AN ABSTRACT OF THE THESIS OF

William Hohenschuh for the degree of Master of Science in Biological and Ecological Engineering presented on May 28, 2014

Title: A Modeling and Experimental Approach to Understanding the Bottlenecks in Xylulose Utilization in *S. Cerevisiae*

Abstract approved: _____

Ganti Murthy

The production of fuel ethanol from lignocellulosic biomass has the potential to replace a significant portion of non-renewable transport fuels. Woody feedstocks are composed of cellulose, hemicellulose, and lignin. Glucose, the monomer of cellulose, is readily utilized by wild-type *S. cerevisiae*, but xylose, which comprises 60% of the sugar in hemicellulose, is not. To make the process economically competitive with conventional fossil fuels, both five and six carbon sugars must be utilized efficiently.

One approach to improving xylose utilization is to convert it to the more readily usable xylulose using an extracellular enzyme. Xylulose is taken up by wild-type *S. cerevisiae* and incorporated into the pentose phosphate pathway. To our knowledge there are no reports that elucidate the kinetics of this pathway, an important hurdle to overcome for strain development.

This thesis documents the work carried out to gain a better understanding of the xylulose utilization pathway in *S. cerevisiae*. The work was comprised of a

series of batch fermentations that identified xylulokinase as a limiting enzyme in wild-type strains and transport through the HXT family of hexose transporters as a possible limiting step in xylulokinase enhanced strains. Batch experiments with HXT knockout strains suggest that alternative modes of xylulose transport are possible and may be up regulated in the knockout. An existing genome scale model for *S. cerevisiae* (iMM904) was used as the basis to develop a dynamic flux balance model. This model was used to verify the batch fermentation findings. The model has strong predictive capacity for xylulose and glucose consumption under anaerobic conditions and sugar levels sufficient to support growth.

A Modeling and Experimental Approach to Understanding the Bottlenecks in Xylulose Utilization in *S. Cerevisiae*

by
William Hohenschuh

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented May 28, 2014
Commencement June 2014

Master of Science thesis of William Hohenschuh presented on May 28, 2014

APPROVED:

_____

Major Professor representing Biological and Ecological Engineering

_____

Head of the Department of Biological and Ecological 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.

_____

William Hohenschuh, Author

ACKNOWLEDGEMENTS

I'd like to expresses my sincere appreciation to …

My advisor, Dr. Ganti Murthy for his insightful guidence, sincere trust, and overwhelming belief that if we just tried enough things and ran enough experiments *something* would work.

Dr. Ronald Hector for both his guidence and work in developing the yeast strains used in this work.

Deepak and Ankita Garg for welcoming me to the lab, guiding my transition to grad school, providing a wealth of experience in the lab and out, and being genuinly good friends.

Allex McDaniel and Hossein Tabatabaie for bringing a boost of fresh energy and new ideas to the lab; key to finishing this work.

My parents for helping keep everything in perspective.

AND

Becca Wolf for gracefully handling the late night panic attacks that have come with a few years of undergrad, more all nighters than we

really neaded, dozens of essays, a handfull of graduate school applications, weeks of failed experiments, and months of thesis writing.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

Chapter 1 - Literature review:

1.1 Introduction: Cellulosic Ethanol as a Viable, Sustainable Fossil Fuel Replacement

Despite increasing prices (British Petroleum, 2012) and warnings of dwindling supplies, liquid fuels consumption is expected to increase 15% by 2030 (British Petroleum, 2013). As evidence mounts supporting the role of fossil fuels in global climate change, research and government policy has increasingly focused on identifying and improving renewable, carbon neutral liquid fuels (2007). Starch and sugar derived ethanol have been leading candidates because of their infrastructure compatibility (Lynd et al. 1991) and, at least in the US, because of the relative abundance of highly subsidized corn (Auld, 2012). However, the emissions savings associated with these technologies can vary greatly. Using efficient processing technologies, the net reduction in GHGs (from both production and use) when gasoline is displaced by corn ethanol can be up to 26%. This value increases to 37% if military GHG emissions are included in the production of crude oil. The use of less efficient processing technologies can push the GHG associated with corn ethanol beyond those of conventional gasoline (Figure 1) (Wang et al., 2011).

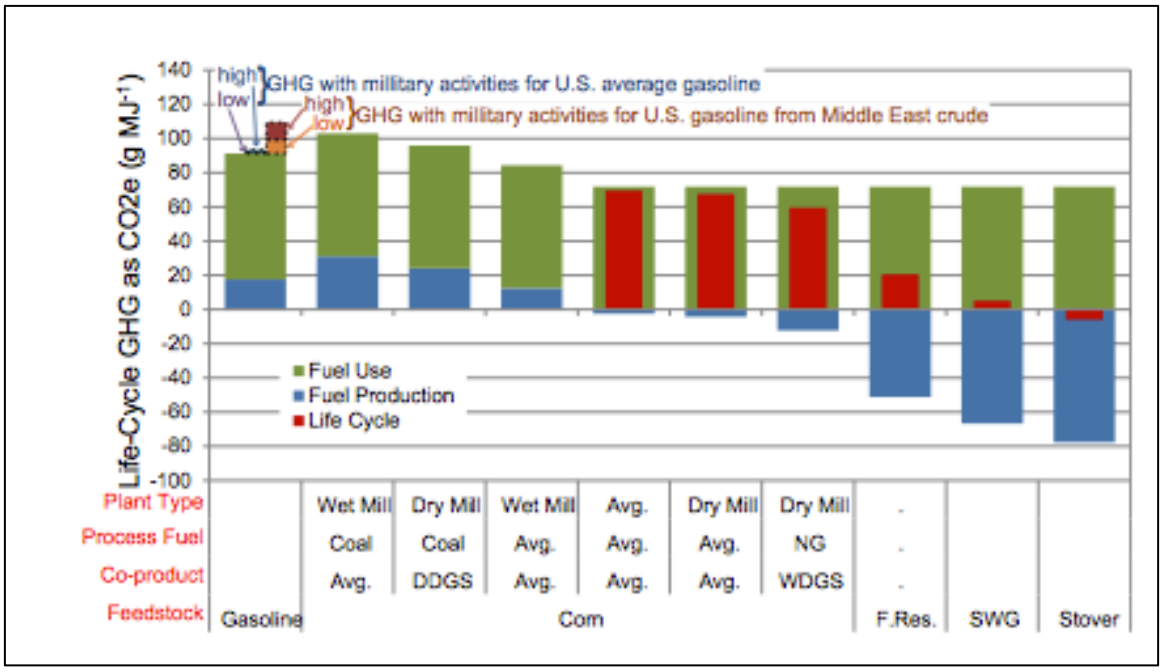


Figure 1: Life-cycle GHG Emissions of Petroleum Gasoline, Corn Ethanol, and Cellulosic Ethanol - (g CO_2 per MJ) (Wang et al., 2011)

Further supporting a case against corn ethanol, corn is grown on agriculturally productive land, often with intensive fertilizer, pesticide, and irrigation usage (Yang et al., 2012). Increasing demand for these resources in combination with increased speculative trading of corn and ethanol is thought to have caused an increase in food cost and price volatility (Demirer et al., 2012; Thompson, 2012).

By comparison, studies have shown that ethanol produced from lignocellulosic biomass (termed a second generation biofuel) could reduce GHG emissions by 46% to 90% in comparison to conventional petroleum (Borrion et al., 2012). Feedstock production is not confined by optimal geographical growing regions, land quality, or the need for irrigation in the same way food crops often are (Gelfand et al., 2013). If produced from low input biomass grown on marginal

land or from waste streams such as agricultural crop residues or forest management residues, lignocellulosic ethanol could provide both a greater supply and environmental benefits than corn-based ethanol (Hill et al., 2006; Lal, 2005).

1.2 General Process for Cellulosic Ethanol Production

The "conventional" process for conversion of lignocellulosic biomass to fuel ethanol can be broken into five main steps:

1) *Pretreatment* – breaking down the lignocellulosic matrix of the biomass to allow enzymes greater access to cellulose
2) *Detoxification* – removal of compounds inhibitory to hydrolysis, microorganism growth, or fermentation
3) *Enzymatic Hydrolysis* – liberating glucose monomers from cellulose chains using a combination of cellulolytic enzymes
4) *Fermentation* – the conversion of glucose to ethanol by any microorganism, but most commonly yeast
5) *Distillation and dehydration* – separation of ethanol from the fermentation broth so that it can be utilized as a fuel

Pretreatment refers to processes meant to disrupt the natural structure of a lignocellulosic feedstock to allow for greater sugar extraction (Hendriks and Zeeman, 2009). Pretreatment methods include dilute acid, dilute alkali, and steam explosion treatments (Table 1) and result in physical size reduction, hydrolysis of hemi-cellulose, structural alteration or removal lignin, and disruption

of cellulose crystals (Margeot et al., 2009). Effective pretreatment is crucial to the economic viability of cellulosic ethanol as it can raise the percentage of sugar extraction during hydrolysis (and there by accessible for fermentation) by greater than three fold (Kumar and Murthy, 2011; Yang and Wyman, 2008).

Table 1: Results of Ethanol Production Models Using Different Pretreatment Processes - All results are per 10,000MJ functional unit unless mentioned otherwise. ** Energy from lignin residue and biogas (Kumar and Murthy, 2011).

	Dilute Acid	Dilute Alkali	Hot Water	Steam Explosion
Ethanol Yield (L/dry ton biomass)	256.65	255.83	255.30	230.25
Thermal Energy Use (MJ)	8935.31	8807.22	9087.42	6349.34
Electrical Use (kWh)	433.61	415.21	439.22	408.85
Water Use (kg)	2801.55	2850.33	2746.31	2050.26
Co-Product Energy (MJ)**	13,270.54	13,145.58	13,696.10	16,366.41
Electricity Produced (kWh)	361.25	361.51	384.04	834.71

After pretreatment the feedstock stream is split often into a solids stream (which contains the majority of cellulose and six carbon sugars) and a liquid stream (which contains the majority of the solubilized hemicellulose and five carbon sugars). The solids stream will be treated with either acid or a cocktail of cellulases and β-glucosidases to release cellobiose from the cellulose and further hydrolyze them to glucose monomers (Katz and Reese, 1968; Reese, 1955). While many microorganisms can easily ferment free glucose monomers, *Saccharomyces cerevisiae* is used predominantly because of its high conversion yield (0.51g ethanol/g glucose) and high utilization rate.

High severity pretreatments (with combinations of high temperature, extreme pH, and duration) often liberate more sugars from cellulose and hemicellulose than gentler pretreatments, but they can also further degrade sugars and lignin monomers to produce compounds such as HMF (5-hydroxymethylfurfural) and 2-furfuralaldehyde (Figure 2) (Kabel et al., 2007; Pedersen and Meyer, 2010).

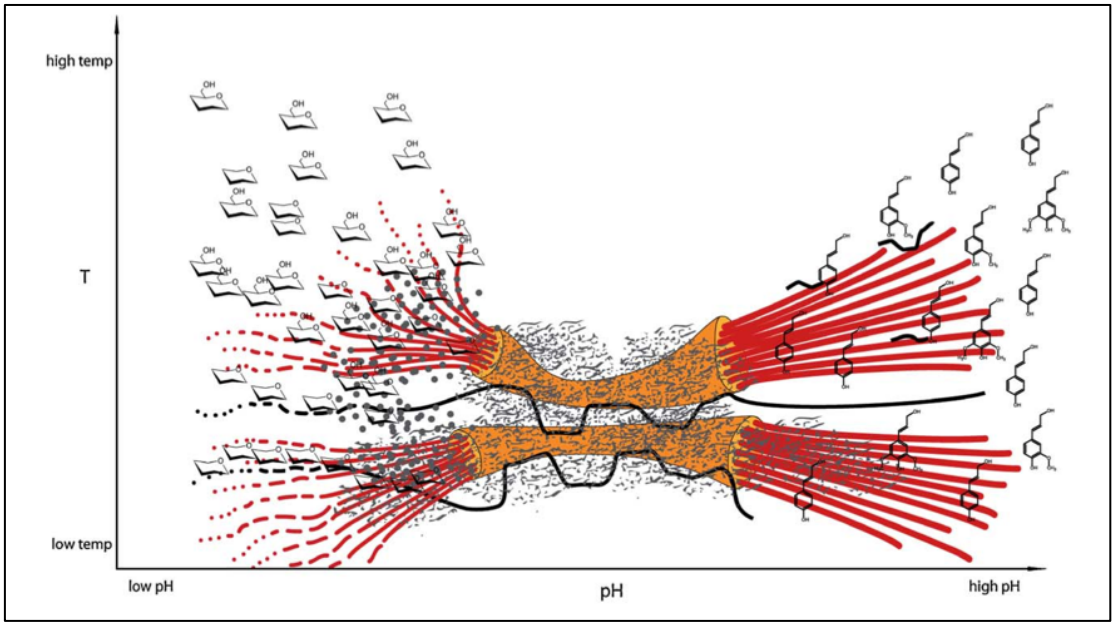


Figure 2: Pretreatment of Lignocellulose as Affected By Temperature and pH - Red and orange illustrate cellulosic fibrils and micro fibrils; Gray represents lignin; Black represents hemicellulose (Pedersen and Meyer, 2010).

These soluble compounds remain in the liquid, five carbon sugar stream and must be removed by detoxification before fermentation (Palmqvist and Hahn-Hägerdal, 2000). Detoxification can be performed through a process known as over-liming in which inhibitory compounds are precipitated through a pH shift or through filtration over activated charcoal.

Following detoxification the remaining sugars are fermented to ethanol. In conventional lignocellulosic ethanol production, the five and six carbon sugar streams are often mixed for fermentation, but some protocols use separate five and six carbon fermentations in hopes of increasing yield or decreasing costs. Still other processes (termed simultaneous saccharification and fermentation or SSF processes) combine the enzymatic saccharification and fermentation steps into one (Olofsson et al., 2008).

Following fermentation, solids are removed via centrifugation. These non-fermentable solids are often burned to produce process energy. Finally, the dilute ethanol stream is distilled and passed through a molecular sieve to produce anhydrous fuel ethanol (Jeong et al., 2012). A generalized process is outlined in Figure 3 (Kumar and Murthy, 2011).

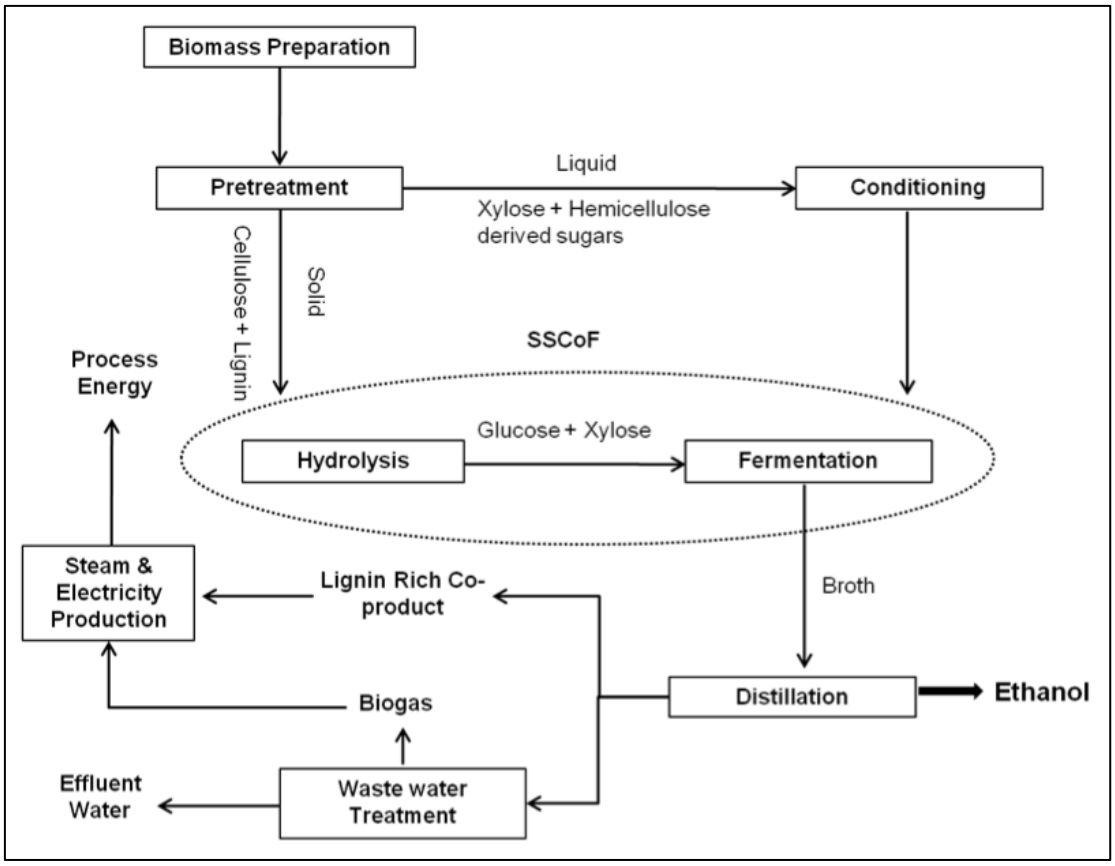


Figure 3: Generic Process Diagram For Ethanol Production From Lignocellulosic Biomass - (Kumar and Murthy, 2011)

While the basic production process is well understood, several aspects (notably feedstock selection, pretreatment, and utilization of five carbon sugars) are still evolving (Kudakasseril Kurian et al., 2013; Kumar and Murthy, 2011; Matsushika et al., 2009). Competing technologies provide different advantages and no technology has a clear path to commercialization. Here we focus on the many options for the utilization of pentose sugars, their advantages, challenges, and potential.

1.3 The Economic Case For Improving Pentose Sugar Utilization

Going forward, the adoption of biofuels will be driven more by economics than by the desire to decrease GHG emissions. Accordingly, the production of fuel ethanol at a production scale must be economically competitive with the production of conventional gasoline. Techno-economic analyses of lignocellulosic ethanol production have identified capital expenditures and the cost of raw feedstocks to be the largest contributing factors to final ethanol cost (Sassner et al., 2008; Wingren et al., 2003). Lignocellulosic feedstocks are comprised mainly of cellulose, hemicellulose and lignin. After pretreatment and saccharification, the hexoses produced from the cellulose fraction are readily fermented to ethanol and CO_2 at efficiencies greater than 90%. The pentose sugar fraction formed from hemicellulose is less readily fermentable by industrial yeast *S. cerevisiae*. Several strategies to improve pentose fermentation are discussed below, currently no industrial yeast strains with greater than 90% xylose conversion efficiencies are commercialized and many techno-economic analyses assume currently infeasible xylose fermentation efficiencies to show the potential cost of ethanol with improved technology (Sassner et al., 2008). Hemicellulose makes up roughly 20 to 50% of common lignocellulosic feedstocks by dry weight (Kumar et al., 2012; Sun and Cheng, 2002). Successful fermentation of hemicellulose sugars would increase ethanol yield by up to 40% (Figure 4) (Kumar et al., 2012).

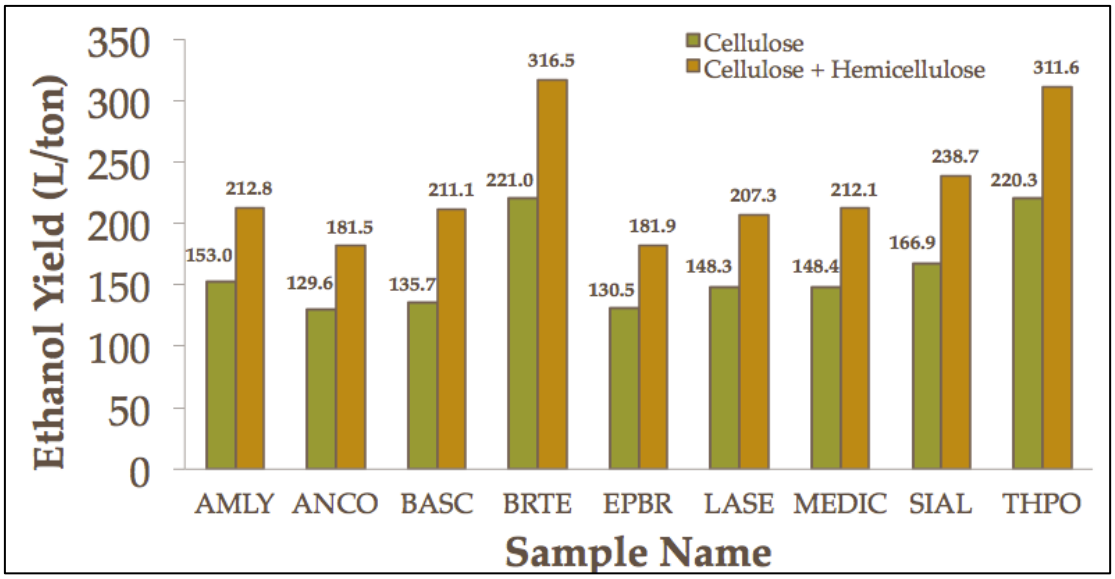


Figure 4: Estimated Ethanol Yields From Biomass Samples Using 5 and 6 Carbon Sugars – Samples were collected in the Pacific Northwest. Green bars represent ethanol produced when only cellulose is utilized. Yellow bars represent ethanol production when cellulose and hemicellulose are utilized (Kumar et al., 2012).

Based on this increased yield, the average production price for a gallon of ethanol is estimated to drop up to 22% depending on the composition of the feedstock (Kumar et al., 2012). It is this very real potential cost savings that drives this work on pentose sugar utilization.

1.4 Five-Carbon Sugar Utilization Strategies

There are three general strategies for optimizing the conversion of five carbon sugars (mainly xylose) from lignocellulosic hydrolysates to fuel ethanol.

1) Modification of a known five carbon fermenter to produce ethanol, increase product yield, increase conversion speed, or increase ability to ferment 6 carbon sugars (Agbogbo and

Coward-Kelly, 2008; Agbogbo et al., 2006; Delgenes et al., 1996; Du et al., 2010; Verduyn et al., 1985)

2) Modification of a known ethanol producing, six-carbon fermenter to be able to directly utilize xylans.(Johansson and Hahn-Hägerdal, 2002; Kötter and Ciriacy, 1993; Kötter et al., 1990; Richard et al., 2006; Träff et al., 2001; Wang and Schneider, 1980)
3) Conversion of five carbon sugars into forms that is more easily utilizable by a given ethanol producing organism (Gong et al., 1981; Liu et al., 1996; Rao et al., 2008; Silva et al., 2012)

1.4.1 Direct Utilization: Uptake of Xylose and Fermentation To Ethanol

D-xylose is the largest component of hemicellulose and the second most abundant sugar on earth. It is no surprise then that many microorganisms have evolved pathways to break down and utilize xylose in their metabolisms. *Pichia stipitis* has been identified as a good candidate for the production of cellulosic ethanol (Agbogbo and Coward-Kelly, 2008; Delgenes et al., 1996).

P. stipitis and most xylose consuming yeasts and fungi, utilize xylose via the same three-step pathway (termed the XR/XDH pathway). Xylose is taken into the cell and converted to xylitol by a xylose reductase (XR). The xylitol is converted to xylulose by a xylitol dehydrogenase (XDH) (Verduyn et al., 1985). This xylulose is finally converted to xylulose-5-phosphate by a xylulose kinase

(XK). Xylulose-5-phosphate is incorporated into the non-oxidative branch of the pentose phosphate pathway where it is converted to pyruvate. In aerobes pyruvate usually oxidized through the citric acid cycle to produce GTP and NADH, but during fermentation pyruvate is converted to acetaldehyde by pyruvate decarboxylase and then ethanol by alcohol dehydrogenase(Sharma and Tauro, 1986). This serves to replenish the stores of NAD+ that act as electron sinks when oxygen is limited.

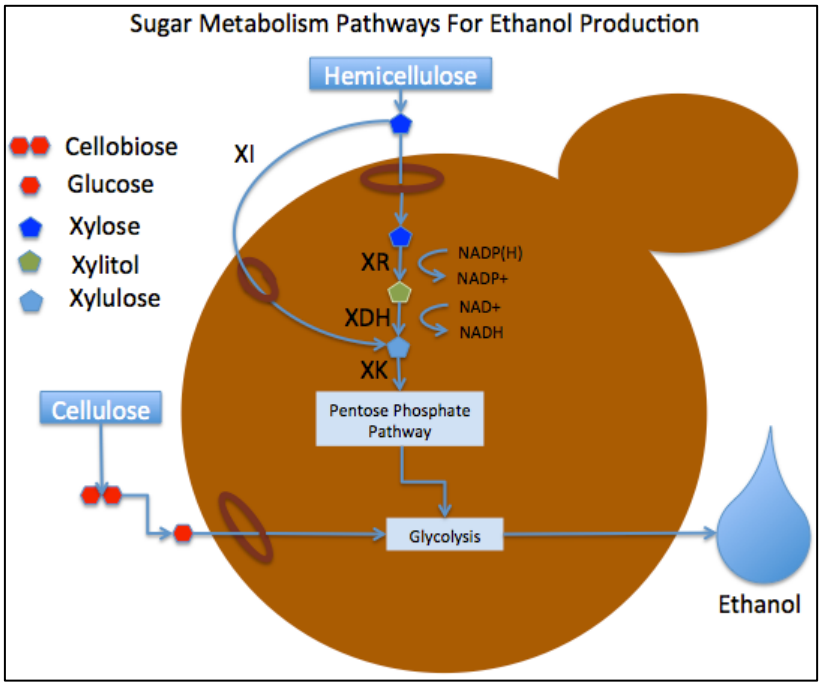


Figure 5: Pentose Utilization Pathways From *P. stipitis* (XR/XDH pathway) and Bacteria (XI pathway) - Here XI is shown as outside the cell to show how an extracellular enzyme could be used to allow *S. cerevisiae* to ferment pentose sugars

Most species that consume xylose in this way have an NADPH dependent XR and a NAD+ dependent XDH. The difference in cofactor preference between the two enzymes often causes cofactor redox imbalance and leads to the preferential accumulation of xylitol (Agbogbo and Coward-Kelly, 2008). *P. stipitis*, however is able to largely avoid this imbalance because its XR has dual cofactor

specificity for NADPH and NADH. In practice, this allows it to metabolize xylose at a sustained high rate, by cycling NAD+ and NADH during the xylose reductase and xylitol dehydrogenase reactions. However, *P. stipitis* is far poorer at fermenting hexose sugars than *S. cerevisiae*. Wild-type *P. stipitis* displays a low specific glucose consumption rate (with a maximum rate approximately 0.3 g/gDCW-h) compared to *S. cerevisiae* has been shown to be as (4 g/gDCW-h) (Agbogbo et al., 2006; Hanly and Henson, 2011; Sonnleitnert, 1986). Additionally, *P. stipitis* requires a tightly controlled oxygen level in order to ferment xylose. A low, but non-zero level oxygen is required for xylose fermentation and ethanol production, but a too high concentration of oxygen will stimulate cell growth (Jeffries and Jin, 2004). Using *P. stipitis* for the fermentation of the hemicellulose stream while fermenting the hexose stream with *S. cerevisiae* is another option, but separate fermentation processes require additional capital investments and produces a more dilute beer that imposes additional distillation costs. Despite the greater yield of ethanol per ton biomass, these factors lead to an increase of production cost of $0.11 per gallon gasoline equivalent when both five and six carbon sugars are utilized in this manner (Kazi et al., 2010).

The cost drawback of separate fermentation has focused research on cloning the genes from *P. stipitis* and other natural xylose fermenters into *S. cerevisiae* in hopes of creating a single organism capable of efficiently fermenting both hexose and pentose sugars. XYL1 and XYL2 (the genes coding for XR and

XDH in *P. stipitis*) were first expressed in *S. cerevisiae* in 1985 (Kötter et al., 1990). This pair of enzymes works well in P. *stipitis* in part because it requires a balanced cofactor pair to operate. XR required NADH to convert xylose to xylitol. This reaction also reduced NADH to NAD+, which was required by XDH to convert xylitol to xylulose (see Figure 6).

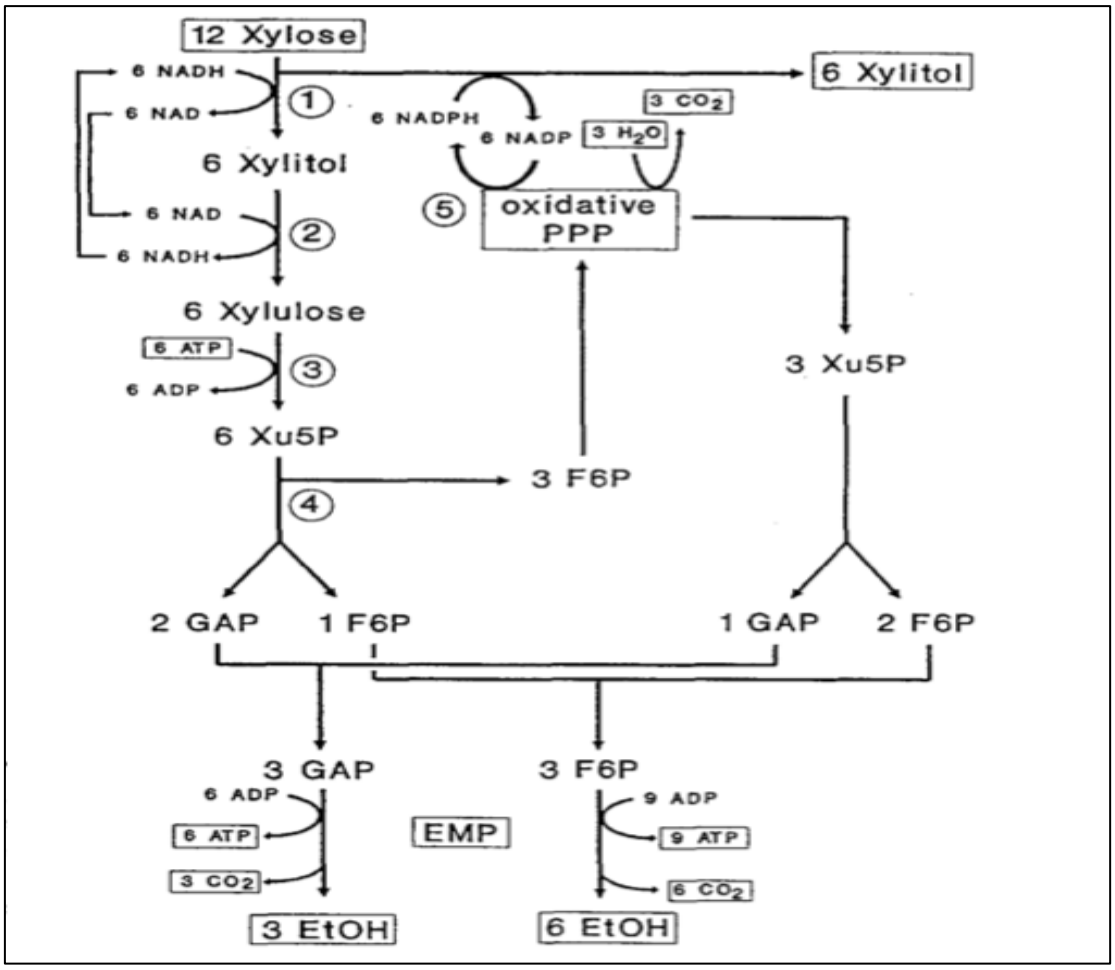


Figure 6: Scheme of Xylose Utilization and Mechanism For Cofactor Regeneration in *S. cerevisiae* - 1=XR (xylose reductase); 2=XDH (xylitol dehydrogenase); 3=XK xylulokinase); 4=ribulose-5-phosphate epimerase, ribose-5-phosphate isomerase, transaldoase, and transketolase; 5= glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase. (Kötter and Ciriacy, 1993)

While the modified strain was able to metabolize xylose, its utilization rate was slow and it appeared to funnel the majority of xylose through an oxidative pathway to produce xylitol and regenerate NADP+ (Kötter and Ciriacy, 1993). By

utilizing this secondary pathway to produce xylitol, *S. cerevisiae* may have uncoupled the cofactor balance between XR and XDH. This uncoupling and resulting cofactor imbalance could have led to the observed low utilization rate for xylose. Recent dynamic flux balance modeling has provided support for this theory by showing that cofactor imbalanced reaction pairs (XR/XDH) produce a slower xylose utilization rate and oxidative regeneration of NADP+ similar to that observed by Kotter (Ghosh et al., 2011).

