NOTE TO USERS

DISSERTATION

Optimization of High Solids Lignocellulosic Biomass Conversion for Ethanol Production

Submitted by

David Hodge

Department of Chemical Engineering

In partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Colorado State University

Ft. Collins, Colorado

Summer 2005

COLORADO STATE UNIVERSITY

Date 02/07/05

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY **DAVID HODGE** ENTITLED "**OPTIMIZATION OF HIGH SOLIDS LIGNOCELLULOSIC BIOMASS CONVERSION FOR ETHANOL PRODUCTION**" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Advisor

Department Head

ABSTRACT OF DISSERTATION

OPTIMIZATION OF HIGH SOLIDS LIGNOCELLULOSIC BIOMASS CONVERSION FOR ETHANOL PRODUCTION

Dilute acid pretreatment followed by enzymatic hydrolysis of cellulose and sugar fermentation is a promising technology for converting lignocellulosic biomass to fuel ethanol. Two of these steps, enzymatic hydrolysis and fermentation, are ultimately catalyzed by protein. The broad goal of this study is to develop a more definite understanding of the physical barriers to performing these two process steps economically. The major foci of this work are high-solids enzymatic saccharification of pretreated lignocellulosic biomass (corn stover) and protein expression profiling of a metabolically engineered bacterium fermenting glucose and xylose, the major biomass sugars.

While high-solids enzymatic hydrolysis of cellulose is advantageous for reducing capital and operating costs, operating an enzymatic saccharification reactor at high insoluble solids levels presents a unique set of physical and reactor-dependent challenges such as mass transfer, temperature control, mixing, pH control, and sugar inhibition. This work partially focused on characterizing the effects of these problems. It was determined in saccharification characterization studies that slurries of pretreated corn stover (PCS) containing between 10% and 15% insoluble solids by weight represent the approximate upper limit for enzymatically hydrolyzing cellulose in a stirred tank reactor (2-10 L scale

fermenter). As this work demonstrates, this maximum limit is ultimately derived from mixing limitations that cause difficulties with temperature control, as well with uniformly distributing enzymes. Using an offline optimal control algorithm in conjunction with a process model, a bench-scale stirred tank reactor was operated at 15% insoluble solids in fed-batch mode, while demonstrating sugar concentrations and yields equivalent to what would be found if operating at 25% initial insoluble solids. Further work applied 2-D gel electrophoresis and hierarchical cluster analysis to track transient protein abundance levels in during the course of mixed substrate fermentation of the two principle sugars contained in lignocellulosic biomass. Of particular significance the saccharification work was able to demonstrate for the first time that high glucose concentrations (>140 g/L or 80% cellulose conversion) are achievable in high-solids enzymatic saccharification reaction systems, and that high-solids enzymatic saccharification presents a viable reaction step option in lignocellulosic biomass conversion technologies.

David Hodge
Department of Chemical Engineering
Colorado State University
Fort Collins, Colorado 80523
Summer 2005

ACKNOWLEDGEMENTS

I am grateful to everyone who has assisted, encouraged, and taught me during the past several years. In particular, I would like to thank my advisor and committee members for taking the time and responsibility for advancing my educational goals. I would also like to thank Dr. Jim McMillan for arranging for me to complete my research in a world-class facility such as NREL. I would also like to thank many of the students, scientists, and engineers at CSU and NREL who helped me, particularly Dr. Ali Mohagheghi for his excellent guidance and training in laboratory techniques for enzymatic hydrolysis and SSF. The DOE Office of the Biomass Program provided the funding and financial support for the enzymatic hydrolysis work, for which I am very grateful.

TABLE OF CONTENTS

APPENDICES

APPENDIX 1: MATLAB SCRIPTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

OVERVIEW OF PROTEIN CATALYSTS IN THE BIOLOGICAL CONVERSION OF LIGNOCELLULOSIC BIOMASS

1.1 Methodological Approach of Dissertation

The biomass conversion process consisting of dilute acid pretreatment, enzymatic hydrolysis, and mixed sugar fermentation is a promising pathway for converting lignocellulosic biomass to fuel ethanol. The US Department of Energy's (DOE) Biomass Program Multiyear Technical Plan identifies a number of technical and economic barriers to the commercialization of renewable lignocellulosic biomass energy (DOE, 2003). For enzymatic hydrolysis, the report specifically points out product yield, product concentration, conversion rate, product quality, and capital investment as areas where technology improvements are needed to improve process economics. High-solids enzymatic hydrolysis has the potential to address these issues by increasing product concentration and decreasing capital costs. Scale-up of a high-solids process is also identified in the report as representing a large knowledge gap and a source of risk for process integration. In addition to cellulose hydrolysis, robust fermentation of all biomass sugars to ethanol presents a metabolic engineering challenge and is another potential technical barrier that is identified in the DOE report.

Two of the reaction steps for the overall process require biological catalysts: cellulase enzymes for enzymatic hydrolysis and a fermentative microorganism for fermentation.

Ultimately, enzymes also catalyze the chemical reactions within the cell during fermentation. For this reason, the study of protein, the physical and physiological factors affecting protein during these reactions, and the barriers to improving process performance must be addressed. The goal of this work is to specifically identify these problems and through detailed analysis demonstrate that the potential bottlenecks can be removed. The experimental course this work will follow is to characterize through experiments, mathematical modeling, and analysis the factors affecting high-solids enzymatic saccharification and scale-up of this reaction to stirred tank reactors (STRs). On the microbial fermentation element of the process, protein expression profiling will be used to determine changes in the metabolic state of the metabolically engineered microorganism during mixing sugar fermentation. This work strives to demonstrate practical implementation concerns along with solid mathematical analyses and models to supplement the laboratory analytical methods, such that an "engineering" perspective can be gained on these problems rather than a purely mechanistic scientific approach.

The purpose of this overview chapter is to develop the necessary background, present the motivating concerns behind this work, and propose the experimental course required to address these concerns. The subsequent chapters each deal with specific components of this research area, and are organized as standalone units of research, with the goal of publication in academic journals. Previously unresolved issues to be specifically addressed by this work include the characterization of high-solids saccharification, identifying the factors affecting scale-up of saccharification for high-solids slurries, and determining metabolic bottlenecks in mixed substrate fermentation. **CHAPTERS 2** and **3** deal with the characterization of high-solids saccharification. In **CHAPTER 2**,

detailed mass balances are developed and analytical methods are compared. For **CHAPTER 3**, a number of substrate and enzyme-related properties are examined and their effects on high-solids saccharification elucidated. **CHAPTERS 4** and **5** identify and expound on the factors affecting the scale-up in stirred tank reactors (STRs) and demonstrate the suitability of scale-up strategies based on this understanding. **CHAPTER 6** examines the proteins involved in the biological conversion of lignocellulosic sugars to ethanol during fermentation, and applies mathematical tools to demonstrate the coordinated expression of related proteins. Appendices include sections for MATLAB scripts (**APPENDIX 1**), *Z. mobilis* genes (**APPENDIX 2**), saccharification experimental designs (**APPENDIX 3**), saccharification experimental data (**APPENDIX 4**), all protein spot data (**APPENDIX 5**) and a list of all abbreviations and symbols (**APPENDIX 6**).

1.2 Biofuels

1.2.1 Biomass Conversion to Energy

The primary source of energy in the world today is fossil fuels: coal, oil, and natural gas. The demand for energy is continually increasing and much of the growth to quench this demand has come from fossil fuels. Illustrating this point, the US has increased oil consumption by 20% in only the last decade (Varchaver, 2004), while global use of fossil fuels has increased by 21-fold in the 20th century and 800-fold since 1750 (Hall *et al.*, 2003). Fossil fuels are nonrenewable and contribute considerably to the increase in atmospheric CO_2, a greenhouse gas that has been linked to global warming (Dickens, 2004). As worldwide and domestic energy demands increase, fossil fuel reserves remain finite, such that it is imperative that more renewable and environmentally responsible

alternatives to fossil fuels be found within the foreseeable future in order to maintain long-term energy sustainability.

Plant (lignocellulosic) biomass as a source of energy is advantageous in that it is renewable, abundant, and is environmentally benign with respect to greenhouse gas emissions. This energy source is ultimately derived from the photosynthetic ability of plants to capture and store solar energy in the form of the chemical bonds within the plant tissues. During growth, plants sequester CO_2 and convert this into more reduced compounds contained in the biomass as part of the biosphere's carbon cycle. In an oxidizing environment such as in the Earth's atmosphere, these reduced compounds represent a large reservoir of potential energy (Hall *et al.*, 2003). Furthermore, the combustion of these plant tissues for energy leads to no net increase in the amount of atmospheric CO_2, whereas fossil fuels release carbon that was sequestered in ancient times (primarily the early Mesozoic and Paleozoic geologic eras). Other benefits of lignocellulosic biomass as an energy source include generating rural income by providing a market for agricultural wastes and improving domestic energy security, which would create a buffer from disturbances in world petroleum markets considering that the US currently imports more than 60% of its oil (Vercharver, 2004).

Lignocellulosic biomass consists of three major fractions: cellulose, hemicellulose, and lignin. Lignin is an irregular polymer based on repeated units of phenyl propane. Hemicellulose is an irregular branched heteropolymer of pentose sugars, primarily xylose, with modest amounts of arabinose and some hexoses. The largest fraction is cellulose, the most common polymer on earth with estimates of 10^{11} tons formed in the biosphere annually (Bailey and Olis, 1986). Only the cellulose and hemicellulose

fractions can be realistically refined to produce fermentable sugars. As depicted in Figure 1.1, cellulose is composed of repeating cellobiose units in a β -1,4 glycosidic bond, and is directional, containing reducing and non-reducing ends. Due to the repeating homogeneous structure of cellulose, hydrogen bonding between chains can result in a strong crystalline structure, which is difficult for potential hydrolyzing agents to penetrate. Compared to starch-based feedstocks containing only glucose sugars, lignocellulosics contain many sugars including glucose in cellulose and a variety of five and six carbon sugars in hemicellulose with a large fraction of xylose in hardwoods and herbaceous crops. These additional sugars further complicate coordinated microbial utilization of cellulosic feedstocks.

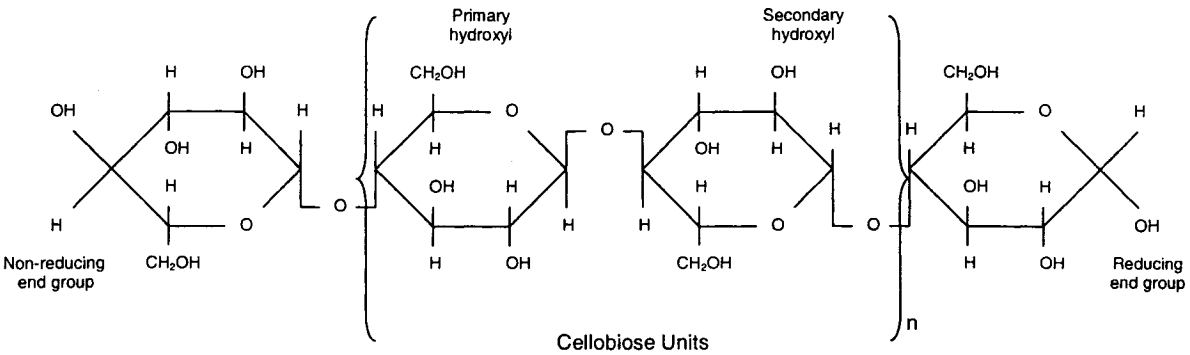


Figure 1.1: Repeating units of glucan in cellulose (Smook, 1992)

Industrial use of lignocellulosic biomass as a fuel source is possible using a number of different technology "platforms". Currently, the most promising technology platforms include the thermochemical conversion platform and the sugar platform (http://www.eere.energy.gov/biomass/). Thermochemical conversion technologies are based on either methods for the direct combustion of biomass such as co-firing with coal, or the conversion to intermediate combustion fuels through pyrolysis or gasification,

which can result in higher efficiency steam generation. The sugar platform consists of breaking down the cellulosic and hemicellulosic fractions of the biomass into individual component sugars that can be utilized for microbial fermentation to ethanol, or alternatively as a precursor for the production of other commodity chemicals. In the paper industry, the industrially unusable fractions (bark and black liquor) are combusted for energy. Similarly for the sugar platform approach, water-insoluble lignin can be separated and burned as a fuel. Sugar platform pathways for converting biomass to its constituent sugars include concentrated acid hydrolysis and any number of "pretreatment" technologies (Sun and Cheng, 2002; Sheehan, 2001). The purpose of these pretreatments is to open up the tight crystalline structure of cellulose, hydrolyze some cellulose and much of the hemicellulose, and solubilization or redistribution of lignin (Lynd *et al.*, 2002). Pretreatment is next followed the hydrolysis of cellulosic polymers to glucose by cellulase enzymes in an overall process that can be broadly categorized as the "enzyme sugar platform" (ESP).

Ethanol can be produced by fermentation of either monomeric glucose only with common industrial fermenters such as *Saccharomyces cerevisiae*, or by utilizing a wider spectrum of monomeric sugars using a microorganism metabolically engineered to ferment these. Ethanol can be used as a transportation fuel as an octane booster and oxygenate in blends up 10% with gasoline or at higher blends with minor modifications to standard gasoline combustion engines. However, ethanol has lower air quality emissions reduction capabilities relative to additives such as methyl tertiary butyl ether (MTBE) (http://www.eia.doe.gov). Because the transportation sector is responsible for nearly 50% of the total CO_2 emissions (Mielenz, 2001), replacement or supplementation

of fossil-based liquid transportation fuels with renewable fuels has the potential to significantly reduce greenhouse gas emissions. The U.S. ethanol industry output has seen amazing growth in the past six years, more than doubling its total annual capacity to 3.5 billion gallons between 1998 to 2004 (Urbanchuk, 2004). The expansion is partially driven by recent regulations requiring the phase-out of the biologically recalcitrant oxygenating additive MTBE coupled with high oil prices making fuel ethanol more competitive. Much of this growth has also been prompted by lower corn prices and increasing oil prices that make corn ethanol competitive with gasoline. This growth is nearing saturation, since currently nearly 10% of domestic corn is used for producing ethanol (Urbanchuk, 2004). Because of this, many argue that future growth for ethanol as a renewable fuel must come from production from nonfood sources such as lignocellulosic biomass.

As an alternative to ethanol, lignocellulosic sugars can be converted to commodity chemicals. Recent industrial examples of biomass-derived polymers include a Cargill-Dow process using corn-derived glucose that produces lactic acid as a precursor for polylactic acid and a DuPont process that produces a plastic from corn-derived 1,3 propanediol (Ritter, 2004). Because of the difficulty in economically competing with mature nonrenewable energy technologies, a multiple-product or "biorefinery" approach is seen to have better potential to provide a more competitive plant (DOE, 2003).

Table 1.1 provides a summary of feedstocks capable of biological conversion to ethanol and the expected ethanol yield based on 100% sugar yield and complete sugar usage. One important observation from this table is the very high conversion yield potential in lignocellulosics which are on par with dedicated starch crops such as corn

and wheat, without requiring the displacement of starch as a food source. An additional advantage to lignocellulosics is the potential for energy cogeneration through the combustion of the lignin-associated byproducts.

Table 1.1 : Ethanol from Biomass Yields

Most values developed from data of Wayman and Parekh, 1990

	Biomass source of substrate	Ethanol yield (g EtOH/g dry biomass)
Sugar crops:	Sugar cane†	0.0583
	Sorghum†	0.0621
	Sugar beet†	0.0738
Starch crops:	Barley grain†	0.3511
	Wheat grain†	0.3884
	Corn grain†	0.3496
Lignocellulosics:	Corn stover‡	0.3496
	Wheat straw‡	0.3806
	Wheat straw† (Ritter, 2004)	0.2799
	Bagasse‡	0.3066
	Hardwood (Aspen)‡	0.3485
	Softwood (Spruce)‡	0.3398

† Based on existing industrial processes
‡ Maximum theoretical yield based on complete utilization of sugars in cellulose and hemicellulose fractions

Corn stover (stalks, leaves, and cobs) represents a large reservoir of potential energy from biomass. The US is the largest single producer of corn in the world, and in 2001 produced ~216 million dry tons of corn stover (Sokhansanj, 2002). Currently, only 6% is harvested for use as animal feed and bedding. In April of 2004, Iogen, in collaboration with Petro-Canada and Shell started production in a cellulosic ethanol plant in Ottawa, Ontario. This plant is capable of processing 3000 tons annually of wheat straw using cellulase enzymes made in an adjacent enzyme manufacturing facility, and the company plans to complete a >150-fold scaled up plant by 2007 (Ritter, 2004). The Iogen wheat

straw conversion yield entry in Table 1.1 is lower than the other wheat straw entry because only glucose sugars from the biomass are utilized.

1.2.2 Challenges, Bottlenecks, and Economic Considerations

Industrial conversion of lignocellulosics to ethanol via the ESP approach involves the general process steps of biomass milling/grinding, pretreatment, enzymatic hydrolysis of cellulose, fermentation, and separation. There are key economic constraints at each stage of the process that dictate the choice of operating conditions. The milling/grinding process involves mechanically decreasing the average particle size of the biomass, increasing surface area available for both pretreatment and enzymatic hydrolysis. Energy costs associated with this process step determine the extent to which this is performed (Wayman and Parekh, 1990).

Pretreatment is the process step in which the biomass is cooked under increased temperature and pressure in the presence of either steam or steam and a mildly hydrolyzing agent such as dilute sulfuric acid. The purpose of this is to solubilize sugars contained in the hemicellulose fraction, redistribute the lignin, and solubilize (depolymerize) a small portion of the lignin and cellulose. As the pressure is dropped at the end of the process, the water-impregnated cellulose swells, exposing more of the cellulose for subsequent enzymatic hydrolysis. Other byproducts such as acetic acid that are inhibitory to microbial growth are also formed is most pretreatment methods (Gutiérrez *et al.*, 2002; Zaldivar *et al.*, 1999; Palmqvist *et al.*, 1999). In addition to depolymerizing hemicellulose and cellulose, acid hydrolysis in pretreatment also degrades hexose and pentose sugars, resulting in HMF and furfural, respectively. Acetic acid is present as acetylated side-chains of pentoses in hemicellulose. The level of acetic

acid residues in the hemicellulose depends on the biomass species, although hardwoods tend to have higher levels (Bierman, 1996). Critical factors (Schell *et al.*, 2003) in this step include temperature, pH and residence time (combined severity of pretreatment) since inhibitory degradation products such as furfural and hydroxymethylfurfural (HMF) are formed from prolonged exposure of hemicellulose-derived monomers to the hydrolyzing conditions of high temperature and acid concontrations.

One of the major drawbacks associated with lignocellulosics is the difficulty in the recovery of the component sugars relative to starch crops, both of which are enzymatically hydrolyzed to sugar monomers. Currently, cellulase enzymes have much lower specific activities (~2 orders of magnitude) and are much more expensive than amylases required in the hydrolysis of starches (Lynd *et al.*, 2002). Another challenge facing cellulosic ethanol is transportation of the raw material, which requires that conversion process plants be located within a relatively close radius of where the feedstock is available (Aden *et al.*, 2002).

Downstream processing steps such as separations make certain demands from other unit operations within the process. Final ethanol concentration, set by either the sugar concentration and microorganism ethanol tolerance, directly affects the overall economics. For ethanol-water distillation, steam usage is a major energy cost, and lower feed ethanol concentrations will increase the steam usage per pound of ethanol recovered. If a higher heating value is required from the steam usage than is available as recovered ethanol, the process will have no value as a renewable energy technology. Since ethanol concentration can potentially define the distillation operating costs, a microorganism's tolerance to ethanol is also of paramount importance when considering the integration of

a fermentation process into the overall lignocellulosic conversion process. For this reason, high sugar yields from pretreatment and saccharification combined with fermentation capable of high ethanol concentrations and yields are desirable process traits.

One option for improving the economics of enzymatic hydrolysis is to operate a saccharification reactor at higher levels of insoluble solids. Potential advantages to this include increasing the product sugar concentrations, which decreases separation costs, while less water in the slurry reduces both capital and operating costs, since smaller reactors are needed for the same product throughput and less energy is required for heating and cooling this smaller volume. Drawbacks and challenges to using higher insoluble solids during saccharification are related to either the reactor or the reaction slurry. As insoluble solids are increased, the slurry becomes much more viscous and increasingly non-Newtonian. The reactor limitations involve mixing and temperature control, both of which become increasingly difficult at higher insoluble solids levels. The reaction limitations involve problems with sugar inhibition and problems with mass transfer, which limits both diffusion of product sugars away from the catalytic site, and diffusion of enzymes to new catalytic sites. All of these issues will be specifically addressed in this work.

1.2.3 Comparison of Integrated Process Options

Many process options exist for the conversion of pretreated lignocellulosic solids and liquors to ethanol or other value-added products using the ESP approach. One major factor in deciding process options is the level of integration between enzymatic hydrolysis and fermentation. These process options include separate (or sequential)

hydrolysis and fermentation (SHF) and coupled schemes such as simultaneous saccharification and (co-)fermentation (SSF or SSCF) and hybrid hydrolysis and fermentation (HHF).

SHF enables both fermentation and enzymatic hydrolysis to be carried out at the respective optimal conditions (T, pH, *etc.*), which has the potential to yield a large benefit if significantly different optima exist. Another advantage to SHF is that this process is capable of producing intermediate sugar products, which can be separated and sold or used to produce co-products. This approach can be used to facilitate a "sugar platform" multi-product biorefinery. Improved mixing during fermentation is another benefit of SHF, since there are lower insoluble solids levels (or a total absence of insoluble solids if a solid-liquid separation is used prior to fermentation) than in fermentation schemes coupled with enzymatic hydrolysis. Disadvantages include inhibition of enzymes due to accumulation of hydrolysis sugars, and higher capital and operating costs due to the additional reaction vessels required for a separate fermentation. SHF could prove an attractive process option for situations where sugar inhibition of enzymatic hydrolysis is not a problem.

Operating a process utilizing SSF or SSCF reduces the reaction equipment complexity and cost by operating both the enzymatic hydrolysis and fermentation steps in a single reactor vessel. There are two additional process-related advantages to the removal of sugar by a microorganism as it is generated by enzymatic hydrolysis. The first advantage is that sugar removal should result in faster hydrolysis rates for enzyme systems and reactor systems where high concentrations of glucose would accumulate and inhibit hydrolysis. The second is that the removal of sugars generates an environment

less amenable to contamination by a foreign microorganism. SSF has the disadvantage of longer residence times due to the slower enzyme reaction kinetics within mesophilic microorganism temperature tolerances, which would be one motivation for HHF. Techno-economic modeling by Wingren *et al.* (2003) found SSF to be economically more favorable than SHF. However, this study assumed the same operating temperatures for the saccharification reactions for both cases and ignores the advantage of operating the saccharification step at a higher temperature since the reaction rate is temperature-dependent in an Arrhenius manner (Kadam *et al.*, 2004; Calsavara *et al.*, 1999).

A variation of SSF is HHF, the goal of which is to run these reactions at their respective temperature and pH optima within a single reaction vessel thereby increasing reaction rates by reducing overall residence times. This is done by first saccharifying the pretreated solids, and at a specified time lowering the temperature to either the fermentation optimum or a higher temperature to allow an moderately accelerated rate of enzymatic hydrolysis to continue, and then inoculating with the fermenting microorganism to commence SSF. Preliminary experimental work at NREL using *S. cerevisiae* has failed to demonstrate any of the anticipated improvement in final cellulose conversion to ethanol using HHF with differing yeast inoculation times relative to conventional SSF.

The fermentative microorganism also determines process conditions such as temperature, pH, and co-fermentation possibilities. Because of this, it is important to identify an enzyme-microorganism system that performs well under these constraints. As a result of advances in metabolic engineering, the fermentation of pentose and hexose sugars to ethanol, which has been one major barrier in the past, can potentially be

performed in a single co-fermentation process step (Aristidou and Penttilä, 2000). The combination of fermentation steps and the removal of a separation step will reduce capital and operating costs for an industrial process. Microorganisms that have had been metabolically engineered to increase the fermentable substrate spectrum to include hexoses and pentose (xylose) sugars cover many strains of enteric bacteria (Ingram *et al.*, 1999) as well as yeast, *S. cerevisiae* (Ho *et al.*, 1998), and *Zymomonas mobilis* (Zhang *et al.*, 1995; Mohagheghi *et al.*, 2002).

The effect of temperature on the fermentative microorganism is an important parameter for SSF consideration due to the strong temperature-dependence of the saccharification reaction. *S. cerevisiae* strains can ferment glucose to ethanol with robust growth at temperatures slightly greater than 37°C (Bollók *et al.*, 2000), while metabolically engineered *Z. mobilis* strains can tolerate temperatures only as high as 35°C (McMillan *et al.*, 1999) with more robust growth at 32°C to ferment mixtures of glucose and xylose. These lower temperatures have important implications for the rate of cellulose hydrolysis, because there are tradeoffs in the rate of conversion with respect to either temperature or sugar inhibition of enzymes. *Kluyveromyces marxianus* is another yeast that ferments glucose to ethanol at high product yields (Bollók *et al.*, 2000), and is more thermotolerant (up to 50°C, robust growth at 42°C) than more commonly used ethanologens. This is noteworthy since these temperatures are closer to the *T. reesei* cellulase optimum temperature. Because of the temperature constraints of most ethanologens, the development of a thermophilic ethanologen would be desirable for rate improvements in SSF or SSCF.

1.3 Protein-Catalyzed Biomass Conversion

Proteins are the catalytic workhorses of the cell, mediate virtually all biologically significant cellular chemical reactions, and are the final expressed complement of the data carried in the genome of an organism. The diverse array of functions performed by proteins is considerable and includes most of the catalytic, transport, structural, and regulatory duties of the cell. Figure 1.2 provides an overview of the diverse range of functions catalyzed by protein for an organism performing a mixed sugar fermentation to ethanol. The work of Tatusov *et al.* (1997) grouped proteins from five microorganisms into generalized categories (defined as clusters of orthologous groups or COGs) based on sequence homology, and which were further categorized at a higher level based on function. In addition to proteins of unknown or suspected function, proteins were classified according the following 13 functional categories:

1. Translation, ribosomal structure, and biogenesis
2. Transcription
3. Replication, repair, and recombination
4. Molecular chaperones
5. Outer membrane, cell wall biogenesis
6. Secretion
7. Inorganic ion transport
8. Energy production and conversion
9. Carbohydrate metabolism and transport
10. Amino acid metabolism and transport
11. Nucleotide metabolism and tranport
12. Coenzyme metabolism
13. Lipid metabolism

For the enzyme sugar approach to lignocellulosic biomass conversion technology, two unit operations require protein-catalyzed reactions, both of which fall in the general category of "carbohydrate metabolism and transport". The enzymatic hydrolysis step uses cell-free cellulase proteins recovered from fungal cultures, while fermentation requires living cells and the cellular machinery involved in the anaerobic metabolism of

pentose and hexose sugars to ethanol. In addition to carbohydrate metabolism, all other physiological functions within the cell are being performed concurrently, which can complicate the understanding of sugar metabolism alone. This requires an understanding of the issues associated with cellular growth in addition to metabolism. Knowledge of the mechanisms involved in protein structural organization, conformational stability, and catalytic function is also essential for the understanding and application of either the characterization of protein catalysis or for protein separations.

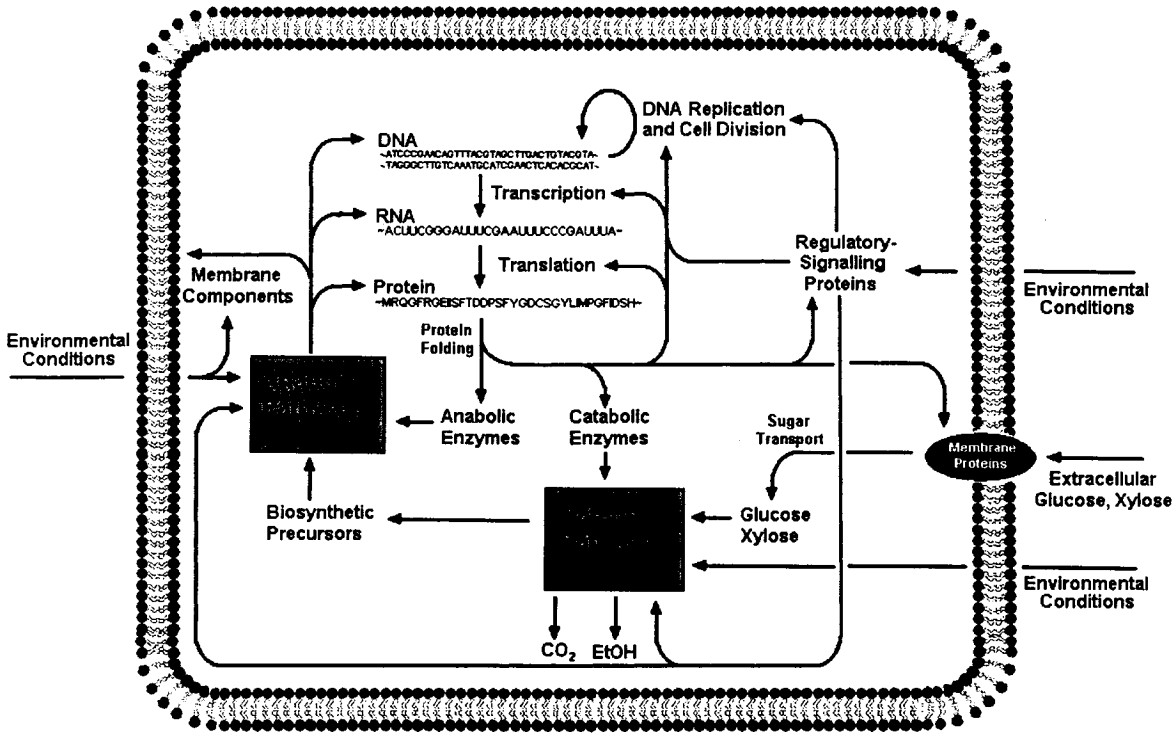


Figure 1.2: The general roles of protein function in cellular physiology

1.3.1 Protein Background

Proteins are heteropolymer chains composed of up to 20 different amino acid residues covalently linked in peptide bonds. Contained on amino acids are both acid and basic groups such that it can exist primarily in the form of a dipolar ion (zwitterions)

(McMurry, 1992). Zwitterions act as salts and have many properties associated with salts such as large dipole moments, which is a significant property that can be exploited for separations.

Protein structure can be considered at three basic levels of organization: primary, secondary, and tertiary. The primary structure consists of the amino acid sequence of a protein linked in peptide bonds. Structural motifs and binding sites that are apparent at the higher levels of organization can often be inferred from knowing the primary protein structure. Frequently encountered short peptide sequences that form structural elements such as α-helices and β-barrels within the larger 3-dimensional organization of the protein are considered components of the secondary structure. The tertiary structure is essentially the complete 3-dimensional structure of a peptide sequence. It contains important conformational information such as the special organization of the catalytic and binding domains, as well as insight into important structural features such as internal packing. Higher order structural elements are present in multiple peptide chain proteins and are involved in protein-protein interactions.

The structural stability of proteins can be examined in terms of the chemical bonds and interactions that make up the internal packing and folding patterns in the tertiary structure of the protein. These interactions between the amino acid residues include covalent bonds, hydrogen bonds, ionic bonds, Van der Waals interactions, and hydrophobic interactions. Covalent bonds, involving two amino acid residues sharing an electron pair, are the strongest of molecular interactions, and are present in both disulfide bridges and peptide bonds. All noncovalent bonds are considered weak interactions due

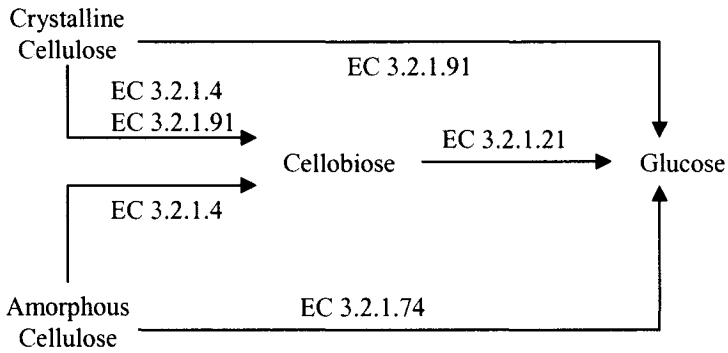
to the comparable difference in bond strength. These weak interactions are considered the driving force behind protein stability, and act with a high degree of cooperativity.

The surface energy of a protein is the energy that results from interactions between the solvent and the protein. The protein disrupts the hydrogen-bonding pattern of water, and results in the water molecules forming an oriented shell that results in an overall unfavorable increase in the free energy. Weak interactions within the protein cannot compensate for the loss of the original hydrogen-bonding pattern of water so that a net free energy increase at the protein-solvent interface will result. Most of the net free energy changes within the protein resulting from weak interactions are derived from the entropy increase in the surrounding aqueous phase (Lehninger *et al.*, 1993). These interactions are important for the unfolding and solubilization of proteins with detergents for SDS-PAGE analysis.

Proteins, nucleic acids, and many other organic compounds exhibit mobility under the application of an electric field due to particle surface charges. Electrophoresis has developed as a tool to separate mixtures of these compounds based on differences in size and isoelectric point. Two-dimensional gel electrophoresis (2-DE) has become one of the core technologies behind proteomics, and is capable of resolving thousands of proteins. The premise behind 2-DE is that proteins can be separated by differences in isoelectric point and molecular weight. Extensive preparation measures are necessary to separate proteins from modifying and degrading compounds, as well as to ensure proper solubilization and protein unfolding. These methods and the principles behind these are further discussed in **CHAPTER 6**.

1.3.2 Enzymatic Hydrolysis of Cellulose

In nature, cellulases are found as either complexed or noncomplexed enzyme systems. Complexed cellulase systems are bound together on a scaffold protein attached to the cell membrane that is collectively known as a cellulosome. These systems are typically found in anaerobic microorganisms where ATP conservation may be more of a priority (Lynd *et al.*, 2002). Noncom-plexed cellulase systems are found primarily in aerobic fungi such as *Trichoderma reesei.* These enzymes are secreted to function outside of the cell and represent the most widely studied class of cellulase enzymes.



EC 3.2.1.74: exo-β-1,4-glucanhydrolase
EC 3.2.1.4: endo-β-1,4-glucanase
EC 3.2.1.91: exo-β-1,4-cellobiohydrolase
EC 3.2.1.21: β-glucosidase

Figure 1.3: Enzymatic pathways of cellulose hydrolysis, after Kadam *et al.* (2002)

Cellulases contain a variety of enzymes that work cooperatively during the depolymerization of cellulose, and belong to the three general classes of endoglucanses, exoglucanases, and β-glucosidases. Endoglucanases cleave random sites within the chain, creating more free ends and oligosaccharides of a variety of lengths. Exoglucanases act processively on both reducing and nonreducing ends of the cellulose chain and can release either glucose or cellobiose as products. Cellobiose and soluble cellodextrans can be depolymerized to glucose by β-glucosidase (Lynd *et al.*, 2002). A

generalized overview of these pathways for cellulose hydrolysis to monomeric glucose along with the enzyme catalysts for each step are given in Figure 1.3.

Many cellulase enzymes are "modular" in structure, containing both a substrate binding domain and a catalytic domain. Much work has been published for cellulase proteins characterizing the 3-dimensional structure and delineating the functional implications of structure (Lynd *et al.*, 2002).

Table 1.2: *T. reesei* Cellulase Nomenclature

Functionally defined nomenclature		**Structurally defined nomenclature**
Common Nomenclature	IUB-MB Enzyme Nomenclature	Based on glycosyl hydrolase family (PFAM nomenclature)
CBH I	EC 3.2.1.91	Cel7A
CBH II	EC 3.2.1.91	Cel6A
EG I	EC 3.2.1.4	Cel7B
EG II	EC 3.2.1.4	Cel5A
EG III	EC 3.2.1.4	Cel12A
EG IV	EC 3.2.1.4	Cel61A
EG V	EC 3.2.1.4	Cel45A
BGL I	EC 3.2.1.21	Cel3A
BGL II	EC 3.2.1.21	Cel1A

A number of nomenclatures have been used for cellulase enzyme systems, with a comparison between some of the common *T. reesei* cellulase components between naming conventions given in Table 1.2. This table was assembled from various sources, and some proteins may be either missing or duplicated since there isn't always perfect compatibility between the naming conventions. The International Union of Biochemistry and Molecular Biology (IUB-MB) provides guidelines for the standardized naming of enzymes based on the reaction catalyzed (Webb, 1992), and is organized at four levels of classification. The structural characteristics of cellulase enzymes can also be used as a

classification criterion, and is used in the most recently developed naming convention (Henrissat *et al.,* 1998). This is based on the classification of enzyme catalytic domains into families derived from the homology of the protein amino acid sequence. This naming convention, analogously to the IUB-MB system, allows protein similarities to be acknowledged between various glycosyl hydrolases regardless of the microorganism source. This naming convention, however, allows more subtle differences in the catalytic domains to be acknowledged. Current information is maintained in the CAZy database (http://afmb.cnrs-mrs.fr/CAZY/) that describes the families of structurally-related enzymes that degrade, modify, or create glycosidic bonds.

1.3.3 Metabolic Pathways for Sugar Fermentation to Ethanol

The enzymes within the metabolic pathways of ethanologens are the ultimate catalysts for the biological conversion of sugars to ethanol during fermentation. There are numerous industrially relevant metabolic pathways, with virtually all of these centered on glucose catabolism. The two primary pathways associated with glucose catabolism are: glycolysis (also known as the Embden-Meyerhof-Parnas or EMP pathway) and the Entner-Douderoff (ED) pathway. The Pentose Phosphate (PP) pathway is used for both anabolism and catabolism of pentose sugars. A summary of how these three pathways are interrelated in the overall metabolic flux from sugar to ethanol is given in Figure 1.4. Aerobic respiration (not shown) shunts pyruvate into the TCA cycle and provides high energy yields (in terms of ATP), whereas fermentation processes are defined as occurring in the absence of terminal electron acceptors (*e.g.* oxygen), and must be internally balanced with respect to redox potential. These fermentative pathways result in much lower energy yields (in terms of ATP) relative to aerobic respiration.

However, from a fuel/energy perspective, this is a positive property since the final oxidized compound (ethanol) represent a relatively high-energy combustion fuel. Ethanol is also advantageous over other biomass components or metabolites with potential for combustion in that this is an easily transportable and stable liquid fuel.

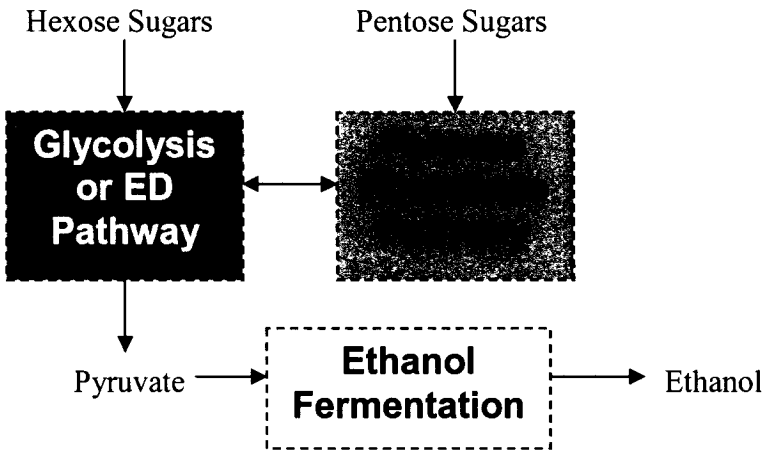


Figure 1.4: Summary of the metabolic fluxes involved in the major pathways for sugar fermentation to ethanol.

The two glucose catabolism pathways are considered to end at pyruvate. The most striking difference between the two glucose catabolic pathways is that the more energetically inefficient ED pathway has a net yield of only 1 mole ATP per mole of glucose, or only 50% of the 2 moles ATP formed using the EMP pathway. The ED pathway is hypothesized to be the more evolutionarily more primitive of the glucose catabolic pathways (Romano and Conway, 1996) and is the major means for glucose catabolism in many pseudomonads as well as *Z. mobilis* (Brock and Madigan, 1991) and branches at 2-keto-3-deoxygluconate-6-phosphate (KDGP) to form 1 mole of glyceraldehyde-3-phosphate and 1 mole of pyruvate. The EMP pathway is found in a taxonomically diverse range of eukaryotes and prokaryotes and uses phosphorylated fructose as the branch point whereby two glyceraldehyde-3-phosphates are created. In

addition to glucose, the metabolism of most hexose sugars can be facilitated through these pathways through prior modification to glucose-6-phosphate or glyceraldehyde-6-phosphate.

The glycolytic flux of the ED pathway of *Z. mobilis* is well characterized, has been reviewed in detail by several authors (De Graaf *et al.*, 1999; Sprenger, 1996; Osman *et al.*, 1987), and is subject to several known types of minimal regulation. The first of these is the differential expression of the GAP operon. The genes encoding glyceraldehyde-3-phosphate dehydrogenase (*gap*) and 3-phosphoglycerate kinase (*pgk*) form a bicistronic operon and constitute the rate-limiting step in glycolysis (Eddy *et al.,* 1989). Glucose-6-phosphate dehydrogenase has been suggested to be the rate-limiting step for glucose fermentation (Snoep *et al.,* 1996). This enzyme has relatively low levels of activity (Viikari, 1988) and is subject to allosteric inhibition by phosphoenolpyruvate (PEP) (Scopes, 1997). Several enzyme activities including fructokinase, phosphofructose isomerase and the glucose transporter (*glf*) increase when cells are grown on fructose instead of glucose (Sprenger, 1996). The stability of mRNA plays a critical role in determining the relative abundance of glycolytic enzymes (Mejia *et al.,* 1992). The expression of genes in the GAP operon and the *glf-zwf-edd-glk* operon have been found to be controlled by differential mRNA processing (Eddy *et al.* 1989, Liu *et al.*, 1992).

Ethanol fermentation from pyruvate to ethanol is considered a distinct pathway (Figure 1.4), since pyruvate can be directed into numerous other aerobic or anaerobic pathways after glycolysis. This ethanol fermentation pathway requires two catalytic steps, whereby pyruvate is first decarboxylated to acetaldehyde and next reduced to ethanol. The first step from pyruvate to acetaldehyde is catalyzed by pyruvate

decarboxylase (PDC), an enzyme rarely found in bacteria (Seo *et al.,* 2005), which requires the cofactors Mg^{2+} and thiamine pyrophosphate (Lehninger, 1992). The isozymes alcohol dehydrogenase I (ADH I) and ADH II further reduce acetaldehyde to ethanol, and require NADH to provide the reducing equivalents. While these two enzymes both catalyze the same reaction, they are unrelated in *Z. mobilis* (Kalnenieks *et al.*, 2002).

For both pentose catabolism and anabolism, the PP pathway is utilized. This pathway requires anabolic capability since ribose is necessary for nucleic acid synthesis. The reactions used in this pathway primarily involve coupled substitutions of ketose and aldose groups. Junctions with glucose catabolic pathways are at glucose-6-phosphate and glyceraldehyde-3-phosphate.

1.3.4 Chaperone Proteins

Heat shock proteins (HSPs) are a highly evolutionarily conserved class of proteins. Major families of HSPs that are found in a wide diversity of organisms include: Hsp100, Hsp90, Hsp70, Hsp60, and the small HSPs (Mager and de Kruijff, 1995), with the mechanisms of the Hsp70s and Hsp60s being the best characterized. Molecular chaperone proteins are a class of HSPs that guide and assist in the folding process for many proteins either following protein synthesis or denaturation.

The higher order structure of proteins is not attained through spontaneous folding; rather, molecular chaperone proteins facilitate this conformational change. DnaK is the prokaryotic homolog of Hsp70 and acts in conjunction with DnaJ and GrpE to fold newly synthesized proteins that would otherwise tend to aggregate due to hydrophobic interactions. This protein ensemble has strong affinities for the hydrophobic domains

exposed on unfolded proteins, and facilitates the posttranslational assembly of growing proteins as they leave the ribosome (Alberts *et al.*, 2002). Another class of chaperone proteins is the Hsp60 family, which includes the protein GroEL in prokaryotes. GroES is a 10 kDa protein located on the same operon as GroEL in *Z. mobilis* (Barbosa *et al.*, 1994) and many other bacteria. In contrast to DnaK, which acts while the protein is still attached to the ribosome, the GroESL complex works later in the protein's life (Alberts *et al.*, 2002). GroEL forms multisubunit rings that combine to form a barrel-like structure (Figure 1.5). This structure surrounds and captures unfolded proteins by binding to exposed hydrophobic sites. The GroES subunit forms a cap for the GroEL barrel (Figure 1.5) and effectively isolates the misfolded proteins. ATP catalyzes the barrel assembly, proper refolding, and release.

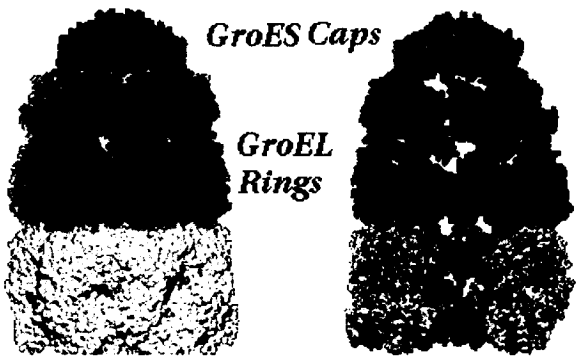


Figure 1.5: Structure of GroESL complex (modified from Walter and Buchner, 2002)

The RNA polymerase (RNAP) in prokaryotes is the enzyme responsible for the DNA transcription to RNA. RNAP consists of the protein subunits α, β, β′, and σ (σ^{70} is the primary σ-factor in *E. coli*, and is named due to its 70-kDa MW), with a two-subunit dimer of α and the 3 remaining subunits comprising the RNAP holoenzyme. RNAP remains active when no σ is present (core enzyme) but with considerable changes to the enzyme's activity. The current model for transcriptional initiation by RNAP (Weaver,

2002) involves the binding of the core enzyme to the σ subunit, whereby the holoenzyme can next bind nonspecifically to the dsDNA. The holoenzyme binds loosely to the dsDNA when a promoter is found forming the closed promoter complex. The holoenzyme now melts the DNA in the region of the promoter forming the open promoter complex, which initiates transcription. After several nucleotides have been added and the promoter is cleared, the core enzyme loses the σ-factor, which can be reused, and elongation begins. It has been found (Wassarman and Storz, 2000) in *E. coli* that 6S RNA binds with β, β′, and σ^{70}, resulting in RNAP holoenzyme activity repression. 6S RNA was found to accumulate during stationary phase in *E. coli* batch cultures, resulting in the repression of σ^{70}-dependent promoters (Wasserman and Storz, 2000).

Prokaryotic promoters can consist of two regions, the core promoter and the upstream promoter (UP) element. The core promoter consists of short sequences centered at –10 and –35 bp upstream of the transcription start site. This core promoter results in the tight binding of the σ-factor to properly align the RNAP at the correct transcription start site. In addition to the core promoter, some strong promoters contain an additional promoter component known as the UP element, which can be bound by the α-subunit to enhance the RNAP-promoter binding (Weaver, 2002).

The crucial control point for the activation of transcription of Hsps in *E. coli* is the replacement of the σ^{70} subunit of the RNA polymerase holoenzyme complex (coded by the *rpoD* gene) with the σ^{32} (*rpoH*) subunit (Cowing *et al.*, 1985). The σ^{32} subunit has a greater specificity for the promoters of the HSPs and results in a marked increase in the transcription of these genes. This type of regulatory approach in prokaryotes, whereby the transcription of a set of genes is induced by the RNAP σ-factor, is used in a number

of circumstances. Examples include the induced response to nitrogen starvation by a σ^{54} subunit (*rpoN*) and the starvation response of nonproliferation in the stationary phase under the control of an σ^{S} (*rpoS*) (Nakahigishi *et al., 1995*) subunit in *E. coli*, the sporulation response to starvation in *B. subtilis* by σ^{E} (*rpoE*) and σ^{C} (Weaver, 2002), and the regulation of flagellar and chemotaxis genes by σ^{F} (*rpoF*) (Kundu *et al.,* 1997).

The σ^{32} subunit is highly unstable under normal cellular conditions, with a half-life of only about 1 minute (Mager and de Kruijff, 1995). This results in very low production of HSPs under normal cellular conditions. Reasons for the high turnover include targeting and degradation by specific ATP-dependent proteases. In addition to the role played in protein folding, the DnaK-DnaJ-GrpE chaperone team modulates the σ^{32} subunit stability through negative regulation by binding to σ^{32} and preventing association with core RNAP. Additional stability is ensured by the binding of σ^{32} to core RNAP (Nakahigashi *et al.*, 2001). Increasing levels of denatured proteins stimulate the activation of σ^{32}, since these proteins are substrates for DnaK and DnaJ, σ^{32} is free to bind to core RNAP and activate HSP transcription. Furthermore, the *rpoH* mRNA has been shown to act as a cellular thermometer whereby the 5′ end melts with increasing temperature and leads to increased translation (Mager and de Kruijff, 1995).

1.3.5 Methods for Analysis of Protein Catalysis

Direct gene expression methods for high throughput technologies such as DNA microarrays and proteomics are suited for the fundamental understanding of the problems associated with metabolic flux, metabolic engineering, and stressed growth. High-affinity epitope tagging can allow very precise quantification of protein expression, and has been used recently to study the global expression of yeast proteins (Ghaemmaghami

et al., 2003). Drawbacks for using this approach, however, are that the complete genomic sequence must be known and the cost may be substantial. On the other hand, rather than using specific experiments and assays to examine particular systems of interest, these approaches take a global perspective by examining all cellular responses at the level of protein expression. Whereas DNA microarrays and epitope tagging experiments require extensive knowledge of the genome as well as large capital investments for their development for novel studies, 2-dimensional gel electrophoresis (2-DE) or enzyme activity assays require neither and can be used to globally examine the expression of less well-characterized genes and gene products. 2-DE has the advantage over enzyme activity assays in that this can detect global and potentially novel physiological responses that would otherwise be overlooked using more directed approaches. This technology has been available since the 1970s although recent advances have made it more practical. Work using a *S. cerevisiae* metabolically engineered to ferment xylose showed major changes in key protein levels for xylose utilization, indicating important changes in metabolic fluxes (Salusjärvi *et al.*, 2003). 2-DE has also been used to characterize protein components of cellulase mixtures (Vinzant *et al.*, 2000).

Besides measurements of protein expression levels, methods for metabolic flux analysis require quantification of metabolites. This can be done either trivially for the substrates and products (Hodge and Karim, 2002) or it can be much more involved, *e.g.* using radio-labeled substrates and quantifying all intermediate metabolite levels and nucleotide phosphates using NMR spectroscopy (De Graaf et al., 1999; Kim et al., 2000). These concentration measurements can ultimately be used to estimate fluxes. Besides

direct analytical methods for the detection and quantification of protein expression, indirect methods are also available. For a cell-free system such as the cellulase enzymes in cellulose saccharification, analysis of the effectiveness of the reaction system is typically performed using tools such as simple sugar measurements (Lynd *et al.*, 2002) and bound/free enzyme quantification (Ooshima *et al.*, 1991).

1.4 References

Aden A, Ruth M, Ibsen K, Jechura J, Neefes K, Sheehan J, Wallace B, Montague L, Slayton A, Lukas J (2002). Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolyiss and Enzymatic Hydrolysis for Corn Stover. NREL Technical Paper 510-32438.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Molecular Biology of the Cell, 4th Ed.* Garland Science, 2002.

Barbosa FS, Yomano LP, Ingram LO (1994) Cloning, sequencing and expression of stress genes from the ethanol-producing bacterium *Zymomonas mobilis*: The groESL operon. *Gene* 148:51-57.

Bierman CJ (1996). *Handbook of Pulping and Papermaking, 2nd Ed.* Academic Press, NewYork.

Bollók M, Réczey K, Zacchi G (2000). Simultaneous Saccharification and Fermentation of Steam-Pretreated Spruce to Ethanol. *Appl Biochem Biotechnol.* 84-86:69-80.

Brock TD and Madigan MT (1991). *Biology of Microorganisms, 6th Ed.* Prentice Hall, Englewood Cliffs, NJ.

Calsavara LP, De Moraes FF, Zanin GM (1999). Modeling cellobiose hydrolysis with integrated kinetic models. *Appl Biochem Biotechnol.* 77-79:789-806.

Cowing DW, Mecsas J, Record MT, Gross CA (1985). Consensus sequence for the *Escherichia coli* heat-shock promoters. *Proc Natl Acad Sci USA.* 82:2679-2683.

De Graaf AA, Spriege, K, Wittig R, Laufer B, Schmitz G, Wiechert W, Sprenger G, Sahm H (1999). Metabolic state of *Zymomonas mobilis* in glucose-, fructose-, and xylose-fed continuous cultures as analyzed by ^{13}C and ^{31}P-NMR spectroscopy. *Arch Microbiol.* 171:371-385.

Dickens GR (2004). Global Change: Hydrocarbon-Driven Warming. *Nature.* 429:513-515.

Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003). Global Analysis of Protein Expression in Yeast. *Nature.* 425:737-741.

Gutiérrez T, Busko ML, Ingram LO, Preston JF (2002). Reduction of Furfural to Furfuryl Alcohol by Ethanologenic Strains of Bacteria and Its Effect on Ethanol Production from Xylose. *Appl Biochem Biotech* 98-100:327-340.

Eddy C, Meija J, Conway T, Ingram L (1989). Differential Expression of *gap* and *pgk* Genes within the *gap* Operon of *Zymomonas mobilis. J Bacteriol.* 171:6549-6554.

Hall C, Tharakan P, Hallock J,Cleveland C, Jefferson M (2003). Hydrocarbons and the evolution of human culture. *Nature* 426, 318 – 322.

Henrissat B, Teeri TT, Warren RAJ (1998). A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS Letters.* 425:352-354.

Ho N, Chen Z, Brainard A (1998) Genetically engineered *Saccharomyces* yeast capable of effective cofermentation of glucose and xylose. *Appl Environ Microbiol* 64:1852-1859.

Hodge D, Karim MN (2002). Modeling and Advanced Control of Recombinant *Zymomonas mobilis* Fed-Batch Fermentation. *Biotech Prog.* 18:572-579.

Ingram LO, Aldrich HC, Borges ZCC, Causey TB, Martinez A, Morales F, Saleh A, Underwood SA, Yomano LP, York SW, Zaldivar J, Zhou S (1999). Enteric Bacterial Catalysts for Fuel Ethanol Production. *Biotechnol Prog* 15: 855-866.

Kadam KL, Rydholm EC, McMillan JD. (2004) Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnol Prog.* 20(3):698-705.

Kalnenieks U, Galinina N, Toma MM, Marjutina U (2002). Ethanol Cycle in an Ethanologenic Bacterium. *FEBS Letters.* 522:6-8

Kim I, Barrow K, Rogers P (2000). Nuclear Magnetic Resonance Studies of Acetic Acid Inhibition of Rec *Zymomonas mobilis* ZM4(pZB5) *Appl Biochem Biotech.* 84-86:357-370.

Lehninger AL, Nelson DL, Cox MM. *Principles of Biochemistry, 2nd Ed.* Worth Publishers, New York, 1994.

Liu J, Barnell WO, Conway T (1992). The Polycistronic Messenger-RNA of *the Zymomonas-mobilis Glf-Zwf-Edd-Glk* Operon is Subject to Complex Transcript Processing. *J Bacteriol.* 174:2824-2833.

Mager WH, de Kruijff AJJ (1995). Stress Induced Transcriptional Activation. *Microbiol Rev* 59(3):506-531.

McMillan JD, Newman MM, Templeton DW, Mohagheghi A (1999). Simultaneous Saccharification and Cofermentation of Dilute-Acid Pretreated Yellow Poplar Hardwood to Ethanol Using Xylose-Fermenting *Zymomonas mobilis*. *Appl Biochem Biotechnol.* 77-79:649-665.

McMurry J. *Organic Chemistry. 3rd Ed.* Brooks/Cole Publishing, Pacific Grove, CA, 1992.

Mejia J, Burnett M, An H, Barnell W, Keshav K, Conway T, Ingram LO (1992). Coordination of Expression of *Zymomonas mobilis* Glycolytic Enzymes: a Simple Hypothesis Based on mRNA Stability. *J Bacteriol.* 174:6438-6443.

Mielenz JR (2001). Ethanol production from biomass: technology and commercialization status. *Curr Opin Microbiol.* 4:324-329.

Miklos GLG, Maleszka R (2001). Protein functions and biological contexts. *Proteomics* 1:169-178.

Mohagheghi A, Evans K, Chou YC, Zhang M (2002). Cofermentation of Glucose, Xylose, and Arabinose by Genomic DNA-Integrated Xylose/Arabinose Fermenting Strain of *Zymomonas mobilis* AX101. *Appl Biochem Biotech* 98-100:885-898.

Nakahigashi K, Yanagi H, Yura T (1995). Isolation and sequence analysis of *rpoH* genes encoding sigma 32 homologs from gram negative bacteria: conserved mRNA and protein segments for heat shock regulation. *Nucleic Acids Res.* 23(21):4383-90.

Nakahigashi K, Yanagi H, Yura T(2001). DnaK chaperone-mediated control of activity of a sigma(32) homolog (RpoH) plays a major role in the heat shock response of *Agrobacterium tumefaciens*. *J Bacteriol.* 183(18):5302-10.

Ooshima H, Kurakak M, Kato J, Harano Y (1991). Enzyme Activity of Cellulase Adsorbed on Cellulose and its Change During Hydrolysis. *Appl Biochem Biotechnol.* 31(3):253-266.

Osman Y, Conway T, Bonetti S, Ingram LO (1987). Glycolytic Flux in *Zymomonas mobilis*: Enzyme and Metabolite Levels during Batch Fermentation. *J Bacteriol.* 169:3726-3736.

Palmqvist E, Almeida JS, Hahn-Hägerdal B (1999). Influence of Furfural on Anaerobic Glycolytic Kinetics of *Saccharomyces cerevisiae* in Batch Culture. *Biotechnol Bioeng.* 62:447-454.

Ritter SK (2004). Biomass or Bust. *C&EN*. 82(22).

Romano AH and Conway T (1996). Evolution of Carbohydrate Metabolic Pathways. Res Microbiol. 147:448-455.

Salusjärvi L, Poutanen M, Pitkänen JP, Koivistoinen H, Aristidou A, Kalkkinen N, Ruohonen L, Penttilä M (2003). Proteome Analysis of Recombinant Xylose-Fermenting *Saccharomyces cerevisiae. Yeast.* 20:295-314.

Schell DJ, Farmer J, Newman M, McMillan JD (2003). Dilute-Sulfuric Acid Pretreatment of Corn Stover in Pilot-Scale Reactor. *Appl Biochem Biotechnol.* 105-108:69-85.

Scopes R (1997). Allosteric control of *Zymomonas mobilis* glucose-6-phosphate dehydrogenase by phosphoenolpyruvate. *Biochem J.* 326:731-735.

Seo JS, Chong H, Park HS, Yoon KO, Jung C, Kim JJ, Hong JH, Kim H, Kim JH, Kil JI, Park CJ, Oh HM, Lee JS, Jin SJ, Um HW, Lee HJ, Oh SJ, Kim JY, Kang HL, Lee SY, Lee KJ, Kang HS (2005). The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat Biotechnol.* 23(1):63-68.

Sheehan J (2001). The Road to Bioethanol: A Strategic Perspective of the U.S Department of Energy's National Ethanol Program. In: *Glycosyl Hydrolases for Biomass Conversion.* (Himmel ME, Baker JO, Saddler JN, Editors). Oxford University Press, Washington D.C.

Smook GA (1992). *Handbook for Pulp and Paper Technologists 2nd Ed.* Angus Wilde Publications, Vancouver, BC.

Snoep JL, Arfman N, Yomano LP, Westerhoff HV, Conway T, Ingram LO (1996). Control of glycolytic flux in *Zymomonas mobilis* by glucose 6- phosphate dehydrogenase activity. *Biotech Bioeng.* 51:190-197.

Sokhansanj S, Turnhollow A, Cushman J, Cundiff J (2002). Engineering aspects of collecting corn stover for bioenergy. *Biomass Bioenergy.* 23:347-355.

Sprenger G (1996). Carbohydrate metabolism in *Zymomonas mobilis*: a catabolic highway with some scenic routes. *FEMS Microbiol Lett.* 145:301-307.

Sun Y, Cheng J (2002). Hydrolysis of Lignocellulosic Materials for Ethanol Production: A Review. *Bioresour Technol.* 83(1):1-11.

Tatusov RL, Koonin EV, Lipman DJ. (1997). A genomic perspective on protein families. *Science.* 278(5338):631-7.

Urbanchuk JM (2004). The Contribution of the Ethanol Industry to the American Economy in 2004. National Corn Growers Association Report.

U.S. Department of Energy (2003). *Biomass Program Multi-Year Technical Plan.*

Varchaver N (2004). How to Kick the Oil Habit. *Fortune.* 150(4).

Viikari L (1988). Carbohydrate Metabolism in *Zymomonas. CRC Crit Rev Biotechnol.* 7:237-261.

Vinzant TB, Adney WS, Decker SR, Baker JO, Himmel ME, Kinter MT, Sherman NE, Fox JW (2000). Fingerprinting *Trichoderma reesei* Hydrolases in a Commercial Cellulase Preparation. *Appl Biochem Biotech.* 91-93:99-107.

Walter S and Buchner J (2002). Molecular chaperones--cellular machines for protein folding. *Angew Chem Int Ed Engl.* 41(7):1098-113.

Wassarman KM, Storz G (2000). 6S RNA Regulates *E. coli* RNA Polymerase Activity. *Cell* 101:613-623.

Wayman M and Parekh SR (1990). Biotechnology of Biomass Conversion: Fuels and Chemicals from Renewable Resources. Prentice Hall, Englewood Cliffs, NJ.

Weaver, R. *Molecular Biology 2nd Ed.* McGraw Hill, Inc., 2002.

Webb EC (1992). *Enzyme Nomenclature.* Academic Press, San Diego, CA.

Zaldivar J, Martínez A, Ingram LO (1999). Effect of Selected Aldehydes on the Growth and Fermentation of Ethanologenic *Escherichia coli. Biotech Bioeng.* 65:25-32.

Zhang M, Eddy C, Deanda K, Finkelstein M, Picataggio S (1995) Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis. Science* 267:240-243.

Zhang YHP and Lynd L (2004). Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. *Biotechnol Bioeng.* 88(7):797-824.

Chapter 2

MASS BALANCES FOR THE ENZYMATIC SACCHARIFICATION OF CELLULOSE

2.1 Overview and Context

This work focuses on obtaining mass balance closures around the enzymatic saccharification of pretreated corn stover (PCS), a lignocellulosic biomass source that has been demonstrated as a promising source of renewable energy and chemicals. The motivation for this work stems from the conflicting results obtained for such an important variable as cellulose conversion using methods that quantify either soluble sugars or insoluble solids phase cellulose (glucan) content. Both methods require the quantification of insoluble solids, and the influence of this measurement using the soluble sugars approach becomes especially pronounced at high levels of insoluble solids. Another issue relevant for process understanding that is assessed is mass balance closure over the course of the saccharification reaction. This work seeks to address these issues by developing analytical methods and mathematical formulations for these mass balance problems. The results were generally successful in achieving mass balance closure, although the standard analytical method for quantifying oligomeric sugars tended to over-predict.

2.2 Background

Mass balancing is necessary in order to relate components within experimental data and kinetic models, and provides a necessary validation of either the accuracy of the analytical methods used to acquire the experimental data or the validity of the assumptions made in the mass balance model. For lignocellulosic biomass conversion technologies, rigorous mass balancing has been performed for cellulase production (Schell *et al.*, 2002; Sáez *et al.*, 2002), pretreatment (Schell *et al.*, 2003; Ehrman and Himmel, 1994), fermentation (Erickson *et al.*, 1978), and simultaneous saccharification and fermentation (SSF) (Ehrman and Himmel, 1994).

Cellulose conversion is commonly used as a measure of the effectiveness of enzymatic hydrolysis of cellulose. There are two general approaches for determining the extent of cellulose conversion. The most common method has been defined by Chung *et al.* (1997) as the "summative approach". This approach relies on measurements of glucose, cellobiose, possibly soluble cellodextrin oligomers, and ethanol in the case of SSF to estimate the cellulose that has been enzymatically hydrolyzed. During SSF, the accuracy of this approach can be limited by the estimation of the yield of ethanol from glucose (Vinzant *et al.*, 1994), since the ethanol yield ($Y_{P/S}$) is dependent on the amount of glucose utilized for cellular maintenance and growth, which is not always constant. The equation for this approach to determine the extent of cellulose conversion (ξ) using closed system saccharification (CSS, defined as batch saccharification) while considering only glucose and cellobiose gives:

$$\xi = \frac{HydrolyzedCellulose}{InitialCellulose} = \frac{\frac{\Delta G}{1.111} + \frac{\Delta CB}{1.056}}{S_{1,0} x_{cell,0}} \qquad (2\text{-}1)$$

The other method is a direct measurement, which relies on quantification of the glucan content of the biomass and the insoluble solids. The equation for this is:

$$\xi = 1 - \frac{C}{C_0} \tag{2-2}$$

or:

$$\xi = 1 - \frac{S_I x_{cell}}{S_{I,0} x_{cell,0}} \tag{2-3}$$

where:

G = liquid phase glucose (g/kg)

CB = liquid phase cellobiose (g/kg)

Δ = change from initial conditions

C = cellulose (g/kg)

S_I = insoluble solids (g solids/kg)

x_{cell} = cellulose fraction in the solid phase (g/g)

$x_{cell,0}$ = initial cell. frac. in the solid phase (g/g)

ξ = extent of cellulose conversion

C_0 = initial cellulose (g/kg)

$S_{I,0}$ = initial insoluble solids (g solids/kg)

The direct measurement approach is only limited by the methodological assumptions and accuracy of the analytical technique used. Lignocellulosic insoluble glucan content is typically assayed using 72% sulfuric acid hydrolysis followed by glucose quantification by HPLC (NREL LAP-002, all standard laboratory protocols can be found at http://www.eere.energy.gov/biomass/analytical_procedures.html).

