization of high-throughput biological data, (b) guidance of metabolic engineering, (c) direction hypothesis-driven discovery, (d) interrogation of multi-species relationships, and (e) metabolic network discovery." (Oberhardt et al, 2009) First developed in 1999 for *Haemophilus influenza*, there are now more than 100 GEMS that have been experimentally validated. (Schilling & Palsson, 2000; UCSD, 2015) Although the availability of sequenced genomes are growing at an exponential pace, the growth of GEMs has continued to lag behind by an order of magnitude, as seen in Figure 2.10. (Milne et al, 2009) The slow increase in GEM production is due to the time intensive nature of manually curating and validating the metabolic network reconstruction.

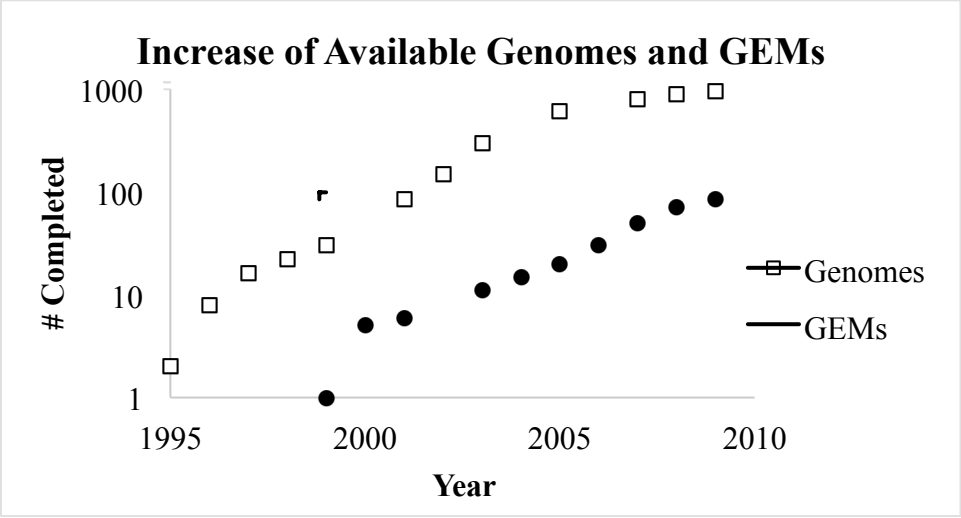

Figure 2.10: Available genome sequences and GEMS (adopted from Milne et al, 2009)

2.7 Genome-Scale Model Construction

GEMS are constructed by utilizing a variety of techniques and depending on the biological and experimental data available, can exist in four distinct phases including (1) draft reconstruction, (2) curated reconstruction, (3) genome-scale metabolic model, and (4) a platform for design and discovery. (Feist et al, 2008) The first, and lowest quality stage is a genome-scale draft reconstruction. A draft reconstruction is produced through the annotation of an organism's genome sequence, where biological functions are attached to genomic elements based on similarities to known genomic elements.

As GEMs have become an active area of research, numerous automated tools, such as The SEED, have been developed. (Overbeek et al, 2005) The SEED employs a process that creates GEM draft reconstructions in conjunction with biological databases such as KEGG, an organism specific genomic and biological pathway database, through the input of an organism's genome sequence. (Kanehisa et al, 2014;

Kanehisa & Goto, 2000) A thorough procedure for constructing an automated GEM by Devoid et al has been summarized below.

2.7.1 Automated GEM Draft Reconstruction

Automated GEM draft reconstruction produced using the SEED network can be broken down into an abbreviated seven step process including (1) submitting a genome sequence to Rapid Annotation using Subsystem Technology (RAST), (2) annotation of the genome, (3) review and curation of the annotation, (4) submitting a RAST annotation to Model SEED, (5) reconstruction of a core metabolic model, (6) generation of a draft biomass composition reaction, (7) review and curation of the metabolic model. (Devoid et al, 2013) A flow chart of the GEM reconstruction process can be seen in Figure 2.11.

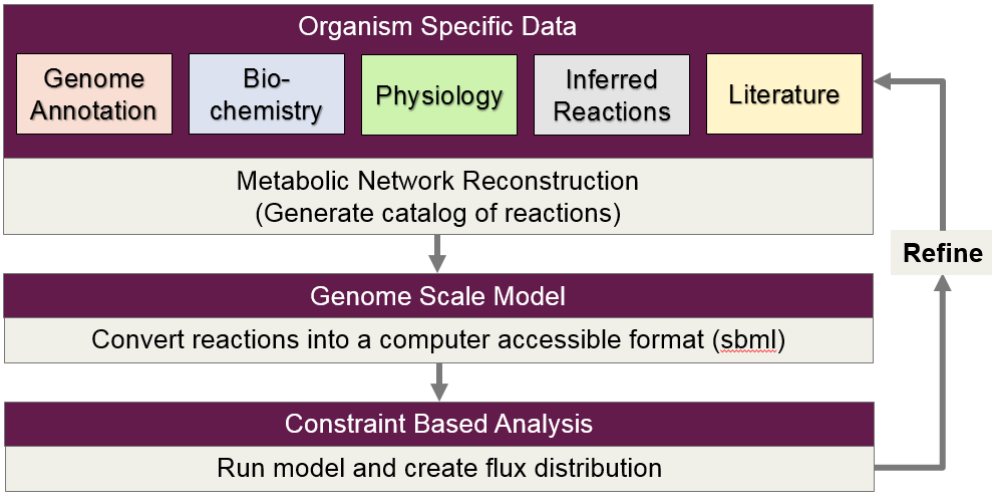

Figure 2.11: Flow chart of GEM reconstruction process

2.7.2 Submitting a Genome to RAST for Genome Annotation

After registering for a SEED account, a genome sequence can be uploaded to the RAST server through SEED. (Overbeek et al, 2005; Aziz et al, 2008) Genomes

are submitted to SEED in a FASTA format, a text-based format used to represent genome sequences, for annotation. Upon submittal to the RAST server for annotation, RAST first completes a targeted search for rRNAs, tRNAs, and genes associated with the synthesis and use of selenocysteine and pyrrolysine. (Devoid et al, 2013)

Following the search for these explicit elements, protein-encoding genes (PEGs) are identified through an iterative process that identifies well known genes and then works to fill gaps and large overlaps based on the knowledge of common genes. After the identification of the PEGs, functions are assigned based on the similarity to well annotated reference genomes. Following the first pass of assigning functions to easily identifiable PEGs, functions are assigned to unknown PEGs utilizing the BLAST program. (Altschul et al, 1990) BLAST identifies these unknown PEGs by comparing their sequences to the DNA sequences of existing PEGs. Functions are then assigned to the unknown PEGs based on these DNA sequence similarities. (Devoid et al, 2013)

2.7.3 RAST Annotation Review

Prior to submitting a RAST annotation for model construction in the SEED environment, it is recommended to review the completed automated RAST annotation. Genome sequences submitted to RAST for annotation are often phylogenetically close to previously annotated genomes, providing an efficient reference point for submitted genomes. While the automated RAST annotation will provide a functional annotation, users are encouraged to manually check for errors.

Annotations can be improved through comparison with curated reference genomes, such as the *E. coli* GEM. Additionally, the sequence quality many induce errors through frameshifts or truncations. (Figures 2.12 and 2.13) Frame shifts can be caused by the addition or deletion of nucleotides, shifting the ORF, which results in an incorrect translation. (Devoid et al, 2013)

After manually reviewing the RAST annotation, the annotated genome can be submitted to the Model SEED for the automated reconstruction of a draft GEM. Upon submittal to Model SEED, a complete automated draft GEM reconstruction will be completed in about 24 hours. (Devoid et al, 2013)

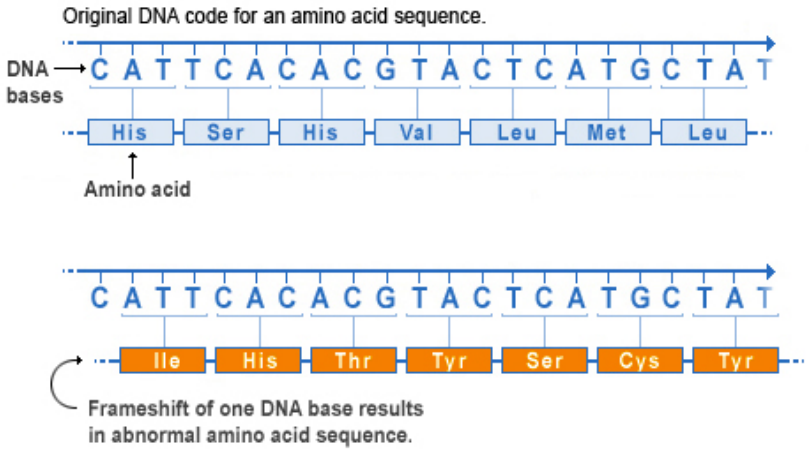

Figure 2.12: Example of a frame shift (adopted from National Library of Medicine, 2015)

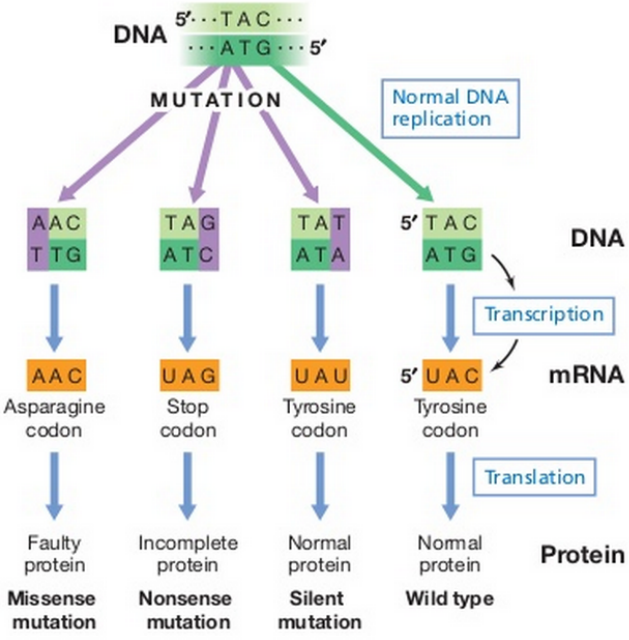

Figure 2.13: Example of a truncation (adopted from Madigan et al, 2010)

2.7.4 Converting RAST Annotation to Draft GEM

The Model SEED calls upon a catalog of over 13,000 metabolic reactions and over 16,000 reactants to complete an automated draft GEM reconstruction. This catalog of biochemical data, populated by KEGG and published GEMs, is then called upon to generate a non-redundant set of reactions. (Kanehisa et al, 2014; Kanehisa & Goto, 2000) Model SEED translates a RAST annotated genome to a draft GEM by constructing a framework that links GPRs with reactions through protein complexes, or groups of proteins with associated polypeptide chains. The Model SEED community continuously updates the mapping framework used in draft GEM reconstruction as new information becomes available. In addition to reactions dictated by the genome, spontaneous reactions that occur in a natural environment are also added to every model constructed in Model SEED. (Devoid et al, 2013)

2.7.5 Developing GEM Draft Biomass Composition Reaction

Following the reconstruction of the metabolic pathways in Model SEED, a draft biomass composition reaction is developed through Model SEED. The biomass composition reaction details the quantity of all metabolites that must be produced to generate one gram of biomass. Biomass composition reactions are composed of DNA, RNA, proteins, lipids, cell walls, and various cofactors. An exact biomass composition reaction cannot be determined from the RAST annotation alone, as the quantity of each metabolite cannot be determined from the genome sequence. Due to the inability to predict an exact biomass composition reaction, four individual biomass compositions have been developed for gram negative bacteria, gram positive bacteria, *Mycoplasma* bacteria, and *Archaea*. A general biomass composition reaction is assigned through interrogation of the RAST annotation or through user assignment. (Devoid et al, 2013)

To assign the general biomass composition reaction, Model SEED queries the RAST annotation for various functional roles associated with cell wall types. After the cell wall type is identified, Model SEED selects one of the four matching generic biomass composition reactions. Upon the selection of the general biomass composition reaction, the remaining metabolites are included.

The first sets of metabolites included are universal metabolites, such as amino acids and nucleotides, which are generally included in all biomass composition reactions. The remaining metabolites are determined through querying the RAST annotation for functional roles associated with the biosynthesis or utilization of the metabolites. Approximate stoichiometric coefficients, or relative quantities for each metabolite are determined based on the general biomass composition reaction applied and an associated reference organism. (Devoid et al, 2013)

The energy cost required for the biomass composition reaction is then determined. The energy category is one of the most important terms in the biomass composition reaction, representing the ATP consumption by the organisms to produce biomass. The stoichiometric coefficient for the energy parameter is also determined by the general biomass composition reaction template used. (Devoid et al, 2013)

Because the stoichiometric coefficient determined during the automated construction of the draft biomass composition reaction is only an approximation, manual curation is required. Manual curation is crucial because the biomass composition reaction is a critical component of the draft GEM construction process and therefor the resulting metabolites and stoichiometric coefficients included in the draft biomass composition reaction can significantly alter the function of the model. Thus, to validate the biomass composition reaction, extensive literature, biological, and experimental data is required to properly determine the stoichiometric coefficients and compounds included in the biomass composition reaction. (Devoid et al, 2013)

2.7.6 Auto-completion of a Draft GEM

After the draft GEM and biomass composition reaction are constructed through the automated protocol in Model SEED, the remaining metabolic gaps (i.e. reactions that do not link with the remaining metabolism) of the draft GEM must be completed. Metabolic gaps present in the draft GEM must be fixed to allow the draft GEM to produce all metabolites required by the biomass composition reaction. Model SEED auto-completes metabolic gaps in a draft GEM through a multi-step process outlined by Figure 2.14. (Devoid et al, 2013)

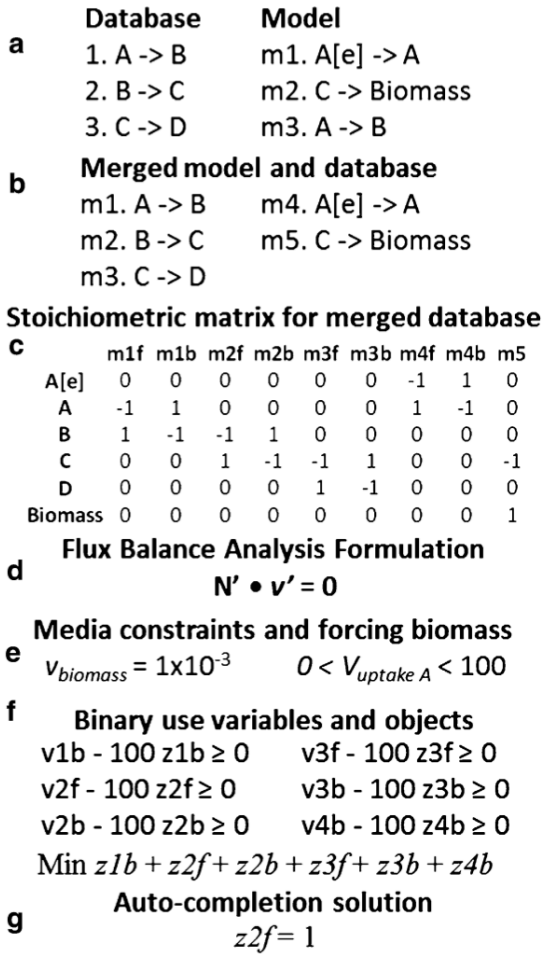

Figure 2.14: Auto-completion process of a draft GEM reconstruction (adopted from Devoid et al, 2013)

Model SEED auto-completes metabolic gaps by first preparing a biochemistry database composed of 10,516 reactions and 8,335 compounds. (Figure 2.14a) To insure that no irrelevant pathways are added by Model SEED, generic reactions, lumped reactions, and unbalanced reactions are removed from the biochemistry database. After the biochemistry database is prepared, it is merged with the draft GEM, and all redundant reactions are removed. (Figure 2.14b) (Devoid et al, 2013)

After composing a merged system of reactions with a uniform reaction name format, Model SEED constructs a stoichiometric matrix with the columns representing the reactions and the rows representing the compounds. (Figure 2.14c) Each reaction is decomposed into unique forward and reverse reactions, regardless of the actual reversibility of the reaction. The elements within the stoichiometric matrix are the stoichiometric coefficients of every compound from each respective reaction. The matrix is then applied to form linear mass balance constraints for a flux balance analysis (FBA) problem by setting the product of the stoichiometric matrix and the vector of fluxes equal to zero. (Figure 2.14d)

Additional constrains are then added to the linear optimization problem. (Figure 2.14e) For example, the biomass composition reaction is constrained to always have a positive flux value. Additionally, exchange fluxes are added for all compounds that occur in the extracellular compartment of the GEM reconstruction. The bounds of the exchange fluxes are adjusted based on the media selected, which is generally always complete media, or media that contains all the chemical compounds found in the Model SEED database. Complete media is usually selected because it allows for the exchange of all transportable compounds in the model. (Devoid et al, 2013)

To identify which reactions should be included when the biomass composition reaction is optimized through FBA, binary variables are created for each candidate reaction that did not appear in the draft GEM reconstruction. (Figure 2.14f) Each binary variable is given a value of one or zero. If the binary variable is equal to one, its associated reaction is active when a flux is applied to the biomass composition

reaction. If the binary variable is equal to zero, its associated reaction is not active when a flux is applied to the biomass composition reaction.

The auto-completion optimization problem is set to have an objective function that minimizes the sum of the binary variables multiplied by a set of cost coefficients. The cost coefficients are calculated for each reaction based on the "thermodynamic feasibility, completion of existing pathways, the confidence in the biochemistry, and the amount of information available for the biochemistry." (Devoid et al, 2013) The cost coefficient may also be calculated by utilizing the BLAST similarity scores for the genes linked to gap filled reactions in related genomes. (Devoid et al, 2013)

Solutions identified for the auto-completion optimization problem signify a set of reactions that must be included or made reversible in the GEM reconstruction. (Figure 2.14g) Each solution to the auto-completion optimization problem will allow the GEM to permit a flux through the biomass composition reaction. Gap filled reactions are selected by Model SEED that minimize the sum of the binary variables multiplied by a set of cost coefficients. Gap filled reactions identified through this process will then be added to the GEM reconstruction. While Model SEED attempts to reliably gap fill the metabolism of the organism, manual curation is required to validate the selected gap filled reactions. (Devoid et al, 2013)

2.7.6.1 Errors Introduced During an Automated Reconstruction

While automated reconstructions of metabolic networks have streamlined GEM construction, there are also errors introduced during this step. Errors introduced through an automated reconstruction include (1) out of date genome annotations, (2) incorrect annotations due to missing genes or errors in the gene finding algorithms, (3) missing functionalities of enzymes, and (4) unknown transporter reactions. Various biological databases, such as KEGG, are also called upon during the assembly of an automated draft reconstruction, introducing the potential for additional errors. Errors stemming from biological databases used during an

automated reconstruction include (1) Gene-Protein-Reaction (GPR) associations, (2) reaction specificity, (3) reaction stoichiometry and directionality, (4) compounds protonation states, (5) coenzyme availability, and (6) organism specific pathways. Detailed tables of common errors and possible solutions are found in Table A.1 and Table A.2 in the Appendix (adopted from Feist et al, 2008).

2.7.7 Completed GEM Review

Following the auto-completion of a draft GEM reconstruction by Model SEED, it is important to review the completed GEM reconstruction. After the automated reconstruction of a GEM, completed automated GEMs may be downloaded in Linear Programming (LP), Microsoft Excel, and Systems Biology Markup Language (SBML) formats. While Model SEED features numerous tools to review and analyze the GEM reconstruction, external tools have also been developed to review, analyze, and simulate GEM reconstructions. Utilizing the COBRA Toolbox and the downloaded SBML GEM reconstruction, the protocol outlined by Devoid et al, 2013 can be followed. (Schellenberger et al, 2011)

Each reaction in the GEM reconstruction is first verified utilizing updated KEGG maps, as KEGG maps utilized by Model SEED may be outdated. Reactions may be verified using experimental, literature, and genomic data. Utilizing KEGG maps, metabolic pathways present in the organism can be easily visualized. KEGG maps are also critical to determine which reactions may be gap-filled or removed. For example, when there are numerous reactions in a metabolic pathway, but only single steps in the pathway missing, the missing reactions may be prime targets for gap filling. Conversely, when lone reactions are found completely detached from the metabolic network, these reactions may be prime examples for removal. When adding or removing a reaction, it is important to consult experimental, literature, and genomic data to verify the addition or removal of a target reaction.

In addition to the manually gap-filled reactions, gap-filled reactions included during the auto-completion process by Model SEED must also be manually verified through a similar process to determine if the proper reactions were included. If after consulting experimental, literature, and genomic data the auto-completed gap-filled reaction appears correct, the gap-filled reaction should be included in the GEM reconstruction. If the auto-completed gap-filled reaction appears to be incorrect, the reaction should be removed.

When removing any gap-filled reactions, it is crucial to understand which metabolic pathways the reaction is used in, and re-verify that the metabolic network is correctly linked. While the individual gap-filled reaction may be incorrect, it still suggests that there is a missing, or incorrectly linked, reaction found somewhere in that metabolic pathway. It is then important to analyze and identify any candidate reactions in the pathway for addition to the GEM reconstruction.

Following the gap-filling process, the biomass composition reaction can be analyzed and reviewed utilizing the COBRA Toolbox in MATLAB. (MathWorks, 2013b) By comparing biomass composition reactions of phylogentically close organisms and applying experimental, literature, and genomic data, each compound included in the biomass composition reaction can be verified. The biomass composition reaction is the primary critical component of a GEM, and influences the operation and validity of the resulting simulations. Thus it is critical that each compound included in the biomass composition reaction is tested one by one to confirm that the compound is produced by the GEM and that it is properly connected within the metabolic network. To confirm that each compound is properly connected, the biomass composition reaction is reconstructed compound by compound and FBA is used to identify if growth is possible at each step.

2.8 Flux Balance Analysis

Various constraint based analysis techniques are applied to GEMs to evaluate to flow of metabolites through a metabolic network. (Orth et al, 2010) The predominant constraint based analysis approach used in genome-scale modeling is FBA. Through FBA, an objective function can be optimized by the constraints imposed by the mass balance and capacity constraints of the GEM. Utilizing the optimization of an object function, a unique solution to the objective function can be identified in a large solution space.

One of the key benefits of FBA is that it does not require kinetic data and solutions can be calculated quickly through various software toolboxes, such as COBRA Toolbox. In addition to predicting the flow of metabolites, constraint based tools can be used in gene knockout simulations, gap-filling, physiological studies, and guided metabolic engineering. (Orth et al, 2010; Feist & Palsson, 2008)

Upon verification that each compound, reaction, and element of the biomass composition reaction is correct, the model can be validated and queried against experimental and literature data using FBA. If the flux distribution produced through FBA does not align with experimental data, the reaction and biomass composition reaction must be re-analyzed to identify any remaining errors. The flux distribution may also be used to identify any reactions that generate excess carbon or energy in an unfeasible manner.