In an attempt to avoid the cofactor intensive XR/XDH pathway all together, some researchers have focused on the bacterial xylose isomerase pathway. The gene encoding the xylose isomerase enzyme (xylA) has been isolated from many bacterial species including *Actinoplanes missouriensis*, *Bacillus subtilis*, *Clostridium thermosulfurogenes*, *Escherichia coli*, and *Lactobacillus pentosus.* Early attempts to incorporate these enzymes in *S. cerevisiae* failed to produce active proteins. The enzyme from *Thermus thermophilus* was expressed in its active form (Walfridsson et al., 1996), but had only modest activity in the yeast as the optimum temperature for the enzyme activity (80°C) was far higher than the optimum temperature for the yeast growth (~30°C). A different xylA from *Piromyces* sp. strain E2[i] with an optimum temperature near 30°C had a strong xylitol inhibition resulting in low xylose utilization (Karhumaa et al., 2007). Research to overcome xylitol inhibition and increase ethanol production is ongoing, but focuses on codon bias matching and removal of enzymes

responsible for xylitol production (specifically GRE3) (Ha et al., 2011; Lönn et al., 2003; Träff et al., 2001).

1.4.2 Direct Utilization: Uptake of Xylose and Conversion To an Alternative Product

The xylose isomerase pathway is less hampered in its native microorganisms, but the metabolisms of these organisms rarely produce ethanol in high quantities. Three of the common metabolic products produced by these organisms include butanol (Jurgens et al., 2012; Lee et al., 2008; Mu et al., 2011), hydrogen (Cheng et al., 2011), and acetic acid. Butanol and hydrogen can be used directly as fuels, but acetic acid must first be catalytically upgraded to ethanol or jet fuel before it can be used. These methods of fuel production have different challenges and will not be discussed at length here.

1.4.3 Indirect Utilization: Extracellular Conversion of Xylose to a More Usable Form, Uptake, and Fermentation to Ethanol

Xylose isomerase can be used in vitro to convert xylose to xylulose for direct utilization of xylulose. Gong et al. (1981) demonstrated that extracellular xylulose could be utilized with greater than 80% efficiencies by *S. cerevisiae*. In vitro conversion of xylose to xylulose is attractive for biofuels production as the enzyme xylose isomerase (sometimes called glucose isomerase) is a well-established (widely available and relatively inexpensive) industrial enzyme used in the production of high fructose corn syrup.

While the promise of indirect xylose utilization pathway is immense, but several hurdles remain for use of this strategy in commercial biofuels production.

1) Production of concentrated xylulose stream (greater than 50g/L):

Perhaps the most obvious challenge is the production of concentrated xylulose. The isomerization of xylose is strongly hampered by product inhibition. The isomerization of pure xylose in the presence of glucose isomerase and an $MgSO_4$ cofactor has been shown to produce xylulose yields of roughly 16% purity (Hochster and Watson, 1954). Assuming a feedstock with 30% hemicellulose comprised of 60% xylose, and a pretreatment solids loading at 10% w/w, 16% conversion efficiency would yield a xylulose concentration of less than 3 grams per liter (a full order of magnitude less than the concentration of the six carbon sugar stream). The conversion rate could be improved by the addition of borate (as sodium tetra borate). Borate removes xylulose from solution by complexing with it allowing for further product formation. By this method purities of up to 80% have been achieved (Hsiao et al., 1982). However borate has been shown to inhibit fermentation in yeasts (Lochhead and Farrell, 1930). This inhibition is likely caused by interfering with its alcohol dehydrogenase activity (Smith and Johnson, 1976). Since ethanol is the desired product, this inhibition is problematic limiting the usefulness of borate.

2) Mismatch between the optimal temperature and pH of *S. cerevisiae* and xylose isomerase.

If yeasts and xylose isomerase could function in the same environment, removal of xylulose via consumption by yeast would limit product inhibition for xylose isomerase and allow further conversion of xylose to xylulose. However, *S. cerevisiae* ferments optimally at a slightly acidic pH near 5.0 and a temperature near 25°C. Even in thermophilic industrial yeast strains, growth is severely inhibited at temperatures at or above 50°C (Lin et al., 2012). Conversely, commercially available glucose isomerases (such as GENSWEET) act optimally at a mildly basic pH of 7-8 and a temperature in excess of 70°C. Research aimed at isolating xylose isomerases with a more acidic optimal pH (Liu et al., 1996) and the fixation of glucose isomerase in pH microenvironments (Jeppsson et al., 1996; Rao et al., 2008; Silva et al., 2012) to allow simultaneous isomerization and fermentation in a single vessel at two different pHs is ongoing and could help increase the yield of ethanol from this strategy.

3) Xylulose utilization rate in *S. cerevisiae*.

Traditionally, pentose utilization rates have been shown to be one to two orders of magnitude slower than glucose utilization in modified strains of *S. cerevisiae*. To make co-fermentation of hexose and pentose sugars viable utilization rates must be more closely compatible. Directed strain modification revolves around the identification and widening of pathway bottlenecks. In

accordance with this need our work will focus on better understanding of the enzymatic and transport bottlenecks in *S. cerevisiae* through bench scale fermentation experiments and subsequent flux balance modeling.

1.4.3.1 Enzymatic Limitation

Competing research groups have focused on different possible bottlenecks in the xylulose utilization pathway. Much of the work has focused on the improvement of utilization rates by over expressing xylulose kinase (Chang and Ho, 1988; Lee et al., 2003; Richard et al., 2006). This has resulted in a 10-fold increase in ethanol production rate from 0.04±0.02 to 0.23±0.05 g/gDCW-h (Lee et al., 2003). This is still significantly lower than the reported ethanol production rate from glucose (4 g/gDW-h) (Sonnleitnert, 1986). Other work has identified a different set of bottlenecks in the pentose phosphate pathway (PPP). When overexpressing the non-oxidative PPP enzymes ribulose 5-phosphate epimerase, ribose5-phosphate keto lisomerase , transaldolase, and transketolase these groups showed up to a 2.4-fold xylulose utilization rate increase from 0.09 to 0.22 g/gDCW-h (Johansson and Hahn-Hägerdal, 2002; Matsushika et al., 2012).

1.4.3.2 Transport Limitation

The transport of xylulose is not well understood in *S. cerevisiae*, but it may also play a role in the limiting xylulose utilization the pathway. The transport of xylose has been more extensively studied (Hotta et al., 2009; Runquist et al.,

2010; Sedlak and Ho, 2004; Tanino et al., 2010). Xylulose transport may use similar transporters and transport mechanisms to its isomer xylose.

Yeasts utilize two main types of sugar transport proteins, namely passive diffusion channels (sometimes called facilitators) and symporters. Most sugars are relatively large, polar, hydrophilic, uncharged molecules and therefore do not pass through the phospholipid cell membrane by simple diffusion as hydrophobic molecules such as O_2, CO_2, steroids, and hormones (Figure 7). In facilitated diffusion, a passive, water filled channel (facilitator) is imbedded in the cell membrane allowing more hydrophilic molecules enter the cell. Solutes utilizing facilitated diffusion must flow down a concentration gradient (from high concentration to low concentration).

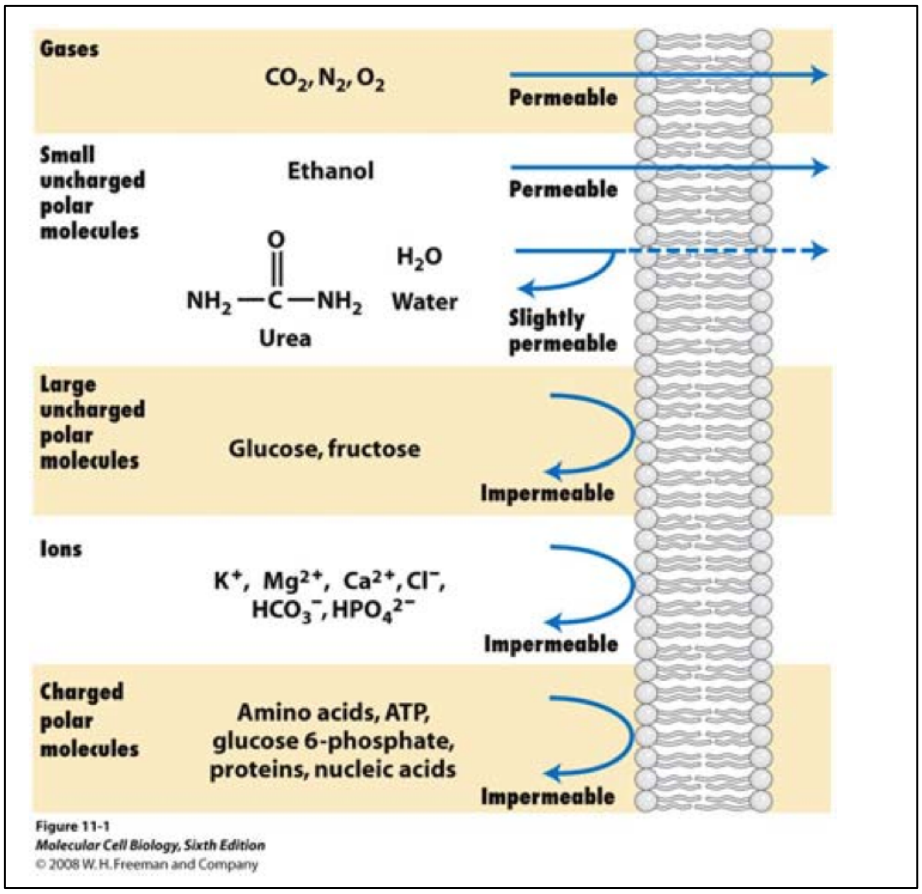


Figure 7: Permeability of the Phospholipid Bilayer to Different Molecules – the diagram examines molecules of different sizes, polarities, and charges

On the other hand, symporters can concentrate a solute by pumping the solute against its concentration gradient (Figure 8). This transport is powered by coupling the concentration of one solute with the dilution of another. In proton/xylose symporters a proton powers the concentration of xylose when it is pumped down its electrochemical gradient. Other symporters utilize the potential energy stored in other ion concentration gradients (notably potassium and sodium gradients), but proton symporters are most common in yeast sugar transport.

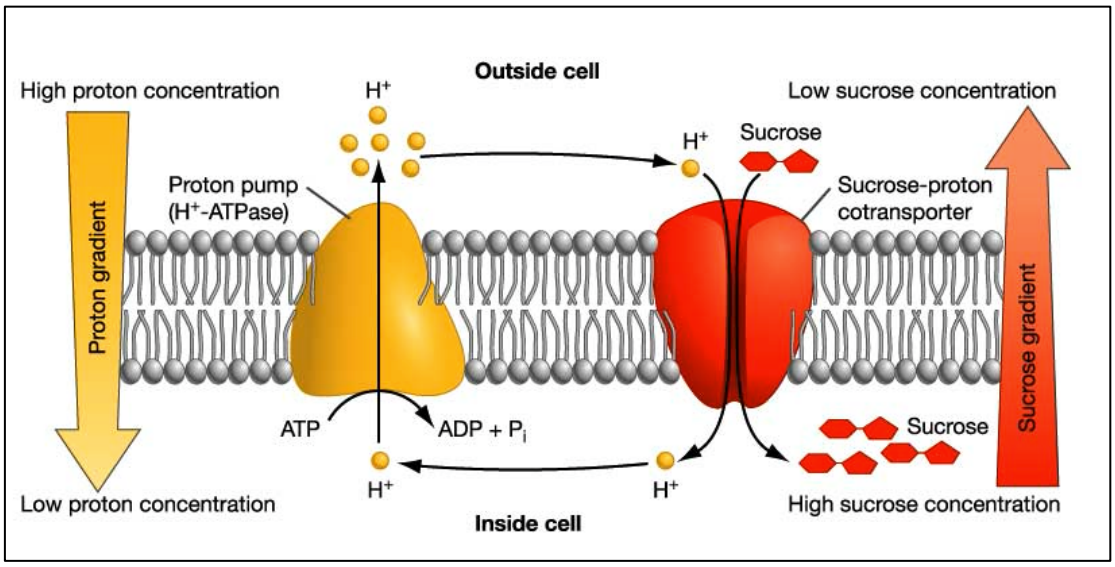


Figure 8: Symport and ATP Driven Transport Across the Cell Membrane - Cells can use ATP fueled proton pumps to create a proton gradient. This gradient can be used with a symporter to transport another molecule (in this case sucrose) against its concentration gradient.

Natural xylose fermenting organisms such as the yeasts *P. stipitis, and C. intermedia,* and the mold *N. crassa*, have transport proteins with relatively high specificity for pentose sugars (Du et al., 2010; Runquist et al., 2010). Most of these transport proteins are not substrate specific, but they often have a high affinity for pentose sugars meaning that they are saturated at low pentose concentrations. Sut1, the nonspecific sucrose/xylose proton symporter identified in *P. stipitis* has been expressed in *S. cerevisiae* leading to an increased rate of xylose transport. Even higher xylose uptake rates were obtained when GXF1 (Table 2 and Figure 9), a glucose/xylose facilitator found in *C. intermedia)* were expressed in *S. cerevisiae.*

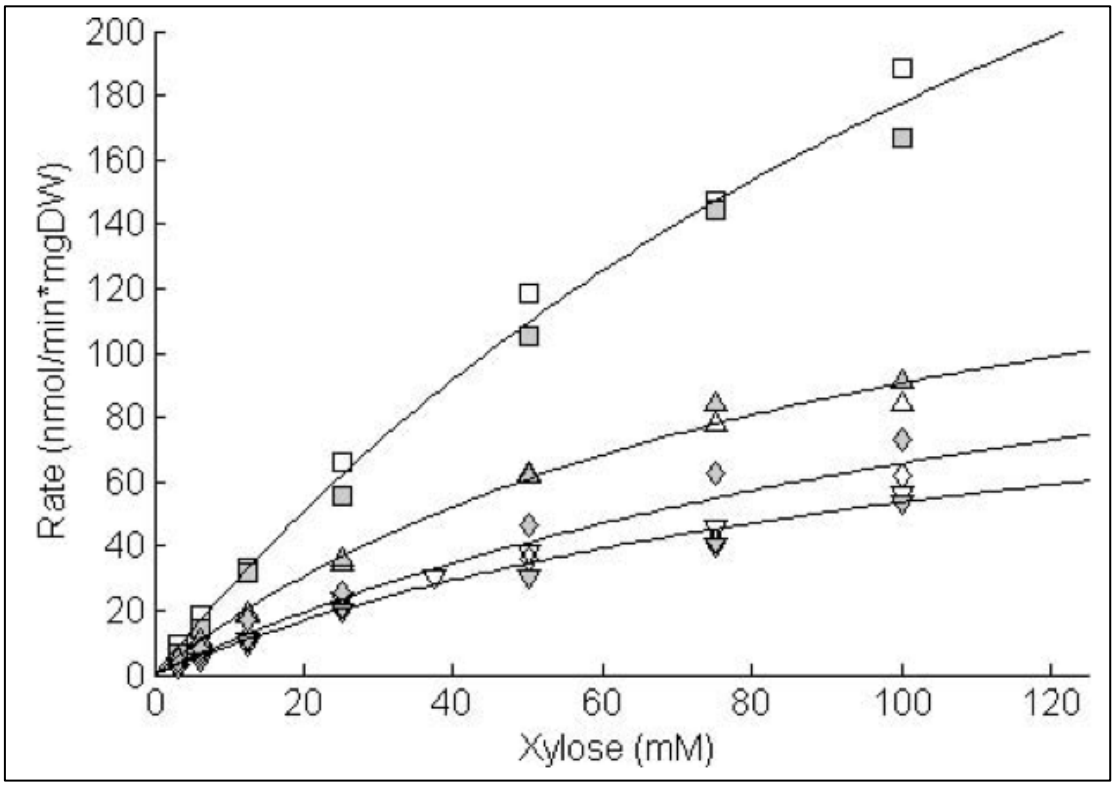


Figure 9: Xylose Uptake Rates of *S. cerevisiae* Strains Expressing Different Transporters - Lines represent the calculated Michaelis-Menten fits. Gxf1=square; Sut1=upward pointing triangle; At5g59250=diamond; control=downward pointing triangle (Runquist et al., 2010).

Table 2: Apparent Kinetic Constraints for Xylose Transport in *S. cerevisiae* - Strains were grown on 20 g/L glucose (Runquist et al., 2010).

	K_m (mM)	V_{max} (nmol/min x mgDW)	V_{max}/K_m (min^{-1} x $mgDW^{-1}$)
TMB 3415 (control)	121±44	119±17	$10x10^{-7}$
TMB 3416 (Gxf1)	166±6	471±10	$28x10^{-7}$
TMB 3418 (Sut1)	96±11	178±16	$19x10^{-7}$
TMB 3419(At5g59250)	148±39	163±7	$11x10^{-7}$

While each of the transporters discussed above transports xylose to some extent, they are all non-substrate-specific. To identify transporters with a singular, xylose specificity researchers used the sequence encoding the protein GXS1 (a transporter with high D-xylose affinity) from *C. intermedia* as a probe in BLAST (the National Center for Biotechnology's Basic Local Alignment Search Tool) to identify transporters likely to show high xylose affinity. This work lead to the discovery of two xylose specific transporters in *N. crassa*: An25 and Xyp29 (Du et al., 2010).

As *S. cerevisiae* is not a natural pentose utilizing species, it is unlikely to have pentose specific transporters and does not have any known proteins with homology matching to known specific pentose transporters. Instead it likely utilizes transporters optimized for hexose transporters to transport xylose under glucose-limited conditions. This is supported by studies that demonstrated the loss of xylose utilization capacity in manipulated *S. cerevisiae* strains (XR and XDH added) with deletion of 18 hexose transporters (Hamacher et al., 2002).

Research has shown that over expression of the Hexose Transporter (HXT) family of proteins (the major family of putative hexose transporters in *S. cerevisiae*) increases xylose utilization in strains containing the XI pathway (Hotta et al., 2009; Tanino et al., 2012). Other studies have suggested that the galactose transporter family (GAL), and specifically the protein GAL2, could play a role in xylose (and perhaps xylulose) uptake (Sedlak and Ho, 2004). Transport protein expression is often strongly linked to the presence or absence of a specific sugar.

The expression of the transport proteins is often not controlled directly through the presence or concentration of the sugar it transports, but is rather linked via a signaling (phosphorylation) cascade. In the case of the HXT family of proteins, glucose levels are sensed by a separate protein (often SNF3), which sets off the signal cascade leading to the differential expression of different high and low affinity transporters (Figure 10).

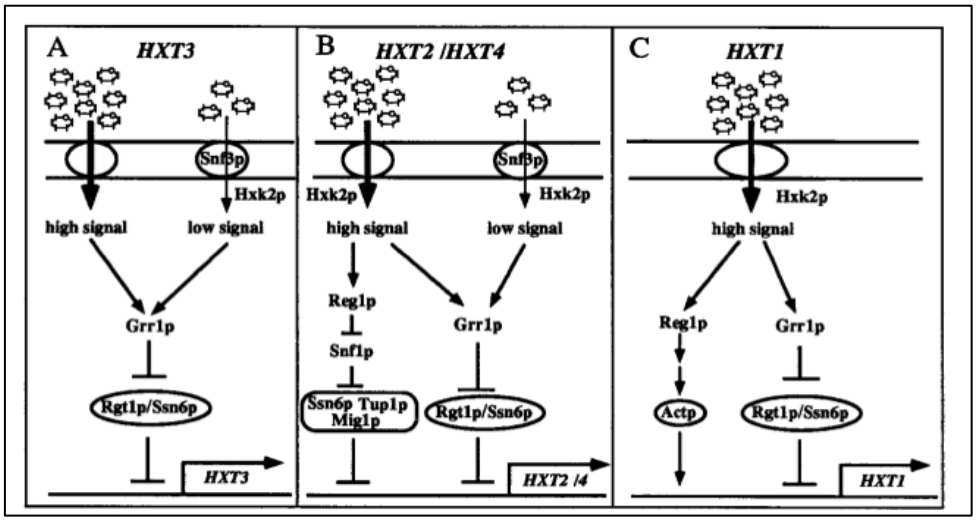


Figure 10: Representation of Three Mechanisms of Induction by Different Glucose Levels - A line with an arrowhead implies positive regulation; a line with a bar denotes negative regulation A=HXT3; B=HXT2 and HXT4; C=HXT1 (Ozcan and Johnston, 1995).

The expression of the HXT family of transport proteins is highly correlated to glucose concentration. Ozcan et al. mapped the expression levels of HXT1, HXT2, HXT3, and HXT4 as a function of glucose concentration and showed that they had different expression patterns (Ozcan and Johnston, 1995). HXT3 was expressed at all glucose levels (Figure 11), while HXT1 was expressed only at high glucose concentration and HXT 2 and 4 were expressed only at low glucose concentration(Ozcan and Johnston, 1995). Other studies have show HXT6 and 7 to be expressed at only at low glucose levels (below 10 g/l) (Diderich, 1999; Liang and Gaber, 1996).

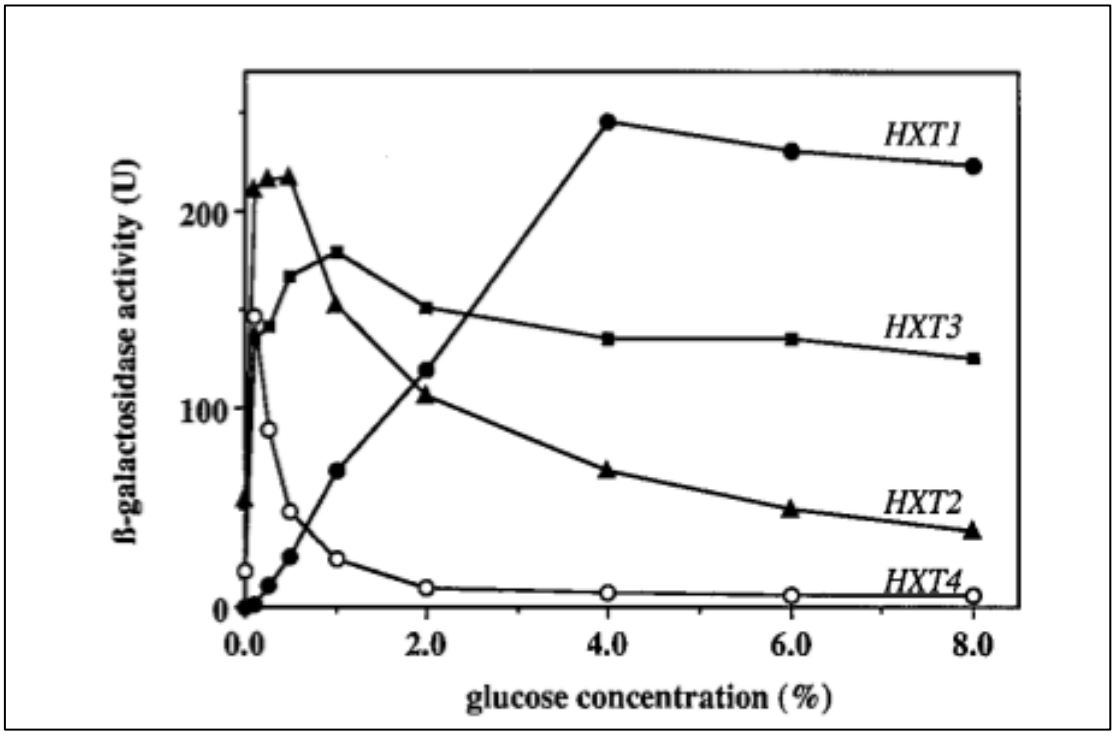


Figure 11: Induction of HXT Gene Expression as a Function of Glucose Concentration - HXT1=closed circle; HXT2=triangle, HXT3=square; HXT4=open circle (Ozcan and Johnston, 1995).

Protein expression level is an important part of determining the rate of a reaction or transport, but it will not provide an absolute transport rate. Michaelis-

Menten kinetic parameters including the maximum throughput (V_{max}) and half saturation constant (K_m) must be estimated before a rate estimate can be determined. These parameters have been calculated for each of the transporters HXT1-HXT7 for consumption of both glucose (Reifenberger et al., 1997) and xylose (Saloheimo et al., 2006). Based on the physiochemical similarity between xylose and xylulose, the kinetics describing their transport through the HXT family of proteins may be similar.

1.4.3.3 Use of Modeling to Understand Xylulose Utilization

The relationship between transport bottlenecks and internal, enzymatic bottlenecks are complex. The understanding of these relationships may best be elucidated through modeling. The study of xylose transport and utilization has been supported through kinetic modeling, flux balance analysis (FBA), and dynamic flux balance analysis (dFBA). For example, Pitkänen et al. was able to use a simplified recombinant *S. cerevisiae* network demonstrate the mechanism behind (and the consequences of) the redox imbalance created under conditions of xylose fermentation using the XR/XDH pathway (Pitkänen et al., 2003). This work confirmed and expanded the redox imbalance which had been identified years earlier (Bruinenberg et al., 1983; Kötter and Ciriacy, 1993). Bertilsson et al. focused on the transport of xylose through the native *S. cerevisiae* hexose transporter family (HXT family). Using a system of ODE's based on the known transport kinetics of the individual HXT transporters; the model was able to accurately represent the consumption of sugars in mixed xylose/glucose

fermentation (Bertilsson et al., 2007). Ghosh et al. expanded upon the work of both Pitkänen and Bertilsson's work by considering the ability of a hypothetical strain of *S. cerevisiae* with balanced cofactors to take utilize xylose. The strain was modeled using a combination of dFBA and kinetic modeling techniques and suggested a potential ethanol yield increase of 24.7% over a strain without balanced cofactors (Ghosh et al., 2011). Similar techniques could help to characterize the mechanism behind xylulose transport and identify opportunities for successful genetic engineering to increase ethanol production via the xylulose isomerase pathway.

1.5 Flux Balance Analysis

Flux balance analysis is a type of constraint-based modeling that uses physical or chemical constraints to define all allowable sets of fluxes through a network (Orth et al., 2010). Constraint-based analysis assumes that the system is at steady state (i.e. that it is not accumulating metabolites and fluxes are constant) and that it is satisfying all phyisio-chemical constraints (reaction rates are within allowable ranges and metabolites used in all reactions are actually available). In any single analysis there may be dozens or thousands of "allowable solutions". To select a relevant solution an optimization is carried out to maximize or minimize a given variable. The maximization of growth rate and the maximization of ATP production are common optimization criteria in biological systems.

The FBA technique was pioneered in the early 1980s by Papoutsakis working on a stoichiometric calculation for butyric acid bacteria fermentation (Papoutsakis, 1984), but standardized mathematical methods weren't developed until 1991 by Savinell and Palsson (Savinell and Palsson, 1992). As computing power, knowledge about metabolic networks, and high throughput genome sequencing technology evolved, the number of reactions captured in flux balance analyses grew as well.

To better explain the concepts underlying basic FBA analysis, the steps are outlined below and illustrated in Figure 12.

1.5.1 System Definition and Boundary

Basic flux balance modeling requires a well-defined system that encompasses all relevant metabolites (A, B, C in Figure 12) and the stoichiometrically defined reactions for their inter-conversion and manipulation (*v1-v4* in Figure 12). The advent of genome sequencing and automated genome annotation provides a convenient starting point from which to build this network. Automated genome annotation looks at sequence data to identify known proteins that the cell may be capable of producing. Once the enzymes are identified, they are annotated with information about the reaction that they catalyze including the identity and stoichiometric coefficients for the products and reactants in the reaction.

Often during this annotation process, reactions in pathways known to exist in an organism may not be identified. This can be because of low genome

sequencing coverage or failure of the genome annotation software to correctly identify protein sequences. To remedy this, unidentified reactions are added to the model in a process known as gap filling. In addition to internal reactions, transport reactions must be considered (*b1-b3* in Figure 12). Some metabolites are transported by simple diffusion while aquaporins, symporters, and ATP driven active transport pumps transport others.

Here it is important to note that no implicit regulation is incorporated in this model. Within the cell, only a small fraction of the proteins required for the reactions defined thus far may be available at any one time. Traditional FBA analysis forgoes attempting to impose proteomic regulation on the system and instead uses mathematic modeling techniques to predict the optimal flux through a given pathway.

1.5.2 Mass Balance

After the list of all reactions and metabolites are compiled, a framework for global mass balance is derived. This framework is composed of a series of ordinary differential equations with one equation representing the stoichiometric flux of all metabolites through a reaction. When compiled, this list of ODEs can be represented in matrix notation. Here ‘S’ is an m x n matrix that represents the stoichiometry of all of the reactions in the cell. Each row in ‘S’ is an individual reaction and each column is an individual metabolite thus the number in element (m,n) is the stoichiometric amount of metabolite ‘n’ associated with reaction ‘m’. ‘V’ represents the molar flux through each reaction. When the elements in ‘V’ are

such that S•V is equal to zero all metabolites in the system are mass balanced and the steady state assumption is met.

1.5.3 Model Constraint

The framework to fulfill the steady state assumption provides the first set of constraints for the model, but because the number of reactions is greater than the number of metabolites the solution set for reaction fluxes remains unbounded. In order to find a unique solution set, further constraints are required. These constraints are formed by identifying measurable fluxes within the network and defining flux rates (values of elements in 'V') for those reactions. Often, exact flux rates are unknown or vary with environmental conditions and an allowable range of fluxes will be defined in the model. Additional constraints can also be imposed based on experimental insights, reaction thermodynamics, or hypothetical situations to be tested (for example a gene knockout).

1.5.4 Identifying Optimal Solution

With constraints in place to produce an allowable solution set, the optimal solution can be identified by employing a linear mathematical solver to optimize a biologically relevant objective function. The objective function, termed 'Z', is most often a weighted combination of reaction fluxes. In matrix notation Z=C•V where 'C' is the weight applied to each of the fluxes defined in 'V'. Commonly used objective functions maximize biomass growth, ATP production, or the synthesis of a specific metabolite.

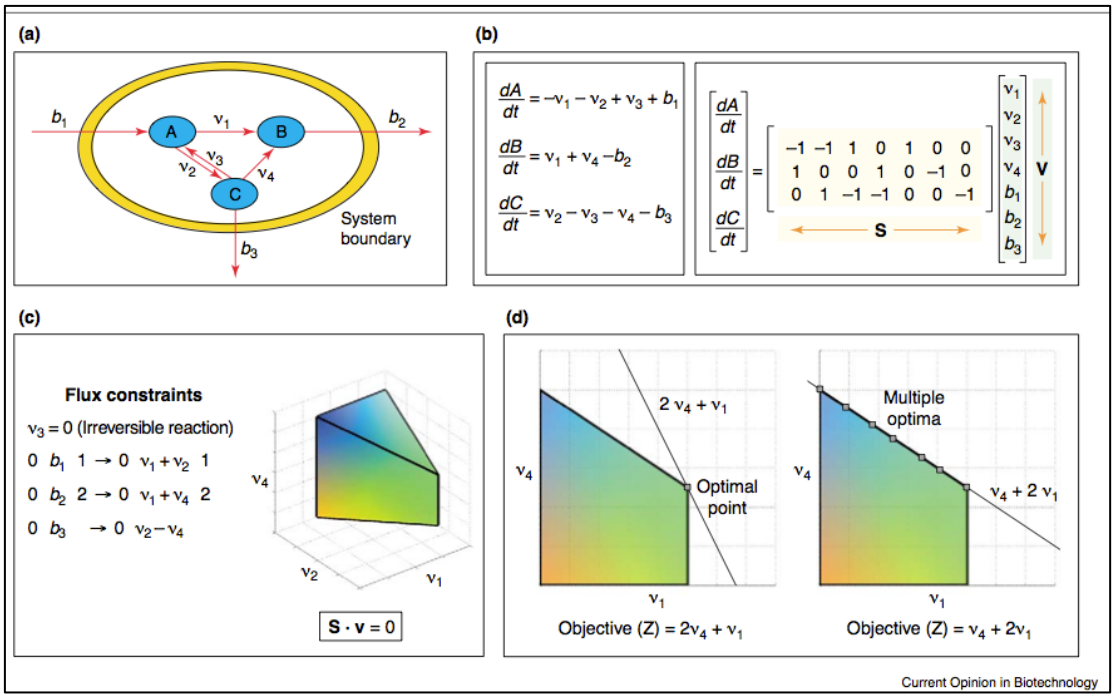


Figure 12: Methodology for Flux Balance Analysis - (a) A model system comprised of 3 metabolites (A, B, and C) with three reactions (including one reversible reaction), and three exchange fluxes *b*. (b) Mass balance equations accounting for all reactions and transport mechanisms are written for each species. These equations are then rewritten in matrix form. At steady state, this reduces to S•V=0 (c) The fluxes of the system are constrained on the basis of thermodynamics and experimental insights. This creates a flux cone corresponding to the metabolic capacity of the organism. (d) Optimization of the system with different objective functions (Z). Case I gives a single optimal point, whereas case II gives multiple optimal points lying along an edge (Kauffman et al., 2003).