2.3 Materials and Methods

2.3.1 Enzymatic Saccharification of PCS

Due to the larger quantities of solids needed for subsequent analysis, 500 mL shake flasks with a working volume of 200 g are used. The Innova 4000 orbital incubator shaker (New Brunswick Scientific) is operated at 130 rpm, the temperature is controlled

at 45°C. Corn stover pretreated with dilute sulfuric acid in NREL's Sunds vertical pulp digester at 25% total solids was used (PS020502CS). This corn stover was pretreated at 190°C at 0.48 g acid / g dry biomass with the reactor operated in "flow-through" mode (residence time of ~1 min). The insoluble solids are composed of 52.0% cellulose by dry mass, are washed with greater than 10 times the mass of DI water as solids and centrifuged at 15800 x g for 10 minutes at about 700 mL/centrifuge tube. Washed solids are used to minimize any inhibition effects on enzymes from pretreatment hydrolyzate liquor entrained in the wet, unwashed solids. The final glucose concentration in the wash water is verified to be less than 0.1 g/L by YSI glucose analyzer. The cellulase mixture used for this study was Spezyme CP (Lot # 301-00348-257, Genencor International). Total protein was assayed at 106.6 mg/mL (BCA assay, Pierce Biotechnology, Inc.) and the activity was determined to be 0.27 FPU/mg of protein using the NREL protocol LAP-006. Enzyme loadings are given in units mg protein / g cellulose, with the enzyme loading of 20 mg/g used for this work corresponding to 5.3 FPU / g cellulose. Citrate buffer (1 M, pH 5.0) is filter sterilized. Citrate buffer (1 M) is prepared by adjusting the pH to 4.8 with NaOH, and is used as 5% of the total mass to yield an effective molality of 0.05 mol/kg. Enzyme, DI water, and citrate buffer are filter sterilized before use. Initial insoluble solids are 10% by weight. Shake flasks are autoclaved empty, and solids are not autoclaved to prevent further degradation of hemicellulose prior to saccharification. Because solids are not autoclaved, tetracycline is added at 4 μL/mL from a stock solution of 10 mg/L to prevent potential bacterial contamination. Solids moisture level is determined after washing and used to determine how much solids to add to the reactor.

Table 2.1 provides an overview of the sampling times and analytical methods performed on each sample used in this study.

Table 2.1: Experimental Design

Flask #	Sample time				
	0 h	**8 h**	**24 h**	**48 h**	**168 h**
1	PSS, IS, TS, ISC	PSS, IS, TS	PSS, IS, TS	PSS, IS, TS	PSS, IS, TS
2		PSS, IS, TS, ISC			
3			PSS, IS, TS, ISC		
4				PSS, IS, TS, ISC	
5					PSS, IS, TS, ISC

Where:

PSS: partial soluble sugars (glucose, cellobiose, and xylose by HPLC)

IS: insoluble solids measurement

TS: total solids measurement

ISC: insoluble solids composition (analytical determination of insoluble solids composition)

2.3.2 Mass Balances for Shake Flask Saccharification

The flasks, from which solids composition analysis is to be performed, are not sampled in order to prevent loss of mass. Instead, the entire flask was sacrificed at the appropriate time, while a control flask (Flask #1) was continuously sampled. Whole shake flask samples for solids analysis had a filtered liquid phase sample for HPLC and an insoluble solids sample taken (in duplicate) prior to washing. Whole flask samples (200 mL) were subjected to 95°C in a boiling water bath for 5 minutes to inactivate enzymes. The samples were next washed and centrifuged to remove soluble solids until

the glucose concentration was less than 0.1 g/L in the wash supernatant as measured by YSI analysis. Washed solids were stored in sealed bags and refrigerated until all samples could be collected for analysis. Analysis of insoluble and soluble solids were performed by NREL laboratory personnel using NREL protocols LAP- 2, 3, 4, 10, 17, and 19. The soluble sugars glucose, xylose, arabinose, and cellobiose were also quantified independently using HPLC. For this method, approximately 1.5 ml of sample was centrifuged at 10000 rpm for 5 minutes. The supernatant was decanted and diluted in water 1:5 (sample:total) and syringe filtered into disposable HPLC vials. An Agilent 1100 series HPLC equipped with a differential refractive index detector was used for analysis of sugar composition together with an Aminex HPX-87H organic acid column operating at 55°C with a 0.01 N H_2SO_4 mobile phase at a flow rate of 0.6 mL/min.

Table 2.2: Pretreated Corn Stover (PCS) Solids Standards

Standard #	Insoluble PCS	Glucose
1	15%	0
2	10%	10%
3	15%	5%

2.3.3 Insoluble Solids Determination

Three standards were prepared as depicted in Table 2.2 based on insoluble solids and sugar levels that were commonly encountered throughout the course of this work. For each standard 100 g was prepared using washed PCS solids from batch PS020502CS. Standard 2 was generated in duplicate to determine the variability associated with the preparation of the standards. The method for the determination of total solids follows

protocol LAP-001. For the determination of insoluble solids, LAP-018 was modified such that smaller samples could be accommodated. For this, approximately 4 g of sample is pipetted into a tared 15 mL Falcon disposable centrifuge tube and weighed. The sample is diluted to about 12 mL with DI water. These are centrifuged at 5000 rpm, decanted and fresh DI water is added. This is repeated 5 times or until glucose in the supernatant is less than 0.1 g/L as measured by YSI analysis. This sample is added to a tared aluminum weighing dish, dried at 105°C in a desiccater, and then reweighed once dry to determine the insoluble solids content.

2.4 Results and Discussion

2.4.1 Mass Balance Approaches for Estimating Insoluble Solids

The physical system of high-solids cellulose saccharification is a multiphase, heterogeneous suspension, in which both insoluble and soluble solids are present in a discontinuous aqueous phase. As a matter of convenience, due to the changing heterogeneous nature of the system, all components are tracked in units of g/kg rather than the usual g/L used for a single liquid phase. For modeling or characterizing specific properties (*e.g.* rates) of a high-solids enzymatic saccharification, the presence of two phases can be problematic. Since all available models make no distinction between a high-solids system and a low-solids system, a mass concentration of 100 g/kg of a given component can have the same reactivity or inhibitory effect at 5% solids and 20% solids. However, the effective liquid phase concentration in these two systems will be 95 g/L and 80 g/L respectively (Figure 2.1). Additionally, inhibitory effects of sugars on enzymes act in the liquid phase. Using models which consider only the overall (*i.e.* mass) concentrations, the inhibitory effects of the sugars are under-predicted.

G = 95 g glucose / kg slurry
g = 100 g glucose / L liquid

G = 80 g glucose / kg slurry
g = 100 g glucose / L liquid

Figure 2.1: Conceptual demonstration of the limitations of using component mass concentrations (g/kg) for modeling at high solids levels. G = glucose total mass concentration; g = glucose liquid phase concentration

Since only the liquid fraction is being measured by HPLC, using a total mass basis requires the determination of the insoluble solids. During the process of saccharification, cellulose in the solid phase is being converted to water-soluble cellodextrin oligomers (degree of polymerization, DP, less than 12 according to Zhang and Lynd, 2004), cellobiose, and glucose, resulting in changes to the insoluble solids concentration. The summative approach outlined in the **Background** (Section 2.2) is often used to estimate the cellulose utilization, although commonly without considering the effects of insoluble solids on the liquid phase concentration. The summative approach can be extended to estimate the insoluble solids assuming that the cellulose is the only (or at least primary) insoluble component that is changing. In the case of high solids systems, the insoluble solids fraction should certainly be considered when using the summative approach to prevent g/L vs. g/kg errors as outlined in Figure 2.1.

While the insoluble solids level can be measured, the summative method permits the solids level to be determined from the change in sugar concentrations relative to the initial level (net sugar produced). This change in sugar levels can be related stoichiometrically to the amount of cellulose removed from the solid phase to estimate an insoluble solids level. A conversion factor is necessary due to the addition of one molecule of water across each β-1,4 bond that is hydrolyzed in the case of cellulosic glucan. Xylan in hemicellulose is also considered to be hydrolyzed due to cross-

reactivity of cellulases. The xylan conversion factor considers polymeric xylose to exist as a linear polymer of xylose. While this is clearly not the case, this assumption provides a close approximation. Furthermore, since the sugars (glucose, xylose, cellodextrins) are in solution, the determination of both component mass concentrations and total insoluble solids by the summative approach at any given sample time can not be estimated independently of each other, and requires the simultaneous solution to the following system of linear equations assuming a known liquid phase density:

$$S_I = S_{I,0} - (\Delta G/1.111 + \Delta CB/1.056 + \Delta X/1.136) \quad (2\text{-}4)$$

$$\Delta G = \Delta g \cdot (1 - S_I) / \rho \quad (2\text{-}5)$$

$$\Delta CB = \Delta cb \cdot (1 - S_I) / \rho \quad (2\text{-}6)$$

$$\Delta X = \Delta x \cdot (1 - S_I) / \rho \quad (2\text{-}7)$$

4 equations

4 unknowns (left-hand side)

Solving in terms of the insoluble solids with only measured variables appearing in the right-hand side of Equation 2-4 gives:

$$S_I = \frac{S_{I,0} - \left(\frac{\Delta g}{1.111} + \frac{\Delta cb}{1.056} + \frac{\Delta x}{1.136}\right)\Big/\rho}{1 - \left(\frac{\Delta g}{1.111} + \frac{\Delta cb}{1.056} + \frac{\Delta x}{1.136}\right)\Big/\rho} \quad (2\text{-}8)$$

where the relevant variables are:

Calculated Variables	Measured Variables
G = liquid phase glucose (g/kg slurry)	g = liquid phase glucose (g/L liquid)
CB = liquid phase cellobiose (g/ kg slurry)	cb = liquid phase cellobiose (g/L liquid)
X = liquid phase xylose (g/ kg slurry)	x = liquid phase xylose (g/L liquid)
C = cellulose (g/ kg slurry)	$S_{I,0}$ = initial insoluble solids (g/kg slurry)
S_I = insoluble solids (g/ kg slurry)	ρ = liquid phase density (g/L liquid)

Δ = change from initial conditions

Liquid fraction density is another important consideration that becomes more significant in high-solids saccharification. The liquid fraction contains high concentrations of sugar, and the density cannot be assumed to be the same as water. A mixture of sugars is present in the liquid phase that varies in composition during enzymatic hydrolysis. Due to this, it would be impractical to determine correlations between liquid phase density and sugar concentrations for a wide range of sugar compositions. Since this work primarily uses washed solids, and glucose is the most abundant sugar, glucose is considered as the primary density-affecting sugar. A correlation based on glucose data developed from Weast (1985) relates density (ρ [g liquid / L]) to sugar concentration [g sugar / L]:

$$\rho = 0.377 \cdot (g + cb + x) + 0.998 \tag{2-9}$$

Recalculating the estimated insoluble solids equation using this correlation yields:

$$S_1 = \frac{\left\{ S_{1,0} - \frac{\left(\frac{\Delta g}{1.111} + \frac{\Delta cb}{1.056} + \frac{\Delta x}{1.136} \right)}{0.000377 \cdot (g + cb + x) + 0.998} \right\}}{\left\{ 1 - \frac{\left(\frac{\Delta g}{1.111} + \frac{\Delta cb}{1.056} + \frac{\Delta x}{1.136} \right)}{0.000377 \cdot (g + cb + x) + 0.998} \right\}} \tag{2-10}$$

The insoluble solids level is one of the most important variables affecting the accuracy in mass balances of solid-liquid phase systems, since all component mass balances in both phases ultimately require this value. Using insoluble solids measurements from a large set of saccharification experiments, the accuracy of the

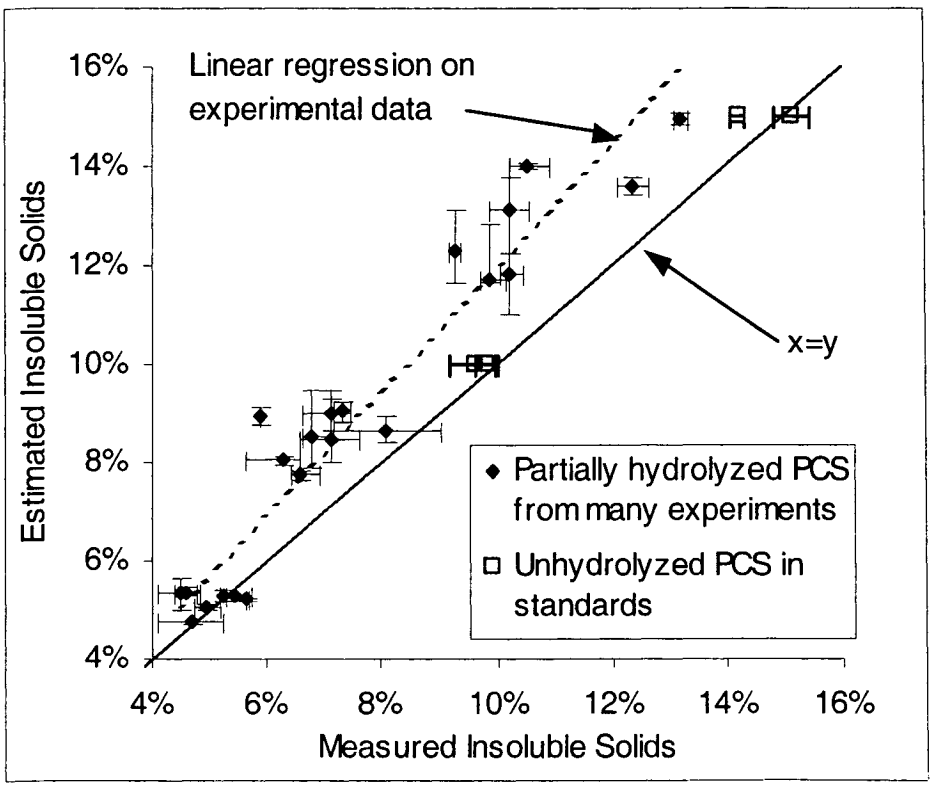


Figure 2.2: Comparison of measurement vs. prediction for insoluble solids level during enzymatic hydrolysis of PCS taken from many experiments.

solution to Equation 2.10 can be assessed relative to the measurement accuracy. The data in Figure 2.2 were generated under a variety of initial conditions and extents of reaction and were quantified using the analytical method outlined in **Materials and Methods** (Section 2.3). The resulting plot of measured vs. estimated has an R^2 value of 0.9214 and an apparent bias (linear regression slope of 1.23). There was uncertainty regarding the accuracy of the analytical procedure for measuring the insoluble solids directly (LAP-018), since the bias in the experimental data consistently show that either the measurement consistently under-predicts the insoluble solids, or the theoretical equation (Equation 2-10) consistently over-predicts. Postulated hypotheses for these inconsistencies include:

1) There are significant levels of other soluble components in the liquid fraction that are not accurately quantified with the current HPLC method (*e.g.* cellodextrin oligomers), which would result in lower measured insoluble solids.
2) The analytical method for insoluble solids measurement relies on repeated cycles of washing samples in DI water, centrifuging, and decanting, during which insoluble solids could be lost, thus the measured value is lower.
3) The method of sampling by tip-less 5 mL pipette selectively removes a greater fraction of liquid than is present in the bulk reactor system (and thereby less insoluble solids), resulting in lower measured insoluble solids.

For Hypothesis 1, soluble cellodextrins with a DP higher than cellobiose are especially apparent in HPLC chromatograms for high-solids saccharifications, where glucose accumulates to the point where cellulases are inhibited. Klesov (1982) considers quantification of soluble glucan polymers to be necessary for determining mass balances

and cellulose degradation kinetics. Because these polymers are not quantified and the error is apparently more acute at higher solids (where there is more product inhibition and accumulation of soluble cellodextrins), it is suspected that this is the most likely reason for the discrepancy between the two methods. Hypothesis 2 was examined by saving the wash supernatant, filtering this through tared filter paper, and drying in a desiccator. The results indicate that very little solids are lost in washing. To address Hypothesis 3 a series of solids standards were generated as described in Table 2.2 in **Materials and Methods** (Section 2.3) and tested using the same analytical method. Using these standards, the disparity between measured and calculated insoluble solids is not present (Figure 2.2). This rules out the possibility of Hypotheses 2 and 3, making cellodextrin quantification the most likely source of the disparities. Because insoluble solids estimates are needed to determine g/kg mass concentrations of components and hence cellulose conversion, the conversions actually calculated are expected to be slightly low due to the bias in the insoluble solids estimates (which do not account for soluble cellodextrin).

2.4.2 Time Course Mass Balances

The sugar concentration time profiles and estimated cellulose conversion are plotted in Figure 2.3. These values were determined by averaging the HPLC results from Flask 1 and the flask being sampled for solids analysis. The results are similar to those obtained from other experimental work (**CHAPTER 3**). The cellulose conversion was estimated using glucose and cellobiose concentrations. Reacting sugars can be tracked in terms of absolute mass concentrations [g/kg] or with respect to a reference component. Because the mass of polymers of glucan or xylan change with the hydrolysis of β-1,4 bonds

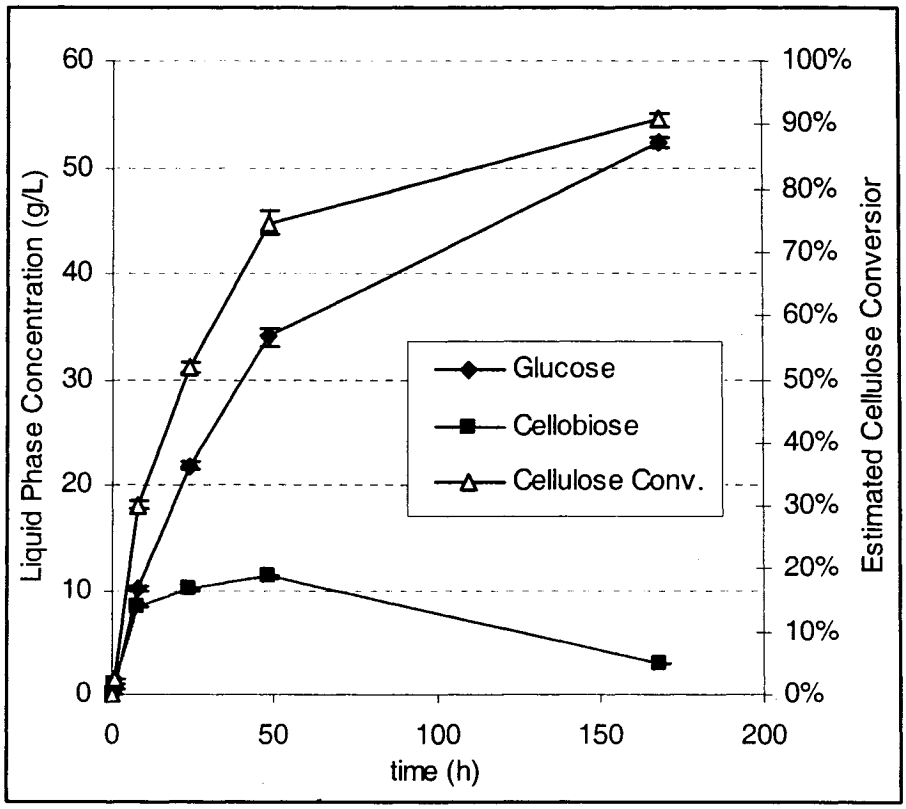


Figure 2.3: Reactor liquid phase composition (by HPLC) and cellulose conversion.

during saccharification, a basis needs to be used to compare mass amounts of unreacted to reacted material in each phase. One alternative to tracking absolute concentrations of components in the reaction is to track relative amounts based on determining the ratio of each component to an "inert" component such as ash. A problem with this approach, however, is that it is not known with absolute certainty that suspected inert components are not changing in composition during the reaction. Also, the value determined in this way is not useful for assessing mass balances. In biological and organic reaction systems, elemental carbon is often used as a reference compound (Schell *et al.*, 2003; Schell *et al.*, 2002; Sáez *et al.*, 2002). This is done to provide a consistent basis for mass balances when individual component masses are altered in chemical reactions. Carbon (as mass) is chosen as a basis since total carbon will not change during hydrolysis (*i.e.* no C loss as CO_2 as is the case for fermentation) and can be used as a consistent measure to track all reacting sugars. This will allow a direct determination of the overall transfer of material between the solid and liquid phase, and can yield a complete mass balance on reactive components. Table 2.3 provides a summary of the conversion factors needed to determine components on a carbon mass basis.

Table 2.3: Carbon Content in Sugars

	Glucose	Cellobiose	Polymeric Glucan (assuming DP>>1)	Xylose	Polymeric Xylan (assuming DP>>1)
Carbon mass fraction (g C / g compound)	0.400	0.421	0.444	0.400	0.455
Molecular weight (g / mol polymeric unit)	180	342	162	150	132

Using time-course data generated from the saccharification experiment, two general classes of mass balances can be made. The first class is for balances at a single time point, while the second compares the totals in both phases between time points. For each of these classes balances can be performed around both the total solids and individual solids components. The total solids balance is independent of insoluble solids composition, but still requires soluble solids composition determined by HPLC, and insoluble solids measurements. The component balances require these same measurements in addition to the insoluble solids composition. When using a carbon mass basis, the mass carbon / mass compound conversion factor (Table 2.3) is needed in the equation. These equations for the first class of mass balances can be written as:

	Total Solids	=	**Solid Phase (Insoluble Solids)**	+	**Liquid Phase (Soluble Solids)**	**Eq. #**
Total Solids Mass Balance	$S_T(t)$	=	$S_I(t)$	+	$G(t) + CB(t) + X(t)$	(2-11)
Glucan Mass Balance* (on a C g/kg basis)	$G_{total}(t)$	=	$S_I(t) \cdot x_{cell}(t) \cdot 0.4444$	+	$G(t) \cdot 0.400 + CB(t) \cdot 0.4211$	(2-12)

* Equivalent balances can be performed for all sugar components

Where:

S_T = total solids [g/kg]

G_{total} = total glucan in both phases [g C / kg slurry]

$x_{cell}(t)$ = mass fraction glucan [g/g]

Considering the conservation of mass, the second class of mass balances compares solids levels between time points. For the total solids, because there are initially no soluble solids, this becomes:

$$S_T(t) = S_{T,0} = S_{I,0} \tag{2-13}$$

Because this balance is not on a carbon basis, there will be unavoidable errors associated with the change in molecular weight of the reacting compounds. For each component in terms of carbon, the balance between time points on glucan, as an example, is:

$$G_{total}(t) = G_{total,0} \quad (2\text{-}14)$$

Many of the items in these mass balance equations can be both measured and calculated from the remaining measured variables in the equations. When this is the case, the two values can be compared to assess the accuracy of either the balance or the quantification method. Mass balances around the total solids were performed using both direct measurements and estimates with Equations 2-10 and 2-11, with the results plotted in Figure 2.4. There is good agreement between the measured and predicted values for the insoluble solids as was expected from the results of Figure 2.2. However, the measured and predicted values for the total solids show differences since the total solids measurements include citrate buffer and enzyme, which are not measured independently. The solids content increases during the course of the reaction due to the increase in molecular weight of cellulose subunits as the degree of polymerization is decreased and water molecules are added across the β-1,4 bond. This trend is apparent in the data for both measured and estimated total solids.

Insoluble solids compositional changes during the course of saccharification are plotted in Figure 2.5. The major trends are that glucan and other sugar components decrease as a fraction of the total composition, while lignin, ash, and protein increase, as would be expected. This plot gives only relative concentrations of each component, so that to determine an absolute mass concentration, the component mass fraction must be multiplied by the insoluble solids. This is done for Figure 2.6, which shows the total

mass concentrations for every measured component in both phases throughout the course of the saccharification reaction. The soluble components appear as shaded bars. This plot is basically all of the components of the total solids mass balance in Equation 2.11, and the total solids concentrations (the top of the bar) correspond to the "estimated total solids" plot in Figure 2.4. These same data (reactive sugars are converted to a C basis) are replotted in Figures 2.7, 2.8, and 2.9 to demonstrate the mass balance closure for individual components. One validation of the mass balance method is that the non-carbohydrate components (lignin, ash, and protein) remain constant throughout the course of the reaction (Figure 2.7), consistent with the hypothesis that these constituents remain inert. A validation of the component sugar mass balance (Equations 2.12 and 2.14) and quantification methods is that the total level of glucan and xylan (tracked as g C / kg slurry) remains approximately constant throughout the reaction (Figure 2.9), while the levels of these components in each phase varies significantly. The total component concentrations only vary from the mean value by an average difference of 1.9% and 8.8% for glucan and xylan, respectively. Arabinan levels in the liquid phase were too low for reliable detection by HPLC (Aminex column HPX-87H) such that only concentrations in the solid phase are given in Figure 2.8. The initial time point for insoluble solids analysis may show higher sugar levels than were present. This trend is apparent for all sugars and may be due to the differing extents of washing performed on the solids used for analysis and those used in the experiment, since some of the sugar may be water-soluble. During the course of the reaction, concentrations of all insoluble hemicellulose sugar polymers show up to 50% decreases, possibly due to cross reactivity of the cellulase system to

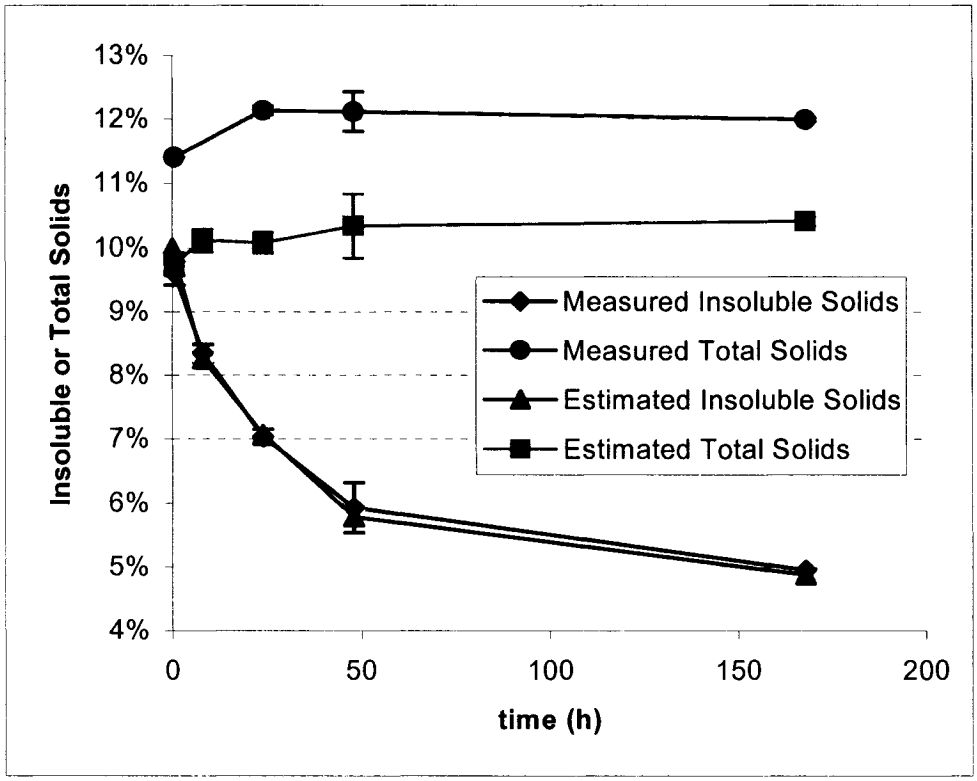


Figure 2.4: Measurement and estimation of soluble and insoluble solids.

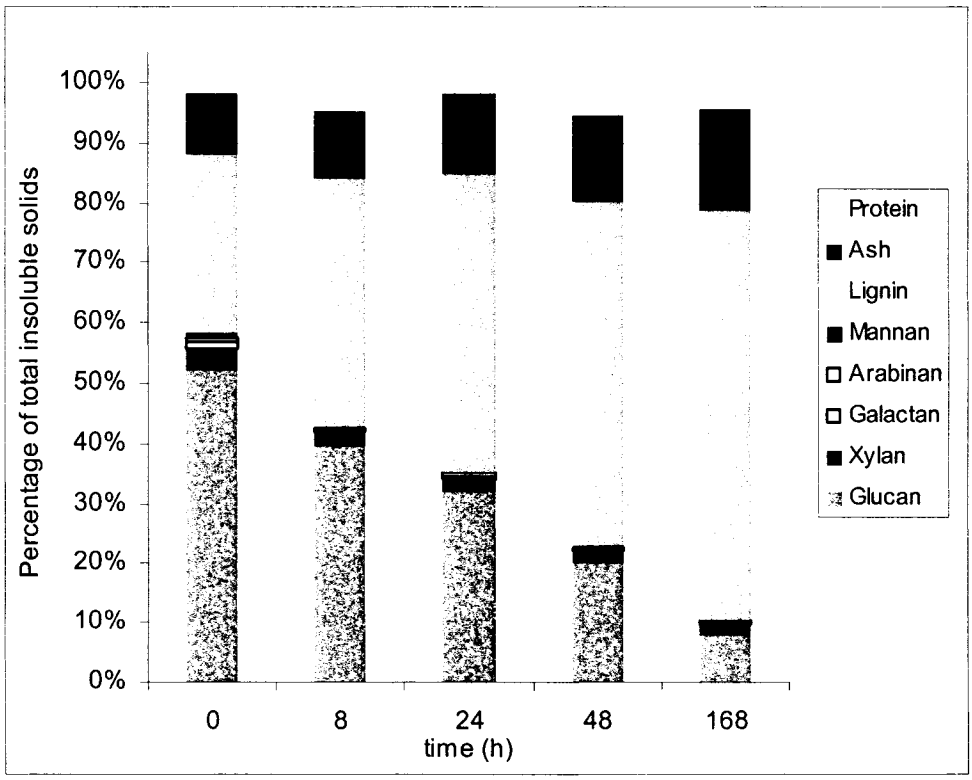


Figure 2.5: Solid phase compositional changes during the course of saccharification.

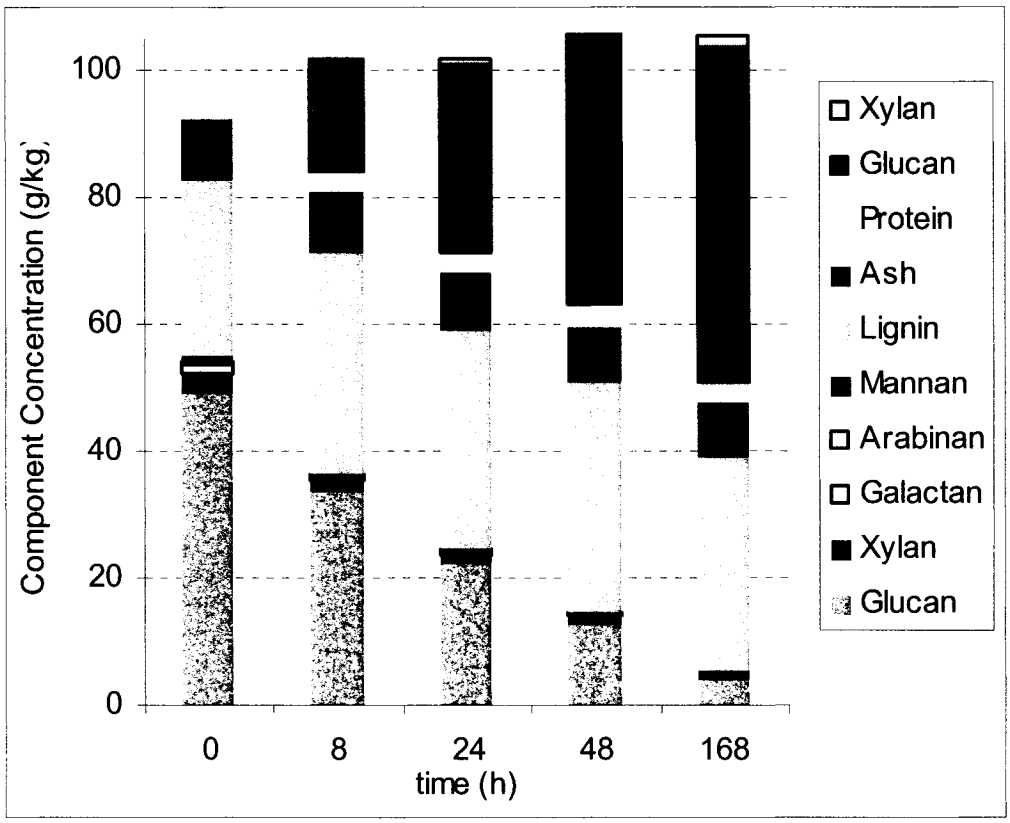


Figure 2.6: Multiphase solids mass balance closure. Soluble liquid phase components are shaded.

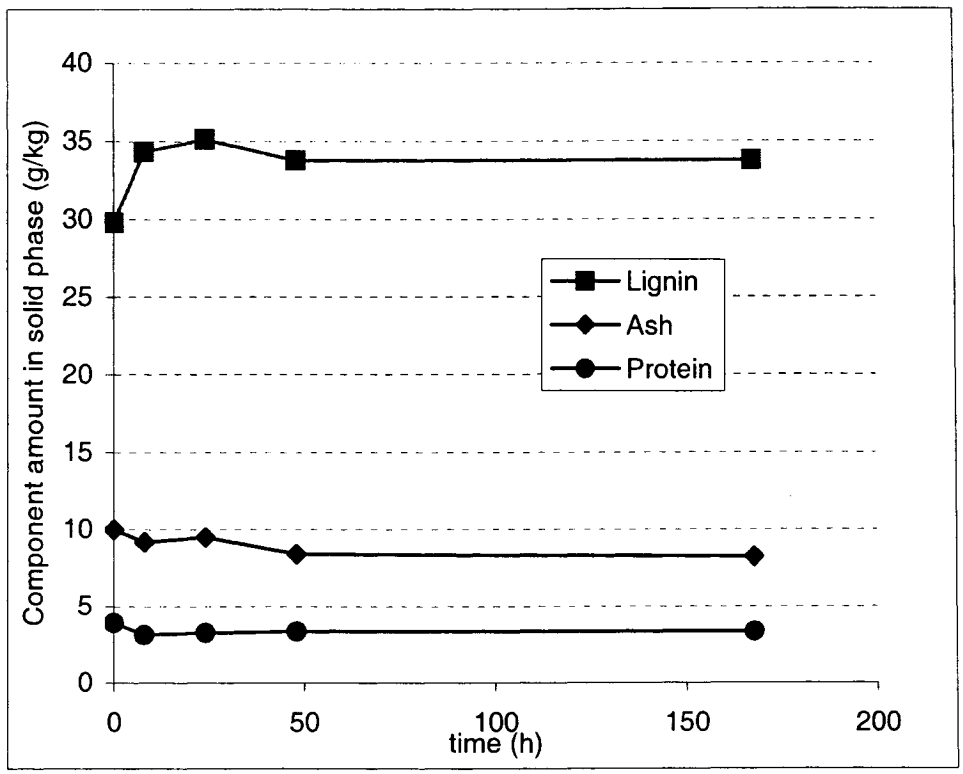


Figure 2.7: Mass concentration of inert/non-carbohydrate components in the solid phase during saccharification.

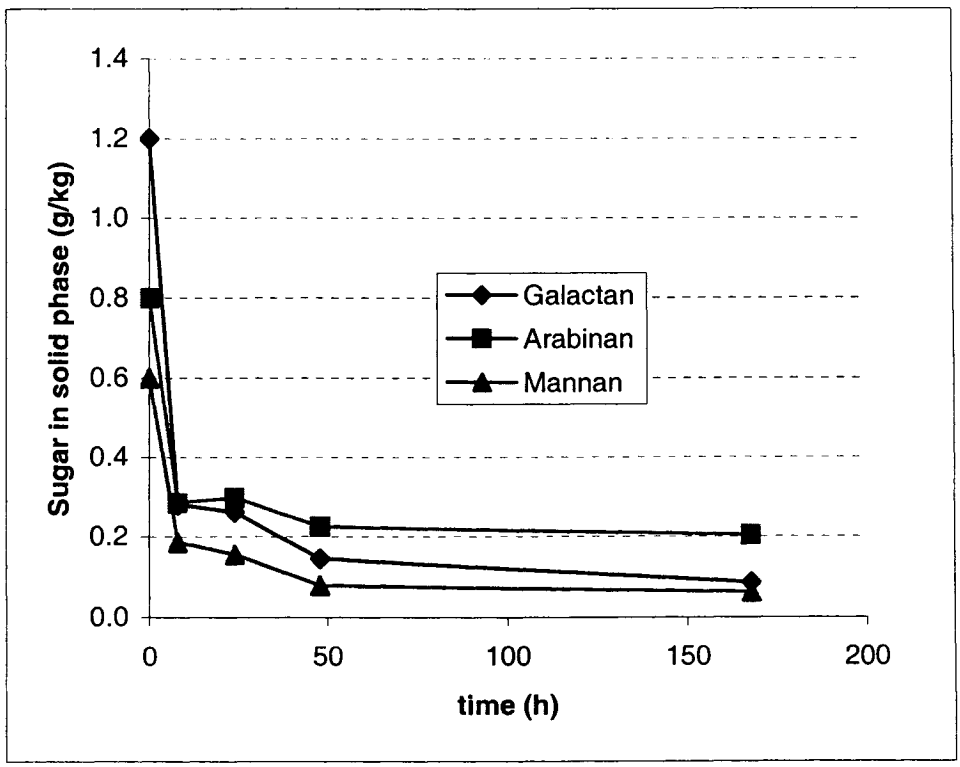


Figure 2.8: Mass concentrations of non-cellulosic sugars in the solid phase during saccharification.

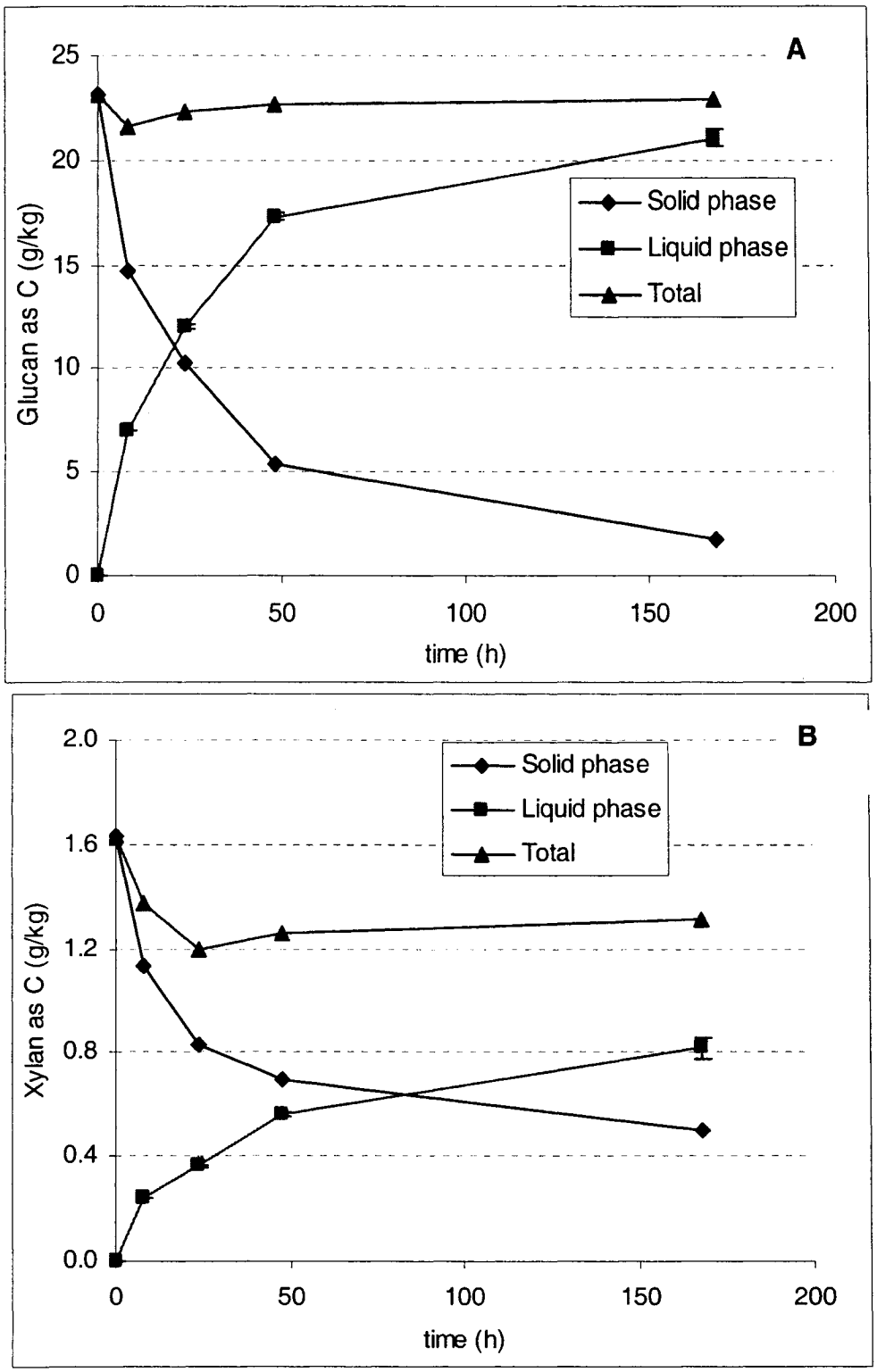


Figure 2.9: Estimated concentrations of glucan and xylan in both phases converted to g/kg carbon.

hydrolyze hemicellulose. Mass balances on these hemicellulose sugars are discussed in the following section.

2.4.3 Comparison of Analytical and Mass Balance Methods

Two separate analyses of the liquid phase composition were made. The first quantified only glucose, cellobiose, and xylose by direct HPLC measurement (Aminex HPX-87H column), and the other method quantified all soluble oligomeric and monomeric sugars using NREL LAP-013, 014, and 015 (performed by Linna Thanh and Justin Sluiter). These two approaches to the quantification of liquid phase composition were compared, and the sensitivity of variables calculated from these values (cellulose conversion and insoluble solids) was evaluated. The two methods are referred to as "partial sugars" and "complete sugars" analysis respectively to simplify the discussion.

Figure 2.10 provides the results of compositional analysis of the liquid phase (soluble) solids using the two analytical methods. The data for the complete sugars analysis show a significant portion of soluble pentose and hexose sugars may exist as oligomers (21 g/L of soluble oligomers). The implications of this are that hemicellulose may be depolymerizing during the saccharification reaction, yielding a lower calculated level of insoluble solids (and higher estimated cellulose conversion) than would be predicted based on the summative approach using partial sugars analysis.

Component mass balances (Equation 2.11) between the two phases for hemicellulose sugars are performed using the results of the complete sugars analysis. For mass balances on the hemicellulose sugars, each component was not converted to a carbon basis since much of the sugar is present as soluble oligomers. Because the DP of these oligomers is not known, conversion to a carbon basis would probably not improve the

accuracy. These multiphase compositions are given in Figure 2.11. Figure 2.11A gives basically the two-phase component balances of Equation 2.12, only without the carbon conversion basis. Comparing these component concentrations to the initial composition in Figure 2.11B shows that values are not consistent over time, indicating that the complete sugars analytical method probably quantifies more sugar than is actually present. This analytical method may be more prone to error due to the number of sample modification steps (acid dilution, pH adjustment) and assumptions (sugar degradation) required in the quantification.

A comparison of the analytical approaches for the calculation of cellulose utilization (summative approach vs. direct cellulose measurement as outlined in the **Background**, Section 2.2) can be made using the two liquid phase analytical methods for the data generated in this work. Estimated and measured insoluble solids are plotted in Figure 2.12 for the final time sample. This shows that the agreement between the measured value and the summative approach using the partial sugars analysis is relatively good, with the same trend that the estimated value is greater than the measurd value, a fact that was apparent in Figure 2.2. When using the summative approach with the total sugars analysis, the insoluble solids are greatly under-predicted (more than 10 g/kg less than the measured value). Applying these same data for estimation of the cellulose conversion (Figure 2.13) results in greater than 100% conversion calculated for the final time point, due to higher levels of soluble glucan from the total sugars analysis results. Using the partial sugars analysis results in calculated cellulose conversions less than the measured value (average of 3% less). As discussed previously, this is most likely due to some soluble oligomer sugars being overlooked. However, since the current complete sugars

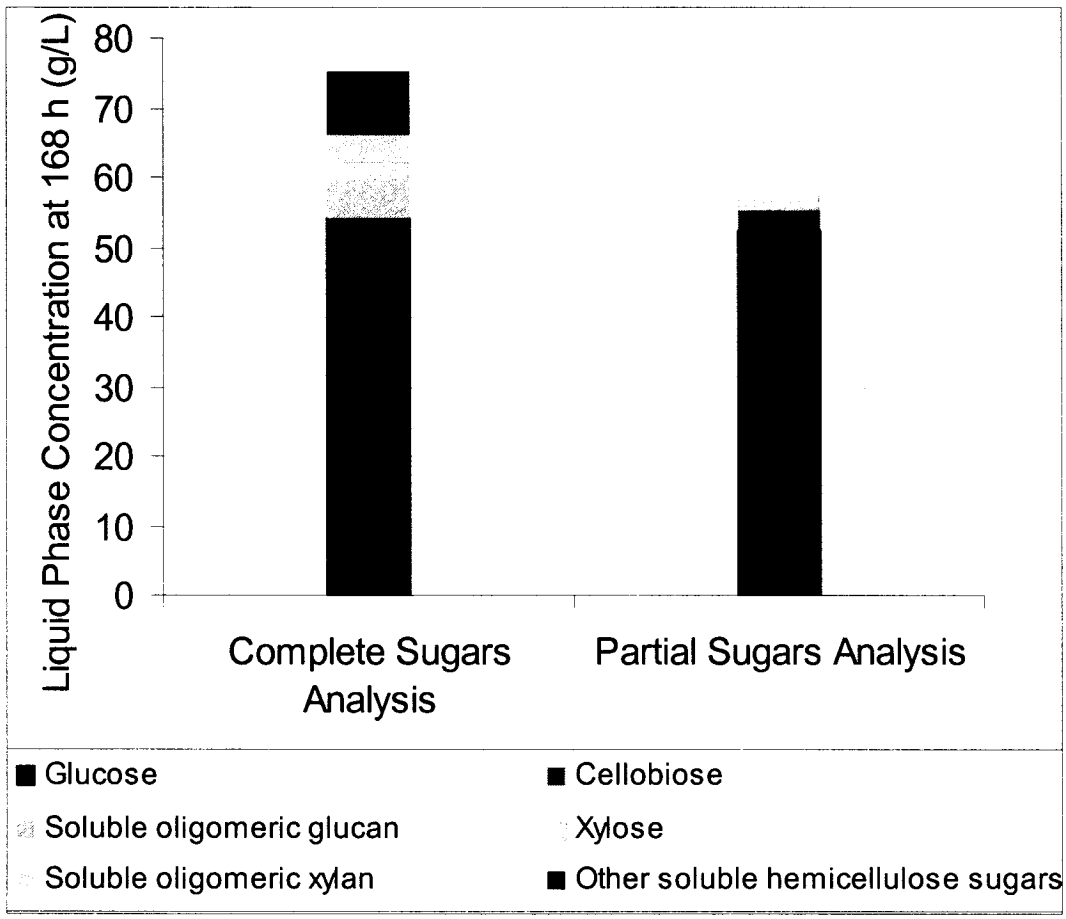


Figure 2.10: Comparison of liquid phase quantification methods.

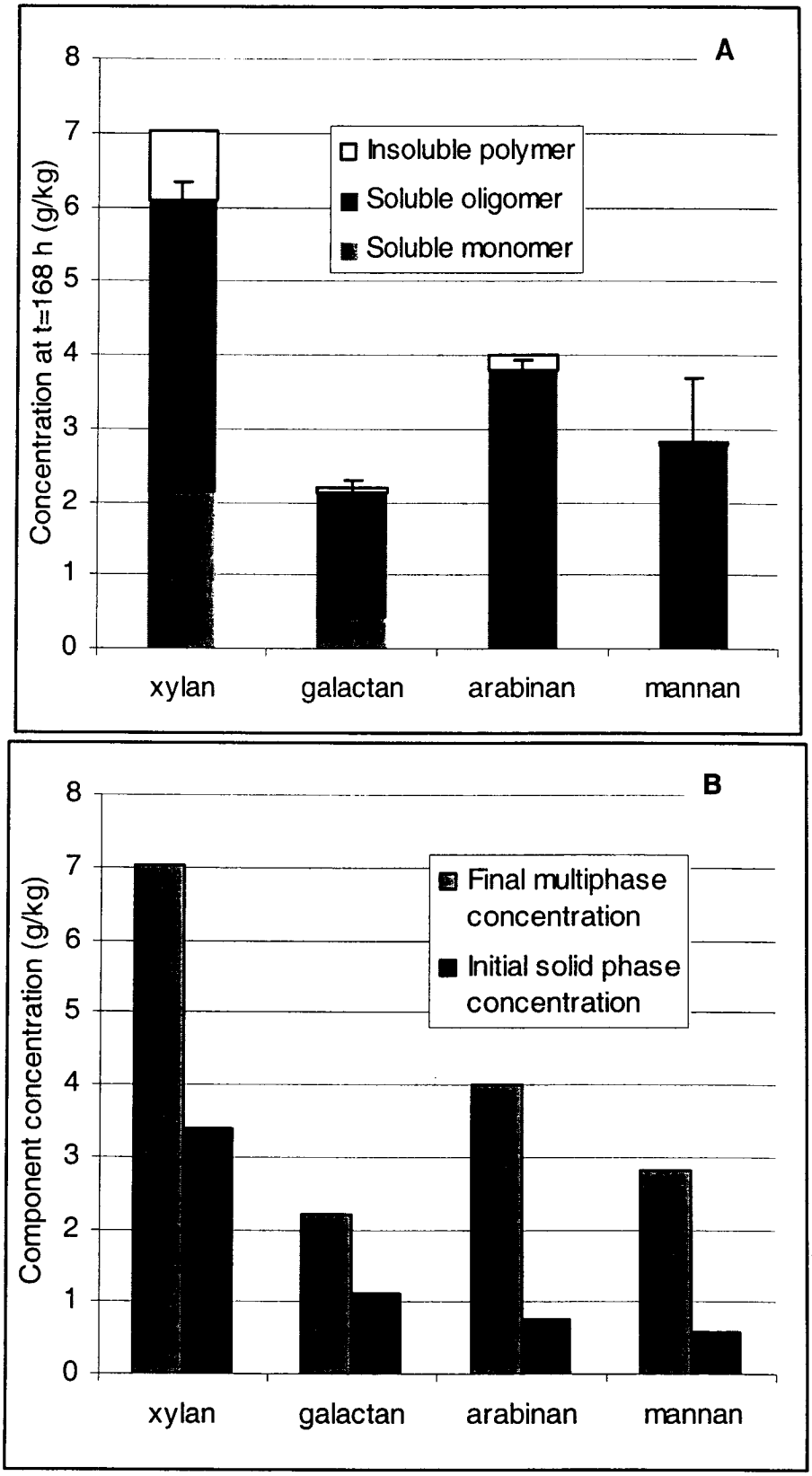


Figure 2.11: Hemicellulose sugars in both liquid and solid phase determined by complete sugars analytical quantification.

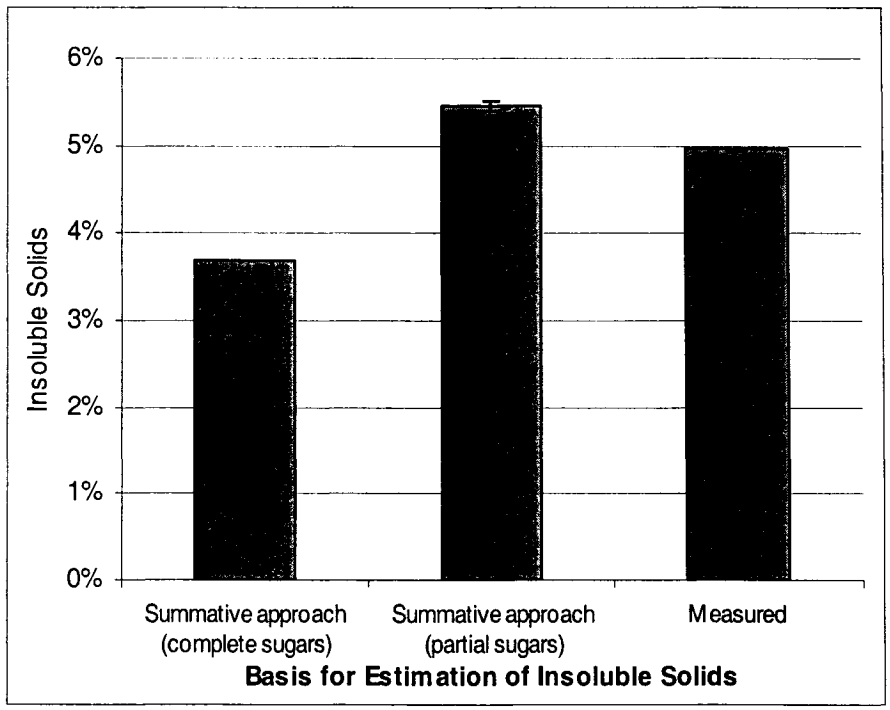


Figure 2.12: Comparison of analytical techniques for determination of insoluble solids.

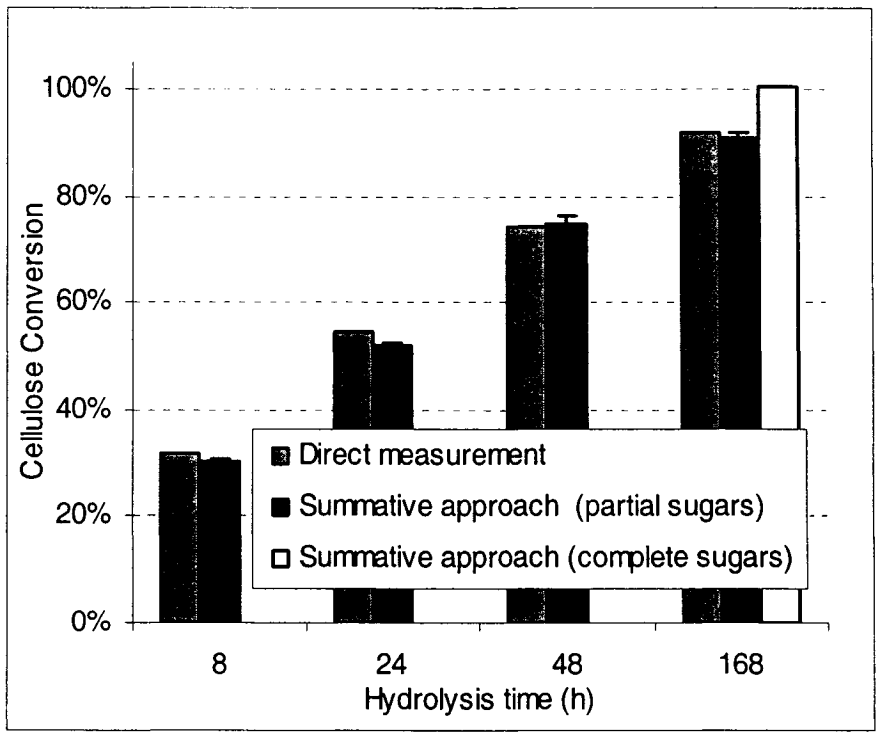


Figure 2.13: Comparison of methods for estimation of cellulose conversion.

quantification methods apparently over-predict the soluble sugars, the partial sugars analysis may for now be the most reliable indirect method for estimating cellulose conversion.

2.5 Conclusions and Recommendations

In the preceding work, a number of issues related to the analytical quantification and the mathematical formulation of mass balances for enzymatic saccharification of PCS. It was determined that calculating insoluble solids based on HPLC measurements of glucose and cellobiose resulted in estimates with a high bias relative to the direct insoluble solids quantification method. This was concluded to be the result of not quantifying soluble cellodextrins formed from cellulose. However, mass balances for both glucan and xylan showed remarkably good closure over the liquid and solid phases even without cellodextrin quantification (average differences over the course of the reaction of only 1.9% and 8.8%, respectively). Comparison of analytical methods indicated that the quantification of total (monomeric and oligomeric) sugars was high, possibly due to the number of sample modification steps required in the procedure. Future work should focus on identifying problems with the current methods for quantification of sugar oligomers that currently over-predict these concentrations. Additionally, while this work demonstrated satisfactory two-phase mass balance closure on xylan in the hemicellulose fraction, other hemicellulosic components (mannan, arabinan, and galactan) need improvement in quantification to allow the calculation of mass balances and estimation of conversions.

2.6 References

Chung YC, Bakalinsky A, Penner MH (1997). Analysis of biomass cellulose in simultaneous saccharification and fermentation processes. *Appl Biochem Biotechnol.* 66(3):249-62.

Ehrman CI and Himmel ME (1994). Simultaneous Saccharification and Fermentation of Pretreated Biomass: Improving Mass Balance Closure. *Biotechnol Tech.* 8(2):99-104.

Erickson LE, Minkevich IG, Eroshin VK (1978). Application of mass and energy balance regularities in fermentation. *Biotechnol Bioeng.* 20(10):1595-1621.

Klesov AA (1982). Regulation of the formation and conversion of intermediate cellooligosaccharides and cellobiose during ezymatic hydrolysis of insoluble cellulose (Article in Russian) *Biokhimiia.* 47(4):608-18.

Sáez JC, Schell DJ, Tholudur A, Farmer J, Hamilton J, Colucci JA, McMillan JD (2002). Carbon mass balance evaluation of cellulase production on soluble and insoluble substrates. *Biotechnol Prog.* 18(6):1400-7.

Schell DJ, Sáez JC, Hamilton J, Tholudur A, McMillan JD. (2002). Use of measurement uncertainty analysis to assess accuracy of carbon mass balance closure for a cellulase production process. *Appl Biochem Biotechnol.* 98-100:509-523.

Schell DJ, Farmer J, Newman M, McMillan JD. (2003). Dilute-sulfuric acid pretreatment of corn stover in pilot-scale reactor: investigation of yields, kinetics, and enzymatic digestibilities of solids. *Appl Biochem Biotechnol.* 105 -108:69-85.

Vinzant T, Panfick L, Nagle NJ, Reynolds JB, Himmel ME. (1994). SSF Comparison of Selected Woods from Southern Sawmills. *Appl Biochem Biotechnol.* 45/46:611-626.

Weast RC (1985). *CRC Handbook of Chemistry and Physics.* 66th Ed. CRC Press Inc. Boca Raton, FL.

Zhang YHP and Lynd L (2004). Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. *Biotechnol Bioeng.* 88(7):797-824.

Chapter 3

CHARACTERIZATION OF RATES AND RATE-ASSOCIATED PARAMETERS IN HIGH-SOLIDS ENZYMATIC SACCHARIFICATION OF PRETREATED CORN STOVER

3.1 Overview and Context

While operating the cellulose saccharification step at high levels of initial insoluble solids is attractive from the perspective of process economics, high-solids enzymatic saccharification of cellulosic biomass presents a number of challenges for both the reaction and the reactor system. Chapters 4 and 5 will characterize and address the parameters affecting the reactor vessel, while this part of the work maintains a special focus on the enzymatic reaction and the variables affecting the rate of cellulose hydrolysis. The variables were characterized as dependent on substrate properties, enzyme properties, or mass transfer. Substrate-related rate properties were isolated by defining a "specific" cellulose hydrolysis rate to minimize the effects of changing substrate concentration. Applying this approach for enzymatic hydrolysis of pretreated corn stover (PCS) showed three distinct rate regimes corresponding to changes in the substrate morphology over the course of the reaction. For closed system saccharification (CSS) in shake flasks, initial insoluble solids were varied from 25% to 32% at several enzyme loadings and resulted in final glucose concentrations ranging between 125 and 160 g/L (7 and 14 day conversions, respectively). Achieving glucose concentrations this

high has significant positive implications for the economic feasibility of a biomass-to-ethanol or biomass-to-monomeric sugars industrial process. Using background sugars and comparisons between CSS and simultaneous saccharification and fermentation (SSF), minimal inhibition by glucose of the cellulose hydrolysis rate was detected for glucose concentrations as high as 120 g/L. Limitations in the reaction rate due to mass transfer were identified at initial insoluble solids concentrations >20%. All mass transfer, both internal and external, was summarized with an effectiveness factor that was found to be linearly correlated to the initial insoluble solids concentrations. At initial insoluble solids concentrations >25%, convective mass transfer (as measured by an increase in reaction rate) was found to be improved by agitation in the high-solids bioreactor (HSBR) reactor relative to shake flask reactors.

3.2 Background

3.2.1 Variables Affecting Enzyme Performance at High Solids

"High-solids" enzymatic saccharification can be roughly defined as beginning at the insoluble solids level (S_I) where free liquid is no longer present in the slurry such that the separation of a liquid and solid phase from the suspension is not spontaneous. The presence of free water in high-solids PCS slurries depends strongly on both the insoluble solids level and the glucan content of the solids, which influences lignocellulose-water interactions and cellulose swelling. Alternatively, this high-solids slurry definition can be regarded as the solids region where the slurry viscosity is highly non-Newtonian (Pimenova *et al.*, 2003). For PCS solids, this region begins at approximately 12%-15% insoluble solids (by weight). Performing the saccharification reaction at high insoluble solids introduces a new set of process-related problems associated with slurry mixing and

cellulase enzyme effectiveness. Factors associated with slurry mixing include temperature control and bulk mixing, which further affect enzyme efficiency. Enzyme-dependent factors include sugar inhibition (resulting from increased solids levels and diffusion of both sugars away from the catalytic site and enzyme to new targets) and enzymatic reaction temperature.

While several saccharification studies have used fed-batch to achieve more than 15% insoluble solids (Varga *et al.*, 2004; Ballesteros *et al.*, 2002), there are also several documented high-solids batch saccharification studies (Mohagheghi *et al.*, 1992; Spindler *et al.*, 1988). All of these studies used SSF to offset the recognized effects of glucose inhibition. High-solids SSF work by Mohagheghi *et al.* (1992), arguably the most notable of these studies, used pretreated wheat straw at initial insoluble solids concentrations as high as 32.3%. SSFs were performed in 1 L horizontally revolving reactors developed from a laboratory ball mill. Key findings included determining that for insoluble solids concentrations greater than 20%, yeast cells lost viability after 4 days (presumably due to the osmotic stress at these low water contents), and that 6-day glucose (equivalent glucose as sugar and EtOH) reached as high as 130 g/L.

For high-solids enzymatic saccharification, the reaction rate is of critical importance to the economic feasibility of a larger scale industrial process, since this unit operation requires by far the longest residence time relative to the other major biomass conversion reactions of pretreatment and fermentation. These longer residence times during saccharification translate into higher operating and capital costs per unit of product output. Important parameters linked to the reaction rate that can become more

pronounced in high-solids CSS studies include: sugar inhibition, thermal inactivation, substrate morphology (reactivity), and mass transfer.

3.2.2 Substrate Derived Rate Factors

There exists a large body of literature describing how the rate of enzymatic hydrolysis of cellulose is influenced by substrate-related factors. Some of the proposed factors affecting substrate reactivity include: cellulose crystallinity, degree of polymerization, particle size, particle surface area, lignin content, pore volume, and acetylation (Lynd *et al.*, 2002; Davis et al., 2002; Chang and Holtzapple, 2000; Mansfield *et al.*, 1999; Mooney *et al.*, 1997; Excoffier *et al.*, 1991; Puri, 1984). All of these factors are strongly dependent on the ultimate biological source of the raw material, the pretreatment method, and the pretreated feedstock composition. Väljamäe *et al.* (1998) used two *T. reesei* exoglucanases (Cel7A and Cel6A) and bacterial microcrystalline cellulose as a substrate to show that rather than following simplified first-order kinetics, the degradation kinetics are compatible with a "surface erosion" model. This approach can apparently account for the more rapid initial rate of cellulose hydrolysis. Eriksson *et al.* (2002) discuss a model for the decrease in hydrolysis rate based on loss of synergism between endo- and exoglucanases based on "inactivation" of Cel7A. This "inactivation" was caused either by the enzyme binding to cellulose at unproductive sites or failure to release from the substrate after catalytically processing a cellulose chain. Additionally, this study was able to rule out thermal deactivation, product inhibition, and the release of soluble lignin components as potential reasons for rate decreases. The effect of enzymatic hydrolysis on the crystalline structure of pretreated cellulosic biomass has been investigated by Ramos *et al.* (1993). This work found that changes in the degree of polymerization and

crystallinity of cellulose were dependent on the substrate pretreatment method, and that a correlation between the decrease in both chain length and crystallinity with extent of cellulose conversion could not be drawn for the various substrates used. Similar work by Puri (1984) could not link crystallinity to rate or extent of saccharification. Davis *et al.* (2002) determined that amorphous cellulose from pretreated yellow poplar is preferentially hydrolyzed by a mixture of CBH I and an endoglucanase.