2.8.1 Mathematics of FBA

Metabolic reactions in a GEM are mathematically represented as a stoichiometric matrix (**S**) of size m x n, where m represents a row for each unique metabolite and n represents a column for each reaction. For each reaction where a metabolite is being consumed, a negative stoichiometric coefficient is applied. When a metabolite is being produced, a positive stoichiometric coefficient is applied. If a

metabolite listed in a row *m* does not participate in a reaction, it is given a value of zero in the **S** matrix. The resulting **S** matrix transcribed from the GEM will be sparse, as each metabolite is only actively involved in a subset of reactions in the GEM. (Orth et al, 2010)

One of the key assumptions made in FBA is that the GEM is at steady state, where metabolite concentrations are no longer changing. This is represented by the equation $d\mathbf{x}/dt = 0$, where the vector **x** represents the concentrations of all metabolites in the system. Given that the GEM is under steady state conditions, the system of mass balance equations can be represented by $\mathbf{S}*\mathbf{v} = 0$, where the vector **v** represents the flux of all the reactions in the system. Because there are generally more reactions than metabolites in a GEM, this system contains more unknown variables than equations, yielding a system of equations with no unique solutions. (Orth et al, 2010) While a unique solution cannot be calculated at this stage, the constraints imposed by the S matrix and capacity constraints imposed by the upper and lower bounds of a reaction yield a defined solution space. (Palsson, 2006)

To identify a unique point on the constraint imposed solution space, an objective function is maximized or minimized. An objective function is formed from the idea that the system, representing a microorganism, is constructed to behave optimally through the process of evolution. A common objective function used in FBA is to maximize the specific growth rate of an organism. Maximizing the growth rate is a valid objective function from the observation that a microorganism has been constructed to grow as efficiently as possible through the process of evolution.

The objective function, **Z**, is mathematically defined as $\mathbf{Z} = \mathbf{c}^{\mathbf{T}}*\mathbf{v}$, where c is a vector of weights on fluxes that contribute to the objective function. In the case of the optimization of the specific growth rate, $\mathbf{c}^{\mathbf{T}}$ would contain a list of zeroes, with a one in the location of the reaction being optimized. Through the use of linear programming, the objective function is optimized and a solution is located on the edge of the previously defined solution space. FBA produces a flux distribution in the units of millimoles per gram of dry cell weight per hour (mmol/gDW-hr). For the

biomass composition reaction, this further reduces down to the specific growth rate (h^{-1}). (Orth et al, 2010) A summary of the FBA process is shown in Figure 2.15.

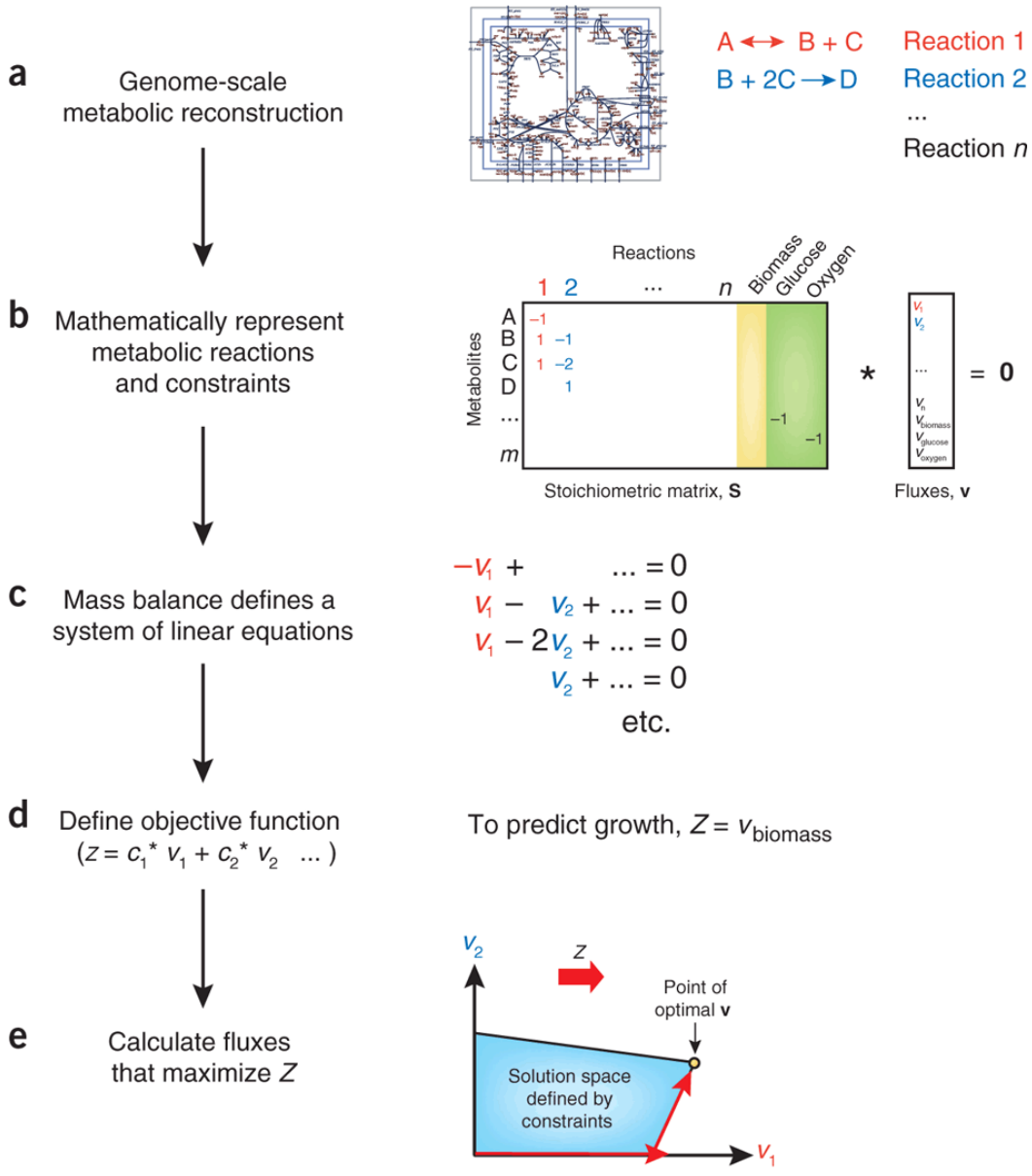

Figure 2.15: Summary of a FBA Problem (adopted from Orth et al, 2010)

2.9 Modeling Microbial Communities With GEMs

Microorganisms are rarely found alone in the environment, and play an important role in many natural cycles. Microbial communities are responsible for the global cycling of energy and nutrients, such as carbon and nitrogen, and are critical in the biodegradation of pollutants, wastewater treatment, production of biofuels, and other biotechnology processes. (Follows et al, 2007; Stephanopoulos, 2007; Peng et al, 2008; Wagner & Loy, 2002; Daims & Wagner, 2006; Sabra et al, 2010; Zomorrodi & Maranas, 2012) Within these microbial communities, individual organisms may operate at various levels of their optimal activity. Some microorganisms may form a syntrophic relationship with other microorganisms where they both benefit from the association while other microorganisms may be negatively impacted due to the competition of scarce resources. The inter-species relationships of the microbial community have been shown to fluctuate as a response from environmental stimuli and plays a critical role in the activity and population of the microbial community. (Hansen et al, 2007; Kerr et al, 2002; Tilman, 2004; Xavier, 2011; Fuhrman, 2009; Zomorrodi & Maranas, 2012)

To better understand the interactions and tradeoffs between species within a microbial community, attempts to model microbial GEM communities have been made. The first microbial community GEMs were constructed by creating separate compartments for each GEM and allowing the transfer of metabolites between compartments. (Mo et al, 2007; Dobson et al, 2010; Bizukoic et al, 2010). Numerous microbial community GEMs have been constructed that model positive inter-species relationships, negative inter-species relationships, or synthetic relationships between mutants of the same species. These models were constructed and analyzed using a variety of methods including FBA, dynamic FBA, minimization of metabolic adjustment (MOMA), evolutionary game theory, non linear systems, and stochastic process. (Segre et al, 2002; Zhuang et al, 2011; Mahadevan et al, 2002; Frey, 2010; Lehmann & Keller, 2006; Schuster et al, 2010; Zomorrodi & Maranas, 2012)

2.9.1 *Development of OptCom to Model Microbial Communities*

Existing techniques to model microbial communities are unable to solve for the multi-level nature of decision making in microbial communities. (Zomorrodi & Maranas, 2012) To address these issues, a flux balance analysis framework, which utilizes a multi-level optimization description, known as OptCom was developed. Applying OptCom allows for the optimization of individual species fitness function versus the microbial community fitness function and evaluates the tradeoffs that each individual species makes for the fitness of the entire community.

Unlike previous attempts to model microbial community interactions, OptCom is able to capture positive interactions, negative interactions, or a combination of both for any number of species in the microbial community. OptCom creates a unique biomass optimization function for each species in the community as the inner level of the framework. The microbial community biomass is optimized in the outer level of the framework. Linking the inner and outer levels of the framework, inter-species exchange constraints and optimality criteria are applied in the outer problem.

Applying the same notation from FBA, this framework can be seen in Figures 2.16 and 2.17. (Zomorrodi & Maranas, 2012) OptCom generates a non-convex bi-linear optimization problem, and calls upon the Branch-And-Reduce Optimization Navigator (BARON) solver in The General Algebraic Modeling Systems (GAMS). (Tawarmalani & Sahinidis, 2005; GAMS Development Corporation, 2013) BARON solves for the global solution of nonlinear and mixed integer nonlinear programs.

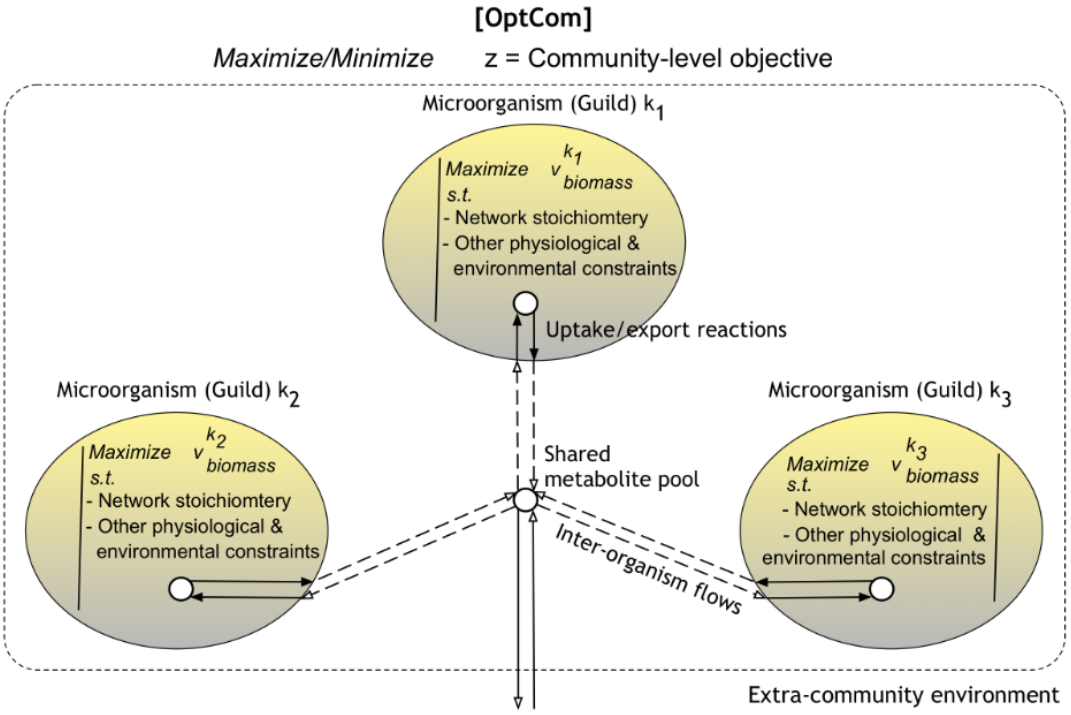

Figure 2.16: Visualization of OptCom (adopted from Zomorrodi & Maranas, 2012)

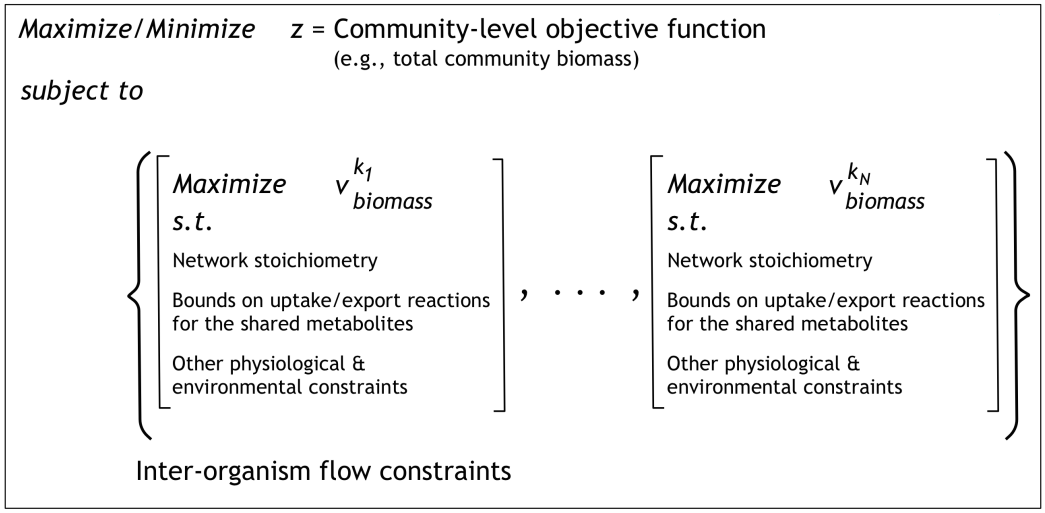

Figure 2.17: Multi-level optimization framework of OptCom (adopted from Zomorrodi & Maranas, 2012)

2.9.2 Descriptive-OptCom

Through the application of descriptive-OptCom, species are able to function below their optimal levels of activity for the benefit of the entire microbial community. To measure an individual species deviation from their optimal activity, a metric for each species ***k*** was developed called the optimality level or c^k. The optimality level for each species is calculated through a variant of OptCom known as descriptive-OptCom. (Zomorrodi & Maranas, 2012) In descriptive-OptCom, the inner framework is composed of all data related to the constraints of individual species of the microbial community while the outer framework includes all data related to the constraints of the community biomass. Descriptive-OptCom allows the biomass flux of individual species to rise above or fall below the maxima ($vopt^k_{biomass}$) calculated using OptCom. The framework for descriptive-OptCom can be seen in Figure 2.18. (Zomorrodi & Maranas, 2012)

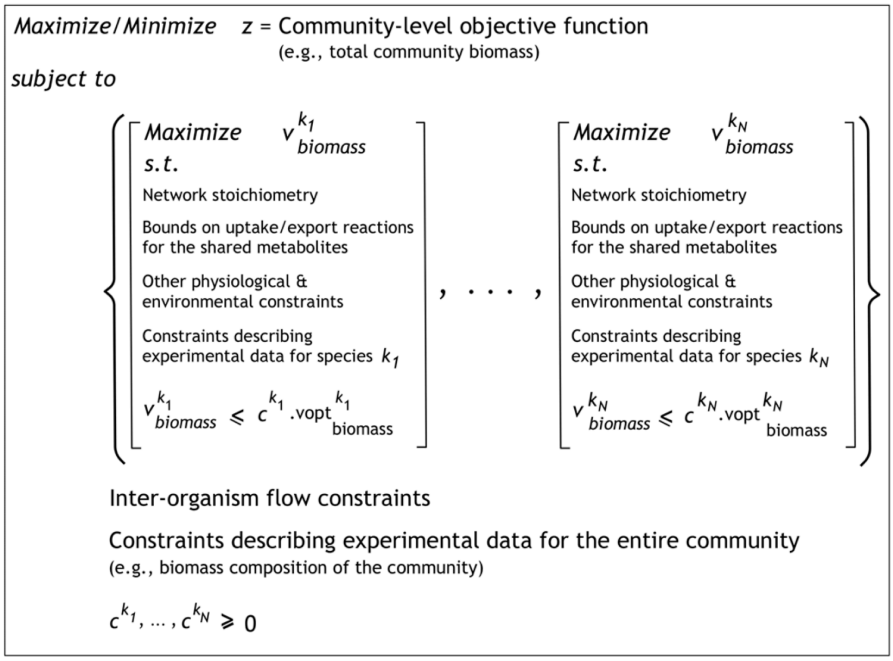

Figure 2.18: Multi-level optimization framework of descriptive-OptCom (adopted from Zomorrodi & Maranas, 2012)

Solving for the optimality level using descriptive-OptCom presents three possible scenarios. 1) An optimality level of less than one for a species in the community suggests that the microorganism grows sub-optimally at the rate 100% x c^k of $vopt^k_{biomass}$, 2) an optimality level of one suggests that the species grows at 100% of the optimum level equal to $vopt^k_{biomass}$, and 3) an optimality level greater than one suggests that the species is growing at a greater rate than the community level specific maximum by depleting resources from other members in the community, resulting in sub-optimal growth for the remainder of the community. (Zomorrodi & Maranas, 2012)

CHAPTER 3
MATERIALS AND METHODS

3.1 Definition of a Pure Culture Anaerobic Digester

To construct an effective and efficient GEM of a pure culture anaerobic digester, the microbial community must be first reduced to a minimal set of microorganisms. The set of microorganisms selected must fit a range of parameters to produce a reliable model that is able to represent the entire community. 1) The first parameter is that the microorganisms must accurately represent the key reactions that occur within the anaerobic digester. 2) To assist future research and validation, the second parameter is that the microorganisms must be commercially available for culture in a laboratory. 3) The final parameter is that the microorganisms must already have a validated GEM available, or at a minimum, have an accurate genome sequence with growth and experimental data available in the literature from which a GEM can be constructed and validated.

In the definition of a pure culture anaerobic digester, the hydrolysis process was not included in the pure culture anaerobic digester GEM. While the hydrolysis of complex organic substrates is a vital step in anaerobic digestion, it is typically performed by a wide array of extracellular enzymes from an even wider array of bacteria. The addition of hydrolysis to a GEM could be accomplished by constructing a separate external compartment and applying known reaction constraints, however, the application of this serves little function to the overall model and creates further modeling complexities. Thus, the critical processes included in the definition of a pure culture anaerobic digester are acidogenesis, acetogenesis, and methanogenesis.

3.2 Acidogenic Organism

The acidogenic bacteria that fulfilled all the necessary parameters for the pure culture anaerobic digester GEM (Section 3.1) was *Clostridium acetobutylicum* ATCC 824. *C. acetobutylicum* is a gram positive strict anaerobic bacteria from the Order *Clostridia* and is commonly found in anaerobic digesters. (Wirth et al, 2012) At the neutral pH commonly found in anaerobic digesters, *C. acetobutylicum* ferments simple sugars, such as glucose, into to acetate, butyrate, hydrogen, and carbon dioxide. (Bahl et al, 1982) When the pH drops below 5.1, *C. acetobutylicum* enters a solventogenic phase and transforms acetate and butyrate to acetone and butanol. In addition to being fully sequenced, *C. acetobutylicum* ATCC 824 has had three GEMs produced, due to interest in *C. acetobutylicum's* ability to produce butanol during solventogenesis. The most recent GEM constructed for *C. acetobutylicum* is iCac802, constructed by Dash et al, 2014. iCac802 features 802 genes and 1,462 reactions that utilize 1,137 metabolites.

3.3 Acetogenic Organism

The acetogenic bacteria that met the bulk of the prescribed parameters in Section 3.1 was identified as *Syntrophomonas wolfei* ATCC BAA-1933. *S. wolfei* is a gram negative bacteria from the Order *Clostridia*. (Sobieraj & Boone, 2006; Sieber et al, 2010) *S. wolfei* is capable of growth on fatty acids of four to eight carbons in length, although it prefers four carbon fatty acids. As a byproduct of beta-oxidation, *S. wolfei* produces acetate and hydrogen, which must be continually removed by a syntrophic partner. (Sobieraj & Boone, 2006) Additionally, methanogens are commonly found in a syntrophic relationship *S. wolfei* in anaerobic digesters, as they utilize the acetate and hydrogen in the production of methane. (Beaty & McInerney, 1989) While a sequence for *S. wolfei* is available, a thorough GEM has not been previously produced. The construction of a GEM for *S. wolfei* is detailed in Section 4.

3.4 Methanogenic Organism

Following the prescribed parameters of the pure culture anaerobic digester in Section 3.1, the methanogen *Methanosarcina barkeri* ATCC BAA-1921 was selected. *M. barkeri* is a gram positive Archaea from the Order *Methanosarcinales*. (Gonnerman et al, 2013) *M. barkeri* was chosen due to its diverse metabolism, which is capable of fermenting acetate, hydrogen, methanol, and methylamines into methane and carbon dioxide. (Balch et al, 1979) *M barkeri* has two high quality GEMs previously constructed, with the most recent being iMG746, constructed by Gonnerman et al, 2013. iMG746 features 746 genes, 816 reactions, and 718 metabolites.

3.5 Construction of a Genome Scale Model for *S. wolfei*

Prior to the formation of a GEM for a pure culture anaerobic digester, a GEM for *S. wolfei* was constructed. The *S. wolfei* GEM was constructed utilizing the automated protocol provided by The SEED in addition to the GEM reconstruction protocols previously outlined. (Thiele & Palsson, 2010; Santos & Teusink, 2011) Following the automated draft metabolic reconstruction, each reaction of the draft GEM was manually verified and curated against the KEGG database and existing literature data.

3.6 Draft Reconstruction

The genome for *S. wolfei* was downloaded from the National Center for Biotechnology Information (NCBI) in a FASTA format. (NCBI, 2010) An overview of the *S. wolfei* genome, adopted from Sieber et al, 2010, can be found in Table 3.1. *S. wolfei* features one circular chromosome composed of 2,936,195 base pairs (bp).

The *S. wolfei* genome is of a similar size to many other syntrophic organisms and features a slightly lower rate of functional assignments of ORFs than many other similar organisms. (Sieber et al, 2010)

The genome was submitted to RAST for an automated annotation. Following the automated annotation of the *S. wolfei* genome, the Model SEED was used to construct a draft GEM reconstruction using the algorithm discussed in Chapter 2. Following the automated draft GEM reconstruction produced by the Model SEED, the *S. wolfei* GEM was manually curated.

Table 3.1: Features of the *S. wolfei* genome (adopted from Sieber et al, 2010)

Category	**Amount**
DNA, total	2,936,195 bp
DNA, coding	2,489,888 bp
G + C content	44.87%
DNA scaffolds	1
Genes total number	2,677
Protein coding genes	2,574 (97.5%)
RNA genes	65 (2.5%)
rRNA genes	19
5S rRNA	13
16S rRNA	3
23S rRNA	3
tRNA genes	46
Genes with function prediction	1,507 (57.1%)
Genes without function prediction	1,067 (40.4%)

3.7 Manual Curation

The completed automated draft GEM was downloaded from the Model SEED and manually curated to further refine the GEM. Due to inaccuracies produced by the algorithm employed by Model SEED, major edits were required. *S. wolfei* was constructed as a gram positive bacteria by the Model SEED, although multiple literature reviews on *S. wolfei* report that it stains gram negative. (Sieber et al, 2010; McInerney et al, 1981; Lorowitz et al, 1989) In addition to transforming the *S. wolfei* GEM from a gram positive to a gram negative model, the biomass composition reaction was re-built. After rebuilding the biomass composition reaction, each reaction in the *S. wolfei* GEM was manually verified with data from the KEGG database. Reactions that were incorrectly added during the automated draft reconstruction were flagged for removal, and subsequently removed after extensive validation. Gap filled reactions, added during the automated draft reconstruction to complete metabolic pathways, were individually reviewed and compared to existing literature data as well.