1.6 Building Upon the Traditional FBA: Second Generation FBA Techniques

1.6.1 Time Course Experiments; dFBA

Traditional FBA captures only a snapshot of the rate of reaction fluxes in a cell or cell population. Varma and Palsson expanded the method in 1994 as to be able to capture a time course experiment, a technique known as dFBA (dynamic flux balance analysis) (Varma and Palsson, 1994). In this technique the time course was divided up into small time intervals. Initial conditions were provided for the first time step and an FBA was carried out producing the vector 'V'

containing all of the fluxes within the cell and the fluxes through transport reactions by which the cell interacts with the environment. Varma assumed that over the small time step the optimal flux through the cell was likely to remain constant. By multiplying the rates of flux through the transport reactions by the length of the time step, Varma could estimate the changes in media composition and use the new composition as initial values for the next FBA iteration. This process could be repeated indefinitely providing a projection of the system.

1.6.2 New Classes of Objective Function; MOMA and ROOM

Traditional FBA predicts the optimal, steady state fluxes within a cell under the assumption that the cell has evolved to favor growth. However, in genetically modified organisms that have acquired or lost genes without prolonged evolutionary pressure and in organisms cultured in pressures unlike those that shaped their evolution, this assumption does not hold true. For example, a model of *E. coli* growth on glycerol (a suboptimal substrate) was shown to match experimental measurements after 40 days of culturing in a glycerol medium, but not at the start of culture (Ibarra et al.). Traditional objective functions did not consider the cost associated with the cell's metabolomic reorganization and were not designed to identify sub optimal solutions. Minimization of Metabolic Adjustments (MOMA) and Regulatory On-Off Minimization (ROOM) constitute a second generation of optimization functions (for use with GMOs) that target solutions with sub-optimal flux through a given pathway while minimizing metabolic changes.

Segre et al. first introduced MOMA in 2002 in an attempt to better model the growth of an *E. coli* knock out mutant. Segre found that the wild-type strain's growth rate closely matched that predicted by traditional FBA analysis (correlation 0.78 to .097). In the knockout strain, the correlation between traditional FBA prediction and experimental data was lower (-0.064 to 0.86). MOMA was based on the assumption that a cell would minimize its adaptation cost after a knockout or change in conditions and that it would do this by minimizing flux changes that would require new protein synthesis. To this end the objective function in MOMA is set to minimize the Euclidean distance between each expected wild-type flux and its corresponding knockout flux. That is $\sqrt{(\Sigma (\Delta V)^2}$) is minimized (Figure 13). Using this objective function produced a suboptimal growth solution that more closely matched that of the knockout experimental data (correlation 0.56-0.94) (Segre et al., 2002).

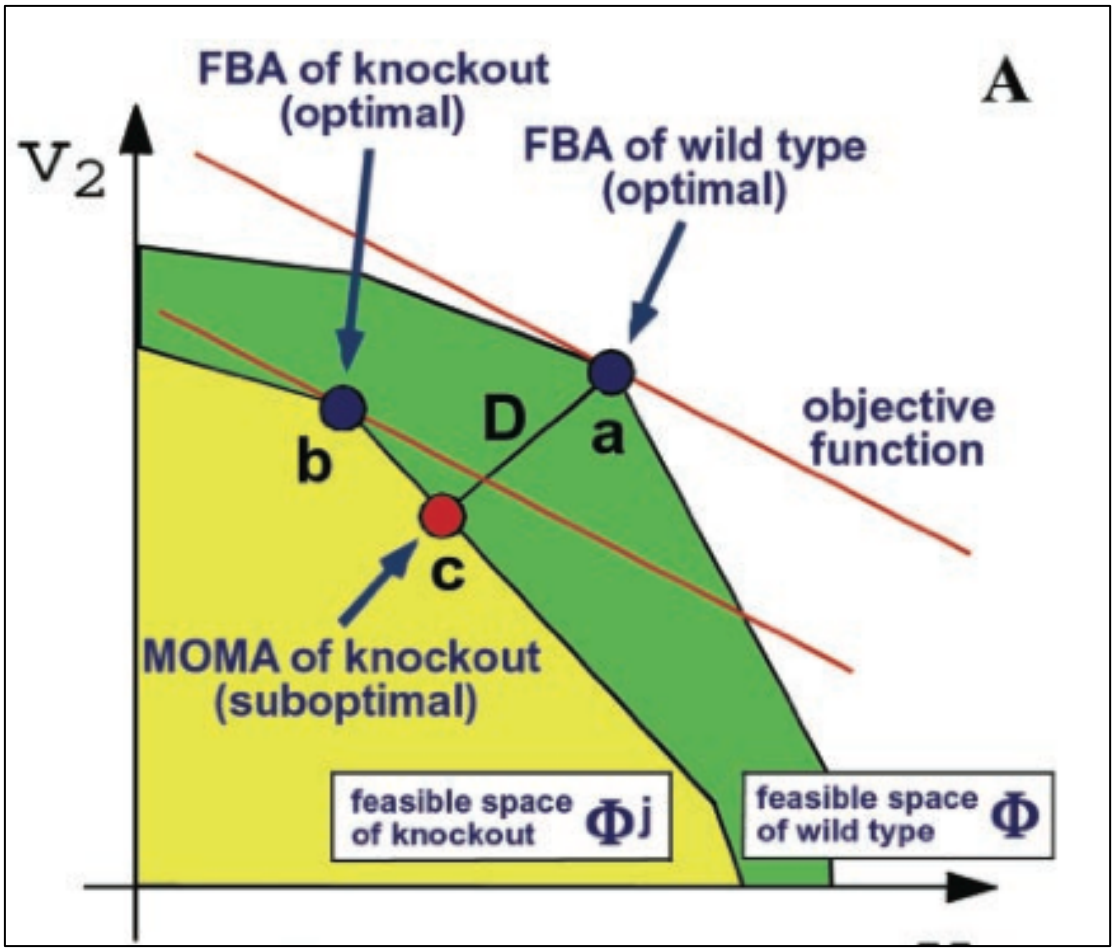


Figure 13: Representation of MOMA Optimization - In MOMA the model attempts to minimize the change in fluxes from a wild-type case (Segre et al., 2002).

MOMA works on the theory that cells metabolisms will avoid all change at the expense of growth and that it will oppose one large change more than several smaller changes. However, other studies have shown that these types of major re-routings do occur to allow for increased flux through an alternative pathway away from the knocked out enzyme (Emmerling et al., 2002). ROOM was developed as an alternative to MOMA in 2005 by Shlomi et al. and was based on the same assumption that cells would minimize their adaptation costs. ROOM differs from MOMA by assuming that the cells major regulatory changes can be

described by Boolean on/off dynamics (changes below a set threshold magnitude are not considered) and that each regulatory change has a fixed cost regardless of its magnitude (Shlomi et al., 2005). This assumption favors a few large magnitude changes over several smaller magnitude changes.

ROOM growth rate estimations have been shown to be similar to FBA estimations both in terms of lethality prediction and growth rate predication. Growth rate measurements for FBA and ROOM tend to be less well correlated with experimental knockout growth data before an adaptation period as compared to MOMA (correlations: 0.772, 0.777, 0.834 respectively), but more well correlated after adaption (correlations: 0.724, 0.727, 0.658 respectively). The lethality prediction data below shows the improved lethality prediction over MOMA (Shlomi et al., 2005).

Table 3: Comparison of Lethality Predictions Using Traditional FBA, MOMA, and ROOM Methods - (Shlomi et al., 2005).

	FBA	MOMA	ROOM
True-positive	449	399	449
False-positive	64	60	62
True-negative	23	27	25
False-negative	19	69	19
Positively predicted genes	96%	85%	96%
Negatively predicted genes	26%	31%	29%
Overall Prediction	85.0%	76.7%	85.4%

In quickly evoloving field of FBA, new formulations of objective functions are regularly published. Growth or product optimization, MOMA, and ROOM constitute the most widly accepted objective functions, but none can yet accurately model all aspects of cell metabolism. As recently as 2013 a new objective formulation termed PSEUDO (Perturbed Solution Expected Under Degenerate Optimality) was released. It is based on the hypothesis that the cell might reorganize its metabolism such that a minimum percentage (90%) of the wild-type growth rate is maintained and has demonstrated more accurate flux redistributions in recombinant *E. coli* (Pearson correlation coefficients of 0.86, 0.84 and 0.91 respectively for FBA, MOMA, and PSEUDO). Additionally, average prediction errors in the PPP and TCA pathways were significantly better using PSEUDO (-17% average erro) versus traditional FBA and MOMA (-41% and -42% respectively) (Wintermute et al., 2013).

1.6.3 Regulatory Incorporation; rFBA and SR-FBA

While incremental increases in prediction accuracy through modified objective functions continue, the trend has now tended towards incorporating regulatory mechanisms within the model as additional constraints to improve model predictions. Traditional FBA and dFBA (including those using alternative objective functions) assumed that all gene products were available at any time and that given the correct reactants, any reaction could happen at any time. This leads to incorrect predictions when the organism's internal regulatory mechanisms are a dominant influencer of behavior. Reactions catalyzed by

enzymes produced only in the presence of certain metabolites or under certain stresses will not be possible without the appropriate stimuli. Several methods for incorporating native intracellular regulation have been proposed

In 2001 Covert, Schilling, and Palsson developed a framework for transcriptional regulation in *E. coli* based on the RegulonDB, a database of transcriptional regulation and operational regulation in *E. coli (Covert and Palsson, 2003; Covert et al., 2001).* RegulonDB contains curated information on gene organization and regulatory mechanisms for each gene including the sigma factors, promoters, terminators, and regulons required for transcription. This framework, later termed rFBA, used Boolean (on/off) logic operators (ex. IF, AND, OR, NOT) to determine a set of "active" reactions (and a set of "inactive" reactions) based on the metabolic profile of the previous time step of the dFBA. If a protein required for a reaction was not "active," the flux through that reaction was limited to zero this constituted a second level of adjustable constraints to compliment the static, physicochemical inherent in the FBA.

Ruppin and colleagues proposed a similar method, Steady-state Regulatory FBA (SR-FBA). This method also used Boolean logic operators to derive a second level set of adjustable constraints, but Ruppin translated the Boolean operators that mapped the presence of proteins to the availability of reactions into a system of linear equations to solve for a steady state regulatory profile (Shlomi et al., 2007). This allowed the regulatory network to be solved within the MILP problem framework of the FBA.

SR-FBA and rFBA share two main weaknesses. First is that they are based solely on Boolean logic which limits proteins to a binary response pattern (they are either fully on or fully off). Cells generally have much finer protein regulation than this. Second, because these methods select a steady state regulation profile they cannot accurately represent feedback responses and signaling feedback loops happening on different time scales than the FBA (Gonçalves et al., 2013).

1.6.4 Signaling Incorporation; idFBA and iFBA

Two advanced techniques have been proposed that incorporate both protein regulation and cell signaling into FBA. Papin et al. proposed Integrated Dynamic FBA (idFBA) in 2008. This method again relied on Boolean (on/off) operators in the derivation of its quasi-steady-state regulatory network, but provides a stoichiometric rate component to differentiate “fast” reactions (which are included in the quasi-steady-state regulatory profile) from “slow” reactions (which can take minutes to hours to reach equilibrium and are incorporated in a time dependent manner). This setup allows the system to exhibit signaling and feedback phenomena seen in vivo (Min Lee et al., 2008), but does require significant additional knowledge about system in terms of protein synthesis rates.

A different approach is taken in iFBA to better model cell signaling along with regulation. iFBA is based on rFBA using Boolean operators to sort reactions into an “active” and an “inactive” group in each time step based on the environmental factors. In addition, iFBA uses a system of ODEs to depict fine regulation of enzymes and pathways with well-characterized kinetic parameters (Covert et al., 2008). These constraints are combined with the regulatory and invariant, physiochemical constraints that must be satisfied in the FBA.

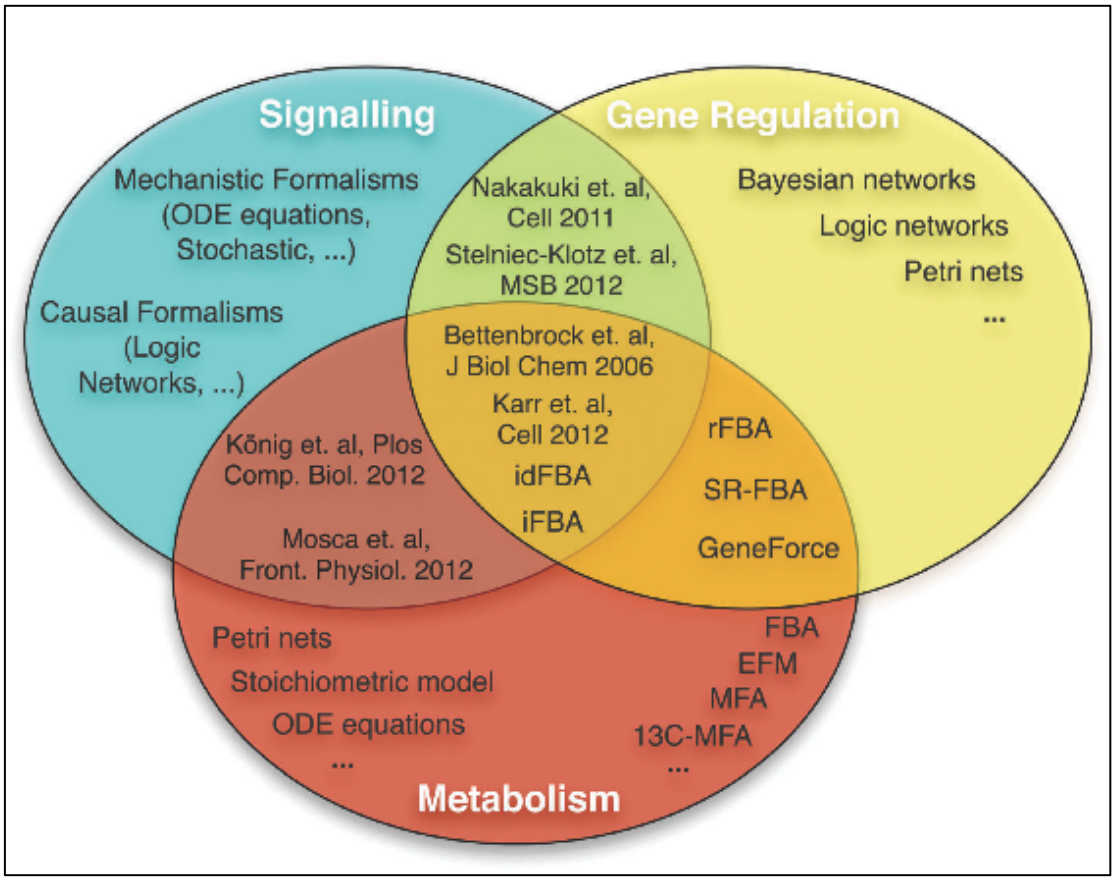


Figure 14: Overview of Formalisms for Modeling Signaling, Gene Regulatory, and Metabolic Networks - Multiple formalisms and simulation methods can be used to model and analyze each biological system. Due to specific biological features, some mathematical formalisms are more suitable for specific systems. Some methods can model different types of systems using different (e.g. SR-FBA) or the same mathematical formalism. Specific references are only used for the cases where a general term is not available (Gonçalves et al., 2013).

1.7 This Work

This work will focus on identifying and modeling the xylulose utilization bottlenecks with specific focus on the transport of xylulose into the cell (in part through the HXT family of transporters) and the conversion of xylulose to xylulo-5-P via the enzyme xylulokinase (XK). A series of batch experiments will be used to determine the kinetic parameters associated with xylulose utilization. These kinetic parameters will be used in a genome scale dFBA to better understand the utilization dynamics in the cell. Regulation will be included for the expression of the HXT family of transporters based on the model of xylose transport previously developed (Bertilsson et al., 2007). The genome scale kinetic modeling will provide insight into metabolic bottlenecks in the xylulose utilization pathway paving the way for directed cell engineering in the future.

Chapter 2 - Batch Fermentation Experiments:

2.1 Introduction:

Ethanol produced from lignocellulosic biomass has the potential to displace the bulk of global, non-renewable transport fuels. However, using current technology, feedstock costs can make up over 40% of total production costs. To justify the high cost of raw materials and to make ethanol economically competitive with fossil fuels it is imperative that the process converts as much of the feedstock sugars to ethanol as possible. While six carbon sugars are readily fermentable by wild-type *Saccharomyces cerevisiae,* five carbon sugars are not. Xylose accounts for roughly 60% of the sugar monomers in hemicellulose and developing a strategy to utilize it efficiently could boost ethanol yields by 40%.

One method to improve xylose utilization is to enzymatically convert it to xylulose. Xylulose can be utilized by wild-type *S. cerevisiae* through the pentose phosphate pathway (PPP). In this process xylulose is transported into the cell, converted to xylulo-5-phosphate (by xylulokinase), and funneled into glycolysis and fermentation through the PPP. To our knowledge no work has attempted to elucidate the full kinetics of this pathway; an understanding of which would improve future strain development aimed at increasing pathway flux.

2.2 Overview of Major Aims:

As a whole the batch fermentation experiments described in this section had three major objectives:

1) To determine the method and rate function of xylulose transport into the cell as a function of xylulose concentration

It was hypothesized that xylulose was most likely transported by the HXT family of proteins, but it was unknown how efficiently these transporters would work for xylulose. It was also unclear if other proteins or simple diffusion participated in this transport process. We examined this aim by looking at the saturation kinetics of xylulose transport as compared to what would be expected in a simple diffusion system. We also utilized HXT knockout strains to consider the possibility of other methods of xylulose transport.

2) To determine the flux capacity of xylulokinase in wild-type strains of *S. cerevisiae* and in strains designed to over express xylulokinase

We hypothesized that the bottleneck for xylulose utilization would most likely be either xylulose uptake or the conversion of xylulose to xylulo-5-phosphate. To test this hypothesis we used strains with different XK expression levels and determined the correlation between increased XK expression and an increase in xylulose utilization rate.

3) To produce a sufficient data set to fit and validate a model of xylulose utilization in *S. cerevisiae.*

Additional findings and the rational for performing different experiments pare described later in sections 2.4.1 to 2.4.8.

2.3 Materials and Methods

A series of batch fermentation experiments were carried out in order to complete the stated aims. This research focused on the utilization of xylulose through the pentose phosphate and glycolysis pathways to produce ethanol. Our initial hypothesis was that this pathway had a bottleneck at xylulose transport (which we hypothesized was carried out in large part by the HXT family of transporters) or at the reaction converting xylulose to xylulo-5-phosphate (carried out by xylulokinase) steps in the pathway. To facilitate this research strains with combinations of increased xylulokinase (XK) activity and knocked out HXT transporters were developed as described below.

2.3.1 Strain Definitions:

A total of six strains were used through the course of this research. All of the strains were developed in Dr. Ronald Hector's USDA lab in Peoria, Illinois from the parent strain CEN.PK2-1C. The CEN.PK family of *S. Cerevisiae* strains is one of the major families of reference strains used for *S. Cerevisiae* research. Its high incidence of use is due in part to the large number of mutants available and the strain's ability to readily express many heterologous protein products (Van Dijken et al., 2000).

YRH524 (later referred to as "wild type") is a control strain consisting of CEN.PK2-1C carrying an empty plasmid with the selection marker allowing growth on media lacking tryptophan. This selection marker was used in all strains to promote plasmid retention. YRH857 (referred to as "low copy") and YRH858

(referred to as “high copy”) contain different plasmids that express either low or high copy xylulokinase under the control of a ubiquitous promoter. Since wild-type *S. Cerevisiae* already contains the gene for XK the plasmids in these strains provide supplemental support to that enzyme. The potential for conversion of xylulose to xylulo-5-phosphate is therefore greatest in YRH858 (high copy), followed by YRH857 (low copy), and least in YRH524 (wild type).

YRH859 contains the same low copy XK vector as YRH857, but has been adapted for growth on xylose. This strain had a vector containing the genes for xylose reductase and xylose dehydrogenase added and was grown on xylose. This plasmid was later removed producing a strain similar to YRH857, but that may have a metabolism better adapted for xylose (and thus xylulose) utilization.

The strains YRH1153 and YRH1154 are knockout strains that do not express the glucose transporters HXT 1-7. YRH1153 has an empty vector with the same tryptophan selection marker used in YRH524. YRH1154 has the same high copy XK plasmid as YRH858. All strain and plasmid definitions are outlined in s 2.4 and 2.5 below. Plasmid maps for unmodified pRS414 and pRS424 from which all plasmids in our study were created are included in the appendix.

Table 4: Strain Definitions - These definitions are based on the parent strain and the incorporated plasmid or knocked out genes

Strain Definitions	
Strain	Genotype
CEN.PK2-1C	*MAT*a *ura3-52 trp1-289 leu2-3_112 his3D1 MAL2-8c SUC2*
YRH524	CEN.PK2-1C + pRS414 (empty TRP1-marked vector)
YRH857	CEN.PK2-1C + pRH195 (low-copy vector + XKS1)
YRH858	CEN.PK2-1C + pRH196 (high-copy vector + XKS1)
YRH859*	CEN.PK2-1C + pRH195 (low-copy vector + XKS1)
YRH1153	CEN.PK2-1C *hxt1*D, *hxt2*D, *hxt3*D, *hxt4*D, *hxt5*D, *hxt6*D, *hxt7*D + pRS424
YRH1154	CEN.PK2-1C *hxt1*D, *hxt2*D, *hxt3*D, *hxt4*D, *hxt5*D, *hxt6*D, *hxt7*D + pRH196

Table 5: Plasmid Definitions – These definitions are based on the parent plasmid and the incorporated vector

Plasmid Definitions	
Plasmid	Description
pRH195	pRS414 P_{HXT7}:*XKS1*:T_{HXT7}
pRH196	pRS424 P_{HXT7}:*XKS1*:T_{HXT7}
pRS414	pBluescript II SK+, *TRP1, CEN6, ARSH4* (low-copy vector)
pRS424	pBluescript II SK+, *TRP1,* 2μ origin (high-copy vector)

2.3.2 Culture Conditions:

2.3.2.1 General Growth/Passage Culture Conditions:

Passage cultures were carried out in temperature controlled incubators set to 25°C (cooling was not available in one of the incubators and temperatures reached as high as 30°C depending on ambient conditions). Approximately 75% of culture media was changed every 2 to 4 days by letting the cells settle and decanting a portion of the spent media. If the duration between passages (media exchanges) exceeded two days glucose feed media was used to boost available sugar to 60g/l if the culture. The cultures were not mechanically stirred during

growth, but were agitated during passaging, cell counts, and feedings (approximately once per day). The headspace of the culture was continuously flushed with pressurized air passed through a 0.2-micron filter to provide a slightly aerobic environment and promote cell growth.

2.3.2.1 General Experimental Culture Conditions:

Immediately prior to experiment inoculation, yeast cells were harvested from passage cultures via centrifugation at 1000 rpm for 10 minutes. The resulting supernatant was decanted and the cells were re-suspended in approximately 5ml experimental media. The required inoculation volume was determined through spectrophotometric analysis of a serial diluted sample of the cell suspension. Flasks (125ml non-baffled Erlenmeyer) containing 50mL of experimental media were inoculated with the required volume of cell suspension (targeting either 10 or 20 g/L depending on experiment) and the first sample was taken immediately thereafter via the vent needle in the flask's stopper. Experimental cultures were maintained at 25°C in a temperature controlled water bath. The cultures were agitated using a built in shaker tray at speed sufficient to keep cells from settling (variable depending upon size of flask and fill level). Flask stoppers were vented to allow escape of CO_2, but the headspace was not flushed (as to maintain an anaerobic environment). A reference table for comparison of specific conditions in each experiment is included below (Table 6).

Table 6: Consolidated Table of Experimental Conditions - All averages are of all strains and all replicates. Individual starting conditions available in the appendices.

	Experiment							
	E1	E2	E200	E202	E203	E300	E2.1	E222
Initial Glucose Concentration (g/l)	22.77±1.89	0.00±0.00	0.82±0.16	0.00±0.00	0.00±0.00	0.08±0.32	3.41±0.37	32.59±1.49 (33.09±2.20)*
Initial Xylulose Concentration (g/l)	11.11±0.54	20.20±0.52	37.24±6.15	60.68±1.18	71.11±0.97	101.29±8.97	8.24±0.43	53.67±3.47 (15.77±10.75)*
Initial Xylose Concentration (g/l)	24.78±1.35	37.51±0.98	102.85±7.35	43.47±1.17	33.45±0.54	77.27±3.41	21.86±1.21	34.84±1.27 (29.40±0.59)*
Initial Biomass Concentration Target (g/l)	11.20±0.31	17.92±0.54	19.58±1.60	9.89±1.35	17.17±0.49	11.26±1.31	9.41±0.58	9.88±0.35 (10.86±0.42)*
Strains Utilized	YRH524, YRH857, YRH858, YRH859	YRH524, YRH857, YRH858, YRH859	YRH524, YRH857, YRH858, YRH859	YRH524, YRH857, YRH858, YRH859	YRH524, YRH857, YRH858, YRH859	YRH524, YRH857, YRH858, YRH859, YRH1153, YRH1154	YRH1153, YRH1154	YRH524, YRH857, YRH858, YRH859, YRH1153, YRH1154
Replicates Per Strain	3	3	3	4	3	3	4	3
Sampling Schedule (hrs)	0, 1, 2, 4, 6, 9, 12, 24, 48, 72, 96	0, 1, 2, 4, 6, 9, 12, 24, 48, 72, 96, 120	0, 2, 4, 6, 9, 12, 24, 34, 48, 72, 96, 120	0, 6, 12, 24, 48, 72, 96, 120, 144, 168	0, 6, 12, 24, 48, 72, 96, 120, 144, 168, 192, 216, 240	0, 12, 24, 48, 72, 96, 117, 141, 171	0, 2, 4, 6, 8, 10, 12, 24	0, 4, 8, 12, 24, 48 (48.1, 52, 56, 60, 72, 96, 120, 144)
Other Notes	* Values after a glucose spike at 48 hr to achieve a target glucose concentration of >30g/L in the media. (E222)							

Experiments E1, E2.1, and E222 were run using media containing glucose. These experiments were designed to show that xylulose consumption is linked to glucose consumption by the glucose-mediated induction of the HXT transporters. All other experiments used glucose free media (the small amounts of glucose in the starting samples of E300 and E200 was carried over from passage during inoculation). These experiments were designed to find the maximum xylulose utilization rate in glucose free medium. By comparing the maximum rates in strains with differing levels of xylulokinase we hoped to demonstrate that XK was the limiting reaction in xylulose utilization.

Experiment E222 was designed as a fed batch experiment with glucose feeding during the 48-hour sample. The values in parentheses in Table 6 denote the concentrations immediately following feeding to raise the glucose level in the cultures by 30g/l. Spiking this culture with glucose allowed us to look into the kinetics of glucose transporter induction and its effect on xylulose utilization.

2.3.3 Measurement Methods

Yeast cell cultures are turbid. The culture medium generally allows most light to pass through while the cells obstruct a portion of the light incident upon a sample. Spectrophotometry takes advantage of this property by measuring the absorbance of a culture sample versus a blank (pure culture medium) and comparing it to a standard curve of known cell densities and corresponding concentrations. In this work cell density was measured via spectrophotometer at 600nm.

In this set of experiment all HPLC runs were performed on a BioRad HPX-87H carbohydrate analysis column and the recommended 0.005M H_2SO_4 mobile phase. The column temperature was maintained at 65°C. The flow rate during all runs was set to 0.6 mL per minute. This column and set of conditions provided reasonable resolution for glucose, xylose, xylulose, ethanol, acetic acid, lactic acid, succinic acid, and glycerol. This yielded good peak separation with approximate residence times listed below.

Table 7: Approximate HPLC Peak Retention Times - As run on an 87H HPLC column with 0.6mL/minute flow rate.

Compound	Approximate Residence Time (minutes)
Glucose	8.9
Xylose	9.5
Xylulose	10.1
Succinic Acid	11.4
Lactic Acid	12.5
Glycerol	13.2
Acetic Acid	14.9
Ethanol	21.8

Several metabolic processes in *S. cerevisiae,* notably ethanol production through fermentation, cause the production and release of CO_2. The measurement of this production is not captured in HPLC because the CO_2 is lost to the atmosphere. One way of quantifying this loss in batch experiments is through the measurement of weight loss. As CO_2 leaves the culture, the weight of the flask decreases. Initially we made this measurement (data not shown), but it was abandoned as the signal to noise ratio of the measurement proved insufficient given the low rate of fermentation. CO_2 production was therefore

estimated for mass balance purposes using the stoichiometric relationships between production of measured metabolites and production of CO_2.

This data was analyzed for scientifically and statistically significant comparisons. P-values were calculated using either one or two tailed student-T tests based on the nature of the null hypothesis to be tested. P-values below 0.05 were considered to be statistically significant.

2.4 Results and Discussion

2.4.1 – Identification of Xylulokinase (XK) as Xylulose Utilization Bottleneck in Wild Type Cells

A major objective of these batch experiments was to identify the bottleneck in xylulose utilization. We had hypothesized that this bottleneck would either be in the transport of xylulose into the cell or in the conversion of xylulose to xylulo-5-phosphate.

To test this hypothesis we compared the experimental results from our wild type, low copy XK, and high copy XK strains. These strains are genetically identical except in the level of XK expression. If, for a given experiment, strains with higher XK capacity showed an increased rate of xylulose utilization we would conclude that XK was the limiting reaction in the utilization of xylulose. Conversely, if the rates of xylulose utilization (for a single experiment) were the same amongst these strains independent of XK capacity, we would conclude that transport was the limiting reaction.

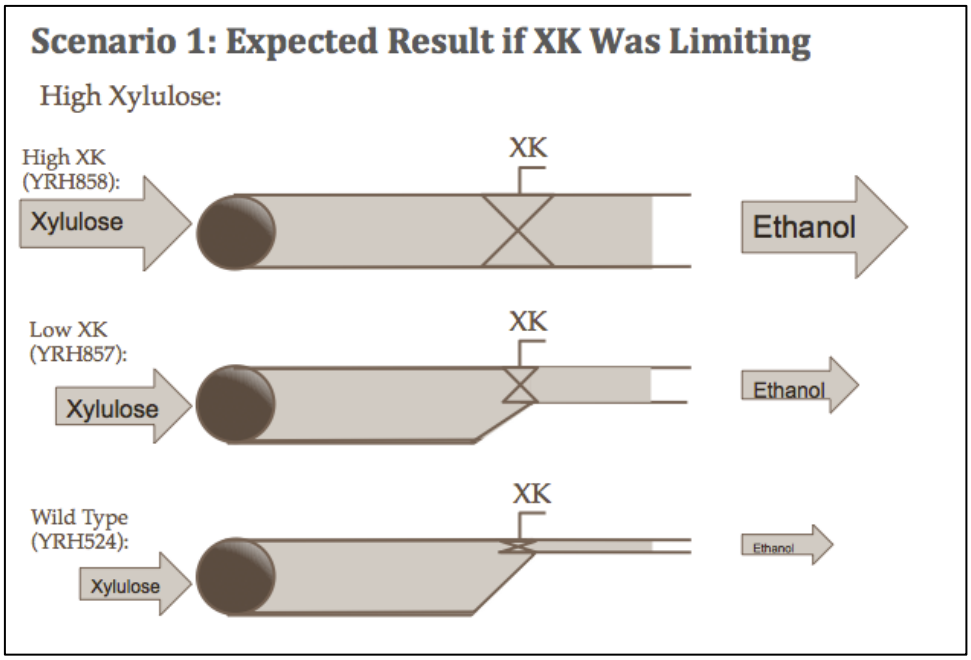


Figure 15: Expected Result if XK Capacity was Limiting Xylose Utilization Rates

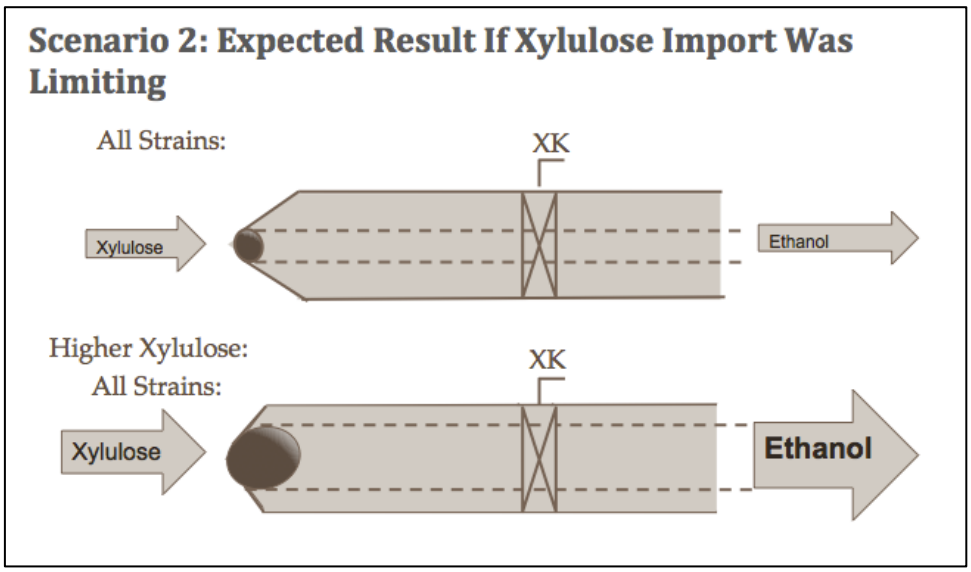


Figure 16: Expected Result if Xylulose Uptake Rate was Limiting Xylulose Utilization

In comparing wild-type (YRH524) *S. cerevisiae* to strains with excess of XK (from either the high copy and low copy XK plasmid), the wild-type strain showed a decreased average xylulose utilization rate over the first 12 hours of

fermentation in four of seven independent experiments. Under experimental treatments E1, E2, E202, and E203 there was a statistically relevant difference (p-value< 0.05) between the wild-type strain (YRH524) and each of the XK enhanced strains. Experimental conditions E200 and E222 showed average consumption rates slightly higher in all XK enhanced strains, but the differences were not large enough to be statistically valid (p-values ranged from 0.096 to 0.226). Data from experimental treatment E300 did not fall in line with this trend. This treatment had the highest starting sugar concentration (>200g/l total sugars) and may have been hampered by sugar inhibition.