Lignin can inhibit cellulose enzymatic hydrolysis rates by interacting directly with the cellulases and sterically blocking access to catalytic sites by either directly covering catalytic sites or preventing cellulose fiber swelling. In several studies (Palonen *et al.*, 2004; Excoffier *et al.*, 1991), the nonspecific binding of cellulases to lignin has been identified as a problem that can inhibit the rate of cellulose hydrolysis. Palonen *et al.* (2004) demonstrated that Cel7A and Cel5A can bind to pretreated spruce lignin. This nonspecific binding of cellulases to lignin has been modeled using a Langmuir adsorption isotherm (Converse *et al.*, 1990; Ooshima *et al.*, 1990). Treatment of lignocellulosic biomass with non-catalytic protein (BSA) during the enzymatic hydrolysis of cellulose has been shown to improve hydrolysis rates, in that the noncatalytic protein can bind to the lignin to prevent nonspecific binding of cellulases (Wyman and Yang, 2004). In contrast to this, Muenier-Goddik and Penner (1999) showed that supplementation of purified lignin to non-lignocellulosic cellulose did not result in decreases in either rate or extent of enzymatic saccharification. It has been speculated that lignin can also affect the digestibility of pretreated cellulosic substrates due to steric hindrance of cellulases by reprecipitation onto the surface of the substrates (Nagle *et al.*, 2002; Lynd *et al.*, 2002). Lignin in relatively unaltered lignocellulosic biomass can reduce cellulose digestibility by

preventing water penetration and swelling of the cellulose. In refiner mechanical pulps Mooney *et al.* (1998) identified lignin content (correlated to fiber swelling) as limiting cellulose digestibility by hindering accessibility to the recalcitrant cellulose. Out of the three parameters examined in one study using pretreated poplar (Chang and Holtzapple, 2000), it was found that lignin content and crystallinity have significant effects on the digestibility of cellulose, while acetyl content in the hemicellulose is not as significant.

Based on the substrate morphological factors relating to the hydrolysis rate outlined above, a broad mathematical treatment of these factors can be developed to demonstrate how this understanding can be rationally formulated using different approaches. The substrate factors affecting the hydrolysis rate have been broadly formulated in mathematical models as either "substrate reactivity" or as dependent on substrate concentration.

The rate of cellulose enzymatic hydrolysis is ultimately dependent on the availability of adsorption binding sites in proximity to sites accessible to enzymatic attack. This is commonly modeled using Langmuir or Freundlich adsorption (Kadam *et al.*, 2004). It has been determined that although highly pretreatment-dependent, the surface area available for enzyme adsorption varies nonlinearly with respect to extent of conversion (Converse et al., 1990), which could partially explain the much more rapid initial rate of hydrolysis encountered for most substrates. Wald *et al.* (1983) approximate the accessible surface area as a shrinking sphere; although this approach ignores the complex pore structure and swelling of cellulose fibers. Many mathematical models of cellulose enzymatic hydrolysis are based on the assumption that the cellulose rate is first order with respect to cellulose (Kadam *et al.*, 2004; Phillipidis and Hatzis, 1997; Gusakov and

Sinitsyn, 1992; Asenjo, 1984). These models ultimately assume (never explicitly, however) that the concentration of catalytic sites (accessible surface area) is proportional to the cellulose concentration. Based on this mathematical formulation, the cellulose utilization rate [g cellulose / kg slurry / h] can be converted to a specific cellulose utilization rate [h^{-1}] by dividing the rate by the estimated cellulose concentration. The motivation for this is to remove the substrate concentration dependence of the rate term, which is probably the largest of the chemical-dependent rate factors, so that the effects of other chemical-dependent rate factors can be isolated:

$$\frac{dC}{dt} = f(C, E, G, CB, X, T, pH) \approx [C] \cdot f(E, G, CB, X, T, pH) \tag{3.1}$$

where rearranging yields:

$$\frac{1}{[C]} \cdot \frac{dC}{dt} \approx f(E, G, CB, X, T, pH) \tag{3.2}$$

This assumption is valid when the cellulose utilization rate is first-order with respect to cellulose concentration. It should be apparent, though, that the rate is still a function of model-considered variables such as temperature and concentrations of glucose, cellobiose, temperature, and enzyme, as well as non-modeled variables such as pH, cellulose crystallinity, substrate accessibility (available surface area), enzyme inactivation, and mass diffusion in high-solids systems.

Other models consider a "substrate reactivity (R_S)" term as in the work of Kadam *et al.* (2004). This term was developed by plotting glucose rate vs. C/C_0 with the intention to identify decreases in rate as the cellulose is depolymerized due to changes in crystallinity and morphology of the substrate. The result was that the glucose rate is exactly linear with respect to C/C_0, which was the basis for the reactivity as:

$$R_S = \frac{C}{C_0} \tag{3.3}$$

which is similar to the extent of cellulose conversion (ξ), or:

$$\xi = 1 - \frac{C}{C_0} \tag{3.4}$$

This reasoning assumes that C_0 was constant for all data used as is the case for hydrolysis "restart" experiments, where hydrolysis is stopped, enzyme removed and fresh enzyme added to the partially hydrolyzed substrate. However, it was not considered that the decrease in rate was due to the approximately first order reaction kinetics, so that adding a reactivity term effectively results in rates that are second order in cellulose concentration. Using a similar methodology, Ooshima *et al.* (1991) detected no decrease in the reactivity of Avicel as a function of conversion, which may be expected for a relatively homogenous processed or manufactured substrate such as Avicel. Other authors have developed empirical equations to relate a change in cellulose utilization rate to the extent of conversion (Ohmine *et al.*, 1983).

A further application of this same model by Desai and Converse (1997) detected little decrease in the reactivity of acid pretreated hardwood using hydrolysis restart experiments. For all these studies, the empirical correlation used was:

$$\frac{d\nu}{d\xi} = -k \cdot \nu \tag{3.5}$$

where:

$$\nu = \frac{dC}{dt} \tag{3.6}$$

and k is an empirical constant. Integrating this equation from an initial condition ($\xi = 0$, v_0 = initial rate) yields:

$$\frac{dC}{dt} = v_0 e^{-k\xi} \tag{3.7}$$

While this is clearly not first order in cellulose and contains no enzyme dependence, it can provide a suitable empirical approximation of the conversion profile, which Zhang and Lynd (2004) define as semi-mechanistic. This model form provides an exponential decrease in the cellulose rate with respect to cellulose conversion (ξ). Since dividing Equation 3.7 by cellulose concentration will still yield a decrease in rate with respect to time and conversion, the model clearly considers an exponentially decreasing substrate reactivity independent of the first order cellulose dependence mechanistic models such as Equation 3.2.

Adsorbed enzyme has been demonstrated to vary as a function of extent of conversion (Ooshima *et al.*, 1990 and 1991; Desai and Converse, 1997). Another model form considers the rate of cellulose utilization as n^{th} order in adsorbed and total enzyme (Converse *et al.*, 1990). This model is also based on the premise that substrate reactivity changes with respect to conversion due to changes in accessible surface area and inhibition by lignin either by steric hindrance or nonspecific binding of enzyme to lignin or lignin degradation products.

3.2.3 Enzyme Derived Rate Factors

With enzyme adsorption arbitrarily categorized as a substrate-derived cellulose hydrolysis rate factor, other major enzyme-related variables relevant to this work include temperature and product inhibition. Two parameters affecting the choice of optimum

enzyme reaction temperature are enzyme activity and stability. While activity typically increases with increasing temperature, stability decreases, such that a compromise must be found. Enzymes derived from mesophilic microorganisms such as *T. reesei* show an optimum temperature for extent and rate of conversion at approximately 45-50°C. More specifically, the *T. reesei* exoglucanase cellobiohydrolase I (Cel7A) has been demonstrated to begin unfolding at temperatures above 55°C over a wide range of pH values (Boer and Koivula, 2003; Jiménez *et al.*, 1995). The inactivation of Novozyme 188 β-glucosidase has been demonstrated after prolonged incubation periods at temperatures greater than 55°C (Calsavara *et al.*, 1999).

The other enzyme-related consideration is the inhibition of cellulases and β-glucosidases by cellobiose and glucose. Since cellulases are coupled to metabolic pathways in naturally occurring cellulase systems, there is no strong evolutionary pressure for enzymes to develop a resistance to glucose and cellobiose inhibition. Numerous literature studies demonstrate quantitatively that both glucose and cellobiose act as either competitive or noncompetitive inhibitors for *T. reesei* cellulases (Gruno *et al.*, 2004; Xiao *et al.*, 2004; Kadam *et al.*, 2004; Holtzapple *et al.*, 1990), with inhibition by cellobiose significantly greater than that of glucose. In particular, it has been demonstrated that these cellulases could retain nearly 37% of uninhibited activity on cellulose (initial rate only) at the remarkably high background glucose concentration of 546 g/L (Holtzapple *et al.*, 1990). This study, however, did not consider the concurrent accumulation of cellobiose, which is the primary inhibitor of cellulases. For practical implementation, the accumulation of high levels of glucose will severely inhibit the activity of β-glucosidases, which in turn leads to cellobiose accumulation and inhibition

of cellobiohydrolases. Because of this, the effect of glucose on β-glucosidase is the most important consideration in sugar inhibition of cellulases. Novel β-glucosidases have been isolated with exceptionally high tolerance to inhibition by glucose (Saha and Bothast, 1996), which could have the potential as a future cellulase supplement for a CSS process operated at high solids levels.

For practical applications designed around sugar inhibition, a number of approaches have focused on removing sugars during saccharification either physically or biologically. Knutsen and Davis (2004) used membrane ultrafiltration to remove sugars during saccharification of PCS. For the conditions studied, however, sugar removal was not found to enhance either the rate or extent of the cellulose conversion. Other work used a diafiltration saccharification assay (DSA) reactor (Baker *et al.*, 1997), which is a membrane reactor that continuously removes sugar to offset the effects of sugar inhibition and simulate SSF. Many other studies have used SSF to counteract the effects of sugar inhibition at higher levels of initial insoluble solids where high sugar concentrations would result if operated in CSS (Varga *et al.*, 2004; Ballesteros *et al.*, 2002; Borden *et al.*, 2000; Mohagheghi *et al.*, 1992; Spindler *et al.*, 1988).

3.2.4 Mass Transfer Limitations

Mass transfer is also a significant issue in reaction systems such as immobilized-enzyme catalyzed reactions in packed or fluidized-bed reactors (Fogler, 1999) and high-solids reaction systems (Lui and Wyman, 2004; Schell *et al.*, 2003). In these reaction systems, the rate of mass transfer can be the rate-limiting step if this rate is slower than the rate of reaction. This limitation in the mass transfer rate can be due to the rate of

diffusion of reactants to catalytic sites, other reactants, or diffusion of products away from the reaction site.

Two dimensionless numbers that are commonly used to characterize the effects of mass transfer in reaction systems are the Damköhler Number (Da) and the effectiveness factor (ε). The value of Da can be considered as a ratio of the rates of reaction and diffusion, typically defined as a ratio of a first order reaction rate constant and the diffusivity (Fogler, 1999). For values of Da > 1, mass transfer becomes the rate-limiting step. The effectiveness factor is often used in applications where the rate data is more empirically-derived, and can include the summative effects of a variety of mass transfer phenomenon other than diffusion. The effectiveness factor can be defined as the ratio of the actual reaction rate to the rate if there were no mass transfer limitations:

$$\varepsilon = \frac{\text{Actual overall rate}}{\text{Rate in the absence of mass transfer limitations}} \qquad (3.8)$$

An application example of ε includes the characterization of mass transfer effects in immobilized-enzyme catalysts in packed bed reactors where both internal and external mass transfer can be reaction rate-limiting (Fogler, 1999). For the modeling of mass transfer effects in high-solids pretreatment, Brennen and Wyman (2004) used separate kinetic models for reaction, biphasic leaching, and branched pore leaching to approximate hot water or dilute acid hydrolysis of hemicellulose. All kinetic parameters were determined empirically for these models. The leaching models are based on the solubilization of compounds within porous media such that chemical reaction and diffusion cannot be considered independently. In a similar study, Lui and Wyman, (2004) demonstrated that increasing the fluid velocity in a continuous pretreatment reactor improves the rate of hemicellulose hydrolysis, indicating that mass transfer limits

the rate of hydrolysis at slower flow rates. Less quantitatively than these approaches, other studies in the literature for the enzymatic hydrolysis of lignocellulosics (Mais et al., 2002; Ingesson et al., 2001) demonstrate that higher conversions and faster hydrolysis rates can be achieved in shake flasks by increasing the shaking rate from nearly stationary to a maximum at ~100-150 rpm for moderate levels of insoluble solids. This indicates that diffusion limits the rate of reaction at the lower shaking rates (minimal convective mass transfer) or where the reactor contents are nearly stagnant. These studies are relevant to high-solids work using shake flask reactors, since a variety of mass transfer conditions are encountered during batch saccharification: initially shaking does not provide convective mass transfer for mixing the slurry, while convective forces become increasingly more dominant as the slurry is saccharified.

3.3 Materials and Methods

3.3.1 Shake Flask CSS

All experiments were performed in 250 mL shake flasks with an initial working mass of 100 g. An Innova 4000 orbital incubator shaker (New Brunswick Scientific) was operated at 130 rpm, the temperature was maintained at 38°C, 45°C, or 50°C. Corn stover was pretreated with dilute sulfuric acid in NREL's Sunds vertical pulp digester. Pretreated corn stover (PCS) from NREL batch PS030312CS (unless otherwise indicated) and composed of 53.2% cellulose by dry mass. This pretreatment batch was using the same corn stover source and identical pretreatment conditions as in batch PS020502CS used for the work in **CHAPTER 2**. These solids were washed with more than 10 times the mass of DI water as solids and centrifuged at 15800 x g. Washed solids were used to

minimize any inhibition effects from background sugars entrained in the solids. The final glucose concentration in the wash water was verified to be less than 0.1 g/L.

The cellulase / β-glucosidase enzyme mixture used for this study was Spezyme CP (Lot # 301-00348-257, Genencor International). Total protein was assayed at 106.6 mg/mL (BCA assay, Pierce Biotechnology, Inc.) and the activity was determined to be 0.27 FPU/mg of protein using the NREL protocol LAP-006. For the modeling work, the ratio for cellulase : β-glucosidase was assumed to be 0.975:0.025. Enzyme loadings are given in units mg protein / g cellulose, with the two enzyme loadings of 20 mg/g and 40 mg/g used for this work, corresponding to 5.3 and 10.7 FPU / g cellulose, respectively. 1 M citrate buffer was prepared by adjusting the pH to 4.8 with NaOH, and was used as 5% of the total mass to yield an effective molality of 0.05 mol/kg. Enzyme, DI water, and citrate buffer were filter sterilized before use. PCS was added to the shake flasks and weighed before autoclaving. After autoclaving the shake flasks were weighed again and any water that has been lost is added back. Water loss to evaporation and reactor mass changes associated with sampling were tracked by weighing each flask before and after sampling. Water loss to evaporation was found to be minimal if shake flask caps are wrapped in Parafilm. For sugar inhibition studies, a high concentration glucose solution (250 g/L) was filter sterilized and substituted into the reactor in place of some of the water to achieve the proper insoluble solids level.

3.3.2 Simultaneous Saccharification and Fermentation

Shake flask reactors were prepared as described in the previous section with the only difference being the addition of 1% yeast extract and the yeast culture. Shake flasks were equipped with bubble traps to maintain an anaerobic environment within the flasks.

Inoculum was prepared from frozen stock cultures of *Saccharomyces cerevisiae* (strain D_5A) in 2% peptone, 1% yeast extract, 5% glucose, and 0.5 M citrate buffer (pH 5.0). Inoculum was incubated in a shaking incubator at 38°C and 130 rpm for 10-14 h. During exponential growth, cells were harvested, centrifuged, resuspended in sterile DI water, and inoculated into the reactors to achieve an optical density equivalent to 0.5 absorbance units at 600 nm.

3.3.3 Sampling and Data Analysis

Shake flask samples were taken approximately every 24 hours or when the slurry has liquefied enough to allow pipette sampling. Samples were removed with 5 mL pipettes (Falcon, Inc.) that have their tips broken off in order to more effectively accommodate the high level of insoluble solids. These were transferred to 2.5 mL microcentrifuge tubes and the pH was measured. The samples were next centrifuged at 10000 rpm for 5 minutes. The supernatant was decanted and diluted 1:5 and syringe filtered into HPLC vials for subsequent HPLC analysis. HPLC analysis was performed on all saccharification samples to determine the glucose, xylose, and cellobiose concentrations. An Agilent 1100 series HPLC equipped with a differential refractive index detector equipped with an Aminex HPX-87H organic acid column operating at 55°C with a 0.01 N H_2SO_4 mobile phase at a flow rate of 0.6 mL/min. Component concentrations obtained from HPLC measurement were converted to g/kg concentrations based on the estimated insoluble solids level in the reactor using the method outlined in the **Results and Discussion** (Section 3.4). Cellulose conversions were based on cellulose converted to glucose and cellobiose, while for SSF conversions were based on cellulose converted to glucose, cellobiose, and ethanol.

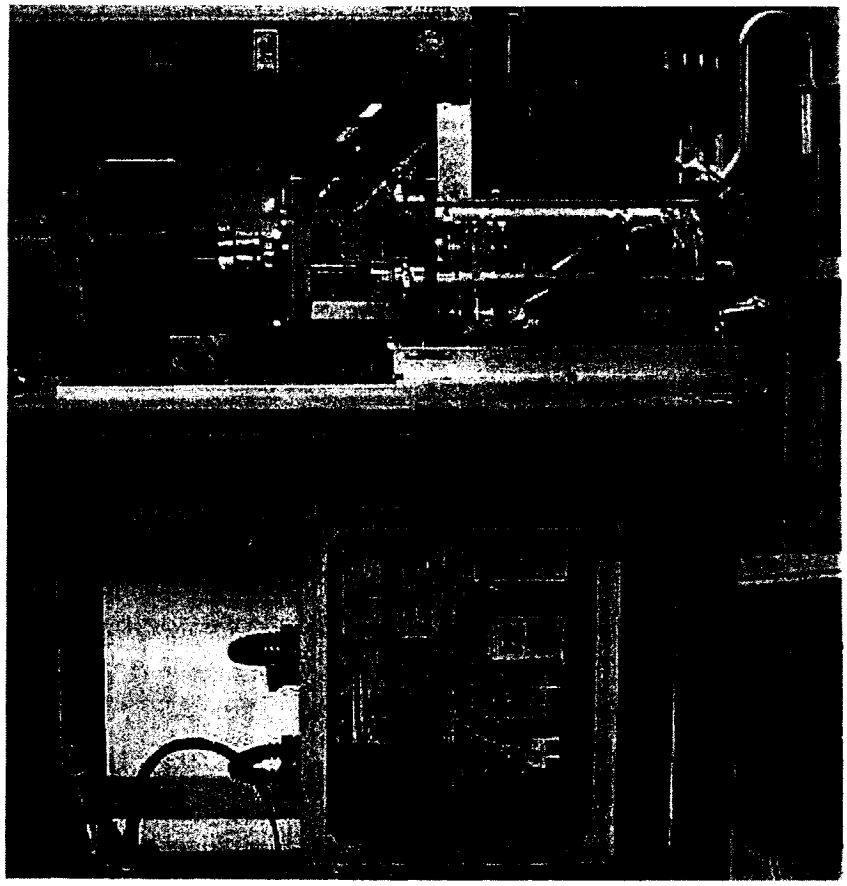


Figure 3.1: HSBR and control instrumentation

3.3.4 High Solids Bioreactor Operation

The high solids bioreactor (HSBR) was designed and built by Jody Farmer at NREL (Figure 3.1). It consists of a 10 L horizontally mounted reactor vessel with 8 blades that continuously clear the walls of the reactor to enable mixing in very high solids slurries. Temperature control was achieved by an external resistance-heating pad, which was prevented from exceeding 52°C to prevent denaturation of enzymes in contact with the inner wall. The entire reactor was insulated to minimize heat loss. The HSBR was autoclaved empty and allowed to cool in a hood. The appropriate amount of solids were measured and autoclaved and the amount of water lost was determined. Solids were added to the reactor and all liquid components and enzyme were mixed before addition to the reactor. Samples for analysis of pH, sugar by HPLC, and insoluble solids content were taken every 24 hours for 10 days.

3.4 Results and Discussion

3.4.1 Cellulose Rates to Characterize Substrate Morphology

Using the specific cellulose utilization variable developed from Equation 3.2, the effects of other variables (independent of concentration) on the rate can be isolated with more confidence. Figure 3.2 shows for PCS batch P020502CS how rate can be strongly correlated to the extent of cellulose conversion, with other enzyme parameters (T and enzyme loading) held constant. The experiments were performed at 10% initial insoluble solids. These plots show that the rates exhibit three distinct phases: initial rate, mid-reaction rate, and final rate. These are proposed to correspond respectively to the utilization of easily accessible cellulose, overall bulk cellulose, and highly recalcitrant cellulose.

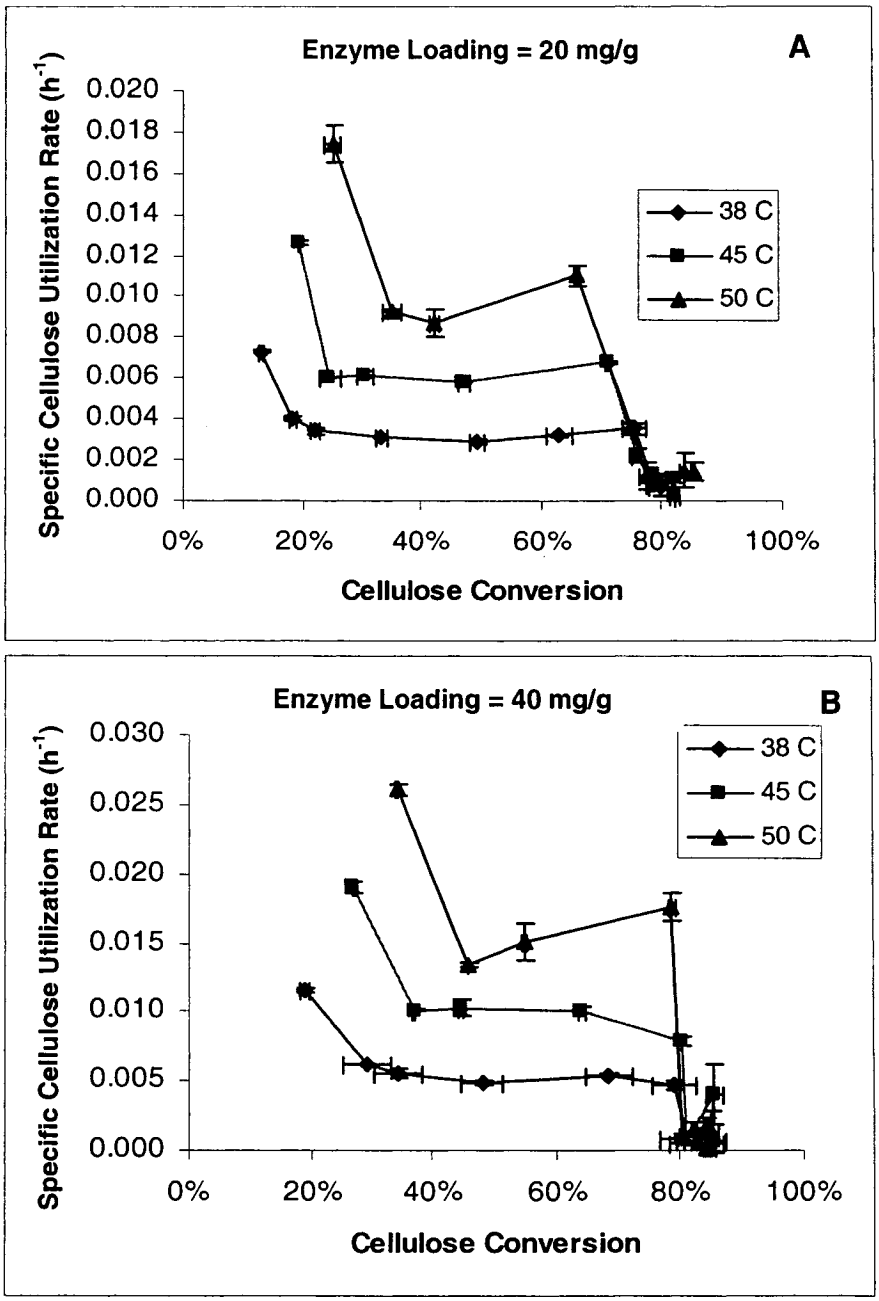


Figure 3.2: Effects of enzyme loading and temperature and conversion on the specific cellulose utilization rate

Many studies have shown that the initial rate of hydrolysis is much higher than the subsequent rate (Zhang and Lynd, 2004), while proposed reasons include selective initial hydrolysis of amorphous cellulose (Davis *et al.*, 2002; Mansfield *et al.*, 1999), decreases in specific enzyme adsorption or subsequent inability of bound cellulases to reach new catalytic sites (Lynd *et al.*, 2002; Eriksson *et al.*, 2002), and steric preferences (Väljamäe *et al.*, 1998; Ooshima *et al.*, 1990; Converse *et al.*, 1990). The specific rate at which the majority of the cellulose is utilized remains approximately constant until a critical value is reached at approximately 80% cellulose conversion. This finding that the rate remains relatively constant calls into question the applicability of models that assume an exponential decrease in the cellulose rate as a function of conversion (Ohmine *et al.*, 1983). This result also indicates that at the glucose levels in this system (maximum of ~50 g/L) product inhibition and enzyme inactivation are not significant considerations. This is in agreement with the work of Zhang *et al.* (1999), which argued that of the three postulated reasons for the decrease in cellulose rate as a function of conversion (enzyme inactivation, product inhibition, substrate heterogeneity) that substrate heterogeneity was the most likely source within the conditions studied. The next phase of cellulose utilization after approximately 80% cellulose conversion was marked by a significant decrease in the specific hydrolysis rate. Due to the abruptness of the transition to this low rate regime and its universality over multiple enzyme loadings and temperatures, it can be postulated that its source is due to morphological factors rendering accessibility to the remaining 20% of the cellulose much more difficult. This increased recalcitrance is apparent in many enzymatic hydrolysis experiments with explanations offered being

either that cellulose is inaccessible to enzymatic attack due to reprecipitated lignin on the surface (Nagle *et al.*, 2002) or that steric hindrance from residual lignin helps to maintain a compact and thus presumably more recalcitrant cellulose structure (Mooney *et al.*, 1998). It should also be considered that using cellulose rate as the characteristic rate causes the transition to the final rate to be much more abrupt than if the glucose rate were used. The reason for this is that cellobiose accumulates during the initial and middle phases, while during the final phase, this cellobiose is converted to glucose due to the much slower conversion of cellulose to cellobiose. This, in effect, allows the β-glucosidases to "catch up" to the cellulases.

3.4.2 High-Solids Saccharification

During closed system saccharification (CSS) in shake flasks, the solids level is one of the most important variables affecting the rate and extent of conversion. Figures 3.3A and 3.3B show the effects of PCS solids loading on rates and extent of conversion for increasing solids levels with the enzyme loading controlled. From the data, which are depicted in Figure 3.3C for the 40 mg/g enzyme loading, it is apparent that increasing the solids loading above 27% significantly decreases the rate of cellulose hydrolysis. The most likely reasons for this decrease in rate are a combination of glucose inhibition of the enzyme system from the correspondingly higher sugar levels reached using higher solids and mass transfer limitations. Figure 3.3C shows that for all cases presented, glucose concentrations in the liquid phase of greater than 140 g/L are achievable, and for the case of 30% solids, approximately 160 g/L of glucose is attained. However, higher solids loadings require significantly longer residence times to achieve these high liquid phase sugar levels. As discussed in the **Background**, the effect of glucose on β-glucosidase

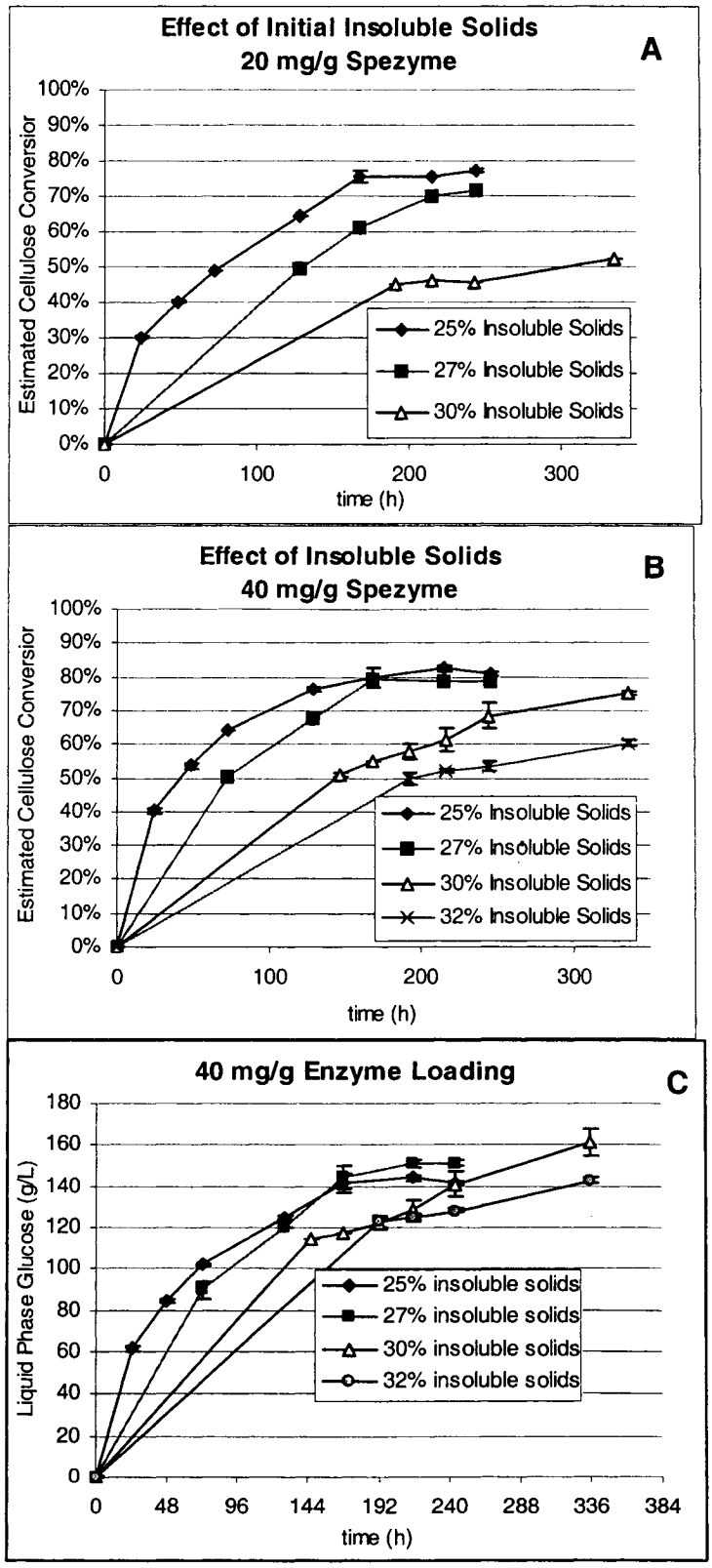
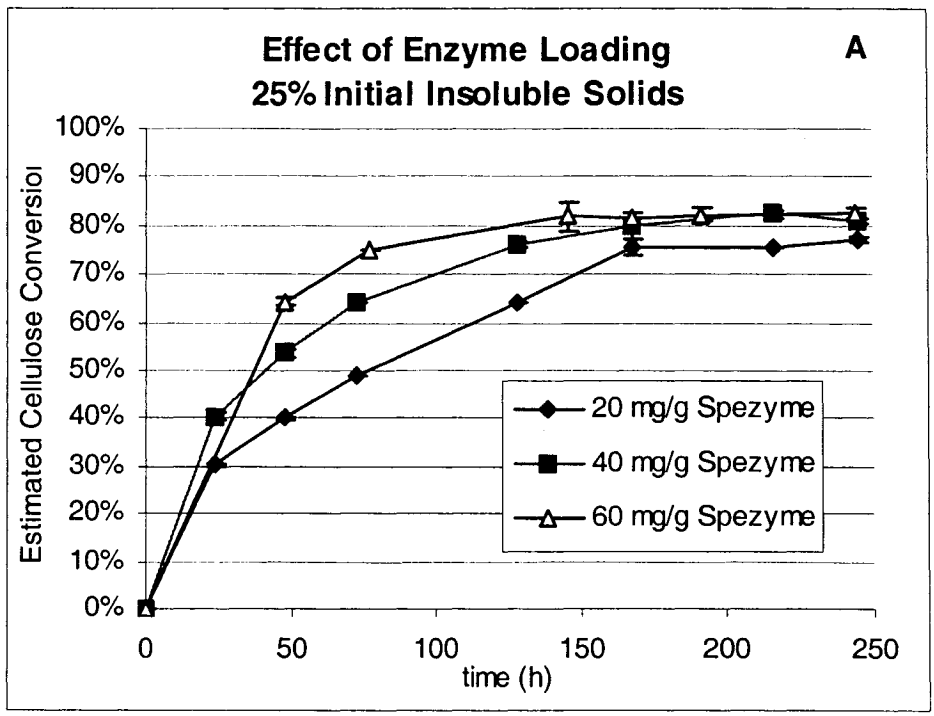


Figure 3.3: Effect of initial insoluble solids and enzyme loading on CSS; T = 45°C

activity is the most important inhibition concern, and this can become the ultimate rate-limiting step when glucose accumulates to very high levels. At 32% solids, the glucose concentration is lower than at 30% solids, indicating that mixing limitations for this level of solids become a significant factor in addition to glucose inhibition. Most importantly, this work also demonstrates that cellulose conversions greater than 80% can be achieved at initial insoluble solids levels as high as 27%. Figures 3.4 and 3.5 depict the effect of enzyme loading on cellulose utilization. The data clearly show that increasing the enzyme loading effects both the rate of hydrolysis and the extent of conversion, although the maximum conversion (digestibility) is apparently fixed at slightly greater than 80%, which was observed for pretreatment batch P020502CS as well. The data from Figure 3.4 is replotted in Figure 3.5 to demonstrate the effects of enzyme loading on cellulose conversion at 168 h and 245 h respectively. This shows that the enzyme is apparently limiting at a loading of 20 mg/g, while 40 mg/g and 60 mg/g do not show significant differences at the two lowest solids levels used.

Cellulose conversion was defined as cellulose converted to glucose and cellobiose (the summative approach of Chung *et al.*, 1997), since one item that should be considered is the difficulty in quantifying soluble β-1,4 glucose oligomers (cellodextrins) by HPLC. Cellodextrins are generally water-soluble for a degree of polymerization < 12 (Zhang and Lynd, 2004), and Klesov (1982) considers quantification of soluble glucan polymers to be necessary in determining mass balances and cellulose degradation kinetics. Using the Aminex 87-H carbohydrate analytical column, cellobiose, cellotriose, and other soluble cellodextrins elute at approximately the same time, rendering quantification more problematic. Due to this, parameters such cellulose conversion, which use cellobiose



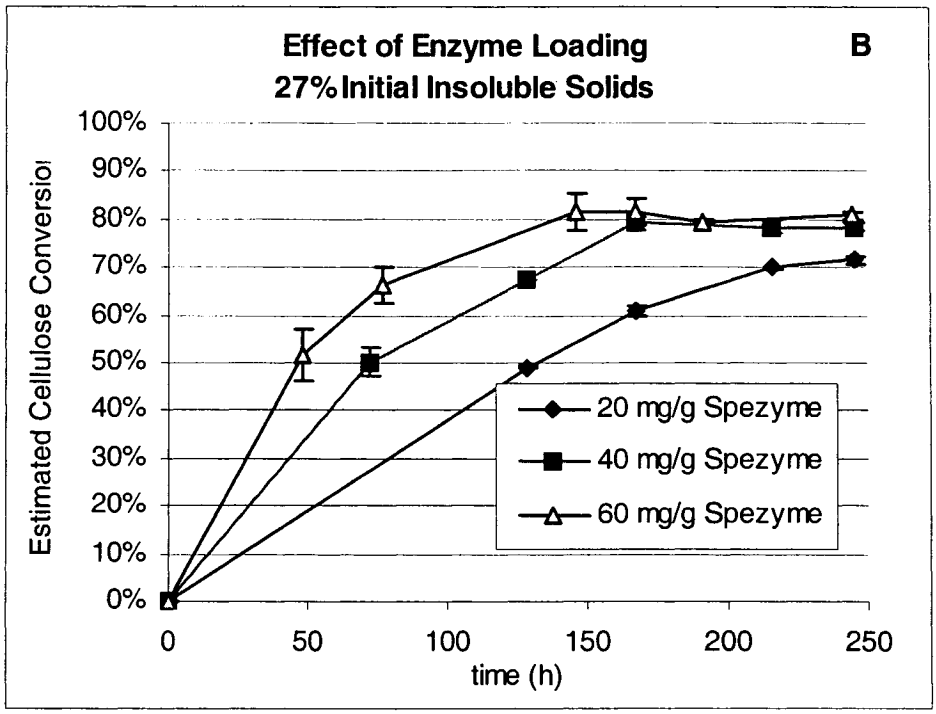


Figure 3.4: Effect of enzyme loading on rate and extent of cellulose utilization for shake flask CSS; T = 45°C

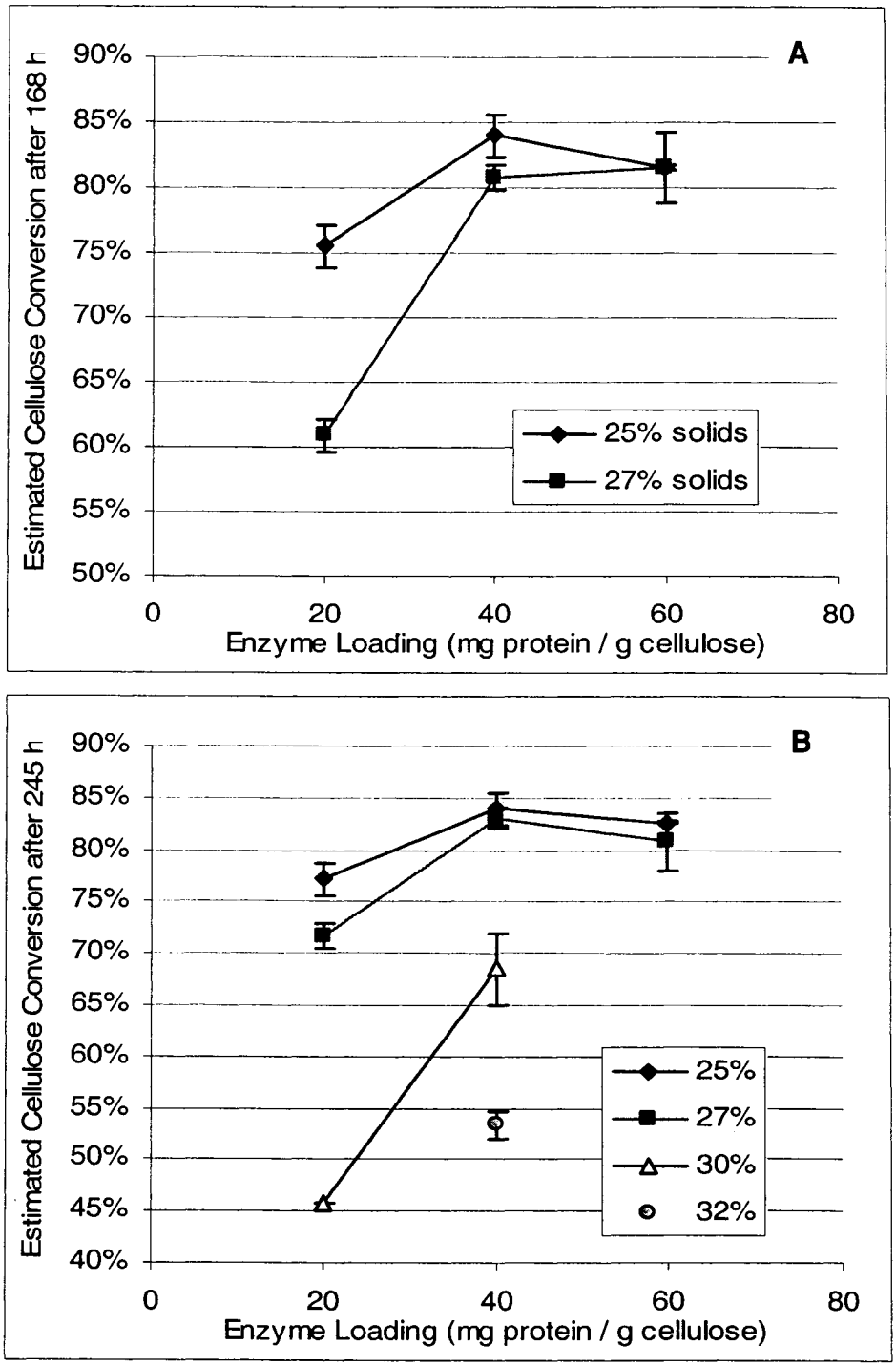


Figure 3.5: Effect of enzyme loading on cellulose utilization for shake flask CSS; T = 45°C

values, should be held somewhat suspect. This problem is further confounded during high-solids saccharifications where inhibition of β-glucosidase by glucose is most notable, and results in the accumulation of oligomers in addition to cellobiose. Additionally, if cellulose conversion were defined as conversion of cellulose to all water-soluble cellodextrins, then higher conversions might be expected.

3.4.3 Effects of Temperature and Sugar on High-Solids Rate

Temperature is another important parameter that strongly affects reaction rate as well as enzyme stability as discussed in the **Background** (Section 3.2). For SSF, it is necessary to compromise with a temperature that is tolerable for microbial fermentation but still provides a reasonable hydrolysis rate. Figure 3.6 shows the effect of temperature on CSS for 25% initial insoluble solids. By comparing the conversion profiles for the 45°C and 50°C experiments it is clear that long-term exposure to a temperature of 50°C does not result in inactivation of enzymes as has been previously suspected. For CSS reactors, operation at 38°C (the temperature that *S. cerevisiae* SSF is typically performed) decreases the rate of hydrolysis by nearly 50% relative to 45°C and 50°C.

SSF shake flask reactors were also run to determine whether SSF would result in improved performance (rate and extent of saccharification) relative to CSS at the same high levels of initial insoluble solids due to the removal of sugars by a fermentative microorganism, since these sugars are inhibitory to cellulases. Figure 3.7 demonstrates that regardless of reaction mode (CSS or SSF) or enzyme loading, at 25% initial insoluble solids the ultimate conversion is apparently constrained to around 80% using this combination of enzyme and microbe (yeast). Figure 3.7B further shows that at a 40 mg/g enzyme loading SSF provides no significant improvement in either rate or extent of

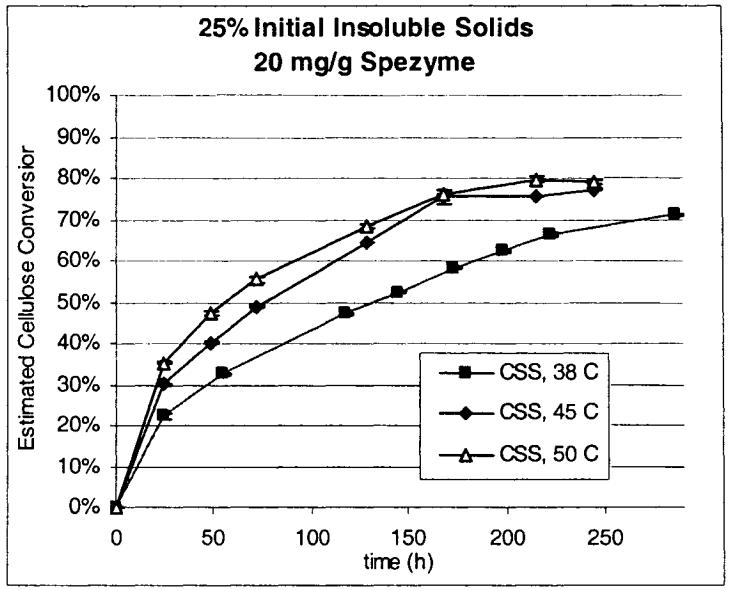


Figure 3.6: Effect of temperature on high-solids CSS conversion

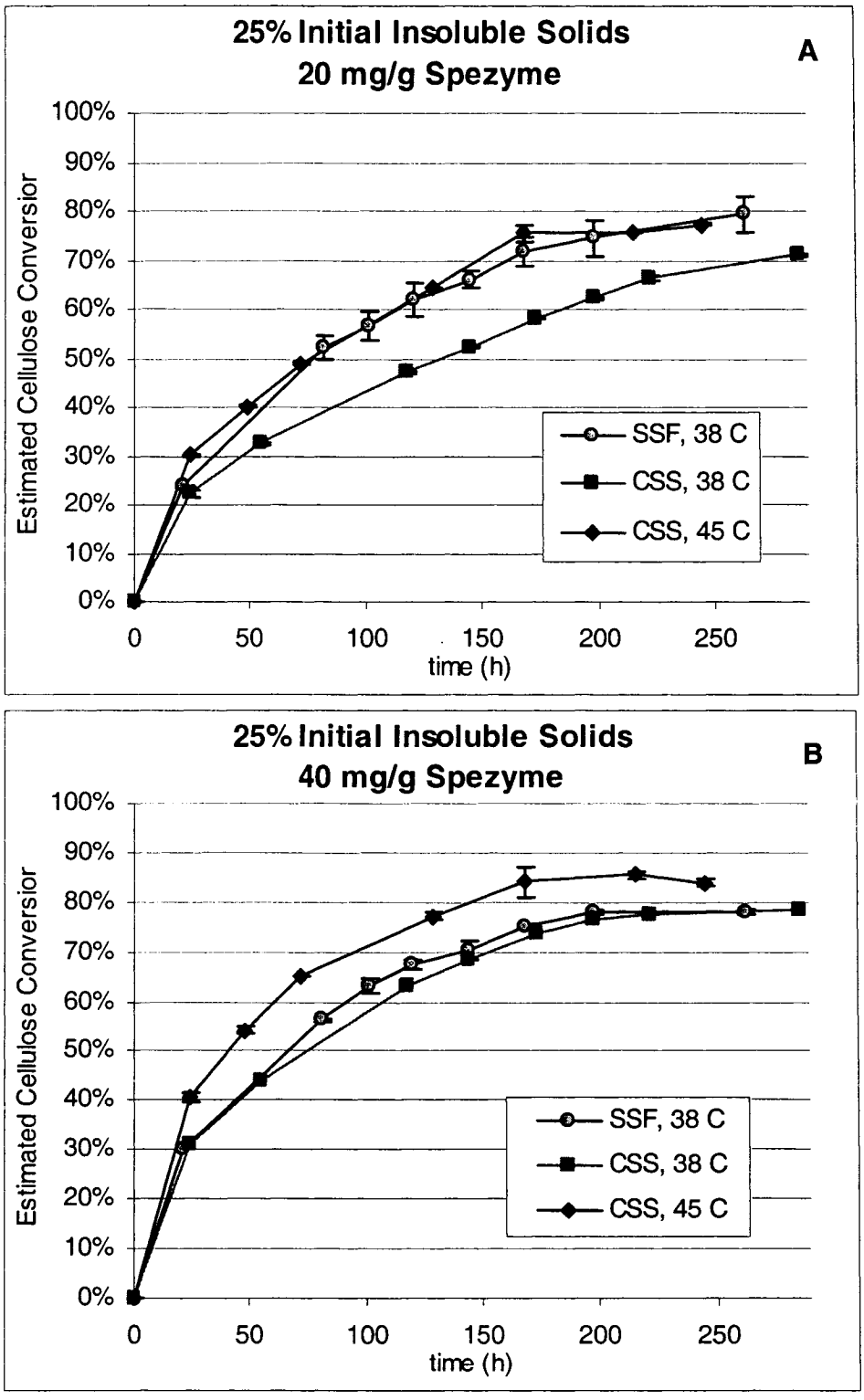


Figure 3.7: Effects of temperature and reaction mode on high-solids cellulose conversion

conversion over CSS. At an enzyme loading of 20 mg/g however, SSF does show an improvement. One explanation for this is that enzyme is limiting at this loading and is more sensitive to sugar inhibition.

An important consideration for the SSF results is that at 38°C the microorganism failed to ferment glucose to ethanol at ethanol concentrations greater than ~45 g/L and reasonable pH (~5.0) after 4 days. After the cells stopped fermenting, glucose accumulated to ~35 g/L for the 25% reactors. Previous high-solids SSF work (Mohagheghi *et al.*, 1992) was able to achieve ethanol concentrations as high as 57 g/L using *S. cerevisiae*, although at initial insoluble solids >20%, the cells lost their viability after 4 days. Postulated reasons for this loss of culture viability in the SSF experiments include a combination of stress factors such as high temperature, slow glucose release, high ethanol concentration, and osmotic stress on the yeast cells due to insufficient liquid. Controls were performed to determine if yeast cell cultures remained viable at 38°C and ethanol concentration >45 g/L in liquid media. The results are given in Table 3.1 and show that final ethanol concentrations reached as high as 58.6 g/L in the fermentation flasks, indicating that stress factors such as high ethanol combined with temperature by themselves are insufficient to cause loss of culture viability.

Another postulated reason for loss of cell viability could be that mass transfer limitations in shake flasks prevent ethanol from dispersing away from the cells. This could explain differences between pure sugar fermentations and high-solids SSFs (this hypothesis is validated by the data in Figure 3.10A, to be discussed in the next section, which shows that mass transfer limitations dominate at 72 h for initial insoluble solids levels greater than 20%). Hybrid hydrolysis and fermentation or HHF (defined as non-

isothermal SSF with a saccharification time longer than the fermentation time) performed by Varga *et al.* (2004) attribute low ethanol yields at 20% insoluble solids (wet oxidation pretreated corn stover) to mass transfer limitations. Alternatively, cells could die from prolonged exposure to osmotic stress (Melvin Tucker, personal communication, 2005).

Table 3.1: Fermentation Control Experiments for *S. cerevisiae*

	Low glucose	High glucose
Initial glucose (g/L)	50.8	136.5
Final ethanol concentration	23.9	58.6
Final glucose utilization	100%	100%
Product Yield, $Y_{P/S}$ (g EtOH/g glucose)	0.471	0.430
Percentage of theoretical product yield	92.1%	84.1%

Experimental ethanol yields were determined from the fermentation control data in Table 3.1, and a value of 92% was used since this is the average value for lower ethanol concentrations. If an 84% experimental yield were used, cellulose conversions greater than 100% would be attained for some cases; therefore 92% can be considered a conservative estimate. This yield may be lower due to either an increased amount of energy diverted to cell maintenance under these more stressful conditions or increased ethanol evaporation due to higher ethanol concentrations. Using this value could introduce error into the SSF vs. CSS comparisons in Figure 3.7.

In order to address other parameters affecting the rate of cellulose hydrolysis, the effects of substrate morphology need to be removed. This is done by only selecting data from less than 80% conversion and not using initial rate data. Using data obtained using

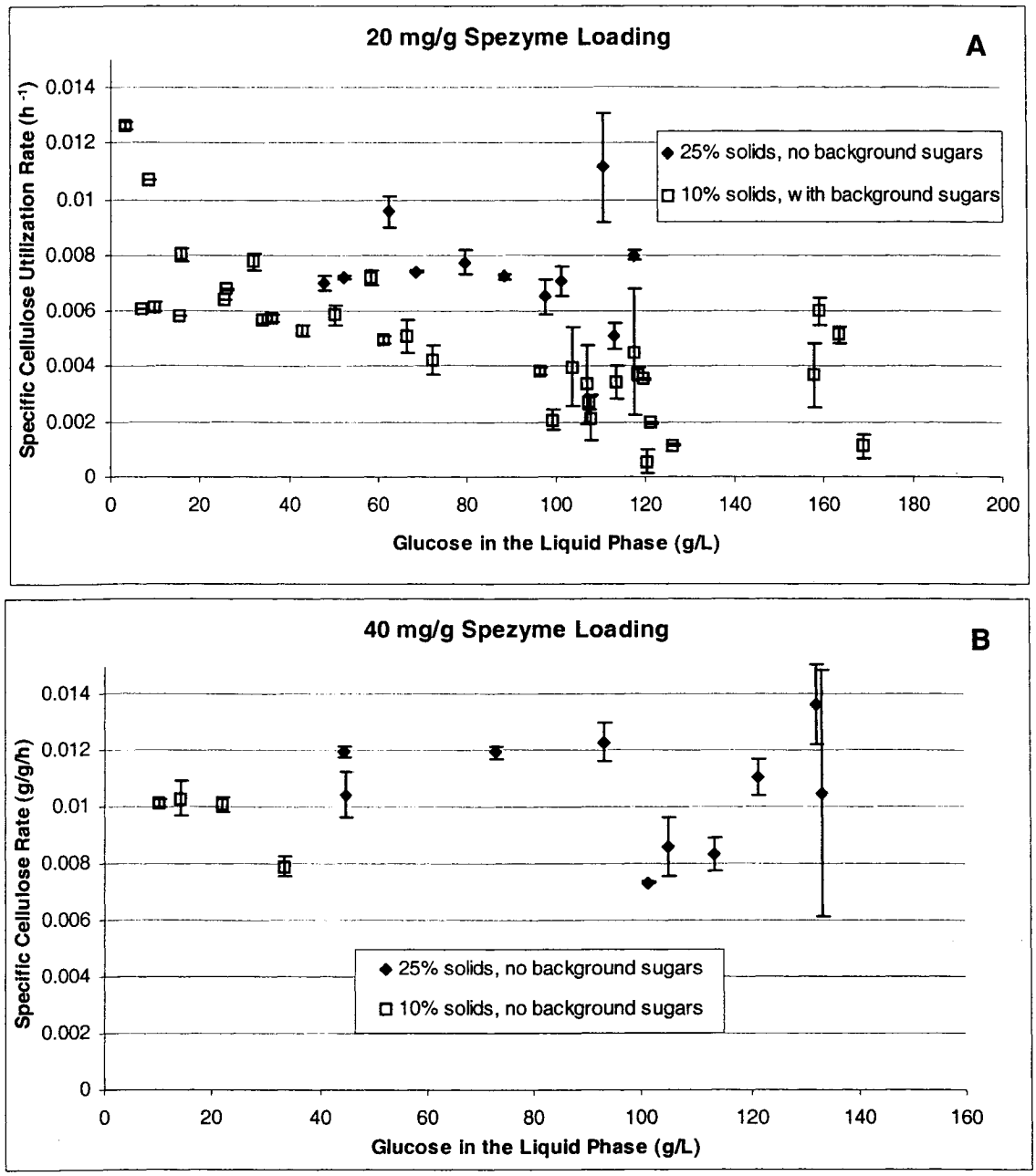


Figure 3.8: Effects of liquid phase glucose on the specific rate of cellulose hydrolysis

several initial solids loadings and initial background glucose concentrations, Figures 3.8A and B were developed to attempt to isolate the effects of glucose inhibition on the specific cellulose utilization rate. For the high-solids shake flask data, the effects of mass transfer limitations were apparently minimal as will be discussed subsequently. The primary inhibitory effect of glucose is its effect on β-glucosidases, which leads to the accumulation of cellobiose and the inhibition of cellulases. These effects are not considered separately, but rather can be combined since during a batch process it is possible for all of these components to accumulate to inhibitory levels. From Figure 3.8A and B it is clear that the glucose inhibition for this enzyme preparation is not a major factor affecting the rate of hydrolysis as might be supposed when using high levels of initial insoluble solids. The values for the specific cellulose utilization rates are comparable to those from Figure 3.7: approximately 0.1 h^{-1} and 0.075 h^{-1} for the enzyme loadings of 40 mg/g and 20 mg/g respectively. Decreases in these rates are apparent at glucose concentrations greater than 120 g/L for the 20 mg/g enzyme loading with background glucose. The results from the high-solids rates indicate that sugar inhibition is not an overriding concern when using initial substrate concentrations this high, which is encouraging for a high-solids saccharification process.

3.4.4 Mass Transfer During High-Solids Enzymatic Hydrolysis

Shake flask reactors provide a good system to attempt to quantify the effects of diffusive mass transfer, since initially, convective mass transfer is minimal above ~20% insoluble solids. One mass transfer derived issue that was encountered in shake flask batch saccharification experiments was based on the differing amount of drying for the PCS solids prior to the experiment. During the reactor preparation, autoclaved PCS

solids are mixing with all liquid phase components: enzymes, water, buffer. Lower moisture content solids require more makeup liquid. At especially high initial reactor insoluble solids levels, this makeup liquid can be very small relative to the total reactor volume. When much more makeup liquid is added to these high-solids systems, there is an opportunity for the liquid phase components to swell the cellulose fibers and penetrate dry pores. This can result in much better initial buffer and enzyme distribution and adsorption onto the substrate. For this work solids were dried using a combination of 105°C oven and air drying. Care was taken to ensure that the exposed solids on the surface of the pan were never completely dried in the oven by frequent mixing. Figure 3.9 gives an indication of the effects of partial substrate drying prior to addition to the reactor for saccharifications at 25% initial insoluble solids. This shows that for the lower enzyme loading (Figure 3.9A) where the enzyme is more rate-limiting, the effects of prior drying are especially pronounced. At the higher enzyme loading (Figure 3.9B) there is still some effect of drying, and it is apparent that drying solids to >40% will not likely improve the rate. The very clear implications are that drying can affect the initial distribution and adsorption of enzymes, which affects the rate of enzymatic saccharification. For this reason, PCS for all high-solids reactions used in this work are dried to 40-45%. Previous work has shown that both air- and oven-drying of PCS solids can result in lower enzymatic digestibility, presumably due to irreversible pore collapse (Mildred Zuccarello, personal communication, 2005). These studies, however, totally dried the solids; while the current work has not been able demonstrate noticeably lower final conversions based on prior partial substrate drying.

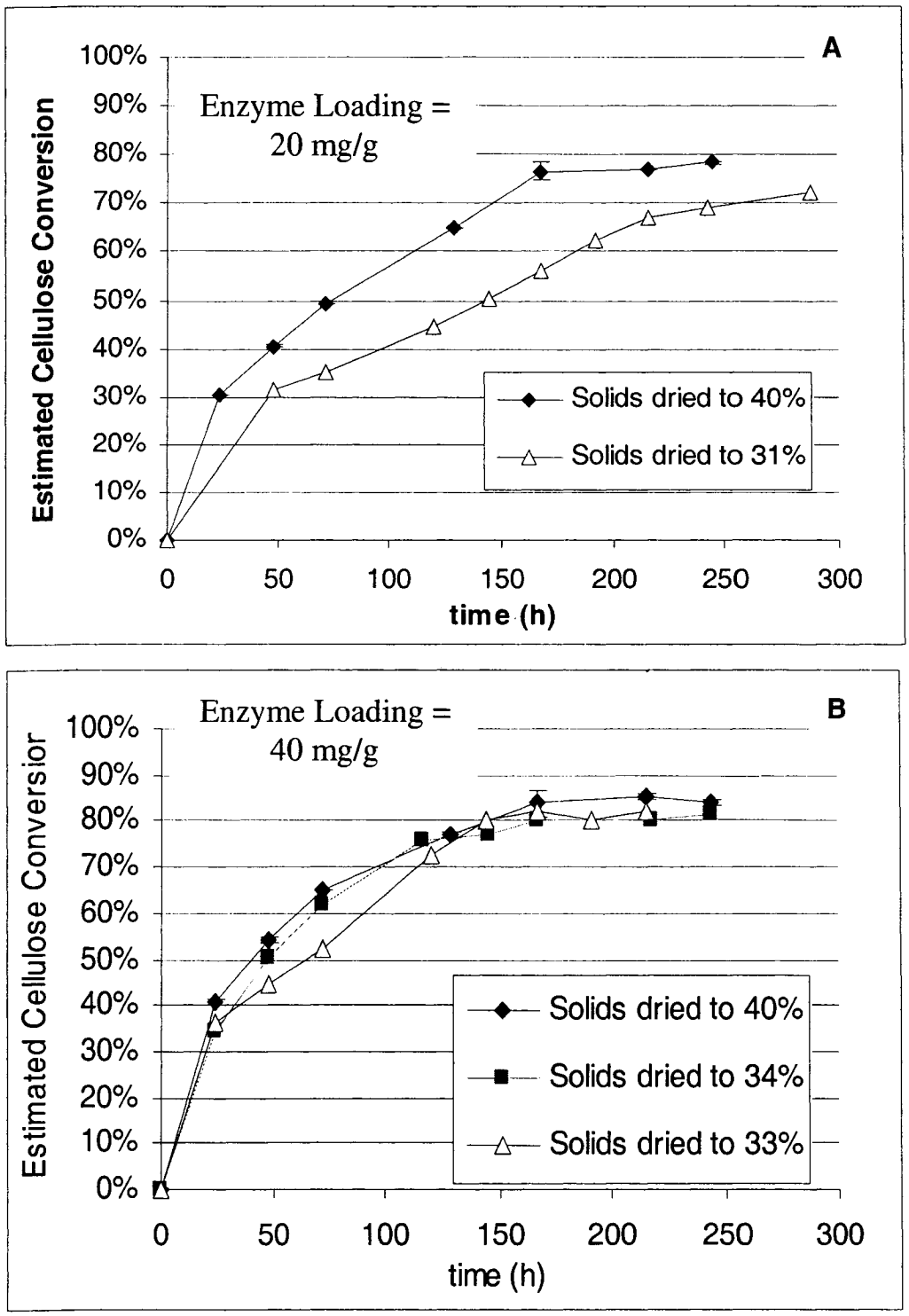


Figure 3.9: Effect of substrate drying on the rate and extent of saccharification in shake flasks for 25% initial insoluble solids.

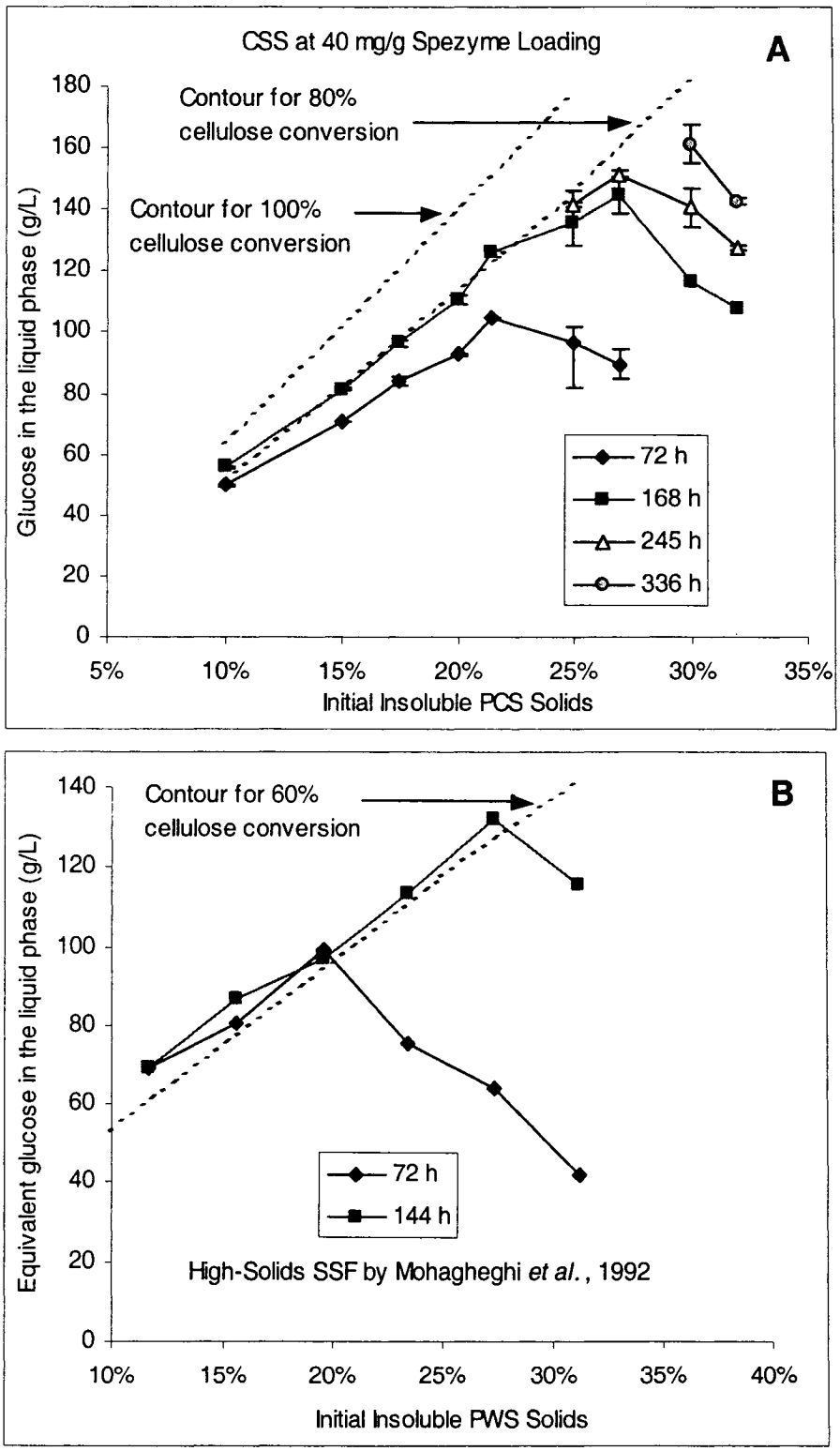


Figure 3.10: Reactant vs. product plot showing the effects of mass transfer rate limitations at high initial insoluble solids.