3.7.1 Growth Media

The Model SEED used complete media during the construction of a GEM to identify candidate reactions and metabolites accurately. Using complete media in the GEM allows for all compounds within the database to be transported across the cell boundary, if required for growth. While complete media may be beneficial during the automated reconstruction of a draft GEM, it may produce a number of errors. (Devoid et al, 2013)

Due to the use of complete media, synthesis of many amino acids and various metabolic intermediates were not included in the reconstruction. These reactions and metabolites were added during the manual curation and gap filling process. Additionally, the uptake of many compounds found in the defined media, listed in

Table 3.2, were not included in the automatic reconstruction and were manually added. (Beaty & McInerney, 1990) Finally, when grown in pure culture, *S. wolfei* utilizes crotonate as a carbon source. (Beaty & McInerney, 1987) However, due to the use of complete media during GEM reconstruction the crotonate uptake pathway was missing. Therefore, the crotonate uptake reactions were manually added to the GEM.

Table 3.2: Defined media for growth of *S. wolfei* (adopted from Beaty & McInerney, 1990)

Defined Media		
Required Growth Factors	**Trace Metal Solution**	**Mineral Solution**
Biotin (B7)	Mn	K
Cyanocobalamin (B12)	SO_4	PO_4
Thiamine (B1)	Fe	Mg
p-Aminobenzoic Acid (Bx)	NH_4	Cl
Lipoic Acid	Co	Na
	Cl	NH_4
	Zn	Ca
	Cu	
	Ni	
	Na	
	MoO_4	
	SeO_4	
	WO_4	

3.7.2 Biomass Composition Reaction

The biomass composition reaction composed by the Model SEED was replaced entirely due fatal flaws in the initial automatic reconstruction. *S. wolfei* was

incorrectly labeled as a gram positive bacteria, resulting in a biomass composition reaction that followed a gram positive template. Numerous studies show that *S. wolfei* forms a gram negative cell wall, complete with peptidoglycan but lacking a lipopolysaccharide layer. (Sieber et al, 2010) This unusual gram negative cell wall construction may have been the reason that Model SEED predicted it to be a gram positive organism.

To replace the incorrect biomass composition reaction, the biomass composition reaction from the highly validated *Escherichia coli* GEM was used as the scaffolding for the manual construction of the *S. wolfei* biomass composition reaction. The *E. coli* biomass composition reaction was selected as the scaffolding since it contained the gram negative biomass composition reaction required by *S. wolfei* and it has been extensively curated and validated. (Feist et al, 2007)

While the general format of the *E. coli* biomass composition reaction was followed, all relevant data from *S. wolfei* was utilized to manually curate the biomass composition reaction. For example, the amino acid profile and DNA composition was modified using the known GC content and relative composition of the cell. (Sieber et al, 2010) The cell wall was revised based on literature data that suggests that *S. wolfei* forms a unique gram negative cell wall. The cell wall of *S. wolfei* was found to have peptidoglycan, but lack a lipopolysaccharide layer. (Sieber et al, 2010) Thus, the lipid profile was adjusted, based on previous phospholipid fatty acid studies, to create *S. wolfei's* unique cell wall. (Henson et al, 1988)

Additionally, polyhydroxybutyrate (PHB) was found to represent up to 20% of the cell by weight. (Beaty & McInerney, 1987) PHB is typically produced when a microorganism is under stress, however, *S. wolfei* uniquely produces PHB under standard growth conditions. It is hypothesized that *S. wolfei* utilizes PHB as an energy storage mechanism and metabolizes the PHB when environmental conditions make the beta-oxidation of butyrate thermodynamically unfavorable. (Sieber et al, 2010; Beaty & McInerney, 1987) The inorganic ions and soluble pool included in the *S. wolfei* biomass composition reaction were based almost entirely on the *E. coli*

biomass composition reaction, except for the substitution of menaquinones, critical in the transfer of electrons, typically found in anaerobic microorganisms. (Unden, 1988)

3.7.3 Model SEED Gap Filled Reactions

Through the automated draft reconstruction process performed by Model SEED, numerous reactions were included in the gap filling process. Each gap filled reaction was manually verified against the KEGG database and existing literature. Gap filled sink and exchange reactions from the complete media were removed and replaced with sink and exchange reactions for the *S. wolfei* defined media. Gap filled sink and exchange reactions for the uptake of amino acids were also removed, as the defined media used for growth in pure culture did not have any amino acids supplemented.

3.8 GEM Evaluation

Upon the manual curation of each reaction included in the automated draft reconstruction, the biomass composition reaction, and the gap filled reactions, the entire GEM was evaluated and debugged. The GEM is evaluated through the application of FBA, with the maximization of the biomass composition reaction serving as the objective function. Utilizing FBA, no solution was obtainable, which required the entire metabolic network to be re-evaluated. Through the manual re-evaluation of the entire metabolic network, new gap filled reactions were identified and added to the *S. wolfei* GEM.

3.8.1 Manual Gap Filling

Gap filled reactions were added based on existing literature data or as a requirement to complete the metabolic network. The first step in identifying the gap

filled reactions was to evaluate the *S. wolfei* GEM for each compound included in the biomass composition reaction using FBA. After the biomass composition reaction was linked to the metabolic network to achieve biomass growth, metabolic reactions were added to the GEM only if explicitly stated in literature. (iTK530 Excel)

3.8.1.1 Linking Biomass Components

The *S. wolfei* GEM was evaluated for each individual compound in the biomass composition reaction. By evaluating the metabolic network for each unique compound included in the biomass composition reaction, reactions required to link the biomass composition reaction to the metabolic network can be more readily identified. When applying FBA for a single compound in the reduced biomass composition reaction, positive growth indicates the GEM will produce the required compound and an infeasible solution indicates that the compound is not linked to the network. The metabolic network can then be examined through the use of metabolic maps provided by KEGG and Model SEED. Utilizing the metabolic maps, the pathway leading to the synthesis of each compound can be examined and missing reactions can be identified. This process was completed for each compound in the biomass composition reaction.

Due to the inclusion of amino acid sinks and exchange reactions, numerous amino acids required gap filled reactions for production. While the amino acid profile has not been sufficiently characterized, the *E. coli* GEM scaffolding used has been extensively validated. (Feist et al, 2007) Gap filling the amino acid profile followed the previously described protocol. (Devoid et al, 2013) After shutting down the exchange of amino acids by the *S. wolfei* GEM, amino acids were checked for connectivity to the metabolic network one by one. Any amino acid that was not produced was analyzed and reactions were gap filled. (iTK530 Excel)

The lipid profile included in the automated draft reconstruction was entirely removed and updated with lipid data for *S. wolfei*. Due to the unique lipid profile and

the incorrect assignment of cell wall type by the Model SEED, the entire lipid profile was not connected to the metabolic network. Additionally, some lipids were not included in the Model SEED compound database, and were manually added to the model. Reactions for compounds manually added were constructed based on the reaction pathway of similar lipids or from literature data. Due to the number of lipids not present in the Model SEED compound database, the lipid profile required a large number of gap filled reactions. (iTK530 Excel)

3.8.2 Futile Cycles and Unbalanced Reactions

After the biomass composition reaction was completely linked to the metabolic network, reactions were analyzed for futile cycles and unbalance reactions. A futile cycle occurs when two metabolic pathways are coupled and run in opposite directions, resulting in a cycle that does not produce any valuable metabolic products or acts as an energy sink. Unbalanced reactions are reactions that have stoichiometric imbalances, resulting in the net production or loss of a compound.

Futile cycles and unbalanced reactions can be identified through the analysis of the flux distribution produced through FBA. Futile cycles and unbalanced reactions will often produce flux values of positive or negative 1000, which can be viewed as positive or negative infinity. (Orth et al, 2011) An example of an unbalanced reaction that was included in the automated draft reconstruction can be seen in rxn05817 (Table 3.3). In rxn05817, dGTP was produced by a single diphosphate. When applying FBA, this reaction would run at its maximum level because it is generating free carbon and energy to use throughout the cell. This reaction generates free carbon and energy because it is converting one mol of diphosphate into a mol of dGTP, which is incorrect. Clearly, this reaction is infeasible and rxn05817 was removed from the *S. wolfei* GEM.

Table 3.3: Unbalanced reaction example from Model SEED (adopted from Overbeek et al, 2005)

Reaction ID	SEED ID: rxn05817 KEGG ID: R00376
Equation	dGTP <=> PPi
Chemical Equation	$C_{10}H_{13}N_5O_{13}P_3$ <=> $H_2O_7P_2$

3.8.3 Substrate Uptake and Product Secretion

Substrate uptake and product secretion rates were constrained based on literature data. Crotonate served as the primary carbon source, and acetate and hydrogen served as the main product secretions, with a small amount of butyrate and hexanoate additionally produced. (Beaty & McInerney, 1987) All compounds described in the defined growth media (Table 3.2) were allowed transport into the cell. Due to the lack of literature data defining required growth co-factor uptake rates, growth media compounds were allowed to enter the cell as required by the GEM for biomass growth.

3.8.4 Continued Evaluation

After each key aspect of the *S. wolfei* GEM was manually curated, FBA was performed to examine the growth rate. The growth rate after the first pass at manual curation was infeasible due to numerous futile cycles and unbalance reactions, which required further manual curation. Utilizing the flux distribution produced through FBA, reactions that produced flux values near or at the upper and lower limit of +/- 1000, representing infinite flux, were examined for validity.

Reactions that continued to produce futile cycles were constrained through the reversibility parameter based on literature data and the logical flow of metabolites through the metabolic network. After a reaction was constrained, FBA was performed

to generate a new set of flux distributions. This process was continued until all major futile cycles and unbalanced reactions were constrained or removed. (iTK530 Excel)

3.9 Construction of a Pure Culture Anaerobic Digester GEM

Upon the completed curation and verification of the *S. wolfei* GEM, the three selected models were linked together through OptCom, performed in GAMS. As detailed in Figure 2.7, the three models included in OptCom share a pool of external metabolites. The uptake of one compound, such as acetate by *M. barkeri*, was programmed to be less than or equal to the combined production of acetate by *C. acetobutylicum* and *S. wolfei.* This process was repeated for all shared metabolites between the three GEMs.

Prior to programing the pure culture anaerobic digester GEM, the *C. acetobutylicum* GEM (iCac802), the *S. wolfei* GEM (iTK530), and the *M. barkeri* GEM (iMG746) were reduced to the active reactions using the removeDeadEnds function in the COBRA Toolbox. By reducing the models to only the active metabolic network, metabolic dead end reactions and metabolites are removed from the pure culture anaerobic digester model, reducing the complexity of the model while preserving the metabolic functions of each GEM.

The pure culture anaerobic digester was programmed in OptCom and descriptive-OptCom based on the protocols and examples provided by Zomorrodi & Maranas, 2012. Through OptCom, the community biomass and individual species biomass were optimized. Descriptive-OptCom parallels OptCom, except descriptive-OptCom allows for each species biomass growth rate to fall above or below a previously determined optimum in order to maximize the community level biomass growth. To fit both OptCom and Descriptive OptCom data to literature results, additional flux constraints were applied to uptake and secretion fluxes. (iTK530 Excel) After programming OptCom and descriptive-OptCom, the flux distribution of the pure culture GEM was solved in GAMS using the BARON global solver.

CHAPTER 4
iTK530 RESULTS AND DISCUSSION

4.1 iTK530 Summary

The *S. wolfei* GEM (iTK530) features 530 genes, 1,175 reactions containing 1,193 metabolites. iTK530 is capable of growth on crotonate and butyrate, which were found to be *S. wolfei's* primary carbon sources. iTK530 produces biomass on defined media specified by Beaty & McInerney, 1990.

The biomass composition reaction was completely rebuilt based on the validated *E. coli* model iAF1260, due to the incorrect cell wall designation by Model SEED, and was reconstructed using literature data. (Sieber et al, 2010; Beaty & McInerney, 1987; Beaty & McInerney, 1989; Amos & McInerney, 1990; Henson et al, 1988) iTK530 features an additional 143 reactions and an additional 83 metabolites when compared to the automated draft reconstruction (Table 4.1). The additional reactions and metabolites were manually added to reconstruct *S. wolfei's* biomass composition reaction and gap-fill *S. wolfei's* metabolism.

Table 4.1: Comparison of original draft *S. wolfei* GEM and iTK530

***S. wolfei* GEMs**		
	Draft Reconstruction	**iTK530**
Genes	528	530
Reactions	1032	1175
Metabolites	1110	1193

4.2 iTK530 Validation

To verify iTK530, FBA was performed and the resulting flux distribution was contrasted with published literature data. Uptake and secretion fluxes of major metabolites were constrained based on literature data from Beaty & McInerney, 1987 (Table 4.2). The application of these literature constraints did not produce growth (Table 4.2). To achieve the specific growth rate produced by Beaty & McInerney, 1987, the secretion of the three measured metabolic byproducts (acetate, butyrate, and hexanoate) were reduced by 30 percent. The secretion fluxes of the three metabolic products were reduced, as opposed to the crotonate uptake flux increased, because the secretion of metabolic byproducts is a function of the model, indicating the overall validity of the GEM. Through the reduction of the secretion fluxes, growth of iTK530 closely matched the specific growth rate detailed by Beaty & McInerney, 1987 (Table 4.3). The constraints applied to reach the validated specific growth rate are detailed in Table 4.3.

Table 4.2: Uptake and secretion rates of critical compounds (adopted from Beaty & McInerney, 1987)

Uptake and Secretion Rates of Critical Compounds (Beaty & McInerney, 1987)					
Compound	Crotonate	Acetate	Butyrate	Hexanoate	Growth[a]
Flux (mmol/gDW-hr)	0.9258	1.9840	0.1824	0.05670	0.01200

[a] Growth is measured as the specific growth rate (h^{-1})

Table 4.3: Uptake, secretion, and specific growth rates of iTK530

Uptake, Secretion, and Specific Growth Rates of iTK530					
Compound	Crotonate	Acetate	Butyrate	Hexanoate	Growth[a]
Flux (mmol/gDW-hr)	0.9258	1.3889	0.1289	0.0397	0.01204

[a] Growth is calculated as the specific growth rate (h^{-1})

The required 30 percent reduction of the secretion fluxes to achieve a literature validated growth rate is likely due to 1) a carbon intensive biomass composition reaction, 2) an overestimation of maintenance requirements, or 3) limitations of the GEM model and the reactions included. The draft biomass composition reaction produced by Model SEED was removed, due to incorrect cell wall assignment, and was replaced with a new biomass composition reaction based off of the *E. coli* iAF1260 model. While the newly constructed biomass composition reaction was constructed using available literature data, the literature data available was not sufficient enough to construct a highly accurate biomass composition reaction and may overestimate the amount of carbon required for biosynthesis by *S. wolfei*. The overestimation of required carbon coupled with *S. wolfei's* inherent metabolic limitations, due to its maximal energy production from the beta-oxidation of butyrate, resulted in the necessity to decrease the secretion fluxes to achieve a validated growth rate 0.012 h^{-1} (Table 4.3).

4.3 Maximum Acetate Production

The major metabolite produced from the beta-oxidation of butyrate by *S. wolfei* is acetate. Acetate is a critical metabolite in the anaerobic digestion process, and accounts for up to 70 percent of all methane produced in an anaerobic digester. (Wang et al, 2009) By increasing the acetate produced in an anaerobic digester, the

production of methane can be increased, resulting in an increase in the economic viability of biogas recovery.

To examine the maximum amount of acetate produced by *S. wolfei*, the lower limits of the fluxes for butyrate and hexanoate production were set to zero. Setting the lower limit of the butyrate and hexanoate production fluxes to zero does not completely shut off production, instead it only allows the metabolite to be produced if it is required to optimize biomass production. The acetate production rate that yielded a literature determined growth rate was 1.85 mmol acetate/gDW-hr, producing a growth rate of 0.012 h^{-1}. (Figure 4.1) (Beaty & McInerney, 1987; Beaty & McInerney, 1989; Amos & McInerney, 1990; Sieber et al, 2010) The maximum acetate production from the uptake of 0.9258 mmol crotonate/gDW-hr was 2.15 mmol acetate/gDW-hr, which resulted in a near zero *S. wolfei* growth rate. (Figure 4.1) A theoretical maximum growth rate of 0.1177 h^{-1} can also be calculated when no acetate is produced from the uptake of 0.9258 mmol crotonate/gDW-hr. The growth rate is greatly increased without acetate production, as acetate represents the major source of carbon leaving the cell. (Figure 4.1)

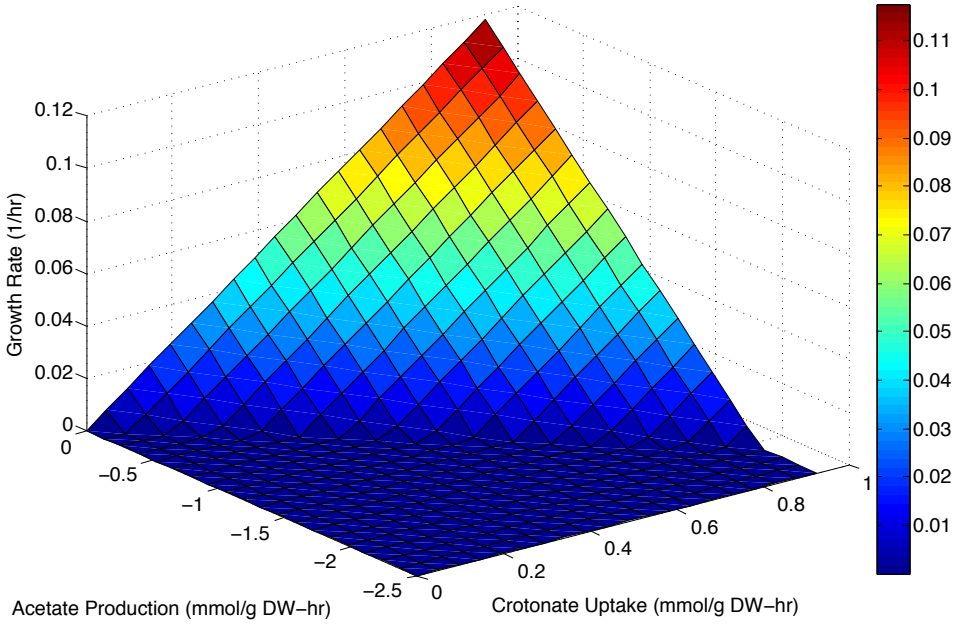

Figure 4.1: Crotonate uptake vs. acetate production vs. growth rate in *S. wolfei*

4.4 Maximum Formate Production

Formate production is a critical metabolic process in *S. wolfei*. (Sieber et al, 2010) Formate production is linked with the reoxidation of NADH and reduced electron transfer flavoproteins. Analysis of the *S. wolfei* genome predicted that there are five formate dehydrogenases, with two being externally oriented. (Sieber et al, 2010) Additionally, numerous methanogenic syntrophic partners are capable of utilizing formate in the production of methane. (Beaty & McInerney, 1989)

While no direct measurements of formate production by *S. wolfei* have been previously conducted, iTK530 predicts that the maximum formate production, coupled with the uptake of 0.9258 mmol crotonate/gDW-hr and secretion of 1.3889 mmol acetate/gDW-hr, is 2.86 mmol formate/gDW-hr. (Figure 4.2) Applying the previous crotonate uptake and acetate secretion fluxes, iTK530 predicts a formate production rate of 1.8 mmol formate/gDW-hr, yielding a biomass growth rate of 0.012 h^{-1},which remains consistent with literature values. (Figure 4.2) (Beaty & McInerney, 1987; Beaty & McInerney, 1989; Amos &McInerney, 1990; Sieber et al, 2010)

The theoretical maximum formate production, without acetate production present, is 8.5 mmol formate/gDW-hr, which results in near zero growth. (Figure 4.3) The production of 1.8 mmol formate/gDW-hr, when coupled with acetate production, and 7.6 mmol formate/gDW-hr when acetate production is absent, both produced a biomass growth rate near the lower limit of literature values. (Figures 4.2 and 4.3) The formate secretion rates increased without acetate production present because formate represents the primary source of carbon leaving the cell when acetate production is absent. (Beaty & McInerney, 1987; Beaty & McInerney, 1989; Amos &McInerney, 1990; Sieber et al, 2010)

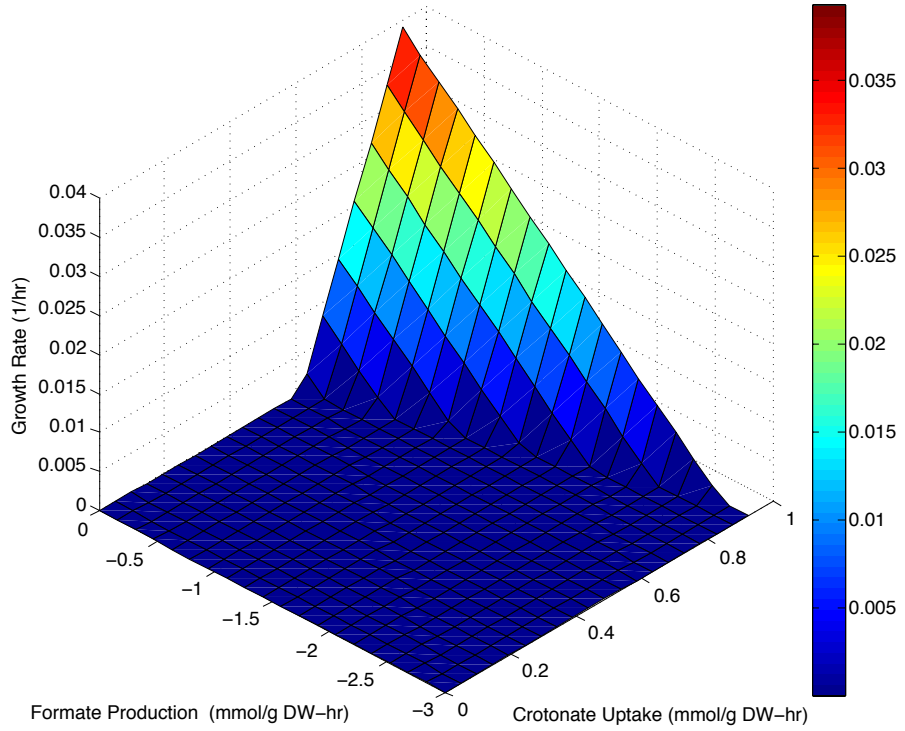

Figure 4.2: Maximum formate production coupled with acetate production in *S. wolfei*

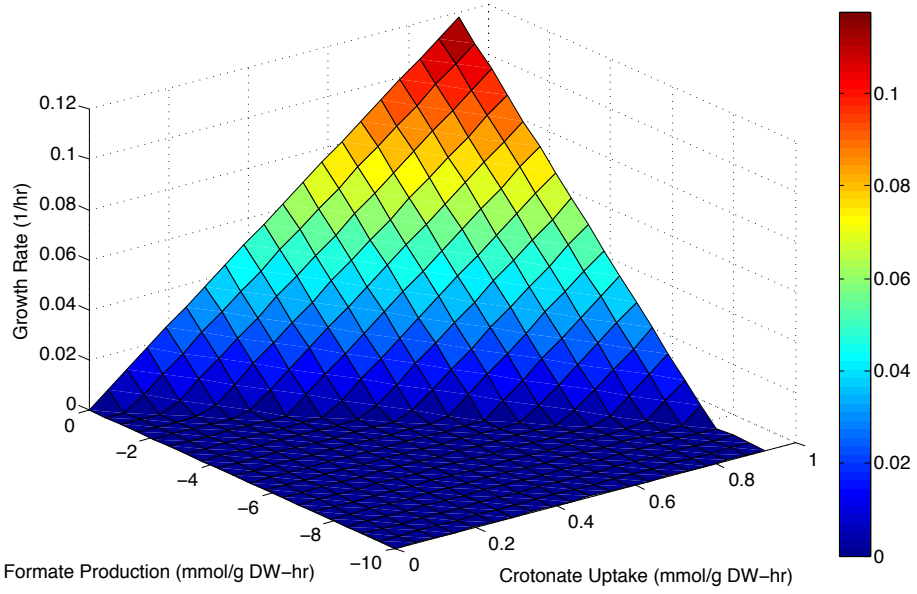

Figure 4.3: Maximum formate production without acetate secretion in *S. wolfei*

4.5 Maximum PHB Production

S. wolfei one of the few members of *Syntrophomonas* that is capable of producing polyhydroxyalkanoates (PHAs), including PHBs. (McInerney et al, 1992) PHB is a valuable bio-product that can be used to produce a sustainable bio-plastic. (Luengo et al, 2003) Typically, PHAs are produced during metabolic stress or nutrient limitation. (Anderson et al, 1990; McInerney et al, 1992) However, unlike most PHA/PHB producers, *S. wolfei* produces PHB, during the exponential phase of growth under no apparent nutrient limitations or metabolic stress. (Amos & McInerney, 1989) Additionally, PHB was found to represent up to 20 percent of the *S. wolfei* cell by weight. (Beaty & McInerney, 1987) It is hypothesized that PHB synthesis in *S. wolfei* may act as an energy storage mechanism, allowing for continued metabolic function when environmental conditions make the beta-oxidation of VFAs unfavorable. (Amos & McInerney, 1989; Beaty & McInerney, 1989; McInerney et al, 1992; Sieber et al, 2010)

FBA was performed to identify the maximum PHB production rate when acetate is produced at a literature derived value as well as the theoretical maximum PHB production rate when acetate is not produced. The maximum PHB production rate when crotonate uptake and acetate production are set to the literature derived rates 0.9258 mmol crotonate/gDW-hr and 1.3889 mmol acetate/gDW-hr, respectively, was calculated to be 0.3 mmol PHB/gDW-hr, yielding a near zero growth rate (Figure 4.4). (Beaty & McInerney, 1987)

Applying the previous literature derived crotonate uptake and acetate production fluxes, iTK530 predicts that PHB is produced at 0.22 mmol PHB/gDW-hr with a growth rate of 0.012 h^{-1}, as suggested by literature (Figure 4.4). (Beaty & McInerney, 1987) The theoretical maximum PHB production rate from a crotonate uptake of 0.9258 mmol crotonate/gDW-hr without acetate production was 0.96 mmol PHB/gDW-hr (Figure 4.5). PHB production is increased when acetate production is

absent because more carbon is now available for the synthesis of PHB, which would have originally been secreted as acetate.