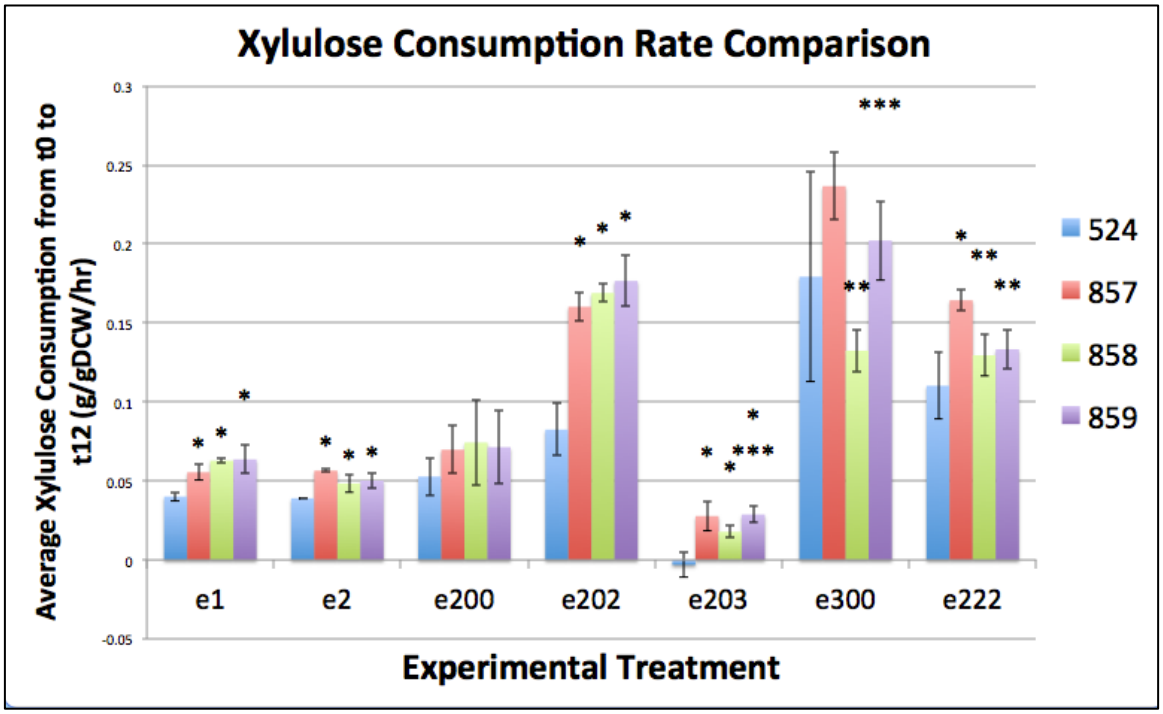


Figure 17: A Comparison of the Average Xylulose Consumption Rates During the Initial 12 Hours of Fermentation Under Each Experimental Treatment - Error bars denote ± one standard deviation. Asterisks above the column bars denote statistical difference (p< 0.05) between strains within an experiment with *, **, * denoting statistical difference to YRH524, YRH857, and YRH858 respectively.**

Additionally, in E300, measurements were taken at less regular intervals (0, 12, 24 hours) making it difficult to judge how the strains initially reacted to the

high sugar concentration. With four of six experimental treatments showing a significantly decreased utilization rate inYRH524 and two additional experiments showing small but statistically insignificant support for this trend, this data as a whole supports the hypothesis that XK was limiting in strain YRH524.

2.4.2 – Diminishing Returns of Expanding XK Capacity and 'Optimizing' Cell Metabolism for Xylose

While expanding XK beyond wild-type expression levels improved xylulose utilization, the difference in xylulose consumption rate between strains with increased levels of XK expression was not significant. Only in two of seven experimental treatments (E300 and E222) was there a statistically significant difference between the low-copy XK enhanced strain (YRH857) and the high copy XK enhanced strain (YRH858). If the XK was the limiting enzyme in the pathway among the strains (YRH524, YRH857, and YRH858) as indicated by scenario I in Figure 15, an expected result of these experiments is higher xylulose consumption rate in high copy strain (YRH858) compared to low copy strain (YRH857) or wild type (YRH524). However, in the two cases that showed statistically significant differences, this property was reversed with the low copy strain (YRH857) consuming xylulose at a higher rate than high copy strain (YRH858) (Figure 17 E222 and E300). The remaining experimental data did not show any significant difference between strains suggesting that transport quickly becomes the limiting reaction in xylulose utilization once the XK capacity is expanded.

Similarly, there seemed to be limited benefit to adapting strain YRH859 for growth on xylose. When comparing the low copy (YRH857) and the low copy xylose adapted strain (YRH859) the xylulose utilization rates were different in only one of the seven experiments (E222). In this instance the adapted strain (YRH859) actually consumed more slowly than the un-adapted strain indicating that adaptation of strain (YRH857→YRH859) did not improve xylulose utilization. While the data in this set of experiments is not in full agreement, the preponderance of the data points towards there being no difference in xylulose uptake rate between the low and high copy enhanced strains nor between the low copy and low copy adapted strains.

2.4.3 – Rate of Xylulose Transport Not Well Correlated With Concentration

Another major goal of this study was to suggest a mechanism for xylulose transport by identifying the correlations between time, xylulose consumption rate and concentrations of xylulose and glucose. A strong, linear correlation between xylulose concentration and consumption rate in the absence of correlation to glucose or time would suggest a simple diffusion as the dominant transport mechanism. A correlation of xylulose consumption rate with the instantaneous glucose levels might suggest transport dominated by glucose transporters that in turn are strongly correlated with glucose availability. Finally, a strong correlation between fermentation time and sugar concentration in the absence of strong correlation between xylulose consumption rate and sugar concentration is likely a

result of transport mediated by a set of transport proteins that are expressed with a time delay between a stimulus and expression.

The experimental results indicate a correlation between maximum xylulose concentration and maximum xylulose consumption rate. This result was expected since in a simple diffusion system or a facilitated transport system scenarios high concentration gradients would be expected to facilitate higher transport rate. However, although these correlations existed within experimental treatments, the correlations were not constant between experiments and no correlation was obvious when data from multiple experimental treatments was viewed in aggregate. This suggests that xylulose concentration at the levels tested does not significantly influence transport rate and that there is a separate mechanism influencing the uptake rate.

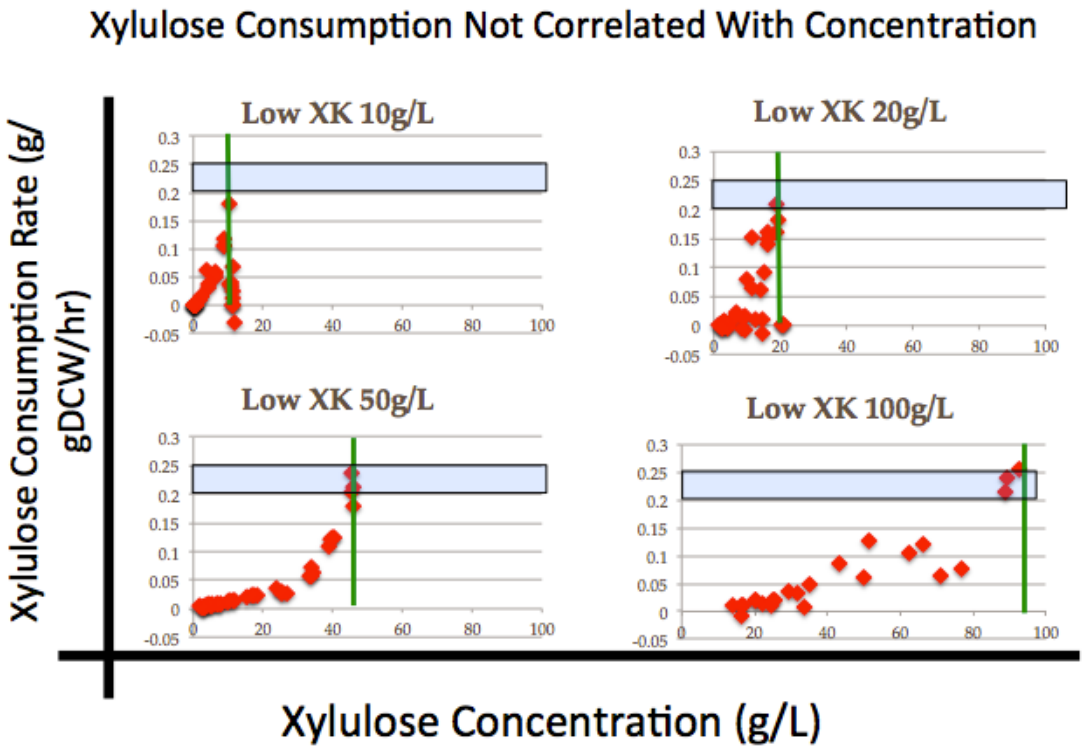


Figure 18: Different Correlations Between Xylulose Concentration and Xylulose Uptake Rate in Different Experimental Treatments - Note that while each treatment reaches a xylulose consumption rate of 0.2-0.25g/gDCW/hr, they do so at different xylulose concentrations.

Our initial hypothesis about transport was that it was carried out in large part by the HXT family of transporters (specifically HXT 1-7). According to Bertilsson's model, the expression level of these transporters in the absence of glucose is constant (Bertilsson et al., 2007). When the expression level was constant, Bertilsson's model showed that xylose consumption was correlated with xylose concentration. Our results presented above indicate that that this correlation was not present in our data. This suggested two new hypotheses. First, there could be a delay in the expression of the HXT transporter family that Bertilsson's model did not capture. Since our strains were grown in a glucose medium prior to experiments, it seemed possible that the HXT family of

transporters could be expressed at non-zero glucose levels and that they remained expressed for some time after the cells were passaged to glucose free medium. Second, there may be a different set of transporters that are expressed under different stimuli causing the unexplained pattern of xylulose utilization.

2.4.4 – Rate of Xylulose Transport Correlated With Time

To test the hypothesis that the observed xylulose utilization pattern is caused by a time-delayed degradation of glucose transporters previously expressed under glucose rich conditions, the correlation between xylulose uptake rate and sample timing was analyzed. High correlation between early samples and high consumption rates would support this hypothesis. The correlation between sample time and consumption rate was clearer than the concentrations correlations we had initially examined, but because different experimental conditions produced slightly different consumption rates the correlation was somewhat unclear.

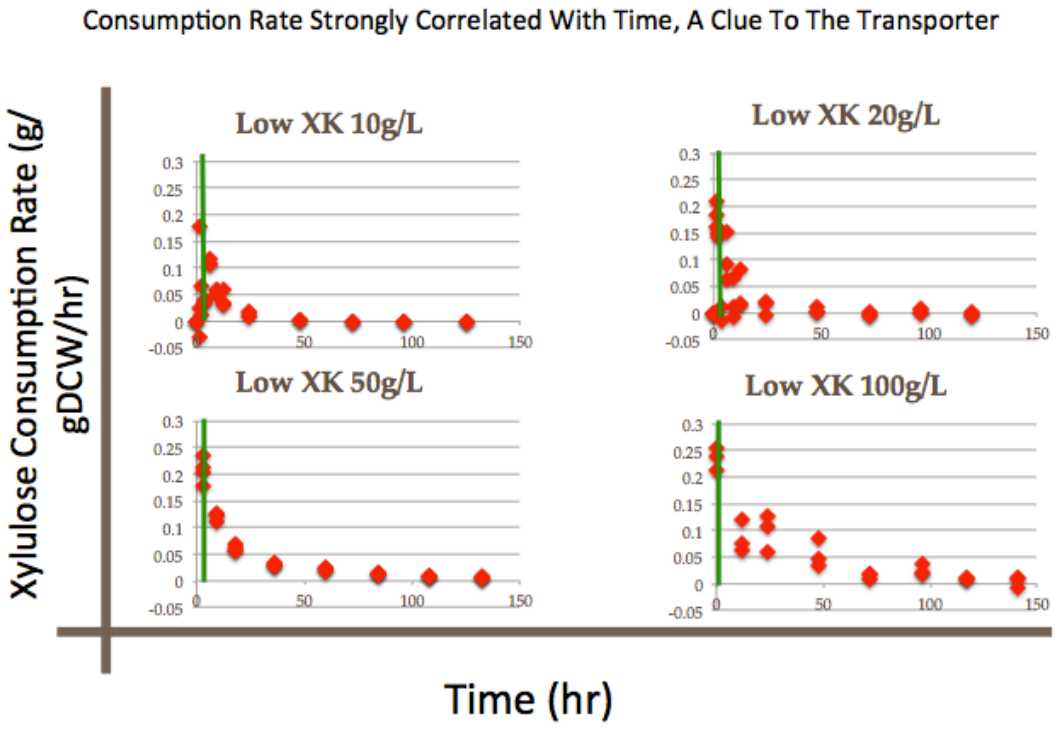


Figure 19: Time Correlation with Xylulose Utilization Rate - The same consumption rate data as in Figure 18, however here it is correlated with the timing of the sample instead of xylulose concentration. This correlation is better standardized amongst experimental treatments

Since this correlation was non-linear (and the strength of the correlation was therefore difficult to quantify) the consumption rate data within an experiment was ranked (where the highest xylulose utilization rate was given the rank 1, the second highest 2, and so on). Comparing this rank to sample time produced a more linear correlation with an R^2 value between 0.113 and 0.526 depending on the strain. While not a strong correlation, this correlation was significantly stronger than those produced in correlations with glucose or xylulose concentration thus supporting the timing hypothesis.

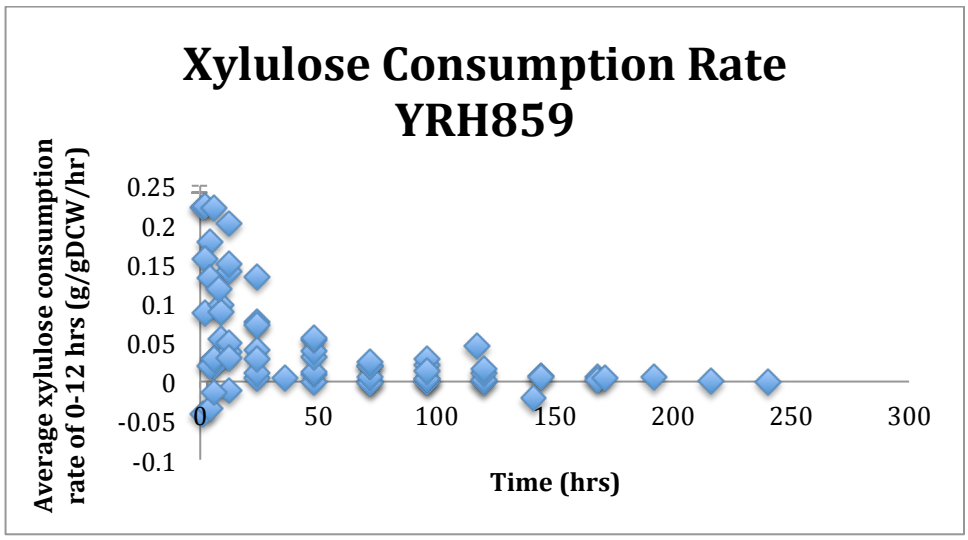


Figure 20: Aggregated YRH859 Data Demonstrating Time/Xylulose Utilization Correlation - This figure shows the correlation between xylulose consumption rate and time using aggregated data from strain YRH859.

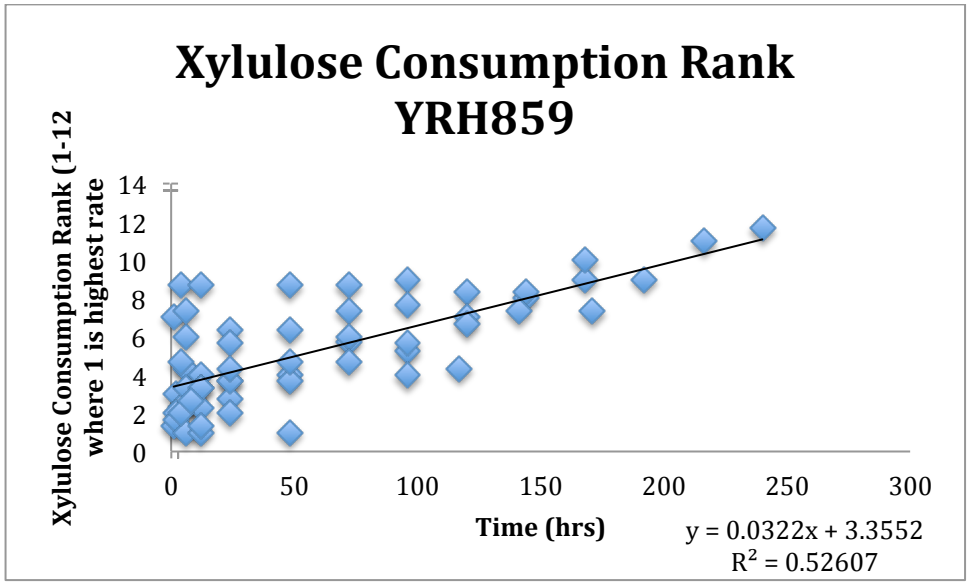


Figure 21: Ranked Aggregated YRH859 Data Further Demonstrating Time/Xylulose Utilization Correlation - To better show the strength of the correlation in Figure 20, the consumption rate data was ranked in each experiment (with 1 being the highest recorded rate and larger numbers denoting slower rates). This allowed the data to be displayed as the linear correlation shown above.

2.4.5 – Xylulose Transport By Transporters Other Than HXT 1-7

To test our secondary hypothesis that additional transporters were active in the transport of xylulose we utilized the knockout strains YRH1153 and YRH1154. Strain YRH1153 is similar to the wild-type strain YRH524 (Figure 22)

with HXT 1-7 knocked out and strain YRH1154 is similar to the high copy XK strain YRH858 (Figure 23). If HXT transporters 1-7 were active in transporting xylulose we would expect to see decreased xylulose uptake rates in these strains. If there were no additional transporters active in xylulose transport we would expect that YRH1153 and YRH1154 would not consume xylulose.

We found that when comparing the 12-hour average consumption rate of an HXT knockout strain to its corresponding HXT positive strain with the same level of XK expressions, the knockout strain showed either no statically valid difference (as in the E222 YRH1153/ YRH524 comparisons) or the knockout strain showed a lower consumption rate for xylulose (Figure 22). This pattern was statistically significant ($p<0.05$) in the E300 YRH1153/ YRH524 comparison and two of three YRH1154/ YRH858 comparisons. In the third YRH1154/ YRH858 comparison the p-value was 0.067; nearly significant given a $p<0.05$ cutoffs (Figure 23).

The statistically significant differences in these experiments suggest that HXT 1-7 play a significant role in xylulose transport. However, because the knockout strains display non-zero xylulose utilization rates, other transporters must be active on xylulose. These transporters seem to be activated to differing degrees under different experimental treatments. Suggesting that the treatments (E300 contained high xylulose, but no glucose while E222 contained low levels of both xylulose and glucose in the initial media as well as a glucose a spike at 48 hours) induce the expression of these transporters differently. Because the HXT

knock out strains were grown on a medium that contained glucose it is also possible that the cells have adjusted their transporter expression to maximize glucose and possibly xylulose uptake through lesser-used transporters. During modeling, this finding might be better studied through MOMA or ROOM FBA techniques.

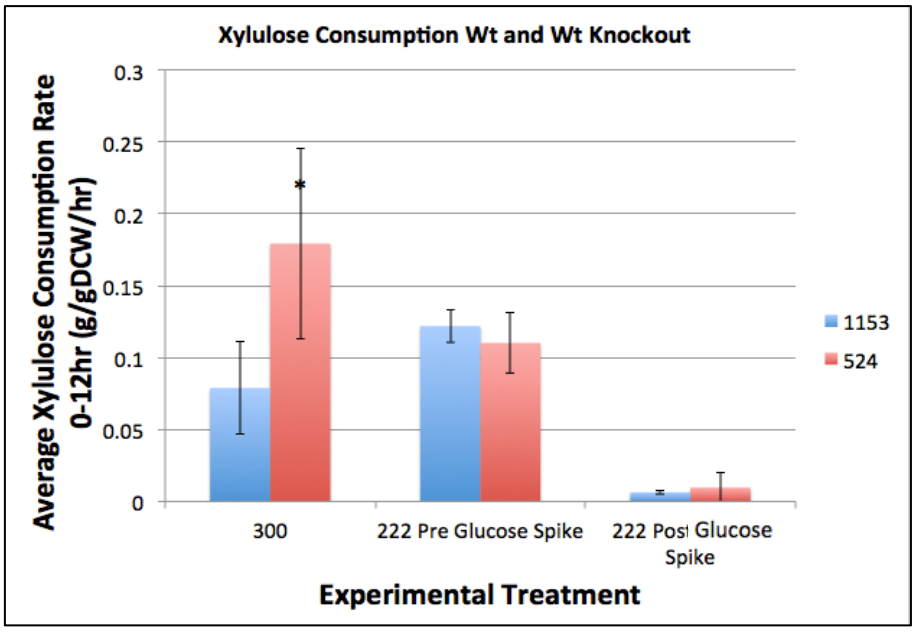


Figure 22: Xylulose Consumption Rate Differences Between YRH1153 (Wild Type XK, HXT-) and YRH858 (Wild Type XK) - Each pair is genetically identical except that YRH1153 (blue bars) has HXT 1-7 knocked out. *denotes a statistically significant P-value <0.05.

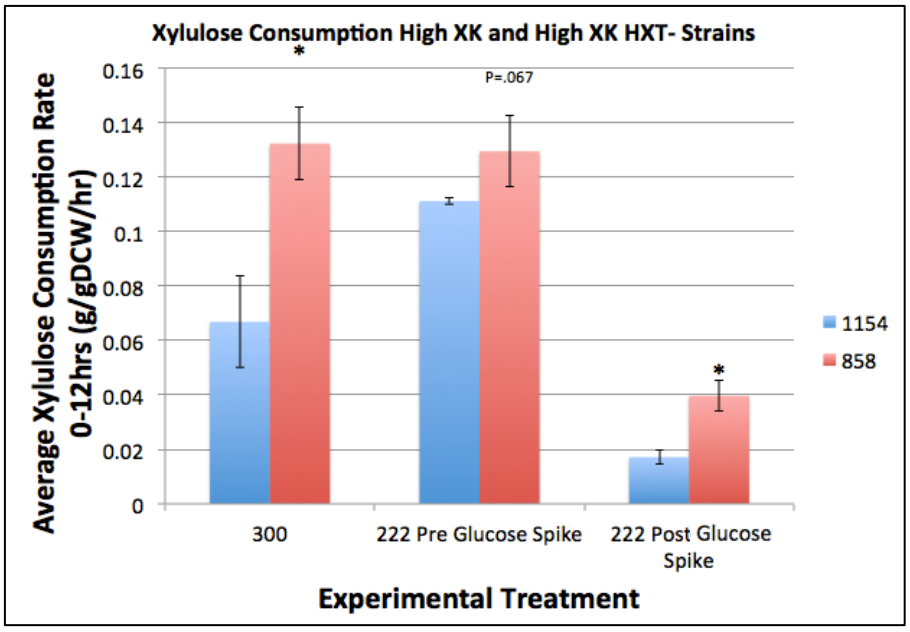


Figure 23: Xylulose Consumption Rate Differences Between YRH1154 (High Copy XK, HXT-) and YRH858 (High Copy XK) - Each pair is genetically identical except that YRH1154 (blue bars) has HXT 1-7 knocked out. *denotes a statistically significant P-value <0.05

2.5 Batch Experiments Conclusions

The batch fermentation section of this work yielded several major conclusions that will be utilized in and further supported through the modeling section of this work. First, XK was identified as the bottleneck for xylulose utilization in the wild-type strain YRH524. A different bottleneck, likely caused by xylulose transport, was identified as a common bottleneck for both high and low copy XK enhanced strains (YRH857 and YRH858). This bottleneck was also found to be the same in the xylose-adapted, low copy XK enhanced strain YRH859. The xylulose utilization rate was found to be better correlated with sample timing than either xylulose concentration or glucose concentration suggesting that xylulose transporters are expressed in a time delayed response to a stimulus. These experiments support the conclusion that it is most likely that

this stimulus is the glucose present in the media used for preparing the inoculum. This conclusion was further reinforced by experiments in which HXT knockout strains had lower xylulose consumption rates. Finally, HXT knockout strains were used to show that transporters other than HXT 1-7 are used to transport xylulose and that those transporters may be overexpressed in adapted strains with HXT 1-7 knocked out.

2.6 Batch Experiments Appendix

2.6.1 Media Formulas

2.6.1.1 Growth (Passage Media):

Table 8: Glucose Passage Media Composition - ᵀGlucose was the only sugar used for passage culture of strains with HXT transporters (YRH534, YRH857, YRH858, YRH859), however in HXT- strains (YRH1153, YRH1154) a mixture of 30g Maltose and 35g Glucose was used.*Tween was not initially included in this media and was added in all passage and experimental media following the second experiment (E2). The inclusion was an attempt to produce higher cell density in passage cultures allowing for higher density experiment inoculation.

Glucose Passage Media Components	
Component	**Amount**
DI-H_2O	<1.0L (split)
Glucose ᵀ	65.0g
Yeast Nitrogen Base without Amino Acids (YNB)	6.7g
Amino Acid Dropout Mix minus tryptophan (AADM)	2.0g
Tween*	1.0mL
Output	**Amount**
Glucose Passage Media	1.0L

Passage media was produced in two parts. In a 500mL volumetric flask, glucose was combined with 300mL DI-H_2O (deionized milli-Q water) and was

stirred to dissolve. Additional DI- H_2O was added to bring the final volume to 500mL. In a second 500mL volumetric flask, YNB, AADM, and Tween were combined with 300mL DI-H_2O and stirred to dissolve. Additional DI- H_2O was added to bring the final volume to 500mL. The contents of the flasks were transferred to cleaned 1L Erlenmeyer flasks. These flasks were capped loosely with foil and autoclaved for a 20 minute wet cycle at 150°C. After cooling to room temperature, the two flasks were combined to produce 1L passage media.

2.6.1.2 Experimental Media:

Table 9: Experimental Media Composition - *Tween was not initially included in this media and was added in all passage and experimental media following the second experiment (E2). The inclusion was an attempt to produce higher cell density in passage cultures allowing for higher density experiment inoculation.

Experimental Media Components	
Constant Components	**Amount**
DI-H_2O	<1.0L (split)
Yeast Nitrogen Base without Amino Acids (YNB)	6.7g
Amino Acid Dropout Mix minus tryptophan (AADM)	2.0g
Tween*	1.0mL
Variable Components	**Amount**
Glucose	0.0-33.0g
Xylose	22.0-103.0g
Xylulose	11.0-101.0g
Output	**Amount**
Experimental Media	1.0L

Experimental media was produced in a similar manner to passage media. Experimental media was produced in two parts. In a 500mL volumetric flask, required amounts of powdered glucose and prepared isomerized xylose/xylulose solution were combined with 300mL DI-H_2O and stirred to dissolve. Additional

DI- H_2O was added to bring the final volume to 500mL. In a second 500mL volumetric flask, YNB, AADM, and Tween were combined with 300mL DI-H_2O and stirred to dissolve. Additional DI- H_2O was added to bring the final volume to 500mL. The contents of the flasks were transferred to cleaned 1L Erlenmeyer flasks. These flasks were capped loosely with foil and autoclaved for a 20 minute wet cycle at 150°C. After cooling to room temperature, the two flasks were combined to produce 1L experimental media.

2.6.1.3 Glucose Feed:

Table 10: Glucose Feed Media Composition

Glucose Feed Components	
Component	**Amount**
Glucose	500.0g
DI-H_2O	<1.0L
Output	**Amount**
Glucose Feed	1.0L

Feed Media was produced in 1L batches and was used to boost available sugar between passage media exchanges. In a 1L volumetric flask 500 grams of Glucose was combined with 400mL DI-H_2O. The mixture was topped to 1L with DI-H_2O. The mixture was autoclaved for a 20 minute wet cycle at 150°C.

2.6.1.4 Xylulose Syrup Production:

Our method for the isomerization of xylose and purification xylulose evolved as experiments required higher concentrations of high purity xylulose, removal of sodium tetra-borate, and minimization of precipitated enzyme. The

initial production method is detailed below along with modifications and the rationale behind them.

Table 11: Initial (E1 and E2) isomerization composition

Initial Isomerization components	
Constant Components	**Amount**
DI-H_2O	<1.0L
Sodium Tetra-borate	2.0g
$MgSO_4$	36.0ppm
Tetracycline	0.01g
Variable Components	**Amount**
Xylose	0-700.0g
Glucose Isomerase Enzyme (Gensweet ™)	.06g/g Xylose
Output	**Amount**
Xylose/Xylulose Solution	1.0L

Xylose was dissolved in 500mL DI-H_2O by stirring. To this, sodium tetra-borate, $MgSO_4$ solution, and tetracycline solution were added. Additional DI-H_2O was added to the mixture to bring the volume to approximately 950mL. The pH of the solution was adjusted to 7.5 using H_2SO_4 and NaOH. The volume of the mixture was brought to 1L by the addition of more DI-H_2O. The mixture was transferred to a 2L Erlenmeyer flask and glucose isomerase was added. This flask was placed in a 60°C water bath set to shake at approximately 40 rpm for a minimum of 7 hours. After 7 hours the solution was autoclaved for 20 minutes on a wet cycle at 150°C and a sample was taken for HPLC analysis. After the xylose and xylulose concentration were identified the syrup could be used in experimental media.

This method produced a xylulose syrup with a low purity (<30% xylulose) and a large amount of precipitated enzyme. It was also noted that sodium tetra-

borate could have growth limiting effects on *S. cerevisiae.* To fix these issues sodium tetra-borate was removed from the isomerization solution, enzyme loading was dropped to approximately 2mL total (not linked to xylose loading), and a xylulose isolation step was added to remove unconverted xylose and enzyme precipitate.