For the enzymatic hydrolysis of cellulose, insoluble solids level is clearly linked to mass transfer considerations. One plot that clearly demonstrates the extent of mass transfer limitations in the shake flask studies used in this work is Figure 3.10A. This compares the level of initial insoluble solids to the glucose level at constant time contours, and can be considered as basically a plot of product versus reactant with the advantage that both glucose time dependence (productivity) and initial insoluble solids can be considered simultaneously. Asymmetric error bars indicate the range of the measured data. At lower solids levels for all time contours the correlation is nearly linear as would be expected for an essentially first order reaction system. One interesting result from this data is that the glucose concentration decreases significantly after a clear critical value is reached for each time contour. A slight decrease is to be anticipated due to the reduction in the rate of hydrolysis by sugar inhibition anticipated at higher solids loadings. The results, however, suggest that there are factors other than sugar inhibition acting on the system, since sugar inhibition should be indicated by a leveling off near an inhibitory sugar concentration value rather than a decrease as solids level increase. Mass transfer limitations are seen as the most likely reason for this result.

The reactions were performed in shake flasks, where at the higher solids loadings there is little convective mixing until the slurry is significantly saccharifiied. Since there is little free water, which is needed to facilitate the mass transfer of water-soluble compounds, it is possible that the reaction rate is diffusion limited, such that sugars can build up around active cellulases and inhibit saccharification, while enzymes are prevented from reaching new catalytic sites.

Data taken from Mohagheghi *et al.* (1992) are replotted in a similar manner to Figure 10A in Figure 3.10B. Since these experiments were actually performed in SSF mode, the "Glucose in the Liquid Phase" y-axis actually represents the hypothetically accumulated the equivalent glucose concentration back-calculated from the actually obtained ethanol concentration. In this plot the same phenomenon is observed at approximately the same level of insoluble solids, with the basically first order rate correlation marked by conspicuous drops in glucose levels at essentially the same solids level as observed in the current work. The lower conversion is due to the differences in substrate and enzyme. The reaction system used a different pretreated substrate, enzyme, and reactor compared to the work presented in Figure 3.10A, while still demonstrating similar trends. This is significant since the abrupt drop-off in glucose cannot be attributed to sugar inhibition since the reaction was performed in SSF, which strongly indicates that there is a more universal phenomenon present such as inhibition of mass transfer by increasing insoluble solids levels.

While SSF experiments show that sugar inhibition under these conditions is not a compelling concern, another way to rule out or isolate the effects of sugar inhibition can be to operate the same reactor at a lower solids level with background sugars. This was performed using an initial background glucose concentration of 100 g/L at an initial insoluble solids level of 10%, which is far below the threshold for mass transfer rate limitation. Based on the cellulose conversion at 168 h, an "equivalent initial insoluble solids" can be determined by calculating the initial insoluble solids required to generate this conversion level assuming that all of the cumulative glucose in the liquid phase was generated from cellulose. Figure 3.11 plots this experiment as well as an experiment run

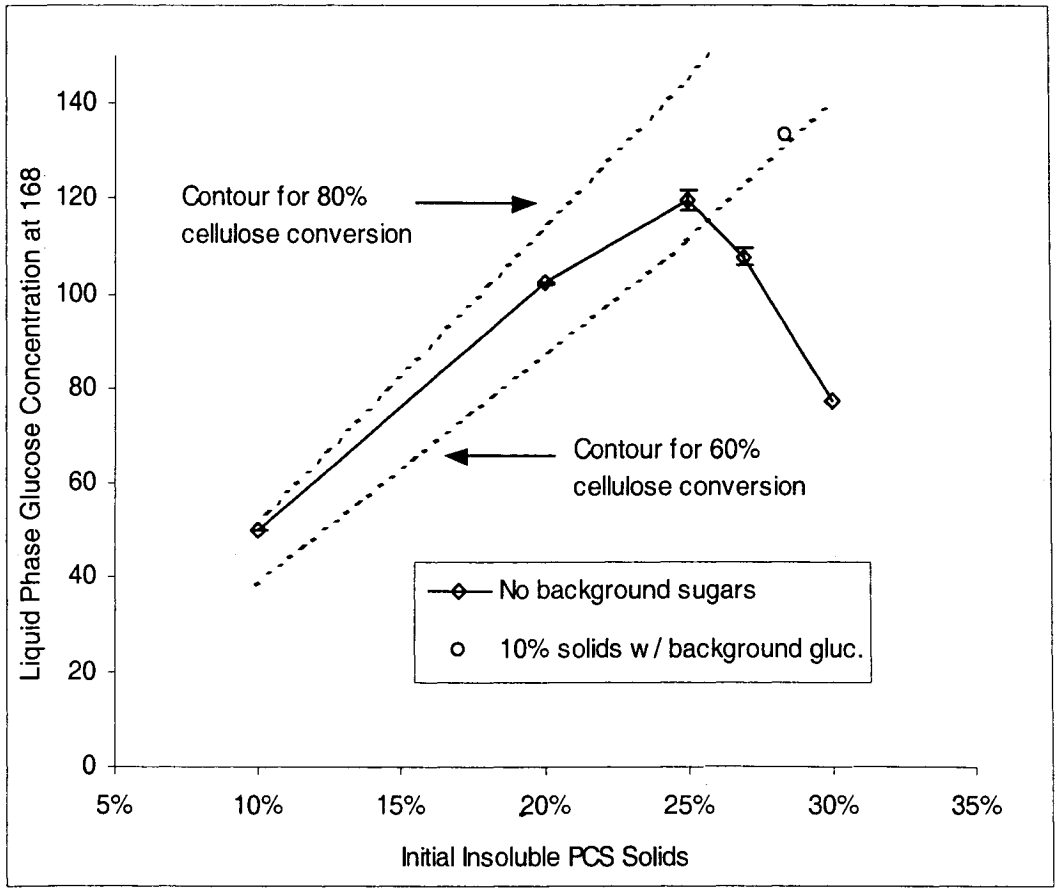


Figure 3.11: Characterization of sugar inhibition using background glucose.

with no background sugars at the same enzyme loading (20 mg/g). These data approximately follow what would be expected for first order hydrolysis kinetics in the absence of mass transfer limitations. Because of this, the effect of sugar inhibition as the source of this drop-off in the reaction rate due to insoluble solids level can be ruled out with more certainty.

Quantitatively addressing these mass transfer limitations for this system in a theoretical manner is challenging. Ideally, high-solids saccharification mass transfer models should consider the cellulase-catalyzed hydrolysis on the surface of solid phase PCS particles, enzyme and sugar diffusion through the porous spaces between and within PCS particles, convective mass transfer, and the liquid phase reaction of soluble cellodextrins to glucose. One method for simplifying all mass transfer and reaction rates is to use an effectiveness factor, which incorporates all of these effects. As described in Equation 3.8, ε can be defined as the ratio of the actual overall rate to the rate in the absence of mass transfer. In this reaction system, ε should be considered an overall effectiveness factor, encompassing both internal (within the PCS particle pores) and external (the interstitial space between PCS particles) mass transfer. Importantly, in considering the external mass transfer limitations, the effectiveness factor also accounts for limitations in the mass transfer due to an insufficient aqueous phase to allow for diffusion, which would be difficult to quantify using other methods. For the simplest case, which considers that all rates are inhibited to the same degree by mass transfer, ε can be determined in terms of glucose for a given insoluble solids level with Equation 3.8 as:

$$\left(\frac{dG}{dt}\right)_{Actual} = \varepsilon\left(\frac{dG}{dt}\right)_{Theo.} \tag{3.9}$$

with a numerical approximation of the derivative:

$$\left(\frac{\Delta G}{\Delta t}\right)_{Actual} = \varepsilon\left(\frac{\Delta G}{\Delta t}\right)_{Theo.} \tag{3.10}$$

When evaluated over the same time interval:

$$\varepsilon = \frac{\Delta G_{Actual}}{\Delta G_{Theo.}} \tag{3.11}$$

The value for $\Delta G_{Theo.}$ can be estimated from either model data or by extrapolation from the plots of Figure 3.10, where it is assumed that the ΔG at initial insoluble solids levels less than 20% is not limited by mass transfer. Additionally, as the Δt becomes much larger, the effects of mass transfer during the initial reaction when the mass transfer limitations are strongest begin to become averaged out over the entire reaction and some of the time-dependent information about ε is lost. However, since sampling a reactor at solids levels greater than 25% is not practical until the slurry is liquefied to a manageable level, this is a necessary shortcoming. Using linear interpolations based on Figure 3.10A, the theoretical $\Delta G_{Theo.}$ can be determined, and used to estimate values of ε. By plotting ε for different values of $S_{I,0}$ over different Δt ranges as in Figure 3.12, the effect of $S_{I,0}$ on ε can be evaluated. The clear trend from this plot is that while the y-intercept shifts at larger time intervals, the effectiveness factor is essentially linear with respect to $S_{I,0}$, with a constant slope, or:

$$\frac{d\varepsilon}{dS_{I,0}} = -5.79 \tag{3.12}$$

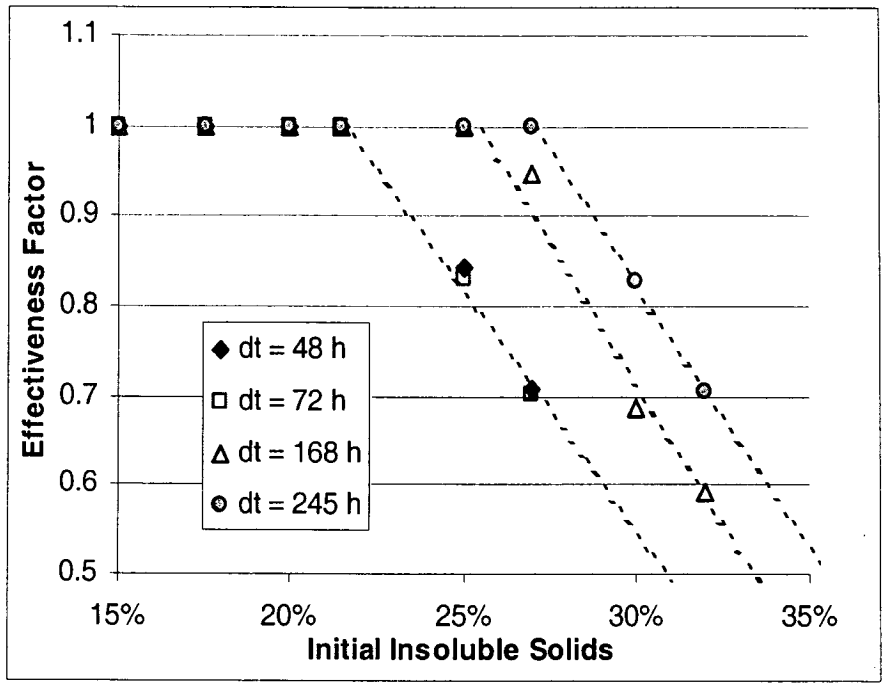


Figure 3.12: Effectiveness factor as a function of initial insoluble solids.

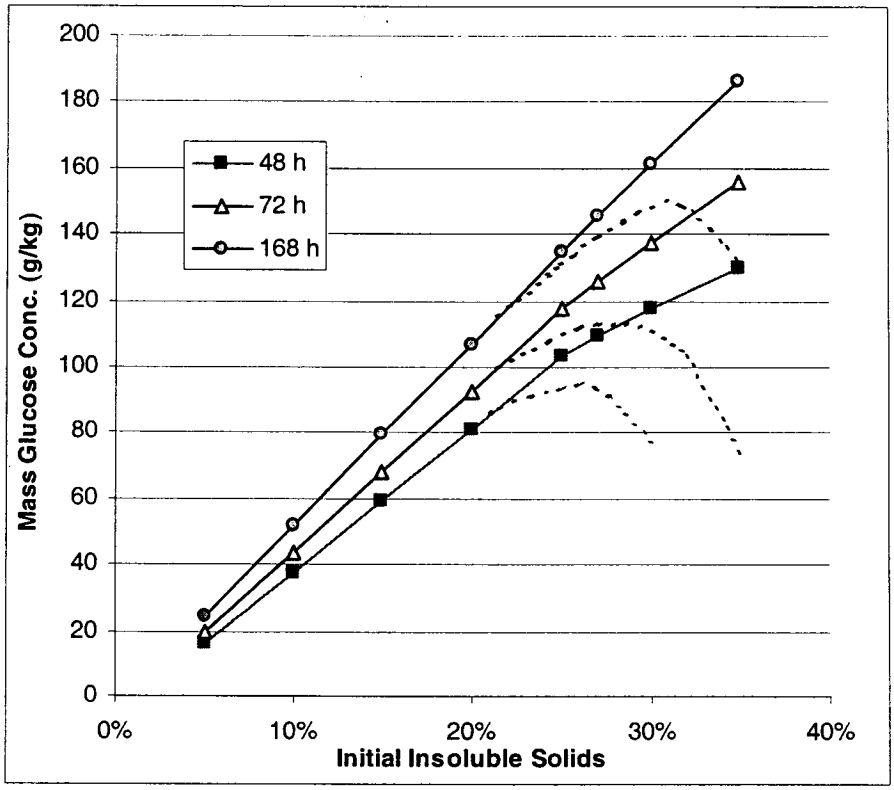


Figure 3.13: MATLAB simulations based on model of Kadam *et al.* (2004) using enzyme loading of 40 mg protein / g cellulose. All points are model simulations.

At the values of $\Delta t < 72$ h, the relation is approximately constant and can probably be considered as best representing the actual rates. This is reinforced by the data for $\Delta t = 48$ h and 72 h lying along the same linear contour. As the time increment is increased, the y-intercept information is lost, while the effect of insoluble solids on the effectiveness factor is preserved. The general correlation derived from this plot is that for $t < 72$ h:

$$\varepsilon(S_{I,0}) = \begin{cases} 1, & 0\% < S_{I,0} < 21.5\% \\ -5.79 S_{I,0} + 2.245, & 21.5\% \le S_{I,0} \le 38.8\% \\ 0, & 38.8\% \le S_{I,0} \end{cases} \quad (3.13)$$

Similar relations can be defined for other time intervals. While it would be tempting to try to define this relation universally in terms of the actual insoluble solids content rather than only the initial value, the available data limit the value of this since the rate of change in insoluble solids is a nonlinear function of itself.

Applying these correlations in a kinetic model can provide a validation for the assumptions used for developing the model. Using the kinetic model of Kadam *et al.* (2004), the g/kg glucose concentration is plotted in Figure 3.13 as a function of $S_{I,0}$ for constant time contours for both a constant ε (dashed lines) and no mass transfer limitation (solids lines). The figure shows that the plot for no mass transfer limitation results in slight inhibition by glucose at higher values of $S_{I,0}$, while the plots for constant ε show decreases similar to those found experimentally. When compared to experimental data (not shown) the results do not precisely match the experimental data since the model kinetics are slightly faster and don't perfectly match reality. The significance of this is that the approach of using an effectiveness factor is demonstrated to provide a useful approximation of the reaction system.

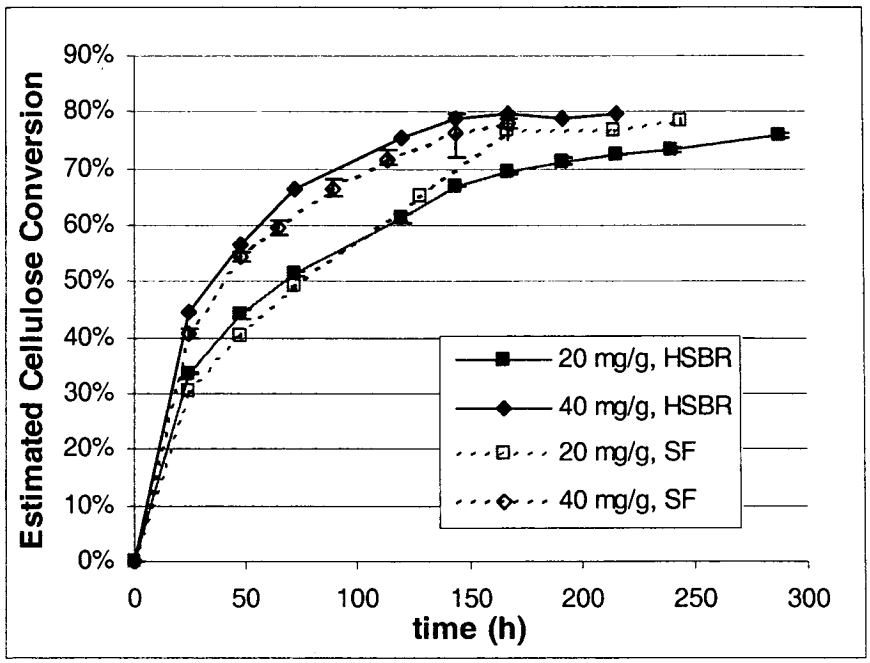


Figure 3.14: Comparison of enzyme loading effects in the HSBR; 7 rpm; 25% initial insoluble solids

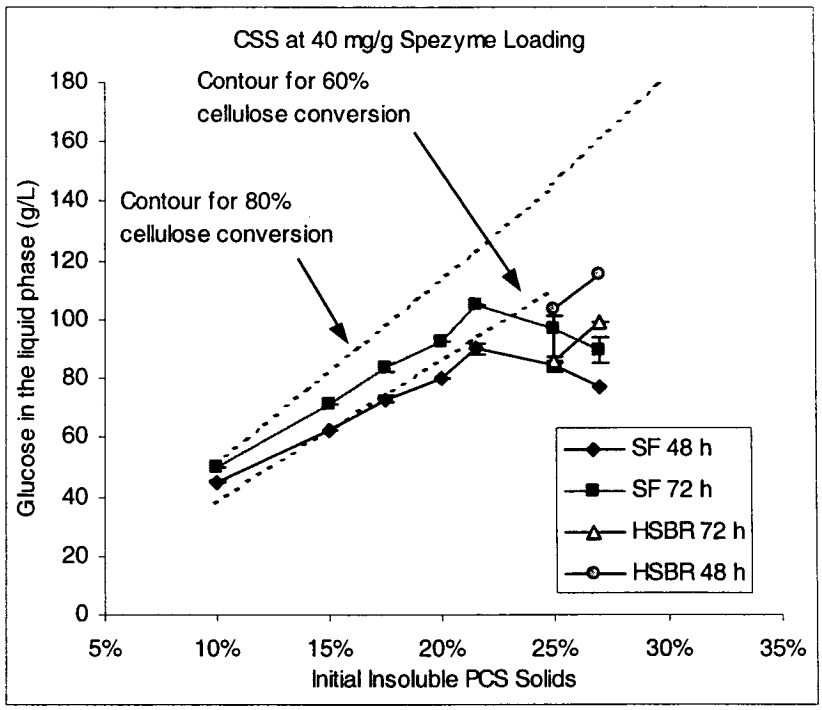


Figure 3.15: Increase in cellulose hydrolysis reaction rate due to improved convective mass transfer provided in the HSBR.

3.4.5 HSBR Scale-up Studies

A series of scale-up studies were performed in the HSBR to determine the effects of stirring at larger reactor scales on the rate of enzymatic hydrolysis of PCS for high-solids slurries. Since it was shown in the shake flask studies that the effectiveness of mass transfer could be correlated to $S_{I,0}$, it might be expected that mild agitation could improve the rate of (external convective) mass transfer. Because the reaction is performed in a larger scale reactor, there are variables other than stirring that may affect the comparison (*i.e.* temperature control). Figure 3.14 shows the results of the HSBR and shake flask saccharifications at 25% initial insoluble solids for two enzyme loadings. These plots show that the rate of cellulose hydrolysis in the HSBR is initially more rapid than for shake flasks. This can possibly be attributed to improved mixing and mass transfer provided by stirring. For the case of 20 mg/g enzyme loading, the final cellulose rate is more rapid in shake flasks than in the HSBR. The postulated reason for this is that for a more saccharified slurry in the HSBR, temperature control becomes more difficult due to the low agitation rate (7 rpm) and the actual temperature in the reactor is several ^{o}C lower than 45^{o}C. This lower temperature will result in a slower cellulose hydrolysis rate. Figure 3.15 depicts the same variables plotted in Figure 3.10 and includes the data for the HSBR. This shows that at both 25% and 27% initial insoluble solids, mixing in the HSBR apparently improves the rate of reaction, particularly at 27%. This is significant in that the scale-up of washed solids in batch mode is demonstrated both at high conversion yields and reaction rates.

3.5 Conclusions and Recommendations

This work demonstrated a number of important novel conclusions related to high-solids saccharification reaction systems. First, it was demonstrated that sugar inhibition of enzymatic saccharification rates is not as compelling a concern as had previously been suspected, and that remarkably high concentrations of glucose are achievable in high-solids enzymatic reactions. Besides sugar inhibition, a number of other substrate and enzyme-related parameters were identified that are important for high-solids enzymatic saccharification. Mass transfer limitations were also quantified in high-solids enzymatic saccharification using an effectiveness factor, and were able to demonstrate quantitatively the approximate upper physical limits for high solids levels in shake flask reactors. From these results it was determined that high glucose concentrations and high-conversion in high-solids saccharification are both physically achievable and scaleable.

While diffusion-limited mass transfer was shown to be a factor in the reaction rate in high-solids shake flask reactors, this work also demonstrated that agitation in larger scale vessels can improve the rate of reaction. It could be postulated that the rate of reaction is intrinsic to the level of insoluble solids limited by the lack of free water for diffusion/convection, which may be validated by the literature data (Mohagheghi *et al.*, 1992) that encountered this limiting mass transfer behavior using a different substrate, enzyme, and reactor system. However, the work here demonstrates that (even moderate) agitation can overcome some of the limitations presented by such a small discontinuous aqueous phase and improve the rate of reaction. This is most likely due to improved external convective mass transfer (a rate-limiting step) imparted by agitation.

Overall, the significance of this work is in demonstrating for the first time that high cellulose conversions are possible in reasonable times (less than 7 days) at very high initial insoluble solids levels and in the presence of high levels glucose. Examining these high glucose concentrations in terms of volumetric productivities can yield insightful comparisons to other reaction systems. In this work 7-day volumetric glucose productivities at 25% initial insoluble solids were found to be about 0.86 g/L/h (or 0.73g/kg/h for the entire slurry), which is nearly 3 times greater than what was obtained at 10% initial insoluble solids (0.33 g/L/h or 0.31 g/kg/L). This remarkable improvement in rate, in addition to the high product concentrations, has the potential to greatly improve the economics of enzymatic saccharification. Future studies should build upon this work under process-relevant conditions by characterizing the effects of pretreatment hydrolyzate liquors on high-solids enzymatic hydrolysis.

3.6 References

Aden A, Ruth M, Ibsen K, Jechura J, Neefes K, Sheehan J, Wallace B, Montague L, Slayton A, Lukas J (2002). Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolyiss and Enzymatic Hydrolyiss for Corn Stover. NREL Technical Paper 510-32438.

Asenjo JA (1984). Modelling the Bioconversion of Cellulose into Microbial Products: Rate Limitations. *Proc Biochem.* 19(6):217-224.

Bhat MK and Bhat S (1997). Cellulose Degrading Enzymes and their Potential Industrial Applications. *Biotechnol Adv.* 15(3/4):583-620.

Baker JO, Vinzant TB, Ehrman CI, Adney WS, Himmel ME (1997). Use of a New Membrane-Reactor Saccharification Assay to Evaluate the Performance of Cellulases Under Simulated SSF Conditions. *Appl Biochem Biotechnol.* 63-65:585-595.

Boer H and Koivula (2003). The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. *Eur. J. Biochem.* 270:841-848.

Borden JR, Lee YY, Yoon HH (2000). Simultaneous Saccarification and Fermentation of Celluloseic Biomass to Acetic Acid.

Brennan MA and Wyman CE (2004). Initial Evaluation of Simple Mass Transfer Models to Describe Hemicellulose Hydrolysis in Corn Stover. *Appl Biochem Biotechnol.* 113-116:965-976.

Calsavara LP, De Moraes FF, Zanin GM (1999). Modeling cellobiose hydrolysis with integrated kinetic models. *Appl Biochem Biotechnol.* 77-79:789-806.

Chang VS and Holtzapple MT (2000). Fundamental Factors Affecting Biomass Enzymatic Reactivity. *Appl Biochem Biotechnol.* 84-86:5-37.

Chung YC, Bakalinsky A, Penner MH (1997). Analysis of biomass cellulose in simultaneous saccharification and fermentation processes. *Appl Biochem Biotechnol.* 66(3):249-62.

Converse AO, Ooshima H, Burns DS (1990). Kinetics of Enzymatic Hydrolysis of Lignocelluloseic Materials Based on Surface Area of Cellulose Accessible to Enzyme and Enzyme Adsorption on Lignin and Cellulose. *Appl Biochem Biotechnol.* 24-25:67-73.

Davis M, Baker JO, Rignall T, Himmel ME (2002). Changes in Cellulose Morphology of Pretreated Yellow Poplar During Enzymatic Hydrolysis. NREL Report No. PO-510-32125.

Desai SG and Converse AO (1997). Substrate reactivity as a Function of the Extent of Reaction in the Enzymatic hydrolysis of Lignocellulose. *Biotechnol Bioeng.* 56(6):650-655.

Eriksson T, Karlsson J, Tjerneld F (2002). A Model Explaining Declining Rate in Hydrolyiss of Lignocellulose Substrates with Cellobiohydrolase I (Cel7A) and Endoglucanase I (Cel7B) of *Trichoderma reesei. Appl Biochem Biotechnol.* 101:41-59.

Excoffier G, Toussaint B, Vignon MR (1991). Saccharification of Steam-Exploded Poplar Wood. *Biotechnol Bioeng.* 38:1308-1317.

Folger HS (1999). *Elements of Chemical Reaction Engineering, 3rd ed.* Prentice Hall PTR, Upper Saddle River, NJ.

Gruno M, Väljamäe P, Pettersson G, Johansson G (2004). Inhibition of the *Trichoderma reesei* Cellulases by Cellobiose is Strongly Dependent on the Nature of the Substrate. *Biotechnol Bioeng.* 86(5):503-511.

Gusakov AV, Arkady PS (1992). A Theoretical Analysis of Cellulase Product Inhibition: Effect of Cellulase Binding Constant, Enzyme/Substrate Ratio, and β-Glucosidase Activity on the Inhibition Pattern. *Biotech Bioeng.* 40(6):665-671.

Gusakov AV, Sinitsyn AP, Klyosov (1985). Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch process. *Enz Microbiol Technol.* 7:346-352.

Holtzapple M, Cognata M, Shu Y, Hendrickson C (1990). Inhibition of *Trichoderma reesei* Cellulase by Sugars and Solvents. *Biotechnol Bioeng* 36:275-287.

Kadam KL, Rydholm EC, McMillan JD. (2004) Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnol Prog.* 20(3):698-705.

Klyosov AA (1990). Trends in Biochemistry and Enzymology of Cellulose Degradation. *Biochem.* 29(47):10577-10585.

Klesov AA (1982). Regulation of the formation and conversion of intermediate cellooligosaccharides and cellobiose during ezymatic hydrolysis of insoluble cellulose (Article in Russian) *Biokhimiia.* 47(4):608-18.

Kruus K, Andreacchi A, Wang WK, Wu JHD (1995). Product inhibition of the recombinant CelS, and exoglucanase component of the *Clostridium thermocellum* cellulosome. *Appl Microbiol Biotechnol* 44:399-404.

Ingesson H, Zacchi G, Yang B, Esteghlalian AR, Saddler JN (2001). The effect of shaking on the rate and extent of enzymatic hydrolysis of cellulose. *J Biotechnol.* 88:177-182.

Jiménez J, Domínguez M, Castillón MP, Acebal C (1995). Thermoinactivation of cellobiohydrolase I from *Trichoderma reesei* QM 9414. *Carb Res.* 268:257-266.

Liu C, Wyman CE (2004). Impact of fluid velocity on hot water only pretreatment of corn stover in a flowthrough reactor. *Appl Biochem Biotechnol.* 113-116:977-87.

Lynd L, Weimer PJ, van Zyl WH, Pretorius IS (2002). Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol Mol Biol Rev.* 66(3):506-577.

Mais U, Estaghlalian AR, Saddler JN (2002). Influence of Mixing Regime on Enzymatic Saccharification of Steam-Exploded Softwood Chips. *Appl Biochem Bioeng.* 98-100:464-472.

Mansfield SD, Mooney C, Saddler JN (1999). Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnol Prog.* 15:804-816.

Meunier-Goddik L and Penner MH (1999). Enzyme-Catalyzed Saccharification of Model Celluloses in the Presence of Lignacious Residues. *J Agric Food Chem.* 47:346-351.

Mohagheghi A, Tucker M, Grohmann K, Wyman C (1992). High Solids Simultaneous Saccharification and Fermentation of Pretreated Wheat Straw to Ethanol. *Appl Biochem Biotechnol.* 33:67-81.

Mooney CA, Mansfield SD, Touhy MG, Saddler JN (1998). The Effect of Initial Pore Volume and Lignin Content on the Enzymatic Hydrolysis of Softwoods. *Biores Technol.* 64:113-119.

Nagle N, Elander RT, Newman MM, Rohrback BT, Ruiz RO, Torget RW (2002). Efficacy of a Hot Washing Process for Pretreated Yellow Poplar to Enhance Bioethanol Production. *Biotechnol Prog.* 18:734-738.

Nguyen QA, Tucker M, Keller F, Eddy F (2000). Two-Stage Dilute Acid Pretreatment of Softwoods. Appl *Biochem Biotechnol* 84-86:561-576.

Ohmine K, Ooshima H, Harano Y (1983). Kinetic Study on Enzymatic Hydrolysis of Cellulose by Cellulase from *Trichoderma viride. Biotechnol Bioeng.* 25(8):2041-2052.

Ooshima H, Burns DS, Converse AO (1990). Adsorption of Cellulase from *Trichoderma reesei* on Cellulose and Lignacious Residue in Wood Pretreated by Dilute Sulfuric Acid with Explosive Decompression. *Biotechnol Bioeng.* 36:446-452.

Ooshima H, Kurakak M, Kato J, Harano Y (1991). Enzyme Activity of Cellulase Adsorbed on Cellulose and its Change During Hydrolysis. *Appl Biochem Biotechnol.* 31(3):253-266.

Palonen H, Tjerneld Zacchi G, Tenkanen M (2004). Adsorption of *Trichoderma reesei* CBH I and EG II and catalytic domains on steam pretreated softwood and isolated lignin. *J Biotechnol.* 107:65-72.

Philippidis GP and Hatzis C (1997). Biochemical Engineering Analysis of Critical Process Factors in the Biomass-to-Ethanol Technology. *Biotech Prog.* 13(3):222-231.

Pimenova NV and Hanley TR (2003). Measurement of Rheological Properties of Corn Stover Suspensions. *Appl Biochem Biotechnol.* 105-108:383-392.

Puri VP (1984). Effect of Crystallinity and Degree of Polymerization of Cellulose on Enzymatic Saccharification. *Biotechnol Bioeng.* 26:1219-1222.

Ramos LP, Nazhad MM, Saddler JN (1993). Effect of Enyzmatic Hydrolysis on the Morphology and Fine Structure of Pretreated Cellulosic Residues. *Enzyme Microbiol Technol.* 15:821-831.

Saha BC and Bothast (1996). Production, Purification, and Characterization of a Highly Glucose-Tolerant Novel β-Glucosidase from *Candida peltata. Appl Environ Microbiol.* 62(9):3165-3170.

Schell DJ, Farmer J, Zuccharello M (2003). Effect of Solids Loading on Pretreatment Kinetics. NREL C-Milestone Report.

Spindler DD, Wyman CE, Mohagheghi A, Grohmann K (1988). Thermotolerant Yeast for Simultaneous Saccharification and Fermentation of cellulose to Ethanol. Appl Biochem Biotechnol. 17:279-293.

Väljamäe P, Sild V, Pettersson G, Johansson G (1998). The initial kinetics of hydrolysis by cellobiohydrolases I and II is consistent with a cellulose surface-erosion model. *Eur J Biochem.* 253(2):469-75.

Varga E, Klinke HB, Réczey K, Thomsen AB. (2004). High solid simultaneous saccharification and fermentation of wet oxidized corn stover to ethanol. *Biotechnol Bioeng.* 88(5):567-574.

Wald S, Wilke CR, Blanche HW (1984). Kinetics of the Enzymatic Hydrolysis of Cellulose. *Biotechnol Bioeng.* 26:221-230.

Wingren A, Galbe M, 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.

Wyman CE and Yang Bin (2004). *Application of Lignin Blockers and Inactive Cellulase to Understand Enzymatic Hydrolysis Kinetics of Pretreated Cellulose.* Presented at AIChE National Meeting, Nov. 11, Austin, TX.

Xiao Z, Zhang X, Gregg DJ, Saddler JN (2004). Effects of Sugar Inhibition of Cellulases and β-Glucosidase During Enzymatic Hydrolysis of Softwood Substrates. *Appl Biochem Biotechnol.* 113-116:1115-1126.

Zhang S, Wolfgang DE, Wilson DB (1999). Substrate heterogeneity causes the nonlinear kinetics of insoluble cellulose hydrolysis. *Biotechnol Bioeng.* 66(1):35-41.

Zhang YHP and Lynd L (2004). Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. *Biotechnol Bioeng.* 88(7):797-824.

Chapter 4

SCALE-UP FROM SHAKE FLASK TO STIRRED TANK REACTORS

4.1 Overview and Context

Past work has indicated that by scaling up the enzymatic hydrolysis of cellulose from shake flask to stirred tank reactors, differences in both the rate and extent of cellulose hydrolysis result. The goal of this work was to determine unambiguously the source of these differences. It was determined in this work that differences in the performance of the enzymatic saccharification reaction between reaction vessels can be minimized if proper attention is given to temperature control and bulk mixing. It was demonstrated that as insoluble solids levels increase in a stirred tank reactor, temperature control becomes more problematic due to mixing inefficiency and the heat transfer properties of the slurry. The high shear forces necessary to ensure mixing in stirred tank reactors at high solids were found not to result in lower extent of reaction due to inactivation of enzyme, but rather impeller agitation was found to improve extent of reaction for some conditions. Statistical analysis using linear correlation models was also applied. This analysis showed that reactor type is not a significant factor that can be linearly correlated to the cellulose conversion. The factors exerting a significant influence on the cellulose conversion were demonstrated (as expected) to be the initial insoluble solids and reaction time.

4.2 Background

4.2.1 Saccharification Reactor Scale-up

Many physical properties scale nonlinearly with reactor volume in biological reaction systems (Bailey and Olis, 1986), which can present some design challenges. Issues usually associated with the scale-up of industrial processes fall into the two general classes of reactor-dependent scaling issues and reaction slurry-dependent scaling issues. The primary physical phenomena associated with these are temperature control and agitation which correspond to reactor-dependent issues, and the reaction slurry-dependent issues of heat and mass transfer. The typical bench scale assay for the enzymatic hydrolysis process uses the shake flask reactor, and much of the literature on enzymatic hydrolysis of lignocellulose describes research performed using this reactor system. At this scale, reactor-dependent scaling issues do not become significant. In shake flask reactors, mixing is typically accomplished through either magnetic stir bars or a rotary incubator, while temperature is maintained using air-temperature controlled incubators or water baths. For industrial or laboratory scale stirred tank reactors (STRs), maintaining precise temperature control requires achieving good mixing in the reactor. This is much harder to achieve in STRs at high solids levels than in shake flasks, since STRs must be effectively mixed to obtain the heat transfer needed for temperature control. Other important issues associated with impeller agitation systems are mixing effectiveness and power usage (Oldshue, 1983) as well as shear-induced effects on enzyme stability (Mukataka *et al.*, 1983; Reese and Ryu, 1980). For batch saccharification, mixing problems are primarily encountered only during the first 8-12 hours when approximately 10% to 20% of the initial insoluble solids are solubilized. Due to this changing rheology

with extent of reaction, it is important to identify mixing systems that are efficient over a range of process-relevant insoluble solids levels and high apparent viscosities.

Because of differences in control approaches used at different scales or solids levels, there are different factors that emerge which affect reaction performance. Some of the factors that are postulated to negatively impact the scale-up of the enzymatic hydrolysis reaction from shake flasks to a larger stirred vessel are listed in Table 4.1. In spite of these differences between reactor systems, issues specifically relating to the behavior of cellulase enzymes in STRs have been examined by relatively few studies, while scale-up specific studies are inadequately represented in the literature.

Table 4.1: Factors Affecting Enzymatic Saccharification Reaction Scale-up

Factor	Postulated Affect(s) on Saccharification
Bulk mixing	Impacts ability to control or minimize gradients within the reaction vessel (*e.g.* T, pH, concentration)
Solids level	Results in non-Newtonian slurry rheology, exacerbates mixing problems and bulk heterogeneity
Shear	Shear inactivation of enzyme (impeller tip speed or gas-liquid interfacial area)
Surface to volume ratio	Influences the amount of interfacial area (liquid-solid, or liquid-gas) available for enzyme adsorption or inactivation events to occur

The enzymatic hydrolysis of cellulose has been widely studied, and a large body of literature exists on reaction dynamics, effects of enzyme source and loading, effects of substrate source and loading, and temperature, which has been summarized in several recent reviews (Zhang and Lynd, 2004; Lynd *et al.*, 2002). Due to the heterogeneous nature and diversity of cellulosic substrates and the differences in the types of enzymes used and biochemical mechanisms involved, it is difficult to generalize about the behavior of the reaction system. Cellulases are primarily obtained from the fungus

Trichoderma reesei, and are sometimes supplemented with β-glucosidase (often from another microbial source) to increase the cellulase preparation activity, since *T. reesei* β-glucosidases less tolerant to glucose inhibition than enzymes from other microorganisms such as *Aspergillus niger* (Lynd *et al.*, 2002). The commercial cellulase preparation Spezyme CP (Genencor International) has been found to exhibit robust activity in the presence of high temperature (50°C) and lignocellulosic hydrolyzate liquors (Kadam and Mohagheghi, unpublished data), and has high specific FPU activity; a Western blot verified that precommercial Spezyme preparations contain *T. reesei* CBH I (Cel7A) and detected no *A. niger* β-glucosidase activity (Nieves *et al.*, 1998). Because of the high glucose tolerance of Spezyme CP β-glucosidases, this enzyme component must be supplemented from another source besides *T. reesei*.

Mass transfer and shear stress effects can be identified initially as having the most potential for scaling differences between shake flasks and STRs. It has been shown (Ingesson *et al.*, 2001; Mais *et al.*, 2002) that increasing the shaking rate from 25 to 150 rpm for shake flasks in a rotary incubator improved cellulose conversion and enzyme adsorption using Celluclast and Novozyme 188 for saccharification of SO_2-catalyzed steam exploded softwood chips with lignin removed by alkali-peroxide treatment. This is most likely due to improved mass transfer. Shear stresses in shake flask reactors are not expected to be significant relative to agitation by impellers in STRs.

Other work has characterized the effect of shear stress on cellulase activity. The study by Reese and Ryu (1980) quantitatively demonstrated that degradation of *T. reesei* cellulase activity could be affected by shear stresses, with significant decreases at shear stresses greater than 15 dyne/cm^2. Other work by Mukataka *et al.* (1983) demonstrated

that agitation rates >200 rpm using a single 6-bladed Rushton impeller lowered the extent of cellulose conversion in slurries of Avicel or paper pulp, with the suspected source being entrapment of air bubbles causing increased shear stress at the liquid-gas interface. The microbial source of the cellulase used in this work was not specified. In more recent studies by Tengborg *et al.* (2001) carried out in a 1-L reactor using a paddle-type impeller operated at speeds of up to 510 rpm, no decrease in the rate of cellulose hydrolysis was observed using Celluclast and Novozyme 188, indicating that inactivation by agitation-induced shear stress might not be a significant concern. For large-scale cellulase production by *T. reesei*, agitation rates >500 rpm are maintained during bench-scale cultivation to achieve aeration rates sufficient to meet the oxygen demand of the culture (Schell *et al.*, 2001). This suggests that cellulases should be tolerant of impeller and air bubble-induced shear stresses.

Fluid viscosities must be known for the determination of standard parameters associated with the characterization of fluids and the scale-up of mixing systems (Wilkens *et al.*, 2003) such as the Reynolds Number (Re_n) and the Power Number (N_P). For cellulosic enzymatic hydrolysis systems, this would require viscosity models as functions of both insoluble solids level and solids composition, since cellulose content decreases with increasing extent of reaction. Currently this information is not available for both variables. Pimenova and Hanley (2003) estimated the viscosities of PCS-water slurries (average fiber length = 120 μm) using a helical ribbon impeller viscometer. Because the high-solids slurries were non-Newtonian, the viscosities varied with shear rate in a power law relation, with order of magnitude increases starting at approximately 50 cP (Newtonian) at a level of 5% solids (w/w, dry basis) and reaching more than 10^6 cP

(highly non-Newtonian) at a level of 30% solids. The viscosities determined from this study are summarized in Table 4.2.

Table 4.2: Viscosities of Corn Stover Solids Suspended in Water
(Pimenova and Hanley, 2003)

Insoluble PCS Solids	Viscosity Range (cP)	Fluid Rheology
5%	~50	Newtonian
10%	~100-1000	Non-Newtonian
20%	~1000-50000	Non-Newtonian
30%	~10000-1000000	Non-Newtonian

Pulp suspensions, physically and rheologically similar to pretreated corn stover (PCS) solids suspensions, act as Bingham plastics in exhibiting a yield stress, in that the shear stress must reach a certain threshold value before the fluid will begin to flow (Bakker and Gates, 1995). Fluids of this type form "caverns" of flowing fluid adjacent to the source of shear in the regions where the shear stresses are large enough to exceed the yield stress. We have observed this behavior at high insoluble solids levels in 2-L, 5-L, and 20-L scales of stirred tank bioreactors.

4.2.2 Statistical Analysis

In order to establish that observed differences in experimental data are either genuine or trivial, a statistical approach should be considered. Statistics seeks to address the relationship between data sets based upon the confidence established by the replication within these data sets. Statistical methodology is based on assumptions about data conforming to given data distributions and confidence limits assigned to these data on the basis of various features of the data (degrees of freedom). In order to describe coherent statistical methods, the metrics, tools, and terminology must be defined, and format for

determining significant differences between data must be developed through techniques such as hypothesis testing, correlation models, and analysis of variance (ANOVA).

Hypothesis testing is a method for using statistics to determine the probability that a given hypothesis is true. This can be expressed by comparing a statistic based on the sample mean and variance to a statistic based on a reference statistic determined from a known probability distribution given the number of degrees of freedom (DOF) and the level of confidence. The sample mean $\overline{X}$ for a set $X = \{x_1,...,x_n\}$ of n sample observations can be defined as:

$$\overline{X} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.1}$$

and can be considered an unbiased estimator of the population mean. The sample variance can be defined as:

$$s^2 \equiv \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2 \tag{4.2}$$

where the n-1 term represents the degrees of freedom (DOF) for this set, and is necessary for this metric to be considered an unbiased estimator of the population variance (http://mathworld.wolfram.com).

The Student's t-distribution is the distribution determined if a small number (< 30) of random samples are taken from a normal distribution and used to estimate the variance (Dowdy and Wearden, 1983). In its simplest form the t-statistic can be defined as the difference between the sample and population mean for the computed variance as:

$$t \equiv \frac{\overline{X} - \mu}{s/\sqrt{n}} \tag{4.3}$$

In the normal distribution, $\overline{X}$ is the only random variable, whereas in the t-distribution, $\overline{X}$ and s are random. Due to this, the two distributions are not identical, although as $n \rightarrow \infty$ Student's t-distribution approximates the normal distribution. To test the null hypothesis H_0: $\overline{X} = \mu$, the confidence region is based on the t-distribution, which for a given confidence level is:

$$t \leq t_{\frac{\alpha}{2}, n-1} \tag{4.4}$$

This only tests whether samples from a single population conform to a known or postulated mean. When comparing samples from two populations, the metric must be modified. The group comparison t-test compares the sample means of two groups of samples, which doesn't require equal sample sizes, but does assume equal variances between the two populations. For testing the null hypothesis H_0: $\mu_1 = \mu_2$, the t-statistic becomes:

$$t \equiv \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} \tag{4.5}$$

with the confidence region determined by the t-statistic with less degrees of freedom:

$$t \leq t_{\frac{\alpha}{2}, n_1 + n_2 - 2} \tag{4.6}$$

When data are available from a number of different conditions, ANOVA can be applied. This method tests for the heterogeneity of a number of means from different populations, but assumes the same variance (Dowdy and Wearden, 1983). For these data sets, there are three types of variability encountered: total variability (between data of all groups), within-group variability, and among-group variability. The method of ANOVA

uses the Fisher Variant Ratio (F-statistic), which is the ratio of the among-group mean squared error (MSE) to within-group MSE. This is compared to a reference F-statistic for a given confidence level and DOF.

Measured variables that are postulated to affect process performance can be defined as "factors", while process performance measures can be called "responses". Linear models can be developed that relate factors and factor interactions to the responses and can be used to determine how variables are correlated. Classical least squares (LS) approaches are widely used to estimate responses ($\hat{Y}$) from factors X, through a linear combination with the set of parameters B:

$$\hat{Y} = XB \tag{4.7}$$

where B can be derived identically through either an optimization approach as in Newton's Method or by direct linear projection as:

$$B = (X^T X)^{-1} X^T Y \tag{4.8}$$

although when measured variables in X are highly correlated, the matrix X could be ill-conditioned, resulting in singularities for the required matrix inversion. Performing nonlinear transformations and adding factor interaction terms or nonlinear combinations of factors generates a model that is nonlinear in X. This only requires the use of linear projection methods to generate a nonlinear model that is capable of predicting more complex behavior. Factor interaction terms may be necessary to include in the model since it is possible that factors have a cooperative effect on the process beyond what can be accounted for in additive linear terms. While increasing the number of parameters usually improves the model prediction, the degrees of freedom (DOF) are also reduced (Deming and Morgan, 1987).

One method of evaluating the strength of a linear relationship between two data sets is with the correlation coefficient (R). This is the measure of how much of the variance in the data is accounted for by the correlation. Values of R of either 1 or –1 indicate a perfect direct or inverse correlation, respectively, while 0 represents no correlation. Using R is not a perfect choice, as this does not consider the DOF (Deming and Morgan, 1987) such that as the number of parameters approaches the number of observations, R will approach 1. Another test of the model effectiveness that considers the DOF is an ANOVA test for the significance of the regression. This determines an F-statistic that is a ratio of the variance (sum of squared errors) due to model factors to the variance due to the regression. The hypothesis tested for ANOVA for regression models is whether the entire or individual parameters are significant relative to the data variance. Using standard statistical analysis, the confidence of the F-statistic can be tested at a given level of confidence.

4.3 Materials and Methods

4.3.1 Mixing and Heat Transfer: Temperature Control

For high solids mixing of PCS solids, metrics of mixing efficiency need to be identified. Two fundamental properties that define a homogenous system are an absence of temperature and concentration gradients. For this case, the component of interest is the solids. Effective bulk mixing of the solids was qualitatively evaluated by visual inspection due to the difficulty of making quantitative measurements. Both the effectiveness of temperature control and temperature measurements obtained at various radial points were used as the primary quantitative metric for characterizing the adequacy of mixing. Effective temperature control was defined as the ability of the temperature

control system to maintain the temperature at a given set point and to adjust to set point changes.

New Brunswick Bioflo 3000 fermentors equipped with 7 L vessels were used in an unoptimized configuration with 2 marine impellers (d/D=0.68) to study mixing of high-solids PCS slurries. Temperature measurements were taken at approximately 50% slurry height at 3 radial points in the bulk solution: at the wall of the reactor (r = R), adjacent to the temperature probe port (r = ~0.5R), and at the center of the reactor above the impellers (r = 0). A thermocouple was used for all measurements in addition to the thermocouple used by the reactor systems for temperature control.

It was expected that achieving a steady-state temperature would be problematic in the reactors at the higher solids levels, and that the temperature would cycle around the set point as mixing limitations increased. Additionally, temperature gradients in both the radial and axial directions were anticipated, since heating in these vessels occurs from the sides and bottom of the tank. Temperature measurements were made only to roughly determine the extent of temperature gradients; there was no attempt to characterize the flow field.

4.3.2 Impeller Speed (Shear) and Solids Effects

A summary of the reactor types studied is given in Table 4.3, where d/D is the ratio of impeller diameter to tank diameter, and H/D is the ratio of working reactor height to tank diameter. The two STR systems used are shown in Figure 4.1. Corn stover was pretreated by dilute sulfuric acid in NREL's continuous pilot-scale vertical Sunds reactor as reported elsewhere (Schell *et al.*, 2003). The resulting PCS solids (P030312CS) have a cellulose content of 53.2% (dry weight basis). The total and insoluble solids content of

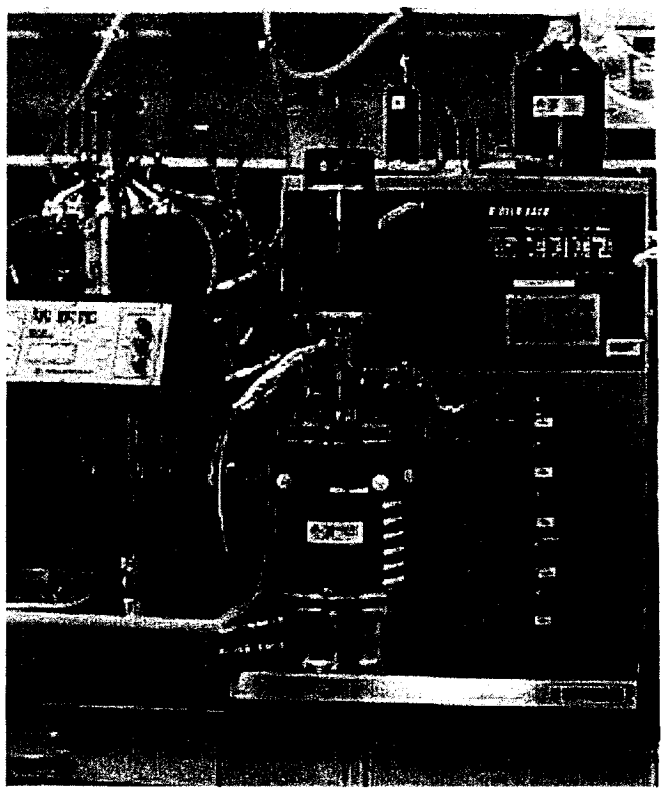


Figure 4.1: The Bioflo 3000 reactor configured for fermentation. *Photo by JC Sáez.*

the unwashed pretreated solids was performed according to NREL protocols LAP-001 and LAP-018, respectively. The appropriate amount of solids needed for each fermentor and shake flask was calculated using the results of these assays. The cellulase enzyme preparation used for this study was Spezyme CP (Lot # 301-00348-257, Genencor International). Total protein was assayed at 106.6 mg/mL (BCA assay, Pierce Biotechnology, Inc.) and the specific activity was determined to be 0.27 FPU/mg of protein using NREL protocol LAP-006. Enzyme loading for all cases was 20 mg protein/g cellulose except where otherwise specified.

Table 4.3: Reactor Systems used for Enzymatic Saccharification of PCS

Reactor System	Reactor Volume (L)	Working Volume (L)	Impeller/Tank Configuration	d/D	H/D
B. Braun	20	13.5	3 marines, no baffles	0.54	1.87
Bioflo 3000	7	3.5	2 marines, no baffles	0.68	1.05
Shake Flask	0.25	0.10			

Citrate buffer (1 M) was prepared, adjusted to pH 5.5 with NaOH, and was used as 5% of the total reactor mass to yield an effective molality of 0.05 mol/kg. Because different levels of solids required differing amounts of pH adjustment, the amount of extra NaOH (4 N) required to bring the pH to 4.8-5.0 was determined on a volume NaOH/g mixture basis. Enzyme, DI water, and citrate buffer were filter sterilized before use. Tetracycline was added to 4 μL/mL from a stock solution of 10 mg/L. The Bioflo 3000 fermentor glass vessel was wrapped with insulating material to minimize temperature gradients across the reactor. Bioflo fermentors and shake flasks were autoclaved empty, and the B. Braun fermentor was steam sterilized *in situ*. Reactors were agitated overnight with all components except enzyme to equilibrate the temperature.

Enzyme was added after the temperature had equilibrated and the zero time samples were taken.

Samples were removed with 5 mL pipettes (Falcon, Inc.) that had the tips broken off in order to accommodate high level of insoluble solids. These were transferred to 2.5 mL microcentrifuge tubes and the pH was measured. The samples were next centrifuged at 10000 rpm for 5 minutes. The supernatant was decanted and diluted 1:5 and syringe filtered into HPLC vials for subsequent HPLC analysis. HPLC analysis was performed on all saccharification samples to determine the glucose, xylose, and cellobiose concentrations. An Agilent 1100 series HPLC equipped with a differential refractive index detector was used. This was equipped with a Bio-Rad HPX-87H organic acid column operating at 55°C with a 0.01 N H_2SO_4 mobile phase at a flow rate of 0.6 mL/min. Insoluble solids were measured using NREL LAP-018 modified for smaller sample sizes. Conversions were calculated using g/kg concentrations of sugar components based on measured or estimated insoluble solids at the time of the sample.

Temperature was controlled for the shake flasks using an air-temperature controlled shaking incubator, while the two types of stirred tank fermentor systems used heating jackets. The Bioflo 3000 heat transfer covers only the dish at the bottom of the reactors, while the jacket for the B. Braun covers the majority of the vertical walls of the reactor. The Bioflo 3000 jacket temperature was controlled by a resistance heater in combination with a feedback control loop based on an internal reactor temperature. This was adjusted for several experiments to control the jacket temperature or connected to a water bath.

The B. Braun reactor controls the jacket temperature by a circulating water loop that is controlled by steam or cooling water addition in a heat exchanger. Shake flasks were

run in triplicate and Bioflo 3000 fermentors in duplicate for most conditions. A summary of conditions used in this series of studies is given in Table 4.4. Experiments #1 and #2 studied the effects of both impeller agitation and solids level in stirred tanks, while Experiments #3 through #6 studied primarily mixing and temperature control in larger vessels at higher solids loadings.

4.4 Results and Discussion

4.4.1 Effect of Solids Levels on Heat Transfer in Bench-Scale STRs

Figure 4.2 summarizes the results of the temperature control experiment, which examined the effects of reactor insoluble solids levels on temperature control, and shows that there are difficulties in using the conventionally configured (*i.e.* unoptimized for mixing PCS slurries) fermentor reactor system (Bioflo 3000) for controlling temperature in high-solids systems. Specifically, temperature gradients within the reactor increase as the insoluble solids level increases. Even at high impeller speeds using larger impellers, inability to adequately control temperature at higher solids levels results in overheating inside the reactor, particularly at the heating interface, and very slow response times to temperature set point changes, which are manifested by wide fluctuations in reactor temperature.

Temperature measurements were also made for the same reactor system using pre-enzymatically hydrolyzed PCS slurry containing about 10% cellulose (~80% cellulose removed). The objective was to determine if insoluble solids level was the most important parameter affecting temperature control, or if solids composition contributed significantly to the resistance to mixing and heat transfer. Since long cellulose fibers act as a barrier to mixing due to the non-Newtonian rheology they impart to PCS slurries, it

Table 4.4: Summary of Experimental Conditions

		Shake Flasks (100 mL working vol.)	Bioflo 3000 (3.5 L working vol.)	B. Braun (13.5 L working vol.)
Exp #1	Initial insoluble solids (w/w, dry basis)	5.0%	5.0%	----
	Agitation (rpm) or rotation rate	130	100 or 250	----
	Temperature control	Incubator (45°C)	Heating jacket with feedback control (45°C)	----
Exp #2	Initial insoluble solids (w/w, dry basis)	10.0%	10.0%	----
	Agitation (rpm) or rotation rate	130	250 or 450	----
	Temperature control	Incubator (45°C)	Heating jacket with feedback control (45°C)	----
Exp #3	Initial insoluble solids (w/w, dry basis)	13.5%	13.5%	----
	Agitation (rpm) or rotation rate	130	400	----
	Temperature control	Incubator (45°C)	Control jacket (45°C) first 24 h. Feedback control w/ internal T after 24 h	----
Exp #4	Initial insoluble solids (w/w, dry basis)	15.0%	15.0%	----
	Agitation (rpm) or rotation rate	130	400	----
	Temperature control	Incubator, stationary and shaking (45°C)	Incubator first 24 h. Feedback control with internal T after 24 h.	----
Exp #5	Initial insoluble solids (w/w, dry basis)	10.0%	10.0%	10.0%
	Agitation (rpm) or rotation rate	130	300	300
	Temperature control	Incubator (45°C)	Heating jacket with feedback control (45°C)	Control jacket T at 45°C
Exp #6	Initial insoluble solids (w/w, dry basis)	15.0%	15.0%	----
	Agitation (rpm) or rotation rate	130	400	----
	Temperature control	Incubator (45°C)	Control jacket T at 47°C- 52°C	----

was anticipated that the use of pre-hydrolyzed PCS might improve mixing and temperature control relative to using "normal" high-cellulose PCS slurries at the same insoluble solids loadings. This work demonstrated that mixing slurry of 14% insoluble solids is not a significant problem for low-cellulose content PCS. This can be compared to the results in Figure 4.2 for 13.5% unhydrolyzed insoluble solids, which was the approximate upper limit for successful temperature control. Near its maximum solids limit, 16% pre-hydrolyzed insoluble solids can be effectively controlled at the desired temperature provided there is vigorous enough agitation.

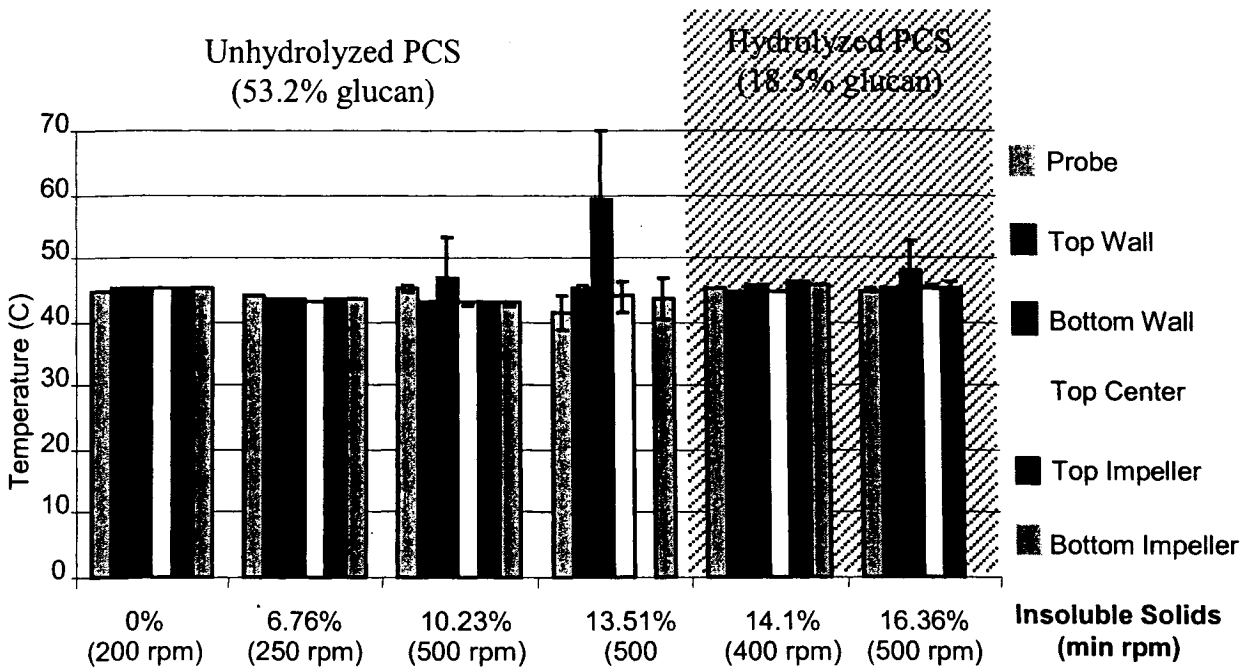


Figure 4.2: Comparison of STR temperature control effectiveness using the New Brunswick Bioflo 3000 reactor.

In addition to temperature gradients, visual observations of mixing efficiency were performed for all experimental conditions. At insoluble solids concentrations ≥13.5% PCS zones of stagnant slurry were present at the highest impeller speed used (500 rpm), indicating that substantial restrictions in mass and heat transfer occur at this solids level. When using enzymatically hydrolyzed PCS containing ~10% cellulose (compared to unhydrolyzed PCS which contains over 50% cellulose), the mixing performance of the fluid was distinctly different. At an insoluble solids level of 14% using enzymatically hydrolyzed PCS, complete mixing could be achieved at relatively low (300 rpm) agitation rates with no apparent dead zones present. Using pre-hydrolyzed PCS at 16%, mixing became more difficult and resulted in some dead zones even at the highest agitation rate used (500 rpm), although mixing and temperature control were still superior to those of unhydrolyzed PCS solids at lower insoluble solids levels. The significance of these findings is that either fed-batch or continuous STRs can be operated at higher insoluble solids levels when they contain PCS slurries at higher extents of conversion.

4.4.2 Effect of Impeller Speed (Shear)

The two postulated influences of impeller speed are fluid mixing and shear inactivation of enzymes. It has been postulated that higher impeller speeds lead to shear-induced inactivation of enzymes and result in lower cellulose conversions in addition to requiring higher power usage. On the other hand, the previous mixing results indicate that relatively high impeller speeds (300-500 rpm) are necessary to sufficiently mix the fluid to allow for proper temperature control. To better understand these issues, this series of enzymatic cellulose saccharification experiments examined the impact of using different

impeller speeds. Variables such as shear rate, shear stress, impeller Reynolds number, and Power number were not calculated since viscosity data on PCS slurries (apparent viscosity as a function of solids level, shear rate, and extent of conversion) is not yet available. Experiments were performed at initial insoluble solids levels of 5%, and 10% in the Bioflo 3000 fermentor as listed in Table 4.3.

Experiments #1 and 2 profiled reactor kinetics using a range of impeller speeds for initial insoluble solids levels of 5% and 10%, respectively. The lower impeller speed used was the minimum required to provide sufficient mixing for temperature control for the given solids level, as determined from Figure 4.2. The data in Figure 4.3 confirms that higher impeller speeds do not result in reduced enzyme reactivity and lower conversions. This indicates that the major concerns related to agitation rate should be power usage and mixing efficiency rather than enzyme stability.

4.4.3 Solids Effects on Extent and Rate of Saccharification

Figure 4.4 shows data from select runs at various unwashed solids loadings carried out in Experiments #1-6, with data shown separately for each of the three reactor systems studied. The final cellulose conversions range from ~85% to 70% as insoluble solids levels increase. The clear trend from this data is that increasing insoluble solids concentrations decreases both the rate and extent of cellulose utilization. The three expected reasons for this behavior are inhibition of cellulases by glucose and cellobiose, lignin, and other soluble components produced during pretreatment, with the levels of these inhibitors increases with increasing unwashed solids loading. A summary of final conversions for all experiments is provided in Table 4.5. The general trend is that for increasing unwashed solids loadings, the ultimate cellulose conversion is decreased.