PHB is a carbon intensive compound that is closely linked to the uptake of the carbon sources utilized by *S. wolfei*, including butyrate and crotonate. (Sieber et al, 2010) As detailed by Figure 4.6, PHB can be efficiently synthesized using either butyrate or crotonate as a carbon source. Due to the carbon intensive nature of PHB, only minimal amounts of PHB can be produced while still achieving a satisfactory growth rate (Figures 4.4 and 4.5).

This is due to the fact that the carbon sources needs to flow through the enzymes associated with PHB production to enter the remaining metabolism of *S. wolfei* (Figure 4.6). Thus, an increase in PHB production would clearly siphon off the carbon entering the remaining metabolism, significantly impacting growth rate and resource availability. Therefore, to increase the PHB production by *S. wolfei*, the rate of carbon uptake would need to be dramatically increased in order to yield a growth rate near literature values. (Beaty & McInerney, 1987; Beaty & McInerney, 1989; Amos &McInerney, 1990; Sieber et al, 2010)

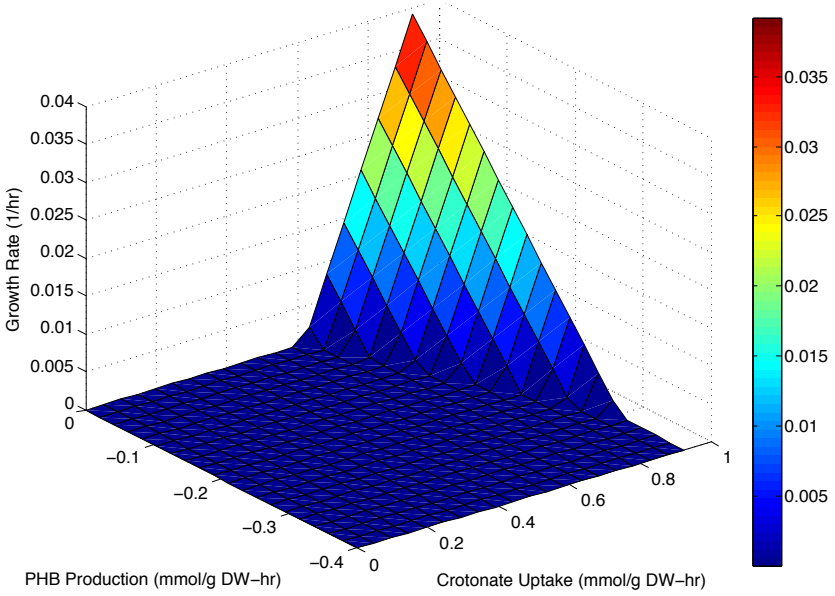

Figure 4.4: PHB production coupled with acetate production in *S. wolfei*

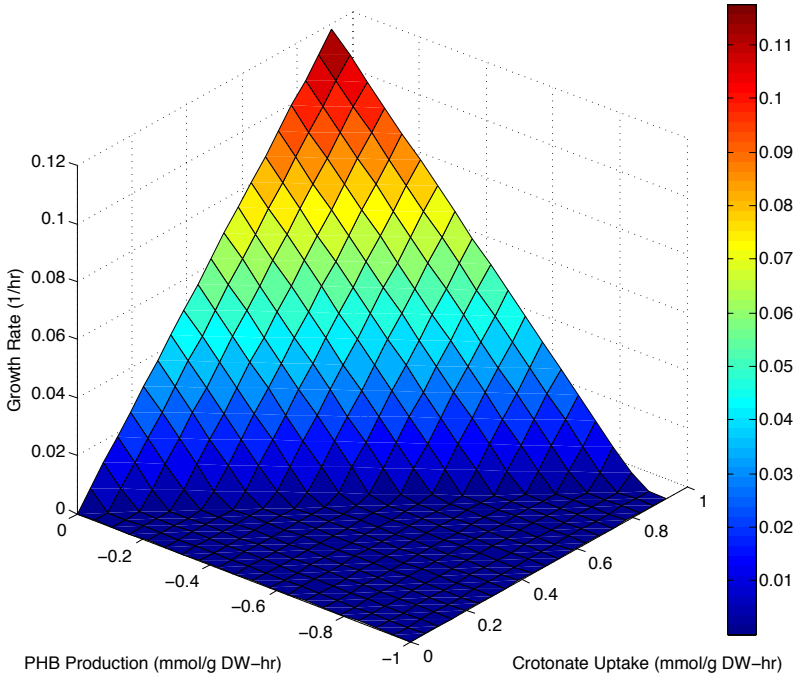

Figure 4.5: Theoretical maximum PHB production without acetate production in *S. wolfei*

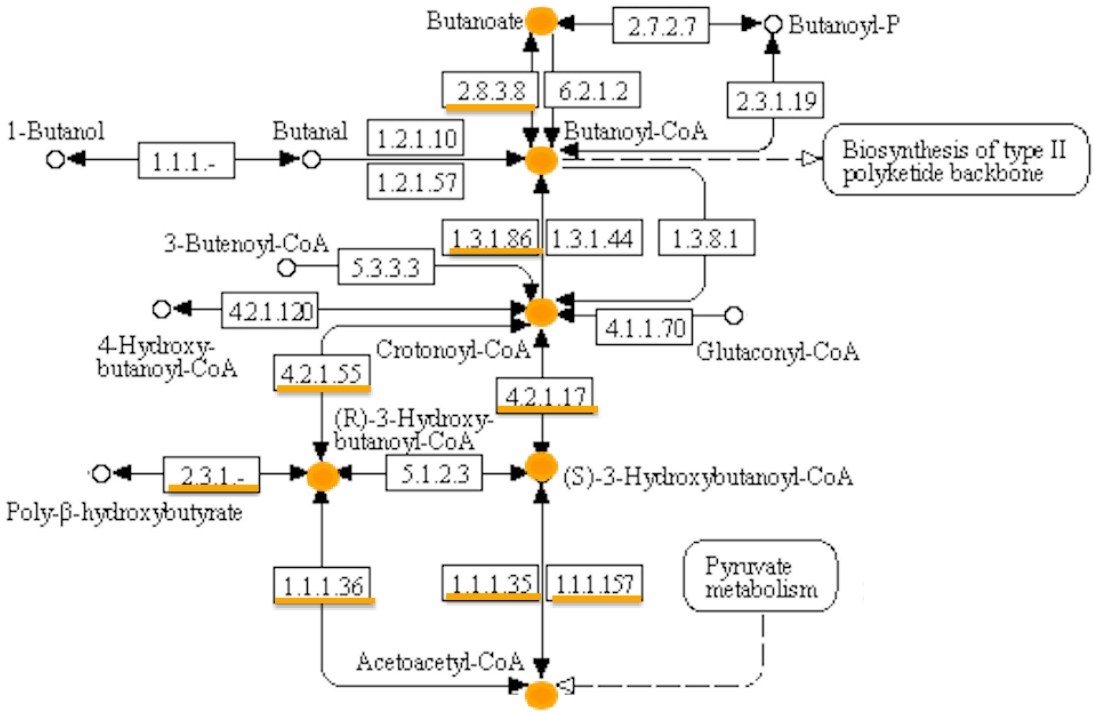

Figure 4.6: Simplified metabolic map of butyrate (butanoate) or crotonate (activated to crotonoyl-CoA during transport) to PHB in *S. wolfei*. Underlined enzymes are active in this process, according to Model SEED, KEGG, and Sieber et al, 2010.

4.6 Maximum H_2 Production

Genomic analysis of *S. wolfei* indicates that *S. wolfei* has three hydrogenases, with one hydrogenase being externally oriented. (Sieber et al, 2010) Hydrogen production in *S. wolfei* is linked to the reoxidation of reduced ferredoxin. (Sieber et al, 2010) While the generation of hydrogen, and subsequent removal by a syntrophic partner, remains critical to the metabolism of *S. wolfei*, little is known about the exact mechanisms of hydrogen production in *S. wolfei.* Subsequently, the hydrogen generation flux is unavailable for model validation.

The theoretical maximum hydrogen production associated with the literature defined rates of crotonate uptake and acetate production, set to 0.9258 mmol crotonate/gDW-hr and 1.3889 mmol acetate/gDW-hr, respectively, is approximately 11 mmol hydrogen/gDW-hr (Figure 4.7). However, when applying the previous literature derived crotonate uptake and acetate production rates, and limiting the growth rate to 0.012 h^{-1}, the hydrogen production rate predicted by iTK530 reduced to 2.1 mmol hydrogen/gDW-hr (Figure 4.7). (Beaty & McInerney, 1987) Hydrogen production may be increased when acetate production is absent due to the availability of greater energy resources. When acetate production is present, *S. wolfei* may produce hydrogen at a lower rate because a greater amount of the NAD^+ and reduced ferredoxin associated with hydrogen production is utilized to produce acetate. However, when acetate production is absent, a greater amount of these compounds required for hydrogen production may be available.

Interestingly, growth production plateaus after the hydrogen production rate reaches 11 mmol hydrogen/gDW-hr even though the hydrogen production continues to increase. Similar trends are observed when acetate is not produced by *S. wolfei* with the maximum theoretical hydrogen production, where the biomass growth rate plateaus at a hydrogen production rate of 37 mmol hydrogen/gDW-hr. (Figure 4.8) The steady hydrogen production rates are likely due to the hydrogen cycling throughout the model due to unknown reaction constraints. Additionally, the growth rates do not decline as hydrogen production is allowed to increase. This suggests that the hydrogen production reactions are not reducing the overall energy available to the cell, and instead endless looping throughout the metabolism is occurring. By leveraging future experimental data, such as the detailed measurement of hydrogen production rates by *S. wolfei*, the hydrogen production bounds could be constrained to more realistic bounds, resulting in a more accurate GEM.

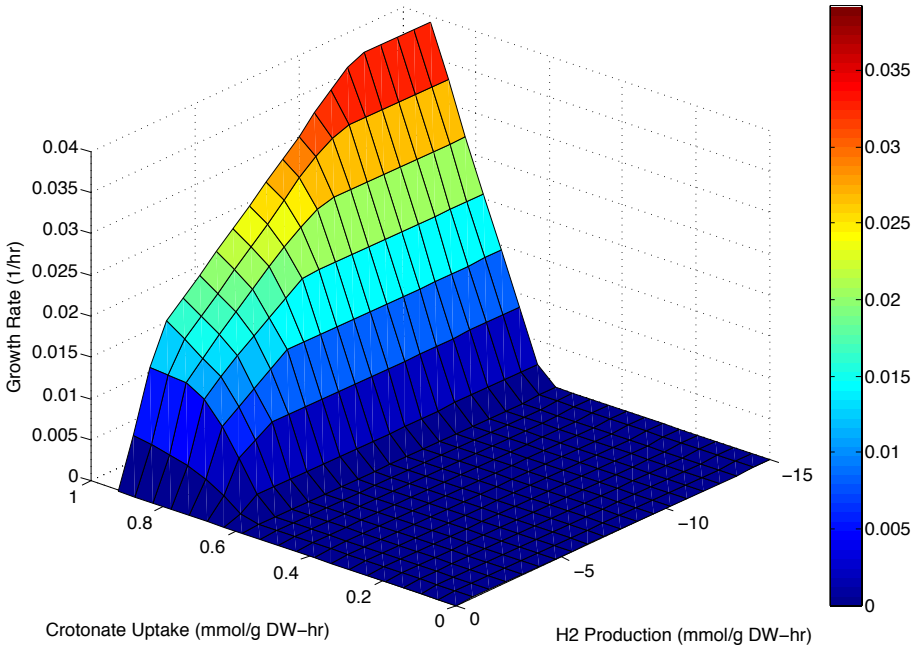

Figure 4.7: Maximum hydrogen production coupled with acetate production in *S wolfei*

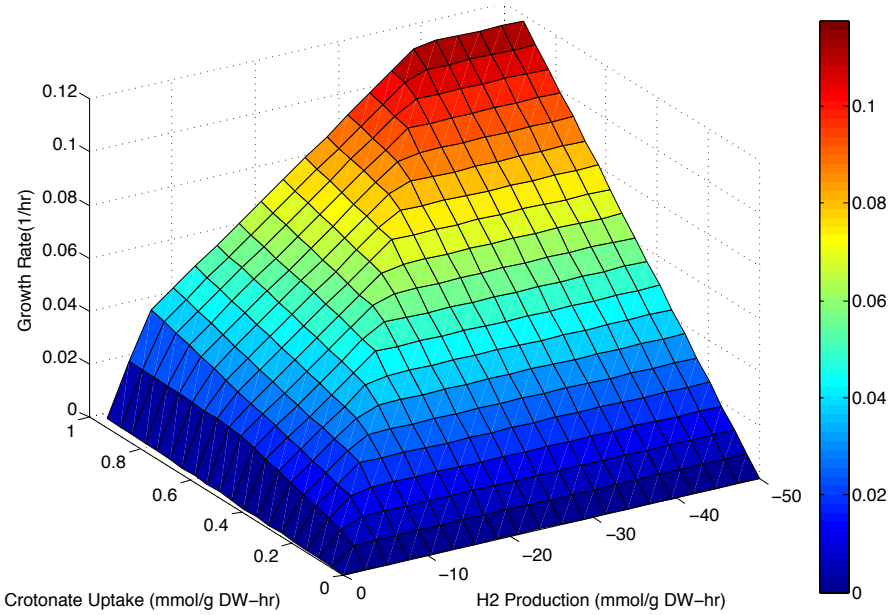

Figure 4.8: Theoretical maximum production of hydrogen without acetate production in *S. wolfei*

4.7 Uptake of Select Growth Media Compounds

S. wolfei grows on a defined medium with five B vitamins as required growth cofactors (Table 3.2). (Beaty & McInerney, 1990) While the five B vitamins required by *S. wolfei* for growth in pure culture have been studied, little research has been conducted to examine the requirement of the other compounds included in the trace metal and mineral solutions. Two compounds supplied in the defined media, ammonium and vitamin B12, were selected for analysis.

Although no nitrogen sources were explicitly specified in literature, ammonium was selected as it represents the most likely dominant nitrogen source. *S. wolfei*, commonly found in anaerobic digesters, would likely see large amounts of ammonium in the wastewater treatment system. Vitamin B12 was selected due to its requirement for growth by *S. wolfei* in addition to its significance in the synthesis of the vitamin B12 coenzyme.

The maximum ammonium uptake by *S. wolfei* when 0.9258 mmol crotonate/gDW-hr was supplied and 1.3889 mmol of acetate/gDW-hr was produced, as derived from literature, was 0.8 mmol ammonium/gDW-hr (Figure 4.9). Applying the previous literature derived crotonate uptake and acetate production rates that produced a growth rate of 0.012 hr^{-1}, iTK530 predicts an ammonium uptake rate of 0.3 mmol ammonium/gDW-hr (Figure 4.9). These uptake rates indicate that *S. wolfei* utilizes a minimal amount of nitrogen to meet its optimum growth conditions. However, it was also observed that no growth occurs when ammonium is not supplied, indicating that ammonium, although minimally utilized, is indeed a critical source of nitrogen for *S. wolfei.*

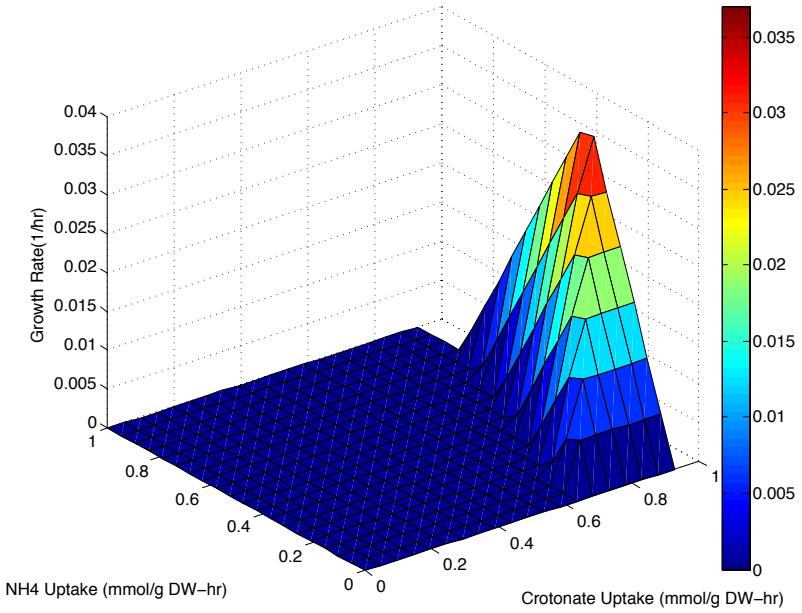

Figure 4.9: Ammonium uptake by *S. wolfei* when acetate is produced at a rate of 1.3889 mmol acetate/gDW-hr

Of the five B vitamins required for growth by *S. wolfei,* vitamin B12 was selected for analysis, as it is required for the synthesis of the vitamin B12 coenzyme. (Kanehisa et al, 2014; Kanehisa & Goto, 2000) It was observed that minimal amounts of vitamin B12 were required for growth when crotonate was supplied at 0.9258 mmol crotonate/gDW-hr and acetate was supplied and 1.3889 mmol of acetate/gDW-hr (Figure 4.10). Applying the previously detailed crotonate uptake and acetate production rates that produced a growth rate of 0.012 h^{-1}, a vitamin B12 uptake rate of 0.27 x 10^{-5} mmol vitamin B12/gDW-hr was predicted (Figure 4.10). Under the same conditions, iTK530 predicts a maximum vitamin B12 uptake rate of 0.8 x 10^{-5} mmol B12/gDW-hr, yielding a near zero growth rate. Because vitamin B12 is directly correlated with biomass growth, uptake above the amount required for growth was infeasible. Figure 4.10 also confirms that vitamin B12 is required for growth, and that the inhibition of vitamin B12 uptake will directly decrease *S. wolfei's* growth rate.

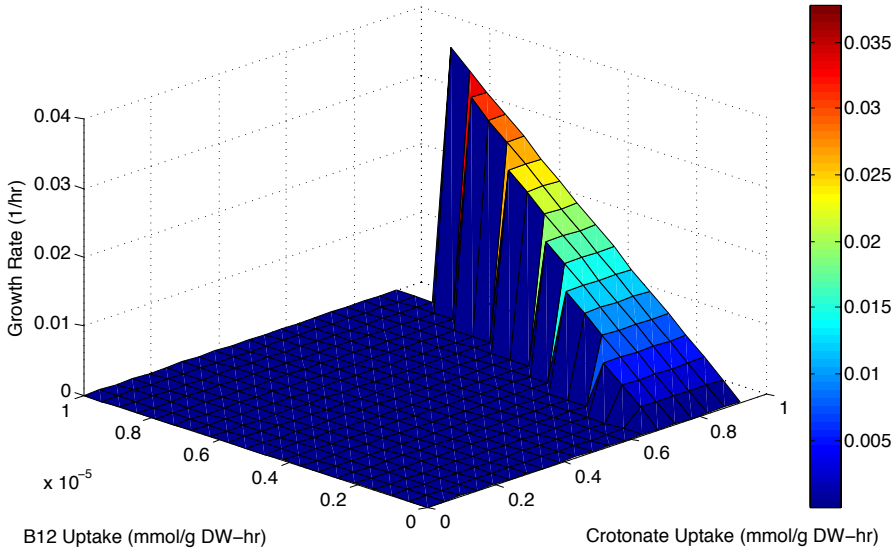

Figure 4.10: Vitamin B12 uptake by *S. wolfei* when acetate is produced at a rate of 1.3889 mmol acetate/gDW-hr

4.8 *S. wolfei* GEM (iTK530) Conclusions

The *S. wolfei* GEM (iTK530) is capable of predicting the specific growth rate of S. *wolfei* through the constraint of the carbon source and acetate production reactions. Due to the minimal amount of available literature data regarding the growth of *S. wolfei,* additional constraints are needed to improve model function. Formate and hydrogen production rates are critically needed, as these compounds have been shown to provide a significant function to the metabolism of *S. wolfei*. (Sieber et al, 2010) Additionally, as formate represents a source of carbon leaving the cell, formate secretion rates are required to accurately perform a carbon balance. (Beaty & McInerney, 1987)

While iTK530 closely matches available literature growth rates for *S. wolfei* grown in pure culture on crotonate, the product secretion rates needed to be reduced by 30 percent to meet these literature expected values. (Beaty & McInerney, 1987) The accuracy of the product secretion rates can be improved through better data on

the *S. wolfei* biomass composition reaction and product formation rates. The current biomass composition reaction for *S. wolfei* was constructed using the *E. coli* biomass composition reaction as the scaffolding, due to the incorrect assignment of cell wall type by Model SEED.