Table 12: Modified Isomerization Composition - (E200, E202, E203, E300, E222, E2.1)

Modified Isomerization Components	
Constant Components	**Amount**
DI-H_2O	<1.0L
$MgSO_4$	36.0ppm
Tetracycline	0.01g
Variable Components	**Amount**
Xylose	0-700.0g
Glucose Isomerase Enzyme (Gensweet ™)	2.1mL (split)
Output	**Amount**
Xylose/Xylulose Solution	1.0L

Under the new procedure, xylose (700g) was dissolved in 500mL DI-H_2O by stirring. To this, the $MgSO_4$ and tetracycline solutions were added. Additional DI-H_2O was added to the mixture to bring the volume to approximately 950mL. The pH of the solution was adjusted to 7.5 using H_2SO_4 and NaOH. The volume of the mixture was brought to 1L by the addition of more DI-H_2O. The mixture was transferred to a 2L Erlenmeyer flask and 2mL of glucose isomerase was added. This flask was placed in a 60°C water bath set to shake at approximately 40 rpm for a minimum of 3 hours. After 3 hours the pH was check and readjusted to 7.5. An additional 0.1mL of glucose isomerase was added and the flask was returned to the water bath. After a minimum of 7 total hours (4 additional hours)

the solution was transferred to a 1L round bottom flask. The flask was attached to a Buchi (Switzerland) RotaVapor R-110 (rotary vacuum evaporator) and heated to 60°C until the mixture in the flask reduced to roughly one-third its original volume. The concentrated xylulose mixture was added slowly to a constantly stirred beaker containing approximately 4 volumes of cold (-20°C) absolute ethanol. The beaker was stirred in an ice bath to keep the temperature at 4°C while a white precipitate (mainly xylose and precipitated enzyme) formed. The beaker was then transferred to a -20°C freezer and the precipitate was allowed to settle for 90 minutes. The supernatant was vacuum filtered and the precipitate washed with an additional 150mL absolute ethanol. The ethanol was collected and returned to the RotoVapor where the volume was reduced by two thirds. DI-H_2O was added to the round-bottom flask and the process was repeated until residual ethanol was no longer detectable via HPLC analysis. At this point the solution was autoclaved for 20 minutes on a wet cycle at 150°C. A sample of the autoclave solution was analyzed via HPLC to verify its high (up to 68%) xylulose concentration. If sufficient purity was not achieved the ethanol extraction process was repeated.

2.6.2 Strain Plasmid Maps

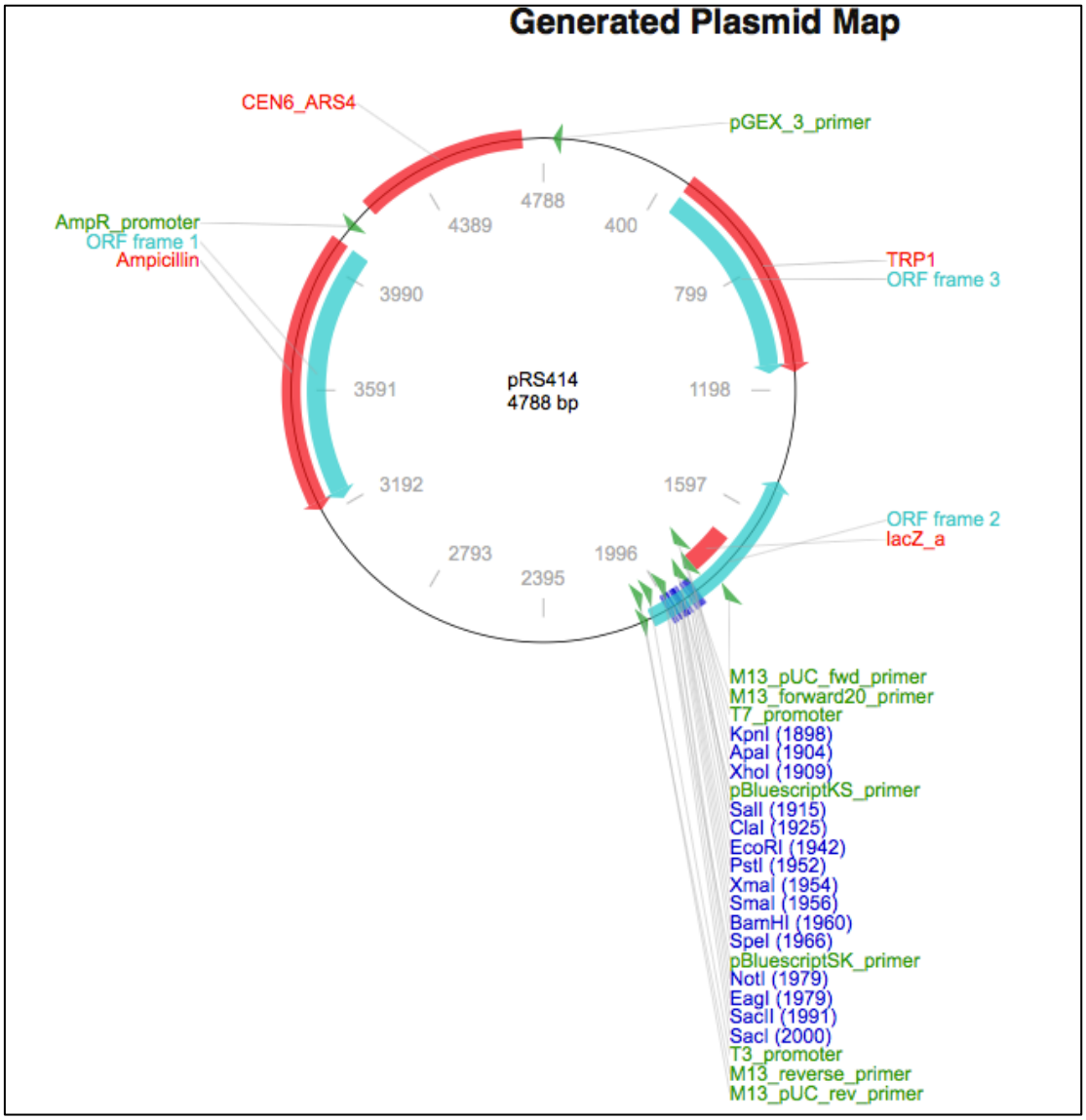


Figure 24: Vector Map for the Parent Vector Used in Plasmid pRH195 - This plasmid is used in YRH857 and YRH859. The empty parent vector was used in YRH524.

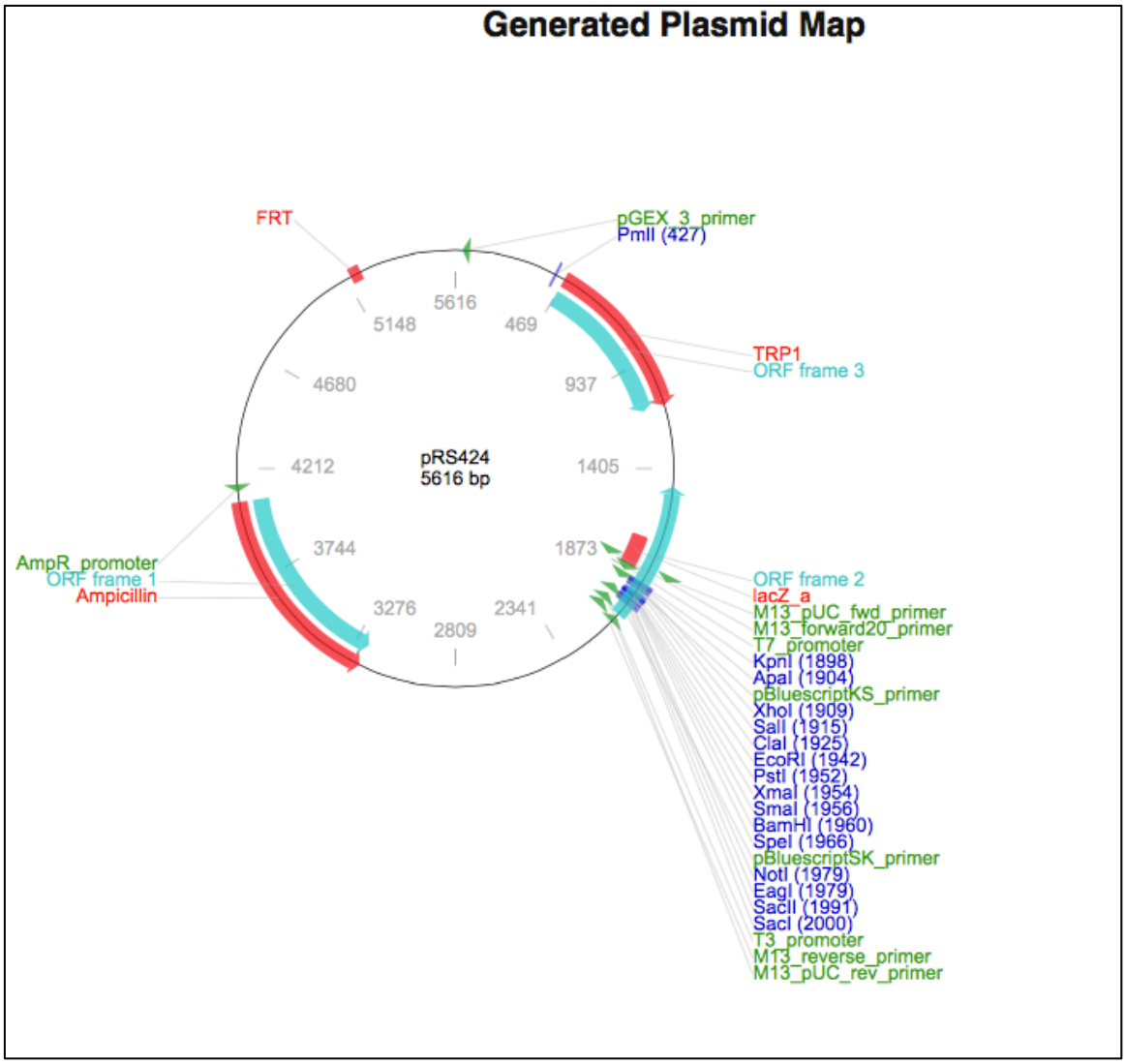


Figure 25: Vector Map for the Parent Vector Used in Plasmid pRH196 - This plasmid was used in the high copy XK strain (YRH858).

2.6.3 – Additional Guiding Experiments

Two small, guiding experiments (2.6.3.1 and 2.6.3.2) were used in this work. They were meant to guide the experimental design by demonstrating that our media composition was not inherently inhibiting to yeast growth. The results are included here to guide future inter lab work; however these experiments were not carried out with enough replicates to support validity statistics.

2.6.3.1 Substrate Inhibition Levels in These Strains

As our sugar levels increased in an attempt to find a saturation point for xylulose uptake we became concerned that sugar inhibition would eventually limit the ability of the yeast to grow and consume xylulose. Here we attempted to find the substrate (glucose) level at which strain YRH857 was inhibited.

Glucose feed media (75g/l) was made following the standard feed media procedure. A stock of 900g/l glucose media was made following the feed media procedure. Five 50ml combinations of feed and media were produced as shown in Table 13 below. T. Flasks were inoculated with 5mL of suspended YRH857 cells from a carry culture in exponential growth. Flasks were weighed after inoculation and at approximately 12 hours. Weight loss was the only measure of sugar consumption considered.

Table 13: Substrate (Sugar) Inhibition Level – The sugar inhibition level was identified via weight loss at different glucose concentrations.

Treatment	Media Volume (ml)	Water Volume (ml)	Feed Volume (ml)	Glucose Concentration (g/l)	Initial Weight (g)	Final Weight (g)	Weight Loss (g)
1	40.0	10.0	0.0	60.0	195.07	194.96	0.11
2	47.0	0.0	3.0	124.5	191.73	191.63	0.10
3	42.0	0.0	8.0	207.0	168.07	167.98	0.09
4	36.0	0.0	14.0	306.0	180.55	180.50	0.05
5	31.0	0.0	19.0	388.5	198.96	198.96	0.00

Sugar consumption appears to drop off substantially in treatments 4 and 5 (from approximately 0.1g/12 hours to 0.05g in treatment 4 and 0.0g in treatment 5. Based on this data we did not attempt to push cells past approximately 200g/l total sugar.

2.6.3.2 – Sodium Tetra-borate Inhibition in These Strains

During the course of our batch experiments we needed increasing concentrations of xylulose without violating the substrate inhibition levels we found above. One way we considered reaching these concentrations was increasing the level of sodium tetra borate in our isomerizations. Sodium tetra borate complexes with xylulose to decrease product inhibition and allows further reaction of xylose to xylulose. This complexing happens in a stoichiometric manner (with one molecule of sodium tetra borate complexing with four xylulose molecules) so it would be theoretically possible to convert 100% of xylose to xylulose if enough sodium tetra borate was present. However, sodium tetra borate has been shown to inhibit the growth of some yeast strains and before increasing our sodium tetra borate concentration it was important to identify the level at which sodium tetra borate inhibition became an issue in our strains.

Determining the inhibition point for sodium tetra borate was carried out in two experiments. The first utilized strains YRH524 and YRH858. Glucose media (60 g/l) was prepared with 15, 10, 7, 5, and 3 g/l sodium tetra borate. Flasks were prepared with 100 ml of media and inoculated with 5 ml of cell suspension from a carry culture in exponential growth phase. Cell weights were taken at inoculation and at 24 hours. This experiment and result is laid out in Table 14 below.

Table 14: Sodium Tetra Borate Inhibition Level - Current and increased levels of sodium tetra borate were shown to be inhibitory via weight loss at varying sodium tetra borate concentrations

Strain	STB concentration (g/l)	Initial Weight (g)	Final Weight (g)	Weight Loss (g)
524	15.0	181.2	181.2	0.0
524	10.0	202.0	202.0	0.0
524	7.0	221.7	221.6	0.1
524	5.0	200.6	200.5	0.1
524	3.0	181.5	181.3	0.2
857	15.0	202.2	202.2	0.0
857	10.0	201.6	201.6	0.0
857	7.0	170.7	170.7	0.0
857	5.0	213.7	213.6	0.1
857	3.0	192.2	192.1	0.1

This experiment suggested that there is increasing inhibition at greater than 3 g/l and that sugar consumption is almost completely stopped above 5 g/l. Previously we had used 2 g/l in our isomerizations and this result suggested that any increase beyond this would have a deleterious effect on sugar consumption, but we also wanted to know if there had been any effect at the 2 g/l level so a second experiment was run measuring sugar consumption in strain YRH859 at levels of STB between 0 and 5 g/l. The media for the experiment was 60g/l glucose passage media. Samples were taken at inoculation, at 5 hours, and at approximately 24 and 72 hours. The experiment and results are detailed in Table 15 below.

Table 15: Low Level Sodium Tetra Borate Inhibition – A second experiment considering whether low level sodium tetra borate was also potentially inhibitory to YRH859

Sample	STB concentration (g/l)	Initial Weight (g)	Weight at 5 hours (g)	Weight at 24 hours (g)	Weight at 72 hours (g)
1	0.0	178.13	178.08	177.30	176.89
2	1.0	219.4	219.4	218.6	218.2
3	1.0	177.22	177.20	176.38	175.91
4	6.0	222.3	222.2	221.6	221.0
5	2.0	224.3	224.3	223.7	223.1
6	2.6	189.76	189.75	189.17	188.57
7	4.0	188.90	188.88	188.49	187.76
8	5.0	189.71	189.69	189.39	188.52

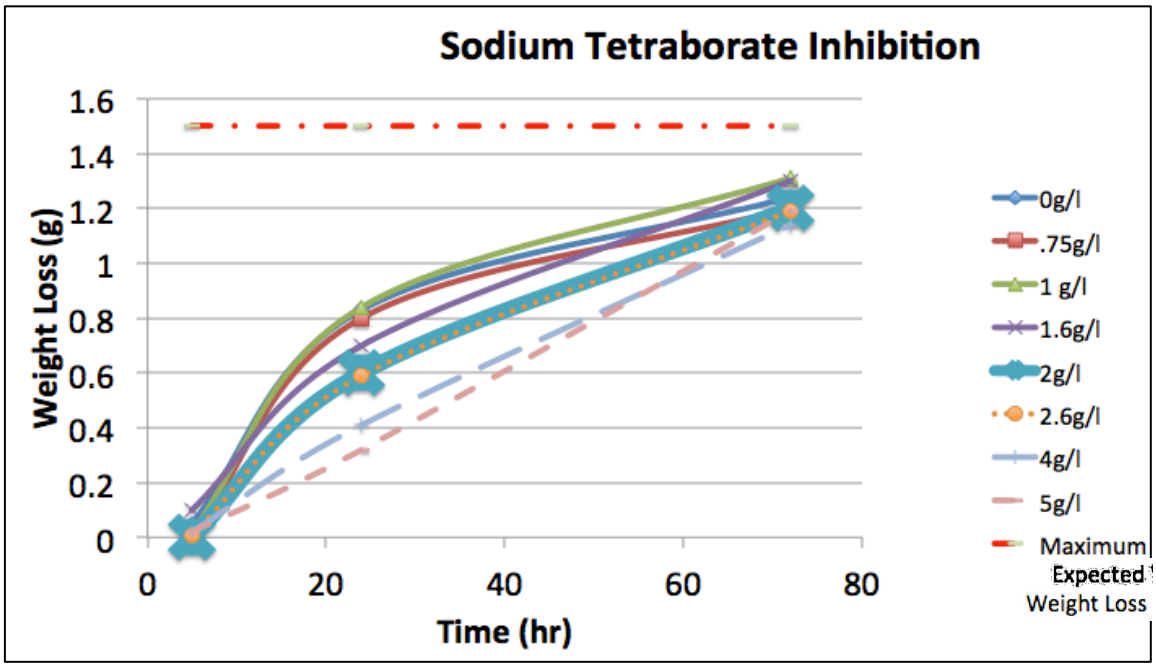


Figure 26: Time Course Experiment Showing Sodium Tetra Borate Inhibition. This figure shows a time course experiment showing decreased weight loss under conditions of increased sodium tetra borate.

There was little difference in weight loss at the first or last time point of this experiment. This is indicative of a lag phase after inoculation and an eventual complete fermentation of glucose in all strains. However, at 24 hours, there is a large (although not statistically verified) difference in weight loss between the

different conditions. Flasks with more than 2.6g/l STB consumed less sugar by hour 24 than the base case (2g/l which we had been using in isomerizations). Conversely, samples from flasks containing less than 1.6 g/l STB consumed slightly more sugar than the base case. Based on this experiment we developed the new method for xylulose purification and eliminated STB in all experimental conditions after E1 and E2.

Chapter 3 - dFBA and Kinetic Modeling of Batch Experiments:

3.1 Introduction

The production of fuel ethanol from lignocellulosic biomass has the potential to replace a significant portion of non-renewable transport fuels. Woody feedstocks are composed of cellulose, hemicellulose, and lignin. While the glucose, the monomer of cellulose, is readily utilized by wild-type *S. cerevisiae*, xylose, which comprises 60% of the sugar monomers in hemicellulose, is not. To make the process economically competitive with conventional fossil fuels, both five and six carbon sugars must be utilized efficiently.

One approach to improving xylose utilization is to convert it to the more readily usable xylulose using an extracellular enzyme. Xylulose is taken up by wild-type S. cerevisiae and incorporated into the pentose phosphate pathway. The key to improving five-carbon sugar utilization by this xylulose pathway is to better understand the bottlenecks in sugar uptake and processing. Dynamic flux balance analysis allows for the identification and quantification of these bottlenecks. In this chapter we have produced a dynamic flux balance model based on the previously developed compartmentalized genome scale model iMM904. This model was adapted with modifications to allow for restricted uptake of five-carbon sugars, expansion of five carbon sugar utilization pathways, and cell death. Model parameters have been fit based on some of the batch fermentation data produced in Chapter 2. The model was validated with separate, independent data sets.

3.2 Overall Aims

The modeling section of this work had three major aims:

1) Develop a model for xylulose uptake and utilization based on the known expression pattern of HXT transporters and the kinetics of xylose uptake to predict xylulose utilization in *S. cerevisiae.*
2) Examine the xylulose utilization bottlenecks discovered in Chapter two and show that they support the model fit.
3) Assess the role of alternative transport mechanisms for xylulose

The potential bottlenecks identified in Chapter 2 of this work act as working theories to test during modeling. Xylulose transport appears to be the limiting factor in xylulose utilization for strains YRH857, YRH858, and YRH859. XK throughput is the apparent bottleneck for strain YRH524. This modeling will test these findings and question whether the expression pattern of the HXT family of proteins is sufficient to predict the observed xylulose uptake.

3.3 Model Development and Methods

3.3.1 Base Genome Scale Model and FBA

The model used in this work was based on a fully compartmentalized genome scale model of *S. cerevisiae* (iMM904) developed by the Palsson group (Mo et al., 2009). This model (iMM904) is itself updated of a previous model (iND750). IMM904 contains 904 genes encompassing 1,412 reactions and has been extensively validated through in silica knockout viability studies (Mo et al., 2009).

IMM904 was read into MATLAB® using Cobra Toolbox scripts where it can be solved as a linear programing problem. As described in the literature review of flux balance analysis (Chapter 1.5-1.5.8), the model is represented as a series of matrices. The matrix 'S' is an m x n matrix that represents the stoichiometry of all of the reactions in the cell. Each row in 'S' is an individual reaction and each column is an individual metabolite thus the number in element (m,n) is the stoichiometric amount of metabolite 'n' associated with reaction 'm'. The matrix 'V' is a column vector with 'm' rows. It represents the molar flux through each reaction. The steady state assumption requires that metabolite concentrations within the cell do not change and thus the viable solution space is restricted to only solutions where S•V is equal to zero. An objective function is used to select a 'best' or biologically valid solution from this set of allowable solutions. Objective functions are defined here as a weighted sum of key fluxes (often growth, ATP production, or ethanol production). The weights are described in a second column matrix 'C' such that C•V gives the value of the objective function. By identifying the allowable solution 'V' that maximizes the objective function, the model predicts the fluxes for a given set of conditions.

3.3.2 Modifications to S. cerevisiae model (iMM904) to incorporate xylulose transport

Xylulose is already an internal metabolite in iMM904. Xylulose is a naturally occurring sugar metabolite and an intermediate in the pentose phosphate pathway, but the iMM904 does not account for xylulose uptake. In

order to facilitate xylulose uptake two reactions must be added. One reaction acts as a transporter for the sugar and the other functions as a sink and is used to track extracellular xylulose. These transporters are non-reversible and the maximum rates for these reactions are calculated based on extracellular conditions. These calculations are explained in section 3.3.4.

Table 16: Model Constraints Restricted or Relaxed to Simulate Anaerobic Growth on Mixed Sugar Media – Sugars include glucose, xylose, and xylulose.

Additional Model Constraints						
Metabolite	**Reaction**	**Type**	**Initial Constraint (mM/hr)**	**Upper, Lower, or Both Bounds**	**Dynamic or Constant**	**Source/ Assumption**
Ergosterol	EX_ergst	Fatty Acid Exchange	-1000	Lower	Constant	http://gcrg.ucsd.edu/InSilicoOrganisms/Yeast/YeastFAQs
Zymosterol	Ex_zymst	Fatty Acid Exchange	-1000	Lower	Constant	
Hexadecenoate (n-C16:1)	Ex_hdcea	Fatty Acid Exchange	-1000	Lower	Constant	
Octadecanoate (n-C18:0)	Ex_ocdca	Fatty Acid Exchange	-1000	Lower	Constant	
Octadecenoate (n-C18:1)	Ex_ocdcea	Fatty Acid Exchange	-1000	Lower	Constant	
Octadecadienoate (n-C18:2)	Ex_ocdcya	Fatty Acid Exchange	-1000	Lower	Constant	
Oxygen	Ex_o2	Fatty Acid Exchange	0	Lower	Constant	
ATP	ATPM	ATP Maintenance	0.9888	Both	Constant	Model Fitting
Glucose	Ex_glc	Sugar Exchange	-21.0	Lower	Dynamic	(Sonnleitnert, 1986)
Xylose	Ex_xyl-(D)	Sugar Exchange	0	Lower	Dynamic	Assume that in the first time step after inoculation from glucose growth medium, only glucose will be consumed
Xylulose	Ex_xylu	Sugar Exchange	0	Lower	Dynamic	
Xylulose -> Xylulo-5-P	XYLK	Internal Xylulose Metabolism	0.4376 (strain YRH524 and YRH1153) 1000 (other strains)	Upper	Constant	Model Fitting
Xylose -> Xylitol	XYLTD_D	Internal Xylose Metabolism	0	Both	Constant	Protein not present

3.3.3 Additional Constraints

Further constraint of the model is based on limiting the allowable range for certain reactions (Table 16). For example, to simulate anaerobic conditions, the

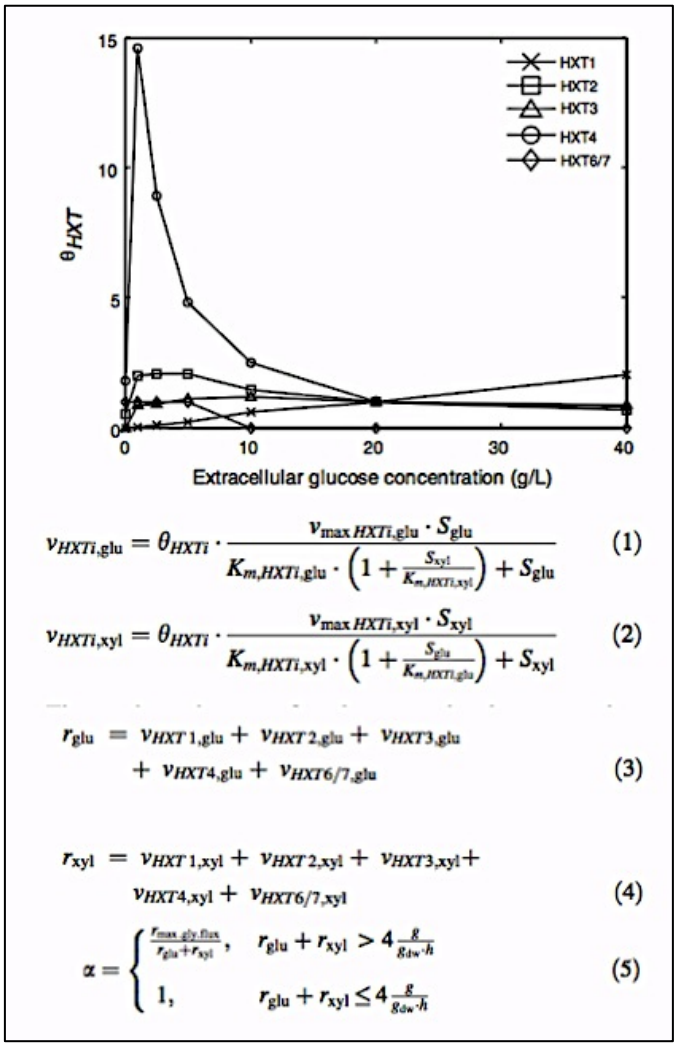
allowable oxygen uptake will be set to zero. Some of these constraints are constant and based on the literature, others are constant and used to fit the model (example ATP maintenance requirements), finally some were dynamic and were updated with each model iteration (example sugar uptake rates are linked to current and previous sugar concentrations and are update each round). The table above outlines the additional constraints added, their source, and whether they are dynamic or constant.

3.3.4 Calculation of Dynamic Constraints (Sugar Uptake Rates)

Sugars are transported into the cell primarily through facilitated transport mechanisms. The expression of each sugar transporter is dictated by stimuli in the extracellular environment. Further, the rate at which sugar is transported through an expressed transporter is constrained by the transport kinetics of the transporter and the concentration of sugar in the media.

Our working hypothesis was that xylulose was transported mainly through the HXT family of transporters and our batch experiments support that view. The HXT family of transporters have a well defined expression profile in relation to glucose concentration in the media (Diderich, 1999; Liang and Gaber, 1996; Ozcan and Johnston, 1995). Bertilsson used this expression profile to help model xylose utilization during mixed glucose xylose fermentation in recombinant *S. cerevisiae (Bertilsson et al., 2007).* Expression levels of the individual HXT transporters were set using the expression profiles. Kinetic parameters for xylose transport through the HXT transporters were taken from work published values (Saloheimo et al., 2006). In the Bertilsson model the expression of each transporter was calculated based on glucose level and relative to a basal expression level (Figure 27 top). The sugar uptake through each transporter (at the previously mentioned basal level) was described for both xylose and glucose with Michaelis-Menten kinetics (Figure 27 equations 1 and 2). The total transport was equal to the sum of transport through each transporter (Figure 27 equations 3 and 4). Bertilsson then imposed an additional constraint limiting the total flux of sugar through the HXT transport family to 4 g/g-DCW/hr (a value previously reported for maximum glucose uptake rate). This produced a scalar (α) used later to calculate uptake (Figure 27 equation 5). If the calculated sugar fluxes

exceeded this rate there were scaled proportionately to be in compliance.



$$v_{HXTi,\mathrm{glu}} = \theta_{HXTi} \cdot \frac{v_{\max HXTi,\mathrm{glu}} \cdot S_{\mathrm{glu}}}{K_{m,HXTi,\mathrm{glu}} \cdot \left(1 + \frac{S_{\mathrm{xyl}}}{K_{m,HXTi,\mathrm{xyl}}}\right) + S_{\mathrm{glu}}} \quad (1)$$

$$v_{HXTi,\mathrm{xyl}} = \theta_{HXTi} \cdot \frac{v_{\max HXTi,\mathrm{xyl}} \cdot S_{\mathrm{xyl}}}{K_{m,HXTi,\mathrm{xyl}} \cdot \left(1 + \frac{S_{\mathrm{glu}}}{K_{m,HXTi,\mathrm{glu}}}\right) + S_{\mathrm{xyl}}} \quad (2)$$

$$r_{\mathrm{glu}} = v_{HXT1,\mathrm{glu}} + v_{HXT2,\mathrm{glu}} + v_{HXT3,\mathrm{glu}} + v_{HXT4,\mathrm{glu}} + v_{HXT6/7,\mathrm{glu}} \quad (3)$$

$$r_{\mathrm{xyl}} = v_{HXT1,\mathrm{xyl}} + v_{HXT2,\mathrm{xyl}} + v_{HXT3,\mathrm{xyl}} + v_{HXT4,\mathrm{xyl}} + v_{HXT6/7,\mathrm{xyl}} \quad (4)$$

$$\alpha = \begin{cases} \frac{r_{\max\ \mathrm{glu\ flux}}}{r_{\mathrm{glu}} + r_{\mathrm{xyl}}}, & r_{\mathrm{glu}} + r_{\mathrm{xyl}} > 4\frac{g}{g_{dw} \cdot h} \\ 1, & r_{\mathrm{glu}} + r_{\mathrm{xyl}} \leq 4\frac{g}{g_{dw} \cdot h} \end{cases} \quad (5)$$

Figure 27: Bertilsson's Glucose/Xylose Uptake Model Assumptions - Assumptions include relative HXT expression level (top figure), kinetic parameters for glucose and xylose uptake rates (equations 1 and 2), the additive transport rate from all HXT transporters (equations 3 and 4), and the derivation of the scalar α (equation 5) (Bertilsson et al., 2007).

Our method is identical to Bertilsson's with two exceptions. First, in addition to the six HXT transporters modeled, an additional aggregate transporter was added to account for unexplained transport of glucose and xylose observed in strains lacking the HXT 1-7. The second way our method differed from that of Bertilsson was in the final scaling constraint. Bertilsson et al. (2007) did not distinguish between fluxes of glucose and xylose when calculating the total sugar

flux into the cell. Maximum possible value for sugar flux was set to 4g/g-DCW/hr based on experimentally observed value of total flux of sugar through the HXT transport family. This implies that the glucose or xylose flux can have a maximum possible value of 4g/g-DCW/hr. However, in all experiments containing glucose (E1, E2.1, and E222) we observed that the highest sugar flux for xylose and xylulose was considerably lower than glucose flux. This observation pointed to differences in the maximum possible flux values through the HXT transporters for glucose, xylose and xylulose, which were not modeled by Bertilsson et al. (2007). It was hypothesized that if the maximum observed transport of glucose was 4 g/g-DCW/hr (in the absence of other sugars) and the maximum level of transport xylose through the same set of transporters was (for example) 2 g/g-DCW/hr (in the absence of other sugars), then the transport capacity of the transporter set should be defined in terms of both glucose and xylose. A brief illustration of the rationale can be explained as follows: Using the a maximal flux rate of 4 g/g-DCW/hr for glucose and 2 g/g-DCW/hr for xylose, and assuming a uniform transportation rate, the time to transport 1 g/g-DCW glucose or xylose would be 15 or 30 min respectively. Therefore the capacity is four units of glucose per hour, or two units of xylose, or some combination of glucose and xylose the sum of which does not violate the transport capacity. For example one xylose and two glucose would be acceptable as the total transport capacity is not violated, but three glucose and one xylose would be unacceptable as the total transport capacity is violated even while individual transport capacities are within allowable

ranges (Figure 28). The maximum level of xylulose transport is not well defined for all conditions. As a result the “time scalar” for xylulose was fit as a model parameter.