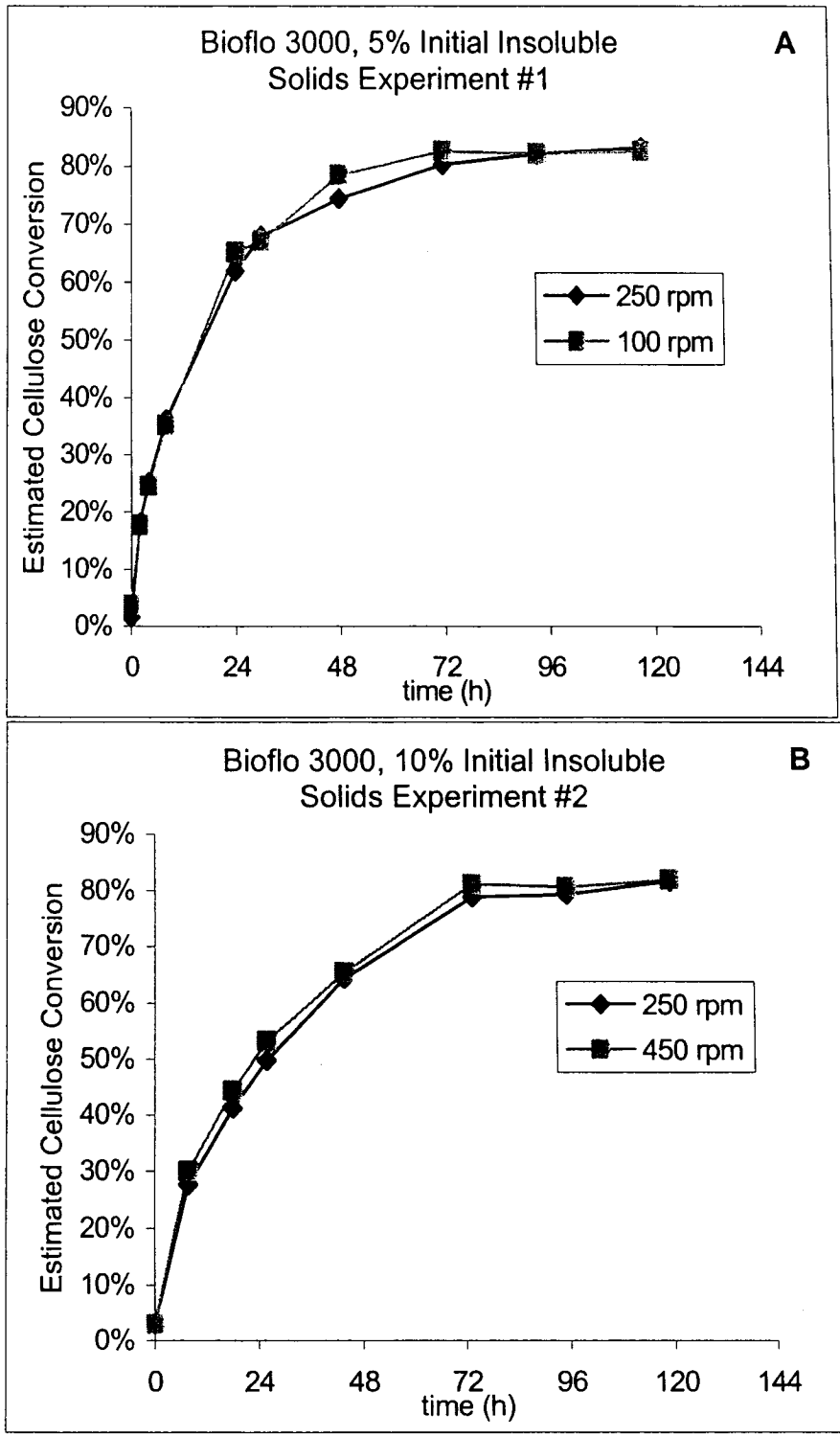


Figure 4.3: Effect of impeller speed on cellulose conversion (unwashed solids, no hydrolyzate liquor)

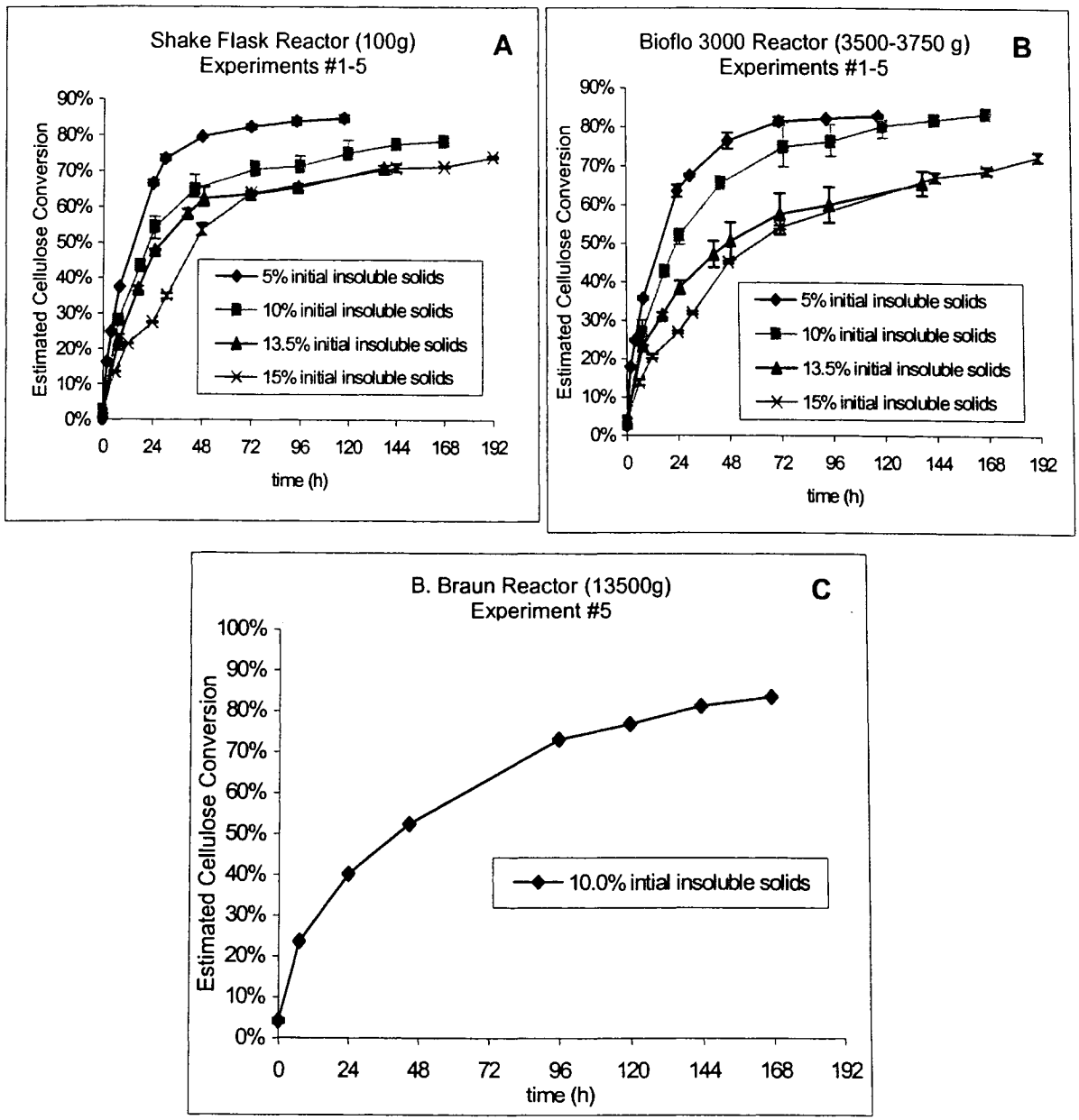


Figure 4.4: Comparison of all solids loadings for each reactor (unwashed solids, no hydrolyzate liquor)

Table 4.5: Summary of Cellulose Conversions for 20 mg/g Spezyme CP Loading					
			Estimated Cellulose Conversion at Final Time Sample (%)		
Exp #	Initial Insoluble Solids (%)	Final Time Sample (h)	Shake Flask (Averages)	New Brunswick Bioflo 3000 Reactor Averages	B. Braun mini-pilot reactor
1	5	118	84.4	82.6	----
2	10	120	74.7	81.6	----
3	13.5	138	70.5	65.5	----
4	15	192	73.5	72.1	----
5	10	168	78.1	82.9	83.5
6	15	192	71.8	68.1	----

Because the concentration of enzymatically produced glucose and cellobiose increases as insoluble solids levels increase and hydrolysis proceeds, it is expected that high accumulated sugar concentrations will result in a decrease in the rate of saccharification due to product inhibition. Since the solids for this set of experiments were not washed, there remain some soluble solids (*i.e.* hydrolyzate liquor) entrained in the PCS. Increasing the insoluble solids level also increases the concentrations of these entrained solids. Even at the highest insoluble solids levels, the initial concentration of glucose (~5 g/L) and xylose (~20 g/L) are not at high enough levels to inhibit of cellulases, so it is suspected that other soluble pretreatment products are contributing to the decrease in the rate and extent of saccharification as insoluble solids levels increase. This phenomenon has been demonstrated in the past when addition of pretreated hydrolyzate liquors to the PCS enzymatic saccharification reaction lowered cellulose conversions by >30% (Kadam and Mohagheghi, unpublished data). Lignaceous residues have been implicated in the inhibition of cellulases due to the nonspecific binding of

cellulases (Vinzant *et al.*, 1997; Mooney *et al.*, 1998; Meunier-Goddik and Penner, 1999), and it is expected that higher concentrations of unwashed solids result in higher concentrations of lignin-derived compounds in the liquid phase.

4.4.4 Effect of Reactor Scale

The rate of cellulose hydrolysis by cellulase enzymes is strongly dependent on the temperature at which the reaction occurs. Since a temperature increase from 38°C to 45°C can almost double the initial rate of cellulose hydrolysis (CHAPTER 3), temperature gradients and zones of poor mixing exacerbate temperature differences and can dramatically affect the rate of cellulose hydrolysis. The important finding of this study was that the problems associated with scale up from shake flasks to STRs are not intrinsically due to scaling, but rather to deficiencies in temperature control caused by the impact of poor mixing on heat transfer.

Figure 4.5A shows that for 5% initial insoluble solids, the type of reactor system used does not affect either the rate of cellulose utilization or the final conversion. The reason for this is that at this level of initial insoluble solids, there are essentially no limitations to mixing or heat transfer in either reactor system used. Figure 4.5B shows the results for the 10% solids loading experiments, which is where the most replicates were run. This is also the experimental condition at which the 20-L steam-in-place mini-pilot scale B. Braun fermentor was operated. In order to prevent excessively high temperatures at the reactor-heating jacket interface for the B. Braun reactor, the jacket water temperature was controlled at 45°C rather than using a feedback control loop to control the internal jacket temperature. For most of the reaction, however, the internal reactor temperature

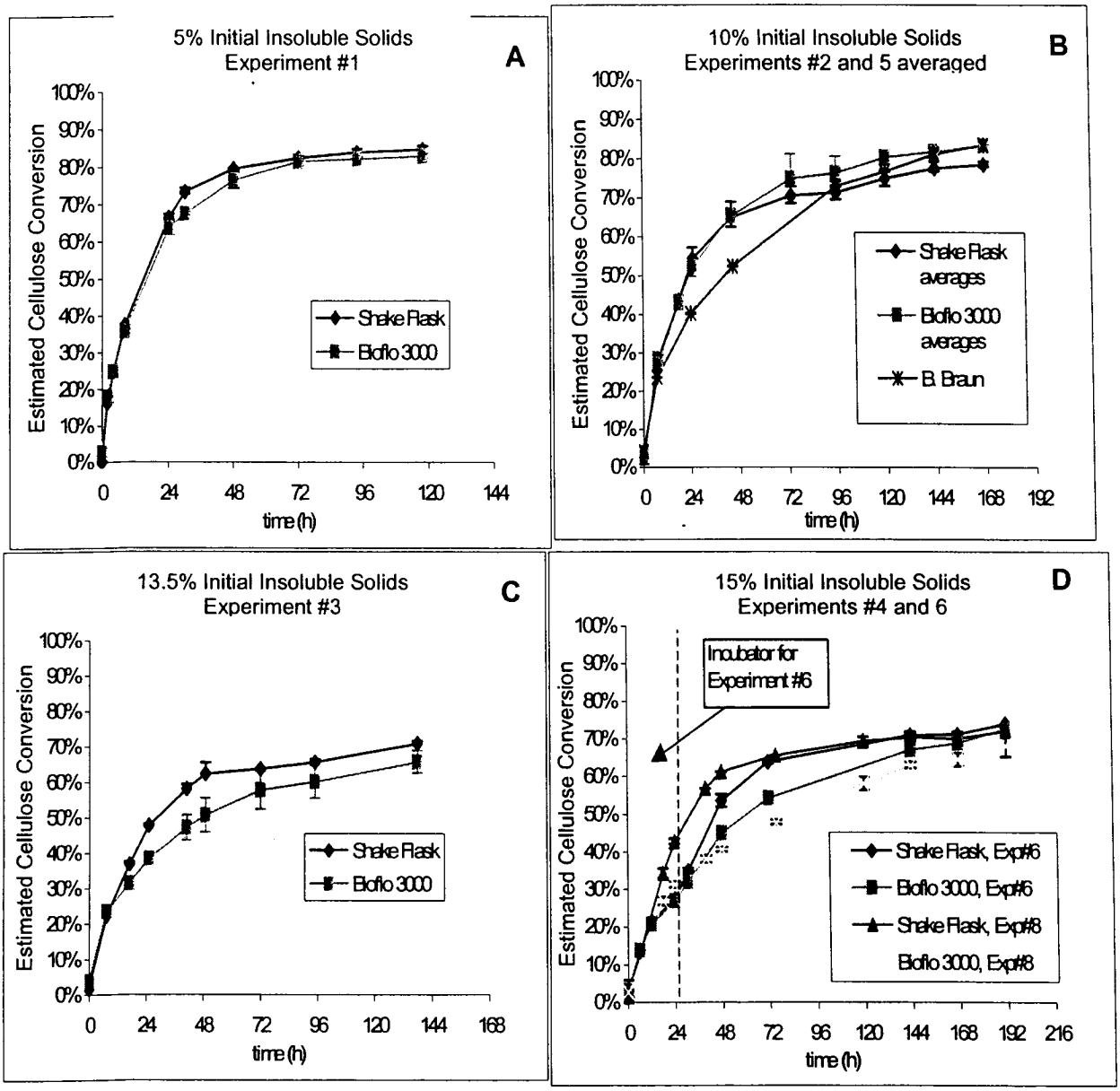


Figure 4.5: Effect of reactor type on enzymatic saccharification (unwashed solids, no hydrolyzate liquor)

was also 45°C, except during approximately the first 24 hours. Figure 4.5B shows that there are not significant differences between the rates of saccharification in the shake flask reactors and the Bioflo 3000 fermentors. Differences between experimental replicates for the Bioflo 3000 fermentors are probably due to the addition and distribution of NaOH added initially for pH control. The initial cellulose utilization rate in the B. Braun fermentor is slightly slower than in the other reactor systems, and is most likely due to the lower initial temperature in the reactor due to the more conservative temperature control approach taken. One interesting result seen in this figure is that the STRs show higher final cellulose conversions than the shake flasks. One postulated reason for this result is that this solids level provides an improvement in the hydrolysis of cellulose due to mechanical fracturing of PCS solids to expose more surface area to hydrolysis, while the solids levels are not so high as to inhibit mixing and temperature control.

At initial insoluble solids loadings greater than 10%, significant differences between STRs and shake flask systems become more apparent, as a comparison of Figures 4.5A and 4.5B with Figures 4.5C and 4.5D shows. The reason for this is primarily the challenging PCS slurry rheology, which at solids levels >10% makes it increasingly difficult to control the temperature (Figure 4.2). Since using a standard feed-back control loop to control temperature in the Bioflo 3000 reactors results in significantly high surface temperatures at the reactor-heating jacket interface, the jacket was controlled at 45°C for the first ~36 hours for the 13.5% initial insoluble solids run. The data in Figures 4.5C and 4.5D show significantly slower cellulose utilization rates for the Bioflo 3000 reactors at these higher solids conditions, although one of the duplicate reactors run at

15% solids was able to reach approximately the same cellulose conversion level as the shake flasks (Figure 4.5D). Another suspected reason for the lower conversions and rates in the experiments run at 13.5% and 15% solids levels is that the heating jacket may have been excessively hot (>55°C) during feedback temperature control, which could have resulted in irreversible thermal inactivation of cellulases.

An initial insoluble PCS loading of 15% was used for two sets of experiments (Exp. #4 and 6, Figure 4.5D). In the first experiment (Experiment #4), both types of reactors were placed in a stationary incubator for the first 24 hours after addition of enzyme. After this, the Bioflo 3000 reactors were operated in a feedback control loop for temperature control while the shake flasks were placed in a shaking incubator at 45°C. In the other experiment (Experiment #6), the heating jackets of the Bioflo 3000 reactors were maintained between 47°C and 52°C to attempt to maintain the internal temperature at 45°C. There were problems initially, however, due to the high solids levels causing problems with heat transfer so that a jacket temperature of 52°C only resulted in a reactor temperature of ~41°C. As the saccharification proceeded, the reactors were ultimately able to reach 45°C without requiring excessively high jacket temperatures.

In addition to glucose, cellobiose levels in the reactors can be used as an indication of the rate at which cellobiohydrolases (CBHs) are acting relative to glucanases and β-glucosidases. Figure 4.7 shows the cellobiose concentration profiles obtained using initial insoluble solids concentrations of 5% or greater. As Figure 4.7A shows, the differences in cellobiose concentrations between reactor types is insignificant at the 5% insoluble

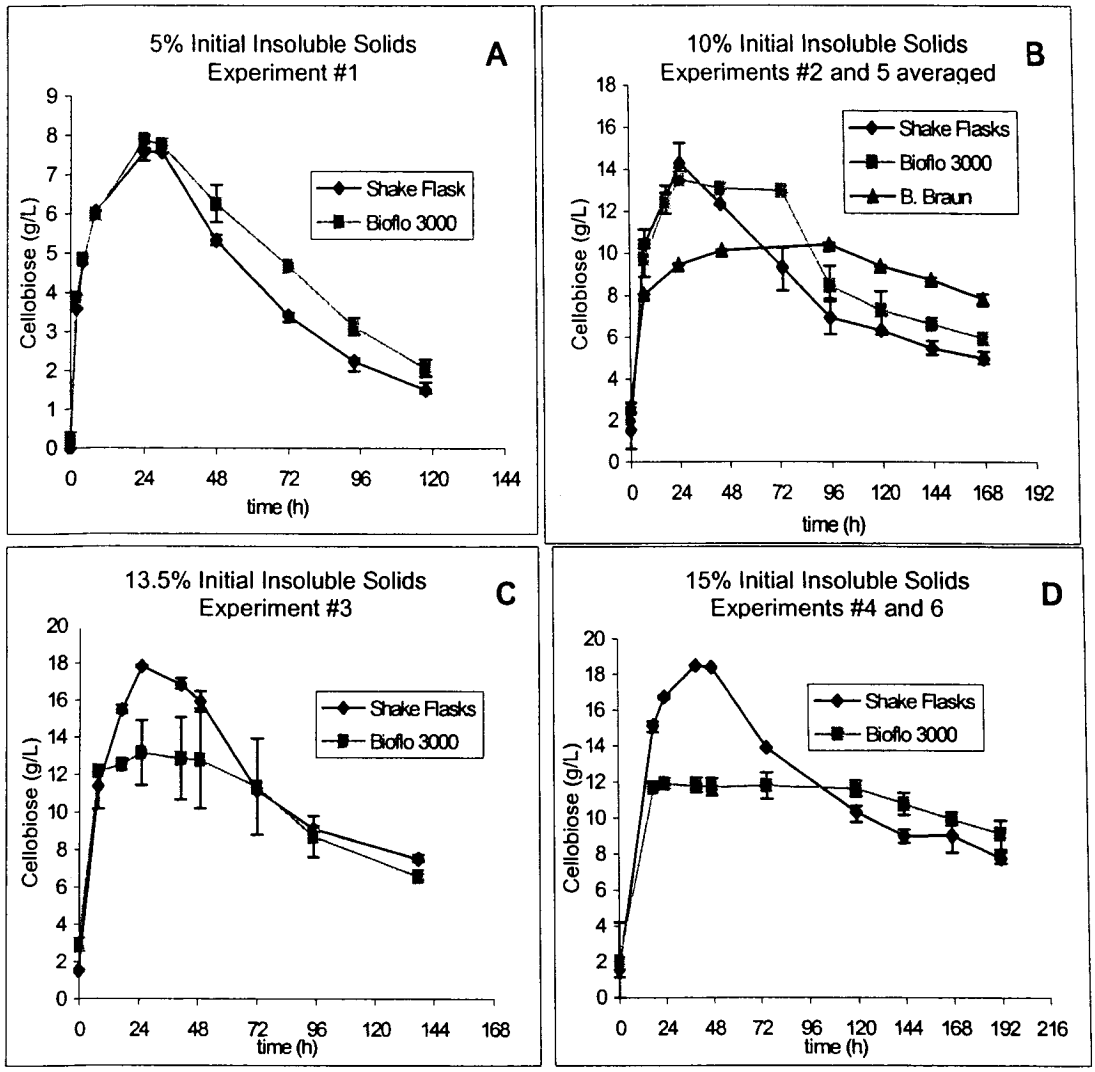


Figure 4.6: Cellobiose levels as a function of reactor type and solids loading (unwashed solids, no hydrolyzate liquor)

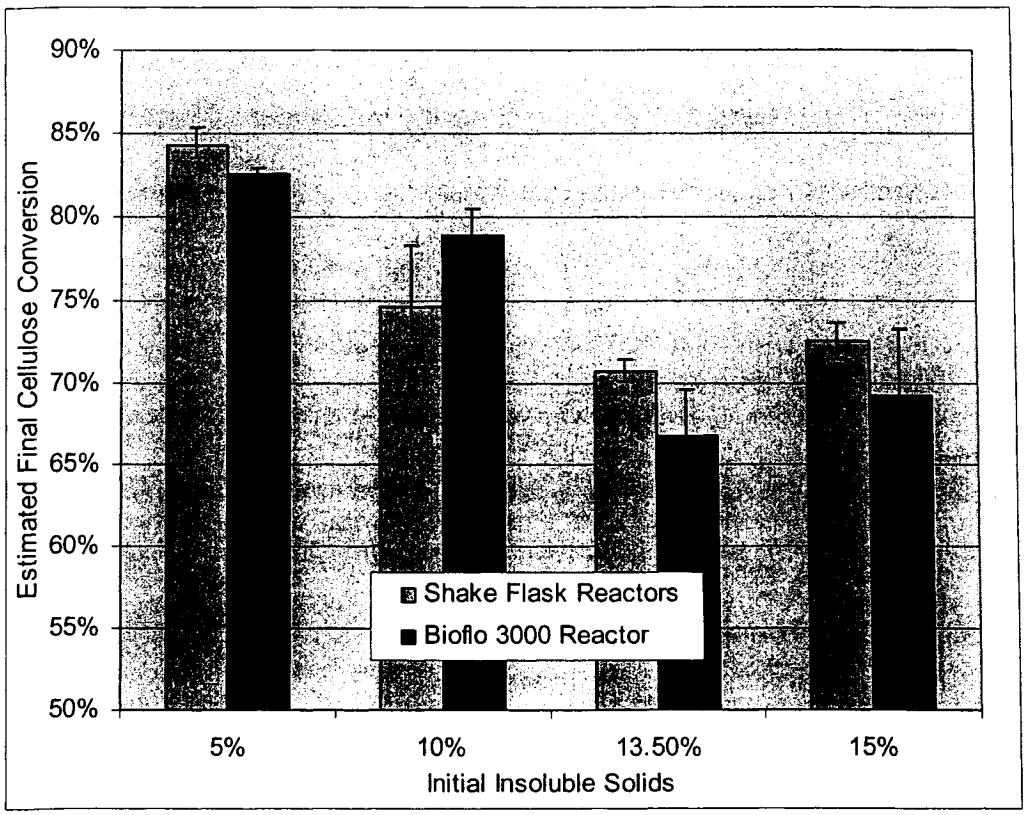


Figure 4.7: Summary of final conversions for scale-up experiments. Error bars indicate total data spread.

solids level. Figure 4.7 shows that at the higher solids levels (≥10%) the STRs generally had lower cellobiose levels during the early phase of the reaction than the shake flasks. The expected reason is that the rates of reaction of CBHs (forming cellobiose) are less than the combined rates of glucanases and β-glucosidase (forming glucose). The slower rates in the STRs probably are a consequence of the lower temperature in the STR systems due to imperfect temperature control at the higher solids levels. As shown previously, the range of agitation rates used for these experiments does not affect cellobiose levels; there is no evidence that these conditions detrimentally affect CBH activity.

4.4.5 Statistical Analysis of Saccharification Scale-up

Several types of statistical analyses were carried out to evaluate the significance of the observed differences in saccharification reaction performance by varying reactor types and initial insoluble solids. The first analysis used the Student's t-test (group comparison for unequal group sizes) to demonstrate that the shake flask system is not inherently superior in conversion efficiency to the STR systems. The data used in the analysis were from the final conversion time points only (Table 4.5), to simplify the comparison. By selecting only the final time point, the time-dependence of the data is removed. The initial insoluble solids level was held constant for each comparison. The data used in this study are plotted in Figure 4.7 with error bars that indicate the data spread. The data were taken from replicate reactors for the Bioflo 3000 reactors, triplicate reactors for the shake flask reactors, and duplicate samples for the B. Braun reactor. The data for 10% and 15% initial insoluble solids levels were from multiple experiments, and were based on similar final time points. A MATLAB script was written

to perform these t-tests, and is given as *ttest2.m* in **APPENDIX 1**. The null hypothesis (H_0) for the t-test comparison is that under similar conditions the mean final level of conversion achieved in STRs is equal to (or greater than) the mean level of conversion in shake flasks. The results of individual t-tests carried out at each solids level indicate that the null hypothesis cannot be rejected at a 95% confidence level for reactor systems containing up to 13.5% initial insoluble solids, such that at these solids levels final extents of cellulose conversion in STRs are equivalent or higher than in shake flask reactors. The null hypothesis can be rejected at the level of 15% initial insoluble solids. Even at this high solids level, however, the null hypothesis is only rejected at the 95% confidence level because of a single outlier: the Bioflo 3000 reactor run that obtained a lower level of cellulose conversion, presumably due to problems with mixing and heat transfer.

To supplement the t-test results models considering much more of the data simultaneously were developed and statistical methods used to determine the validity of the correlations uncovered by the model. The strength of these correlations was evaluated, and statistical ANOVA was used to determine whether the correlation explained by the model is significant relative to the noise in the data. This retrospective 2-level factorial ANOVA was carried out on the entire data set simultaneously using both Design Expert 6.0.1 statistical analysis software (Stat-Ease, Inc., Minneapolis, MN) and scripts developed in MATLAB to identify significant variables and construct models that were linear in factors and factor interaction terms. This approach has more statistical power than the single condition t-tests described above, since it considers variances and correlations across the entire data set simultaneously. Two models (Models I and II)

were used to determine correlations within the data. Model I used only the final time points, as in the data used for the t-tests, while Model II utilized all data throughout the course of the saccharification. The data from the two types of STRs are classified as a single category of reactor. The factors in Model I include the reactor type (as discrete binary states) and initial insoluble solids level, while the experimental response is the final cellulose conversion. In Model II, time is included as an additional factor, while the response is the cellulose conversion at the corresponding time.

The data for all statistical models are plotted in Figure 4.7 and the MATLAB scripts used to generate the models are given in **APPENDIX 1** as *pcadata.m* for Model I and *pcadata2.m* and *totaldata.m* for Model II. For Model I, using either the entire data set or different data subsets excluding one or more potential outliers, solids level was found to explain between 64% and 72% of the observed variance (R-value for each parameter as a percentage of the total sum of squared errors) in the data, while reactor type could only explain up to 2% (Table 4.6). The significance of solids level is expected, since increasing solids concentrations results in a proportional decrease in the extent of cellulose conversion. However, this analysis identified no statistically significant influence of reactor type on the extent of conversion. Using the complete data set and both inputs in the linear model, the prediction is given in Figure 4.8 and shows that reactor type is not well predicted by the model. The R-squared value for this model is 0.719. Further ANOVA analysis of the data set and model determined at a >95% confidence level that solids level influences cellulose conversion. Regardless of the data set used, the confidence level for the correlation between reactor type and cellulose conversion never exceeded 43%.

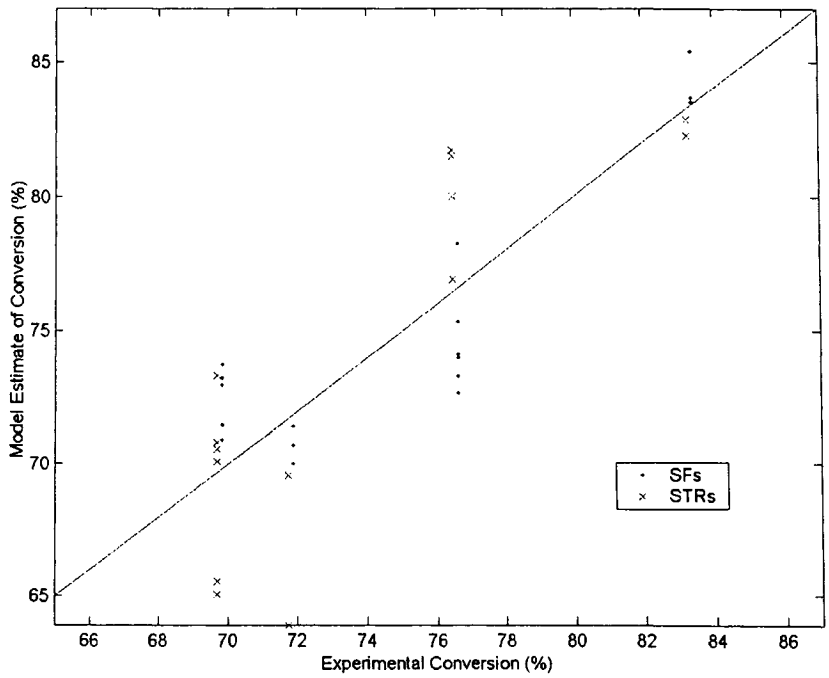


Figure 4.8: Linear correlation model based on final sample (Model I).

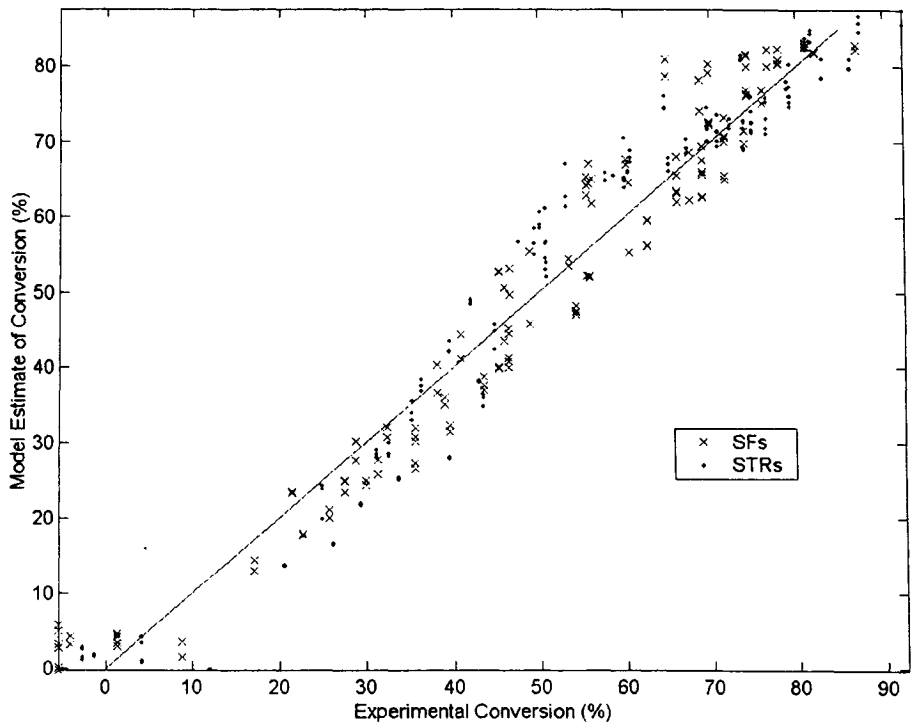


Figure 4.9: Linear correlation model based on data taken at all times throughout many saccharification reactions (Model II).

Table 4.6: Variable Contributions to Linear Models		
Variable	**Variable Contribution to Model I**	**Variable Contribution to Model II**
Time	----	86.98%
Initial Insoluble Solids	72.12%	7.73%
Reactor Type	1.77%	0.58%
Time/Initial Insoluble Solids interaction term	----	0.99%
Time/Reactor Type interaction term	----	0.13%
Initial Insoluble Solids/Reactor Type interaction term	0.31%	0.10%

The data used in Model II represent a much larger data set and a potentially more productive source for data mining, and were transformed with the $\log_{10}$ function. This nonlinear transformation allowed the nonlinear (approximately logarithmic) time-conversion correlation to be characterized by a linear model. As given in Table 4.6, the variables that contribute the most to the model correlation to cellulose conversion are reaction time and initial insoluble solids, while reactor type and the interaction terms contribute less than the model noise. The fact that time is correlated so strongly to conversion is not surprising, while the initial insoluble solids were found to be the strongest predictor of conversion in the previous example. The data again show that reactor type is less important than the other two variables, and again validates the findings of the t-test comparison and Model I ANOVA analysis. The Model II prediction using the three input variables is given in Figure 4.9, while the R-squared term for this model is 0.8698. This again shows that reactor type is approximately normally

distributed throughout the plot, indicating that this variable is not well correlated within the model. ANOVA analysis of this model shows that with >>99.99% confidence that the correlation described by the model is significant.

Considering all three statistical studies together, the implications are that the extent and rate of cellulose utilization are not necessarily dependent on the reactor scale, but more dependent on the ability of the reactor system to effectively mix the slurry and control its temperature, upon which the rate of enzymatic hydrolysis is strongly dependent. Furthermore, it was shown that as initial insoluble solids levels increase, it becomes increasingly more difficult to precisely control the temperature and that 15% initial insoluble solids represents the approximate upper limit for temperature control and mixing in the bench and mini-pilot scale STRs of the type used in this work.

4.5 Conclusions

The three major areas that this work examined were mixing and temperature control in STRs, the effect of agitation rate on enzymatic hydrolysis, and the issues involved in scaling up the enzymatic saccharification reaction from shake flasks to laboratory-scale STRs. The first part of the study found that approximately 10% to 13.5% insoluble PCS solids represents the maximum solids level where effective mixing and temperature control can occur in the stirred tank reactor systems tested. Additionally it was found that enzymatically hydrolyzed PCS can be used at higher insoluble solid levels (up to 16%) than unhydrolyzed PCS while still maintaining effective mixing and temperature control. The implications of this are that either continuous or fed-batch STR systems can be fed to ultimately achieve higher cumulative insoluble solids levels as saccharification proceeds and still remain mixable. However, the higher impeller speeds and/or power

requirements required to maintain adequate mixing at higher solids levels may be impractical.

The second portion of this work examined the effect of agitation rate (shear rate) on the rate and extent of cellulose hydrolysis during the saccharification reaction. Using two 4.5-inch diameter marine impellers in Bioflo 3000 fermentors and a range of impeller speeds of 100-400 rpm, no significant differences in the rate or extent of cellulose utilization were found. These experiments indicate that enzyme inactivation by impeller-induced shear is not a compelling concern.

The final part of this study investigated the effect of the vessel size/reactor system on saccharification performance. At insoluble PCS solids levels ≤10%, the Bioflo 3000 fermentor system achieved similar performance levels as obtained in shake flasks. As solids levels increased above 10%, saccharification rates and extents decreased in the STRs due to mixing problems, which resulted in difficulties in maintaining or reaching the desired reactor temperature set point. Although controlling the heating jacket temperature was effective for maintaining the reactor temperature, this resulted in slower reaction rates since reactor temperatures were usually slightly lower than the desired set point of 45°C. A decrease in the rate of cellobiose formation in the STRs relative to shake flasks at the higher solids levels provides evidence of this point. This work shows that these differences in performance between reactor types can be attributed to poor mixing and temperature control in specific, suboptimal STR configurations. Statistical analyses also confirmed that differences in shake flask reactors and STRs are not significantly different except at 15% initial insoluble solids. Overall, it can be concluded from the experiments summarized here that the apparent effects of scale upon the rate and extent

of cellulose hydrolysis result from differences in the effectiveness of mixing and temperature control between different reactor systems.

4.6 References

Aldenderfer MS and Blashfield RK. *Cluster Analysis*. Sage Publications, Beverly Hills, CA, 1984.

Bakker A and Gates LE. (1995). Properly Choose Mechanical Agitators for Viscous Liquids. *CEP* 91(12):25-34.

Bailey JE and Olis DF (1986). *Biochemical Engineering Fundamental, 2nd Ed.* McGraw-Hill, New York.

Deming SN and Morgan SL (1987). *Experimental Design: A Chemometric Approach.* Elsevier, New York.

Dowdy S and Wearden S (1983). *Statistics for Research.* Wiley & Sons, New York.

Ingesson H, Zacchi G, Yang B, Esteghlalian AR, Saddler JN (2001). The effect of shaking on the rate and extent of enzymatic hydrolysis of cellulose. *J Biotechnol.* 88:177-182.

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002). Microbiol Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol Mol Biol Rev* 66(3):506-577.

Mais U, Estaghlalian AR, Saddler JN (2002). Influence of Mixing Regime on Enzymatic Saccharification of Steam-Exploded Softwood Chips. *Appl Biochem Bioeng*. 98-100:464-472.

Makutaka S, Masazumi T, Takahashi J (1983). Effects of Agitation on Enzymatic Hydrolysis of Cellulose in a Stirred-Tank Reactor. *J Ferment Technol* 61:615-621.

Meunier-Goddik L and Penner MH (1999). Enzyme-Catalyzed Saccharification of Model Celluloses in the Presence of Lignaceous Residues. *J Agric Food Chem*. 47:346-351.

Mooney CA, Mansfield SD, Touhy MG, Saddler JN (1998). The Effect of Initial Pore Volume and Lignin Content on the Enzymatic Hydrolysis of Softwoods. *Biores Technol.* 64:113-119.

Nieves RA, Ehrman CI, Adney WS, Elander RT, Himmel ME (1998). Survey and analysis of commercial cellulase preparations suitable for biomass conversion to ethanol. *World J Microbiol Biotech*. 14:301-304.

Nomikos P and MacGregor JF (1994). Monitoring Batch Processes Using Multiway Principal Component Analysis. *AIChE J* 40:1361-1375.

Oldshue JY (1983). *Fluid Mixing Technology.* McGraw-Hill, New York.

Pimenova NV and Hanley TR (2003). Measurement of Rheological Properties of Corn Stover Suspensions. *Appl Biochem Biotechnol.* 105-108:383-392.

Reese ET and Ryu DY (1980). Shear inactivation of cellulase of *Trichoderma reesei. Enzyme Microb Technol.* 2:239-240.

Schell DJ, Farmer J, Hamilton J, Lyons B, McMillan JD, Sáez JC, Tholudur A (2001). Influence of Operating Conditions and Vessel Size on Oxygen Transfer During Cellulase Production. *Appl Biochem Biotechnol.* 91-93:627-642.

Schell DJ, Farmer J, Newman M, McMillan JD (2003). Dilute-Sulfuric Acid Pretreatment of Corn Stover in Pilot-Scale Reactor: Investigation of Yields, Kinetics, and Enzymatic Digestibilities of Solids. *Appl Biochem Biotech.* 104:69-85.

Tengborg C, Galbe M, Zacchi G (2001). Influence of Enzyme Loading and Physical Parameters on the Enzymatic Hydrolysis of Steam-Pretreated Softwood. *Biotechnol Prog.* 17:110-117.

Vinzant TB, Ehrman CI, Adney WS, Thomas SR, Himmel ME (1997). Simultaneous Saccharification and Fermentation of Pretreated Hardwoods. Effect of Native Lignin Content. *Appl Biochem Biotech.* 62:99-104.

Wilkens RJ, Henry C, Gates LE (2003). How to Scale-Up Mixing Processes in Non-Newtonian Fluids. *CEP* 99(5):44-52.

Zhang YHP and Lynd L (2004). Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. *Biotechnol Bioeng.* 88(7):797-824.

Chapter 5

SCALE-UP OF HIGH-SOLIDS SACCHARIFICATION OF PCS USING A FED-BATCH METHODOLOGY

5.1 Overview of Results

While many kinetic models have been developed for the enzymatic hydrolysis of cellulose, few have been extensively applied for process design, optimization, or control. This work utilizes both insights obtained from experimental work and kinetic modeling to develop an optimization strategy for high-solids cellulose saccharification, which presents a problem for the operation of stirred tank reactors (STRs) at insoluble solids levels greater than ~15%. In this work, a previously developed model for enyzmatic hydrolysis of cellulose was modified to consider the effects of feeding a stream of lignocellulosic solids and cellulase for a fed-batch process. By solving the set of model differential equations while controlling for a feed stream, it was possible to develop a feeding profile that controls the insoluble solids at a constant or manageable level throughout the course of the reaction. This prevents any difficulties that would otherwise be encountered in mixing high-solids slurries. Using this approach, feeding a stream of relatively concentrated pretreated solids can increase the final sugar concentration within the reactor without requiring the high initial levels of insoluble solids that would be required in batch mode. Experimental application of two fed-batch feeding policies in bench-scale STRs resulted in similar cellulose conversion profiles to batch shake flasks,

with final cellulose conversions reaching ~80% for fed-batch STRs fed at the equivalent of 25% initial insoluble solids.

5.2 Background

5.2.1 Models for Enzymatic Hydrolysis

Process modeling is ultimately motivated by one of two objectives. The first objective is to demonstrate an understanding of a physical process by proposing a model structure that suitably fits the experimental data. Another goal is to develop an application-based model for process design, simulation, or control. An application-based model could be either mechanistic or empirical. Many kinetic models of both types have been developed for the enzymatic hydrolysis of cellulose or cellulosic biomass to glucose and cellobiose (Kadam *et al.*, 2004; Chang and Holtzapple, 2000, Schell *et al.*, 1999; Philippidis and Hatzis, 1997; Gusakov and Arkady, 1992; Ooshima *et al.*, 1990 and 1991; Converse *et al.*, 1990; Wald *et al.*, 1984; Asenjo, 1984; Gusakov *et al.*, 1985; Ohmine *et al.*, 1983). Mechanistic models of cellulose hydrolysis are typically based on factors relating to both the substrate structure and the mechanisms/interactions of enzymes as the major limiting factors in the decrease in hydrolysis rate as time progresses (Mansfield *et al.*, 1999). Physical factors that have been used in models can include: adsorption of enzymes, sugar inhibition, synergy of enzyme components, enzyme-substrate mass transfer limitations, lignin distribution, inhibition by other components in biomass, particle size, and crystallinity. Empirical models are based solely upon experimental data and typically consider factors over a wide range of conditions that are more difficult to quantify mechanistically such as acetyl and lignin content, crystallinity (Chang and

Holtzapple, 2000) and enzymatic hydrolysis in the presence of complex pretreatment hydrolyzate liquors (Schell *et al.*, 1999).

5.2.2 Reactor Considerations for High-Solids

High-solids reactor configurations have been developed in the past for lignocellulosic conversion technologies in the areas of high-solids pretreatment (Zhu *et al.*, 2004; Schell *et al.*, 2003b; Nguyen *et al.*, 2000), high-solids saccharification or SSF (Varga *et al.*, 2004; Mohagheghi *et al.*, 1992; Spindler *et al.*, 1988), and solid-state or high-solids submerged culture cellulase production (Zhang *et al.*, 2003; Schell *et al.*, 2001; Tsao *et al.*, 2000). All of these technologies are motivated by favorable economics. Positive economic advantages associated with a high-solids saccharification process over a conventional low solids process include: lower capital costs due to the reduced volume; lower operating costs due to less energy required for heating and cooling; lower downstream processing costs due to higher product concentrations; reduced disposal and treatment costs due to lower water usage (Mohagheghi *et al.*, 1992). Techno-economic models have been used to approximate these cost savings. The work of Wingren *et al.* (2003) estimated that increasing the insoluble solids from 5% to 8% in an SSF process for bioconversion of softwood to ethanol reduces the operating costs by 19%. Recent design and modeling work at NREL considers a 2000 ton/day corn stover to ethanol plant using dilute-acid pretreatement followed by enzymatic hydrolysis and pentose and hexose fermentation (Aden *et al.*, 2002). Due to energy and capital costs, this model predicts a decrease in the minimum ethanol selling price (MESP) of $0.10/gal by increasing the total solids to the saccharification unit operation from 20% to 30%. Simulation predictions are depicted in Figure 5.1.

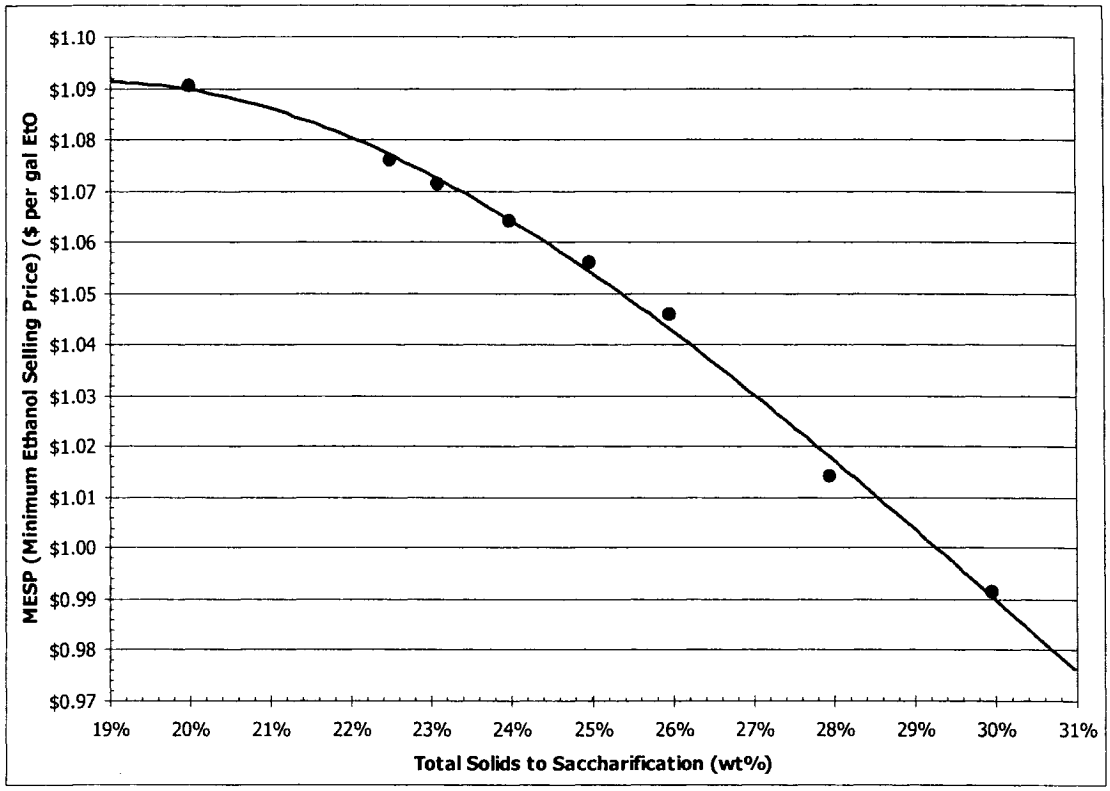


Figure 5.1: Prediction of the Aden *et al.* (2002) model showing impact of change in concentration of total solids to saccharification on MESP. Simulations courtesy of John Jechura.

High-solids slurries of pretreated lignocellulosic biomass are challenging to manage due to the increasingly high viscosities and non-Newtonian rheological properties (Pimenova and Hanley, 2003). As discussed in **CHAPTER 4**, this rheology results in increased problems with mixing (mass transfer) and temperature control (heat transfer). A variety of reactors have been used to deal with these problems during high-solids enzymatic hydrolysis. Examples include fermentation shake flasks (Varga *et al.*, 2004; Ballesteros *et al.*, 2002; Spindler *et al.*, 1988), the paddle-impeller reactor of Tengborg *et al.* (2001), a horizontally revolving reactor developed from a laboratory ball mill (Mohagheghi *et al.*, 1992), and a theoretical continuous tower reactor design (Nguyen, 1998). Currently, all of these laboratory-scale reactors require either a large reservoir of temperature controlled air (incubator) or water (water bath) to circumvent the problem of accurate temperature measurement and feedback control. Additionally, high power requirements for agitation may also be a limiting factor in many types of high-solids reactor.

Fed-batch processes combine the benefits of both batch and continuous processes to improve reaction conditions. A fed-batch methodology is commonly applied in fermentation for the purposes of minimizing effects of inhibitors, maximization of growth associated products, and induction of desired products with a feed stream (Hong, 1986). Several examples of fed-batch fermentation applied to biomass conversion technologies include the optimization of recombinant *Zymomonas mobilis* fermentation of glucose-xylose mixtures (Hodge and Karim, 2002) and both the pilot-scale and bench-scale fermentation of dilute-acid pretreated softwood hydrolyzates by *S. cerevisiae* (Rudolf *et*

al., 2004; Nilsson *et al.*, 2002; Tehzadeh *et al.*, 2000), which were performed in fed-batch to minimize the effect of inhibitory compounds in the hydrolyzate. Since HMF and furfural can be biologically detoxified by *S. cerevisiae* over time, fed-batch fermentation can result in slurries with higher final equivalent levels of these inhibitors than could be tolerated by the microorganism if the entire slurry were added initially.

Several approaches have been applied in the past for the fed-batch enzymatic saccharification of cellulose, although all of these were based on *ad hoc* approaches to feeding rather than a rigorous design methodology. These process-derived motivations fall into three general categories, the first of which is enzyme recycle. Pristavka *et al.* (2000) used an approach similar to fed-batch for high-solids saccharification of steam-exploded willow. For this, a high-solids slurry was saccharified, the liquid phase removed, and more solids added to the solids fraction such that a large portion of enzyme adsorbed to the solids could be reused. Total sugar concentrations averaged 120 g/L including glucose and hemicellulosic sugars from pretreatment. Another enzyme recycle approach is to add only pretreated solids but no enzymes. As the cellulose is hydrolyzed, enzymes should be released to back into liquid such that more cellulosic biomass could be added to reactor as these enzymes are released.

The second category uses fed-batch SSF (Söderström *et al.*, 2004; Wingren *et al.* 2004) to mitigate the inhibitory effects of compounds in hydrolyzate liquors on the fermentative microorganism, which can be considered as an extension of the motivation for fermentation-only fed-batch of hydrolyzates. Effective insoluble solids were also increased from 3% to 5% or to 8.4% for the two studies, respectively. Borden *et al.*

(2000) also used a fed-batch approach to mitigate the effects of inhibitors in the hydrolyzate during SSF to acetic acid by *Clostridium thermoaceticum*.

The third motivation for fed-batch saccharification is to utilize feeding to increase the equivalent insoluble solids level during saccharification to overcome rheological limitations in the reactor as in the work of Ballesteros *et al.* (2002) for SSF by *Kluyveromyces marxianus*. In this work, the insoluble solids were increased from 10% to 20%, which the reactor would have been incapable of handling initially, as well as reducing the effects of inhibitors in the hydrolyzate. Due to problems encountered in mixing wet oxidation pretreated corn stover at insoluble solids greater than 12%, Varga *et al.* (2004) used a fed-batch hybrid hydrolysis and fermentation (HHF) approach to achieve 17% insoluble solids with high ethanol yields. Mohagheghi *et al.* (1992) compared batch to fed-batch SSF and determined that the feeding of various combinations of pretreated biomass, media nutrients, and fresh inocula did not improve either ethanol yield or productivity in high-solids SSF for the ball-mill reactor system used.

5.2.3 Mathematical Approaches to Fed-Batch

For a more robust approach to developing a fed-batch process, Kleman *et al.* (1991) identify four open- and closed-loop control options for developing a feeding policy based on controlling substrate levels for fed-batch fermentations. Listed in order of increasing accuracy and subsequent complexity these include:

1. Open-loop control schemes (off-line control) where a feeding policy is developed based on either model predictions or historical data (Hong, 1986).

2. Indirect closed-loop control (on-line control) based on secondary variables indirectly correlated to the feeding variables. For typical bioreactors this includes variables such pH, oxygen uptake rate, and concentrations of organic products (Rudolf *et al.*, 2004; Tehzadeh *et al.*, 2000).
3. Indirect closed-loop control (on-line control) based on mass balance equations using values calculated from on-line sensors (Nilsson *et al.*, 2002).
4. Direct closed-loop control (on-line control) based on on-line substrate measurements (Hodge and Karim, 2002).

The indirect closed-loop control approach to fed-batch has been applied to the fermentation of dilute acid pretreated hydrolyzate liquors of hardwoods and softwoods based on measurement of off-gas composition and calculation of the of the CO_2 evolution rate (CER). Using this variable, control strategies have been developed that determine changes in feed rate based on changes in the CER (Rudolf *et al.*, 2004; Tehzadeh *et al.*, 2000), and to estimate and control the sugar concentration based on this variable (Nilsson *et al.*, 2002). The direct closed-loop approach to fed-batch control used in our previous work (Hodge and Karim, 2002) was based on HPLC measurements of reactor conditions coupled to a model predictive controller using a kinetic model of the process.

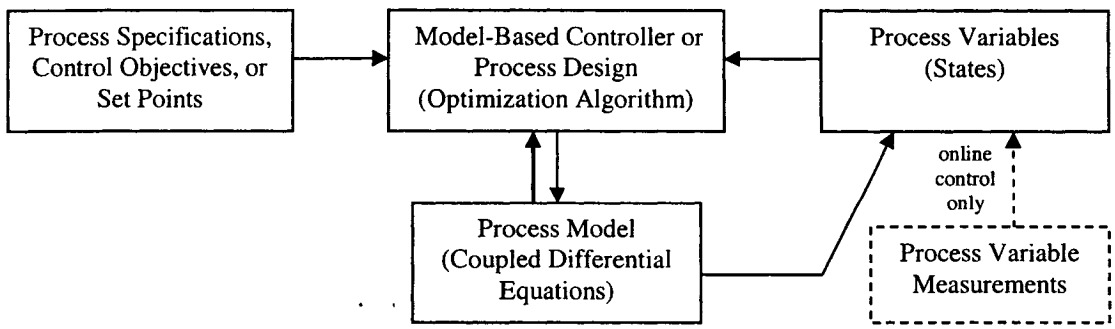


Figure 5.2: Development of Open-Loop Control Configuration for Model Process Design

Since open-loop control schemes rely solely on process models or speculation on the physical parameters governing the process, these control methods are only as good as the model. In order to develop only a feeding policy without requiring online measurements of insoluble solids, an open-loop approach can be used. This doesn't rely on feed-back from process measurements, but rather is based only on the model predictions. An overview of this approach is given in Figure 5.2.

Using a system of model differential equations, an exact analytical solution for a feeding profile can be obtained by developing and solving an "optimal control" problem. Optimal control fed-batch operation has been applied to cell cultures in bioreactors for a variety of objectives such as to maximize the production of cell mass, penicillin, amino acids, enzymes, and bioengineered products such as penicillin, ethanol, and recombinant proteins (Lee *et al.*, 1999; Hong, 1986; Modak *et al.*, 1985). There are two general classes of methods to determine a numerical solution to this problem (Griesse and Walther, 2004). The first of these applies Pontryagin's maximum principle for determining the necessary conditions for the optimizer, which takes the form of a multi-point boundary problem for state and adjoint equations. The points at which the solution enters the boundary of the set of feasible conditions must be known *a priori* resulting in either singular or nonsingular control schemes. The second class of methods involves the discretization of the optimal control problem into a nonlinear programming (NLP) problem (Griesse and Walther, 2004), which is solvable by optimization methods. For this method, as the number of time steps is increased, the calculated profile approaches the continuous optimal profile (Lee *et al.*, 1999). While both of these methods are open-

loop control schemes, the second class of optimal control problem can also be implemented in combination with online variable measurements to yield a model predictive control (MPC) problem (Zheng, 2000; Rawlings and Muske, 1993; Bequette, 1991).

A general form for the continuous-time nonlinear system that is approximated by the model system in the optimal control problem can be described by the following system of equations as:

$$\dot{x} \equiv \frac{dx}{dt} = f(x(t), u(t)) \tag{5.1}$$

where $x(t)$ is the vector of state variables and $u(t)$ contains the controlled variables. A continuous form for the objective function for this algorithm seeks to minimize the difference between the desired process trajectory and the predicted process trajectory giving the minimization algorithm as:

$$\min_{u(t)} \Phi(x) = \int_0^{t_f} [r(t) - x_i(t)]^2 \, dt \tag{5.2}$$

subject to:

$$\begin{cases} \dot{x} = f(x(t), u(t), t) \\ t \in [0, t_f], x(0) = x_0 \end{cases}$$

where $r(t)$ is the desired output trajectory of the controlled variable and $x_i(t)$ is the controlled variable. The optimal path for $u(t)$ can be determined by reformulating the problem in terms of a Hamiltonian equation and developing a set of adjoint equations in order to solve the optimization problem under the conditions of the Pontryagin maximum

principle (Gamkrelidze, 1978). Another approach is to discetize this continuous form of the objective function by subdividing the time region into K equal intervals as:

$$\Phi = [\phi_1, \ldots, \phi_{k-1}, \phi_k, \ldots, \phi_K] \tag{5.3}$$

and the discretized form of the minimization algorithm generated becomes a series of optimization tasks as:

$$\min_{u(k)} \phi_k = [r(k) - x_i(t_k)]^2 \tag{5.4}$$

subject to:

$$\begin{cases} \dot{x} = f(x(t), u(k), t) \\ t \in [t_{k-1}, t_k], x(t_{k-1}) = x_{k-1} \end{cases}$$

Using this approach, only K values for the control variable *u(k)* are required, and these are fixed over each time interval. This results in a straightforward optimization problem.

5.3 Materials and Methods

5.3.1 Shake Flask CSS

Corn stover was pretreated with dilute sulfuric acid in NREL's Sunds vertical pulp digester as the batch designated PS030312CS. The solids are composed of 53.2% cellulose by dry mass, and are washed with >10 times the mass of DI water as solids and centrifuged at 15800 x g. Washed solids are used to minimize any inhibition effects from background sugars entrained in the solids. The final glucose concentration in the wash water is verified to be less than 0.1 g/L.

The cellulase / β-glucosidase enzyme mixture used for this study was Spezyme CP (Lot # 301-00348-257, Genencor International). Total protein was assayed at 106.6 mg/mL (BCA assay, Pierce Biotechnology, Inc.) and the activity was determined to be 0.27 FPU/mg of protein using the NREL protocol LAP-006. For the modeling work,

cellulase : β-glucosidase is assumed to be 0.975:0.025. 1 M citrate buffer is prepared by adjusting the pH to 4.8 with NaOH, and is used as 5% of the total mass to yield an effective molality of 0.05 mol/kg. Enzyme, DI water, and citrate buffer are filter sterilized before use. PCS is added to the shake flasks and weighed before autoclaving. After autoclaving the shake flasks are weighed again so that any water lost could be made up. Water loss to evaporation in addition to reactor mass changes associated with feeding and sampling are tracked by weighing each flask before and after sampling and feeding.

5.3.2 Bench-Scale STR Fed-Batch Saccharification

Bench-scale fed-batch saccharification is performed using the same PCS substrates and enzymes as the shake flask reactors. Washed PCS is air-dried at ambient conditions to ~45% insoluble solids and the insoluble solids content is determined in triplicate using the 105°C oven. New Brunswick Bioflo 3000 (7 L vessel size) reactors are autoclaved empty and allowed to cool in a sterile biological hood. The appropriate amounts of solids are measured and autoclaved and the amount of water lost is determined. Solids and other sterile components are next added to the reactors and allowed to mix and reach steady-state temperature before enzyme is added. Feeding policies are determined based on the solution to the optimal control problem as outlined in the **Results and Discussion** (Section 5.4). Feeding is performed by adding both autoclaved solids and filter-sterilized enzyme at discrete times. The solids and enzyme feeding policy is determined from the model kinetic equations in MATLAB to control the reactors using two feeding policies (F1 and F2). For F1, the actual insoluble solids level is increased from 12% to 15% as the reaction proceeds proportionally to the sugar produced. For F2 the insoluble solids are maintained at 15% insoluble solids. For both cases the feeding is continued until the

reactors reach the equivalent of 25% initial insoluble solids. HPLC measurements for the reactor are taken both prior to and after feeding. Temperature is maintained at 45°C by water circulating to a heating jacket from a temperature controlled water bath (maximum T = 52°C). Enzyme loading is 40 mg protein / g cellulose (10.7 FPU / g cellulose), and enzyme is fed proportionally to the amount of PCS solids added. This can be considered as equivalent to the same amount of enzyme required for a given amount of PCS as would be used in an equivalent batch reaction. New Brunswick reactors are insulated and equipped with two marine impellers (d/D = 0.68) maintained at 400 rpm. From the work presented in **CHAPTER 4**, this high impeller speed did not cause shear-induced inactivation of cellulases. Cellulose conversion for fed-batch reactors is determined based on the percentage of the final cumulative cellulose converted to glucose and cellobiose rather than only the initial or currently fed cellulose.

5.3.3 Sampling and Data Analysis

For the batch mode reactors, samples are taken approximately every 24 hours. Samples are removed with 5 mL pipettes (Falcon, Inc.) that had the tips broken off in order to accommodate the high level of insoluble solids. These are transferred to 2.5 mL microcentrifuge tubes and the pH is measured. The samples are next centrifuged at 10000 rpm for 5 minutes. The supernatant is decanted and diluted 1:5 and syringe filtered into HPLC vials for subsequent HPLC analysis. HPLC analysis is performed on all saccharification samples to determine the glucose, xylose, and cellobiose concentrations. An Agilent 1100 series HPLC equipped with a differential refractive index detector is used. This is equipped with a Bio-Rad HPX-87H organic acid column operating at 55°C with a 0.01 N H_2SO_4 mobile phase at a flow rate of 0.6 mL/min.

Component concentrations obtained from HPLC measurement are converted to g/kg concentrations based on the estimated insoluble solids level in the reactor using the method outlined in **CHAPTER 2.** For fed-batch and CSS, cellulose conversions are based on cellulose converted to glucose and cellobiose, while for SSF conversions are based on cellulose converted to glucose, cellobiose, and ethanol.

5.4 Results and Discussion

5.4.1 Mass Balance Equations for Fed-Batch Prediction

Mass balances can be performed on the reaction system such that the limitations or potential of using a fed-batch approach can be evaluated. This approach assumes that the insoluble solids (S_I) are fed at the same rate as they are removed resulting in a constant level of insoluble solids during the feeding ($S_I(t_0) = S_I(t_{final})$). Performing a mass balance on the insoluble solids and normalizing all volumetric values with respect to the initial reactor volume results in:

Solids fed + Initial Solids = Final Solids

$$\frac{F}{V_0}(1 - x_{cell}\xi)S_{IF} + (1 - x_{cell}\xi)S_I = (1 + \frac{F}{V_0})S_I \tag{5.5}$$

with the solution for the feeding rate:

$$\frac{F}{V_0} = \frac{S_I x_{cell}\xi}{(1 - x_{cell}\xi)S_{IF} - S_I} \tag{5.6}$$

and the equivalent solids (S_{equiv}) can be determined as:

$$S_{equiv} = \frac{S_{IF}\frac{F}{V_0} + S_I}{\frac{F}{V_0} + 1} \tag{5.7}$$

where:

x_{cell} = cellulose fraction in the solid phase

ξ = Extent of cellulose conversion

S_I = Insoluble solids (g solids/kg)

S_{IF} = Insoluble solids in the feed (g solids/kg)

F = Feed mass flow rate (kg/h)

V_0 = Initial working "volume" of the reactor (kg)

It should also be considered that the solutions to these equations are constrained only by the physical limitations of the reaction system. Since these results are determined independent of reaction rate, some of the solutions obtained may be physically unrealizable due to the constraints of reaction rate. Using these mass balance equations, the plots in Figures 5.3 and 5.4 were developed. Figure 5.3 shows how feeding can increase the level of equivalent insoluble solids for constant conversion contours; up to 27% at 85% conversion by maintaining a reactor at 15% insoluble solids. Another important consideration for the fed-batch reactor is the change in reactor volume associated with the feeding of solids. Figure 5.4 depicts how the most important parameter for the change in reactor volume, S_{IF}, affects the reactor volume for feeding as given in Figure 5.3. The major result of this is that both the reaction volume and the level to which reactors can be fed would be limited by low values of S_{IF}.

5.4.2 Development of Optimal Control Problem

The mechanistic model of Kadam *et al.* (2004) was used in this study. This model is based on the enzymatic hydrolysis of dilute acid pretreated corn stover (PCS) and

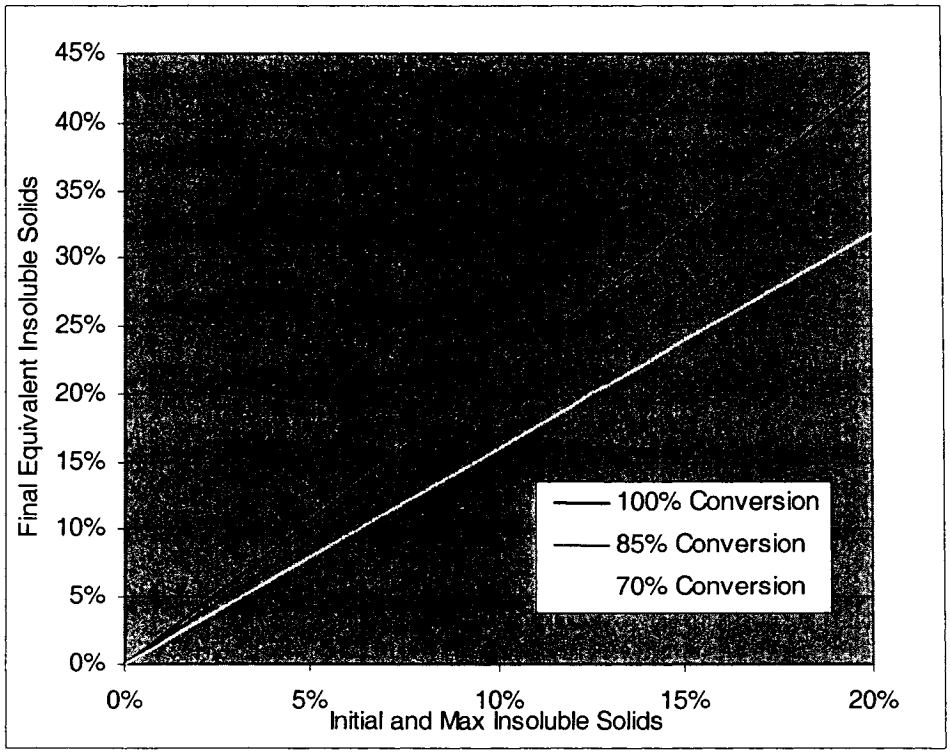


Figure 5.3: Mass balance predictions of the solids feeding capabilities of fed-batch reactors.