In addition to the in-depth analysis of the composition of the *S. wolfei* cell, better data is needed to better understand the maintenance energy requirements of *S. wolfei*. One of the most critical aspects of the biomass composition reaction is the maintenance energy required to produce biomass. The maintenance requirements used in iTK530 were taken from the *E. coli* GEM, and do not fully represent the maintenance requirements of *S. wolfei*. Through additional experiments, the biomass composition reaction can be refined to further improve the accuracy and validity of iTK530.

S. wolfei is a model organism for syntrophic organisms, and represents a critically understudied aspect of the ecology and function of anaerobic digestion. iTK530 provides a high-resolution *in silico* platform for the study of *S. wolfei,* and through the application of FBA (and other constrain based analysis techniques not used in this study), the metabolism of *S. wolfei* can be efficiently simulated and analyzed. While additional experiments are required to further improve iTK530, iTK530 presents a critical step forward in the elucidation of the genome, activity, and capabilities of *S. wolfei.*

CHAPTER 5
PURE CULTURE ANAEROBIC DIGESTER GEM RESULTS AND DISCUSSION

5.1 Pure Culture Anaerobic Digester GEM

While current *in silico* anaerobic digestion models approach the system at a macroscopic level, genome-scale modeling allows for the in-depth analysis of a microorganism, down to the genomic level. Recently, OptCom was developed to link multiple GEMs together to allow for the detailed analysis of a microbial community. (Zomorrodi & Maranas, 2012) Applying OptCom and descriptive-OptCom to the constructed pure culture anaerobic digester, presented in this work, provided a high-resolution *in silico* platform to study the complex processes of anaerobic digestion.

The pure culture anaerobic digester GEM was first analyzed using OptCom, a multi-level framework that allows for the optimization of both the community and individual species biomass growth rate. To examine the impact of the variation of the percent methanogens found in an anaerobic digester, the percent of methanogens was manually varied using both OptCom and descriptive-OptCom. Applying descriptive-OptCom allows for the optimization of the community and individual species biomass growth rate, but additionally allows the individual species growth rate to rise above or fall below the originally predicted optimum value. By allowing each species to rise above or fall below the previously determined optimum value, the tradeoffs and syntrophic relationships that occur between species can be examined.

5.2 Pure Culture Anaerobic Digester GEM Validation with OptCom

The pure culture anaerobic digester GEM was first examined using OptCom. Through the application of OptCom, the model predicted a total community biomass

growth rate of 0.457 h^{-1} from a glucose feed rate of 10 mmol glucose /gDW-hr (Figure 5.1). The glucose feed rate was constrained to a maximum uptake rate of 10 mmol glucose/gDW-hr as determined by experimental results and *in silico* validation. (Dash et al, 2014) The growth rates for *C. acetobutylicum*, *S. wolfei*, and *M. barkeri* were calculated to be 21 percent, 110 percent, and 4 percent greater than their literature measured growth rates of 0.32, 0.020, and 0.025 h^{-1}, respectively. (Dash et al, 2014; Beaty & McInerney, 1987; Gonnerman et al, 2013) The calculated growth rates for each organism were likely elevated from the predicted literature growth rates as OptCom does not allow for the organisms to rise above or fall below their optimum growth rates.

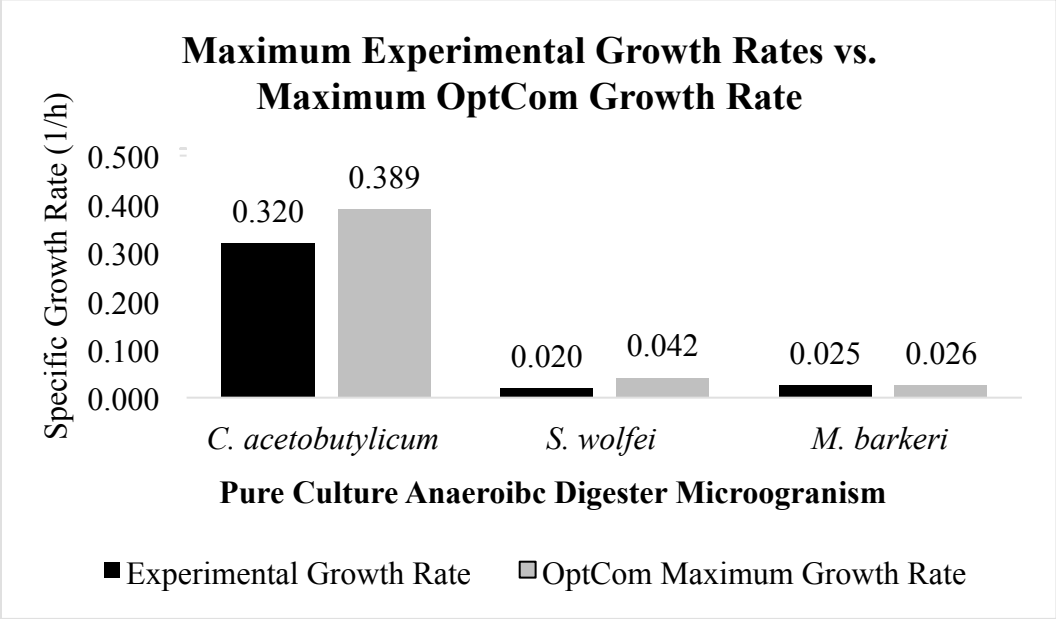

Figure 5.1: Validation of OptCom model results

Methane production of the pure culture anaerobic digester was also analyzed using OptCom (Figure 5.2). Applying the Buswell & Mueller, 1952 equation, as modified by Richards et al, 1991, yielded a maximum theoretical methane yield of 3 mmol methane per mmol glucose. Previous batch anaerobic digestion of glucose experiments yielded an experimental methane production yield of 2.3 mmol methane

per mmol glucose. (Kalyuzhnyi & Davlyatshina, 1997a) A lower literature limit was selected from the analysis of the anaerobic digestion that occurs in the rumen, with an expected methane yield of 0.56 mmol methane per mmol glucose. (Czerkawski, 2013) OptCom calculated a methane production rate of 0.045 mmol methane per mmol glucose, which fell well below literature determined range.

The minimal methane production yield was likely due to the relatively low percentage of methanogens (5.7 percent) calculated by OptCom to be in the pure culture anaerobic digester. This calculated percentage of methanogens was within the literature determined values of 2 to 30 percent, but was on the lower end of the spectrum. (Reyes et al, 2015)

One possible explanation for the low methanogen percentage may be attributed to the methanogen selected for the model. The methanogen used in this study, *M. barkeri*, was chosen, in part, due to its capability of growth on a wide range of substrates including acetate, hydrogen, methanol, and methylamines. (Balch et al, 1979) Although *M. barkeri* has a wide range of compatible substrates, specialized methanogens that grow on a sole substrate, such as acetate and hydrogen, have been shown to have a specific growth rate up to 0.163 h^{-1}, which is 552 percent increase from *M. barkeri's* specific growth rate. (Kalyuzhnyi, 1997b). Additionally, OptCom does not allow for interactions between organisms in the community, the application of OptCom may not yield literature supported results. (Zomorrodi & Maranas, 2012)

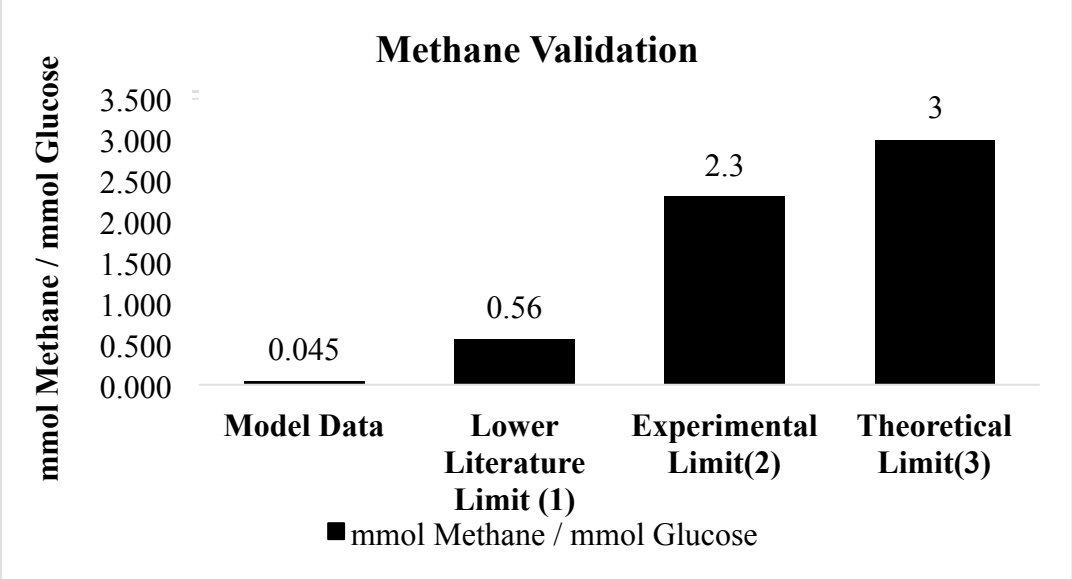

Figure 5.2: Analysis of methane production rates

[1](Czerkawski, 2013); [2](Kalyuzhnyi & Davlyatshina, 1997a); [3](Richards et al, 1991)

5.3 Varying Methanogenic Population with OptCom

The total percent methanogens was varied to examine the impact of the methanogenic community on the growth and metabolic byproduct production rates of the microorganisms in the pure culture anaerobic digester (Figure 4.13). Through the variation of the ratio of total methanogens and holding the glucose feed rate at 10 mmol glucose/g DW-hr, the total community biomass increased by 5.1 percent from 0.435 h^{-1} to 0.457 h^{-1} at 1 percent and 5.7 percent methanogens, respectively. (Figure 5.3) OptCom was unable to determine a solution outside of these narrow bounds, as it does not allow the species to rise above or fall below the optimal growth rate calculated by OptCom.

The increase in the methanogenic growth rate as *C. acetobutylicum's* and *S. wolfei's* growth rate stayed constant suggests that methanogenic growth was not limited by acetate and hydrogen production by the community. This was confirmed through the analysis of the flux distribution calculated by OptCom, which showed an

excess of acetate, butyrate, formate, hydrogen, and carbon dioxide in the system. While growth was not limited by acetate uptake at such low methanogen percentages, acetate uptake was nearing its literature validated limit of 6.4 mmol acetate /gDW-hr (Figure 5.11). (Gonnerman et al, 2013)

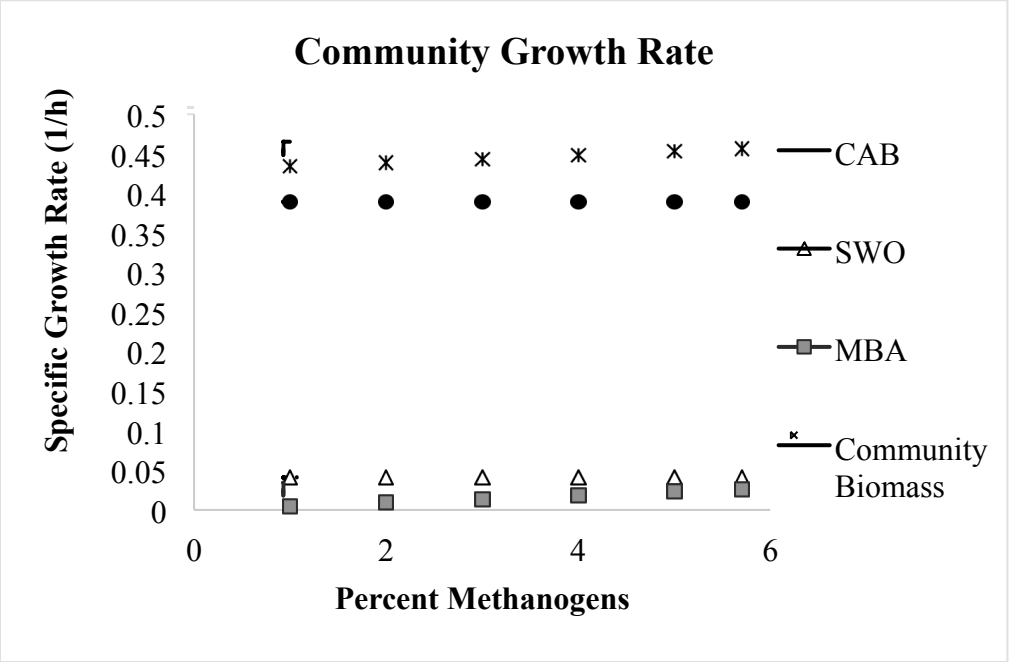

Figure 5.3: Community and species growth rate with methanogenic population variation. CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

The methane production yield as a function of the methanogen percentage was also analyzed through OptCom (Figure 5.4). The methane production yield as, mmol of methane per mmol of glucose, remained far below literature values (0.56 – 3 mmol methane / mmol glucose). At only 5.7 percent methanogens, the methane yield topped out at 0.045 mmol methane per mmol glucose. (Figure 5.4) Aside from the low percentage of total methanogens within the community, the low methane yield may be a result of the low literature validated acetate uptake limit, as up to 70 percent of all methane produced is through acetoclastic pathways. (Gonnerman et al, 2013; Avery et al, 2003)

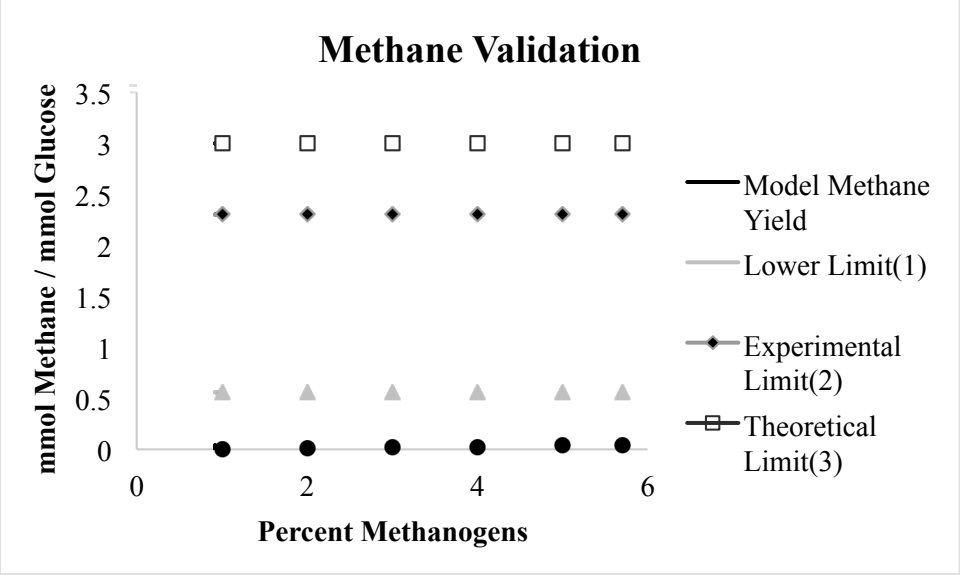

Figure 5.4: Validation of mmol methane produced per mmol of glucose supplied. [1](Czerkawski, 2013); [2](Kalyuzhnyi & Davlyatshina, 1997a); [3](Richards et al, 1991)

5.4 Varying Methanogenic Population with Descriptive-OptCom

Throughout nature, microorganisms often operate under less than optimal conditions to benefit the community, and may form syntrophic partnerships with surrounding microorganisms. Through descriptive-OptCom, the growth rate of each individual organism is allowed to deviate from the optimum value in order to maximize the community biomass growth.

To analyze the community dynamics of the pure culture anaerobic digester, descriptive-OptCom was applied with the optimum growth rate for each organism set at the literature derived rates of 0.32 h^{-1}, 0.012 h^{-1}, and 0.025 h^{-1} for *C. acetobutylicum*, *S. wolfei*, and *M. barkeri*, respectively. The glucose uptake limit was set at 10 mmol glucose/gDW-hr, as specified by Dash et al, 2014, and the acetate uptake rate was doubled from 6.4 mmol acetate/gDW-hr to 12.8 mmol acetate/gDW-hr. The doubling of the acetate uptake rate was necessary due to the failure of the model at 10% total methanogens and an acetate uptake rate of 6.4 mmol

acetate/g DW-hr. The model failed, in part, due to the methanogen reaching its acetate uptake limit, stagnating its growth. As a response, the model did not allow for the growth of *S. wolfei* in order for the methanogen to meet the percent methanogen limit specified in the model. To work around these limits, the acetate uptake rate was doubled from the original literature validated rate due to *M. barkeri* quickly approaching the acetate uptake limit (Figure 5.3).

5.5 Growth Dynamics with Descriptive-OptCom

The growth dynamics of the pure culture anaerobic digester indicate that both *S. wolfei* and *M. barkeri* operated at an increased specific growth rate when compared to their literature values when the percent methanogens were greater than 15 percent (Figures 5.5 and 5.6). In contrast, *C. acetobutylicum* operated under its optimal literature specific growth rate when the percent methanogens was greater than 15 percent (Figures 5.5 and 5.6).

The greatest community biomass occurred at 10 percent methanogens, which was followed by a plateau of methanogenic growth, and an overall decrease in the community biomass (Figure 5.5). This drop in community level biomass growth rate is due to *M. barkeri* reaching its acetate uptake limit, stagnating its growth. As the model is operated at increased percent methanogens, the growth of *C. acetobutylicum* must be decreased to meet this condition, as *M. barkeri* growth can no longer increase to meet the designated percent methanogen parameter of the model. This can be more clearly seen through the examination of each species' optimally levels (Figure 5.6). At a percent methanogen value less than 5 percent methanogens, *M. barkeri* is operating at a reduced activity level while *C. acetobutylicum* is operating an increased activity level. From 5 percent methanogens to 15 percent methanogens, *M. barkeri* increases its optimality level until its acetate uptake constrain is reached and growth stagnates. As the percent methanogens increases past 15 percent, *C. acetobutylicum* must decreased is activity level in order for the methanogen to meet the percent

methanogen parameter defined in the model. In other words, since the methanogen growth rate cannot increase further, the only way for their relative abundance to increase is for *C. acetobutylicum* to decrease.

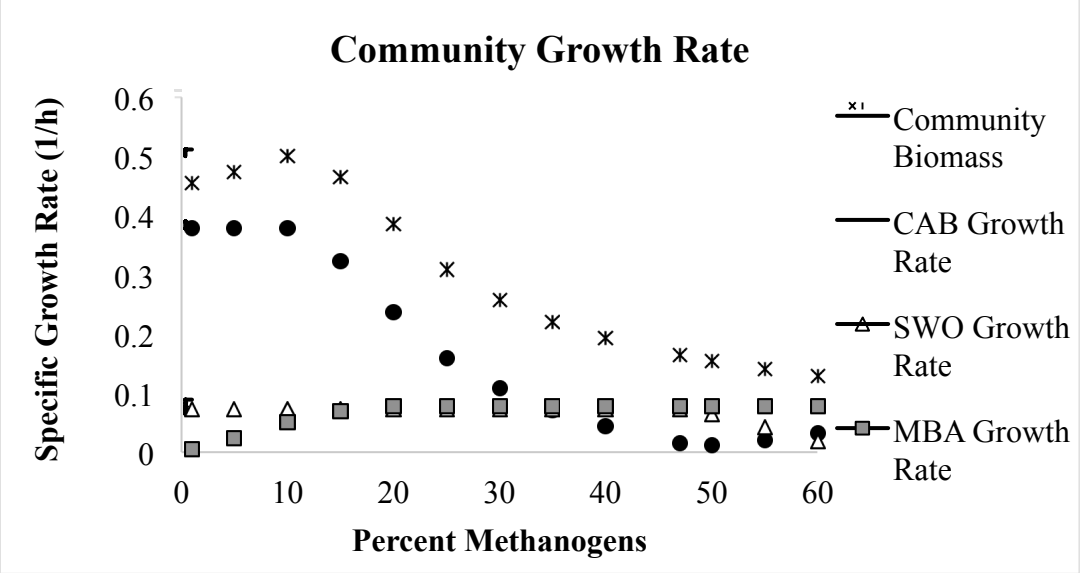

Figure 5.5: Community growth dynamics as a function of percent methanogens CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

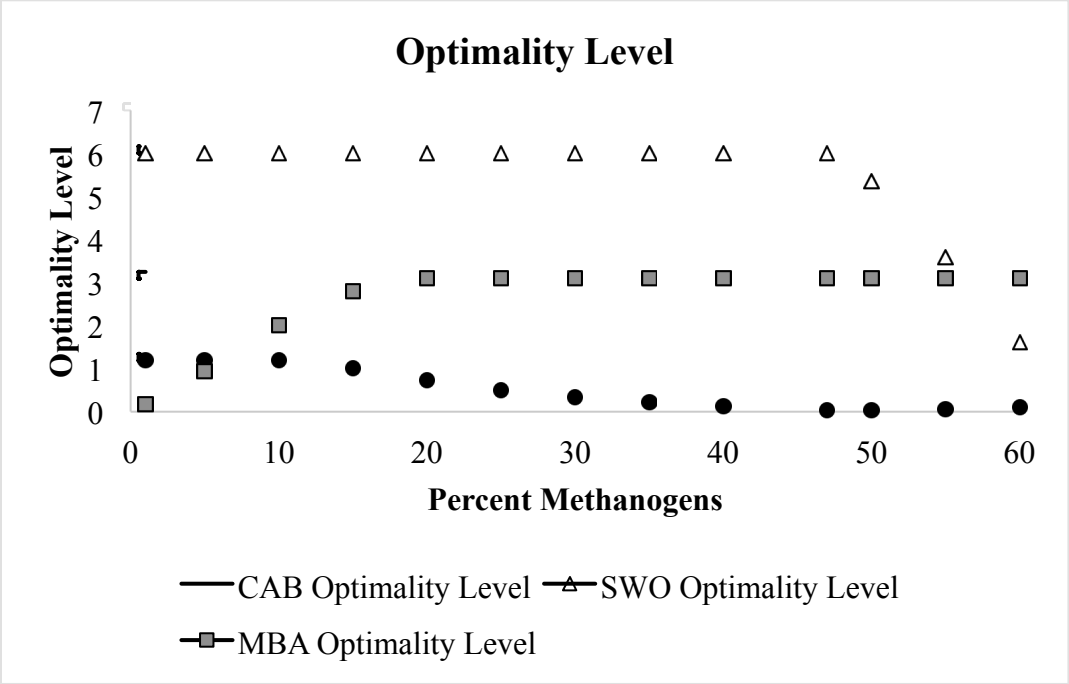

Figure 5.6: Optimality level of each organism versus the total percent methanogens CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

5.5.1 Rate Limiting Step Identification with Descriptive-OptCom

Applying descriptive-OptCom to the pure culture anaerobic digester GEM indicated that digester performance is limited by the activity of the methanogens, specifically their acetate uptake rate. The *M. barkeri* growth rate quickly reached a plateau due to the acetate uptake rate approaching its limit of 12.8 mmol acetate/gDW-hr, even though this was twice the literature reported value of 6.4 mmol acetate/gDW-hr (Figure 5.5). Throughout the model, the methanogen was identified as the key rate limiting step due to its elevated optimality level and acetate uptake limitations. (Figures 5.6 and 5.11)

While *S. wolfei* is also operating at an increased optimality level, it's metabolic activity is not considered rate limiting as it is metabolizing very low amounts of butyrate (0.7244 mmol butyrate/gDW-hr) and producing minimal

amounts of acetate (1.24 mmol acetate/gDW-hr). However, *S. wolfei's* increased optimality level does indicate that the organism is performing at many times its expected activity to support overall digester growth. *S. wolfei* likely remains at an elevated optimality level throughout the model due to minimal amount of supporting experimental data available. Because minimal supporting experimental data is available, the growth, metabolite uptake, product secretion, and maintenance energy requirement rates may not be properly validated.