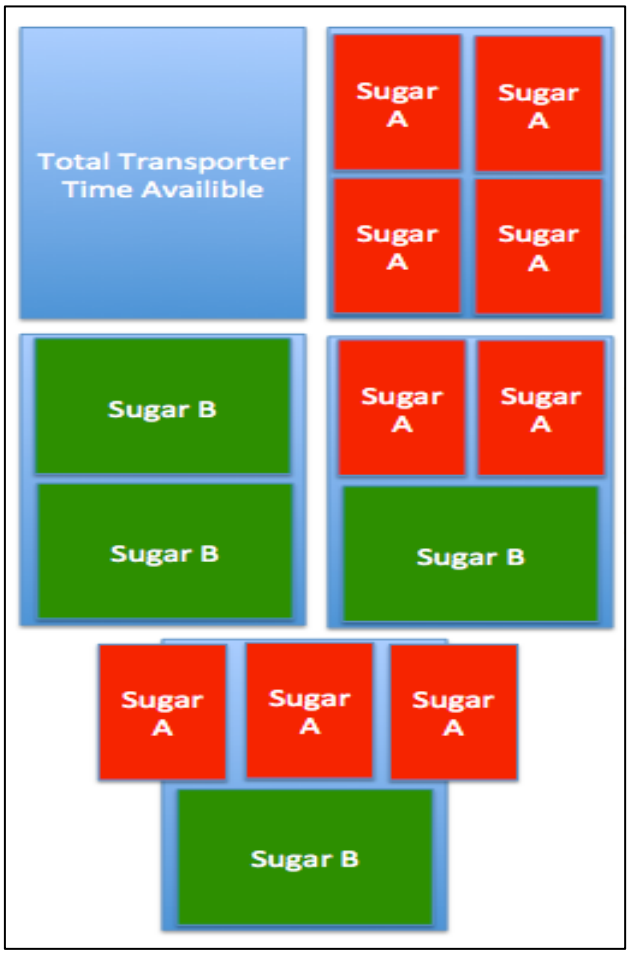


Figure 28: Rational For a Sugar Time Scalar - If a transporter has an allowable throughput (an “allowable transporter time”) sufficient to transport four units of sugar ‘A’ (ex. Glucose) or two units of sugar ‘B’ (ex. Xylose) then three units of sugar ‘B’ and one unit of sugar ‘A’ would require too much transporter time and should not be an allowable throughput. Our “time scalar” differentiates between sugars and how quickly they can be transported while this is ignored in the Bertilsson model.

3.3.5 Death Reaction Explains Batch Fermentation End Dynamics

Modeling the end dynamics of batch culture was a challenge in this work. In the batch experiments there were several runs in which fermentation stalled prematurely. This rarely happens in glucose fermentation in complex media and

when ethanol is low (below 100g/l), but because of the low efficiency with which xylulose is taken up we hypothesized that cells may have difficulty reaching the minimum ATP production required for maintenance and that this may have led to cell death. Modeling provided an opportunity to test this hypothesis in-silica.

To this end a cell death equation was added to the model. When this reaction was active it decreased viable biomass and returned a percentage the metabolites required for cell growth to the surviving cells. This recycling of nutrients allowed a portion of the cells to produce the minimum ATP required for maintenance, survive, and continue to slowly consume xylulose (Figure 29). After several iterations of cell death, this surviving fraction would be sufficiently small to halt fermentation.

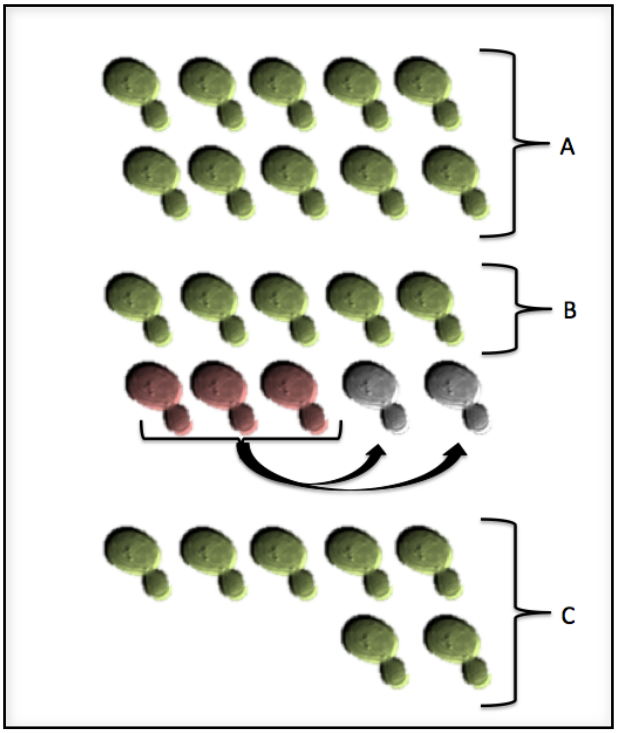


Figure 29: Yeast Cell Death Function Representation - The yeast cell death function imagines a population of cells (A). In this example, there is only enough available sugar in the media to support the ATP maintenance needs of five of the ten cells. Without the cell death function, the model would return an error and the run would stop, but by including the cell death function a fraction of the under fed cells will die and provide nutrients available for other cells to take up (B). The result in this example is that five cells are fed immediately from the available sugars in the media. Another three cells die to provide nutrients that support two additional cells.

In addition to the death equation itself, several new sink equations were added so that the surviving cells could get rid of excess metabolites that they could not use. The required level of ATP for maintenance was fit in the model and allowed for control over the minimum sugar concentration at which death would begin. The percentage of metabolites recovered by death was also fit and allowed for further control of the rate at which cell death occurred.

3.3.6 Changing the Objective Function

With the death equation in use, the initial objective function maximizing growth was no longer adequately representative of the net growth rate. We

therefore updated the objective function to maximize the net growth rate (growth minus death rates) and used this objective function through the course of the analysis.

3.3.7 Model Algorithm Explanation

The model algorithm is outlined in the flow charts below. The genome scale model (iMM904) is read into cobra and the appropriate reactions are added for sugar consumption, death, and death metabolite tracking. The reaction bounds are changed to model anaerobic growth and initial conditions are defined. The first dFBA loop then begins and the counter is set to one. The dynamic flux bounds for glucose, xylose, xylulose, O_2, and CO_2 are changed based on the initial conditions. Next the model is solved to optimize the objective function (maximize growth and minimize death). This produces a solution matrix 'V' and flux values for the tracked metabolites (glucose, xylose, xylulose, ethanol, biomass, acetate, succinate, lactate, and glycerol) are extracted from this matrix. These values are multiplied by the time step (the time step in our model was set to 30 seconds) and added to the current concentrations to derive the new metabolite concentrations. These concentrations are then used to calculate new dynamic flux bounds for the next iteration of the FBA (Figure 31). The simulation loop was run until the target time was reached (Figure 30).

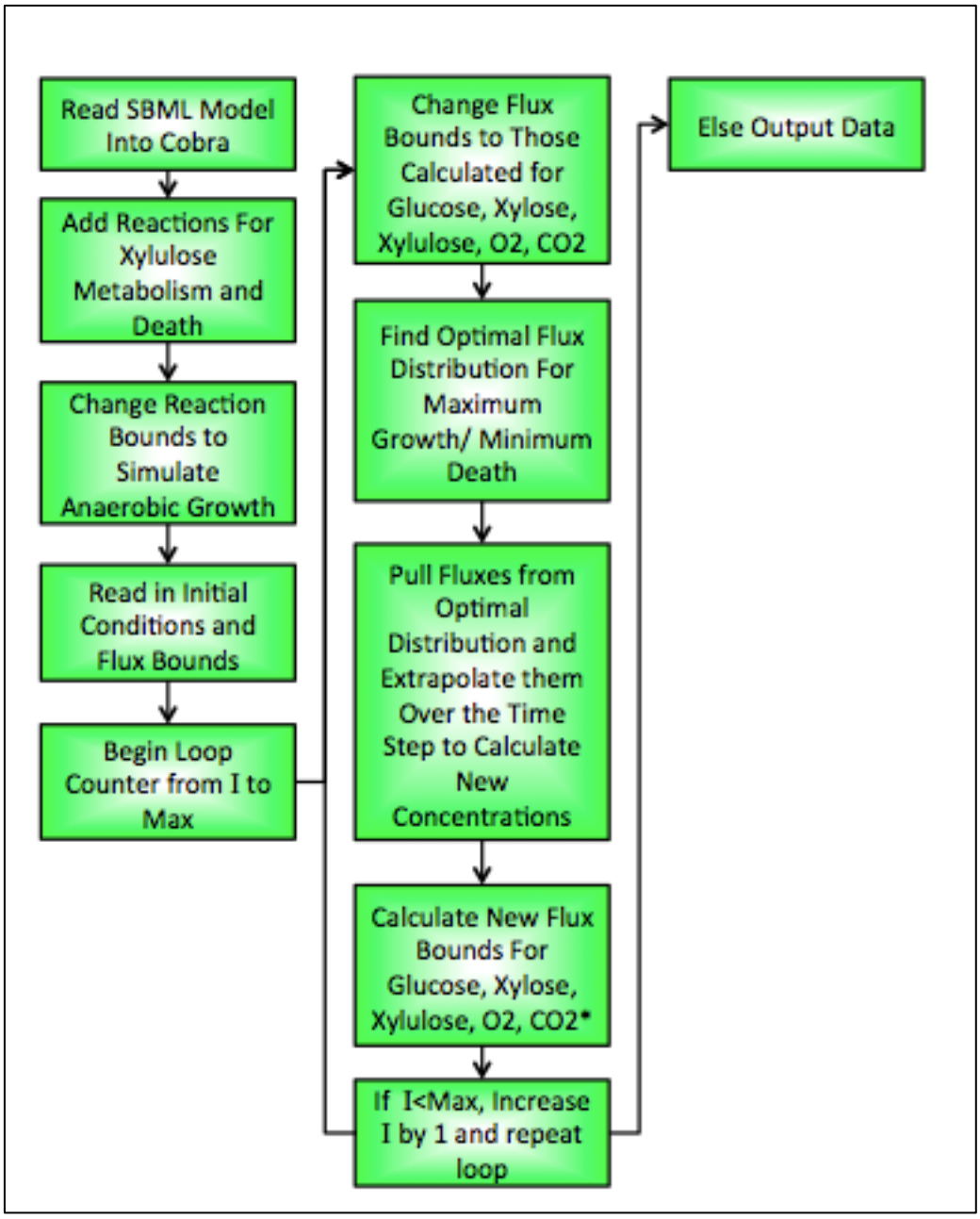


Figure 30: Flow Chart of dFBA Model Loop - This flowchart depicting the dFBA loop used in this model. The sub algorithm for calculating new flux bounds is described below in Figure 31

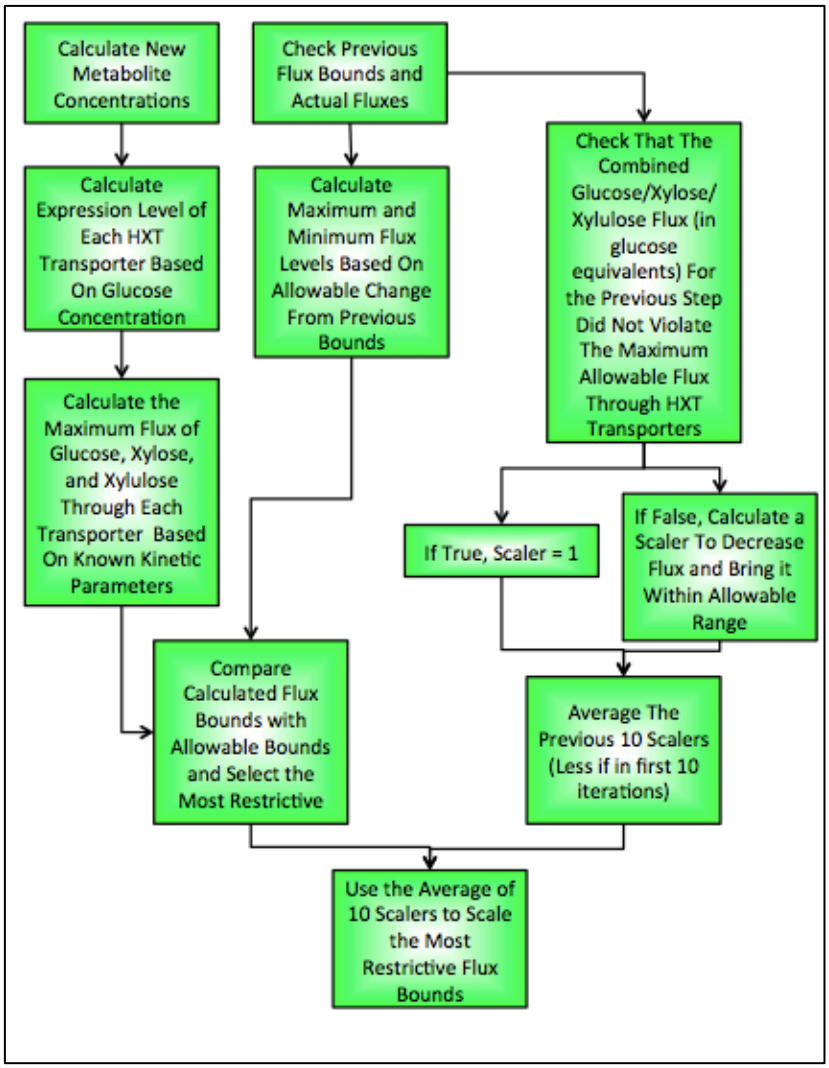


Figure 31: Flowchart of Flux Bounds Calculation Model Sub-loop - This flowchart depicts the sub-algorithm used to calculate new flux bounds for sugar uptake after each modeled time step. The algorithm uses a moving average of the scalar to smooth sugar consumption and avoid major consumption swings.

3.3.8 Model Fitting Technique

Model parameters were fit in three steps. First, the bounds for the aggregate transporter (accounting for previously unexplained transport of glucose and xylulose in HXT negative strains) were set directly from observed transport in HXT negative strains.

Second, the parameters common to all strains (dv/dt for HXT transport change, xylose size scalar, xylulose scalar, ATP maintenance, death reaction recovery percent) were set using a Monte Carlo method. The parameters were randomly generated to be within specified bounds for each parameter value. Using this randomly selected parameter set, two dFBA model runs were initiated with initial conditions for the experiments E1 and E222. These experiments using strain YRH858 (high copy XK) were chosen to allow for determination of maximum flux for the XK mediated reaction. Since the XK reaction flux determined for the strain YRH858 (high copy XK) would reflect the maximum possible XK reaction flux, this maximum value was used for all other strains in the simulations. The individual fitting runs were selected because of their high degree of mass closure. At each experimental sample time point the predicted media composition was compared via least squares method to the experimental data for biomass, glucose, xylose, xylulose, acetic acid, lactic acid, succinic acid, and glycerol concentrations. Finally, the sums of least squares for each measurement in the two experiments were added to produce a fitness objective function. This score and the parameter set that produced it were output and the process was repeated with a new set of randomly selected parameter set. After 400 repetitions, the bounds for randomly selecting each of the parameters were adjusted to cover the range of parameter values that produced the top 20 solutions. This process was repeated three times. During the final run there was little (less than 0.1%) improvement in fitness objective function (See Figure 32).

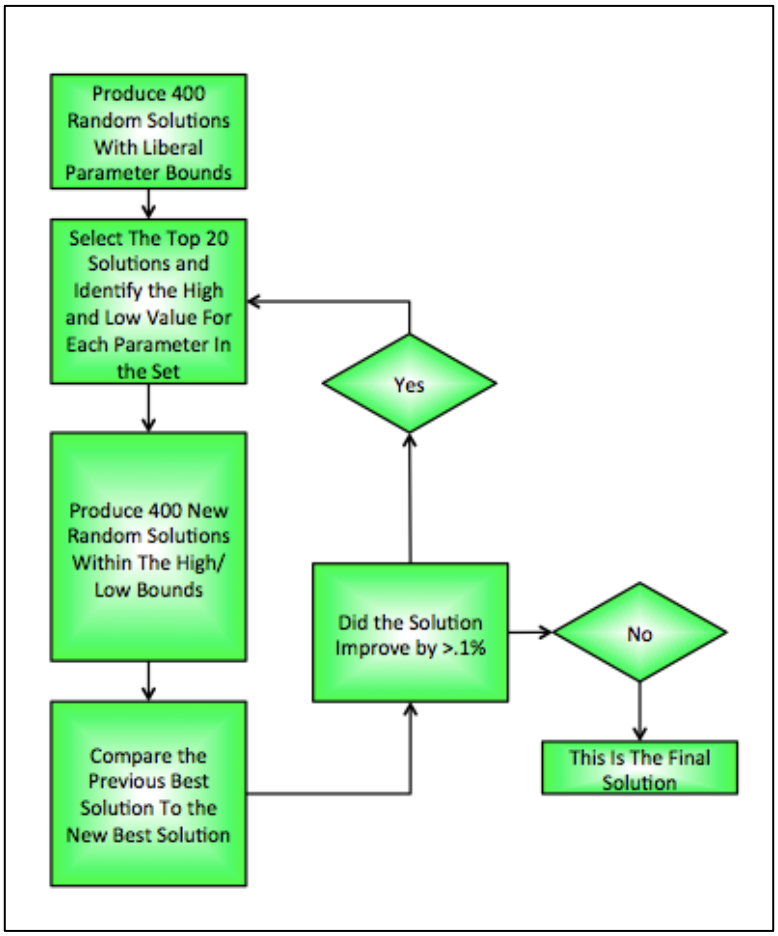


Figure 32: Flowchart of Monte Carlo Fitting Model Loop - This flowchart depicts the Monte Carlo fitting technique initially used to fit our model to selected experimental data.

Following fitting with the Monte Carlo method, a second method termed particle swarm analysis was utilized to verify the result. In particle swarm analysis a set of random solutions is created and their fitness objective functions are calculated. Each solution (or particle) is "aware" of its personal best solution and of the best solution ever found by any particle. Competing forces (one pulling towards the global best solution, one pulling towards a particle's personal best solution, and one pulling in a random direction) pull on the particle and give it a "velocity." Each particle's unique velocity transports it (and its associated parameters) to a new solution (unique set of variables) and the process is

repeated. Over several iterations the particles in the swarm should converge to the global best solution.

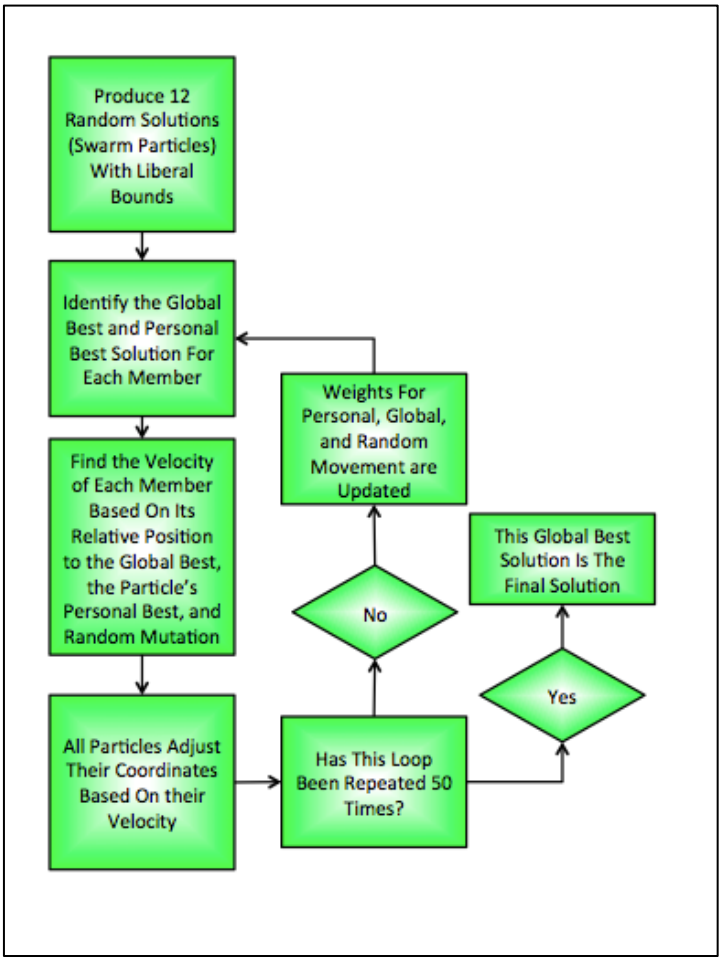


Figure 33: Flow Chart of Particle Swarm Fitting Loop - The flow chart above depicts the Particle Swarm Analysis method used to refit the model and verify the previously identified Monte Carlo fitting.

Part of the challenge in implementing this meta-heuristic was in the length of time required to run the dFBA. Normally, a swarm would be made up of tens or hundreds of members and the heuristic would iterate several hundred or thousand times. This analysis was time constrained and thus a swarm size of 12 was selected and run over 50 iterations. This allowed ample time for the solutions to converge and a high probability that a strong solution (as good or slightly

better than the one identified in the Monte Carlo method) would be found. Over 16 runs, this algorithm produced 11 solutions with better (lower) fitness objective function values than the Monte Carlo method. The lowest of these solutions found a fitness objective function of 136.056 while the Monte Carlo method identified a best solution of 136.58. The particle supports the findings of the Monte Carlo method, but this minor improvement did not visibly improve model predictions substantially. The solution from the Monte Carlo method was used in the model below.

In the third step of fitting, the allowable flux through xylulokinase was constrained for strain YRH524 (which batch experiments suggested was the only strain limited by xylulokinase capacity). This fitting used a Monte Carlo method similar to that used to fit the common model parameters, however, with only a single parameter to fit, this required far fewer runs to produce non improving fitness objective function. The XK parameter was fit using data from the first 48 hour of experiment E222 (before the 48 hour glucose spike).

Table 17: Best Parameter Fits - This table lists parameters, their description, and the best-identified fit using the Monte Carlo method described above. These values were used in modeling throughout the rest of this work.

Parameter Fits Used In Modeling		
Parameter	Description	Value
XK Capacity	Maximum Capacity of Xylulokinase (strain YRH524 and YRH1154, all other strains)	0.4376, 1000 mM/gDCW/hr
Theta max growth ratio	The maximum scalar by which the sugar consumption rate could increase in one time step	1.2932
Theta max degradation ratio	The minimum scalar by which the sugar consumption rate could decrease in one time step	0.935
Xylulose Time Scalar	The time scalar used to compare xylulose and glucose for transport	600.0975
Xylose Time Scalar* *Since xylose consumption was not possible in the model this parameter had no real meaning	The time scalar used to compare xylose and glucose for transport	55.556
Death Recovery Percentage	The percentage of cell metabolites recovered by living cell population after cell death	66.391%
ATP Maintenance	The ATP cost to maintain a cell	0.988819 mM/gDCW/hr

3.4 Results and Discussion

There were two overarching objectives in the modeling section of this work. The first was to corroborate our findings from the batch fermentation portion of this study. For reference, these findings included that

1) XK capacity was limiting in strain YRH524 (3.4.1 and 2.4.1)

2) Xylulose was transported in correlation with time (and thus its previous exposure to glucose) suggesting possible action by the HXT family of proteins (3.4.2 and 2.4.2-2.4.3)

3) Other transporters (outside HXT 1-7) were active in the transport of xylulose (3.4.3 and 2.4.5)

4) The expression of these additional transporters was non-constant. (3.4.3 and 2.4.5)

The second overarching goal for the modeling section of this work was to develop a predictive model capable of forecasting the metabolomic outcome of experiments given a limited set of metabolomic inputs. These goals are discussed at length in the subsections below.

3.4.1 XK is limiting only in strain YRH524

In the experiments discussed in Chapter 2, XK capacity was identified as the limiting reaction for xylulose utilization in strain YRH524. The experimental results supported the assessment that xylulose utilization was not limited due to XK in strains YRH857, YRH858, and YRH859 due to over expression of the XK compared in these strains. Model parameters were initially fit to strain YRH858 (high copy xylulokinase XK). In this fitting XK capacity was left unbounded (1000 mM/gDCW/hr). After the initial fitting, XK was fit using a Monte Carlo method for strains YRH524 (the wild-type strain) and YRH857 (low copy XK). If XK was the limiting factor for YRH524 or YRH857 strains the fitness objective function for the fit between the model predictions and the experimental values would be unaffected as XK limit is decreased from unbounded (1000 mM/gDCW/hr) to the maximum utilized capacity (0.44 mM/gDCW/hr for strain 524, 1.24 mM/gDCW/hr in strain 857). On further lowering the XK reaction bounds, the fitness objective

function value would decrease due to better match between the theoretical model and the experimental values. Finally, decreasing the X reaction flux below its optimum value would lead to an increase in the fitness objective value caused by an increasing mismatch between experiments and the theoretical model constrained by sub-optimum values of the XK reaction flux. This was the pattern identified with strain YRH524 suggesting that XK is limiting as shown in Figure 34. The best fit produced here was with an XK reaction flux constraint of 0.4376 mM/gDCW/hr. This fitting was used for strain YRH524 throughout the remainder of the modeling work.

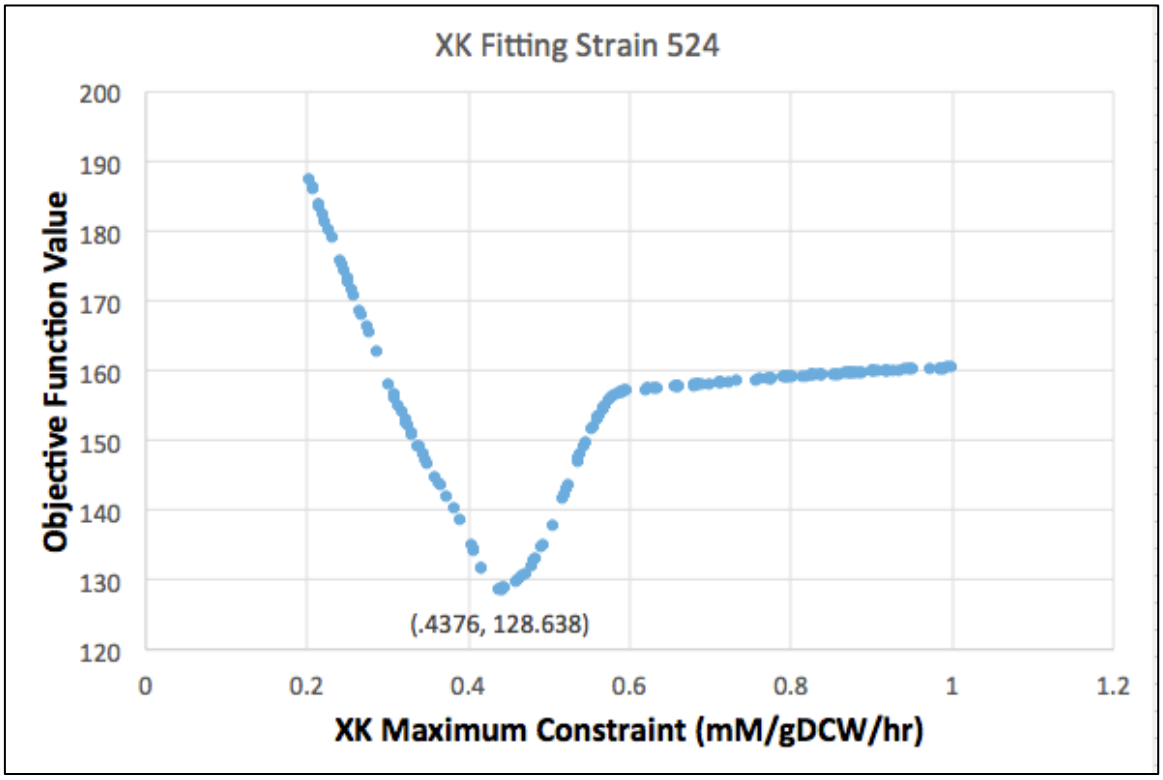


Figure 34: Xylulokinase Fitting for YRH524 (Wild Type) - The fitting of xylulokinase (XK) capacity for strain YRH524 (wild type). The inflection point here represents the XK capacity (x axis) that produces the lowest fitness objective. Further constraint produces a worsening solution.

In strain YRH858 a similar pattern was seen during the initial XK capacity constrain with no change in fitness objective value as XK reaction flux was progressively decreased from 1000mM/gDCW/hr to the maximum utilized value. However, when XK reaction flux was constrained beyond this value the value of the fitness objective function increased (indicating a poor fit) as shown in Figure 35. No fit with constrained XK proved to be better than the fit with unconstrained XK. This corroborates the finding from the fermentation experiments that XK is not limiting in strains with enhanced XK. As such, no constraint was placed on XK for YRH857 (or the strains with higher XK expression: YRH858, YRH859, and YRH1154)

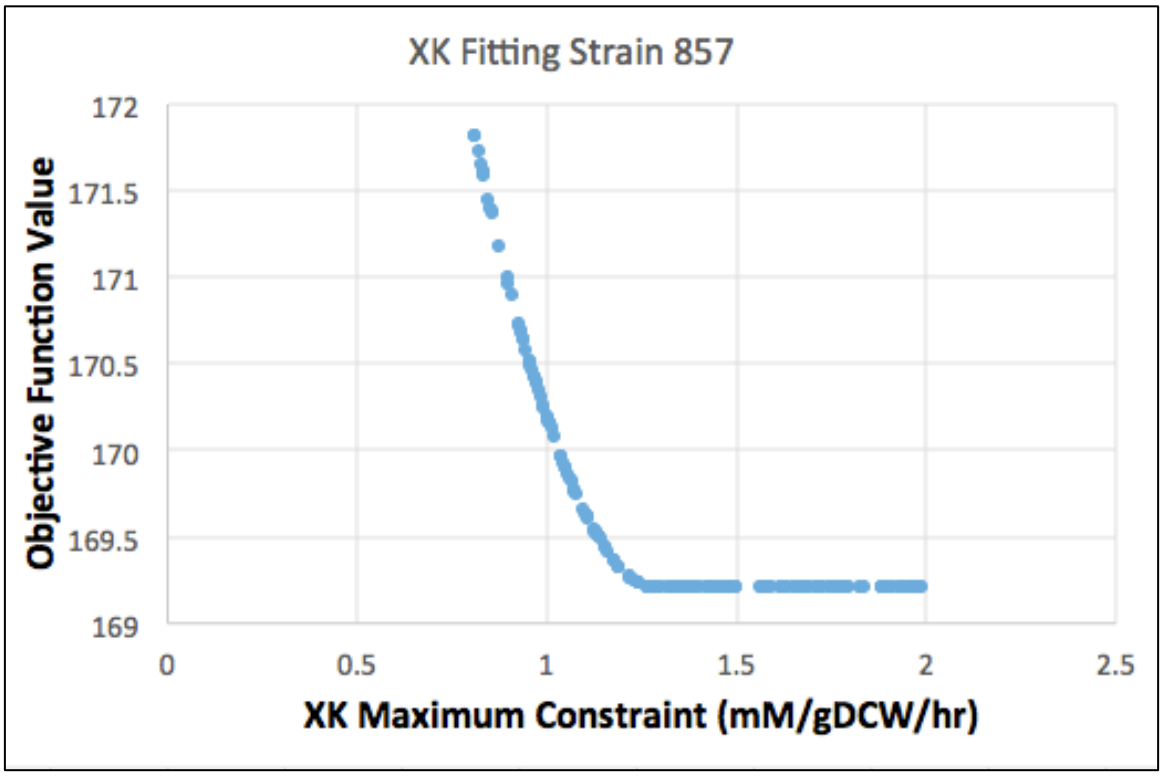


Figure 35: Xylulokinase Fitting for YRH857 (Low Copy XK Enhanced) - The fitting of xylulokinase (XK) for strain YRH857 (low copy XK enhanced). The fitness objective here never decreases (improves) with increased constraint. This suggests that XK is not a constraining factor in strain 857.

3.4.2 HXT transport explains uptake of xylulose

Our initial hypothesis was that either XK capacity or transport rate would be the limiting reaction for xylulokinase. To validate this claim, our model would need to correctly predict xylulose consumption rate in all strains based on the developed constraint of XK in strain YRH524 and the set of constant parameters used to fit the model.

There were difficulties fitting the end dynamics of batch fermentation (premature fermentation termination or over consumption of sugars). This issue and a strategy used to partially overcome it are discussed later. Since the yeast cell death is not the predominant feature in the early fermentations, the focus of this section is on the initial hours of fermentation before our yeast cell death function plays a major role in determining fermentation rate. All model simulations in this section are based on the average starting conditions of the fermentation replicates for a given strain.

Overall, modeled xylulose uptake rates during the initial fermentation closely approximated those observed during batch experiments. The internal dynamics of xylulose consumption were not strictly constrained (except in strain YRH524 where XK capacity was constrained). This suggests that the constraint of the xylulose transport pathway was sufficient to approximate xylulose utilization. This points to xylulose transport as the bottleneck for strains YRH857, YRH858, and YRH859 (all of which do not have excess XK constraint due to overexpression of XK enzyme).