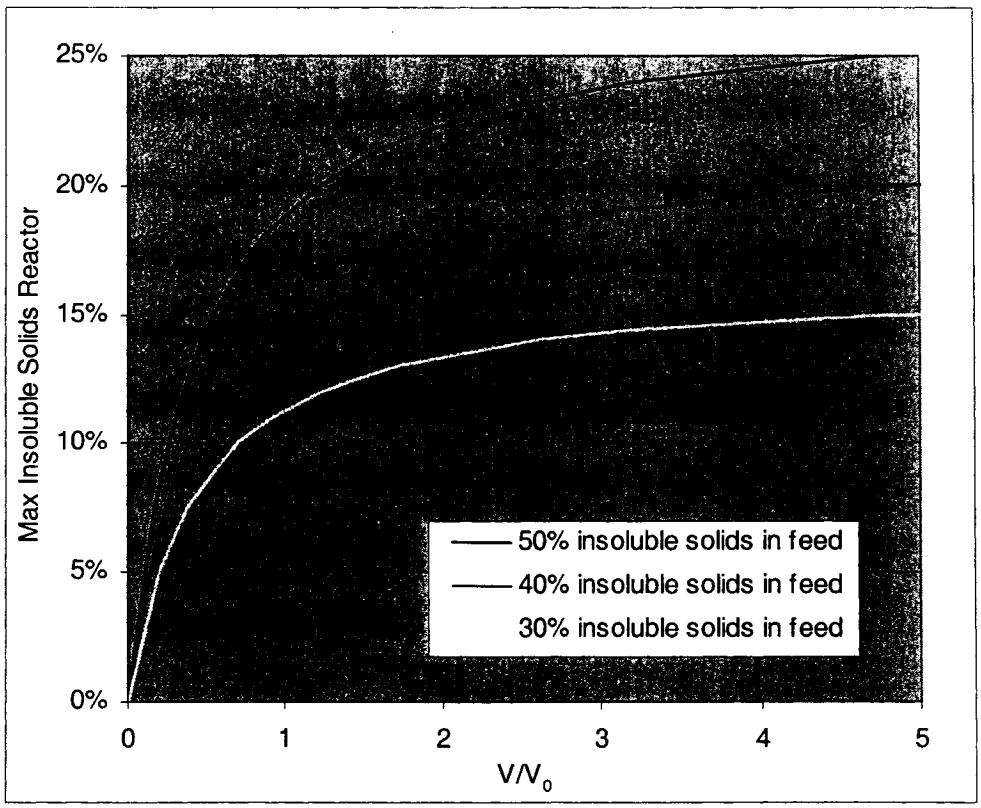


Figure 5.4: Mass balance predictions of the reactor volume changes required for fed-batch operation as in Figure 5.3.

considers the adsorption of β-glucosidase and that of CBH/EG enzymes separately while including an inhibition term for xylose. The kinetic model was implemented in MATLAB, and was modified to consider changes in volume for fed-batch operation, as well as introducing several new state variables. New state variables that need to be added include insoluble solids (determined from Equation 2.8 in **CHAPTER 2),** enzyme concentration, and total reactor working mass. In addition to these new variables, all of the other differential equations needed modification to account for the rate changes associated with feeding and the change in reactor working mass for fed-batch operation. A new variable that needs to be introduced for systems where reactor working mass and/or concentrations is changing due to a feed stream is the dilution rate (D). This term can be defined as the feed flow rate / reactor volume, and can be considered the inverse of the reactor residence time. For fed-batch reactor configurations, the modified model equations are given as follows:

State Variable	Equation	Units	Eq. #
Insoluble solids	$\frac{dS_I}{dt} = D \cdot (S_{IF} - S_I) - r_1 - r_2$	[g/kg/h]	(5.8)
Cellulose	$\frac{dC}{dt} = D \cdot (C_{feed} - C) - r_1 - r_2$	[g/kg/h]	(5.9)
Glucose	$\frac{dG}{dt} = D \cdot (G_{feed} - G) + 1.111 \cdot r_1 + 1.053 \cdot r_3$	[g/kg/h]	(5.10)
Cellobiose	$\frac{dCB}{dt} = D \cdot (CB_{feed} - CB) + 1.056 \cdot r_1 - r_3$	[g/kg/h]	(5.11)
Xylose	$\frac{dX}{dt} = D \cdot (X_{feed} - X)$	[g/kg/h]	(5.12)
Enzyme 1 (all EGs and CBHs)	$\frac{dE_1}{dt} = D \cdot (E_{1\,feed} - E_1)$	[g/kg/h]	(5.13)
Enzyme 2 (β-glucosidase)	$\frac{dE_2}{dt} = D \cdot (E_{2\,feed} - E_2)$	[g/kg/h]	(5.14)
Volume	$\frac{dV}{dt} = F$	[kg/h]	(5.15)

where:
r_1 = Rate of cellulose conversion to glucose (g/kg/h)
r_2 = Rate of cellulose conversion to cellobiose (g/kg/h)
r_3 = Rate of cellobiose conversion to glucose (g/kg/h)

Variables:
G = Glucose mass concentration (g/kg)
G_{feed} = Glucose mass concentration in the feed (g/kg)
X = Xylose mass concentration (g/kg)
X_{feed} = Xylose mass concentration in the feed (g/kg)
CB = Cellobiose mass concentration (g/kg)
CB_{feed} = Cellobiose mass concentration in the feed (g/kg)
C = Cellulose mass concentration (g/kg)
C_{feed} = Cellulose mass concentration in the feed (g/kg)
V = Working "volume" of the reactor (kg)
D = F/V (h^{-1})
E_1 = Enzyme 1 (all EGs and CBHs) mass concentration (g/kg)
E_2 = Enzyme 2 (β-glucosidase) mass concentration (g/kg)
$E_{1\,feed}$ = Enzyme 1 (all EGs and CBHs) mass concentration in the feed (g/kg)
$E_{2\,feed}$ = Enzyme 2 (β-glucosidase) mass concentration in the feed (g/kg)

An illustrative summary of these reaction rates is given in Figure 5.5. All of the expressions for rate terms used in this model and the parameters within the rate terms can be in found in Kadam *et al.* (2004). It is important to note that enzyme inactivation is not considered in this model, and should be more significant for fed-batch operation when there is a large disparity between the ages of the enzymes and fractional utilization of substrates in the reactor. Additionally, the substrate reactivity (R_S) term is not utilized due to the problems outlined in **CHAPTER 3,** and since the substrate age varies with feeding. It should also be considered that the model was developed using CPN cellulase (Iogen Corp., Ottawa, Ontario), while this experimental work used Spezyme CP (Genencor International Corp., Palo Alto, CA).

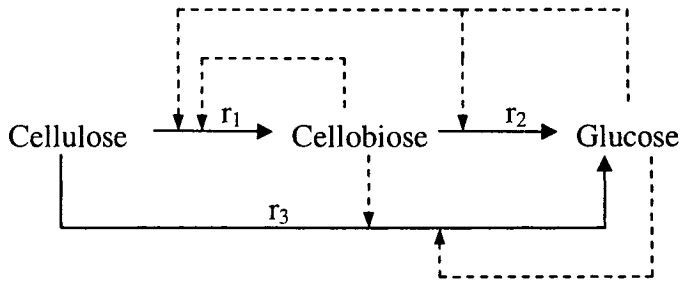


Figure 5.5: Rates for the enzymatic hydrolysis of cellulose used in Equations 5.8 through 5.11. Dashed lines indicate sugar inhibition of these rates.

5.4.3 Fed-Batch Saccharification Model Simulation

In order to develop a fed-batch saccharification process, variables and parameters that are affected by the feeding and a strategy on which the optimization can be based must be identified. Many options can be considered for the development of a fed-batch strategy, and some of those applied in the past are discussed in the **Background** (Section 5.2). Table 5.1 gives an overview of some of these considerations in which independent variables can be considered as controlled variables that can be manipulated in order to achieve a control objective (optimization parameters).

Table 5.1: Model-Based Process Design for Fed-Batch Saccharification

Independent Variables for Fed-Batch Model-Based Control:

- Feed Rate (kg/h)
- Duration of Feeding (h)
- Insoluble Solids Levels in Feed (g insoluble solids / g feed)
- Initial Insoluble Solids Level in Reactor (g insoluble solids / g feed)
- Enzyme Loading in Feed (mg protein / g cellulose)
- Initial Enzyme Loading in Reactor (mg protein / g cellulose)

Optimization Parameters (Control Objectives):

- Reactor Conditions
  - – Insoluble Solids Levels
  - – Inhibitory Compound Levels
  - – Sugar Levels
- Chemical Reaction
  - – Cellulose Conversion (Extent of Reaction)
  - – Enzyme Levels

One barrier to developing a high-solids saccharification process in stirred tank reactors (STRs) is rheological limitations manifested as problems in mass and heat transfer. The work outlined in **CHAPTER 4** demonstrated that the saccharification reaction is scalable from the shake flask to bench scale STRs at insoluble solids levels of 15% or less using PCS solids. Above this level, STRs encounter problems with mixing and temperature control, which is primarily correlated to the amount of free water in the slurry. Based on this finding, the fed-batch approach control objective is to optimize reactor conditions (insoluble solids levels) to facilitate sufficient mixing and temperature control. A large number of variables exist (Table 5.1) that can be altered to achieve this objective. During the process of feeding, the feed rate is the only variable that is free to be changed.

Using the kinetic model, a feeding policy can be developed based on controlling the insoluble solids below a defined critical value during the saccharification reaction. This is possible by feeding a stream of PCS solids and cellulase at a rate that approximately matches the rate at which cellulose in the reactor is depolymerized and solubilized. Using the model equations, the rate of change for insoluble solids can be determined based on a set of initial operating conditions and a feed rate.

To determine a solution for the feeding policy, the set of equations must be integrated over time with the value for feed rate free in order to satisfy the control objectives of insoluble solids. For this, Equation 5.8 can be either set to 0 as:

$$\frac{dS_I}{dt} = D \cdot (S_{IF} - S_I) - r_1 - r_2 = 0 \qquad (5.16)$$

or set the insoluble solids level (S_I) to a specified trajectory $(S_{I,SP}(t))$ as a function of the sugar level:

$$S_{I,SP}(t) = f(G, CB, S_I) \tag{5.17}$$

since sugar level can be correlated to the substrate conversion and consequently to slurry rheology. To achieve the control objectives of both Equations 5.16 and 5.17, the objective function of the optimization problem of Equation 5.4 becomes:

$$\min_{F(k)} \phi_k = \left[S_I(k \mid k-1) - S_{I,SP}(k)\right]^2 \quad . \tag{5.18}$$

while the constraints become the system of model equations in Equations 5.8 through 5.15.

The model is implemented in the MATLAB script *sacchmodel4.m* while the value for the feed over each time interval is determined in an optimization routine in the script *optimfunction.m*, and the overall feeding profile is generated in the script *mpc.m*. All of these MATLAB scripts are located in **APPENDIX 1**. The optimization routine solves the model differential equations over a defined minute time interval and determines a value for the feed rate for each discrete time interval that minimizes the objective function. The constrained nonlinear minimization function from the MATLAB optimization toolbox *fmincon.m* used in *optimfunction.m* utilizes the standard Newton's method for optimization with the BFGS method for Hessian matrix updates and the mixed quadratic/cubic polynomial line search process.

Using this optimal control algorithm, fed-batch feeding policies can be developed, and a related set of feeding curves over various process objectives and initial conditions can be generated to determine the theoretical physical limitations and potential for using this type of fed-batch approach. Ultimately, physical limitations exist that limit the equivalent insoluble solids level to which a reactor under given initial conditions can be fed as was discussed previously. The two important variables for this are reactor volume

(V, or in dimensionless form as V/V_0) and insoluble solids level in the feed (x_{sf}). As the reactor is fed, the level of equivalent initial insoluble solids increases. The cumulative or equivalent initial insoluble solids is defined as the level of initial insoluble solids that would be present if all of the solids were added initially and the reactor was operated in batch mode. Based on this variable, fed-batch performance can be compared to batch reactor performance.

Simulations were performed for batch and fed-batch using an enzyme loading of 40 mg protein / g cellulose (10.7 FPU / g cellulose for Spezyme CP). Figure 5.6 shows four feeding policies developed from the MATLAB simulation for 45% insoluble solids in the feed while controlling insoluble solids at a constant level by feeding. This illustrates how the equivalent insoluble solids are affected by the feeding to maintain the reactor at a given level of solids in the reactor. An important result from this figure is that the model predicts that controlling insoluble solids at levels below 15% results in unreasonably long times (>300 h) required to complete the feeding. However, previous scale-up work in **CHAPTER 4** showed 15% insoluble solids is near the upper limit for effective control of temperature.

Figure 5.7 gives the predicted conversion profiles for batch (25% initial insoluble solids) and the fed-batch conversion profile for maintaining the insoluble solids at 15% as in the profile in Figure 5.6. While this shows batch saccharification as capable of reaching 80% cellulose conversion in about 66 h, such rapid rates are physically unrealistic and emphasize the limitations of kinetic models in their predictive power. There are two reasons for this higher predicted rate in batch operation. The first is that

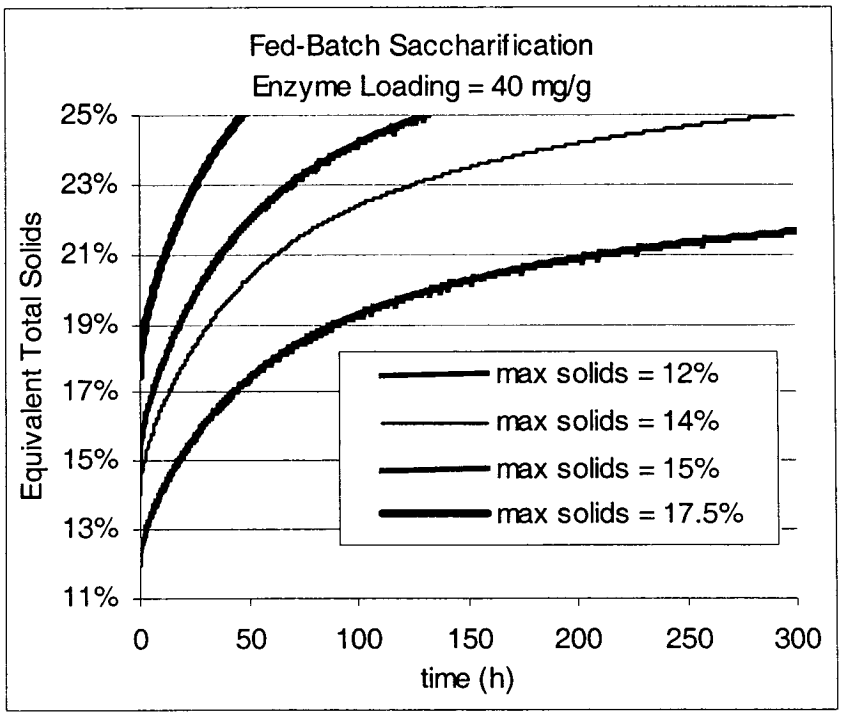


Figure 5.6: Equivalent total solids based on calculated feeding profiles for fed-batch operation at an enzyme loading of 40 mg/g and 45% insoluble solids in the feed stream

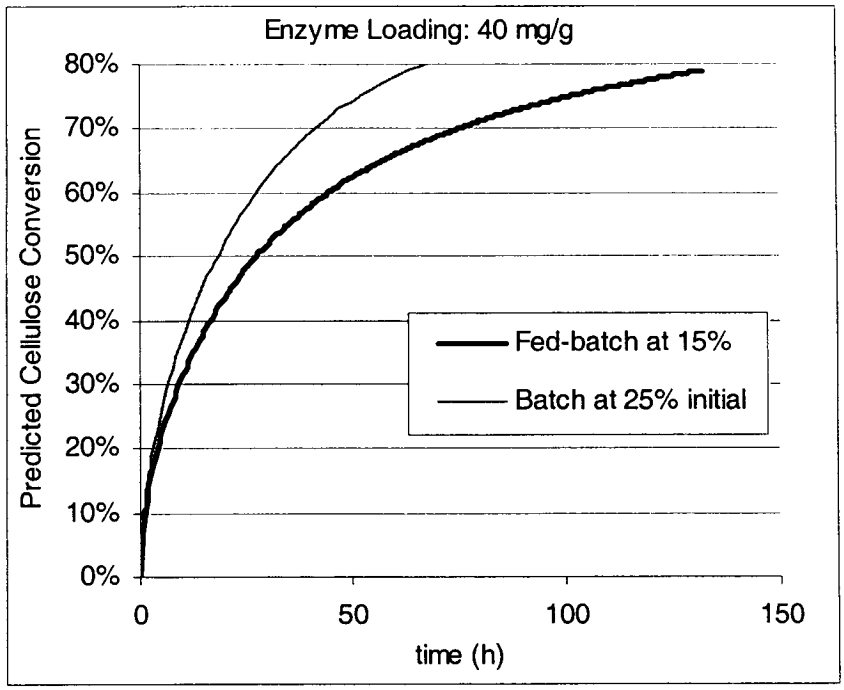


Figure 5.7: Comparison estimated cellulose conversion in batch and fed-batch reactors. Fed-batch uses control of insoluble solids at 15%.

initially batch has higher concentrations of cellulose, and since the rates are first order in cellulose this will result in faster predicted hydrolysis rates. The other reason is that in fed-batch, the enzymes and cellulose have shorter reactor residence times relative to batch operation.

Figure 5.8 gives the results of fed-batch simulations for these same feeding requirements that show the effect of the moisture level of the insoluble solids in the feed on cellulose conversion, reactor volume, and feeding duration. These results show that the insoluble solids level in the feed has a strong impact on the rate of conversion and especially the time required to attain an equivalent level of insoluble solids. The simulation results in Figure 5.8B show that higher solids levels in the feed result in both smaller reactor volumes and shorter residence times to achieve the same feeding objectives. The reason for this is that feeding a lower solids stream adds more water and is effectively diluting the product in the reactor.

Another option for a feeding policy is to use a non-constant insoluble solids level for the control trajectory. This could be based on the finding from **CHAPTER 4** that cellulose removal affects the mixing behavior in STRs by allowing higher levels of glucan-depleted solids to be mixed effectively relative to the same level of glucan-rich solids. For this scheme, insoluble solids are initially maintained at 12% and are allowed to increase to about 15% during feeding as the equivalent cumulative solids approach 25%. The rate of increase in the insoluble solids set point ($S_{I,\,SP}$) is proportional to the amount of glucose produced. Estimating the final sugar concentration when the feeding was complete resulted in the development of the following empirically-derived relation:

$$S_{I,\,SP} = S_{I0} + (G + CB) / 4800 \qquad (5.19)$$

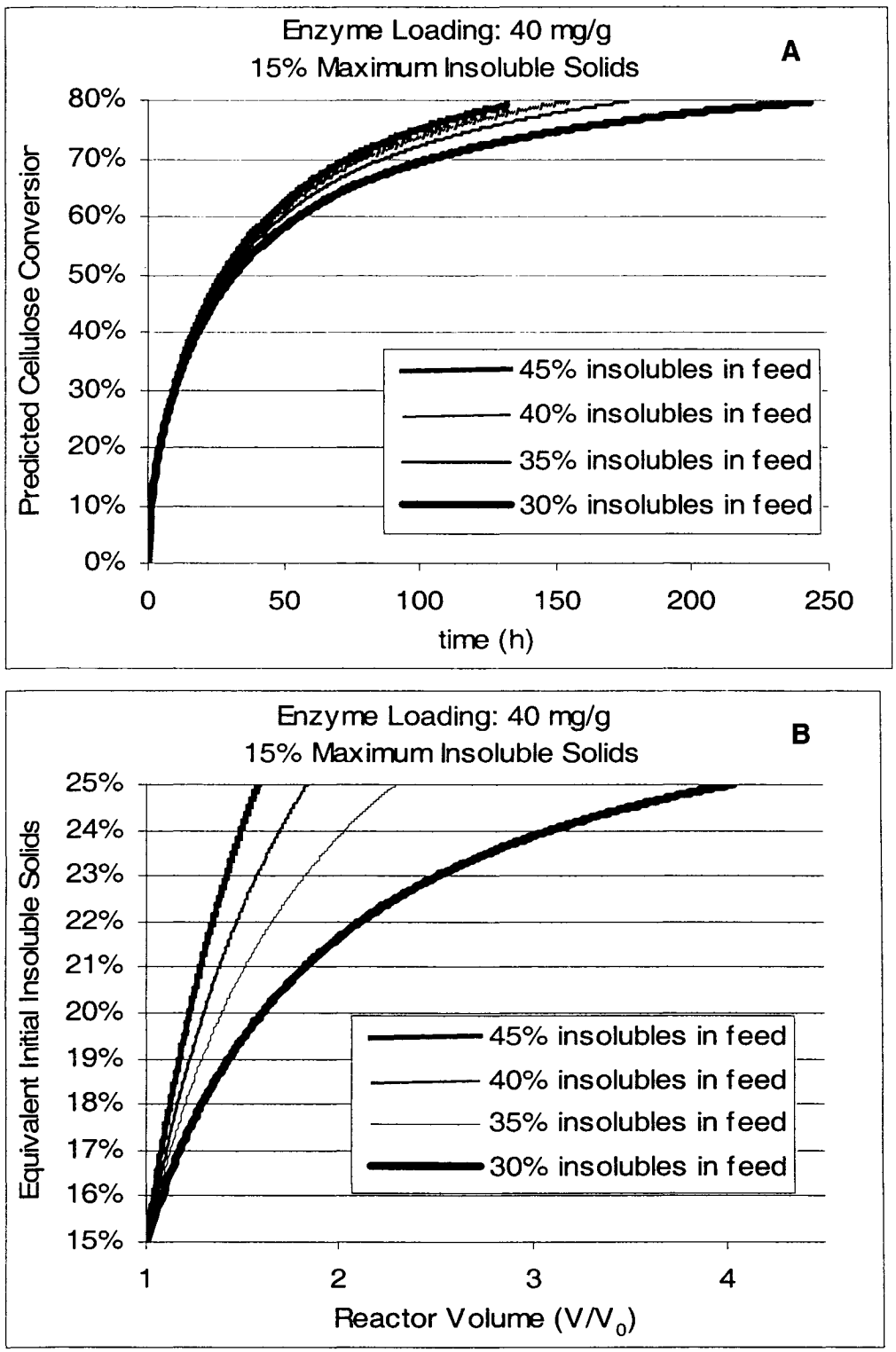


Figure 5.8: Effect of insoluble solids in the feed on reactor feeding variables; Reactors are fed to maintain insoluble solids at 15%

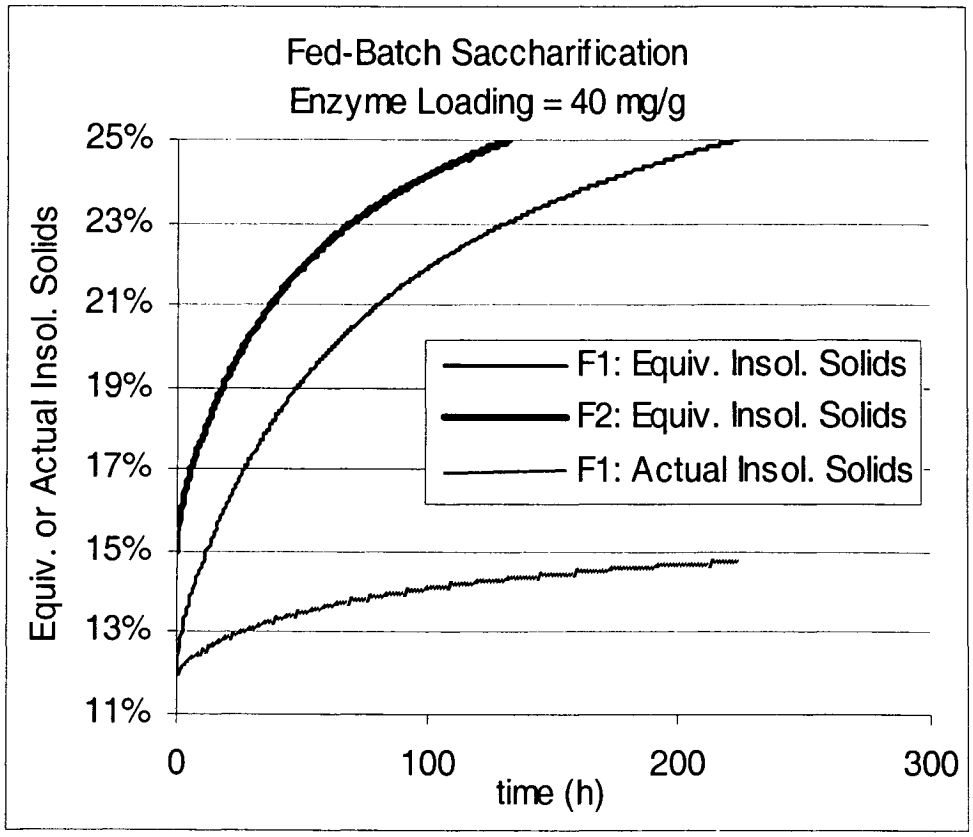


Figure 5.9: Equivalent total solids based on calculated feeding profiles for fed-batch operation at an enzyme loading of 40 mg/g and 45% insoluble solids in the feed stream

This approach (F1) was one of the two feeding policies selected for experimental validation of the modeling results. The other (F2) was based on controlling the insoluble solids at 15%, which was shown in simulation to have relatively rapid performance for 45% insoluble solids in the feed. Figure 5.9 shows the two feeding policies (F1 and F2) chosen for experimental testing in the bench-scale STRs (New Brunswick Bioflo 3000). As seen in Figure 5.9, by maintaining the insoluble solids level at 15% as in F2, solids can be fed up to 25% equivalent within a reasonable timeframe. Important characteristics of the feeding policies are outlined in Table 5.2.

Table 5.2: Fed-Batch Reactor Feeding Policies

	Batch Shake Flasks (SFs)	Fed-Batch Reactor 1 (F1)	Fed-Batch Reactor 2 (F2)
Initial insoluble solids	25%	12%	15%
Insoluble solids at completion of feeding	----	15%	15%
Equivalent insoluble solids at completion of feeding	25%	25%	25%
Time required to complete feeding	----	216 h	120 h
Initial mass charged to the reactor	100 g	2503 g	2789 g
Final mass in the reactor	100 g	4500 g	4500 g

5.4.4 Experimental Validation of Theoretical Results

The two feeding policies chosen in Figure 5.9 were performed experimentally in the bench-scale STRs. Batch saccharification shake flasks are also performed as controls in triplicate at 25% initial insoluble solids under similar conditions to determine maximum saccharification potential. A summary of the 3 reactor configurations is given in Table

5.2. The shake flask data is meant to be applied as a "best case" condition for these reaction conditions since these can be assumed to have demonstrated uniform temperature throughout the saccharification due to the reaction being performed in an air-temperature controlled shaking incubator. However, concentration gradients and rate limitations due to problems with diffusion of sugars and enzymes cannot be ruled out in shake flasks. Figure 5.10 shows the estimated insoluble solids for both feeding policies as a function of time, and demonstrates that both reactors are capable of reaching ~80% cellulose conversion in almost the same time required in shake flasks. Both fed-batch reactors were able to maintain low levels of insoluble solids such that the control of temperature was not a problem.

Figure 5.11 shows that the glucose and conversion time profiles for both F1 and F2 meet the profiles for the batch shake flask reactors at approximately the end of the feeding, indicating that this fed-batch feeding policy will not require significantly higher residence times than a batch reaction. The "sawtooth" pattern for the fed-batch reactors is due to the change in volume and subsequent dilution of reactor components after addition of the feed. The reactor using feeding policy F1 is slower due to the slower feeding rate, while F2 shows more rapid rates as was predicted in simulation. From this work it was concluded that the fed-batch approach could allow STRs to saccharify the equivalent of 25% insoluble solids without significantly different residence times than shake flask reactors operating in batch mode at 25% solids. Interestingly, the much faster rate in batch predicted by the kinetic model was found not be consistent with experimental data, which show approximately equivalent rates for both batch and fed-batch reactors.

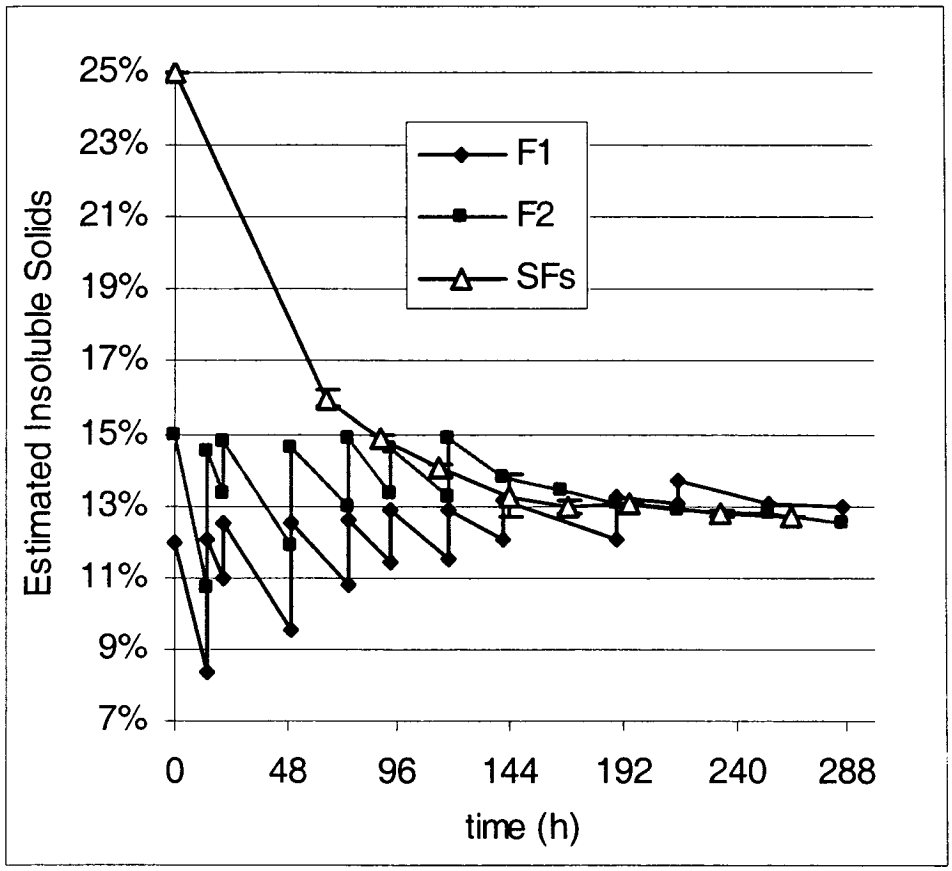


Figure 5.10: Estimated insoluble solids for fed-batch STRs and batch shake flasks; T = 45°C

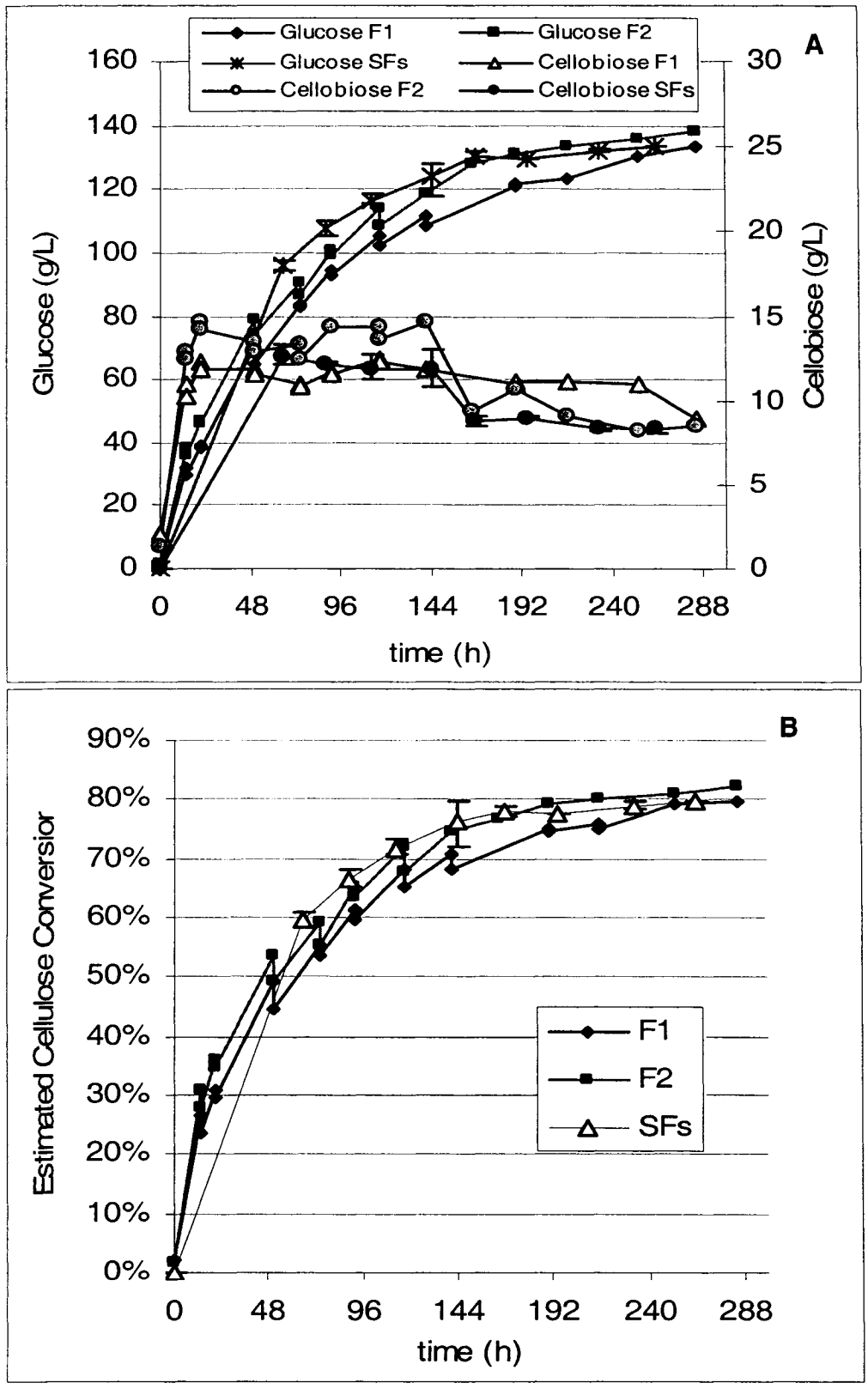


Figure 5.11: Glucose, cellobiose, and estimated cellulose conversion for the bench-scale STRs using the fed-batch feeding policies F1 and F2 and compared to batch CSS in shake flasks; T = 45°C

The implications of this work are that fed-batch saccharification enables high solids levels to be hydrolyzed at high conversions without the drawbacks associated with mixing, temperature, and pH control that would limit level of solids used in stirred tank reactors. Final glucose concentrations >130 g/L and estimated cellulose conversions of 80% are achieved in both reactors, which is very promising economically. These high sugar levels are significant in that previous high-solids enzymatic saccharification work (Varga *et al.*, 2004; Mohagheghi *et al.*, 1992) assumes that sugar inhibition will be a severely limiting factor, necessitating an SSF approach. Not only does this work show that current enzyme preparations are capable of robust activity at these high glucose levels, but that fed-batch can be used to perform these reactions at reasonable insoluble solids levels in STRs.

5.5 Conclusions

The results of this work were significant in that it was demonstrated that through fed-batch, a reactor capable of handling approximately 15% insoluble solids, was able to achieve high cellulose conversions at 25% equivalent initial insoluble solids. It was also demonstrated that a kinetic model could be applied in an offline optimal control scheme to determine a fed-batch feeding policy capable of facilitating mixing and temperature control within the reactor, while achieving these high equivalent insoluble solids levels. When compared to batch saccharification results obtained in shake flask reactors, the overall rates in fed-batch reactors were slower due to the shorter cumulative residence times, although the final conversion results were similar. This is significant since slurry handling by process equipment (pumping, mixing, temperature, and pH control) is greatly

simplified by operating at a lower solids level while gaining the economic advantages of a high-solids process.

5.6 References

Aden A, Ruth M, Ibsen K, Jechura J, Neefes K, Sheehan J, Wallace B, Montaue L, Slayton A, Lukas J (2002). Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolyiss and Enzymatic Hydrolyiss for Corn Stover. NREL Technical Paper 510-32438.

Asenjo JA (1984). Modelling the Bioconversion of Cellulose into Microbial Products: Rate Limitations. *Proc Biochem.* 19(6):217-224.

Ballesteros I, Oliva JM, Negro MJ, Manzanares P, Ballesteros M (2002). Ethanol Production from Olive Oil Extraction Residue Pretreated with Hot Water. *Appl Biochem Biotechnol.* 98-100:717-731.

Bequette B. (1991) Nonlinear Control of Chemical Processes: A Review. *Ind Eng Chem Res* 30:1391-1413.

Borden JR, Lee YY, Yoon HH (2000). Simultaneous Saccarification and Fermentation of Celluloseic Biomass to Acetic Acid.

Calsavara LP, De Moraes FF, Zanin GM (1999). Modeling cellobiose hydrolysis with integrated kinetic models. *Appl Biochem Biotechnol.* 77-79:789-806.

Chang VS and Holtzapple MT (2000). Fundamental Factors Affecting Biomass Enzymatic Reactivity. *Appl Biochem Biotechnol.* 84-86:5-37.

Converse AO, Ooshima H, Burns DS (1990). Kinetics of Enzymatic Hydrolysis of Lignocellulosic Materials Based on Surface Area of Cellulose Accessible to Enzyme and Enzyme Adsorption on Lignin and Cellulose. *Appl Biochem Biotechnol.* 24-25:67-73.

Gamkrelidze, RV (1978). *Principles of Optimal Control Theory.* Plenum Press, New York.

Griesse R and Walther A (2004) Evaluating Gradients in Optimal Control: Continuous Adjoints versus Automatic Differentiation. *J Optim Theo Control.* 122(1):63-86.

Gusakov AV, Arkady PS (1992). A Theoretical Analysis of Cellulase Product Inhibition: Effect of Cellulase Binding Constant, Enzyme/Substrate Ratio, and β-Glucosidase Activity on the Inhibition Pattern. *Biotech Bioeng.* 40(6):665-671.

Gusakov AV, Sinitsyn AP, Klyosov (1985). Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch process. *Enz Microbiol Technol.* 7:346-352.

Ingesson H, Zacchi G, Yang B, Esteghlalian AR, Saddler JN (2001). The effect of shaking on the rate and extent of enzymatic hydrolysis of cellulose. *J Biotechnol.* 88:177-182.

Hodge D and Karim MN (2002) Modeling and Advanced Control of Recombinant *Zymomonas mobilis* Fed-Batch Fermentation. *Biotech Prog.* 18:572-579.

Hong J (1986) Optimal Substrate Feeding Policy for a Fed Batch Fermentation with Substrate and Product Inhibition Kinetics. *Biotech Bioeng* 28:1421-1431.

Kadam KL, Rydholm EC, McMillan JD. (2004) Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnol Prog.* 20(3):698-705.

Kleman GL, Chalmers JJ, Luli GW, Strohl WR (1991). A Predictive and Feedback Control Algorithm Maintains a Constant Glucose Concentration in Fed-Batch Fermentations. *App Environ Microbiol.* 57(4), 910 -917.

Klesov AA (1982). Regulation of the formation and conversion of intermediate cellooligosaccharides and cellobiose during ezymatic hydrolysis of insoluble cellulose (Article in Russian) *Biokhimiia.* 47(4):608-18.

Lee JH, Lim HC, Yoo YJ, Park YH (1999). Optimization of feed rate profile for the monoclonal antibody production. *Bioproc Eng* 20:137-146.

Mais U, Estaghlalian AR, Saddler JN (2002). Influence of Mixing Regime on Enzymatic Saccharification of Steam-Exploded Softwood Chips. *Appl Biochem Bioeng.* 98-100:464-472.

Mansfield SD, Mooney C, Saddler JN (1999). Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnol Prog.* 15:804-816.

Modak J, Lim H, Tayeb Y. (1986) General Characteristics of Optimal Feed Rate Profiles for Various Fed-Batch Fermentation Processes. *Biotech Bioeng* 18:1396-1407.

Mohagheghi A, Tucker M, Grohmann K, Wyman C (1992). High Solids Simultaneous Saccharification and Fermentation of Pretreated Wheat Straw to Ethanol. *Appl Biochem Biotechnol.* 33:67-81.

Nguyen QA, Tucker M, Keller F, Eddy F (2000). Two-Stage Dilute Acid Pretreatment of Softwoods. *Appl Biochem Biotechnol* 84-86:561-576.

Nguyen QA (1998). Tower Reactors for Bioconversion of Lignocelluose Material. US patent no. 5733758.

Nilsson A, Taherzadeh MJ, Lidén G (2002). On-line estimation of sugar concentration for control of fed-batch fermentation of lignocellulosic hydrolyzates by *Saccharomyces cerevisiae. Bioproc Biosys Eng.* 25:183-191.

Ohmine K, Ooshima H, Harano Y (1983). Kinetic Study on Enzymatic Hydrolysis of Cellulose by Cellulase from *Trichoderma viride. Biotechnol Bioeng.* 25(8):2041-2052.

Ooshima H, Burns DS, Converse AO (1990). Adsorption of Cellulase from *Trichoderma reesei* on Cellulose and Lignacious Residue in Wood Pretreated by Dilute Sulfuric Acid with Explosive Decompression. *Biotechnol Bioeng.* 36:446-452.

Ooshima H, Kurakak M, Kato J, Harano Y (1991). Enzyme Activity of Cellulase Adsorbed on Cellulose and its Change During Hydrolysis. *Appl Biochem Biotechnol.* 31(3):253-266.

Philippidis GP and Hatzis C (1997). Biochemical Engineering Analysis of Critical Process Factors in the Biomass-to-Ethanol Technology. *Biotech Prog.* 13(3):222 – 231.

Pimenova NV and Hanley TR (2003). Measurement of Rheological Properties of Corn Stover Suspensions. *Appl Biochem Biotechnol.* 105-108:383-392.

Pörtner R, Schwabe JO, Björn F (2004). Evaluation of selected control strategies for fed-batch cultures of a hybridoma cell line. *Biotechnol Appl Biochem.* 40:47-55.

Pristavka AA, Salovarova VP, Zacchi Z, Berezin IV, Rabinovich ML. (2000) Enzyme regeneration during hydrolysis of steam-pretreated willow and requirement for cellulase complex composition (Article in Russian) *Prikl Biokhim Mikrobiol.* 36(3):278-86.

Rawlings JB and Muske KR (1993). The Stability of Receding Horizon Control. *IEEE Trans Aut Control* 38:1512-1516.

Rudolf A, Galbe M, Lidén G (2004). Controlled Fed-Batch Fermentations of Dilute-Acid Hydrolysate in Pilot Development Unit Scale. *Appl Biochem Biotech* 113-116:601-617.

Ryszczuk A and Emborg C (1996). Evaluation of mammalian fed-batch cultivations by two different models. *Bioproc Eng* 16:185-191.

Schell DJ, Farmer J, Newman M, McMillan JD. (2003a). Dilute-sulfuric acid pretreatment of corn stover in pilot-scale reactor: investigation of yields, kinetics, and enzymatic digestibilities of solids. *Appl Biochem Biotechnol.* 105 -108:69-85.

Schell DJ, Farmer J, Zuccharello M (2003b). Effect of Solids Loading on Pretreatment Kinetics. NREL C-Milestone Report.

Schell DJ, Farmer J, Hamilton J, Lyons B, McMillan JD, Sáez JC, Tholudur A (2001). Influence of Operating Conditions and Vessel Size on Oxygen Transfer During Cellulase Production. Appl Biochem Biotechnol. 91-93:627-642.

Schell DJ, Ruth M, Tucker M (1999). Modeling the Enzymatic Hydrolysis of Dilute-Acid Pretreated Douglas Fir. *Appl Biochem Biotechnol.* 77-79:67-81

Söderström J, Galbe M, Zacchi G (2004). Effect of washing on yield in one- and two-step steam pretreatment of softwood for production of ethanol. *Biotechnol Prog.* 20(3):744-9.

Spindler DD, Wyman CE, Mohagheghi A, Grohmann K (1988). Thermotolerant Yeast for Simultaneous Saccharification and Fermentation of cellulose to Ethanol. Appl Biochem Biotechnol. 17:279-293.

Taherzadeh MJ, Niklasson C, Lidén G (2000). On-Line Control of Fed-Batch Fermentation of Dilute-Acid Hydrolyzates. Biotechnol Bioeng. 69(3):330-338.

Tengborg C, Galbe M, Zacchi G (2001). Influence of Enzyme Loading and Physical Parameters on the Enzymatic Hydrolysis of Steam-Pretreated Softwood. *Biotechnol Prog.* 17:110-117.

Tsao GT, Gong CS, Cao NJ (2000). Repeated solid-phase fermentation and extraction for enzyme production. Appl Biochem Biotechnol. 84-86:505-24.

Varga E, Klinke HB, Réczey K, Thomsen AB. (2004). High solid simultaneous saccharification and fermentation of wet oxidized corn stover to ethanol. *Biotechnol Bioeng.* 88(5):567-574.

Wald S, Wilke CR, Blanche HW (1984). Kinetics of the enzymatic Hydrolysis of Cellulose. *Biotechnol Bioeng.* 26:221-230.

Weast RC (1985). CRC Handbook of Chemistry and Physics 66th Ed. CRC Press, Inc. Boca Raton, FL.

Wingren A, Söderström J, Galbe M, Zacchi G (2004). Process Considerations and Economic Evaluation of Two-Step Steam Pretreatment for Production of Fuel Ethanol from Softwood. *Biotechnol. Prog.* 20(5):1421-1429.

Zhang X, Mo H, Zhang J, Li Z. (2003). A solid-state bioreactor coupled with forced aeration and pressure oscillation. *Biotechnol Lett.* 25(5):417-20.

Zheng, A (2000) Some Practical Issues and Possible Solutions for Nonlinear Model Predictive Control. In: *Nonlinear Model Predictive Control* (F. Allgower and A. Zheng (ed)), 130-143. Birkhauser Press, Boston.

Zhu Y, Lee YY, Elander RT. (2004) Dilute-acid pretreatment of corn stover using a high-solids percolation reactor. *Appl Biochem Biotechnol.* 117(2):103-14.

Chapter 6

PROTEIN PROFILING OF MIXED SUBSTRATE FERMENTATION BY METABOLICALLY ENGINEERED *ZYMOMONAS MOBILIS*

6.1 Overview and Context

The ultimate catalysts for converting sugars to ethanol during fermentation are the proteins of the catabolic pathways. Whereas other work in this study has focused solely on characterization and optimization of the hydrolysis of cellulose to glucose through enzyme catalysis, this part of the work focuses on understanding *in situ* enzyme catalysis within the native and recombinant fermentative pathways of live cell cultures.

Two dimensional gel electrophoresis (2-DE) was the analytical tool used to profile protein expression patterns in a metabolically engineered *Zymomonas mobilis* strain performing ethanol fermentation using both its native fermentative pathway and a recombinant pentose phosphate pathway. Comparison to literature and predicted pIs and MWs resulted in the tentative identification of 15 proteins belonging to three functional classes. Using densitometry of protein spots during several time points of the fermentation resulted in the development of protein relative abundance profiles. Comparisons with previous work reported in the literature using either 2-DE or enzyme activity assays for both non-recombinant and metabolically engineered strains, prove difficult. It was, however, demonstrated in this work that functionally related proteins show similar expression trajectories that are separable by hierarchical clustering analysis

at the higher levels of agglomeration, and that both cluster analysis and correlation maps can be useful tools for extracting biologically relevant information from proteomic data.

6.2 Background

6.2.1 Biology of *Z. mobilis*

Z. mobilis utilizes the anaerobic Entner-Douderoff (ED) pathway to produce 1 mole of ATP per mole of glucose. Due to this low energy yield relative to glycolysis, it is necessary for the bacterium to sustain large glycolytic fluxes in order to obtain the energy equivalents required to support growth. Moreover, approximately 50% of the cellular protein in the *Z. mobilis* consists of ED pathway enzymes (An *et al.*, 1991), making the microorganism advantageous from an industrial perspective. The natural substrate range of *Z. mobilis* is limited to glucose, fructose, and sucrose. Although limited levels of enzymes for the pentose phosphate pathway (PP) have been detected (Sprenger, 1996; Bringer-Meyer and Sahm, 1989) for the purpose of synthesizing biosynthetic intermediates, growth or fermentation using pentose sugars is not possible. Metabolically engineered strains *of Z. mobilis* CP4 (also called ZM4) have the pentose phosphate pathway introduced from *Escherichia coli* K12 (Zhang *et al.*, 1995) to enable growth on xylose as a sole substrate. Xylose is transported into the cell via a high velocity carrier mediated facilitated diffusion mechanism for glucose, which might not be a rate-limiting step for xylose utilization (DiMarco and Romano, 1985). Simultaneous metabolism of both substrates occurs with no diauxic effects apparent, although glucose utilization is favored.

The genomes of strains ZM1 (ATCC 10988) and ZM4 (ATCC 31821) of *Z. mobilis* have recently been sequenced (Integrated Genomics, Inc., Chicago, IL; MacroGEN,

Seoul, South Korea). The MacroGEN sequencing for the genome of strain ZM4 was released publicly in January 2005 (Seo *et al.,* 2005), and found a total of 1998 genes, 1346 of which could be assigned putative functions. The genomic sequence was used to design probes for use on a DNA microarray, which were used to compare expressed genes between strains ZM1 and ZM4. Before this was released, publicly available information on the genes of *Z. mobilis* was available on the NCBI database, a search of which yielded the complete sequences of 17 fosmid and 2 cosmid clones from Kang and Kang (1998), as well as 3 plasmids and numerous sequences of single and polycistronic genes from other work. Overall these represent 367 ORFs of 253 identified and 114 suspected genes covering approximately 20% of the 2.05 Mbp genome. These protein sequences are used for this work and the gene names, determined MWs, and pIs are listed in **APPENDIX 2** if these are available.

Two-dimensional gel electrophoresis (2-DE) proteomic analysis of *Z. mobilis* strain ZM4 (ATCC 31821) showed the strong induction of a set of proteins under heat and alcohol stress (Michel and Starka 1986) as well as during the stationary growth phase (Michel and Starka, 1987) although all proteins remained unidentified. Subsequent work (Michel, 1993) postulated the identity of some of these proteins to be the heat shock proteins (HSPs) homologous to the *E. coli* proteins GroEL and DnaK. Work by An *et al.* (1991) identified 13 proteins, and determined relative time course expression levels for several ED proteins during glucose fermentation. Other work has focused specifically on relating the flux of metabolites through the ED pathway in *Z. mobilis* to enzyme levels and other cellular physiological conditions. The objective of the work of Osman *et al.* (1987) using strain CP4 was to determine the physiological reasons for decrease in the

glycolytic flux from glucose to ethanol in uncoupled stationary or late phase batch cultures of 200 g/L initial glucose. This work used comprehensive analysis of specific enzyme activities, metabolite levels, intracellular pH gradients, magnesium, and nucleotide measurements. The results indicate that these decreases are not the result of ethanol accumulation, but rather proposed that based on experimental results, reasons for this may be due to decreased abundance of nucleotides, decrease in internal cellular pH, decrease in the specific activities of pyruvate kinase (PYK) and glyceraldehyde-3-phosphate dehydrogenase (GAP). The specific fermentative activity of the cell culture was measured separately as the CO_2 evolution rate per mg of cellular protein. This value was used as an independent measure of catabolic activity and the decreasing fermentative activity was found to correlate well to the activities of PYK and GAP, indicating that these enzymes may represent the rate-limiting steps. Other ED enzymes were found to increase with increasing cell density. An *et al.* (1991) used 2-DE to track enzyme abundance through the course of a batch fermentation of glucose by strain CP4. Sixteen major protein spots representing enzymes of the ED pathway and heat shock proteins were identified by several methods. One method used co-elution on the same gel using the protein that has been purified from *E. coli* or *Z. mobilis*, or overexpressed in *E. coli* or *Z. mobilis*. Other methods include Western blots and tentative assignment based on comparison of apparent molecular weights with literature values. The enzymes GAP, PGK, ADH II, ENO and PGM remained abundant into stationary phase, GDH, GDT, and PDC decreased, while GLK, PYK and ADH I were absent or greatly reduced. It was also determined that ADH II acts as a stress protein and increases in heat and ethanol stressed cells along with DnaK, GroES, and GroEL. These three proteins are associated with the

folding or refolding of unfolded proteins that would otherwise tend to aggregate due to hydrophobic interactions.

In metabolically engineered strains, the regulation of fluxes through the recombinant pentose phosphate pathway is dependent on the methods chosen for the expression of these enzymes. In the genomically integrated strain (C25) and plasmid-bearing strain CP4:pZB5 (Zhang *et al.*, 1995), two distinct operons have been inserted for the pentose phosphate pathway. In this, xylose isomerase *(xylA)* and xylulokinase *(xylB)* gene expression are under the regulation of the constitutive *gap* promoter, while transketolase *(tkt)* and transaldolase *(tal)* are regulated by the native enolase *(eno)* promoter. These combined pathways are illustrated in Figure 6.1. There are incomplete components of the TCA cycle present, which are used for generating biosynthetic precursors (Seo *et al.*, 2005; Sprenger, 1996).

Problems associated with these metabolically engineered strains are slower xylose utilization, with a μ_{max} up to 7 times lower and a maximum specific productivity up to 8 times lower for xylose metabolism relative to that of glucose for the strain CP4:pZB5 (Hodge and Karim, 2002). The work of De Graaf (1999) and Kim *et al.* (2000) have focused on determining the underlying problems and inefficiencies associated with xylose fermentation in plasmid-bearing constructs. One bottleneck identified by De Graaf *et al.* (1999) using *in vivo* NMR spectroscopy was the underexpression of the recombinant xylulokinase enzyme in one of their strains. Their work has also shown lowered amounts of nucleotide triphosphates when xylose is used as the sole substrate,

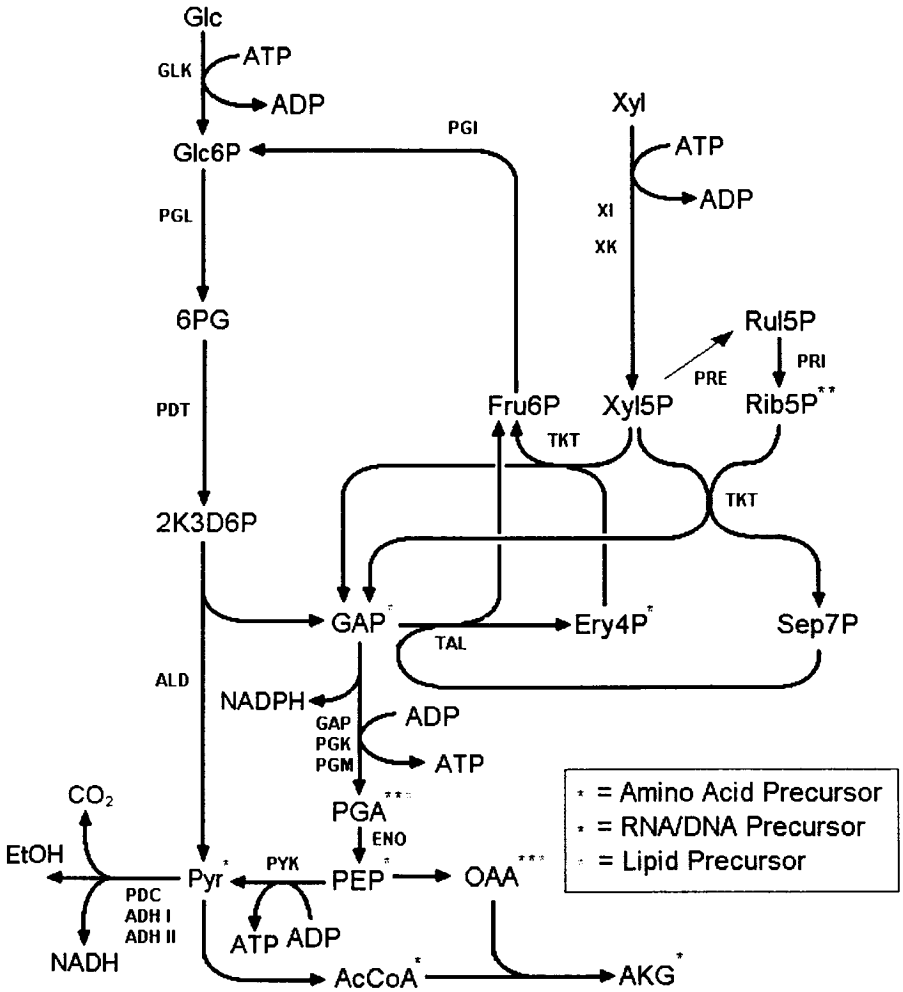


Figure 6.1: Combined Entner-Douderoff and Pentose Phosphate Pathways (after De Graaf *et al.*, 1999)

Metabolites

Glc	Glucose	GAP	Glyceraldehyde-3-phosphate
Xyl	Xylose	2K3D6P	2-keto-3-deoxy-6-phosphogluconate
Glc6P	Glucose-6-phosphate	PGA	Phosphoglycerate
Xyl5P	Xylulose-5-phosphate	PEP	Phosphoenolpyruvate
Fru6P	Fructose-6-phosphate	Pyr	Pyruvate
Rul5P	Ribose-5-phosphate	EtOH	Ethanol
Rib5P	Ribulose-5-phosphate	OAA	Oxaloacetate
Ery4P	Erythrose-5-phosphate	AcCoA	Acetyl-Coenzyme A
Sep7P	Sedoheptulose-7-phosphate	AKG	2-oxoglutarate
6PG	6-phosphogluconate		

Enzymes

PRI	Ribose 5-phosphate isomerase	PGI	Phosphoglucose Isomerase
PRE	Ribulose-5-phosphate epimerase		

All other enzymes given in Table 6.1

leading to a decrease in the overall energy available in the cell. Additionally, they detected no redox imbalances based on the pools of available reducing equivalents in anaerobic cultures and showed that the biomass yield was identical for both glucose and xylose growth. The work of Gao *et al.* (2002) has focused on the measurement of key enzyme levels in *Z. mobilis* strain C25 and comparing these with the similar plasmid-bearing construct 39676/pZB4L and the non-recombinant host strain. Their work concluded that the carbon flux through the ED pathway for glucose is not influenced by the presence of the four heterologous genes, and that the enzyme concentration profiles are divergent in the two metabolically engineered strains used. It was also determined that the enzyme levels do not show major differences between growth on glucose and growth on xylose as the primary substrate, although the recombinant enzymes TAL and TKT showed significant decreases over the course of fermentation

6.2.2 Sample Preparation

Sample preparation ultimately depends upon the method utilized for electrophoresis and visualization, although the primary goal of sample preparation should be the separation and solubilization of proteins. Several requirements need to be met for sample preparation and the formulation of a protein solubilization medium (Rabilloud and Chevallet, 2000). The first of these is the disruption of macromolecular interactions. This includes disulfide bonds and weak interactions within the protein, protein-protein interactions, and protein-other biomolecular (lipid, nucleic acid) interactions. The next requirement is the prevention of artifactual modifications of the polypeptide chain, which will affect the repeatability of the 2-DE process. The deactivation of protein-modifying enzymes such as proteases is necessary to satisfy this requirement. Next is the removal

of interfering substances such as salts, lipids, nucleic acids, and polysaccharides that could lead to protein modifications. Finally, proteins must be kept in solution for the duration of the electrophoresis process.

Given these previous requirements, compounds necessary for protein solubilization can be identified and compared. The disruption of disulfide bridges is usually achieved by the addition of a thiol compound. Dithiothreitol (DTT) is one of the most popular disulfide reducing compounds used in proteomics that does not have some of the drawbacks of certain other thiols (Rabilloud and Chevallet, 2000). One method for disrupting the noncovalent interactions within the protein is the use of a chaotrope. Chaotropes alter solvent parameters such as dielectric constant, hydrogen bonding, and polarizability which in turn alters the protein-solvent boundary and hence the net free energy for weak interactions within the protein. This results in protein unfolding becoming thermodynamically favorable. Urea is a commonly used chaotrope in protein solubilization medium and is efficient in disrupting hydrophobic interactions from the exposed protein core. Surfactants are often used in conjunction with the chaotrope to overcome hydrophobic interactions. Only nonionic and zwitterionic surfactants (*e.g.* CHAPS) can be used during isoelectric focusing (IEF) since these have no charge (Rabilloud and Chevallet, 2000).

6.2.3 Isoelectric Focusing

Since amino acids and hence proteins are amphoteric, *i.e.* exhibiting both acidic and basic functional groups, there is a pH value at which the anionic and cationic forms exist in equilibrium as neutral, dipolar zwitterions known as the isoelectric point. In the presence of an electric field, a charged particle is subjected to a force proportional to the

charge of the particle (Görg and Weiss, 2000). At the isoelectric point, the charge is zero and hence the force is zero and the particle ceases to migrate. The consequences of this for electrophoretic separations are that particles can be separated based on differences in isoelectric point during the process known as IEF.

Immobilized pH gradient (IPG) strips (Görg and Wiess, 2000) are currently the most widely utilized means for performing IEF. Commercially available IPG strips are available for a wide range of pH gradients and spans. Important considerations for IPG strips involve the pH range type of gradient. In analyzing global protein expression changes, where observed changes in the most proteins would be desirable, either an IPG strip covering a wide range of pHs (*e.g.* 3-10) should be utilized, or several gels should be run on the same sample with IPG strips of differing ranges covering the same overall range. A drawback to the second approach is that while providing a better resolution picture of protein, more gels will be required for each sample, and between gel variances must be taken into account that would not emerge if only a single pH range were used. The pH gradient is the next parameter of importance for IPG electrophoresis, and can be either linear or nonlinear in pH with respect to distance. Advantages in using linear pH gradients are that these simplify the identification of proteins on gels based on the position of the protein. Nonlinear pH gradients can have a sigmoidal functional relationship between the pH and strip position. Advantages to this include improved resolution of proteins in the mid-pH region where the majority of protein isoelectric points occur, while still profiling a large pH range.

6.2.4 SDS-PAGE

All surfactants contain a hydrophobic domain and hydrophobic/hydrophilic (amphiphilic) regions. Sodium dodecyl sulfate (SDS) is a surfactant that is commonly used in polyacrylamide gel electrophoresis (PAGE). SDS is capable of establishing hydrophobic interactions with all combinations of amino acids in the protein with its alkyl tail without interfering with peptide bonds. This results in considerable protein unfolding, with SDS-protein complexes formed with the approximate ratio of 1.4 g SDS to 1 g protein (Righetti *et al.*, 2001). The ionic head group found on SDS serves to break internal ionic and hydrogen bonding interactions and causes electrostatic repulsion between SDS-protein complexes. Electrophoretic migration of the SDS-protein complex is approximately proportional to the protein molecular weight. The migration is based only on SDS-protein complex size, and is independent of any inherent protein charges (Righetti *et al.*, 2001).

6.2.5 Protein Quantification from 2-DE Gels

Reliable quantification in 2-DE is necessary if differentially expressed proteins are to be identified between gels. Obtaining reliable and repeatable quantification of data from 2-DE gels can be problematic due to the potential for large variability between samples and gels. Image densitometry, the measurement of the absorbance of a stained gel at either a given wavelength or a wide spectrum, is one of the more common tools for quantification in electrophoresis gels. Recent research has shown that several methods of staining and image analysis can provide suitable protein quantification from 2-DE gels (Lilley *et al.*, 2001; Yan 1999).

The process of staining is a necessary step in the visualization of proteins in order to differentiate protein from background. This requires the selective binding of visualization agents to reactive groups on the protein. Commonly used visualization techniques include staining with organic dyes, precipitation of salts, metal ion reduction, fluorescence staining, and radioisotope labeling (Rabillioud and Charmont, 2000).

The silver stain is based on the affinity of protein for the silver cation (Rabilloud and Charmont, 2000). SDS, chloride, and amino acids also have a high affinity for silver so it is necessary that these be removed during a fixation step before the silver is added. The fixation step, however, prevents proteins from being excised from the gel for further analysis. To gain a higher sensitivity, a sensitizing step can be performed prior to silver impregnation. In this step, a sensitizing agent is bound to the protein that is capable of reacting with or binding silver. With the sensitizing agent bound the protein, these sites are already "primed" for autocatalytic silver reduction, resulting in improved sensitivity (Rabilloud and Charmont, 2000). Silver staining has been found to be generally unsuitable for the quantitative analysis of proteins, with very limited dynamic range for correlations between relative spot volumes and protein concentrations (Lilley *et al.,* 2001; Yan *et al.,* 1999). Additionally, matched spots in duplicate gels from the same samples vary in intensity when using the silver stain (Voss and Haberl, 2000).

Another common staining method relies on the covalent attachment of organic dyes to reactive groups on the protein (thiols, amines). Since this covalent modification changes the pI, the staining must be performed after IEF. Compounds containing amines or thiols that are used in the solubilization media (*e.g.* Tris) must be removed prior to staining. This results in the lower binding yields and decreased sensitivity of organic

dyes such as Coomassie dyes and Amido Black. Because of the covalent modification, proteins must be de-stained prior to subsequent MS analysis.

Fluorescence stains are similar to organic dyes except that these rely on the binding (either covalently or not) of a fluorophore to the protein to provide the image signal intensity. When light at a given wavelength is applied, the bound fluorophore fluoresces, that is, emits photons at wavelengths longer than those used for excitation. These emissions can be measured to quantify the abundance of bound fluorophore-protein complexes. The advantages of fluorescence as a quantification technique include its noninvasive nature and its sensitivity. The Sypro Ruby (Molecular Probes, Eugene, OR) stain is a commercial fluorescent stain utilizing a ruthenium chelate, that compared to the silver stain has been shown to have a much more dynamic range of sensitivity (Lilley *et al.*, 2001), although the overall sensitivity is ~2-5 times less (Rabilloud *et al.*, 2001). The cost of this product, however, could prove to be a limiting factor for extensive quantitation projects requiring numerous gels, although a cheaper equivalent has been proposed by Rabilloud *et al.* (2001).