As the methanogenic population starts to increase, the growth of the acidogen, *C. acetobutylicum*, begins to decrease. This decrease in *C. acetobutylicum* growth can be attributed to the increased acetate production needed to support the methanogenic growth at higher percentages, which leaves little remaining resources for *C. acetobutylicum's* biomass growth. The continued decrease in growth from *C. acetobutylicum* after 30 percent of total growth is methanogenic may also be attributed to the constructed model forcing a lower acidogen growth to meet the required percent methanogens in the system. This reasoning can also be applied to the decrease in *S. wolfei* growth once methanogens reach 50 percent of the digester community population. Due to the inability for *M. barkeri* to continue increasing its growth rate as a result of its acetate uptake limitation, the growth rates of the other organisms, *C. acetobutylicum* and *S. wolfei*, must decrease to allow the total percent methanogens to increase.

5.5.2 Methane Production Validation with Descriptive-OptCom

Through the application of descriptive-OptCom, the yields of methane produced per mmol of glucose supplied can be analyzed. As indicated by Figure 5.4 and Figure 5.7, methane production remained well below literature supported values when the methanogens accounted for less than 20 percent of the microbial community. As the percent of methanogens in the digester increases from 20 percent to 40 percent, the methane production per mmol of glucose falls within literature

derived limits of 0.56 to 3 mmol methane/ mmol glucose (Figure 5.7). This closely aligns with real world anaerobic digesters, which typically have a distribution of methanogens ranging from 2 to 30 percent of all microorganisms. (Wirth et al, 2012; Jaenicke, 2011; Reyes et al, 2015) While the biomass growth rate for *M. barkeri* and *S. woflei* both deviated above their literature derived growth rates, *C. acetobutylicum* fell within its expected literature growth range. The elevated *M. barkeri* and *S. wolfei* growth rates can be accounted for due to *M. barkeri's* increased acetate uptake rate and the lack of literature data for the uptake of butyrate and production of acetate, formate, and hydrogen by *S. wolfei,* which made validation difficult to achieve.

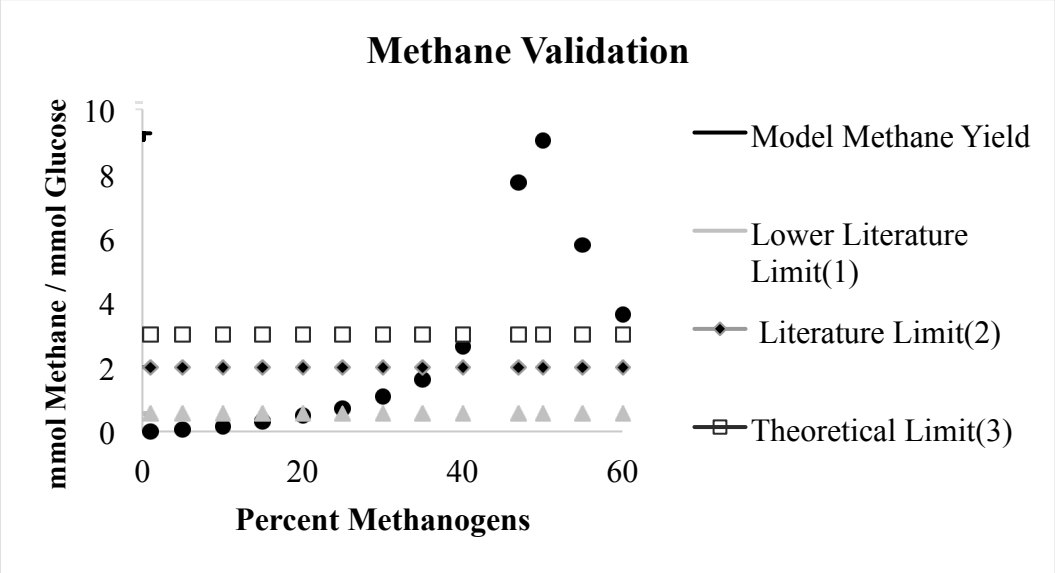

Figure 5.7: Methane yield as a function of percent methanogens at double the literature support acetate uptake rate

[1](Czerkawski, 2013); [2](Kalyuzhnyi & Davlyatshina, 1997a); [3](Richards et al, 1991)

5.6 Model Improvement Though Constraint of *S. wolfei*

The previous simulations were conducted using literature derived butyrate uptake and acetate production constraints applied to *S. wolfei.* (Beaty and McInerney,

1989) These constraints yielded an elevated *S. wolfei* growth rate as compared to the literature data (Figure 5.6). (Beaty and McInerney, 1987) These constraints likely yielded an elevated growth rate due to the unknown secretion rate of formate, which would serve as another major sink for carbon, further lowering the growth rate of *S. wolfei.* Applying the crotonate uptake and acetate secretion constraints established from growth on crotonate by Beaty and McInerney, 1987, *S. wolfei* growth aligned closely with published studies. Crotonate is a four carbon short chain acid that is activated to Crotonoyl-CoA, where it is integrated into the butyrate metabolism pathway. (Sieber et al, 2010) The application of these constraints was made based on the assumption that crotonate and butyrate would likely be metabolized at similar rates, as they enter the metabolism of *S. wolfei* in nearly identical positions (Figure 5.6). By applying the values associated with the crotonate uptake and acetate production rates from pure cultures of *S. wolfei* grown on crotonate to the model, the *S. wolfei* growth fell by 68.9 percent from 0.0721 h^{-1} to 0.0224 h^{-1}, which was more closely aligned to its literature expected growth rate of 0.012 h^{-1}. The *S. wolfei* growth rate decreased due to an increase in the overall net carbon secreted as acetate.

5.6.1 Growth Dynamics of Modified Model Through Descriptive-OptCom

Upon updating the uptake and production constraints of *S. wolfei*, growth dynamics and optimality levels were examined using descriptive-OptCom (Figures 5.8 and 5.9). While the maximum growth of the community biomass slightly decreased from 0.5 to 0.45 h^{-1} at 10 percent methanogens, the growth of each individual species included in the pure culture anaerobic digester GEM were closer to their literature expected values, with the only outlier being growth of *M. barkeri.* The 0.05 h^{-1} reduction of overall community biomass growth rate can be solely attributed to the decrease in growth by *S. wolfei*, which had a 0.05 h^{-1} decrease in growth due to the newly applied constraints (Figure 5.8).

M. barkeri had an elevated growth rate compared to its literature derived growth rate after 5 percent total methanogens due to the increased acetate uptake rate of 12.8 mmol acetate/gDW-hr, twice the literature value. Figure 5.8 and Figure 5.9 suggest that the activity of *M. barkeri* is the rate limiting step of the pure culture anaerobic digester, with S. wolfei contributing very little to the overall metabolism of the community. The lack of influence by *S. wolfei* on the community was determined through the analysis of the flux distribution calculated by OptCom. Due to the low literature derived limits of butyrate uptake and acetate production by *S. wolfei*, a maximum of 15 percent of all acetate consumed by *M. barkeri* is derived from the production of acetate by *S. wolfei*.

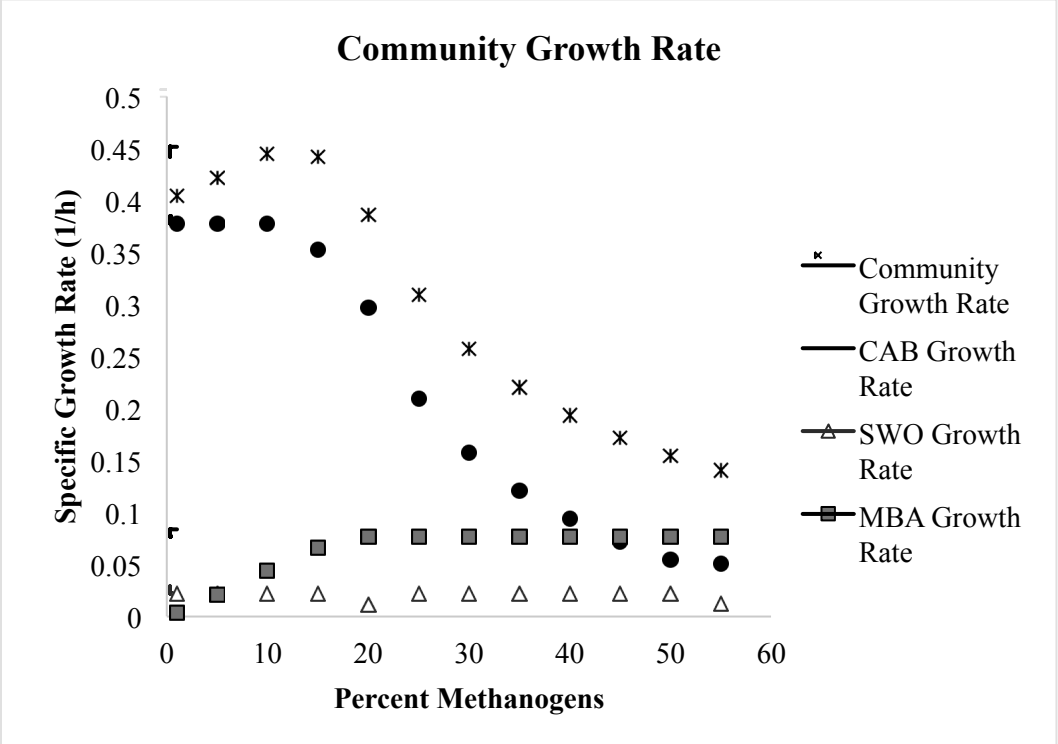

Figure 5.8: Community specific growth rate as a function of percent methanogens with updated *S. wolfei* metabolism constraints and doubled *M. barkeri* acetate uptake constraints

CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

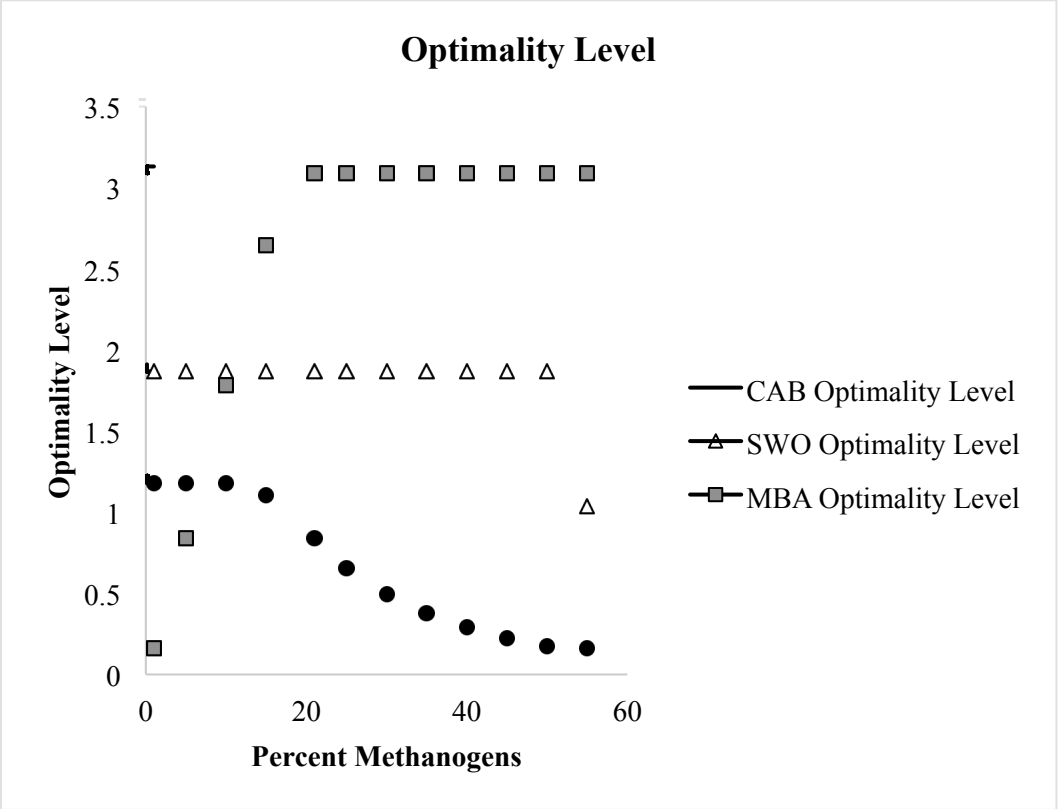

Figure 5.9: Optimality levels as a function of percent methanogens with updated *S. wolfei* metabolism constraints and doubled *M. barkeri* acetate uptake constraints
CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

5.6.2 Methane Validation of Modified Model Through Descriptive-OptCom

Similar to the previous simulations conducted, the mmol methane produced per mmol of glucose consume remained below literature data at percentages where the methanogens were less than 25 percent (Figure 5.10). The methane production yield fell in the range of literature rates from 25 to 60 percent methanogens. The methane production rate begins to plateau when methanogens comprise 50 percent of the digester community, as *M. barkeri's* acetate uptake limit is reached, even though it was set at twice the expected literature value. The fact that *M. barkeri* can reach its

new acetate uptake limit that was twice the literature limit reinforces the idea that *C. acetobutylicum's* metabolism dominates the pure culture anaerobic digester.

While literature data suggests that methanogens contribute between 2 and 30 percent of overall biomass in an anaerobic digester, a wider variety of methanogens are typically found in real world anaerobic digesters, which may increase the methane production rate of the digester community. (Reyes et al, 2015) Although *M. barkeri* is capable of metabolizing a wide range of single carbon substrates, many methanogenic specialist which degrade a single compound, such as acetate or hydrogen, also comprise a large percentage of the methanogenic population. (Reyes et al, 2015) Due to *M. barkeri's* inability to completely metabolize the available metabolic byproducts, the methane yields remains low. Through the addition of other acetate, carbon dioxide, hydrogen, and formate degrading methanogens, the methane yield should increase so that it correlates more closely with literature supported methane yields.

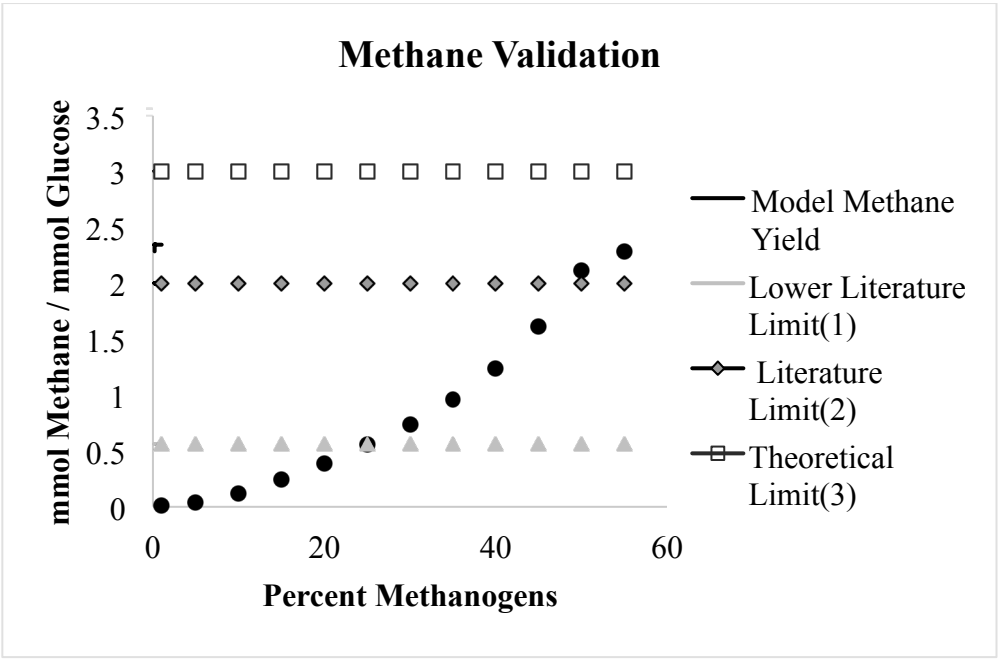

Figure 5.10: Methane yield versus percent methanogens with updated *S. wolfei* metabolism constraints and doubled *M. barkeri* acetate uptake constraints [1](Czerkawski, 2013); [2](Kalyuzhnyi & Davlyatshina, 1997a); [3](Richards et al, 1991)

5.7 Effect of Acetate Uptake Rate Through Descriptive-OptCom

Methanogens are a diverse group of *Archaea*, with many species that specialize in growth on single substrates, such as acetoclastic methanogens and hydrogenotrophic methanogens. While *M. barkeri* was selected due to its range of available substrates, it is unlikely that a single species of methanogens would accurately represent a set of sensitive and diverse microorganisms. While the acetate uptake of M. barkeri was originally set to the literature constraints of 6.4 mmol acetate/g DW-hr, the acetate uptake in a real world anaerobic digester would likely be increased due to the diverse consortia of microorganisms involved. (Gonnerman et al, 2013) As suggested in Figures 5.3, 5.5, 5.10, the acetate uptake limit remains the rate-limiting reaction of the pure-culture anaerobic digester, limiting the growth of *M. barkeri.* Although *M. barkeri* remained above its literature determined growth rate of 0.025 h^{-1} throughout each simulation, the methanogen acetate uptake rate was varied to examine the effect of removing the rate limiting step of the pure culture anaerobic digester GEM.

The methane production flux of *M. barkeri* increased as the acetate uptake rate increased due to the increased amount of carbon entering *M. barkeri* (Figure 5.11). At the original literature acetate uptake rate of 6.4 mmol acetate/g DW-hr, the methane production rate quickly plateaus as the methanogenic population increases before the model produces a zero growth rate. (Gonnerman et al, 2013) As the acetate rate is increased to 12.8 mmol acetate/g DW-hr, the methane flux increases before plateauing due to the defined acetate uptake limit. When the acetate uptake rate is unconstrained, the methane production flux drastically increases before beginning to plateau. The plateau when acetate uptake is unconstrained is likely due to internal constraints of required growth cofactors and maintenance energy requirements.

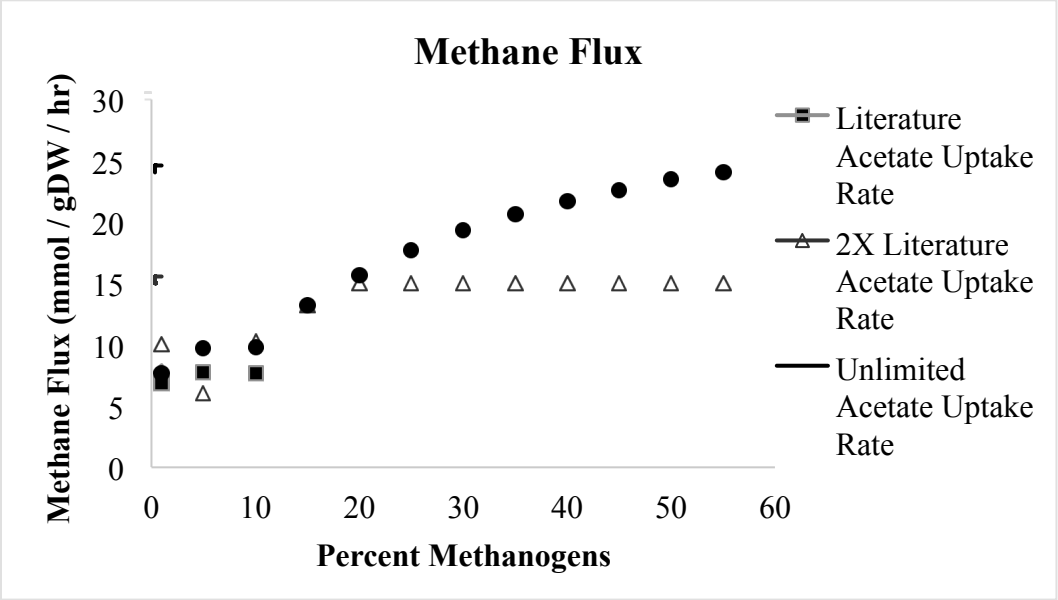

Figure 5.11: Methane flux as acetate uptake is increased

5.7.1 Population Dynamics With Acetate Unconstrained

Removing the limits of the acetate growth rate did not greatly impact overall biomass growth of the total community when compared to the total community growth when the acetate uptake limit was set to double the literature value (Figures 5.12 and 5.13). However, the removal of the acetate uptake limits did significantly impact the growth rate of *M. barkeri*. With the acetate uptake rate constrained at 12.8 mmol acetate /gDW-hr, the *M. barkeri* specific growth rate plateaued at 0.072 h^{-1} (Figure 5.8). But, when the acetate uptake flux was unconstrained, the *M. barkeri* specific growth rate increased to 0.146 h^{-1} before plateauing (Figure 5.12). The growth plateau with the acetate uptake rate unconstrained is a function of the internal constraints of the organism, likely related to the uptake of required growth co-factors and various limits places on the ATP maintenance energy requirements.