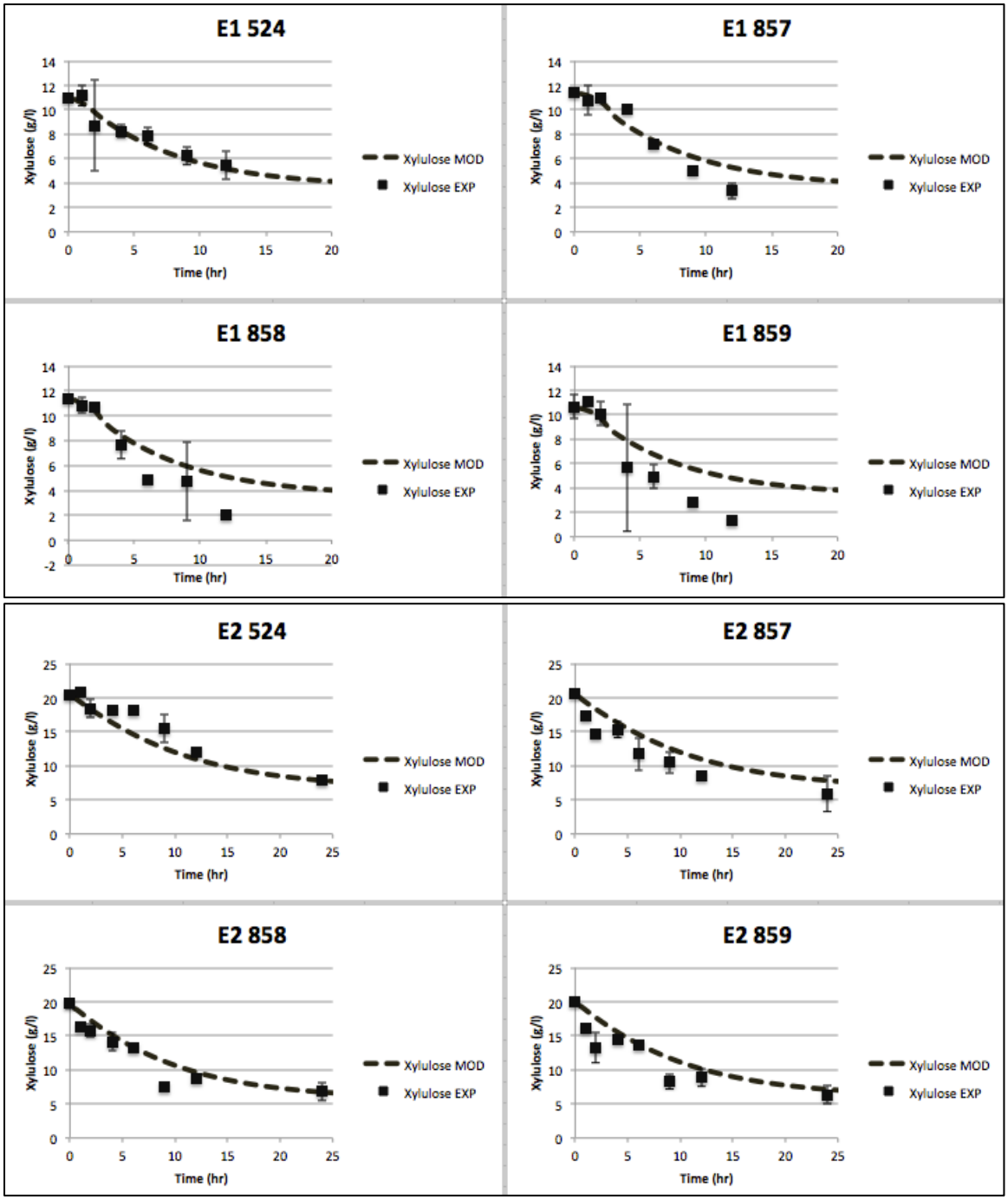


Figure 36: Modeled and Experimental Xylulose Consumption During the First 24 Hours of E1 and E2 - Low staring xylulose experiments E1 and E2 show strong agreement between model and experimental data for xylulose consumption in the first 24 hours.

Under conditions with low sugar concentrations (for example E1 and E2 above), sugars were quickly exhausted and the yeast cell death function was implemented earlier in the model. Therefore, there was divergence between

model predictions and the experimental results at earlier times (note the shorter X-axis in the figures). In contrast, the experiments with higher sugar concentrations and longer fermentation periods (as in E200, E202, and E300 below) have better agreement with the model and the experimental data.

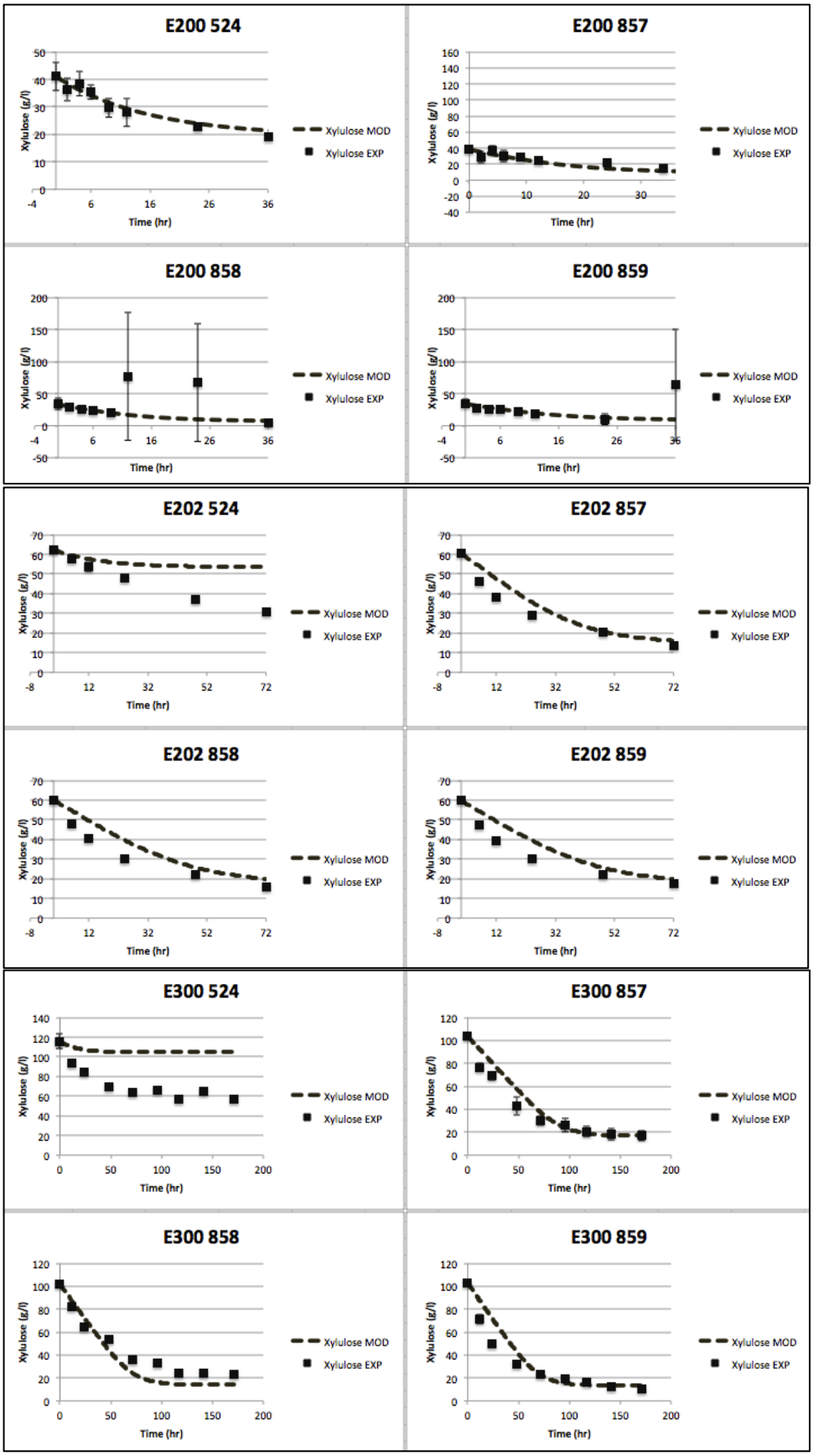


Figure 37: Modeled and Experimental Xylulose Consumption During the First 72-168 Hours of E200, E202 and E203 - Experiments with higher initial xylulose concentrations (E200, E202, E203) show strong agreement between modeled and experimental xylulose consumption over a longer time frame (note the longer x-axis shown)

In experiments with high total sugar (>40g/L), but low glucose (<20 g/L) the fitting for stain YRH524 was particularly poor. In these runs (E300 and E202) the model predicts premature termination of strain YRH524, but experimental data shows further consumption of xylulose. This may be related to the fitting of XK using a run from condition E222 (Figure 38) where high concentrations of glucose supported the high concentration of xylulose. *S. cerevisiae* is known to show diauxic growth during co-fermentation of sugars. Our model does not artificially 'rank' sugars as having a higher or lower consumption preference, but by over-constraining XK, the model could push consumption of glucose first (and thus maintain total sugar flux below 4 g/g-DCW-hr limit) resulting in a better model fit. However, when glucose is not included in the media (as for experiments E1, E2, E202, E300) and XK is over-constrained, the model may not be able to sustain cell growth through xylulose fermentation alone. This results in modeled cell death and premature termination when compared to experimental data.

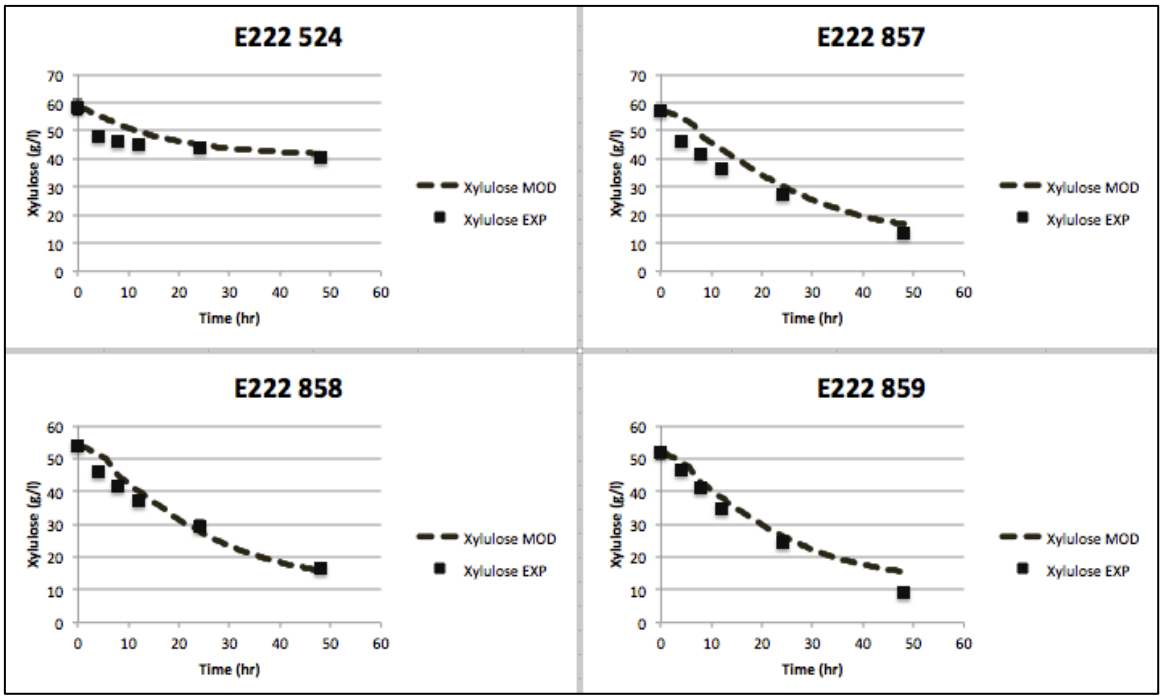


Figure38: Modeled and Experimental Xylulose Consumption During the First 48 Hours of E222 - E222 contained both high xylulose and high glucose. A run from E222 was to fit XK for strain YRH524. With glucose present, XK constraint may have pushed the model to better mirror diauxic growth known to occur in *S. cerevisiae* under mixed sugar growth. While beneficial here, this constrained XK will limit xylulose consumption under glucose free conditions possibly resulting in the premature termination of xylulose consumption present in several YRH524 runs.

3.4.3 Xylulose transport supported by other transporters not captured in model

During batch experiments we considered that transporters other than HXT 1-7 may be active in transporting xylulose into the cell. To test this theory we utilized strains YRH1153 and YRH1154 which were homologous to YRH858 and YRH524 respectively, but contained knockouts for HXT transporters 1-7. These strains showed consumption of xylulose that was non-zero, but statistically slower than YRH858 and YRH524. This suggested that other transporters were playing a role in xylulose transport and that it may be important to incorporate a basal level of xylulose transport in the model (beyond that present with HXT 1-7).

It was also noted that although the uptake of xylulose in HXT- strains was well correlated with xylulose concentration in individual experiments, the correlations were not the same between experiments (Figure 39 top). This may be due to transporter expression patterns stimulated by a different environmental condition. This made it difficult to fit a baseline for xylulose uptake from sources other than the HXT family of transporters. Ultimately the base level (incorporated in all strains in the model) was based on the lower level uptake pattern shown in experiments E222 and E300.

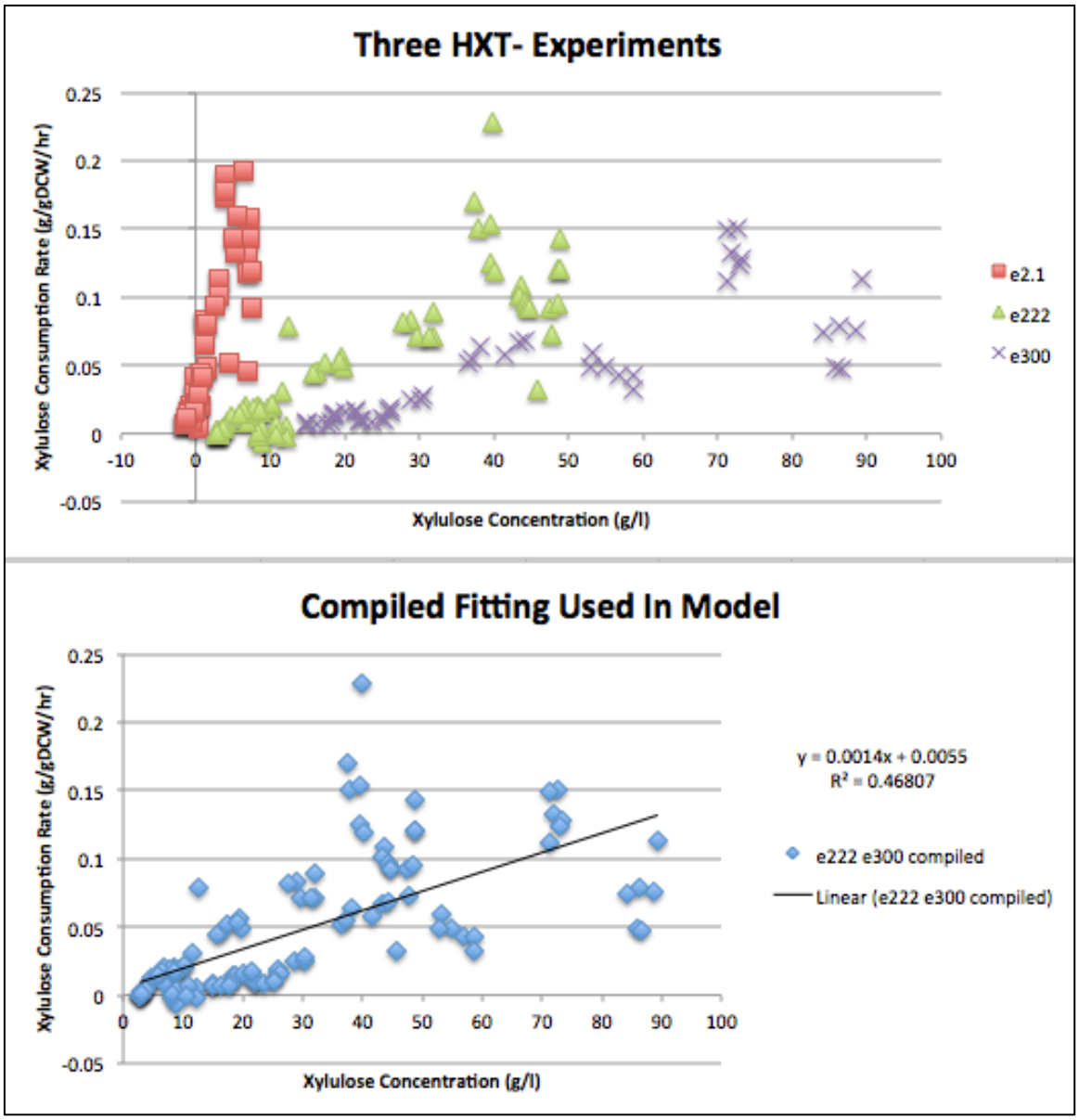


Figure 39: Xylulose Consumption Rates of HXT(-) Strains in E2.1, E222, and E300 Vs. Xylulose Concentration - (note the different correlations in the three experiments). Below, the compiled fit of HXT- xylulose consumption rates vs. xylulose concentration for E222 and E300. This relationship was used in the model as a basal level to supplement HXT based xylulose transport.

It should also be noted that protein expression patterns in knock out strains are often observed to adjust over time leading to increased cell growth (Ibarra et al.). While proteomic data was not collected for these strains, it is possible that the relatively high levels of xylulose uptake in the HXT knockout

strains (as in E2.1) is not indicative of a high level transport through non-HXT transporters in the wild-type transport strains.

Based on our base level fitting of non-HXT xylulose transport we attempted to model the experimental runs of E2.1, E222, and E300 for the knockout strains YRH1153 and YRH1154. The results are in Figure 40 below.

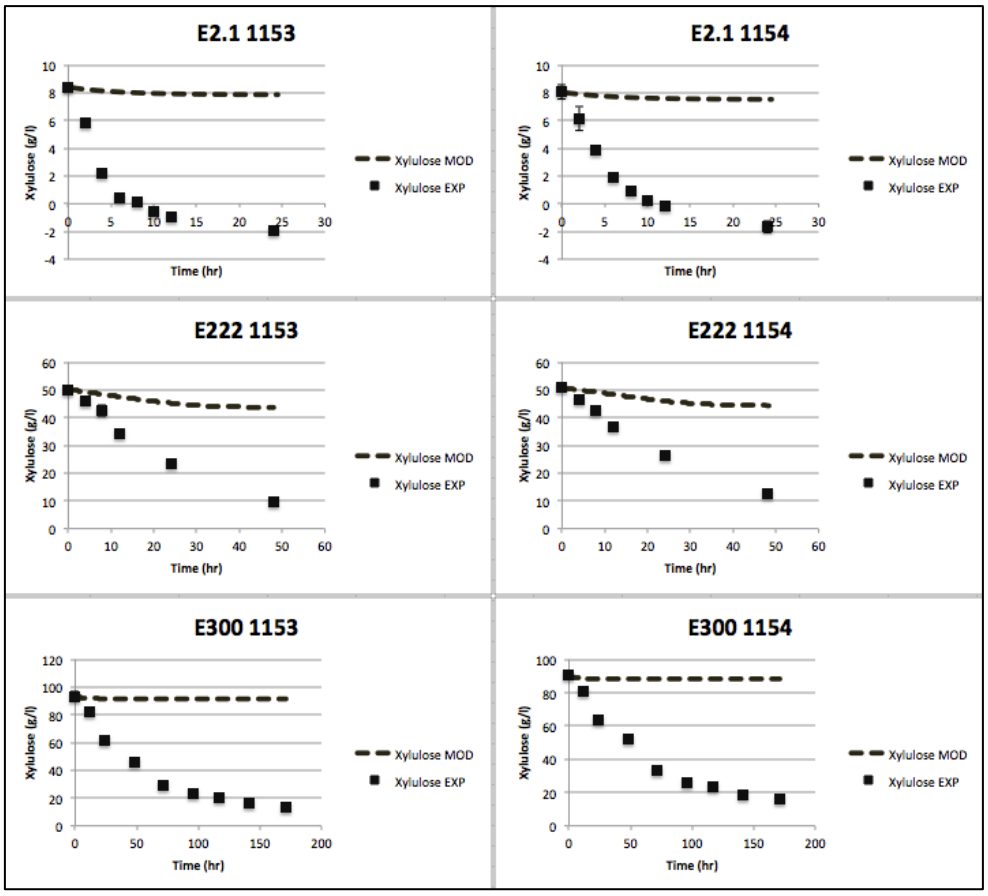


Figure 40: Poor Model Agreement Between Modeled and Experimental HXT(-) Strains - Model agreement with experimental data for HXT- strains was poor. This suggests that inducible transporters may be active. This is exacerbated by the cell death function with takes effect almost immediately as very little sugar is being taken up.

In these simulations the fitting was very poor. The poor fit in E2.1 was expected as the transporters for the HXT- strains were fit to E222 and E300 (where E2.1 had a much more positive correlation between xylulose concentration and uptake rate). The fit was expected to be better in experiments

E222 and E300, but the model predicts almost no consumption. This highlights the model deficiency in simulating end dynamics of batch fermentation where ATP produced at current sugar consumption levels fails to meet ATP maintenance. Here, the constrained sugar uptake level is predicted to lead to swift cell death even in high xylulose cases (E300) and cases in which glucose was present (E222). In future simulations identifying other potential xylulose transporters and their expression patterns will be important to more fully predicting xylulose utilization. The only conclusion that could be drawn from this portion of is that the approximation of secondary xylulose transporters was poor and that this issue exacerbates our existing difficulty modeling the end dynamics of batch fermentation.

3.4.4 Yeast Cell Death Approximation Insufficient to Explain Xylulose Attenuation

During analysis of data collected during fermentation experiments it was noted that xylulose consumption tended to terminate while xylulose concentration was still non-zero. A challenging aspect of these results from a modeling perspective was that this termination did not necessarily correspond with a similar xylulose level in all experiments. In E1 xylulose level approached zero and E2 final xylulose levels fell below 2 g/l, but in E300 consumption stopped at between 10 and 56 g/l (depending on the strain). In the original flux balance model, if an ATP maintenance requirement was enforced and could not be met the model would return no solution. Conversely, if no ATP maintenance

requirement was in place the model would always predict complete sugar consumption. In an attempt to balance these extremes we implemented a yeast cell death function that was activated to decrease the cell concentration when ATP maintenance requirement could not be met. This was based on the hypothesis that if ATP maintenance requirements were not met cell death would occur and that dying cells might release nutrients to the culture allowing other cells to fulfill ATP maintenance and survive. This hypothesis is corroborated by recent results (Bren et al. 2013) that demonstrated that bacteria maintain maximum possible growth rate on limiting nutrients until the growth abruptly stops. Strains that had higher xylulose consumption abilities might then be able to survive for longer at lower sugar levels producing lower final sugar levels.

The yeast cell death hypothesis was developed during the analysis of the batch fermentation data and during the modeling portion of this work. Consequently we were not able to retroactively collect the measurements (cell viability) needed to properly support this hypothesis or suggest a correlation between insufficient sugar and cell death for the model.

In an attempt to support the death hypothesis the end dynamics of batch fermentations were compared to predictions modeled with the death hypothesis. The yeast cell death parameter was moderately successful in predicting fermentation rates in high sugar concentration experiments (E202, E222, E300). In these experiments, the end dynamics of fermentation by strains YRH857, YRH858, and YRH859 showed strong agreement between the model and

experimental data. However, the model results for strain YRH524 showed much poorer agreement with the experiments. Strain YRH524 was predicted to prematurely terminate in both E202 and E300 compared to experiments. Early termination suggests that in the model insufficient sugar was being consumed to generate sufficient ATP to meet the ATP maintenance needs. This could be due to over-constraining of XK reaction flux as discussed earlier or an unidentified issue with the death hypothesis. The strain showed close agreement to the experimental data from E222 which was not surprising as data from this experiment was used to fit XK capacity for strain YRH524. In certain situations XK capacity has a direct effect on cell death (where XK constraint forces insufficient sugar consumption). It is possible that this is contributing to the premature termination here.

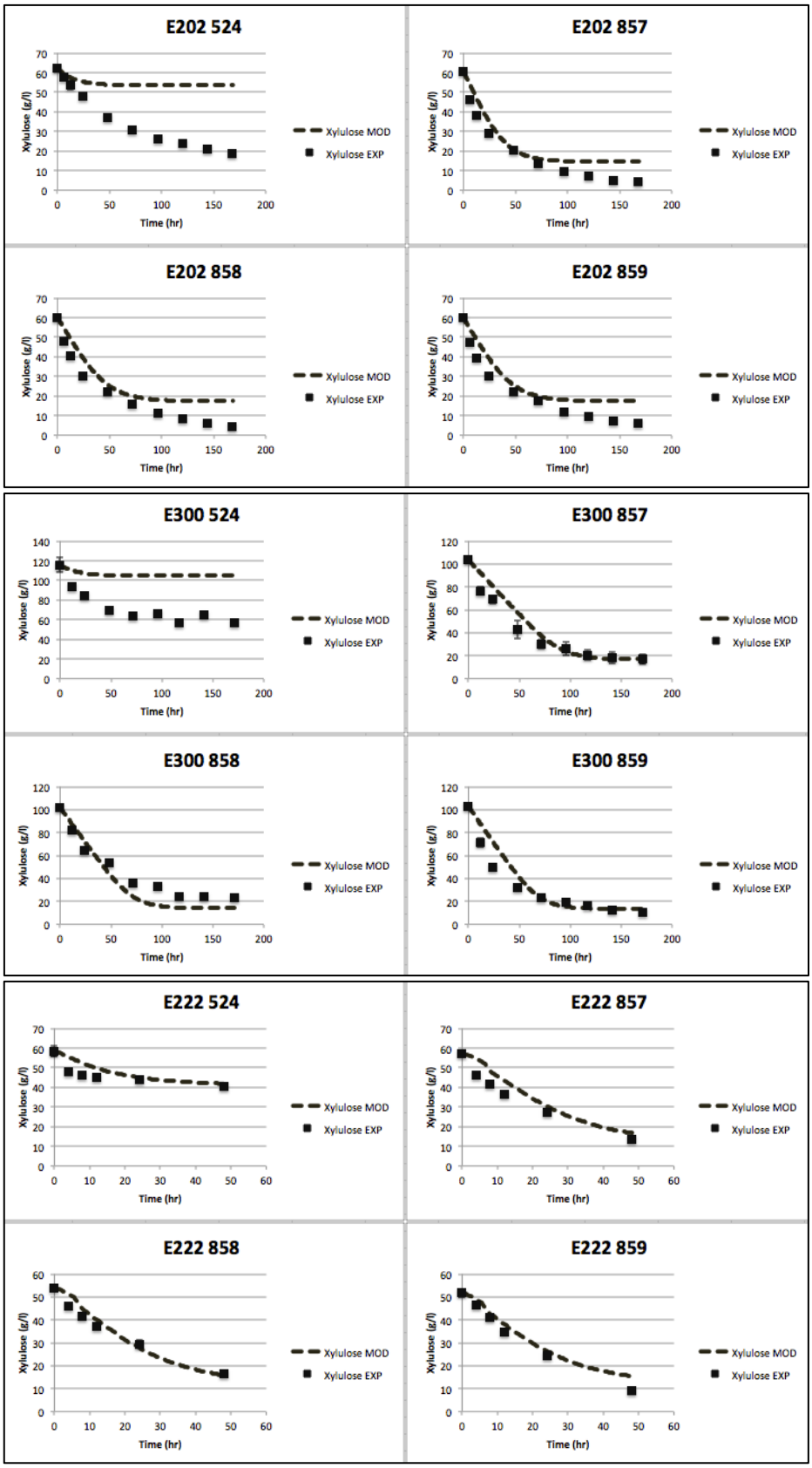


Figure 41: Modeled and Experimental End Dynamics of E202, E300, and E222 – These experiments shows good agreement for strains YRH857, YRH858, and YRH859, but premature termination for YRH524.

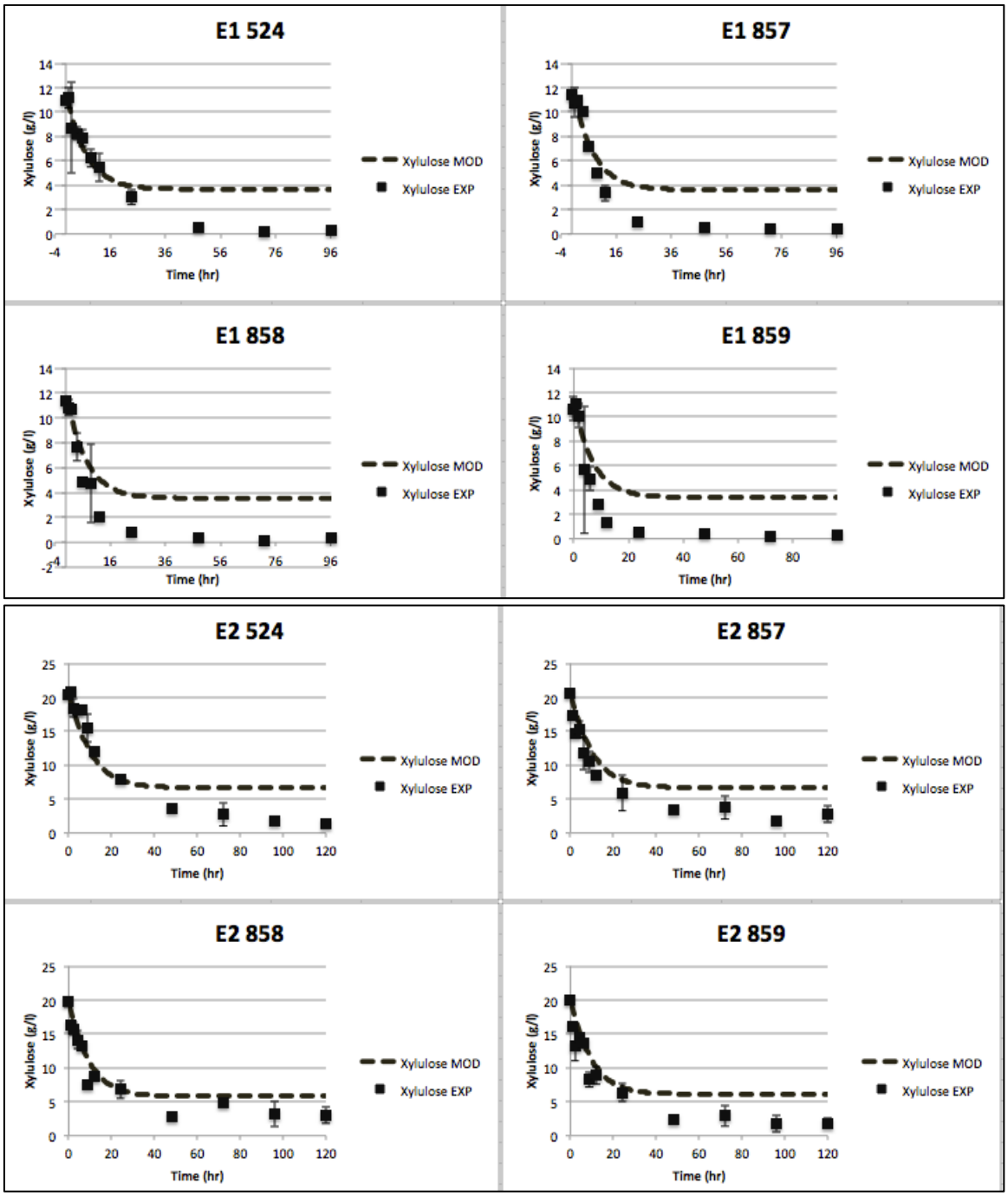


Figure 42: Modeled and Experimental End Dynamics of E1 and E2 - Experimental conditions with lower starting xylulose (E1 and E2) tend to terminate slightly prematurely for all strains

In experiments with lower xylulose concentrations (E1 and E2) the yeast cell death function tends to terminate the xylulose consumption in all strains prematurely (Figure 42). This might suggest that ATP maintenance is variable

over time or with different metabolite concentrations (lower sugar experiments would have lower final ethanol). Although our model could simulate varying ATP maintenance needs, existing data is insufficient to test this hypothesis and is outside the scope of this work. Real time measurements of cell viability over time and cell survival rates in different conditions would be needed in elucidating these relationships.