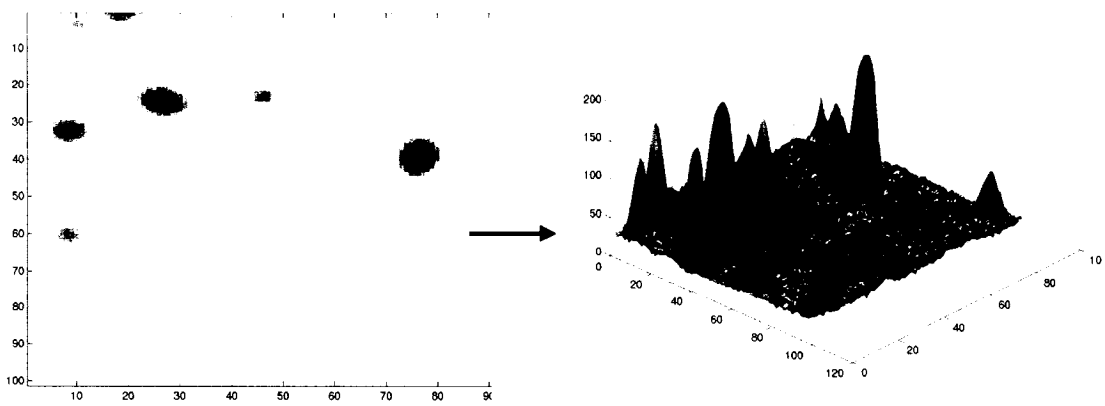


Figure 6.2: Conceptual rendering of 8-bit 2-DE gel to three dimensional image based on image pixel intensity (range of pixel intensities is reversed for clarity)

Scanned images are often digitized by assigning each pixel in a matrix corresponding to the image an 8-bit or 12-bit integer signal intensity. This yields a dynamic range for intensities between 0 and 256 for 8-bit images (2^8) and 0 to 4096 (2^{12}) for 12-bit images. A example of how this can be visualized is given in Figure 6.2, which shows how spot volume are essentially the volume under each peak. Colors or grayscale values associated with the pixel intensity are determined by assigning the image a corresponding color map. Alternatively, many color image formats utilize three intensities (red, green, blue) for each pixel. Several metrics can be developed for spot volume or density determinations. For 2-DE gels, these spot densities can be taken as either the relative spot volume or the percent of optical density (Yan *et al.*, 1999). The relative spot density can be defined as the sum of the spot intensities within the identified spot boundaries divided by the sum of all spot intensities. The percent optical density can be defined as the maximum intensity within the spot boundaries divided by the sum of all maximum spot intensities. These values can then be summed for all spots corresponding to known isoforms of a protein to yield a parameter that should ideally be correlated with protein abundance. Yan *et al.* (1999) correlated the various metrics for spot density to known protein concentrations and found that the relative volume provided the best linear match.

6.2.6 Cluster Analysis

Clustering is useful in determining groups of similarly expressed genes, and is a standard technique to group gene expression patterns for transcriptional profiling, while at the translational level this approach has yet to be applied extensively. This is most likely due to the difficulty in both spot quantification as well as identification of all protein spots. The fundamental premise behind the clustering gene or protein expression

profiles is that genes with similar expression patterns are co-regulated and may share a common function or contribute to a common pathway. Clustering techniques are generally defined as unsupervised or self-organizing in the sense that the classification is based solely on the significant features and patterns found in the data rather than approximating a specified input-output mapping. In mechanistic terms, clustering can be considered as determining similarities between objects within a data set. These similarities can be used to classify the objects into subsets or classes based on shared properties or behavior. Typically, these similarities can take the form of a correlation coefficient, distance metric, association coefficient, or a probabilistic similarity measure (Aldenderfer and Blashfield, 1984). Clusters can involve the statistical inference of cluster boundaries (Horimoto and Toh, 2001) or optimally bounded temporal clusters. Optimal temporal clustering can be done for transient expression profiles, and has been applied to group genes that are involved in the different phases of cell cycle in yeast (Lukasin and Fuchs, 2001). The temporal evolution in clusters is important in that this provides clues to changing regulatory responses to the environment.

One of the most widely applied clustering techniques is hierarchical clustering analysis (HCA), and was the first clustering technique applied to DNA microarray data (Eisen *et al.,* 1998). For this approach, data points are agglomerated into clusters according to their proximity to each other based on a distance metric. The Euclidian distance is one of the most common and is simply the geometric distance between the data vectors in Cartesian coordinates. For protein expression data set, this takes the form of:

$$Y_{jk} = \sqrt{\sum_{i=1}^{m} \left(x_{ji} - x_{ki}\right)^2} \qquad (6.1)$$

for the distance between transcripts or proteins j and k over conditions i. Other distance metrics include the Minkowski distance, the Manhattan (city block) distance, and the Mahalanobis distance, which incorporates the within group variance-covariance matrix in the distance estimate, and is equivalent to the Euclidian distance when there is no correlation.

Correlation within the data is an important property that establishes the linear dependencies between data vectors. Using the same subscript notation, the correlation coefficient can be defined as:

$$r_{jk} = \frac{\sum_{i=1}^{m} (x_{ji} - \overline{X}_j)(x_{ki} - \overline{X}_k)}{\sqrt{(x_{ji} - \overline{X}_j)^2 (x_{ki} - \overline{X}_k)^2}} \qquad (6.2)$$

where $\overline{X}_j$ represents the mean of transcript or protein j. The correlation matrix assembled from this equation can be regarded as the covariance matrix scaled with respect to the two variances in each matrix entry. This term defines the strength of the linear relation between two data vectors, with a value of 1 representing perfect linear correlation, -1 a perfect inverse linear correlation, and 0 no correlation.

Pretreatment of data is important for the performance of the clustering algorithm, where parameters such as scaling and weighting can be critical. Since the shape of the data profile rather than magnitude is often the property around which clustering is based, scaling such as autoscaling can be used. Autoscaling involves mean centering and scaling to unit variance and can be useful when magnitude is not as important a consideration as information profile. An alternative to data scaling is clustering with the data reconfigured as the correlation matrix. For this method, the Euclidian distance can be utilized with similar clustering properties to the Mahalanobis distance since variance-

covariance correlations are considered. Another alternative is to reorganize the data matrix through principal component analysis (PCA) and perform the hierarchical agglomeration based on the distance metric of the principle components (Yueng and Ruzzo, 2001). This is also similar to using the Mahalanobis approach since the data is processed to remove linear correlations between variables.

The agglomeration of clusters and data points in hierarchical clustering is dependent on the linkage method used. Linkage methods are based on the assumptions required for cluster membership and are defined in terms of cluster-data or cluster-cluster distances. The most popular linkage methods include single, complete, average, and Ward linkage (Aldenderfer and Blashfield, 1984). The single linkage method joins clusters based on the next closest member of the group to an outside point. This type of linkage method is defined as a space contracting linkage, since this tends to reduce the space between groups of data. A property of this method is that new data points are more often added to existing clusters rather forming new clusters, resulting in the chaining together of data to form a single large cluster. The average linkage method uses the average between the candidate data point and all members of a cluster to determine the agglomeration order. This linkage method is space conserving since the original nature of the distance space is conserved. The complete linkage method requires that a candidate data point must have similarity to all members of a cluster, and forms hyperspherical clusters. The Ward linkage method is formulated as an optimization problem that seeks to minimize within cluster variance. Both the Ward and complete linkage methods are considered space-dilating linkages since these have the property of forming small distinct clusters that "recede" upon formation (Aldenderfer and Blashfield, 1984).

Problems associated with hierarchical agglomerative clustering techniques usually are related to the large data matrices required for determining similarity (Aldenderfer and Blashfield, 1984). Iterative partitioning techniques, such as K-means, are available that are less computationally expensive algorithms. Cluster stability or robustness is another important factor when examining clusters. Genuine natural groupings of data should be considered highly stable and insensitive to clustering method (provided the clustering method can identifies the similar properties), data reordering, or data omission.

6.3 Materials and Methods

6.3.1 Cell Culture

Metabolically engineered *Zymomonas mobilis* (C25) and a non-recombinant strain (ATCC 39676) were obtained through a material transfer agreement with the National Renewable Energy Laboratory (NREL). The metabolically engineered strain contains four *E. coli* genes *xylA, xylB, tktA*, and *talB* introduced into its genome for a functional pentose phosphate pathway in addition to its native Entner-Douderoff (ED) pathway. The microorganism was routinely grown in culture tubes incubated at 30°C with no precautions taken to provide anaerobic conditions. The basal semi-defined culture medium (Joachimstal *et al.*, 1999) consisted of the following: KH_2PO_4, 2 g/L; $MgSO_4$ 1 g/L; $(NH_4)_2SO_4$ 1 g/L; Yeast Extract (DIFCO Laboratories Inc, Detroit, MI) 10 g/L; glucose/xylose, specified concentration. Stock solutions of the mineral salts were prepared to 10 times the desired concentration (g/L). Yeast extract was mixed with the mineral salt solutions and autoclaved separately from the carbohydrate solutions to prevent carmalization of the media. Initial pH was not adjusted, although measurement confirmed that the initial pH was consistently in the range of 5.7 to 5.8.

Cells were stored in stock cultures preserved in 10% glycerol, 90% basal media at –70°C. These were revived by incubating in culture tubes at 30°C in basal media with 5 g/L glucose and 5 g/L xylose. These cultures were kept active by intermittent transfers and were subsequently used as starter culture for all primary inocula used in shake flask and bioreactor fermentation studies.

The batch fermentation was performed in a 1 L fermenter (Applikon) with an agitation rate of 200 RPM and a temperature of 30°C. The pH was controlled at 5.75 by addition of 1 M NaOH and initially controlled at a low DO by sparging with N_2 and later by CO_2 evolution by the microorganism. Samples were taken at regular intervals at approximately every few hours during exponential growth. Biomass was measured through OD_{600} readings by spectrophotometer and correlated to the dry cell weight. HPLC analysis was performed on all fermentation samples to determine the glucose, xylose, and ethanol concentrations. A Waters HPLC equipped with a Waters differential refractive index detector was used. This was equipped with a Bio-Rad HPX-87H organic acid column operating at 65°C with a 0.01 N H_2SO_4 mobile phase at a flow rate of 0.6 mL/min. Fermentation samples were analyzed to determine glucose, xylose, and ethanol concentrations.

6.3.2 Dry Cell Weight/Optical Density Calibration Curve

Dry cell weight (DCW) is necessary for the determination of several key parameters in these fermentation studies. This value was found by determining a linear correlation between the total cell mass and the absorbance of visible light at 600 nm (OD_{600}) across a 1 cm path. Samples of *Zymomonas mobilis* CP4:pZB5 were collected in the late exponential phase. A portion of these samples was saved for OD analysis. The

remainder was centrifuged (Sorvall RC-5B, Du Pont Instruments) for 10 minutes at 12000 RPM. The resulting supernatant was removed and the biomass pellet was resuspended in DI water and centrifuged again. This washing procedure was to ensure the removal of all noncellular material that could be potentially reactive with the biomass during drying. The washing procedure was repeated twice and the biomass pellet resuspended in DI water was added to previously dried and tarred aluminum dishes. These were dried in a desiccator at 40°C for 12 hours and allowed to equilibrate to room temperature in a dry box for 3 hours before the mass was measured. The total cell mass in these was then determined and divided by the original sample volume to give the cell concentration.

The samples that were set aside for DCW measurement were diluted to various concentrations and the OD_{600} was taken with the spectrophotometer (Perkin-Elmer) zeroed on a cuvette blank containing DI water as a reference. Any OD_{600} measurement larger than 0.6 was considered out of the linear range of the calibration with more dilution required. The resulting correlation was determined to be:

$$[\text{DCW g/L}] = 0.2198*[OD_{600}] + 0.0038 \quad (6.3)$$

This correlation was verified periodically with samples from various fermentations. The relation only provides only the total DCW and does not consider differences between viable and dead cells.

6.3.3 Cell Harvesting for Protein

Cells are harvested by centrifuging ~70 mL of fermentation broth in 2 35 mL centrifuge bottles at 4000*g and 4°C for 20 minutes. Supernatant is decanted and cell mass is combined and washed 3 times with centrifugation at 20000*g and 4°C for 10

minutes in salt washing solution consisting of: KCl, 3.0 mM; KH_2PO_4, 1.5 mM; NaCl, 68 mM; NaH_2PO_4, 9.0 mM.

The biomass pellet is next resuspended in 10 mL of cell lysis solution consisting of the following: Urea, 9.0 M; CHAPS, 4.0% w/v; Tris base, 40 mM; Dithiothreitol, 0.5 mM. Cell suspension is sonicated on ice for 15 minutes (2 seconds on, 1 second off). Lysate is next centrifuged at 6000*g and 15°C for 30 minutes to remove cell debris. Supernatant is stored at -80°C until use. The total protein in this solution is next quantified.

6.3.4 Total Protein Assay

For total protein quantification, the BioRad Protein Assay is used. Bovine serum albumin (BSA) is used as a standard. A 1:5 dilution of the protein dye reagent (BioRad) is performed with nanopure water, and 5.0 mL of this dilution is pipetted into clean test tubes (six for standards and two for each unknown). Six dilutions of bovine plasma γ globulin (IgG) between 0 and 100 μL are pipetted into the standard tubes. From the known concentration of the standard, the mass of standard in each tube is determined. Samples are prepared by microcentrifuging 1 mL of the protein extraction lysate at 13000 RPM for 15 minutes. The supernatant is transferred to a new microcentrifuge tube to remove cell debris. This sample supernatant is added to the test tubes for the unknown as a volume that yields an approximate color corresponding to the range found in the standards, with the volume recorded. Each tube is vortexed for 15 seconds and allowed to incubate at room temperature for 5 minutes to 1 hour before absorbance measurements at 595 nm are taken on the spectrophotometer (Perkin-Elmer). The standards are utilized

to develop a calibration curve (Figure 6.3) from which the determination of unknown sample protein concentration is determined.

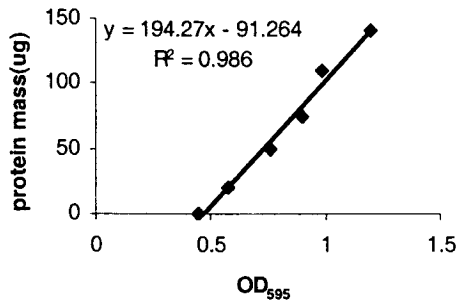


Figure 6.3: Example of a protein calibration curve.

6.3.5 Immobilized pH Gradient (IPG) Strip Rehydration

IPG rehydration stock solution is prepared containing in nanopure water: Urea 9.0 M; CHAPS, 4.0% w/v; Bromophenol Blue, Trace; DTT, 75 mM; IPG buffer, 0.5 mM; with the DTT and IPG buffer added only prior to use. A total of 100 μg of protein sample is mixed with the appropriate amount of rehydration stock solution to yield a total of 250 μL. Each sample is vortexed for 20 s to ensure mixing, and next microcentrifuged at 13000 RPM for 10 s to remove any bubbles that may form during mixing. All of the sample-rehydration stock mixture is added to the Rehydration Tray. Using clean gloves and forceps, one IPG strip is removed from the package and the protective plastic backing is removed. The pointed end of the strip is placed in the trough gel side down and added slowly so that the mixture evenly coats the strip and most bubbles are driven toward the blunt end. This process is repeated for each IPG strip and all are topped with mineral oil. The Rehydration Tray is leveled and allowed to remain undisturbed for more than 10 hours.

6.3.6 Isoelectric Focusing

One hundred μg of protein from each sample are mixed with urea, CHAPS, IPG Buffer (Amersham-Pharmacia), DTT, and Bromphenol blue in nanopure water. Samples are then applied to an IPG strip (Amersham-Pharmacia, pH range of 3 – 10 nonlinear gradient) by in-gel sample loading during rehydration. The 1st dimension unit (Immobiline Strip Tray, Amersham-Pharamacia Biotech) is attached to a cooling unit. One blotting strip is soaked in 6 mM H_3PO_4 and the other in 0.5 mM NaOH. The acid-soaked strip is placed under the anode with the pointed end of the IPG strip under this, gel-side up. The base-soaked strip is placed under the cathode and the blunt end of the strip is positioned under this. The IEF program that is used on the electrophoresis power supply (Amersham Pharmacia Biotech) consists of four steps. In the first step the voltage is linearly increased to 500 V over 1 minute, the second step holds the voltage at 500 V for 3 hr, the third step linearly increases the voltage from 500 V to 3500 V over 5 hours, and for the fourth step, the voltage is held at 3500 V for 9 hours.

6.3.7 SDS-PAGE

These are cast as 12% w/v consisting of (per 50 mL): H_2O, 16.5 mL; 30% acrylamide mix; 20.0 mL; 1.5 M Tris (pH is adjusted to 8.8 using 1 M HCl), 12.5 mL; 10 % (w/v) SDS, 0.5; 10% (w/v) ammonium persulfate, 0.5 mL; TEMED, 0.02 mL. After this is pipetted into the mold it is topped with 50% butanol to prevent drying and form a smooth, level surface. These are allowed to polymerize for at least 1 hour.

IPG strips are removed from the IEF unit and excess mineral oil is removed from the IPG strips with a Chem-wipe. Care is taken not to disturb the gel during this step. The IPG strips are placed in the rehydration tray and each strip is topped with ~2 mL of SDS

equilibration buffer consisting of: Tris-Cl, 50 mM (pH is adjusted to 6.8 using 1 M HCl); Urea, 6 M; Glycerol, 30% (v/v); SDS, 2% (w/v); Bromophenol Blue, trace. The IPG strips are allowed to equilibrate for 15 minutes. These are next transferred to the top of the SDS gel in the mold. These are placed in the electrophoresis unit (PROTEAN II xi multi-cell, Bio-Rad) and topped with SDS buffer. The SDS buffer is made in large batches (3 L) and consists of the following: Tris, 25mM; Glycine, 192 mM; SDS, 0.1% (w/v). The SDS-PAGE program used on the electrophoresis power supply consists of 3500 V for 2:15 h. The power is stopped if the dye from the equilibration buffer approaches the bottom of the gel.

6.3.8 Imaging

Gels are stained based on the procedure for the PlusOne Silver Staining Kit from Amersham Biosciences, modified by to the procedure of Yan *et al.* (2000) not to include gluteraldehyde and ethanol in place of methanol. For each 13 x 8 cm gel, 250 mL of staining solution was used. Gels were scanned using an HP DeskJet 3600 II digital scanner. Densitometry was performed using these digitally-rendered images with the software Z3 (Compugen, Inc.). This software determines spot volumes using the percent optical density measurement method outlined in Section 6.2.5.

6.4 Results and Discussion

6.4.1 Protein Identification

As is true for large-scale biochemical reactors as well as individual cells, protein is the primary catalyst for all chemical reactions, such that global protein expression measurement should yield significant insights into the condition of the culture. Analysis of the proteome is valuable because it can identify important proteins that may have

significant functions in cellular physiology that would not be detected by more specific approaches. In order for biological meaning to be addressed, protein identities must be known. Proteins spots visualized on gels can typically be identified using MALDI-TOF mass spectrometry or Western blots. Due to laboratory limitations, for this work information about *Z. mobilis* proteins was obtained from online databases and the literature. This was done for the purpose of comparison to laboratory-generated 2D-PAGE gels, such that a "proteome reference map" can be developed for the identification of proteins.

Using the online NCBI database (http://www.ncbi.nih.gov), a complete set of all publicly available nucleotide sequences belonging to the *Z. mobilis* genome, with annotations, was collected. Using the sequence annotations, 367 ORFs were found and consist of 253 identified and 114 suspected genes. The genes corresponding to all 15 proteins of the ED and fermentative pathways were present in the assembled data set as a result of the considerable body of literature that exists on these enzymes of *Z. mobilis*. Genomic sequence data of the 4 cloned enzymes for the PP pathway were determined from the *xylAB, tktA*, and *talB* genes of *E. coli K12*. The MW and theoretical pI for each of the translated sequences for this set of genes were determined using the "Compute pI/MW tool" available from the ExPASy website (http://www.expasy.ch/spot/). The results of this for the available catabolic enzymes and heat shock proteins are provided in Table 6.1, and are grouped according to the three functional categories: ED pathway, PP pathway, and heat shock proteins. Using this data, a simulated 2-DE protein gel can be generated (Figure 6.4). A summary of all proteins with pI/MW predictions is given in **APPENDIX 2**. Besides the theoretical activity of proteins on gels, protein identifications

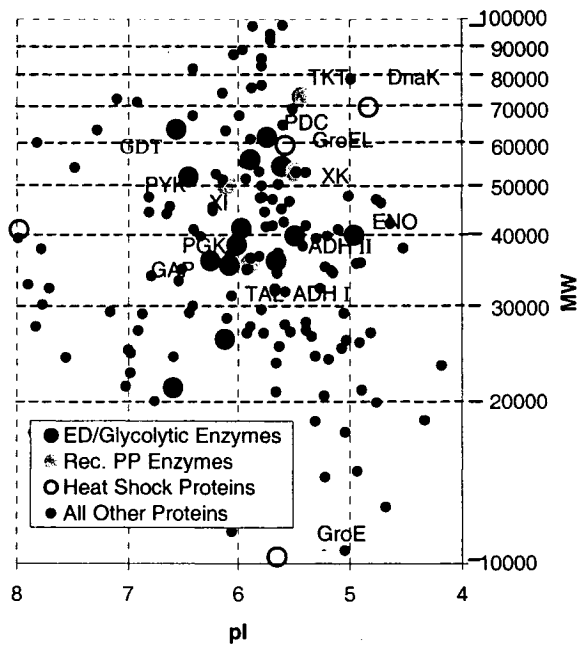


Figure 6.4: Simulated 2-D gel generated using MWs and pIs predicted from the protein sequence

have been published based on experimental methods such as Western blots or reference gels containing the protein of interest. Previously published material (Michel, 1993; An *et al.*, 1991; Michel and Starka, 1987) includes information about how a total of 16 identified proteins for strains CP4 and ZM4 partition on 2-DE gels.

Table 6.1: Major Catabolic and Heat Shock Proteins of *Z. mobilis*

Gene	Enzyme	Enzyme Name	EC Number	pI	MW (Da)	
adhA	ADH I	Alcohol dehydrogenase I	*1.1.1.1*	5.66	36094	Native ED Pathway Enzymes
adhB	ADH II	Alcohol dehydrogenase II	*1.1.1.1*	5.49	40155	
eda	ALD	2-dehydro-3-deoxyphosphogluconate aldolase	*4.1.2.14*	6.59	21121	
edd	PDT	Phosphogluconate dehydratase	*4.2.1.12*	6.56	62925	
eno	ENO	Enolase	*4.2.1.11*	4.96	40247	
gap	GAP	Glyceraldehyde 3 phosphate dehydrogenase	*1.2.1.12*	6.25	36102	
glf	GLF	Glucose facilitator transporter protein	----	9.07	50200	
glk	GLK	Glucokinase	*2.7.1.2*	6.08	35423	
pdc	PDC	Pyruvate decarboxylase	*4.1.1.1*	5.74	60912	
pgi	PGI	Phosphoglucose isomerase	*5.3.1.9*	5.89	55399	
pgk	PGK	Phosphoglycerate kinase	*2.7.2.3*	5.97	41388	
pgl	PGL	6-phosphogluconolactonase	*3.1.1.31*	5.33	27315	
pgm	PGM	Phosphoglucomutase	*2.7.5.1*	6.12	*25938*	
pyk	PYK	Pyruvate kinase	*2.7.1.40*	6.45	51445	
zwf	GPD	Glucose-6-phosphate dehydrogenase	*1.1.1.49*	5.61	53859	
*xylA**	XI	Xylose isomerase	*5.3.1.5*	6.1	49742	PP Pathway Enzymes
*xylB**	XK	Xylulokinase	*2.7.1.17*	5.49	52618	
*talA**	TAL	Transaldolase	*2.2.1.2*	5.89	35659	
*tktA**	TKT	Transketolase	*2.2.1.1*	5.43	72274	
groES	groES	10 kDa heat shock protein	----	5.65	10291	Heat Shock Proteins
groEL	groEL	60 kDa heat shock protein	----	5.58	58920	
*dnaK**	dnaK	70 kDa heat shock protein	----	4.83	69115	

* data for *E. coli* K12 proteins taken

Using metabolically engineered *Z. mobilis* C25 and the nonrecombinant host strain (ATCC #39676), 60 proteins of significant spot volume were apparent in 2D-PAGE gels under various conditions out of >300 visualized proteins. Also apparent were differences in protein expression patterns between the two strains, with major spots appearing in the

metabolically engineered strain which were absent from the non-recombinant strain, assumed to consist of some of the four recombinant proteins. Using the theoretical MWs and pIs as well as comparing gels to published data, protein spot identifications for 8 of the ED pathway proteins, 3 recombinant PP pathway proteins, and 3 heat shock proteins were inferred. These identifications are given in Figures 6.5 and 6.6 for the nonrecombinant and metabolically engineered *Z. mobilis* strains, respectively. As is apparent from the gel images, several proteins of significant spot volume exhibit multiple stable isoforms or charge chains across the pH gradient. For spot quantification work, the sum of charge chains is used.

6.4.2 Protein Profiling

The fluxes of metabolites through the catabolic pathways should be the primary point of interest for streamlining the fermentation process. Ultimately fluxes of metabolites can be regulated at four levels: through transcription, translation, protein turnover, and enzyme level. 2-DE only measures the relative abundance of enzymes, while quantification can be problematic.

For this work, a batch fermentation in a controlled bioreactor was performed with a glucose-xylose mixture concentration of 35-35 g/L. Four protein samples for 2-DE analysis were taken throughout the course of the fermentation; two during glucose consumption (t = 8 h and 11 h) and two during xylose consumption (t = 15 h and 19 h) as given in Figure 6.7. 2-DE was performed in duplicate for each time point except 11 h, in which the duplicate gel was irreparably damaged during preparation. Protein spot intensities were taken as a fraction of the total spot density of the gel (as described in the **Materials and Methods**, Section 6.3) and averaged among duplicate gels. Figure 6.8

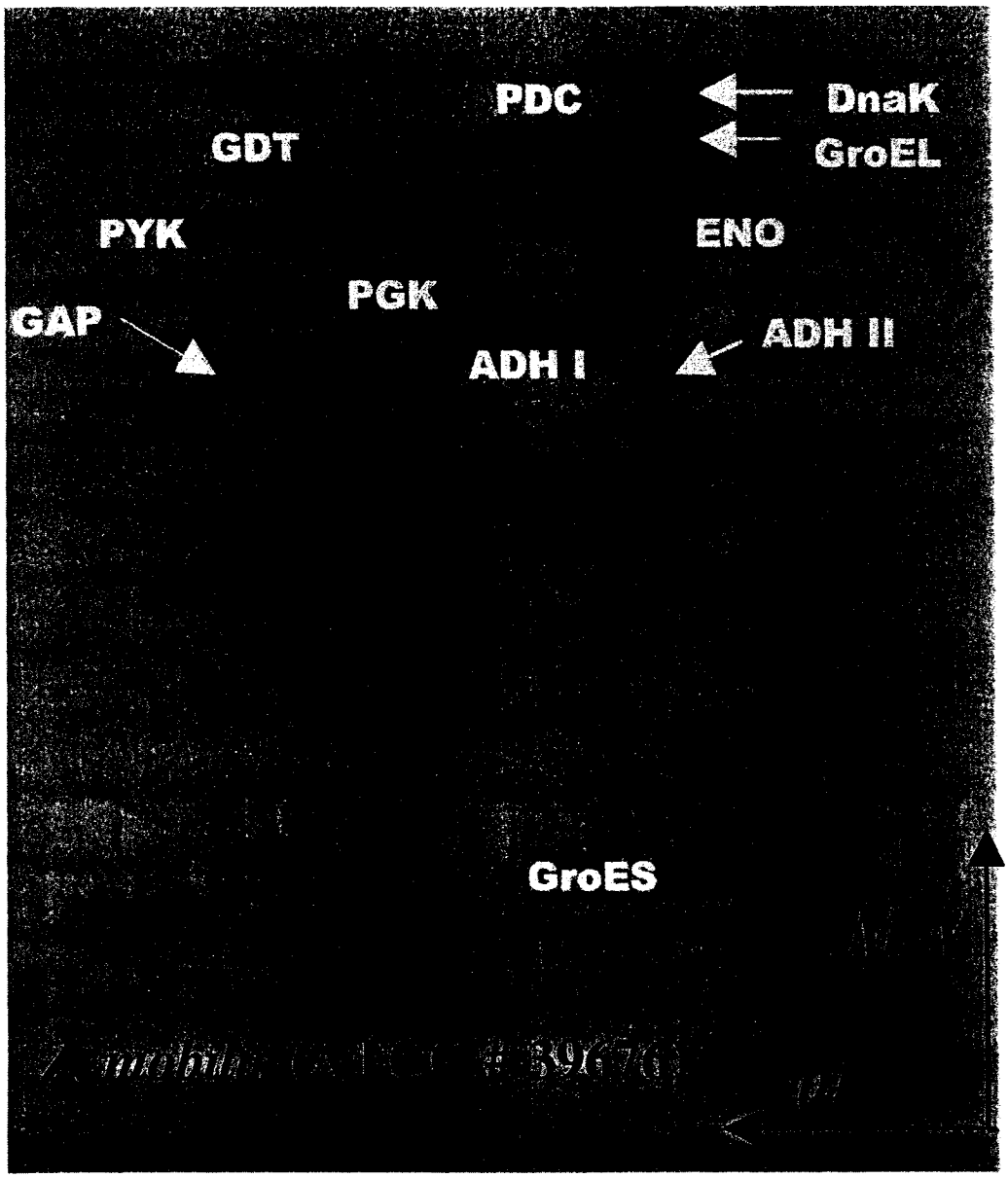


Figure 6.5: 2-D SDS-PAGE gel of nonrecombinant *Z. mobilis* with tentative protein identification.

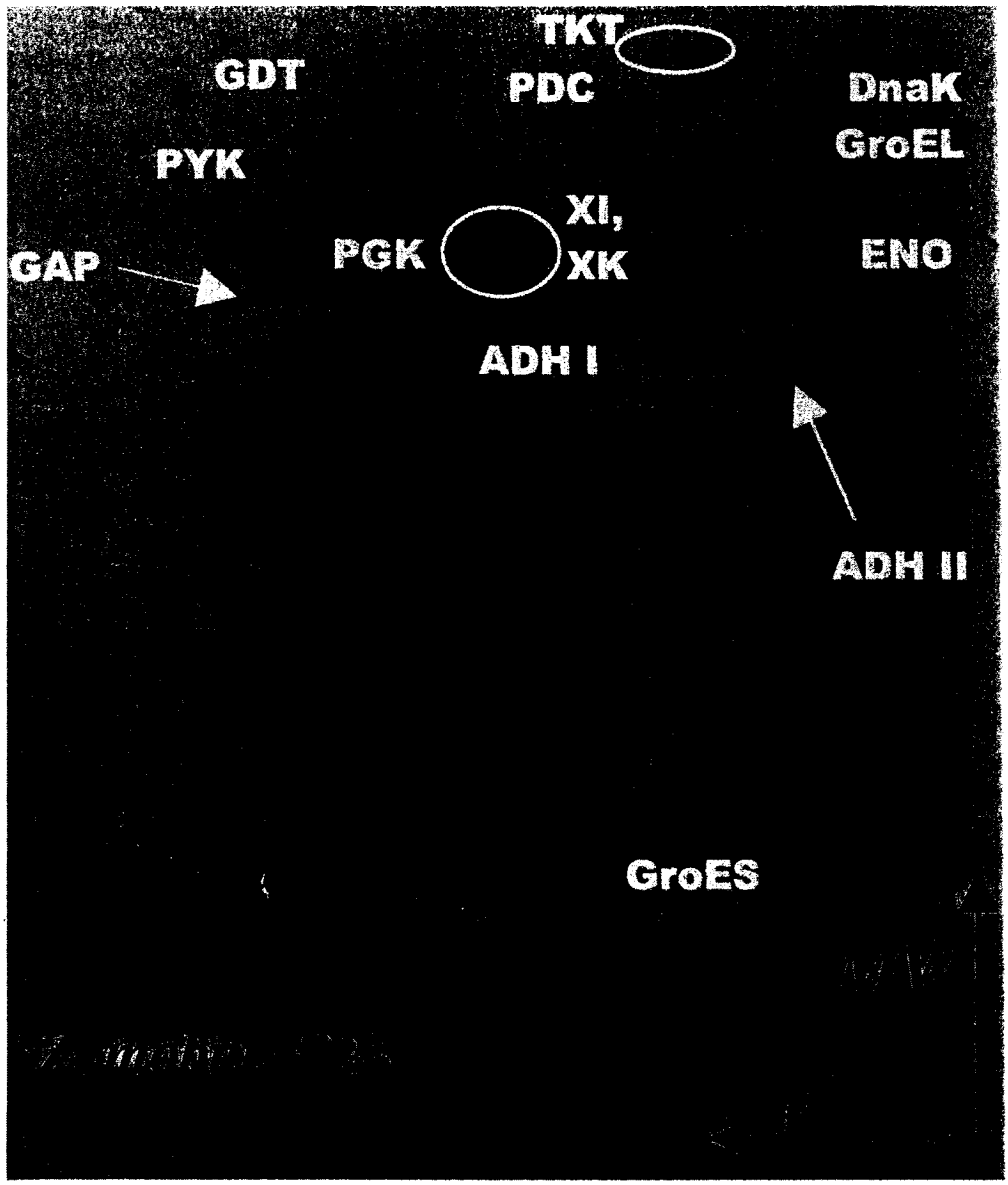


Figure 6.6: 2-D SDS-PAGE gel of *Z. mobilis* C25 with tentative protein identification. Circled spots indicate recombinant proteins.

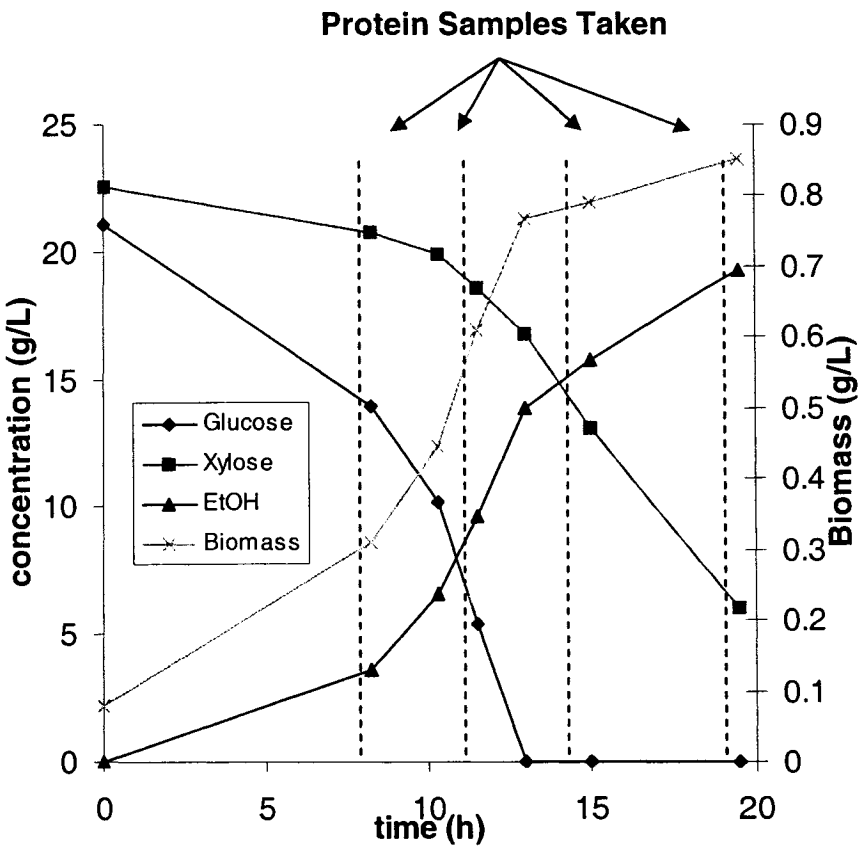


Figure 6.7: Mixed substrate batch fermentation by metabolically engineered *Z. mobilis* corresponding to protein profiles of Figure 6.9.

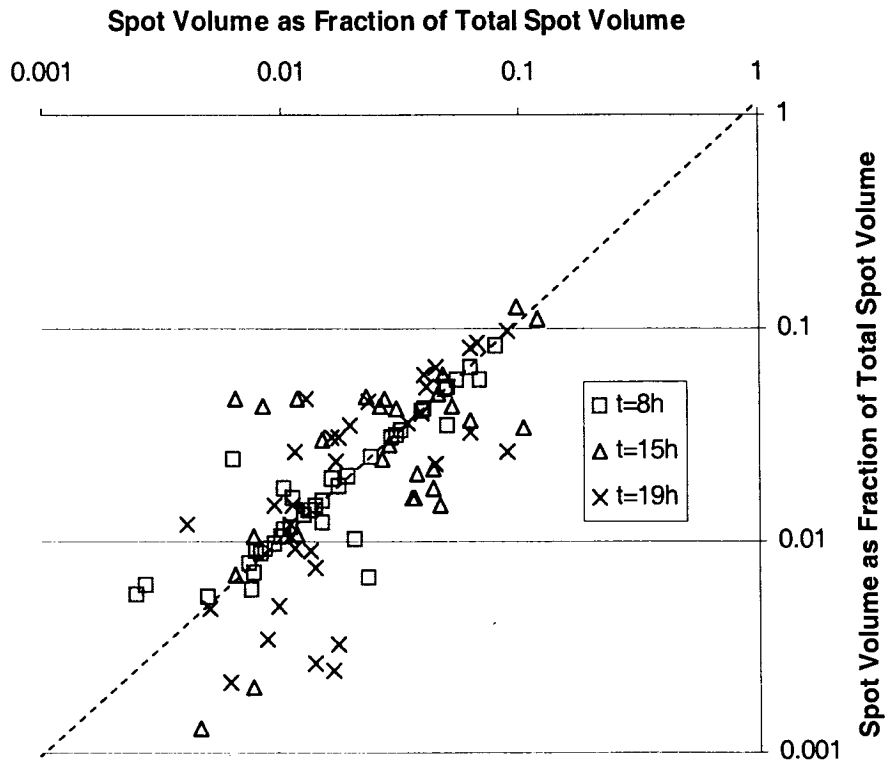


Figure 6.8: Repeatability of protein spot volumes taken from gel duplicates.

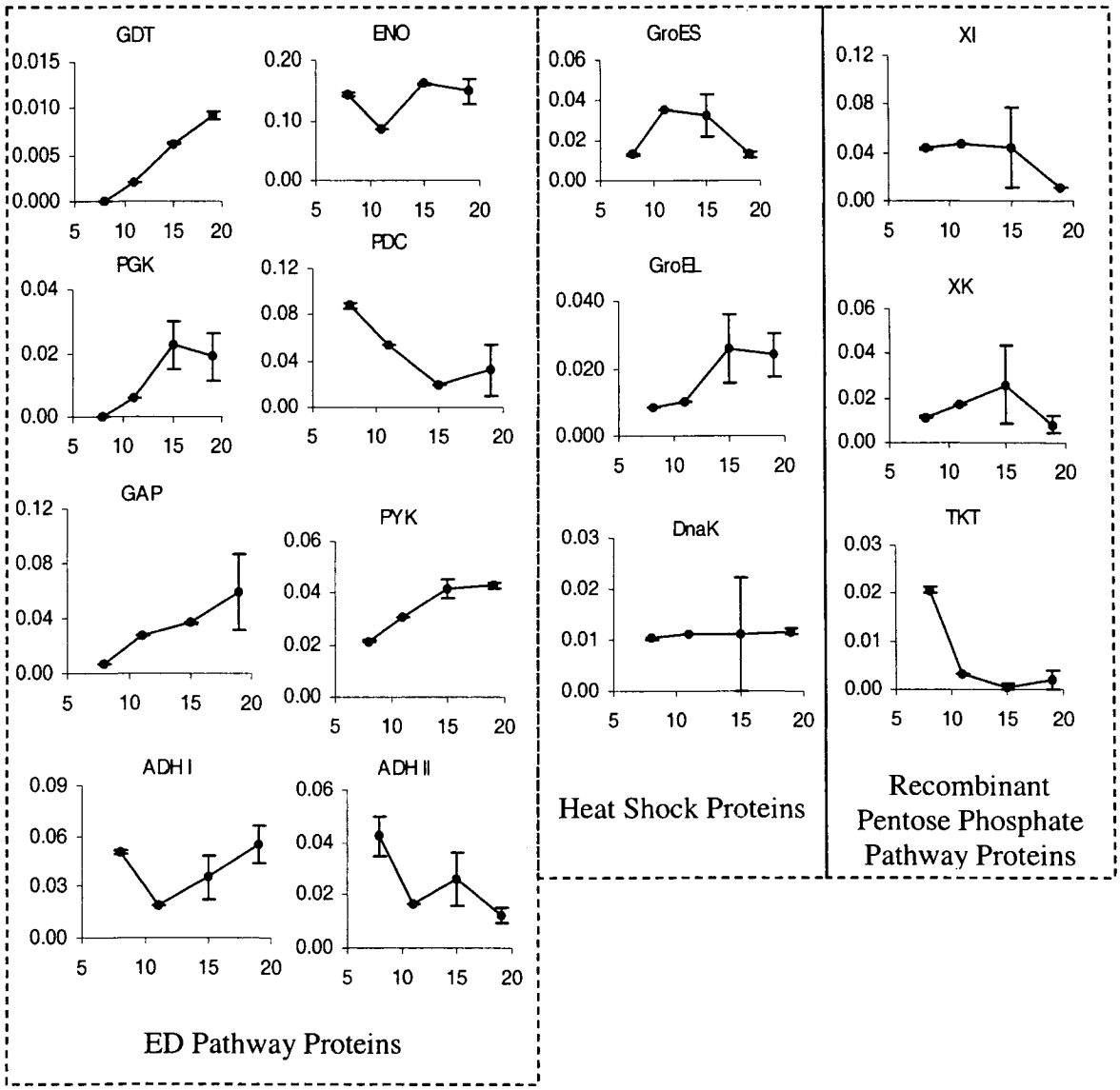


Figure 6.9: Expression profiles for tentatively identified proteins. x-axis: time (h); y-axis: spot volume as fraction of total.

Quantification Method	Specific Enzyme Activity	SDS-PAGE	Specific Enzyme Activity	SDS-PAGE
Z. mobilis Strain	Non-recombinant CP4	Non-recombinant CP4	Recombinant CP4	Recombinant CP4
Protein	Osman *et al.*, 1987	An *et al.*, 1991	Gao *et al.*, 2002	This work
ADH I	↑	↓		↔
ADH II	↑	↑		↓
ENO	↑	↑		↔
GAP	↓	↑		↑
GDH	↑	↓	↔	
GDT		↓		↑
GLK	↑	↓	↔	
PDC	↑	↓		↓
PGI			↔	
PGK	↑	↑		↑
PGM	↑	↑		
PYK	↓	↓		↑
XI			↔	↓
XK			↔	↔
TAL			↓	
TKT			↔	↓
dnaK		↑		↔
groES		↑		↔
groEL		↑		↑

Table 6.2: Trends in Protein Expression. ↑: greater than 2-fold increase; ↓ : greater than 2-fold decrease; ↔: less than 2-fold change

shows the variability encountered both between gels and between preparations. The protein spot volumes for the time point at 8 h show little variability between duplicates since the gel duplicates were performed in parallel, while the samples at 15 h and 19 h were processed in different batches. Some variability may also result from incorrect matching of protein spots between gels, which was performed manually. Using the commercial software Z3 (Compugen, Inc.), spot densitometry was performed and protein volumes were determined as the spot volume as a fraction of the total spot volume. More than one protein may elute for the same spot, which could result in added error to this

already imprecise quantification method of silver staining. The time course spot density profiles for all tentatively identified proteins are given in Figure 6.9.

A summary of the trends is given in Table 6.2, and compares these results to the findings of other researchers for trends in protein expression during batch fermentation. Gao *et al.* (2002), performed enzyme assays on several key enzymes during mixed substrate fermentation by the same microorganism (strain C25) and was given as μmol / min / g dry cell weight, An *et al.* (1991) carried out 2-DE for time course fermentation of glucose by a non-recombinant *Z. mobilis* (CP4), and Osman *et al.* (1987) also performed enzyme assays during glucose utilization by strain CP4 with the results given in units of μmol / min / mg protein. The data presented in Table 6.2 shows the trends found between the four works during the batch fermentation course of either glucose (nonrecombinant strains) or glucose-xylose (metabolically engineered strains). The arrows indicate greater than 2-fold change in expression level for the data. The limitations for using the different methods of protein quantification should be considered when comparing these results. Because the four studies used two different methods, neither of which are precise quantification methods, it is understandable that there may be some dissimilarity between the data. Additionally, considering that the spot identifications for this work are inexact can cast further doubt on any generalizations concluded from this study.

6.4.3 Cluster Analysis

Hierarchical cluster analysis (HCA) was next applied to quantify similarity between protein expression profiles. For standard HCA, each protein spot was normalized with respect to its mean and standard deviation throughout the time course such that a

consistent basis of relative expression could be used. HCA was next applied separately to both the normalized expression level and the correlation coefficient to develop a dendrogram linking patterns of similar expression. Both of these dendrograms are given in Figure 6.10, and include the normalized expression profile along the x-axis with increasing intensities black or white indicating the deviation from the protein mean. The postulated protein identities associated with each numbered protein spot or protein number are given in the figure. Another way to visualize data ordered by HCA is with a correlation map (Figure 6.11), and has an advantage over a standard dendrogram in that correlations between all protein spots can be visualized simultaneously. The strength of the linear correlation is indicated by the shade of brightness of the map.

A positive aspect of hierarchical clustering is that class membership can be evaluated at any level of agglomeration from the sample level to the largest supercluster. This characteristic of hierarchical clustering is ideal when analyzing data of unknown inherent data groupings that emerge from exploratory technologies such as 2-DE protein analysis where many aspects of the system are either unknown or poorly understood. At the highest level, the two largest superclusters in Figure 6.10 show the general trends for proteins of either increasing or decreasing relative abundance throughout the course of the fermentation using both clustering approaches. These same trends are captured in the correlation map as the light and dark regions of the four corners of the correlation map (Figure 6.11), while light regions along the diagonal represent the clustered proteins. At the closest level of similarity, isoforms of several proteins are correctly associated with each other as in case of TKT, PDC, and GDT.

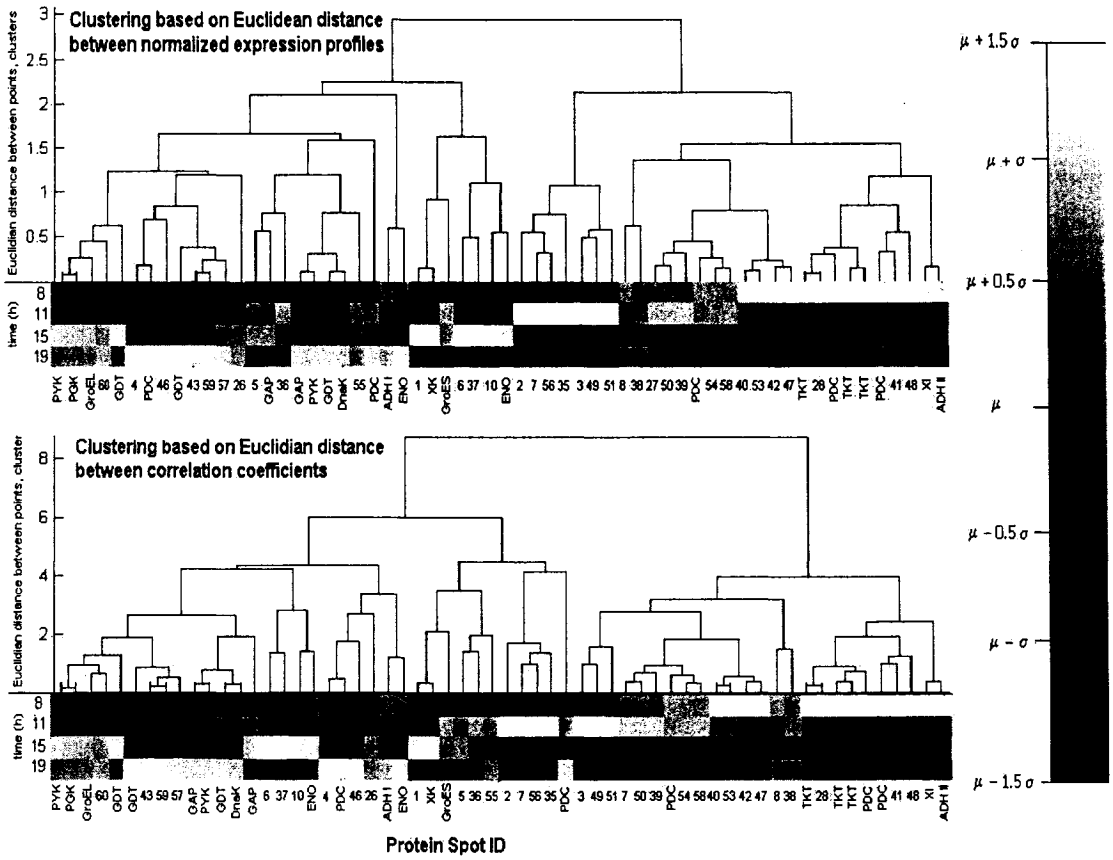


Figure 6.10: Comparison of clustering dendrograms of protein spot abundance profiles based on either Euclidian distance or correlation coefficient. μ: mean for each profile; σ: standard deviation of each profile

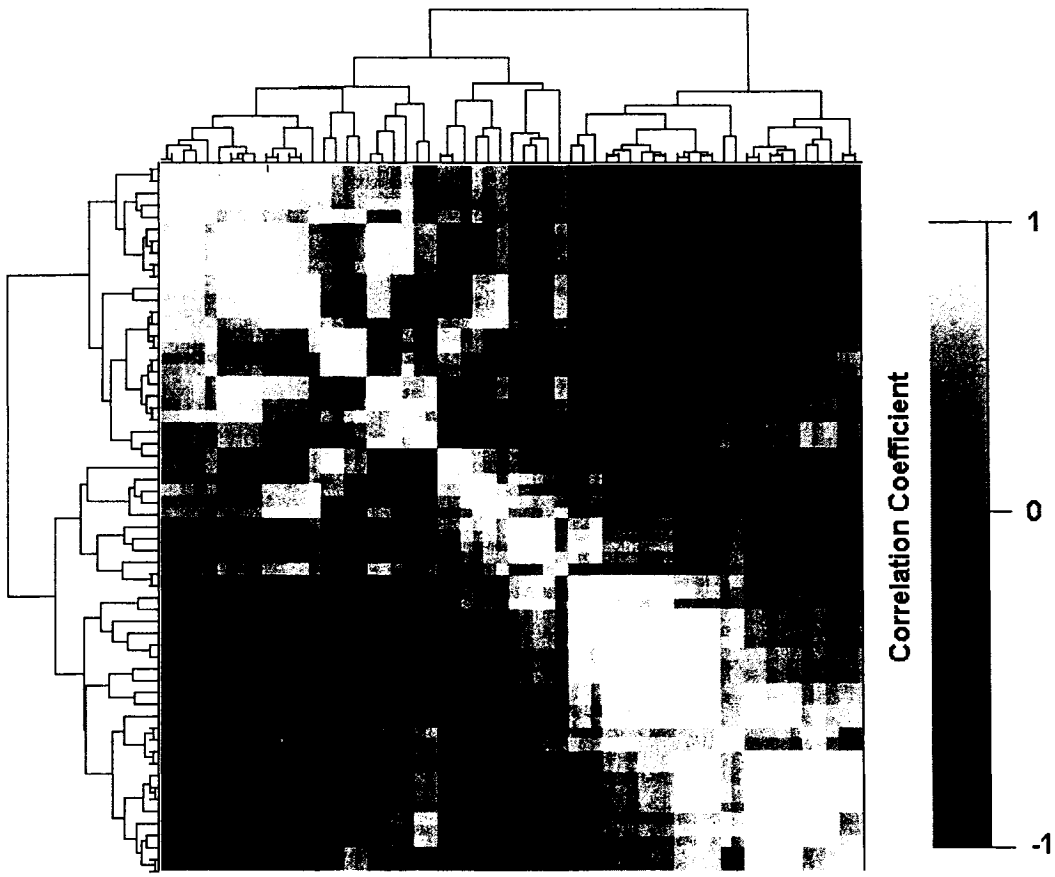


Figure 6.11: Correlation map organized based on clustering of correlation coefficient as given in Figure 6.10.

While the three classes of proteins do not necessarily show absolutely unambiguous coordinated patterns of expression, some general trends do appear in the clustering dendrograms and correlation map. One trend is that heat shock proteins and ED pathway enzymes (with the exceptions PDC and ADH II) are classified in the supercluster on the left of the figure which contains either continually increasing or approximately stationary expression throughout the course of the fermentation. The other trend is that all PP pathway proteins are classified in the right supercluster containing proteins that show the general trend of decreasing relative abundance. The biological implications of this result are that the foreign proteins of the metabolically engineered pathway show decreased expression when the foreign PP pathway is utilized for the carbon and energy source. However, it is difficult to implicate only the foreign pathway promoters in the expression decreases since the expression of ENO and GAP, whose promoters are used for the recombinant genes, remains high.

6.5 Conclusions and Recommendations

This work focused on the application of 2-DE to study protein expression in both a nonrecombinant and metabolically engineered bacterium *Z. mobilis*. Fifteen proteins including 3 of the 4 metabolically engineered PP pathway enzymes were tentatively identified based on comparison to reference gels and theoretical MWs and pIs. Samples from a time course profile for an ethanol fermentation using the metabolically engineered pathway were used to develop protein expression profiles for a large set of visualized proteins. General trends such as similarities in expression for functional related proteins were detectable using both hierarchical cluster analysis and correlation maps, and

demonstrated the value of applying this class of data analysis methods in order extract biologically significant information from data sets at the level of the proteome. Specifically, it was demonstrated that many functionally similar proteins can show coordinated expression that can be separated by agglomeration. Future work in this area should focus on using both more reliable identification and quantification methods to provide more definite conclusions about levels of the cellular protein. For determining spot identities MALDI-TOF MS would be a sound alternative, while one of many fluorescence stains would be useful in quantifying relative abundance of protein.

6.6 References

Aebersold R, Rist B, Gygi SP (2000). Quantitative proteome analysis: methods and applications. *Ann NY Acad Sci.* 919:33-47.

Aldenderfer MS and Blashfield RK (1984). *Cluster Analysis.* Sage Publications, Beverly Hills, CA.

An H, Scopes RK, Rodriguez M, Keshav KF, Ingram LO (1991). Gel electrophoretic analysis of *Zymomonas mobilis* glycolytic and fermentative enzymes: Identification of alcohol dehydrogenase II as a stress protein. *J. Bacteriol.* 173, 5975-82.

Aristidou AA, Penttilä M. (2000). Metabolic engineering applications to renewable resource utilization. *Curr Opin Biotech.* 11:187-198.

Bringer-Meyer S, Sahm H (1989). Junctions of catabolic and anabolic pathways in *Zymomonas mobilis*: phosphoenolpyruvate carboxylase and malic enzyme. *Appl Microbiol Biotechnol.* 31:529-536.

De Graaf AA, Spriege, K, Wittig R, Laufer B, Schmitz G, Wiechert W, Sprenger G, Sahm H (1999). Metabolic state of *Zymomonas mobilis* in glucose-, fructose-, and xylose-fed continuous cultures as analyzed by ^{13}C and ^{31}P-NMR spectroscopy. *Arch Microbiol.* 171:371-385.

DiMarco A, Romano A (1985). D-Glucose Transport System of *Zymomonas mobilis. Appl Environ Microbiol.* 49:151-157.

Eisen M, Spellman P, Brown P, Botstein D (1998). Cluster analysis and display of genome-wide expression patterns. *PNAS.* 95:14863-14868.

Eliasson A, Christensson C, Wahlbom CF, Hahn-Hägerdal B (2000). Anaerobic Xylose Fermentation by Recombinant *Saccharomyces cervisiae* Carrying *XYL1, XYL2,* and *XKS1* in Mineral Medium Chemostat Cultures. *Appl Environ Microbiol.* 66:3381-3386.

Feldmann S, Sahm H, Sprenger G (1992). Pentose metabolism in *Zymomonas mobilis* wild-type and recombinant strains. *Appl Microbiol Biotech.* 38:354-361.

Gao Q, Zhang M, McMillan J, Kompala D (2002). Characterization of Heterologous and Native Enzyme Activity Profiles in Metabolically Engineered *Zymomonas mobilis* Strains During Batch Fermentation of Glucose and Xylose Mixtures. *Appl Biochem Biotech.* 98-100:341-356.

Görg A and Weiss W (2000). Two-Dimensional Electrophoresis with Immobilized pH Gradients. In: *Proteome Research: Two-Dimensional Gel Electrophoresis and Identification Methods.* (Rabilloud T, Ed) Springer-Verlag, Berlin.

Gygi S, Rist B, Aebersold R (2000). Measuring gene expression by quantitative proteome analysis. *Curr Opin Biotech* 11:396-401.

Ho N, Chen Z., Brainard A (1998). Genetically engineered *Saccharomyces* yeast capable of effective cofermentation of glucose and xylose. *Appl Environ Microbiol.* 64:1852-1859.

Hodge D, Karim MN (2002). Modeling and Advanced Control of Recombinant *Zymomonas mobilis* Fed-Batch Fermentation. *Biotech Prog.* 18:572-579.

Horimoto K, Toh H (2001). Statistical estimation of cluster boundaries in gene expression profile data. *Bioinformatics* 17:1143-1151.

Ingram LO, Aldrich HC, Borges ZCC, Causey TB, Martinez A, Morales F, Saleh A, Underwood SA, Yomano LP, York SW, Zaldivar J, Zhou S (1999). Enteric Bacterial Catalysts for Fuel Ethanol Production. *Biotechnol Prog.* 15:855-866.

Joachimsthal E, Haggett K, Rogers P (1999). Evaluation of Recombinant Strains of *Zymomonas mobilis* for Ethanol Production from Glucose/Xylose Media. *Appl Biochem Biotech.* 77-79:147-157.

Kang HL, Kang HS (1998). A physical map of the genome of ethanol fermentative bacterium *Zymomonas mobilis* ZM4 and localization of genes on the map. *Gene.* 206:223-228.

Kim I, Barrow K, Rogers P (2000). Nuclear Magnetic Resonance Studies of Acetic Acid Inhibition of Rec *Zymomonas mobilis* ZM4(pZB5) *Appl Biochem Biotech.* 84-86:357-370.

Lilley KS, Razzaq A, Dupree P (2002). Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. Curr *Opin Chem Biol.* 2002 6(1):46-50.

Lukashin A, Fuchs R (2001). Analysis of temporal gene expression profiles: clustering by simulated annealing and determining the optimal number of clusters. *Bioinformatics* 17:405-414.

McAtee CP, Hoffman PS, Berg DE (2001). Identification of differentially regulated proteins in metronidozole resistant *Helicobacter pylori* by proteome techniques. *Proteomics.* 1(4):516-21.

Michel GP, Starka J (1986). Effect of ethanol and heat stresses on the protein pattern *of Zymomonas mobilis. J Bacteriol.* 165:1040-1042.

Michel GP, Starka J (1987). Preferential synthesis of stress proteins in stationary *Zymomonas mobilis* cells. *FEMS Microbiol Lett.* 43:361-365.

Michel GP (1993). Cloning and Expression in *Escherichia coli* of the dnaK Gene of *Zymomonas mobilis. J Bacteriol.* 175:3228-3231.

Miklos GLG, Maleszka R (2001). Protein functions and biological contexts. *Proteomics.* 1:169-178.

Mohagheghi A, Evans K, Chou YC, Zhang M (2002). Cofermentation of Glucose, Xylose, and Arabinose by Genomic DNA-Integrated Xylose/Arabinose Fermenting Strain of *Zymomonas mobilis* AX101. *Appl Biochem Biotech.* 98-100:885-898.

Osman Y, Conway T, Bonetti S, Ingram LO (1987). Glycolytic Flux in *Zymomonas mobilis*: Enzyme and Metabolite Levels during Batch Fermentation. *J Bacteriol.* 169:3726-3736.

Pollard JW (1984). Two-Dimensional Polyacrylamide Gel Electrophoresis of Proteins. In: *Methods in Molecular Biology: Proteins.* (Walker JM, Ed) Humana Press, Clifton, NJ.

Rabilloud T and Charmont S (2000). Detection of Proteins on Two-Dimensional Electrophoresis Gels. In: *Proteome Research: Two-Dimensional Gel Electrophoresis and Identification Methods.* (Rabilloud T, Ed) Springer-Verlag, Berlin.

Rabilloud T and Chevallet M (2000). Solubilization of Proteins in Two-Dimensional Electrophoresis. In: *Proteome Research: Two-Dimensional Gel Electrophoresis and Identification Methods.* (Rabilloud T, Ed) Springer-Verlag, Berlin.

Rabilloud T, Strub JM, Luche S, van Dorsselaer A, Lunardi J (2001). A comparison between Sypro Ruby and ruthenium II tris (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels. *Proteomics*. 1(5):699-704.

Righetti PG, Stoyanov AV, and Zhukov MY. (2001). *The Proteome Revisited. Theory and Practice of all Relevant Electrophoretic Steps*. Elsevier, Amsterdam.

Salusjärvi L, Poutanen M, Pitkänen J, Koivistoinen H, Aristidou A, Kalkkinen N, Ruohonen L, Penttilä M (2003). Proteome analysis of recombinant xylose-fermenting *Saccharomyces cerevisiae*. *Yeast*. 20:295-314.

Seo JS, Chong H, Park HS, Yoon KO, Jung C, Kim JJ, Hong JH, Kim H, Kim JH, Kil JI, Park CJ, Oh HM, Lee JS, Jin SJ, Um HW, Lee HJ, Oh SJ, Kim JY, Kang HL, Lee SY, Lee KJ, Kang HS (2005). The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat Biotechnol*. 23(1):63-68.

Sprenger G (1996). Carbohydrate metabolism in *Zymomonas mobilis*: a catabolic highway with some scenic routes. *FEMS Microbiol Lett*. 145:301-307.

Voss T and Haberl P (2000). Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels: consequences for comprehensive data analysis. *Electrophoresis*. 21(16):3345-50.

Yan JX, Sanchez JC, Tonella L, Williams KL, Hochstrasser DF (1999). Studies of quantitative analysis of protein expression in *Saccharomyces cerevisiae*. *Electrophoresis*. 20:738-42.

Yan JX, Harry RA, Spibey C, Dunn MJ (2000). Postelectrophoretic staining of proteins separated by two-dimensional electrophoresis using SYPRO dyes. *Electrophoresis*. 21(17):3657-3665.

Yeung KY, Ruzzo WL (2001). Principle component analysis for clustering gene expression data. *Bioinformatics*. 17:763-774.

Zhang M, Eddy C, Deanda K, Finkelstein M, Picataggio S (1995). Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. *Science*. 267:240-243.