The optimality levels of each organism under unconstrained acetate uptake rate conditions (Figure 5.13) closely followed the optimality levels seen under an acetate uptake constraint of 12.8 mmol acetate/gDW-hr (Figure 5.9). The exception is

the increased optimality level of *M. barkeri* and higher total percent methanogens observed when the acetate rate is unconstrained. This follows closely with the growth data seen in Figure 5.12, as the increase in acetate uptake only truly benefits *M. barkeri.*

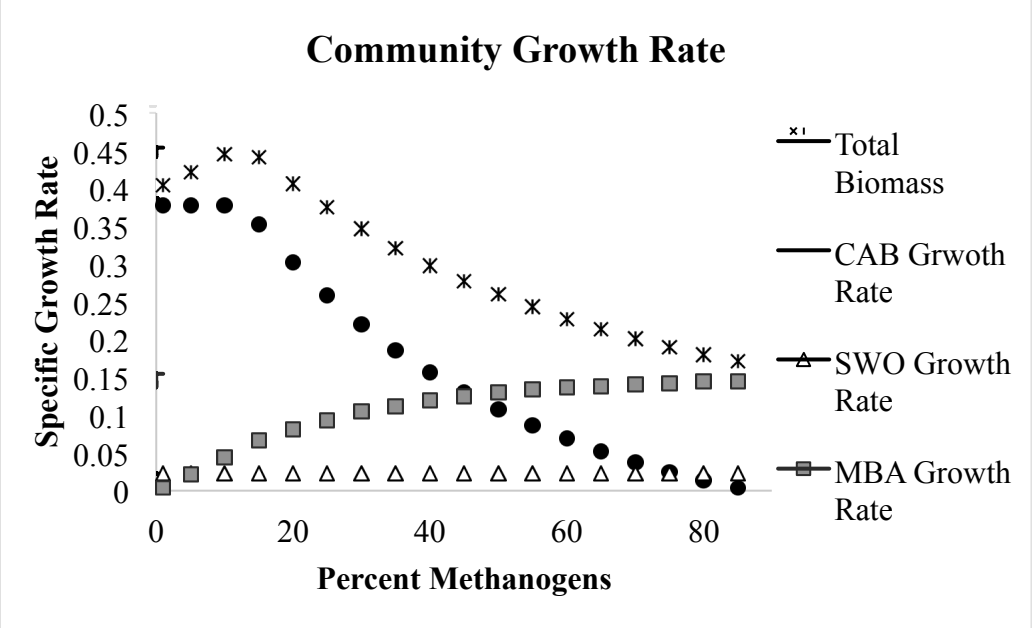

Figure 5.12: Community biomass growth with unconstrained acetate uptake as a function of percent methanogens

CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

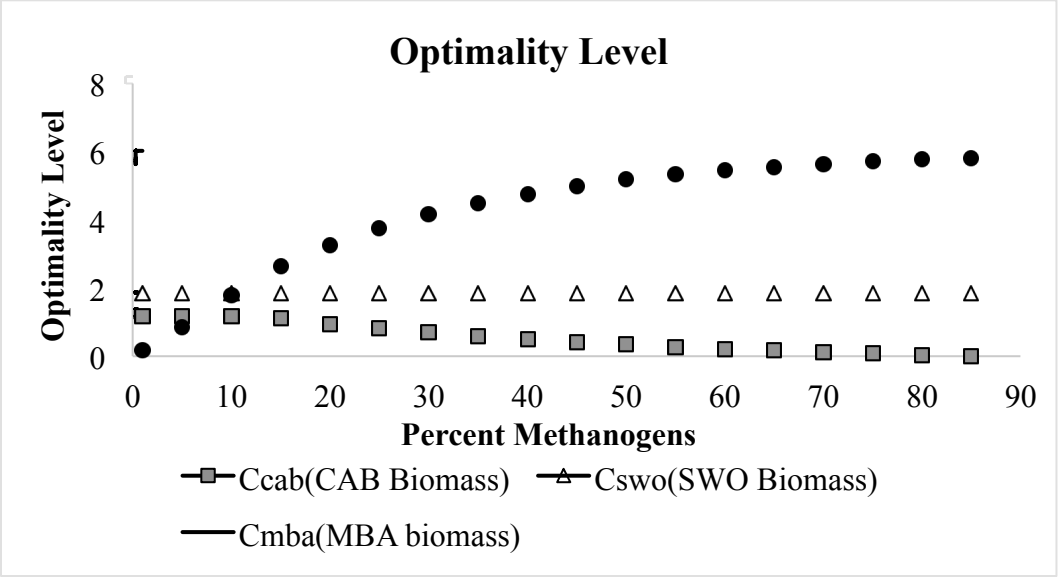

Figure 5.13: Optimality levels of a pure culture anaerobic GEM with acetate unconstrained as a function of the percent methanogens
CAB: *C. acetobutylicum*; SWO: *S. wolfei*; MBA; *M. barkeri*

5.7.2 Methane Validation of the Modified GEM

The methane production yield, modeled through descriptive-OptCom, closely followed the methane production when the acetate uptake rate was constrained to 12.8 mmol acetate/gDW-hr (Figures 5.14 and 5.10). While methane yields began to level off at 50 percent methanogens when the acetate uptake rate was constrained to 12.8 mmol acetate/gDW-hr, the methane yield continued to increase exponentially when the acetate uptake rate was unconstrained and did not reach a plateau. Additionally, the methane production yield also increased past the theoretical methane production yield of 3 mmol methane per mmol glucose when the acetate uptake was unconstrained. (Richards et al, 1991) This increase in the methane production yield beyond the theoretical methane yield is likely due to the acidogen producing large amounts of acetate at the expense of its growth. This resultins in an

increase in acetate uptake by the methanogen, which is ultimately removed from the system as methane as the methanogen's growth rate begins to level off.

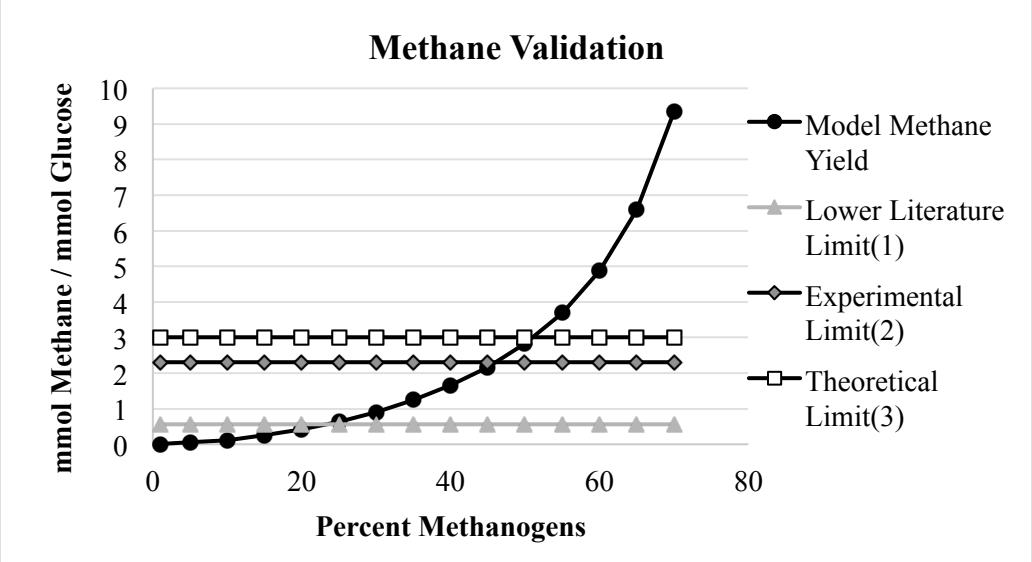

Figure 5.14: Methane yield with acetate unconstrained as a function of percent methanogens

[1](Czerkawski, 2013); [2](Kalyuzhnyi & Davlyatshina, 1997a); [3](Richards et al, 1991)

5.8 Comparison of the Pure Culture Anaerobic Digester GEM to a Real World Anaerobic Digester

Using data supplied by the Gresham Wastewater Treatment Plant (Gresham, Oregon), the mmol of methane per mmol glucose yield was calculated. Assuming a conversion factor of 1.42 mg COD per mg of volatile suspended solids, 1.07 mg COD per mg glucose, 10^{16} milliliters of cells per milliliter of anaerobic digestate, and that methane accounts for 70% of biogas formed, the methane yield was calculated to be 14.5 mmol methane per mmol glucose (Tables 5.1 and 5.2). (Grady et al, 2011) This methane yield falls above the batch scale digestion of glucose value of 2.3 and the theoretical maximum yield of 3 due to the incorporation of fats, oils, and greases (FOG) and other high strength wastes at the Gresham Wastewater Treatment Plant. (Kalyuzhnyi & Davlyatshina, 1997a; Richards et al, 1991) The Gresham Wastewater

Treatment Plant produced an elevated methane yield due to the high loading rates and volatile solids destruction rates they are able to achieve through the addition of FOG.

Theoretically, a typical real world digester produces 4.485 to 6.788 mmol methane per mmol glucose. (Metcalf & Eddy, 2003), This theoretical methane yield is likely greater than the methane yields observed during the anaerobic digestion of glucose in batch anaerobic digesters is due to the complexity and variety of the influent wastes. Primary and secondary sludge is a relatively complex waste, which is typically a mixture of carbohydrates, fats, and proteins which are reduced down to simple sugars, VFAs, and amino acids, respectively. An increase in the diversity of substrates allows for the wide array of microorganisms in an anaerobic digester to produce a greater amount of single carbon compounds used by the methanogens. Additionally, the methane yields were derived from the batch anaerobic digestion of glucose, as opposed to a continuous stirred tank reactor commonly using in real world anaerobic digesters, which may alter the activity of the microbial community. Batch scale digesters may not capture the activity of real world digesters because they are only loaded once at the beginning of the experiment, which may overload or underload the microbial community with a substrate, potentially inhibiting or starving the microorganisms. (Metcalf & Eddy, 2003)

Applying the real world anaerobic digestion glucose loading rates to the pure culture anaerobic digester GEM produced methane yields that were much less than the expected methane yields of real world anaerobic digesters (Figure 5.15). The disparity between the results may be a result of the effective glucose loading rate, which is orders of magnitude lower in real world digesters than the pure culture anaerobic digester GEM glucose loading rate of 10. The difference between the effective glucose loading rates may be derived from the difference of culturing organisms in pure culture versus an active full-scale anaerobic digester. Additionally, the conversion of the volatile solids loading rate of an anaerobic digester to glucose loading rate involves the assumption that all VS can be effectively converted to glucose.

Table 5.1: Glucose loading rates of real world anaerobic digesters.

Source	Metcalf & Eddy Design Parameters		Gresham, Oregon Anaerobic Digestion
	Lower limit	Upper Limit	
Glucose Uptake (mmol Glucose/gDW/hr)	2.12E-07	1.06E-06	7.97E-07

A methane yield curve was created by varying the glucose loading rate in descriptive-OptCom while holding the percent methanogens constant at 15 percent (Figure 5.15). 15 percent methanogens was selected as it in the middle of the reported range of methanogens. (Reyes et al, 2015) To linearize the data, the log was taken of both the methane yield and the glucose loading rate. A linear fit was applied a generate a line with a R squared value of 0.9802 (Figure 5.16). Utilizing the data from a real world anaerobic digester, the methane yields were estimated based on the calculated glucose loading rate of the anaerobic digester (Table 5.2).

According to Metcalf & Eddy, 2003, methane yields as, mmol methane per mmol glucose, are estimated to range 4.49 to 6.79. The fitted data yields a slight underestimate due to limitations within the model when operating at very low glucose loading rates. The pure culture anaerobic digester GEM is comprised of three GEMs, which all have unique constraints. Each model has various internal constraints related to ATP yield or growth co-factor uptake. Decreasing the glucose loading rate far below a typical value that each cell would see when grown in pure culture during exponential growth, results in an infeasible solution as calculated by the model. As a result, the data at low glucose loading rates is sparse which leads to errors in the fitted model yields and an underestimated methane yield.

Table 5.2: Calculated glucose loading rates and predicted methane yields

Glucose Loading Rate (mmol/gDW-hr)	Model Predicted Methane Yield (mmol Methane / mmol Glucose)	Actual Methane Yield (mmol Methane / mmol Glucose)
0.000000212[a]	4.22	6.79
0.000000797[b]	3.87	14.5
0.00000106[a]	3.79	4.49

[a] Lower and upper limits of anaerobic digestion loading (Metcalf & Eddy, 2003)

[b] Data from an anaerobic co-digester, digesting high strength wastes.

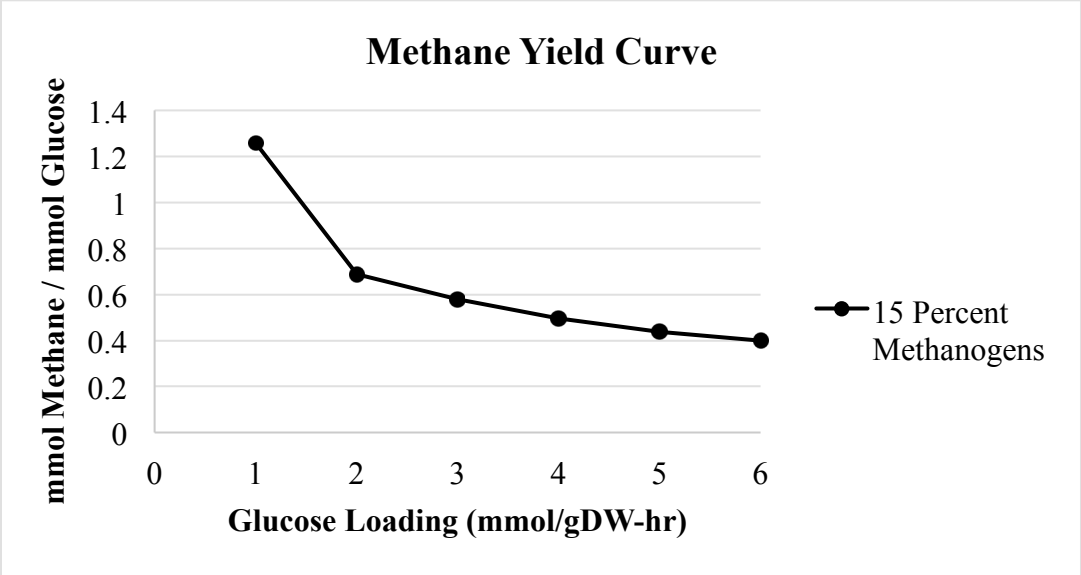

Figure 5.15: Methane yield as a function of the glucose loading rate at 15 percent methanogens

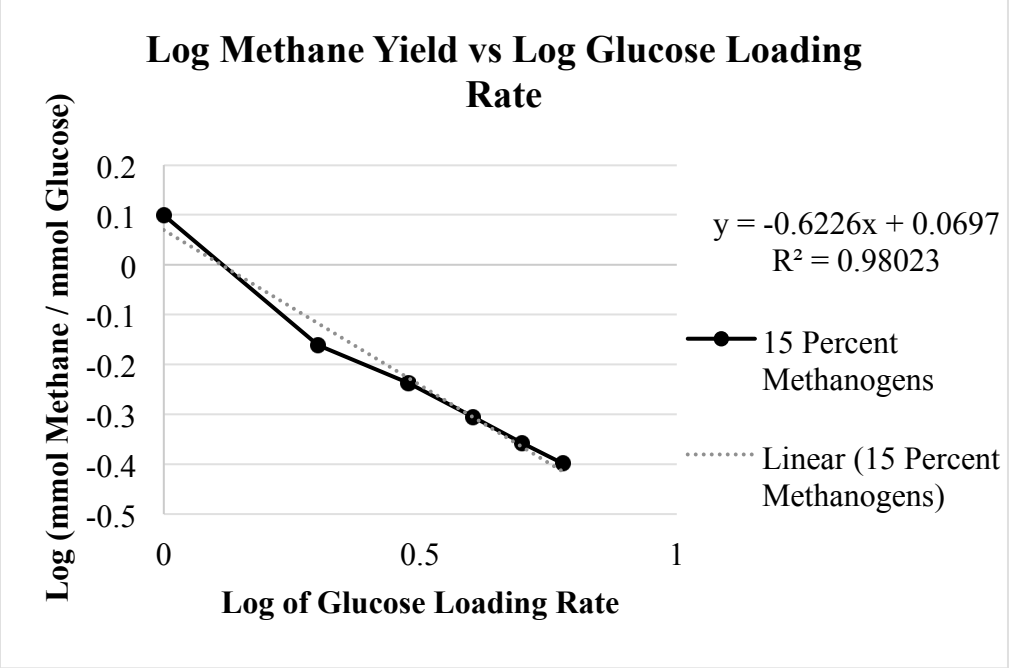

Figure 5.16: Linearized methane yield as a function of glucose loading rate at 15 percent methanogens

Applying the real world and theoretical anaerobic digester data to simulations of a pure culture anaerobic digester through the application of descriptive-OptCom produced methane yields that fell below real world methane yields. While the microorganisms selected in our anaerobic digester were selected due do their appearance in anaerobic digester sequences and apparent substrate connectivity, anaerobic digesters feature a large and diverse array of microorganisms. A greater diversity of microorganisms may allow specific organisms to metabolize the substrates that produce the most efficient biomass growth. Additionally, detailed data relating to the population diversity and operating parameters from real world anaerobic digesters would increase the number of data points to accurately constrain and validate the pure culture anaerobic digester GEM.

5.9 Pure Culture Anaerobic Digester Conclusions

Due to limited literature data available, validation of the individual and combined models was difficult to achieve. While each of the organisms selected for the pure culture anaerobic digester are commonly found in an anaerobic digester, no studies were conducted that cultured each of the organisms together. Culturing a tri-culture anaerobic digester and analyzing the resulting flux distribution, growth rates, and population dynamics would allow for the accurate validation of each organism, as well as provide additional constrains for the models included. While literature growth data exists for each organism, the growth studies were often conducted in pure culture, and may not fully represent the capability of each organism.

The methanogen selected for the pure culture anaerobic digester was chosen due to its capability to utilize a wide array of substrates. However, many methanogens are often only capable of using a limited range of substrates, or even only a single substrate. The inclusion of a greater array of methanogens may yield more accurate results, as a wide variety of methanogenic organisms are typically found in anaerobic digesters. (Reyes et al, 2015; Gerardi 2003)

While the constructed pure culture anaerobic digester GEM is limited by scarce data, the model represents the first community level GEM constructed for an anaerobic digester. The pure culture anaerobic digester GEM presents a high-resolution platform for the *in silico* analysis of an anaerobic digester. Through the application of OptCom and descriptive-OptCom, community dynamics and the flux distribution through an entire anaerobic digester can be examined. Additionally, as more manually curated GEMs become available, they can be quickly included in the pure culture anaerobic digester GEM, increasing the real world significance and applicability of the developed platform.

CHAPTER 6
CONCLUSIONS

6.1 iTK530 and Pure Culture Anaerobic Digester GEM Conclusions

Anaerobic digestion presents a unique opportunity for the sustainable treatment of wastewater and for the generation of a renewable energy source. Although many anaerobic digesters do not currently produce enough biogas suitable for economic recovery, a growing number of anaerobic digesters are able to produce sufficient quantities of biogas to reduce or completely offset the wastewater treatment plants energy consumption. Through the continued research and optimization of anaerobic digestion, wastewater treatment plants could see a significant reduction in net energy consumption or the addition of new sources of income through providing energy back to the electrical grid.

While there have been numerous attempts to model the anaerobic digestion process, current models approach the system at a macroscopic level. To address the limits presented by current *in silico* anaerobic digestion models, genome scale modeling was applied to the anaerobic digestion process. The application of genome-scale modeling presents a high-resolution platform for *in silico* analysis. Through the manual curation of a pure culture anaerobic digester GEM, population dynamics and operating parameters can be critically examined. The pure culture anaerobic digester GEM presented represents the first GEM constructed of an anaerobic digester. To complete the pure culture anaerobic digester GEM, a novel GEM for *S. wolfei* was constructed and validated.

Through the application of OptCom and descriptive-OptCom, the pure culture anaerobic digest GEM was analyzed. Through OptCom, the system was dominated by the acidogenic bacteria *C. acetobutlyicum*, while the methanogen *M. barkeri* only comprised 5.7 percent of the total community biomass. At a lower percent of total

methanogens, the methane yield fell below literature expected values, as the methanogens could not keep pace with the rapid growing acidogen.

Through the application of descriptive-OptCom, the total percent of methanogens was varied to examine the effects of methanogenic population on digester activity. Through the additional constraint of the *S. wolfei* GEM, the model produced methane yields that fell within the literature expected range when the methanogenic population comprised 20 to 40 percent of the total biomass and the glucose loading rate was set to 10 mmol/g DW-hr. Although the application of the pure culture anaerobic digester GEM did not accurately match data from real world anaerobic digesters, the construction of the pure culture anaerobic digester presents a unique opportunity for the continued exploration of the anaerobic digestion process.

Anaerobic digestion is a complex process that may not be accurately represented by three model microorganisms. Through the expansion of the pure culture anaerobic digester platform developed, anaerobic digestion can be continued to be analyzed and optimized. As more species are included in the pure culture anaerobic digester GEM, a high-resolution model will be developed that may be used by wastewater treatment plant operators to optimally control an anaerobic digester for by researchers for the continued exploration and elucidation of the anaerobic digestion process.

6.2 Future Work

Future work is needed to further optimize and explore both the *S. wolfei* GEM and the pure culture anaerobic digester GEM. To increase the validity of the *S. wolfei* GEM, in depth studies of the biomass composition and substrate uptake and secretion rates are required. In addition to the pure culture analysis of *S. wolfei*, further experiments need to be conducted that utilize *S. wolfei* with additional syntrophic partners to elucidate the interaction between multiple species.

To improve upon the pure culture anaerobic digester GEM, each organism included in the GEM needs to be culture separately and together. While data for each organism grown in pure culture was relatively abundant, no experiments have been conducted that examine how the three species included would function and interact together. In addition to the further understanding of the original three organisms included in the pure culture anaerobic digester GEM, additional species need to be included in the pure culture anaerobic digester to increase its applicability to real world anaerobic digesters. Increasing the diversity of acidogens in the model would allow the GEM to metabolize a greater range of substrates. Additional acetogens are required to further understand the role that fatty acid oxidizers play in anaerobic digestion. Increasing the number of methanogenic species will vastly improve the applicability of the model, as numerous methanogens species specialize in the degradation of a single substrate.

BIBLIOGRAPHY

Abbasi, T., Tauseef, S. M., & Abbasi, S. A. (2011). *Biogas energy* (Vol. 2). Springer Science & Business Media.

Ahring, B. K., Ibrahim, A. A., & Mladenovska, Z. (2001). Effect of temperature increase from 55 to 65 C on performance and microbial population dynamics of an anaerobic reactor treating cattle manure. *Water research*, *35*(10), 2446-2452.

Ali Shah, F., Mahmood, Q., Maroof Shah, M., Pervez, A., & Ahmad Asad, S. (2014). Microbial ecology of anaerobic digesters: the key players of anaerobiosis. *The Scientific World Journal*, *2014*.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

Amos, D. A., & McInerney, M. J. (1990). Growth of Synthrophomonas wolfei on unsaturated short chain fatty acids. *Archives of Microbiology*, *154*(1), 31-36.

Andrews, J. F. (1969). Dynamic model of the anaerobic digestion process. *Journal of the Sanitary Engineering Division*, *95*(1), 95-116.

Avery, G. B., Shannon, R. D., White, J. R., Martens, C. S., & Alperin, M. J. (2003). Controls on methane production in a tidal freshwater estuary and a peatland: methane production via acetate fermentation and CO2 reduction.*Biogeochemistry*, *62*(1), 19-37.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., ... & Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, *9*(1), 75.

Bahl, H., Andersch, W., Braun, K., & Gottschalk, G. (1982). Effect of pH and butyrate concentration on the production of acetone and butanol by Clostridium acetobutylicum grown in continuous culture. *European journal of applied microbiology and biotechnology*, *14*(1), 17-20.

Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R., & Wolfe, R. S. (1979). Methanogens: reevaluation of a unique biological group. *Microbiological reviews*, *43*(2), 260.

Beaty, P. S., & McInerney, M. J. (1987). Growth of Syntrophomonas wolfei in pure culture on crotonate. *Archives of microbiology*, *147*(4), 389-393.

Beaty, P. S., & McInerney, M. J. (1989). Effects of organic acid anions on the growth and metabolism of Syntrophomonas wolfei in pure culture and in defined consortia. *Applied and environmental microbiology*, *55*(4), 977-983.

Beaty, P. S., & McInerney, M. J. (1990). Nutritional features of Syntrophomonas wolfei. *Applied and environmental microbiology*, *56*(10), 3223-3224.

Boone, D. R., Whitman, W. B., & Rouvière, P. (1993a). Diversity and taxonomy of methanogens. In *Methanogenesis* (pp. 35-80). Springer US.

Bryers, J.D. (1985). Structured modelling of the anaerobic digestion of biomass particulates. *Biotechnology and Bioengineering*, **27**, 638-649.

BIBLIOGRAPHY (Continued)

Buswell A. M., & Mueller H. F. (1952). Mechanics of methane fermentation. J Ind Eng Chem 44: 550

Chawla, O. P. (1986). Advances in biogas technology. *Advances in biogas technology*.

Chen, Y., Cheng, J. J., & Creamer, K. S. (2008). Inhibition of anaerobic digestion process: a review. *Bioresource technology*, *99*(10), 4044-4064.

Costello, D.J., Greenfield, P.F. and Lee, P.L. (1991a). Dynamic modelling of a single-stage high-rate anaerobic reactor - I. Model derivation. *Wat. Res*., **25**, 847-858.