3.4.5 – Model predictions on secondary metabolites poor

One of the major objectives with this model were to better predict xylulose sugar utilization under various mixed glucose-xylose conditions, which was achieved. A perfect model would go beyond the primary objective and should predict secondary results corroborating the primary objective. In the current modeling scheme this implies the ability to accurately predict the formation and consumption of metabolites such as glycerol, acetic acid, lactic acid, and succinic acid. Here we discuss the accuracy with which our model fit these metabolites and some of the possible factors driving the experiments and model to disagree.

In general, our model did a poor job of predicting secondary metabolite production. Only a small portion of the data will be shown here, but it is representative of the secondary metabolite data as a whole. The model predicted a measurable level (0.5-2.5 g/l) of succinic acid production in every model run. This was contrary to the experimental data that never showed succinic acid production beyond 0.26 g/l and only 32.1% of experimental runs produced a measureable concentration of succinic acid. In contrast to succinic acid, the

model predicted that no measureable amounts of glycerol, acetic acid, or lactic acid would be produced. However, experimental data showed that these compounds (glycerol, acetic acid, and lactic acid) were produced in concentrations as high as 5.1, 2.6, and 1.25g/l respectively. Across all strains, glycerol was produced in 64.9% of experimental runs; compared to 57.1% with acetic acid, and 42.9% with measurable lactic acid production.

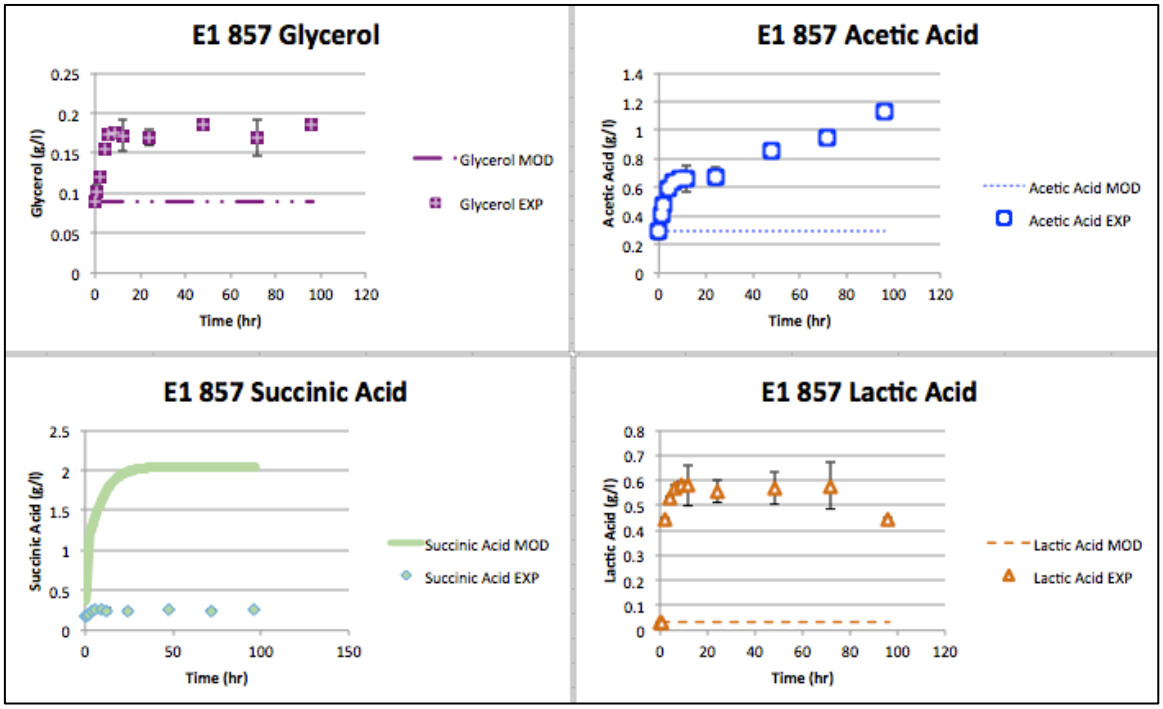


Figure 43: Experimental and Modeled Correlation of Secondary Metabolites - Secondary metabolites were not well predicted by the model. Above is the secondary metabolite comparison between the model and experimental data for E1 strain YRH857.

Elucidating the exact reason for the production of these metabolites was difficult, but it likely stems from multiple optimal solutions and the lack of regulatory elements in FBA modeling. Often secondary metabolite production is used as a method for regenerating cofactors vital to glycolysis. For example, lactic acid production from pyruvate recycles NADH to produce NAD+. The

regenerated NAD+ is required for pyruvate decarboxylation to produce acetyl CoA. In other instances secondary metabolites are produced as an alternative to cofactor regeneration. An example here would be the production of acetic acid which produces 2 ATP but does not regenerate cofactors. FBA modeling is able to capture some of this through its insistence on mass balance, but there are often multiple pathways capable of regenerating the same cofactors. This results in multiple best solutions and the model is forced to select one of the multiple optimal solutions (often the first feasible solution found by the solver in the solution space). While there are approaches such as flux variability analysis to assess this case, performing FVA for current dFBA model would be prohibitively expensive in terms of computer time. In this case the model selects succinic acid production (producing NADH in the process).

The lack of regulatory elements may also play a large role in the poor prediction of secondary metabolite production. Some aspects of cell regulation are not well captured through mass balance alone. For example, studies have shown glycerol to be produced and accumulated to regulate osmotic stress. FBA models cannot capture this accumulation because of their steady state assumption, but in addition, most basic FBA models do not account for environmental osmolality (or any signal induction) at all. Inclusion of signaling pathways would act to further constrain the model and force it to produce certain metabolites in certain conditions. Without these constraints the model finds the

solution that maximizes the objective function. In this case that solution did not include production of glycerol, acetic acid, or lactic acid.

3.4.5 Other Limitations of dFBA

The field of flux balance analysis (including dFBA and all its variants) has great potential to explain complex cell behavior, but this work has highlighted some of the issues inherent in the base method. Here we these shortcomings will be outlined and suggestions for future work will be provided.

A model is only as powerful as the inputs used to formulate the model. Our model was fit and validated using data produced in the batch fermentation section of this work. In some of this data, achieving a reliable mass closure was a challenge. This issue could be a result of a wide range of reasons ranging from ineffective sample and standard storage leading to partial evaporation of ethanol (the most volatile metabolite measured) to experimental errors. While data for ethanol was collected and is reported in the appendix of this work, it was excluded from the model fitting because of questions surrounding its accuracy. The remaining data was sufficient to fit the model and validate the fit, but it is possible that a different set of parameters would have been developed with perfect data.

A model is only as powerful as its assumptions. In the development of our model we made a series of assumptions to simplify the complexity of a living cell to a series of equations and constraints. These assumptions may not be universally applicable and may not fit all experimental conditions.

- The HXT family of proteins may have very different kinetics for xylulose than xylose and the ‘time scalar’ we fit for xylulose may be different for different transporters.
- ATP maintenance is not constant, but linked to environmental stimuli.
- Cell death may introduce new nutrients to the media, but they may be introduced in a more nuanced, time-delayed manner.
- Cells do not obey strict steady state assumption and cells may not optimize to maximize cell growth.

These assumptions are in place to create a problem that is solvable and to fill in knowledge gaps with plausible mechanisms regulating cell activity. As these assumptions improve, the model will have increased ability to capture cell dynamics under a variety of environmental conditions. Until these assumptions are near perfect the model will not accurately predict all experimental outcomes as seen in the results here for yeast cell death and end dynamics of fermentation approximation.

Protein expression patterns are not static and our model does not capture knockout adaptation. One of the assumptions with our model was that knockout HXT strains could provide insight into the level of xylulose and glucose consumption and that this base level would be present in strains containing HXT 1-7. The results do not support this assumption and it is speculated that remaining transporters in the knock out strains were up regulated to compensate

for the loss of the HXT transporters, thus biasing our parameters to higher consumption. The base model that we utilized did not show specific protein transporters and we chose to only incorporate HXT 1-7 as per our hypothesis and the past work by Bertilsson et al. To better capture cell adaptation and transport dynamics in knock out organisms a full transporter network and expression profile would need to be developed.

3.5 Modeling Conclusions

The modeling portion of this work produced data supporting several of the conclusions we formulated during the batch fermentation section. Model fit was found to improve with restriction of XK capacity in strain YRH524, but not in strains YRH857, YRH858, or YRH859. This supports the XK bottleneck identified for strain YRH524 during the fermentation experiments. The constraint of HXT transporters was sufficient to approximate the xylulose utilization rate in strains YRH857, YRH858, and YRH859 without further internal xylulose utilization constraints. This suggests that xylulose transport was the bottleneck in xylulose utilization in these strains. Finally, we showed that basal level xylulose transport in strains YRH1153 and YRH1154 was insufficient to explain observed xylulose transport in those strains suggesting that these strains adapted to increase xylulose utilization capacity.

The model produced showed strong fit for glucose and xylulose consumption, but poor fit of secondary metabolite formation (glycerol, acetic acid, lactic acid, and succinic acid). The yeast cell death function helped to stop

xylulose fermentation before xylulose concentration reached zero, however the assumptions used were not sufficient to accurately model when the fermentation would stop. Further experimental data (including cell viability) would be helpful fitting this function in the future.

Chapter 4 - Conclusions And Future Work:

This work consisted of a wet lab, experimental portion and an in silica modeling portion jointly aimed at better understanding the bottlenecks of xylulose utilization in wild type and xylulokinase enhanced strains of *S. Cerevisiae*. This dual approach yielded several major findings, identified during the wet lab phase and supported during modeling. By comparing xylulose utilization rates of YRH524 (wild type) and the XK enhanced strains YRH857 and YRH858 we were able to identify XK as the rate limiting reaction in the wild-type strain. During model fitting, this finding was supported by the constraint of XK capacity. Similarly, batch experiments showed that strains with a high-level increase in XK capacity (YRH858) did not perform better than those with a low-level increase in XK capacity. This suggested that XK was not limiting in these strains. The flux balance model again supported this finding through XK fitting. Accordingly, future work on strain development should focus on increasing capacity in enzymes other than xylulokinase.

Xylulose transport was considered as another possible xylulose utilization bottleneck. By taking a novel approach to adapting Bertillson's HXT transport, xylose consumption model for use with xylulose a robust agreement was

achieved between our experimental data and model suggesting that HXT 1-7 were a likely candidate for xylulose transporters. This model assumed that the kinetic parameters are similar to those for xylose transport. Future work should include identifying the xylulose specific kinetic parameters for each of these transporters.

HXT knockout strains were used to determine if other transporters were active in transporting xylulose. The knockout strains showed a decreased rate of xylulose utilization, but the level was inconsistent, poorly correlated with xylulose concentration, and higher than the background level expected. This suggests that HXT knockout cells may adapt by up regulating the expression of other transporters to facilitate sufficient sugar transport.

This work focused only on a small group of HXT transporters, but our data suggests a larger number of hexose transporters are active in xylulose transport. The model could be improved by understanding and incorporating the expression patterns of these transporters and their xylulose specific kinetics. Since these transporters appear to be expressed differently in knock out cells MOMA or ROOM may be a more appropriate method for predicting expression patterns. Alternatively the expression of transporters could be described as a Boolean, cost based optimization problem associating the expression of the transporter with a cost and weighing that cost against the improvement in energy production.

The model itself is the final deliverable for this thesis. It performs well to predict glucose and xylulose uptake under anaerobic conditions with ample

sugar. It performs less well in terms of secondary metabolite predictions and predicting end dynamics of batch fermentations during conditions of starvation and yeast cell death. To our knowledge this was the first attempt to model cell death via flux balance analysis. The yeast cell death function used in the model was a best guess approximation that attempted to explain the experimental observations. Cell death was not quantified in the experiments and collecting this data would give better insight into the shortcomings of the yeast cell death function.

Chapter 5 – Bibliography:

Agbogbo, F.K., and Coward-Kelly, G. (2008). Cellulosic ethanol production using the naturally occurring xylose-fermenting yeast, Pichia stipitis. Biotechnol. Lett. *30*, 1515–1524.

Agbogbo, F.K., Coward-Kelly, G., Torry-Smith, M., and Wenger, K.S. (2006). Fermentation of glucose/xylose mixtures using Pichia stipitis. Process Biochem. *41*, 2333–2336.

Auld, D. (2012). The Economics of Ethanol, Agriculture and Food. J. Sustain. Dev. *5*.

Bertilsson, M., Andersson, J., and Lidén, G. (2007). Modeling simultaneous glucose and xylose uptake in Saccharomyces cerevisiae from kinetics and gene expression of sugar transporters. Bioprocess Biosyst. Eng. *31*, 369–377.

Borrion, A.L., McManus, M.C., and Hammond, G.P. (2012). Environmental life cycle assessment of lignocellulosic conversion to ethanol: A review. Renew. Sustain. Energy Rev. *16*, 4638–4650.

British Petroleum (2012). BP statistcal review of world energy 2012.pdf.

British Petroleum (2013). BP_World_Energy_Outlook_booklet_2013.pdf.

Bruinenberg, P.M., Dijken, J.P., and Scheffers, W.A. (1983). The role of redox balances in the anaerobic fermentation of xylose by yeasts. Appl. Microbiol. Biotechnol. *18*, 287–292.

Chang, S.-F., and Ho, N.W. (1988). Cloning the yeast xylulokinase gene for the improvement of xylose fermentation. Appl. Biochem. Biotechnol. *17*, 313–318.

Cheng, C.-L., Lo, Y.-C., Lee, K.-S., Lee, D.-J., Lin, C.-Y., and Chang, J.-S. (2011). Biohydrogen production from lignocellulosic feedstock. Bioresour. Technol. *102*, 8514–8523.

Covert, M.W., and Palsson, B.O. (2003). Constraints-based models: Regulation of Gene Expression Reduces the Steady-state Solution Space. J. Theor. Biol. *221*, 309–325.

Covert, M.W., Schilling, C.H., and Palsson, B. (2001). Regulation of Gene Expression in Flux Balance Models of Metabolism. J. Theor. Biol. *213*, 73–88.

Covert, M.W., Xiao, N., Chen, T.J., and Karr, J.R. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli. Bioinformatics *24*, 2044–2050.

Delgenes, J.P., Moletta, R., and Navarro, J.M. (1996). Effects of lignocellulose degradation products on ethanol fermentations of glucose and xylose by< i> Saccharomyces cerevisiae, Zymomonas mobilis, Pichia stipitis</i>, and< i> Candida shehatae</i>. Enzyme Microb. Technol. *19*, 220–225.

Demirer, R., Kutan, A.M., and Shen, F. (2012). The effect of ethanol listing on corn prices: Evidence from spot and futures markets. Energy Econ. *34*, 1400–1406.

Diderich, J.A. (1999). Glucose Uptake Kinetics and Transcription of HXT Genes in Chemostat Cultures of Saccharomyces cerevisiae. J. Biol. Chem. *274*, 15350–15359.

Van Dijken, J.P., Bauer, J., Brambilla, L., Duboc, P., Francois, J.M., Gancedo, C., Giuseppin, M.L.F., Heijnen, J.J., Hoare, M., and Lange, H.C. (2000). An interlaboratory comparison of physiological and genetic properties of four< i> Saccharomyces cerevisiae</i> strains. Enzyme Microb. Technol. *26*, 706–714.

Du, J., Li, S., and Zhao, H. (2010). Discovery and characterization of novel d-xylose-specific transporters from Neurospora crassa and Pichia stipitis. Mol. Biosyst. *6*, 2150.

Emmerling, M., Dauner, M., Ponti, A., Fiaux, J., Hochuli, M., Szyperski, T., Wuthrich, K., Bailey, J.E., and Sauer, U. (2002). Metabolic Flux Responses to Pyruvate Kinase Knockout in Escherichia coli. J. Bacteriol. *184*, 152–164.

Gelfand, I., Sahajpal, R., Zhang, X., Izaurralde, R.C., Gross, K.L., and Robertson, G.P. (2013). Sustainable bioenergy production from marginal lands in the US Midwest. Nature *493*, 514–517.

Ghosh, A., Zhao, H., and Price, N.D. (2011). Genome-Scale Consequences of Cofactor Balancing in Engineered Pentose Utilization Pathways in Saccharomyces cerevisiae. Plos One *6*, e27316.

Gonçalves, E., Bucher, J., Ryll, A., Niklas, J., Mauch, K., Klamt, S., Rocha, M., and Saez-Rodriguez, J. (2013). Bridging the layers: towards integration of signal transduction, regulation and metabolism into mathematical models. Mol. Biosyst. *9*, 1576.

Gong, C.-S., Chen, L.-F., Flickinger, M.C., Chiang, L.-C., and Tsao, G.T. (1981). Production of ethanol from D-xylose by using D-xylose isomerase and yeasts. Appl. Environ. Microbiol. *41*, 430–436.

Ha, S.-J., Kim, S.R., Choi, J.-H., Park, M.S., and Jin, Y.-S. (2011). Xylitol does not inhibit xylose fermentation by engineered Saccharomyces cerevisiae expressing xylA as

severely as it inhibits xylose isomerase reaction in vitro. Appl. Microbiol. Biotechnol. *92*, 77–84.

Hamacher, T., Becker, J., Gardonyi, M., Hahn-Hagerdal, B., and Boles, E. (2002). Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization.

Hanly, T.J., and Henson, M.A. (2011). Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. Biotechnol. Bioeng. *108*, 376–385.

Hendriks, A.T.W.M., and Zeeman, G. (2009). Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour. Technol. *100*, 10–18.

Hill, J., Nelson, E., Tilman, D., Polasky, S., and Tiffany, D. (2006). Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. Proc. Natl. Acad. Sci. *103*, 11206–11210.

Hochster, R.M., and Watson, R.W. (1954). Enzymatic Isomerization of d-xylose to d-xylulose. Arch. Biochem. Biophys. *48*, 120–129.

Hotta, A., Tanino, T., Ito, T., Hasunuma, T., Ogino, C., Kondo, A., and Ohmura, N. (2009). Bioethanol production from mixed sugars using sugar uptake ability enhanced yeast strain by overexpression of transporters. J. Biosci. Bioeng. *108*, S53.

Hsiao, H.-Y., Chiang, L.-C., Chen, L.-F., and Tsao, G.T. (1982). Effects of Borate on Isomerization and Yeast Fermentation of High xylulose solution and acid hydrolysate of hemicellulose. Enzyme Microb. Technol. *4*, 25–31.

Ibarra, R., Edwards, J.S., and Palsson, B. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth.

Jeffries, T.W., and Jin, Y.-S. (2004). Metabolic engineering for improved fermentation of pentoses by yeasts. Appl. Microbiol. Biotechnol. *63*, 495–509.

Jeong, J., Jeon, H., Ko, K., Chung, B., and Choi, G.-W. (2012). Production of anhydrous ethanol using various PSA (Pressure Swing Adsorption) processes in pilot plant. Renew. Energy *42*, 41–45.

Jeppsson, H., Yu, S., and Hahn-Hägerdal, B. (1996). Xylulose and glucose fermentation by Saccharomyces cerevisiae in chemostat culture. Appl. Environ. Microbiol. *62*, 1705–1709.

Johansson, B., and Hahn-Hägerdal, B. (2002). The non-oxidative pentose phosphate pathway controls the fermentation rate of xylulose but not of xylose in Saccharomyces cerevisiae TMB3001. Fems Yeast Res. *2*, 277–282.

Jurgens, G., Survase, S., Berezina, O., Sklavounos, E., Linnekoski, J., Kurkijärvi, A., Väkevä, M., Heiningen, A., and Granström, T. (2012). Butanol production from lignocellulosics. Biotechnol. Lett. *34*, 1415–1434.

Kabel, M.A., Bos, G., Zeevalking, J., Voragen, A.G.J., and Schols, H.A. (2007). Effect of pretreatment severity on xylan solubility and enzymatic breakdown of the remaining cellulose from wheat straw. Bioresour. Technol. *98*, 2034–2042.

Karhumaa, K., Sanchez, R.G., Hahn-Hägerdal, B., and Gorwa-Grauslund, M.-F. (2007). Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant Saccharomyces cerevisiae. Microb. Cell Factories *6*, 5.

Katz, M., and Reese, E.T. (1968). Production of glucose by enzymatic hydrolysis of cellulose. Appl. Microbiol. *16*, 419.

Kauffman, K.J., Prakash, P., and Edwards, J.S. (2003). Advances in flux balance analysis. Curr. Opin. Biotechnol. *14*, 491–496.

Kazi, F.K., Fortman, J.A., Anex, R.P., Hsu, D.D., Aden, A., Dutta, A., and Kothandaraman, G. (2010). Techno-economic comparison of process technologies for biochemical ethanol production from corn stover. Fuel *89*, S20–S28.

Kötter, P., and Ciriacy, M. (1993). Xylose fermentation by Saccharomyces cerevisiae. Appl. Microbiol. Biotechnol. *38*, 776–783.

Kötter, P., Amore, R., Hollenberg, C.P., and Ciriacy, M. (1990). Isolation and characterization of the Pichia stipitis xylitol dehydrogenase gene, XYL2, and construction of a xylose-utilizing Saccharomyces cerevisiae transformant. Curr. Genet. *18*, 493–500.

Kudakasseril Kurian, J., Raveendran Nair, G., Hussain, A., and Vijaya Raghavan, G.S. (2013). Feedstocks, logistics and pre-treatment processes for sustainable lignocellulosic biorefineries: A comprehensive review. Renew. Sustain. Energy Rev. *25*, 205–219.

Kumar, D., and Murthy, G. (2011). Life cycle assessment of energy and GHG emissions during ethanol production from grass straws using 45 various pretreatment processes.

Kumar, D., Juneja, A., Hohenschuh, W., Williams, J.D., and Murthy, G.S. (2012). Chemical composition and bioethanol potential of different plant species found in Pacific Northwest conservation buffers. J. Renew. Sustain. Energy *4*, 063114.

Lal, R. (2005). World crop residues production and implications of its use as a biofuel. Environ. Int. *31*, 575–584.

Lee, S.Y., Park, J.H., Jang, S.H., Nielsen, L.K., Kim, J., and Jung, K.S. (2008). Fermentative butanol production by clostridia. Biotechnol. Bioeng. *101*, 209–228.

Lee, T.-H., Kim, M.-D., Park, Y.-C., Bae, S.-M., Ryu, Y.-W., and Seo, J.-H. (2003). Effects of xylulokinase activity on ethanol production from d-xylulose by recombinant Saccharomyces cerevisiae. J. Appl. Microbiol. *95*, 847–852.

Liang, H., and Gaber, R.F. (1996). A novel signal transduction pathway in Saccharomyces cerevisiae defined by Snf3-regulated expression of HXT6. Mol. Biol. Cell *7*, 1953.

Lin, Y., Zhang, W., Li, C., Sakakibara, K., Tanaka, S., and Kong, H. (2012). Factors affecting ethanol fermentation using Saccharomyces cerevisiae BY4742. Biomass Bioenergy *47*, 395–401.

Liu, S.-Y., Wiegel, J., and Gherardini, F.C. (1996). Purification and cloning of a thermostable xylose (glucose) isomerase with an acidic pH optimum from Thermoanaerobacterium strain JW/SL-YS 489. J. Bacteriol. *178*, 5938–5945.

Lochhead, A.G., and Farrell, L. (1930). The Effect of Preservatives on Fermentation by Sugar-Tolerant Yeasts from Honey. Can. J. Res. *3*, 95–103.

Lönn, A., Träff-Bjerre, K.., Cordero Otero, R.., van Zyl, W.., and Hahn-Hägerdal, B. (2003). Xylose isomerase activity influences xylose fermentation with recombinant Saccharomyces cerevisiae strains expressing mutated xylA from Thermus thermophilus. Enzyme Microb. Technol. *32*, 567–573.

Lynd, L., Cushman, J., Nicholas, R., and Wyman, C. (1991). Fuel Ethanol from cellulosic biomass.

Margeot, A., Hahn-Hagerdal, B., Edlund, M., Slade, R., and Monot, F. (2009). New improvements for lignocellulosic ethanol. Curr. Opin. Biotechnol. *20*, 372–380.

Matsushika, A., Inoue, H., Kodaki, T., and Sawayama, S. (2009). Ethanol production from xylose in engineered Saccharomyces cerevisiae strains: current state and perspectives. Appl. Microbiol. Biotechnol. *84*, 37–53.

Matsushika, A., Goshima, T., Fujii, T., Inoue, H., Sawayama, S., and Yano, S. (2012). Characterization of non-oxidative transaldolase and transketolase enzymes in the pentose phosphate pathway with regard to xylose utilization by recombinant Saccharomyces cerevisiae. Enzyme Microb. Technol. *51*, 16–25.

Min Lee, J., Gianchandani, E.P., Eddy, J.A., and Papin, J.A. (2008). Dynamic Analysis of Integrated Signaling, Metabolic, and Regulatory Networks. Plos Comput. Biol. *4*, e1000086.

Mo, M.L., Palsson, B.Ø., and Herrgård, M.J. (2009). Connecting extracellular metabolomic measurements to intracellular flux states in yeast. Bmc Syst. Biol. *3*, 37.

Mu, X., Sun, W., Liu, C., and Wang, H. (2011). Improved efficiency of separate hexose and pentose fermentation from steam-exploded corn stalk for butanol production using Clostridium beijerinckii. Biotechnol. Lett. *33*, 1587–1591.

Olofsson, K., Bertilsson, M., and Lidén, G. (2008). A short review on SSF – an interesting process option for ethanol production from lignocellulosic feedstocks. Biotechnol. Biofuels *1*, 7.

Orth, J.D., Thiele, I., and Palsson, B.Ø. (2010). What is flux balance analysis? Nat. Biotechnol. *28*, 245–248.

Ozcan, S., and Johnston, M. (1995). Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose. Mol. Cell. Biol. *15*, 1564–1572.

Palmqvist, E., and Hahn-Hägerdal, B. (2000). Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. Bioresour. Technol. *74*, 17–24.

Papoutsakis, E.T. (1984). constraint based modeling.pdf. Biotechnol. Bioeng. *26*, 174–187.

Pedersen, M., and Meyer, A.S. (2010). Lignocellulose pretreatment severity – relating pH to biomatrix opening. New Biotechnol. *27*, 739–750.

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

Rao, K., Chelikani, S., Relue, P., and Varanasi, S. (2008). A Novel Technique that Enables Efficient Conduct of Simultaneous Isomerization and Fermentation (SIF) of Xylose. Appl. Biochem. Biotechnol. *146*, 101–117.

Reese, E. (1955). Enzymatic Hydrolysis of Cellulose.

Reifenberger, E., Boles, E., and Michael, C. (1997). Kinetic characterization of individual hexose transporters of Saccharomyces cerevisiae and their relation to the triggering mechanisms of glucose repression. Eur. J. Biochem. *245*, 324–333.

Richard, P., Toivari, M.H., and Penttilä, M. (2006). The role of xylulokinase in Saccharomyces cerevisiae xylulose catabolism. Fems Microbiol. Lett. *190*, 39–43.

Runquist, D., Hahn-Hagerdal, B., and Radstrom, P. (2010). Comparison of heterologous xylose transporters in recombinant Saccharomyces cerevisiae. Biotechnol Biofuels *3*.

Saloheimo, A., Rauta, J., Stasyk, O.V., Sibirny, A.A., Penttilä, M., and Ruohonen, L. (2006). Xylose transport studies with xylose-utilizing Saccharomyces cerevisiae strains expressing heterologous and homologous permeases. Appl. Microbiol. Biotechnol. *74*, 1041–1052.

Sassner, P., Galbe, M., and Zacchi, G. (2008). Techno-economic evaluation of bioethanol production from three different lignocellulosic materials. Biomass Bioenergy *32*, 422–430.

Savinell, J.M., and Palsson, B.O. (1992). Optimal selection of metabolic fluxes for< i> in vivo</i> measurement. I. Development of mathematical methods. J. Theor. Biol. *155*, 201–214.

Sedlak, M., and Ho, N.W.Y. (2004). Characterization of the effectiveness of hexose transporters for transporting xylose during glucose and xylose co-fermentation by a recombinantSaccharomyces yeast. Yeast *21*, 671–684.

Segre, D., Vitkup, D., and Church, G.M. (2002). Analysis of optimality in natural and perturbed metabolic networks. Proc. Natl. Acad. Sci. *99*, 15112–15117.

Sharma, S., and Tauro, P. (1986). Control of ethanol production by yeast: Role of pyruvate decarboxylase and alcohol dehydrogenase. Biotechnol. Lett. *8*, 735–738.

Shlomi, T., Berkman, O., and Ruppin, E. (2005). Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proc. Natl. Acad. Sci. U. S. A. *102*, 7695–7700.

Shlomi, T., Eisenberg, Y., Sharan, R., and Ruppin, E. (2007). A genome-scale computational study of the interplay between transcriptional regulation and metabolism. Mol. Syst. Biol. *3*.

Silva, C.R., Zangirolami, T.C., Rodrigues, J.P., Matugi, K., Giordano, R.C., and Giordano, R.L.C. (2012). An innovative biocatalyst for production of ethanol from xylose in a continuous bioreactor. Enzyme Microb. Technol. *50*, 35–42.

Smith, K.W., and Johnson, S.L. (1976). Borate inhibition of yeast alcohol dehydrogenase. Biochemistry (Mosc.) *15*, 560–565.

Sonnleitnert, B. (1986). Growth of Saccharomyces cerevisiae is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis.

Sun, Y., and Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour. Technol. *83*, 1–11.

Tanino, T., Hotta, A., Ito, T., Ishii, J., Yamada, R., Hasunuma, T., Ogino, C., Ohmura, N., Ohshima, T., and Kondo, A. (2010). Construction of a xylose-metabolizing yeast by genome integration of xylose isomerase gene and investigation of the effect of xylitol on fermentation. Appl. Microbiol. Biotechnol. *88*, 1215–1221.

Tanino, T., Ito, T., Ogino, C., Ohmura, N., Ohshima, T., and Kondo, A. (2012). Sugar consumption and ethanol fermentation by transporter-overexpressed xylose-metabolizing Saccharomyces cerevisiae harboring a xyloseisomerase pathway. J. Biosci. Bioeng. *114*, 209–211.

Thompson, P. (2012). The Agricultural Ethics of Biofuels: The Food vs. Fuel Debate. Agriculture *2*, 339–358.

Träff, K.L., Cordero, R.O., Van Zyl, W.H., and Hahn-Hägerdal, B. (2001). Deletion of the GRE3 Aldose Reductase Gene and Its Influence on Xylose Metabolism in Recombinant Strains of Saccharomyces cerevisiae Expressing thexylA and XKS1 Genes. Appl. Environ. Microbiol. *67*, 5668–5674.

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

Verduyn, C., Van Kleef, R., Frank, J., Schreuder, H., Van Dijken, J.P., and Scheffers, W.A. (1985). Properties of the NAD (P) H-dependent xylose reductase from the xylose-fermenting yeast Pichia stipitis. Biochem. J. *226*, 669.

Walfridsson, M., Bao, X., Anderlund, M., Lilius, G., Bülow, L., and Hahn-Hägerdal, B. (1996). Ethanolic fermentation of xylose with Saccharomyces cerevisiae harboring the Thermus thermophilus xylA gene, which expresses an active xylose (glucose) isomerase. Appl. Environ. Microbiol. *62*, 4648–4651.

Wang, P.Y., and Schneider, H. (1980). Growth of yeasts on D-xylulose. Can. J. Microbiol. *26*, 1165–1168.

Wang, M.Q., Han, J., Haq, Z., Tyner, W.E., Wu, M., and Elgowainy, A. (2011). Energy and greenhouse gas emission effects of corn and cellulosic ethanol with technology improvements and land use changes. Biomass Bioenergy *35*, 1885–1896.

Wingren, A., Galbe, M., and Zacchi, G. (2003). Techno-Economic Evaluation of Producing Ethanol from Softwood: Comparison of SSF and SHF and Identification of Bottlenecks. Biotechnol. Prog. *19*, 1109–1117.

Wintermute, E.H., Lieberman, T.D., and Silver, P.A. (2013). An objective function exploiting suboptimal solutions in metabolic networks. Bmc Syst. Biol. *7*, 98.

Yang, B., and Wyman, C.E. (2008). Pretreatment: the key to unlocking low-cost cellulosic ethanol. Biofuels Bioprod. Biorefining *2*, 26–40.

Yang, Y., Bae, J., Kim, J., and Suh, S. (2012). Replacing Gasoline with Corn Ethanol Results in Significant Environmental Problem-Shifting. Environ. Sci. Technol. *46*, 3671–3678.

(2007). US energy independence act.pdf.