Czerkawski, J. W. (2013). *An introduction to rumen studies*. Elsevier.

Daims, H., Taylor, M. W., & Wagner, M. (2006). Wastewater treatment: a model system for microbial ecology. *Trends in biotechnology*, *24*(11), 483-489.

Dash, S., Mueller, T. J., Venkataramanan, K. P., Papoutsakis, E. T., & Maranas, C. D. (2014). Capturing the response of Clostridium acetobutylicum to chemical stressors using a regulated genome-scale metabolic model.*Biotechnology for biofuels*, *7*(1), 144.

Demirel, B., & Scherer, P. (2008). Production of methane from sugar beet silage without manure addition by a single-stage anaerobic digestion process. *Biomass and Bioenergy*, *32*(3), 203-209.

Demirel, B., & Scherer, P. (2008). The roles of acetotrophic and hydrogenotrophic methanogens during anaerobic conversion of biomass to methane: a review. *Reviews in Environmental Science and Bio/Technology*, *7*(2), 173-190.

Devoid, Scott, et al. "Automated genome annotation and metabolic model reconstruction in the SEED and Model SEED." *Systems Metabolic Engineering*. Humana Press, 2013. 17-45.

Diekert, G., & Wohlfarth, G. (1994). Metabolism of homoacetogens. *Antonie van Leeuwenhoek*, *66*(1-3), 209-221.

Drake, H. L. (2012). *Acetogenesis*. Springer Science & Business Media.

Eastman, J.A., and Ferguson, J.F. (1981). Solubilization of particulate organic carbon during the acid phase of anaerobic digestion. Journal WPCF, 53(3), 352–366.

Edwards, J. S., and B. O. Palsson. "The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities." *Proceedings of the National Academy of Sciences* 97.10 (2000): 5528-5533.

European Bioplastics. (2015). Anaerobic Digestion Fact Sheet. Retrieved from http://en.european-bioplastics.org/wp-content/uploads/2011/04/fs/FactSheet_Anaerobic_Digestion.pdf

Feist, Adam M., et al. "Reconstruction of biochemical networks in microorganisms." *Nature Reviews Microbiology* 7.2 (2008): 129-143.

Ferry, J. G. (1992). Methane from acetate. *Journal of Bacteriology*, *174*(17), 5489-5495.

BIBLIOGRAPHY (Continued)

Follows, M. J., Dutkiewicz, S., Grant, S., & Chisholm, S. W. (2007). Emergent biogeography of microbial communities in a model ocean. *science*, *315*(5820), 1843-1846.

Frey, E. (2010). Evolutionary game theory: Theoretical concepts and applications to microbial communities. *Physica A: Statistical Mechanics and its Applications*, *389*(20), 4265-4298.

Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature*, *459*(7244), 193-199.

GAMS Development Corporation. (2013). General Algebraic Modeling System (GAMS) Release 24.2.1. Washington, DC, USA.

Gerardi, M. H. (2003). *The microbiology of anaerobic digesters*. John Wiley & Sons.

Goldman, A. D., Leigh, J. A., & Samudrala, R. (2009). Comprehensive computational analysis of Hmd enzymes and paralogs in methanogenic Archaea. *BMC Evolutionary Biology*, *9*(1), 199.

Gonnerman, M. C., Benedict, M. N., Feist, A. M., Metcalf, W. W., & Price, N. D. (2013). Genomically and biochemically accurate metabolic reconstruction of Methanosarcina barkeri Fusaro, iMG746. *Biotechnology journal*, *8*(9), 1070-1079.

Grady Jr, C. L., Daigger, G. T., Love, N. G., & Filipe, C. D. (2011). *Biological wastewater treatment*. CRC Press.

Griffin, M. E., McMahon, K. D., Mackie, R. I., & Raskin, L. (1998). Methanogenic population dynamics during start-up of anaerobic digesters treating municipal solid waste and biosolids. *Biotechnology and bioengineering*, *57*(3), 342-355.

Gujer, W., & Zehnder, A. J. B. (1983). Conversion processes in anaerobic digestion. *Water Science & Technology*, *15*(8-9), 127-167.

Hansen, S. K., Rainey, P. B., Haagensen, J. A., & Molin, S. (2007). Evolution of species interactions in a biofilm community. *Nature*, *445*(7127), 533-536.

Hedderich, R., & Whitman, W. B. (2006). Physiology and biochemistry of the methane-producing Archaea. In *The prokaryotes* (pp. 1050-1079). Springer New York.

Henze, M. (Ed.). (2008). *Biological wastewater treatment: principles, modelling and design*. IWA publishing.

IPCC. (2001). *Climate change 2001: impacts, adaptation, and vulnerability: contribution of Working Group II to the third assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Jackson, B. E., & McInerney, M. J. (2002). Anaerobic microbial metabolism can proceed close to thermodynamic limits. *Nature*, *415*(6870), 454-456.

BIBLIOGRAPHY (Continued)

Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Dröge, M., Gartemann, K. H., ... & Goesmann, A. (2011). Comparative and joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-pyrosequencing. *PLoS One*,*6*(1), e14519.

Kalyuzhnyi, S. V. (1997b). Batch anaerobic digestion of glucose and its mathematical modeling. II. Description, verification and application of model.*Bioresource technology*, *59*(2), 249-258.

Kalyuzhnyi, S. V., & Davlyatshina, M. A. (1997a). Batch anaerobic digestion of glucose and its mathematical modeling. I. Kinetic investigations. *Bioresource Technology*, *59*(1), 73-80.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, *28*(1), 27-30.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research*, *42*(D1), D199-D205.

Kapdi, S. S., Vijay, V. K., Rajesh, S. K., & Prasad, R. (2005). Biogas scrubbing, compression and storage: perspective and prospectus in Indian context. *Renewable energy*, *30*(8), 1195-1202.

Karakashev, D., Batstone, D. J., & Angelidaki, I. (2005). Influence of environmental conditions on methanogenic compositions in anaerobic biogas reactors. *Applied and environmental microbiology*, *71*(1), 331-338.

Kerr, B., Riley, M. A., Feldman, M. W., & Bohannan, B. J. (2002). Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature*, *418*(6894), 171-174.

Kotsyurbenko, O. R., Friedrich, M. W., Simankova, M. V., Nozhevnikova, A. N., Golyshin, P. N., Timmis, K. N., & Conrad, R. (2007). Shift from acetoclastic to H2-dependent methanogenesis in a West Siberian peat bog at low pH values and isolation of an acidophilic Methanobacterium strain. *Applied and environmental microbiology*, *73*(7), 2344-2348.

Kythreotou, N., Florides, G., & Tassou, S. A. (2014). A review of simple to scientific models for anaerobic digestion. *Renewable Energy*, *71*, 701-714.

Leang, C., Ueki, T., Nevin, K. P., & Lovley, D. R. (2013). A genetic system for Clostridium ljungdahlii: a chassis for autotrophic production of biocommodities and a model homoacetogen. *Applied and environmental microbiology*, *79*(4), 1102-1109.

Lehmann, L., & Keller, L. (2006). The evolution of cooperation and altruism–a general framework and a classification of models. *Journal of evolutionary biology*, *19*(5), 1365-1376.

Li, Y., and Noike, T. (1987). Characteristics of the degradation of excess activated sludge in anaerobic acidogenic phase. Jpn. J. Water Poll. Res., 10, 740–795.

BIBLIOGRAPHY (Continued)

Liu, Y., & Whitman, W. B. (2008). Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences*, *1125*(1), 171-189.

Liu, Y., Kong, Y., Zhang, R., Zhang, X., Wong, F., Tay, J., ... & Liu, W. (2010). Microbial population dynamics of granular aerobic sequencing batch reactors during start-up and steady state periods.

Lorowitz, W. H., Zhao, H., & Bryant, M. P. (1989). Syntrophomonas wolfei subsp. saponavida subsp. nov., a Long-Chain Fatty-Acid-Degrading, Anaerobic, Syntrophic Bacterium; Syntrophomonas wolfei subsp. wolfei subsp. nov.; and Emended Descriptions of the Genus and Species. *International Journal of Systematic Bacteriology*, *39*(2), 122-126.

Luengo, J. M., García, B., Sandoval, A., Naharro, G., & Olivera, E. R. (2003). Bioplastics from microorganisms. *Current opinion in microbiology*, *6*(3), 251-260.

Lyberatos, G., & Skiadas, I. V. (1999). Modelling of anaerobic digestion–a review. *Global Nest Int J*, *1*(2), 63-76.

Madigan, M. T., Clark, D. P., Stahl, D., & Martinko, J. M. (2010). *Brock Biology of Microorganisms 13th edition*. Benjamin Cummings.

Mahadevan, R., Edwards, J. S., & Doyle, F. J. (2002). Dynamic flux balance analysis of diauxic growth in Escherichia coli. *Biophysical journal*, *83*(3), 1331-1340.

MATLAB. (2013b). The MathWorks, Inc., Natick, Massachusetts, United States.

McCarty, P. L. (1982). One hundred years of anaerobic treatment. In *Anaerobic digestion 1981: proceedings of the Second International Symposium on Anaerobic Digestion held in Travemhunde, Federal Republic of Germany, on 6-11 September, 1981/editors, DE Hughes...[et al.]*. Amsterdam; New York: Elsevier Biomedical, 1982.

McInerney, M. J., Bryant, M. P., Hespell, R. B., & Costerton, J. W. (1981). Syntrophomonas wolfei gen. nov. sp. nov., an anaerobic, syntrophic, fatty acid-oxidizing bacterium. *Applied and Environmental Microbiology*, *41*(4), 1029-1039.

McLellan, S. L., Huse, S. M., Mueller-Spitz, S. R., Andreishcheva, E. N., & Sogin, M. L. (2010). Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environmental microbiology*, *12*(2), 378-392.

McMahon, K. D., Stroot, P. G., Mackie, R. I., & Raskin, L. (2001). Anaerobic codigestion of municipal solid waste and biosolids under various mixing conditions—II: microbial population dynamics. *Water Research*, *35*(7), 1817-1827.

BIBLIOGRAPHY (Continued)

Milne, C. B., Kim, P. J., Eddy, J. A., & Price, N. D. (2009). Accomplishments in genome-scale in silico modeling for industrial and medical biotechnology. *Biotechnology journal*, *4*(12), 1653-1670.

Mladenovska, Z., Dabrowski, S., & Ahring, B. (2003). Anaerobic digestion of manure and mixture of manure with lipids: biogas reactor performance and microbial community analysis. *Water Science & Technology*, *48*(6), 271-278.

Moigno, A. F. (1881). Mouras' automatic scavenger.

Montero, B., Garcia-Morales, J. L., Sales, D., & Solera, R. (2008). Evolution of microorganisms in thermophilic-dry anaerobic digestion. *Bioresource technology*, *99*(8), 3233-3243.

National Library of Medicine. (2015). What kinds of gene mutations are possible? Retrieved August 13, 2015.

NCBI. (2010). Syntrophomonas wolfei subsp. wolfei str. Goettingen G311. http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=335541

Oberhardt, M. A., Palsson, B. Ø., & Papin, J. A. (2009). Applications of genome-scale metabolic reconstructions. *Molecular systems biology*, *5*(1), 320.

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. Ø. (2011). A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011. *Molecular systems biology*, *7*(1), 535.

Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., ... & Vonstein, V. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic acids research*, *33*(17), 5691-5702.

Parkin, G. F., & Owen, W. F. (1986). Fundamentals of anaerobic digestion of wastewater sludges. *Journal of Environmental Engineering*, *112*(5), 867-920.

Peng, R. H., Xiong, A. S., Xue, Y., Fu, X. Y., Gao, F., Zhao, W., ... & Yao, Q. H. (2008). Microbial biodegradation of polyaromatic hydrocarbons. *FEMS microbiology reviews*, *32*(6), 927-955.

Ragsdale, S. W., & Pierce, E. (2008). Acetogenesis and the Wood–Ljungdahl pathway of CO 2 fixation. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1784*(12), 1873-1898.

Rapport, J., Zhang, R., Jenkins, B. M., & Williams, R. B. (2008). Current anaerobic digestion technologies used for treatment of municipal organic solid waste. *University of California, Davis, Contractor Report to the California Integrated Waste Management Board*.

Reece, J., Urry, L. A., Meyers, N., Cain, M. L., Wasserman, S. A., Minorsky, P. V., ... & Cooke, B. N. (2011). *Campbell biology*. Pearson Higher Education AU.

BIBLIOGRAPHY (Continued)

Reyes, M., Borrás, L., Seco, A., & Ferrer, J. (2015). Identification and quantification of microbial populations in activated sludge and anaerobic digestion processes. *Environmental technology*, *36*(1), 45-53.

Richards, B. K., Cummings, R. J., White, T. E., & Jewell, W. J. (1991). Methods for kinetic analysis of methane fermentation in high solids biomass digesters. *Biomass and Bioenergy*, *1*(2), 65-73.

Riviere, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., ... & Sghir, A. (2009). Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *The ISME journal*, *3*(6), 700-714.

Sabra, W., Dietz, D., Tjahjasari, D., & Zeng, A. P. (2010). Biosystems analysis and engineering of microbial consortia for industrial biotechnology. *Engineering in Life Sciences*, *10*(5), 407-421.

Sanapareddy, N., Hamp, T. J., Gonzalez, L. C., Hilger, H. A., Fodor, A. A., & Clinton, S. M. (2009). Molecular diversity of a North Carolina wastewater treatment plant as revealed by pyrosequencing. *Applied and environmental microbiology*, *75*(6), 1688-1696.

Santos, F., Boele, J., & Teusink, B. (2011). A practical guide to genome-scale metabolic models and their analysis. *Methods enzymol*, *500*, 509-532.

Sarmiento, F. B., Leigh, J. A., & Whitman, W. B. (2011). Genetic systems for hydrogenotrophic methanogens. *Methods in enzymology*, *494*, 43-73.

Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., ... & Palsson, B. Ø. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2. 0. *Nature protocols*, *6*(9), 1290-1307.

Scherer, P., Vollmer, G., Fakhouri, T., & Martensen, S. (2000). Development of a methanogenic process to degrade exhaustively the organic fraction of municipal grey waste under thermophilic and hyperthermophilic conditions. *Water science and technology*, *41*(3), 83-91.

Schilling, Christophe H., and Bernhard Ø. Palsson. "Assessment of the Metabolic Capabilities of *Haemophilus influenza* Rd through a Genome-scale Pathway Analysis." *Journal of theoretical biology* 203.3 (2000): 249-283.

Schmidt, J. E., Mladenovska, Z., Lange, M., & Ahring, B. K. (2000). Acetate conversion in anaerobic biogas reactors: traditional and molecular tools for studying this important group of anaerobic microorganisms. *Biodegradation*, *11*(6), 359-364.

Schöcke, L., & Schink, B. (1997). Energetics of methanogenic benzoate degradation by Syntrophus gentianae in syntrophic coculture. *Microbiology*,*143*(7), 2345-2351.

BIBLIOGRAPHY (Continued)

Scholten, J. C., & Conrad, R. (2000). Energetics of syntrophic propionate oxidation in defined batch and chemostat cocultures. *Applied and environmental microbiology*, *66*(7), 2934-2942.

Schuster, S., Kreft, J. U., Brenner, N., Wessely, F., Theißen, G., Ruppin, E., & Schroeter, A. (2010). Cooperation and cheating in microbial exoenzyme production–theoretical analysis for biotechnological applications. *Biotechnology journal*, *5*(7), 751-758.

Segre, D., Vitkup, D., & Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences*, *99*(23), 15112-15117.

Sieber, J. R., Sims, D. R., Han, C., Kim, E., Lykidis, A., Lapidus, A. L., ... & McInerney, M. J. (2010). The genome of Syntrophomonas wolfei: new insights into syntrophic metabolism and biohydrogen production. *Environmental microbiology*, *12*(8), 2289-2301.

Siegrist, H., Renggli, D. and Gujer, W. (1993). Mathematical modelling of anaerobic mesophilic sewage sludge treatment. *Wat. Sci. Technol.*, **27**, 25-36.

Smith, P.H., Bordeaux, F.M., Goto, M., Shiralipour, A., Wilke, A., Andrews, J.F., Ide, S. and Barnett, M.W. (1988). Biological production of methane from biomass. In: Methane from biomass. *A treatment approach.* Smith, W.H. and Frank, J.R. (Eds.), Elsevier, London, 291-334.

Sobieraj, M., & Boone, D. R. (2006). Syntrophomonadaceae. In *The Prokaryotes* (pp. 1041-1049). Springer US.

Speece, R. E. (1983). Anaerobic biotechnology for industrial wastewater treatment. *Environmental science & technology*, *17*(9), 416A-427A.

Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. *Science*, *315*(5813), 801-804.

Stroot, P. G., McMahon, K. D., Mackie, R. I., & Raskin, L. (2001). Anaerobic codigestion of municipal solid waste and biosolids under various mixing conditions—I. Digester performance. *Water Research*, *35*(7), 1804-1816.

Tawarmalani, M., & Sahinidis, N. V. (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, *103*(2), 225-249.

Thayanukul, P., Zang, K., Janhom, T., Kurisu, F., Kasuga, I., & Furumai, H. (2010). Concentration-dependent response of estrone-degrading bacterial community in activated sludge analyzed by microautoradiography-fluorescence in situ hybridization. *Water research*, *44*(17), 4878-4887.

Thiele, I., & Palsson, B. Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols*, *5*(1), 93-121.

BIBLIOGRAPHY (Continued)

Tietjen, C. (1975). From biodung to biogas—historical review of European experience. *Energy, Agriculture, and Waste Management, WJ Jewell Ed. Ann Arbor Science Publishers, Inc. Ann Arbor, Michigan*, 247-259.

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(30), 10854-10861.

Tomei, M. C., Braguglia, C. M., Cento, G., & Mininni, G. (2009). Modeling of anaerobic digestion of sludge. *Critical Reviews in Environmental Science and Technology*, *39*(12), 1003-1051.

UCSD. "Other Organisms." *Other Organisms*. Web. 7 Jan. 2015. <http://systemsbiology.ucsd.edu/InSilicoOrganisms/OtherOrganisms>.

United States. Environmental Protection Agency. (2012). *Summary Report: Global Anthropogenic Non-CO2 Greenhouse Gas Emissions: 1990-2030.* Environmental Protection Agency.

Vogels, G. D., Keltjens, J. T., & Van Der Drift, C. (1988). Biochemistry of methane production. *Biology of anaerobic microorganisms*, 707-770.

Wagner, M., & Loy, A. (2002). Bacterial community composition and function in sewage treatment systems. *Current Opinion in Biotechnology*, *13*(3), 218-227.

Wang, A., Liu, W., Cheng, S., Xing, D., Zhou, J., & Logan, B. E. (2009). Source of methane and methods to control its formation in single chamber microbial electrolysis cells. *international journal of hydrogen energy*, *34*(9), 3653-3658.

Wirth, R., Kovács, E., Maróti, G., Bagi, Z., Rákhely, G., & Kovács, K. L. (2012). Characterization of a biogas-producing microbial community by short-read next generation DNA sequencing. *Biotechnol Biofuels*, *5*(1), 41.

Xavier, J. B. (2011). Social interaction in synthetic and natural microbial communities. *Molecular systems biology*, *7*(1), 483.

Yu, Y., Lee, C., & Hwang, S. (2005). Analysis of community structures in anaerobic processes using a quantitative real-time PCR method. *Water science & technology*, *52*(1), 85-91.

Zheng, D., & Raskin, L. (2000). Quantification of Methanosaeta species in anaerobic bioreactors using genus-and species-specific hybridization probes. *Microbial ecology*, *39*(3), 246-262.

Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., & Lovley, D. R. (2011). Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *The ISME journal*, *5*(2), 305-316.

BIBLIOGRAPHY (Continued)

Zomorrodi, A. R., & Maranas, C. D. (2012). OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput Biol*, *8*(2), e1002363.

APPENDIX

Table A.1: Errors introduced during from genome annotations during automated GEM reconstruction (adopted from Feist et al, 2008)

Problem	Description	Methods
	Genome annotations	
Annotations are not continuously updated with new information	As new genes are found, older genome annotations are not updated, resulting in incorrectly annotated genes. For example, in most databases, *slr0788* in *Synechocystis* spp. is annotated as a pre-B-cell enhancing factor (a mammalian function assigned to a bacterial gene), but in SEED21, is correctly annotated as nicotinamide phosphoribosyltransferase.	Automated annotation pipelines can be used to reanalyse older genome annotations.
Incorrect annotations	Incorrect annotations can be due to either missing genes (from sequencing or gene-finding algorithm errors) or incorrect gene annotations. This can occur for a number of reasons. For example, when new sequences are not used to update older genome annotations or when weak homology is used as sole evidence for functional assignment.	Analysis of reconstructed networks can help identify some of these errors.
Missing functionalities	Approximately 30% of enzyme activities with enzyme commission numbers lack sequence data118.Therefore, not all reactions will be associated with gene or protein sequences. For example, in 2005, the 6-phosphogluconlactonase gene (*pgl*) in *Escherichia coli* was discovered119. Prior to this, there was no *pgl* gene in the genome annotation even though the enzymatic activity was observed in cell extracts.	Automated tools have been developed to find missing reactions (for example, SMILEY algorithm, GapFind (or GapFill) PathoLogic114 and topology-based methods).
Transporter specificity	Annotations for transporters often lack sufficient detail to determine what substrate (or substrates) they transport, even though the mechanism (for example, proton symport or ATP hydrolysis) is known.	Methods for improving transporter functional annotations are needed.

Table A.2: Errors introduced from databases during automated GEM reconstruction (adopted from Feist et al, 2008)

Problem	Description	Methods
	Databases	
Gene-protein-reaction (GPR) associations	Relationships between genes, enzymes and reactions are not always clearly defined (for example, subunits compared with isozymes).	Can be automated based on comparisons of sequences and known GPRs.
Reaction specificity	Reactions are often characterized through their actions on a general class of compounds, which can result in ambiguous connections in a network. Common general classes include electron carriers (for example, quinones, NAD compared with NADP) and alcohols (for example, ethanol and methanol compared with butanol).	Changes in databases are needed or automated tools need to be developed.
Reaction imbalances	Reactions are not elementally balanced for H, C, P, N, O or S. This means that substrates and products are missing from imbalanced reactions. For example, analysis of the KEGG database in 2004 found that only 51% of the reactions were balanced for C, P, N, O, H and S.	Automated procedures are available to check elemental reaction balancing.
Reaction directionality	Reactions are generally defined as reversible. This can be a problem; for example, if cycles between reactions allow the free conversion of ADP into ATP (free-energy equivalents).	Automated procedures have been developed.
Compound protonation states	Reactions are generally written for the neutral form of molecules and do not account for the protonation state of compounds (for example, carboxylic acid groups are deprotonated at pH 7). This affects the stoichiometric coefficients for protons across the network.	pKa prediction software is available, and therefore automation is possible.
Coenzyme availability	Enzymes often need coenzymes (for example, pyridoxal 5-phosphate, vitamin B12 and biotin). For enzymes to be functional, the cell must be able to produce them or get them from the environment. BRENDA contains this type of information, and is available for download.	Automation is possible now that data are becoming available.
Organism-specific pathways	The cell membrane (or membranes) is composed of macromolecules (for example, phospholipids and peptidoglycans) that can vary across organisms and species. As a result, the biosynthesis pathways for these compounds are often unique.	Would require experimental data and is therefore unlikely to become subject to automation.