APPENDIX 1: MATLAB Scripts

A. Statistical Models

```
%-------------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  ttest2.m
%Created:               5/12/04
%Last Modified:         12/20/04
%Purpose:               Uses the group comparison t-test for two groups
%                       of unequal sample sizes.  Assumes equal
%                       variances.
%
%-------------------------------------------------------------------------

function f = ttest2(x,y)

n1 = length(x);
n2 = length(y);

y1 = mean(x);
y2 = mean(y);

s1 = var(x);
s2 = var(y);

sp = ((n1-1)*s1+(n2-1)*s2)/(n1+n2-2);

t = abs((y1-y2)/((s1/n1)+(s2/n2))^0.5);

DOF = n1+n2-2

% DOF  t-crit at 95% conf
%  1    6.3140
%  2    2.9200
%  3    2.3530
%  4    2.1320
%  5    2.0150
%  6    1.9430
%  7    1.8950
%  8    1.8600
%  9    1.8330
% 10    1.8120
% 11    1.7960
% 12    1.7820
```

```
% 13     1.7710
% 14     1.7610
% 15     1.7530

f = t;
```

```
%-------------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  linearmodel1.m
%Created:               12/13/04
%Last Modified:         12/21/04
%Purpose:               Generates a linear model correlating conversion
%                       to initial solids and reactor type.  Compares
%                       model terms to determine the correlation
%                       captured by each parameter.
%-------------------------------------------------------------------------

clear all;
x=[15.0000    1.0000   73.0000
    5.0000    1.0000   85.4100
    5.0000    1.0000   83.6700
    13.5000    1.0000   70.0000
    15.0000    1.0000   70.9200
   13.5000    1.0000   70.7000
   13.5000    1.0000   71.4300
   15.0000    1.0000   71.4900
   15.0000    1.0000   73.2600
    5.0000    1.0000   83.5100
   10.0000    1.0000   74.0200
   10.0000    1.0000   74.1200
   10.0000    1.0000   72.6700
   10.0000    1.0000   73.3200
   10.0000    1.0000   78.2500
   10.0000    1.0000   75.3700
   15.0000    1.0000   73.7400
   15.0000   -1.0000   70.5700
   10.0000   -1.0000   81.4900
   15.0000   -1.0000   73.3350
   15.0000   -1.0000   70.0900
   15.0000   -1.0000   65.5550
   15.0000   -1.0000   65.0500
   15.0000   -1.0000   70.8100
   10.0000   -1.0000   76.9500
   10.0000   -1.0000   80.0200
   13.5000   -1.0000   63.9100
   13.5000   -1.0000   69.5500
    5.0000   -1.0000   82.8700
    5.0000   -1.0000   82.2700
    10.0000   -1.0000   81.7200];

%linear model
p = 3;
n = length(x);

a = [x(:,1:2)];
a(:,3)=1;
y = [x(:,3)];

Ymean=mean(y);

m1 = inv(a'*a)*a'*y;
```

```
Ypred = a*m1;

SSR = (y-Ypred)'*(y-Ypred);

SSF = (Ypred-Ymean)'*(Ypred-Ymean);

MSR=SSR/(n-p);
MSF=SSF/(p-1);
F = MSF/MSR
plot(Ypred(1:17),y(1:17),'.')
hold on
plot(Ypred(18:31),y(18:31),'x')
plot([65 87],[65 87]);
legend('SFs','STRs')
axis tight
xlabel('Experimental Conversion (%)')
ylabel('Model Estimate of Conversion (%)')
```

```
%-------------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  linearmodel2.m
%Created:               12/13/04
%Last Modified:         12/21/04
%Purpose:               Generates a log10-transformed linear model
%                       correlating conversion to reaction time,
%                       intitial solids, and reactor type.  Compares
%                       model terms to determine the correlation
%                       captured by each parameter.  Calls the data
%                       file "totaldata.m".
%-------------------------------------------------------------------------

clear all;

totaldata;
x=[STR;SF];

a=log(1.1+x(:,[1,3:4]));

a(:,4)=1;

aa = a(:,[1 4]);

y=log(1+x(:,2));

%linear model
p = 4;                  %number of parameters
n = length(x);          %number of samples

Ymean=mean(y);

m1 = inv(a'*a)*a'*y;
m2 = inv(aa'*aa)*aa'*y;

Ypred = a*m1;
Ypred2 = aa*m2;

SSR = (y-Ypred)'*(y-Ypred);
SSR2 = (y-Ypred2)'*(y-Ypred2);

SSC = (y-Ymean)'*(y-Ymean);

SSF = (Ypred-Ymean)'*(Ypred-Ymean);
SSF2 = (Ypred2-Ymean)'*(Ypred2-Ymean);

MSR=SSR/(n-p);
MSF=SSF/(p-1);
F = MSF/MSR                         %F-statistic for model
significance.
F2 = (SSF2/SSR2)*((n-p)/(p-1))      %F-statistic for parameter
significance.

r1 = SSF/SSC                        %Correlation coefficient for model
r2 = SSF2/SSC                       %Correlation coefficient for each
param.
```

```
plot((exp(Ypred(1:150))-1)*100,(exp(y(1:150))-1)*100,'x')
hold on
plot((exp(Ypred(151:308))-1)*100,(exp(y(151:308))-1)*100,'.')
plot([0 85],[0 85]);
legend('SFs','STRs')
axis tight
xlabel('Experimental Conversion (%)')
ylabel('Model Estimate of Conversion (%)')
```

```
%-------------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  totaldata.m
%Created:               12/13/04
%Last Modified:         12/13/04
%Purpose:               Generates a log10-transformed linear model
%                       correlating conversion to reaction time,
%                       initial solids, and reactor type.  Compares
%                       model terms to determine the correlation
%                       captured by each parameter.
%                       Called by "linearmodel2.m".
%-------------------------------------------------------------------------

clear all;

SF=[0          0.000192073          0.05
2              0.16606609           0.05
4              0.254605904          0.05
8              0.383263201          0.05
24             0.673164782          0.05
30             0.762113875          0.05
48             0.815462778          0.05
72             0.834320601          0.05
94             0.847257862          0.05
118            0.856727939          0.05
0              0.00048745           0.05
2              0.168366566          0.05
4              0.252057252          0.05
8              0.381135783          0.05
24             0.689442379          0.05
30             0.746078197          0.05
48             0.816276174          0.05
72             0.84903172           0.05
94             0.868277957          0.05
118            0.876314567          0.05
0              -0.000679523         0.05
2              0.168296852          0.05
4              0.255280812          0.05
8              0.383001906          0.05
24             0.679045412          0.05
30             0.744811097          0.05
48             0.810174714          0.05
72             0.845071851          0.05
94             0.858457991          0.05
0              0.008434668          0.10
7.75           0.283054488          0.10
18             0.449969865          0.10
26             0.540794585          0.10
44             0.640161251          0.10
74             0.701996715          0.10
96             0.715931314          0.10
120            0.747260739          0.10
0              0.01105365           0.10
7.75           0.286330718          0.10
18             0.426026319          0.10
26             0.521136364          0.10
```

44	0.647869986	0.10
74	0.718256643	0.10
96	0.727741642	0.10
120	0.753045181	0.10
0	0.010984675	0.10
7.75	0.300693939	0.10
18	0.458871519	0.10
26	0.568142369	0.10
44	0.705984217	0.10
74	0.746193958	0.10
96	0.760440275	0.10
120	0.80412684	0.10
0	0.019066233	0.14
7.5	0.240187435	0.14
17	0.384633182	0.14
25	0.491693171	0.14
41	0.607187766	0.14
49	0.627834729	0.14
72	0.652818559	0.14
95	0.679368056	0.14
138	0.71910463	0.14
0	0.016695914	0.14
7.5	0.200813123	0.14
17	0.376419959	0.14
25	0.487466426	0.14
41	0.59063618	0.14
49	0.670412173	0.14
72	0.650221946	0.14
95	0.671102717	0.14
138	0.72373566	0.14
0	0.017775915	0.14
7.5	0.245135068	0.14
17	0.369373085	0.14
25	0.485915579	0.14
41	0.585469726	0.14
49	0.615200999	0.14
72	0.651282998	0.14
95	0.661641887	0.14
138	0.732068056	0.14
0	0.012605673	0.15
6	0.137295983	0.15
12	0.219008967	0.15
24	0.279213165	0.15
31	0.348294105	0.15
48	0.531356198	0.15
72	0.658584319	0.15
144	0.736373968	0.15
168	0.726789959	0.15
192	0.759879374	0.15
216	0.781491522	0.15
240	0.838858938	0.15
0	0.016223175	0.15
6	0.138234551	0.15
12	0.220187281	0.15
24	0.282286201	0.15
31	0.360947878	0.15
48	0.565455046	0.15

72	0.648182451	0.15
144	0.714866147	0.15
168	0.730171048	0.15
192	0.753713869	0.15
216	0.771341163	0.15
240	0.822598201	0.15
0	0.013019357	0.15
6	0.139598293	0.15
12	0.218384983	0.15
24	0.280711882	0.15
31	0.36448087	0.15
48	0.546980046	0.15
0	0.035094965	0.10
7	0.286018633	0.10
24	0.550989842	0.10
45	0.661140014	0.10
96	0.725685913	0.10
120	0.761102672	0.10
144	0.785284956	0.10
168	0.799633909	0.10
0	0.044035571	0.10
7	0.29178308	0.10
24	0.585424804	0.10
45	0.673551627	0.10
96	0.741303466	0.10
120	0.774516182	0.10
144	0.810552212	0.10
168	0.797441118	0.10
0	0.042034849	0.10
7	0.281213775	0.10
24	0.565611819	0.10
45	0.65923266	0.10
96	0.71233768	0.10
120	0.761943613	0.10
144	0.785741142	0.10
168	0.811655837	0.10
0	0.027163143	0.15
18	0.340827605	0.15
24	0.422281152	0.15
120	0.692229544	0.15
144	0.694579214	0.15
168	0.693104112	0.15
192	0.711252546	0.15
0	0.016282426	0.15
18	0.330752475	0.15
24	0.42288878	0.15
120	0.704948499	0.15
144	0.701622831	0.15
168	0.690487606	0.15
192	0.718492506	0.15
0	0.029419614	0.15
18	0.355555092	0.15
24	0.435781058	0.15
40	0.567418774	0.15
48	0.611953131	0.15
75.75	0.655382304	0.15
120	0.685055281	0.15

```
144          0.713621793          0.15
168          0.718323409          0.15
192          0.73221762          0.15];

STR=[0       0.015796866          0.05
2            0.180047769          0.05
4            0.250475672          0.05
8            0.361166725          0.05
24           0.617708606          0.05
30           0.676986862          0.05
48           0.741469737          0.05
72           0.800285954          0.05
94           0.818140406          0.05
118          0.828726047          0.05
0            0.036659249          0.05
2            0.177259783          0.05
4            0.244999107          0.05
8            0.351745032          0.05
24           0.650256991          0.05
30           0.669382433          0.05
48           0.782541533          0.05
72           0.82300092           0.05
94           0.819371335          0.05
118          0.822687268          0.05
0            0.029971734          0.10
7.75         0.301604604          0.10
18           0.44432114           0.10
26           0.531921823          0.10
44           0.653560181          0.10
74           0.809943706          0.10
96           0.804192965          0.10
120          0.817216298          0.10
0            0.030290039          0.10
7.75         0.27699932           0.10
18           0.412556072          0.10
26           0.497586688          0.10
44           0.642676071          0.10
74           0.787015381          0.10
96           0.791646794          0.10
120          0.814916652          0.10
0.00         0.042292531          0.10
7.00         0.250272295          0.10
24.00        0.528616178          0.10
45.00        0.670796722          0.10
96.00        0.72467884           0.10
120.00       0.769481677          0.10
144.00       0.803939998          0.10
168.00       0.83236821           0.10
0.00         0.044412489          0.10
7.00         0.24943177           0.10
24.00        0.527756142          0.10
45.00        0.644660254          0.10
96.00        0.72278398           0.10
120.00       0.800175964          0.10
144.00       0.823660993          0.10
168.00       0.826049169          0.10
0            0.032681789          0.14
```

7.5	0.236854604	0.14
17	0.321378185	0.14
25	0.367369753	0.14
41	0.43659643	0.14
49	0.45871304	0.14
72	0.522622404	0.14
95	0.553382867	0.14
138	0.623085344	0.14
0	0.0432678	0.14
7.5	0.234042784	0.14
17	0.308468707	0.14
25	0.403508998	0.14
41	0.506936691	0.14
49	0.554773426	0.14
72	0.629448835	0.14
95	0.645546367	0.14
138	0.686168844	0.14
0	0.032710457	0.15
6	0.130786641	0.15
12	0.201386476	0.15
24	0.265831598	0.15
31	0.315848935	0.15
48	0.44588252	0.15
72	0.544922107	0.15
144	0.68042968	0.15
168	0.695156914	0.15
192	0.733539322	0.15
216	0.715604564	0.15
240	0.769247375	0.15
0	0.028472099	0.15
6	0.145535201	0.15
12	0.212057532	0.15
24	0.273144276	0.15
31	0.323436205	0.15
48	0.453497908	0.15
72	0.535793076	0.15
144	0.655807489	0.15
168	0.675281939	0.15
192	0.708082887	0.15
216	0.698849647	0.15
240	0.752299481	0.15
0	0.001340276	0.15
18	0.278670689	0.15
24	0.319992718	0.15
40	0.388886405	0.15
48	0.408329106	0.15
75.75	0.482879613	0.15
120	0.59541809	0.15
144	0.634612418	0.15
168	0.657104529	0.15
192	0.705652439	0.15
0	0.0581514	0.15
18	0.278408777	0.15
24	0.320264444	0.15
40	0.38811553	0.15
48	0.413721842	0.15
75.75	0.476320909	0.15

```
120             0.596759482             0.15
144             0.632293193             0.15
168             0.660569937             0.15
192             0.700883805             0.15
0               -0.001340276            0.15
18              0.259305066             0.15
24              0.302660381             0.15
40              0.370299198             0.15
48              0.400901895             0.15
75.75           0.470618694             0.15
120             0.563262788             0.15
144             0.620439405             0.15
168             0.628408933             0.15
192             0.650513217             0.15
0               0.050553896             0.15
18              0.258770227             0.15
24              0.3079574               0.15
40              0.376603361             0.15
48              0.400675482             0.15
75.75           0.474838466             0.15
120             0.562553937             0.15
144             0.620401003             0.15
168             0.626727023             0.15
192             0.655530879             0.15
0               0.04704227              0.10
7               0.23562847              0.10
24              0.402002266             0.10
45              0.522986107             0.10
96              0.72561968              0.10
120             0.764777358             0.10
144             0.809934684             0.10
168             0.828055353             0.10
0               0.035234755             0.10
7               0.234618105             0.10
24              0.399311671             0.10
45              0.520413116             0.10
96              0.727259387             0.10
120             0.762365966             0.10
144             0.804706888             0.10
168             0.830969336            0.10];

SF(:,4) = -1;
STR(:,4) =1;
```

B. Saccharification Models and Optimal Control

```
%-------------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  optimfunction.m
%Date:                  8/1/04
%Purpose:               Contains system of ODEs for the model system in
%                       fed-batch operation
%
%-------------------------------------------------------------------------

%     This file contains the system of ODEs that is called by the
program mpc.m

function J = optimfunction(feed)

global initcond;
global sugarvalue;
Ssp = 0.12+sugarvalue/4800;
[t x] = ode45('saccmodel4',[0 1], initcond',[],feed);

[m n] = size(x);

J = (10000*(Ssp - x(m,8)))^2;
```

```
%-------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  sacchmodel4.m
%Date:                  8/1/04
%Purpose:               Contains system of ODEs for the model system in
%                       fed-batch operation
%
%-------------------------------------------------------------------

%     This file contains the system of ODEs that is called by the
program optimfunction.m

function varargout = saccmodel4(t,x,flag,F)

 switch flag
 case ''                          % Return dx/dt = f(t,x).
    varargout{1} = f(t,x,F);
 otherwise
   error(['Unknown flag ''' flag '''.']);

end
  % ------------------------------------------------------------------
-------

function dxdt = f(t,x,F)
%Fedbatch operation

xsf = 0.45;                      %xsf = 0.30/(1+Evolume);
Eloading = 0.040;                %Enzyme loading in g protein / g
cellulose
Econc = 106.6;                   %Enzyme strength in protein solution
(mg protein/mL)
Etotal = xsf*0.532*1000*Eloading; %Total enzyme concentration (g/kg)
%Evolume = Etotal/Econc;          %Total enzyme volume (
Evolume = 0;
Efeed1 = Etotal*0.975;           %Enzyme 1 in feed (g/kg) (97.5% of
total enzyme)
Efeed2 = Etotal*0.025;           %Enzyme 2 in feed (g/kg) (2.5% of
total enzyme)
Xfeed = 0*(1-xsf)/(1+Evolume);   %Pentose sugars in feed (g/kg)
CBfeed = 0*(1-xsf)/(1+Evolume);  %Cellobiose in feed (g/kg)
Gfeed = 0*(1-xsf)/(1+Evolume);   %Hexose sugars in feed (g/kg)
         %Fraction of insoluble solids in feed

Cfeed = xsf*0.532*1000/(1+Evolume);   %Cellulose in the feed based on
solids of 53% cellulose (g/kg)

competitive = 1;
noncompetitive = 0;

%Variables
S = x(1);     % Cellulose (g/kg)
CB = x(2);    % Cellobiose (g/kg)
G = x(3);     % Glucose (g/kg)
X = x(4);     % Xylose (g/kg)
Et1 = x(5);   %Enzyme 1 (g/kg)
```

```
Et2 = x(6);  %Enzyme 2 (g/kg)
V = x(7);    %Volume (kg)
xs = x(8);   %Insoluble solids (kg/kg)

D = F/V;

% Enzyme adsorption parameters
Emax1 = 0.06;
Kad1 = 0.4;
Emax2 = 0.01;
Kad2 = 0.1;

%-------%quadratic equation to solve Ebi---------
A = Emax1*Kad1*S;
B = Kad1;
C = Et1;
a = B;
b = -(1+B*C+A);
c = A*C;

Eb1 = (-b-(b^2-4*a*c)^0.5)/(2*a);
Ef1 = Et1-Eb1;
%-------------------
AA = Emax2*Kad2*S;
BB = Kad2;
CC = Et2;
aa = 1;
bb = -(1+BB*CC+AA);
cc = AA*CC;

Eb2 = (-bb-(bb^2-4*aa*cc)^0.5)/(2*aa);
Ef2 = Et2-Eb2;
%-----------------------------
if competitive == 1

    %Parameters----------------------------
    k1r   = 35.7;
    K1icb = 0.008;
    K1ig  = 0.049;
    K1ix  = 0.039;
    k2r   = 2.59;
    K2icb = 110;
    K2ig  = 0.000024;
    K2ix  = 93.6;
    k3r   = 38;
    K3m   = 3.1;
    K3ig  = 15.1;
    K3ix  = 32;

    %Kinetic rates ----------------------
    r1 = (k1r*Eb1*S)/(1+(CB/K1icb)+(G/K1ig)+X/(K1ix));
    r2 = (k2r*(Eb1+Eb2)*S)/(1+(CB/K2icb)+(G/K2ig)+X/(K2ix));
    r3 = (k3r*(Ef2)*CB)/(K3m*(1+(G/K3ig)+X/(K3ix))+CB);
end
if noncompetitive == 1
    %Parameters----------------------------
    k1r   = 73.7;
```

```
    K1icb = 3.80;
    K1ig  = 0.015;
    K1ix  = 23.0;
    k2r   = 54.3;
    K2icb = 4050;
    K2ig  = 0.0044;
    K2ix  = 57.7;
    k3r   = 360;
    K3m   = 29.9;
    K3ig  = 4.1;
    K3ix  = 192;

    %Kinetic rates ----------------------
    r1 =(k1r*Eb1*S)*K1icb*K1ig*K1ix/((K1icb+CB)*(K1ig+G)*(K1ix+X));
    r2 =
k2r*(Eb1+Eb2)*S*K2icb*K2ig*K2ix/((K2icb+CB)*(K2ig+G)*(K2ix+X));
    r3 = k3r*Ef2*CB*K3ig*K3ix/((K3m+CB)*(K3ig+G)*(K3ix+X));
end

Rs = 0.001*(-r1-r2)+D*(xsf/(1+Evolume)-xs);
(-r1-r2)/S;
% putting it all together
dxdt = [D*(Cfeed-S)-r1-r2;1.056*r1-r3+D*(CBfeed-CB); 1.111*r2 +
1.053*r3 + D*(Gfeed - G); D*(Xfeed-X); D*(Efeed1 - Et1); D*(Efeed2 -
Et2); F; Rs];
```

```
%-------------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  mpc.m
%Date:                  8/1/04
%Purpose:               Contains system of ODEs for the model system in
%                       fed-batch operation
%
%-------------------------------------------------------------------------

%     This file contains the system of ODEs that is called by the
%     program getflndata.m

clear all;
initialsolids = 0.12;
xsf = 0.45;
V0 = 1;
C0 = initialsolids*0.532*1000;
Eloading = 0.040;    %Enzyme loading in mg protein/ g cellulose
E0 = C0*Eloading;

initcond = [C0 0*(1-initialsolids) 0*(1-initialsolids) 0*(1-
initialsolids) E0*0.975 E0*0.025 V0 initialsolids];

t0 = 0;
x = initcond;
t = t0;
global initcond;

sugarvalue = 0;

global sugarvalue;
for i = 1:300
   i
   %save sss;
   D = fmincon('optimfunction',[0.001],[],[],[],[],[0],[1]);
   [tt xx] = ode45('saccmodel4',[t0 t0+1],initcond,[], D);
   initcond = xx(length(xx),:);
   t0 = tt(length(tt));
   t = [t;tt(length(tt),:)];
   x = [x;xx(length(xx),:)];
   dd(i) = D;
  % if D < 0.001
  %     break
  % end
  sugarvalue = xx(length(xx),3) + xx(length(xx),2);


end

%[tt xx] =ode45('saccmodel4',[t0 t0+1],initcond,[],0);

% t = [t;tt];
% x = [x;xx];

% conversion of cellulose
for i = 1:length(x)
```

```
    Vfed(i) = x(i,7)-V0;
    Cfed(i) = xsf*(Vfed(i)/(1+0.532*1000*xsf*Eloading/106.6))*0.532;
%Cellulose in the feed based on solids of 53% cellulose (g/kg)
    solidsfed(i) = (Cfed(i)/0.532);
    totalsolids(i) = (solidsfed(i)+initialsolids)/x(i,7);
    Conv(i) = 1-x(i,1)/((Cfed(i)*1000+C0*V0)/x(i,7));
end
```

C. Hierarchical Cluster Analysis

```
%-------------------------------------------------------------------
%
%Created by:            David Hodge
%File:                  proclust2.m
%Date:                  10/1/03
%Purpose:               Contains system of ODEs for the model system in
%                       fed-batch operation
%
%-------------------------------------------------------------------

%     This file contains the system of ODEs that is called by the
program getflndata.m

clear all;

data = [0.056258031          0.085699066    0.115414297    0.047483168
0.015354656  0.051704573     0.030488624    0.020398676
0.010821579  0.021513048     0              0.003423627
0.064179884  0.062771128     0.069587331    0.093761451
0.013529306  0.028671295     0.030758461    0.023436897
0.024532209  0.024820445     0.036335659    0.027146835
0            0.03218937      0.006502143    0.003549675
0.051665352  0.044255446     0.029014357    0.046985806
0.0504198    0.018582715     0.035665901    0.055099764
0.040496652  0.031058176     0.047686738    0.034953051
0            0.018026095     0.025377075    0.034748699
0.014497874  0.037820206     0.047822016    0.039273593
0.013040856  0.034772564     0.032541965    0.013000135
0.030020143  0.035991142     0.05478553     0.049841055
0.012884071  0.024820445     0.028490772    0.033645398
0            0               0.002984089    0.007373443
0            0.006286809     0.006691391    0.010459223
0            0               0.009119969    0.004957169
0            0.005817572     0.022471361    0.019028983
0.031190486  0.005537467     0.000610835    0.00042229
0.019874621  0.002154656     0.000993272    0.003503787
0.010932424  0.001666268     0              0.00202432
0.063035835  0.045776873     0.049829153    0.071920338
0.080706699  0.038896337     0.111141256    0.075819315
0.008484833  0.010305243     0.025903984    0.024027374
0.008414354  0.005007182     0.010338864    0.010805992
0.00740856   0.008525257     0.001892813    0.003441634
0.024569475  0.009409864     0.005318457    0.00514822
0.010274055  0.011158726     0.011152503    0.011617241
0.040127113  0.038142207     0.004825133    0.008370811
0.032500269  0.010985157     0.011327553    0.009650543
0.015355863  0.005260953     0.002697635    0.010442467
0            0               0.000458127    0.004167646
0            0.004784534     0              0.006090128
0            0.00599593      0              0.002049878
0            0.00528609      0.004062775    0.003143069
0.004116215  0.00464927      0.029587549    0.002421052
0.013668138  0.013903519     0              0.011136545
```

```
0.00760882   0.009488868     0              0.003396907
0.005254561  0.003542016     0              0
0.013510146  0.003770649     0              0.006462721
0.003125821  0.001522624     0              0
0            0               0.034641709    0.057660346
0.041289731  0.017476658     0.026702037    0.010819844
0.011324877  0.017359349     0.02607058     0.007996157
0.014072212  0.003045248     0.011812362    0.029652504
0.015012515  0.008648552     0              0
0.008917994  0.004916208     0.005156248    0.006517706
0.013615453  0.021841034     0.003586666    0.001027792
0.009601174  0.011288006     0              0.001440276
0.006775979  0.019466124     0.003180992    0.002594928
0.042496954  0.016344266     0.025995946    0.012102792
0.017947176  0.016752454     0.014709847    0.014427832
0.00449634   0.004543931     0.003137171    0.003020326
0            0.006083313     0.004288054    0.006116524
0            0.006892507     0              0
0            0               0.00397994     0.005581546
0.017936803  0.020245391     0              0
0            0               0.004811384    0.008766437
0            0               0.007999952    0.0086804];


d = corrcoef(data');
Y = pdist(d,'euclid');
X = linkage(Y,'average');
dendrogram(X,0)
```

APPENDIX 2: *Z. mobilis* Genes from NCBI Database

A. ED and fermentative Pathway Proteins

pI	MW	Gene	Gene Product
5.66	36094	adhA	alcohol dehydrogenase I
5.49	40155	adhB	alcohol dehydrogenase II
6.59	21121	eda, kdpg	2-keto-3-deoxy-6-phosphogluconate aldolase
6.56	62925	edd	phosphogluconate dehydrogenase
	?????	edp, gapB	glyceraldehyde 3 phosphate dehydrogenase -A complex
4.96	40247	eno	enolase
6.25	36102	gapA	glyceraldehyde 3 phosphate dehydrogenase 2
9.07	50200	glf	glucose facilitator
6.08	35423	glk	glucokinase
5.74	60912	pdc	pyruvate decarboxylase
6.01	38594	pdhA	pyruvate dehydrogenase
4.63	49964	pdhB	pyruvate dehydrogenase
5.89	55399	pgi	phosphoglucose isomerase
5.97	41388	pgk	phosphoglycerate kinase
		pgl	6-phosphogluconolactonase
6.12	25938	pgm	phosphoglucomutase
6.45	51445	pyk	pyruvate kinase
5.61	53859	zwf	glucose-6-phosphate dehydrogenase

B. Heat Shock Proteins

pI	MW	Gene	Gene Product
5.58	58920	groEL	60 kDa heat shock protein
5.65	10291	groES	10 kDa heat shock protein

C. All Other Proteins

A current list of all known proteins can be found at:
http://www.nature.com/nbt/journal/v23/n1/extref/nbt1045-S1.pdf

pI	MW	Gene	Gene Product
			16S ribosomal RNA
			16S ribosomal RNA
			16S ribosomal RNA
			23S ribosomal RNA
			23S ribosomal RNA
			23S ribosomal RNA
			5S ribosomal RNA
			5S ribosomal RNA
			5S ribosomal RNA

5.87	97016	alaS	alanyl-tRNA synthetase
5.39	26958	arcA	aerobic respiration control protein
5.93	51095	argH	arginosuccinate lyase
5.6	64095	argS	arginyl-tRNA synthetase
5.03	25837	asp	aspartate racemase
5.07	24981	asr	aspartate racemase
6.34	40094	carA	carbamoylphosphate synthetase small subunit
5.22	117626	carB	carbamoylphosphate synthetase large subunit
7	24774	cat	chloramphenicol acetyltransferase
9.62	23424	cgpA	GTP-binding protein CgpA
9.77	45303	CLC-b	chloride channel protein
9.64	29526	comF	competence protein F
6.14	50955	copp	oxygen-independent coproporphyrinogen III oxidase
6.41	81141	cpxA	chemotaxis protein
8.72	53263	cycC	cytochrome C peroxidase
9.47	52440	cydA	cytochrome oxidase D subunit A
8.38	37206	cydB	cytochrome oxidase D subunit B
5.65	35878	cysz	cysteine synthase
6.54	33180	dapA	dihydrodipicolinate synthetase
6.06	11448	dcp	putative oxidoreductase
		ddh	2-hydroxyacid dehydrogenase homologue
5.22	35215	ddlB	D-alanine:D-alanine ligase
6.62	45382	dfp	pantothenate metabolism flavoprotein
6.23	44600	dgtP	dGTP triphosphohydrolase
10.04	46938	ditE	putative transporter
5.67	31808	dkg	2,5-diketo-D-gluconate reductase
9.27	34092	dlh	putative carboxymethylenebutenolidase
7.71	32145	dnaA	chromosomal replication initiation protein
		dnaA	chromosomal replication initiator protein
8.92	41425	dprA	DNA processing chain A
9.27	15254	dut	deoxyuridine 5'triphosphate nucleotidohydrolase
5.69	41858	dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase
4.89	20925	efp	elongation factor P efp
5.89	60470	ETF-QO	electrotransfer ubiquinone oxidoreductase
8.61	52759	eutP	ethanolamin permease
9.25	82315	fcuA	ferrichrome receptor FcuA precursor
9.34	29162	fla	flagellin
8.86	19878	flgA	flagellar protein
4.68	12724	flgB	flagellar basal-body rod protein
5.22	14435	flgC	flagellar basal-body rod protein
4.18	23208	flgD	basal-body rod modification protein
4.64	42235	flgE	flagellar hook protein
4.81	26677	flgF	flagellar basal-body rod protein
		flgG	flagellar basal-body rod protein
9.52	23096	flgH	flagellar L-ring protein precursor
7.99	39721	flgI	flagellar P-ring protein precursor
5.88	74985	flhA	flagellar biosynthesis protein
5.58	27631	fliA	sigma-F factor
6.13	11991	fliE	flagellar hook-basal body complex protein
6.13	11991	fliE	flagellar hook-basal body complex protein
8.6	59641	fliF	flagellar M-Ring protein
6.41	66596	fliF	flagellar M-Ring protein
4.52	38066	fliG	flagellar motor switch protein
5.8	47146	fliI	probable H+-transporting ATP synthase
8.94	32801	fmf	methionyl-tRNA formyltransferase
9.37	28698	fpg	formamidopyrimidine-DNA glycosylase

5.05	28941	fpr	ferredoxin-NADP+ reductase
6.98	22486	fprA	pyridoxamine 5'-phosphate oxidase
5.61	44969	ftsA	cell division protein FtsA
6.87	28887	ftsQ	cell division protein FtsQ
		ftsW	cell divisin protein FtsW
4.94	35726	ftsZ	cell division protein FtsZ
7.1	71426	fyuA	outer membrane protein FyuA
6.14	73380	gaa	glutaryl-7-ACA acylase precursor
5.79	49667	gabD	succinic semialdehyde dehydrogenase
		gfo	glucose-fructose oxidoreductase
6.00	168965	gltB	glutamine-pyruvate aminotransferase gltB
6.17	167519	gltB	glutamate synthase large subunit
7.48	53553	gltS	glutamate synthase small subunit gltS
8.43	47561	gntP	gluconate permease
5.04	17470	greA	transcription elongation factor
		grp	glutamate uptake regulatory protein
5.48	52545	gsh	gamma-glutamylcysteine synthetase
5.71	91678	gyrB	DNA gyrase subunit B
5.17	34666	hdh	histidol dehydrogenase
5.58	31724	hdl	hydrolase
5.64	37354	hemA	aminolevulinic acid synthase
10.32	12757	himA	
6.41	14150	hisE	phosphoribosyl c-AMP cyclohydrolas
6.98	24442	hisG	ATP phosphoribosyltransferase
		hpnA	related to dihydroflavonol 4-reductases
		hpnB	related to glucosaminyl transferases
		hpnC	squalene synthase
		hpnD	squalene synthases
		hpnE	phytoene desaturases
5.99	66658	ilvB	acetolactate synthase large subunit
5.89	36575	ilvC	acetohydroxy acid isomeroreductase
5.31	18302	ilvN	acetolactate synthase small subunit
4.72	46100	invB	extracellular sucrase
7.28	62743	kefC	glutathione-regulated potassium efflux system protein
9.1	64888	kpsC	capsule polysaccharide export protein
4.76	46746	levU	levansucrase
		lig	DNA ligase
8.8	35666	lipA	lipoic acid synthase
9.08	33468	ltrA	transcriptional regulator
6.23	45249	lysA	diaminopimelate decarboxylase
6.15	17271	mad	hydroxymyristol acyl carrier protein dehydrolase
6.41	29849	map	methionine aminopeptidase
5.39	27898	mdh	morphine 6-dehydrogenase
5.66	20776	metF	5,10-methylenetetrahydrofolate reductase
9.72	33230	mexT	mexT
6.06	31146	modD	molybdenum transport system protein
6.91	26905	moeB	molybdopterin biosynthesis protein
8.89	24834	motA	motility protein A
		mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase
6.79	33905	murB	UDP-N-acetylenolpyruvoylglucosamine reductase
5.81	52602	murC	UDP-N-acetylmuramate-alanine ligase
		murD	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase

8.23	51828	murE	UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase
9.1	38866	murF	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase
		murG	UDP-N-acetylglucosamine:N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
5.31	24146	murI	glutamate racemase
5.3	39665	mviM	mviM rrnB operon
7.78	37940	nifR3	NifR3-like protein
6.4	41257	nifV	homocitrate synthase
8.72	34283	nosX	nosX
5.1	41225	ntrC	nitrogen assimilation regulatory protein
5.39	52478	ntrX	nitrogen assimilation regulatory protein
9.28	85551	ntrY	nitrogen regulation protein
4.93	14814	nusB	N-utilization substance protein B
9.3	19969	omp	20kD outer membrane protein
6.54	102199	omp	outer membrane protein
10	43331	opdE	transcriptional regulatory protein
5.19	23831	parA	partitioning protein
7.82	59419	pbp	penicillin-binding protein
9.32	33172	pchR	regulator protein PchR
5.23	20424	pdf	polypeptide deformylase
5.71	93988	pepN	aminopeptidase N
5.6	97501	pepN	membrane alanyl aminopeptidase
6.04	86113	pfl	pyruvate formate lyase, formate acetyltransferase
6.44	29015	pflA	pyruvate formate lyase activating enzyme
5.53	26759	phoB	phosphate regulatory protein PhoB
5.69	46839	phoR	phosphate regulon sensor protein PhoR
5.34	26271	phoU	transcriptional regulator PhoU
5.01	47488	pmbA	
5.64	49903	pmm	phosphomannomutase pmm
7.77	29986	pol	DNA polymerase
5.98	17132	PPIASE	peptidyl prolyl cis trans isomerase
4.76	19850	psp	phosphoserine phosphatase Psp
9.12	24792	pssA	CDP-diacylglycerol--serine O-phosphatidyltransferase
5.94	54613	purF	amidophosphoribosyltransferase
4.99	77679	pur-q	phosphoribosylformyl glycinamidine synthetase
		pyk	pyruvate kinase
7.56	24029	pyrH	uridylate kinase
5.42	38458	queA	S-adenosylmethionine tRNA ribosyltransferase
5.81	36807	recA	recombinase
6.03	35162	recF	DNA/ATP binding protein
5.04	10589	relB	negative regulator of translation
5.65	34320	ribF	riboflavin kinase ribF
6.59	24081	RNaseIII	ribonuclease III
8.43	17251	rnh	ribonuclease H
5.89	27380	rnsB	30S ribosomal protein S2
9.27	42153	rpfN	rpfN
5.96	87974	rpoD	RNA polymerase sigma factor
9.96	23367	rpsD	30S ribosomal protein S4
8.71	20652	rrf	ribosome recycling factor
		sacA	sucrase
6.81	47220	sah	succinylarginine dihydrolase
8.18	47243	sah	succinylarginine dihydrolase

4.9	35804	serB	
		shc	squalene-hopene cyclase
6.08	36586	tal	low specificity L-threonine aldolase
		tgt	tRNA guanine transglycosylase
6.51	34845	thiL	thiamin-monophosphate kinase
9.68	29755	thyA	thymidylate synthetase
6.16	131474	topA	topoisomerase I
8.56	43573	torf	transcriptional activator
			tRNA-Ala
			tRNA-Ala
			tRNA-Ala
			tRNA-Ala
			tRNA-Asn
			tRNA-Il
			tRNA-Ile
			tRNA-Ile
			tRNA-Lys
			tRNA-Met
			tRNA-Met
			tRNA-Met
6.76	19963		tRNA-Pro
7.02	21236	trpR	trp repressor binding protein
5.67	36749	trpS	tryptophanyl-tRNA synthase
9.5	27412	truA	tRNA-pseudouridine synthase I
5.15	34280	trxB1	thioredoxin reductase
5.27	32123	tsf	elongation factor Ts
5.54	46406	tylC	hemolysin
7.83	27351	ubm	ubiquinone methyltransferase
7.16	29096	udk	undecaprenol kinase udk
5.67	32142	UDPGP	UTP-glucose-1-phosphate uridyltransferase
5.78	47312	ugd	UDP-glucose dehydrogenase
6.74	101649	uvrA	excinuclease ABC subunit A
9.85	27989	vacJ	lipoprotein precursor
9.52	60753	xseA	exodeoxyribonuclease VII
6.65	43993	ygcA	RNA methyltransferase
5.51	68621	yheS	hypothetical ABC transporter ATP-binding protein
8.99	65970	yidC	60KD inner-membrane protein yidC
9.68	51864	yjcC	tetracenomycin C resistance and export protein
9.25	30140	ylxH-1	ATP-binding protein
5.59	42560	yugJ	NADH-dependent butanol dehydrogenase"
		zliE	expression activator protein
		zliS	secretion activator protein
7.85	17411	zrp	
9.23	17377		nitrogen fixation protein
5.77	26621		beta-hydroxysteroid dehydrogenase
9.33	44543		aspartate aminotransferase A
6.81	44287		tRNA guanine transglycosylase
5.79	84643		tRNA guanine transglycosylase
5.79	82089		DNA ligase
8.94	57859		partial ATP-dependent protease Lon
5.76	17207		aspartokinase subunit B
5.92	34761		trans-2-enoyl-ACP reductase II
5.79	84643		phosphotransferase system enzyme I
10.69	20866		2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase
			Mfd protein

8.93	75018	ATP-dependent DNA helicase RecG
10.37	57879	virulence factor mviN
4.33	18384	protein-export protein secB (secB)
9.41	51982	outer membrane lipoprotein
5.2	40062	succinyl-diaminopimelate desuccinylase
9.02	83047	ferrichrome-iron receptor
		putative membrane protein
8.95	31209	prolipoprotein diacylglyceryl transferase
		metabolite transport protein
11.06	10416	50S ribosomal protein L28
5.66	23473	NADH dehydogenase
5.76	44388	phosphoribosylamine-glycine liase
9.39	28342	plasmolipin
9.2	37761	ferrochelatase
8.31	48437	mannosyltransferase B
6.1	28318	ribitol dehydrogenase
5.79	29399	pyrroline-5-carboxylate reductase
5.75	41676	amino acid amido hydrolase
9.32	52533	cardiolipin synthase
5.92	26640	3-oxoacyl-(acylcarrier protein) reductase
5.63	25171	oxidoreductase
9.26	66865	lipoprotein inner membrane ABC-transporter
6.11	62576	inner membrane ABC transporter
7.9	32646	ABC binding protein
5.39	41889	NADH dehydrogenase
		aspartate aminotransferase A
4.91	25614	thiamine-phosphate pyrophosphorylase
5.79	75802	cystathionine-gamma-lyase
5.57	106288	isoleucyl-tRNA synthetase
		dihydrofolate reductase typeIII
		3-oxoacyl-(acyl carrier ptn) reductase
6.2	52096	metabolite transport protein
8.85	35162	IPP-transferase
9.16	35087	IPP-transferase
		30S ribosomal protein S20
		tRNA delta(2)-isopentenylpyrophosphate transferase
6.92	70596	DNA primase
6.07	37059	triacylglycerol lipase

APPENDIX 3: SACCHARIFICATION EXPERIMENTAL DESIGNS

A. Washed Solids Experiments

Fed-batch saccharification-stirred tank					Cellulase Activity:		Enzyme	Activity	Protein mg/n FPU/mg protein			
Experiment #1							Spezyme	28.40	106.60	0.27	Enhanced BCA protocol	
Corn stover PS020502CS		Glucan	52.00	Solids	0.303	corn stover						
Tank Number	Sample Size g	Total Insoluble Solids wt%	Dry Corn Stover g	Wet Corn Stover g	Glucan g	Cellulase Loading mg protein/g	Cellulase Loading FPU/g cellulo	Initial Insoluble Solids wt%	Initial Dry Corn Stover g	Initial Wet Corn Stover g	4 increments Incremental Wet Corn Stover g	4 increments Incremental 1:1 Undiluted Cellulase mL
1	2750	20.0	550.0	1815.78	286.00	20	5.3	8.0	220.0	726.31	272.4	8.0
2	2000	20.0	400.0	1320.57	208.00	20	5.3	20.0	400.0	1320.57	0.0	0.0
3	2750	20.0	550.0	1815.78	286.00	40	10.7	8.0	220.0	726.31	272.4	16.1
4	2000	20.0	400.0	1320.57	208.00	40	10.7	20.0	400.0	1320.57	0.0	0.0
				6272.7						4093.76	544.73	24.15

Tank Number	Cellulase mL	1.0 1:1 Undiluted Cellulase mL	1M Citrate Buffer mL	Water mL	Weight Before Autoclave g	Actual Water Added g	Error In Water Volume g	Weight After Autoclave g	Water Correction g	Actual Wet Corn Stover g	Actual Glucan Added g	Initial 1:1 Undiluted Cellulase mL
1	53.66	53.66	137.5	743.1					0.00		0.00	21.46
2	39.02	39.02	100.0	540.4					0.00		0.00	39.02
3	107.32	107.32	137.5	689.4					0.00		0.00	42.93
4	78.05	78.05	100.0	501.4					0.00		0.00	78.05
	278.0	278.0	475.0	2474.3								181.5

Tank Number	Time, h 0	24	32	48	72	Initial Insoluble Solids wt%
1	8.0	11.0	14.0	17.0	20.0	13.5
2	20.0					20.0
3	8.0	11.0	14.0	17.0	20.0	13.8
4	20.0					20.0
Total daily additions		2	2	2	2	

Fed-batch saccharification-shake flasks					Cellulase Activity:		Enzyme	Activity	Protein mg/n FPU/mg protein			
Experiment #2							Spezyme	28.40	106.60	0.27	Enhanced BCA protocol	
Corn stover PS020502CS	Glucan	52.00	Solids	0.301	corn stover							

Shake Flask	Initial Sample Size g	Final Sample Size g	Initial Insoluble Solids wt%	Initial Dry Corn Stover g	Initial Wet Corn Stover g	Initial Glucan g	Initial Cellulase Loading mg protein/g c	Cellulase Loading FPU/g cellulo	Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	Final equiv Insoluble Solids wt%
1	100	155.01	15.0	15.00	49.80	7.80	20	5.3	1.46	7.75	40.99	20.1%
2	100	155.01	15.0	15.00	49.80	7.80	20	5.3	1.46	7.75	40.99	20.1%
3	100	176.66	15.0	15.00	49.80	7.80	40	10.7	2.93	8.83	38.44	20.8%
4	100	176.66	15.0	15.00	49.80	7.80	40	10.7	2.93	8.83	38.44	20.8%
7	100	133.20	16.6	16.60	55.11	8.63	20	5.3	2.60	6.66	35.63	20.0%
8	100	133.20	16.6	16.60	55.11	8.63	20	5.3	2.60	6.66	35.63	20.0%
9	100	100.00	20.3	20.27	67.30	10.54	20	5.3	1.98	5.00	25.72	20.3%
10	100	100.00	20.3	20.27	67.30	10.54	20	5.3	1.98	5.00	25.72	20.3%
11	100	100.00	21.3	21.30	70.72	11.08	40	10.7	4.16	5.00	20.13	21.3%
12	100	100.00	21.3	21.30	70.72	11.08	40	10.7	4.16	5.00	20.13	21.3%
					585.5				26.24	66.49	321.82	

Shake Flask	Feed #1 t = 8 hr total mass (g)	Feed #1 Wet Corn Stover (g)	Feed #1 Undiluted Cellulase (mL)	Feed #2 t = 12 hr total mass (g)	Feed #2 Wet Corn Stover (g)	Feed #2 Undiluted Cellulase (mL)	Feed #3 t = 24 hr total mass (g)	Feed #3 Wet Corn Stover (g)	Feed #3 Undiluted Cellulase (mL)	total PCS	total enzyme
1	9.75	9.47	0.28	8.49	8.25	0.24	9.46	9.19	0.27	26.91	0.79
2	9.75	9.47	0.28	8.49	8.25	0.24	9.46	9.19	0.27	26.91	0.79
3	14.56	13.75	0.81	13.46	12.71	0.75	14.75	13.93	0.82	40.40	2.37
4	14.56	13.75	0.81	13.46	12.71	0.75	14.75	13.93	0.82	40.40	2.37
7	0.00	0.00	0.00	0.00	0.00	0.00	8.30	8.30	0.00	8.30	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	8.30	8.30	0.00	8.30	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	48.62	46.45	2.17	43.90	41.92	1.98	65.02	62.84	2.18	151.21	6.33

Shake Flask	Feed #4 t = 36 hr total mass (g)	Feed #4 Wet Corn Stover (g)	Feed #4 Undiluted Cellulase (mL)	Feed #5 t = 54 hr total mass (g)	Feed #5 Wet Corn Stover (g)	Feed #5 Undiluted Cellulase (mL)	Feed #6 t =72 hr total mass (g)	Feed #6 Wet Corn Stover (g)	Feed #6 Undiluted Cellulase (mL)	total PCS	total enzyme
1	7.85	7.63	0.22	9.87	9.59	0.28	9.59	9.32	0.27	26.53	0.78
2	7.85	7.63	0.22	9.87	9.59	0.28	9.59	9.32	0.27	26.53	0.78
3	11.20	10.58	0.62	12.37	11.68	0.69	10.32	9.75	0.57	32.01	1.88
4	11.20	10.58	0.62	12.37	11.68	0.69	10.32	9.75	0.57	32.01	1.88
7	8.30	8.30	0.00	8.30	8.30	0.00	8.30	8.30	0.00	24.90	0.00
8	8.30	8.30	0.00	8.30	8.30	0.00	8.30	8.30	0.00	24.90	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	54.70	53.01	1.69	61.08	59.14	1.94	56.42	54.73	1.69	166.88	5.32
										318.09	11.65

Fed-batch saccharification-shake flasks — **Cellulase Activity:** — Enzyme: Spezyme — Activity: 28.40 — Protein mg/n: 106.60 — FPU/mg protein: 0.27 — Enhanced BCA protocol

Experiment #3

Corn stover PS020502CS

Glucan: **52.00** — Solids: **0.271** — corn stover

Shake Flask	Initial Sample Size g	Final Sample Size g	Initial Insoluble Solids wt%	Initial Dry Corn Stover g	Initial Wet Corn Stover g	Initial Glucan g	Initial Cellulase Loading mg protein/g c	Cellulase Loading FPU/g cellulo	Initial Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	Final equiv Insoluble Solids wt%
1	100.00	133.03	10.0%	10.0	36.94	5.20	20	5.3	0.98	6.65	55.43	14.1%
2	100.00	133.03	10.0%	10.0	36.94	5.20	20	5.3	0.98	6.65	55.43	14.1%
3	100.00	144.53	10.0%	10.0	36.94	5.20	40	10.7	1.95	7.23	53.88	14.8%
4	100.00	144.53	10.0%	10.0	36.94	5.20	40	10.7	1.95	7.23	53.88	14.8%
5	100.00	101.01	20.1%	20.1	74.11	10.43	20	5.3	0.94	5.05	19.90	20.1%
6	100.00	101.01	20.1%	20.1	74.11	10.43	20	5.3	0.94	5.05	19.90	20.1%
7	100.00	102.41	20.8%	20.8	76.97	10.83	40	10.7	1.66	5.12	16.25	20.8%
8	100.00	102.41	20.8%	20.8	76.97	10.83	40	10.7	1.66	5.12	16.25	20.8%
9	100.00	100.00	14.1%	14.1	52.09	7.33	20	5.3	1.38	5.00	41.54	14.10%
10	100.00	100.00	14.1%	14.1	52.09	7.33	20	5.3	1.38	5.00	41.54	14.10%
11	100.00	100.00	14.8%	14.8	54.67	7.70	40	10.7	2.89	5.00	37.44	14.80%
12	100.00	100.00	14.8%	14.8	54.67	7.70	40	10.7	2.89	5.00	37.44	14.80%
total	1200.00	1361.96	1.80	179.59	663.44			95.91	19.58	68.10	448.88	

Shake Flask	Feed #1 t = 8 hr total mass (g)	Feed #1 Wet Corn Stover (g)	Feed #1 Undiluted Cellulase (mL)	Feed #2 t = 12 hr total mass (g)	Feed #2 Wet Corn Stover (g)	Feed #2 Undiluted Cellulase (mL)	Feed #3 t = 24 hr total mass (g)	Feed #3 Wet Corn Stover (g)	Feed #3 Undiluted Cellulase (mL)	total PCS	total enzyme
1	0.00	0.00	0.00	0.00	0.00	0.00	14.41	14.04	0.37	14.04	0.37
2	0.00	0.00	0.00	0.00	0.00	0.00	14.41	14.04	0.37	14.04	0.37
3	0.00	0.00	0.00	0.00	0.00	0.00	22.02	20.92	1.10	20.92	1.10
4	0.00	0.00	0.00	0.00	0.00	0.00	22.02	20.92	1.10	20.92	1.10
5	0.18	0.00	0.18	0.16	0.00	0.16	0.17	0.00	0.17	0.00	0.51
6	0.18	0.00	0.18	0.16	0.00	0.16	0.17	0.00	0.17	0.00	0.51
7	0.46	0.00	0.46	0.42	0.00	0.42	0.46	0.00	0.46	0.00	1.34
8	0.46	0.00	0.46	0.42	0.00	0.42	0.46	0.00	0.46	0.00	1.34
Total	1.27	0.00	1.27	1.16	0.00	1.16	74.14	69.91	4.23	69.91	6.66

Shake Flask	Feed #4 t = 36 hr total mass (g)	Feed #4 Wet Corn Stover (g)	Feed #4 Undiluted Cellulase (mL)	Feed #5 t = 48 hr total mass (g)	Feed #5 Wet Corn Stover (g)	Feed #5 Undiluted Cellulase (mL)	Feed #6 t =72 hr total mass (g)	Feed #6 Wet Corn Stover (g)	Feed #6 Undiluted Cellulase (mL)	Feed #7 t =96 hr total mass (g)	Feed #7 Wet Corn Stover (g)	Feed #7 Undiluted Cellulase (mL)
1	0.00	0.00	0.00	7.75	7.55	0.20	6.05	5.89	0.16	4.82	4.70	0.12
2	0.00	0.00	0.00	7.75	7.55	0.20	6.05	5.89	0.16	4.82	4.70	0.12
3	0.00	0.00	0.00	10.73	10.19	0.54	6.98	6.63	0.35	4.80	4.56	0.24
4	0.00	0.00	0.00	10.73	10.19	0.54	6.98	6.63	0.35	4.80	4.56	0.24
5	0.14	0.00	0.14	0.18	0.00	0.18	0.18	0.00	0.18	0.00	0.00	0.00
6	0.14	0.00	0.14	0.18	0.00	0.18	0.18	0.00	0.18	0.00	0.00	0.00
7	0.35	0.00	0.35	0.39	0.00	0.39	0.32	0.00	0.32	0.00	0.00	0.00
8	0.35	0.00	0.35	0.39	0.00	0.39	0.32	0.00	0.32	0.00	0.00	0.00
Total	0.99	0.00	0.99	38.10	35.48	2.62	27.06	25.05	2.01	19.24	18.51	0.73

Shake Flask	total PCS	total enzyme
1	18.14	0.48
2	18.14	0.48
3	21.38	1.13
4	21.38	1.13
5	0.00	0.50
6	0.00	0.50
7	0.00	1.06
8	0.00	1.06
Total	60.53	5.62
Grand total	130.44	12.28

Experiment #4
2/19/2004
David Hodge

D/d = tank diameter/impeller diameter
All work performed in Bioflo 3000 reactors with 4000 g of slurry

D/d	impellers	rpm	T_1	T_1 corrected	T_2 top wall	bot wall	top center	top impell	bot impell	% insol	Observations
1.48	2 marine	50	44.1	42.19	40	60	40	40.7	42.3	6.76%	Difficult to control T, 1.5 cm of liquid phase at top, some
1.48	2 marine	50	44.2	42.29	41.7	55.6	41.3	42	41.9	6.76%	
1.48	2 marine	100	44.6	42.69	42.4	60.9	42.2	42.2	42.5	6.76%	Complete mixing, T control OK
1.48	2 marine	125	44.5	42.59	42.5		42.5	42.5		6.76%	Complete mixing, T control OK
1.48	2 marine	250	45.8	43.90	43.4	43.7	43.3	43.5	43.5	6.76%	Complete mixing, T control OK

D/d	impellers	rpm	T_1	T_1 corrected	T_2 top wall	bot wall	top center	top impell	bot impell	% insol	Observations
2.23	2 marine	50	43.3	41.38	38.2	62	40.4	40.5	41.2	6.76%	Completely stagnant, 3 cm of liquid phase at top, T contr
2.23	2 marine	50	44	42.09	39.2	59.4	38.8	42.8	43.3	6.76%	
2.23	2 marine	175	50.2	48.33	39.9	70	47.1	48	48.3	6.76%	Continuous turnover of liquid, 1.5 - 3 cm of liquid phase i
2.23	2 marine	250	45	43.09	43	43.3	43	43	43	6.76%	Complete mixing, T control OK

D/d	impellers	rpm									Observations
2.23	1 mar, 1 rushton	50								6.76%	Completely stagant, liquid phase at top
2.23	1 mar, 1 rushton	100								6.76%	Complete mixing, no liquid phase at top
2.23	1 mar, 1 rushton	175								6.76%	Complete mixing, no liquid phase at top

D/d	impellers	rpm	T_1	T_1 corrected	T_2 top wall	bot wall	top center	top impell	bot impell	% insol	Observations
1.48	2 marine	100	28.6	26.58	29.44	72.78	29.44	28.33	28.33	13.51%	Completely stagnant
1.48	2 marine	300	31.3	29.30	35.56	71.67	32.78		31.67	13.51%	some movement of solids along walls, T control problem
1.48	2 marine	400	36.4	34.44	46.67	75.00	37.22	37.22	31.67	13.51%	Solids moving up walls until 7 cm from top, T control OK
1.48	2 marine	400	42.2	40.27	43.33	68.89	42.78		43.33	13.51%	
1.48	2 marine	500	46.3	44.40	45.56	48.33	46.11		46.67	13.51%	Solids moving up walls until 3 cm from top, T control OK
1.48	2 marine	500	40.6	38.66	44.44	70.00	41.67		40.56	13.51%	

D/d	impellers	rpm	T_1	T_1 corrected	T_2 top wall	bot wall	top center	top impell	bot impell	% insol	Observations
1.48	2 marine	100	20.7		17.6	61.1	16.8	16.8	16.6	10.23%	Completely stagnant
1.48	2 marine	200	24.1		23.6	63.3	21.9	21.9	21.1	10.23%	Movement of solids up walls until 5 cm from top, entire s
1.48	2 marine	300	29.9		28.9	70.3	28.5	28.2	25.2	10.23%	Some mixing, T control possible
1.48	2 marine	400	37.2		32.2	59.3	34.9	35.2	34.5	10.23%	Complete mixing, small stagnant zone by T probe
1.48	2 marine	500	46		43.2	40.7	43.2	43.2	43.3	10.23%	Cooling, Complete mixing, small stagnant zone by T prol
1.48	2 marine	500	44.7		42.9	53.4	42.6	42.6	42.5	10.23%	Heating, Complete mixing, small stagnant zone by T pro

Water only

D/d	impellers	rpm	T_1	corrected	top wall	bot wall	top center	top impell	bot impell	% insol	Observations
1.48	2 marine	200	22.9		23.33	23.33	23.33	23.33	23.33	0%	No problem with temperature control
1.48	2 marine	200	26.7		26.67	27.22	26.67	26.67	26.67	0%	Heating jacket T not hot enough to burn skin
1.48	2 marine	200	31.3		31.67	31.67	31.67	31.67	31.67	0%	
1.48	2 marine	200	38.1		38.33	38.33	38.33	38.33	38.33	0%	
1.48	2 marine	200	42.7		42.78	43.33	42.78	42.78	42.78	0%	
1.48	2 marine	200	44.8		45.00	45.00	45.00	45.00	45.00	0%	

Enzymatically hydrolyzed PCS solids, ~80% cellulose conversion gives <10% cellulose by dry weight fraction

D/d	impellers	rpm	T_1	corrected	top wall	bot wall	top center	top impell	bot impell	% insol	Observations
2.23	2 marine	300	45.0		44.44	45.56	44.44	46.11	45.56	14.10%	No problem with temperature control

Enzymatically hydrolyzed PCS solids, ~80% cellulose conversion gives <10% cellulose by dry weight fraction

D/d	impellers	rpm	T_1	corrected	top wall	bot wall	top center	top impell	bot impell	% insol	Observations
1.48	2 marine	400	45.0			48.89	45.00	45.00		16.36%	Difficulty in maintaining T
1.48	2 marine	400	44.7		45.00	52.78	45.56	46.11		16.36%	
1.48	2 marine	500	45.0			45.00	45.00	45.00		16.36%	T control possible
1.48	2 marine	500	45.0			45.56	45.00	45.00		16.36%	

Fed-batch saccharification-shake flasks				Cellulase Activity:		Enzyme	Activity	Protein mg/n	FPU/mg protein	
Experiment #5						Spezyme	28.40	106.60	0.27	Enhanced BCA protocol
Corn stover PS020502CS			Glucan	Solids						
			52.00	0.255	corn stover					
Shake Flask	Sample Size g	Insoluble Solids wt%	Dry Corn Stover g	Wet Corn Stover g	Glucan g	Cellulase Loading mg protein/g c	Cellulase Loading FPU/g cellulo	Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	Temperature degrees C
1	100.00	10.0%	10.0	39.23	5.20	20	5.3	0.98	5.00	54.79	38.0
2	100.00	10.0%	10.0	39.23	5.20	20	5.3	0.98	5.00	54.79	38.0
3	100.00	10.0%	10.0	39.23	5.20	40	10.7	1.95	5.00	53.82	38.0
4	100.00	10.0%	10.0	39.23	5.20	40	10.7	1.95	5.00	53.82	38.0
5	100.00	10.0%	10.0	39.23	5.20	20	5.3	0.98	5.00	54.79	45.0
6	100.00	10.0%	10.0	39.23	5.20	20	5.3	0.98	5.00	54.79	45.0
7	100.00	10.0%	10.0	39.23	5.20	40	10.7	1.95	5.00	53.82	45.0
8	100.00	10.0%	10.0	39.23	5.20	40	10.7	1.95	5.00	53.82	45.0
9	100.00	10.0%	10.0	39.23	5.20	20	5.3	0.98	5.00	54.79	50.0
10	100.00	10.0%	10.0	39.23	5.20	20	5.3	0.98	5.00	54.79	50.0
11	100.00	10.0%	10.0	39.23	5.20	40	10.7	1.95	5.00	53.82	50.0
12	100.00	10.0%	10.0	39.23	5.20	40	10.7	1.95	5.00	53.82	50.0
total	1200.00	1.20	120.00	470.77				17.56	60.00	651.67	

Experiment #6 — **Cellulase Activity:** — **Enzyme** Spezyme — **Activity** 28.40 — **Protein mg/n** 106.60 — **FPU/mg protein** 0.27 Enhanced BC

Corn stover PS030312CS

Shake Flask	Sample Size g	Insoluble Solids wt%	Glucan **53.20** Dry Corn Stover g	Solids **0.260** Wet Corn Stover g	corn stover Glucan g	glucose (g/kg)	cellobiose (g/kg)	dens = 1.07 glucose (200 g/L) (mL)	dens = 1.07 cellobiose (200 g/L) (mL)
1	100.00	10.0%	10.0	38.46	5.32	0.0	0	0.00	0.00
2	100.00	10.0%	10.0	38.46	5.32	0.0	0	0.00	0.00
3	100.00	10.0%	10.0	38.46	5.32	25.0	0	13.42	0.00
4	100.00	10.0%	10.0	38.46	5.32	25.0	0	13.42	0.00
5	100.00	10.0%	10.0	38.46	5.32	50.0	0	26.84	0.00
6	100.00	10.0%	10.0	38.46	5.32	50.0	0	26.84	0.00
7	100.00	10.0%	10.0	38.46	5.32	100.0	0	53.69	0.00
8	100.00	10.0%	10.0	38.46	5.32	100.0	0	53.69	0.00
9	100.00	10.0%	10.0	38.46	5.32	80.0	18.94	42.95	10.17
10	100.00	10.0%	10.0	38.46	5.32	80.0	18.94	42.95	10.17
11	100.00	10.0%	10.0	38.46	5.32	90.0	9.47	48.32	5.08
12	100.00	10.0%	10.0	38.46	5.32	90.0	9.47	48.32	5.08
total	1200.00	1.20	120.00	461.54				370.43	30.50

Shake Flask	Cellulase Loading mg protein/g c	Cellulase Loading FPU/g cellulo	Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	Total (g)
1	20	5.3	1.00	5.00	55.54	100.00
2	20	5.3	1.00	5.00	55.54	100.00
3	20	5.3	1.00	5.00	42.12	100.00
4	20	5.3	1.00	5.00	42.12	100.00
5	20	5.3	1.00	5.00	28.70	100.00
6	20	5.3	1.00	5.00	28.70	100.00
7	20	5.3	1.00	5.00	1.86	100.00
8	20	5.3	1.00	5.00	1.86	100.00
9	20	5.3	1.00	5.00	2.42	100.00
10	20	5.3	1.00	5.00	2.42	100.00
11	20	5.3	1.00	5.00	2.14	100.00
12	20	5.3	1.00	5.00	2.14	100.00
total			11.98	60.00	265.55	

			Cellulase Activity:		Enzyme	Activity	Protein mg/n	FPU/mg protein		
Experiment #7, 3/30/04-4/6/04					Spezyme	28.40	106.60	0.27	Enhanced BCA protocol	
Corn stover PS020502CS		Glucan	Solids							
		52.00	0.248	← Pretreated corn stover						
		36.03	0.935	← Raw corn stover						

Shake Flask	Sample Size g	Insoluble Solids wt%	Dry Corn Stover g	Wet Corn Stover g	Glucan g	tetracycline (mL)	Cellulase Loading mg protein/g	Cellulase Loading FPU/g cellulo	Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	Total (g)
1	200.00	10.0%	20.0	80.68	10.40	0.8	20	5.3	1.95	10.00	106.57	200.00
2	200.00	10.0%	20.0	80.68	10.40	0.8	20	5.3	1.95	10.00	106.57	200.00
3	200.00	10.0%	20.0	80.68	10.40	0.8	20	5.3	1.95	10.00	106.57	200.00
4	200.00	10.0%	20.0	80.68	10.40	0.8	20	5.3	1.95	10.00	106.57	200.00
5	200.00	10.0%	20.0	80.68	10.40	0.8	20	5.3	1.95	10.00	106.57	200.00
total	1000.00		100.00	403.39		4.00			9.76	50.00	532.86	
6	200.00	10.0%	20.0	21.38	7.21	0.8	20	5.3	1.35	10.00	166.46	200.00
7	200.00	10.0%	20.0	21.38	7.21	0.8	20	5.3	1.35	10.00	166.46	200.00
total	400.00		40.00	42.77		1.60			22.22	0.00	0.00	

Fed-batch saccharification-shake flasks				Cellulase Activity:			Enzyme	Activity	Protein mg/n	FPU/mg protein			
Experiment #8, 4/20/04-4/30/04							Spezyme	28.40	106.60	0.27	Enhanced BCA protocol		
Corn stover PS030312CS				Glucan	Solids								
				53.20	0.409	corn stover							
Shake Flask	Initial Sample Size g	Final Sample Size g	Initial Insoluble Solids wt%	Initial Dry Corn Stover g	Initial Wet Corn Stover g	Initial Glucan g	Initial Cellulase Loading mg protein/g	Cellulase Loading FPU/g cellulo	Initial Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	Final equiv Insoluble Solids wt%	Temperature (deg C)
1	200.00	304.38	17.5%	35.0	85.51	18.62	20	5.3	3.49	15.22	95.78	25.0%	45
2	200.00	304.38	17.5%	35.0	85.51	18.62	20	5.3	3.49	15.22	95.78	25.0%	45
3	200.00	316.16	17.5%	35.0	85.51	18.62	40	10.7	6.99	15.81	91.69	25.0%	45
4	200.00	316.16	17.5%	35.0	85.51	18.62	40	10.7	6.99	15.81	91.69	25.0%	45
5	100.00	100.00	25.0%	25.0	61.08	13.30	20	5.3	2.50	5.00	31.42	25.0%	45
6	100.00	100.00	25.0%	25.0	61.08	13.30	20	5.3	2.50	5.00	31.42	25.0%	45
7	100.00	100.00	25.0%	25.0	61.08	13.30	40	10.7	4.99	5.00	28.93	25.0%	45
8	100.00	100.00	25.0%	25.0	61.08	13.30	40	10.7	4.99	5.00	28.93	25.0%	45
9	100.00	100.00	25.0%	25.0	61.08	13.30	20	5.3	2.50	5.00	31.42	25.0%	50
10	100.00	100.00	25.0%	25.0	61.08	13.30	20	5.3	2.50	5.00	31.42	25.0%	50
11	100.00	100.00	27.0%	27.0	65.97	14.36	20	5.3	2.69	5.00	26.34	27.0%	45
12	100.00	100.00	27.0%	27.0	65.97	14.36	20	5.3	2.69	5.00	26.34	27.0%	45
13	100.00	100.00	27.0%	27.0	65.97	14.36	40	10.7	5.39	5.00	23.64	27.0%	45
14	100.00	100.00	27.0%	27.0	65.97	14.36	40	10.7	5.39	5.00	23.64	27.0%	45
total	1800.00	1641.08		398.00	972.39			106.57	35.93	82.05	495.65		

Shake Flask	Feed #1 t = 6 hr total mass (g)	Feed #1 Wet Corn Stover (g)	Feed #1 Undiluted Cellulase (mL)	Feed #2 t = 11 hr total mass (g)	Feed #2 Wet Corn Stover (g)	Feed #2 Undiluted Cellulase (mL)	Feed #3 t = 22.5 hr total mass (g)	Feed #3 Wet Corn Stover (g)	Feed #3 Undiluted Cellulase (mL)	total PCS	total enzyme
1	0.00	0.00	0.00	26.68	25.63	1.05	0	0.00	0.00	25.63	1.05
2	0.00	0.00	0.00	26.68	25.63	1.05	0	0.00	0.00	25.63	1.05
3	28.64	26.48	2.16	12.48	11.54	0.94	22.30	20.62	1.68	58.63	4.79
4	28.64	26.48	2.16	12.48	11.54	0.94	22.30	20.62	1.68	58.63	4.79
Total	57.28	52.95	4.33	51.64	74.34	3.98	44.60	41.23	3.37	168.52	11.68

Shake Flask	Feed #4 t = 30.5 hr total mass (g)	Feed #4 Wet Corn Stover (g)	Feed #4 Undiluted Cellulase (mL)	Feed #5 t = 45.5 hr total mass (g)	Feed #5 Wet Corn Stover (g)	Feed #5 Undiluted Cellulase (mL)	Feed #6 t =72 hr total mass (g)	Feed #6 Wet Corn Stover (g)	Feed #6 Undiluted Cellulase (mL)	Feed #7 t =93.5 hr total mass (g)	Feed #7 Wet Corn Stover (g)	Feed #7 Undiluted Cellulase (mL)
1	0.00	0.00	0.00	38.06	36.57	1.49	18.5	17.77	0.73	21.14	20.31	0.83
2	0.00	0.00	0.00	38.06	36.57	1.49	18.5	17.77	0.73	21.14	20.31	0.83
3	0.00	0.00	0.00	32.64	30.17	2.47	20.10	18.58	1.52	0.00	0.00	0.00
4	0.00	0.00	0.00	32.64	30.17	2.47	20.10	18.58	1.52	0.00	0.00	0.00
Total	0.00	0.00	0.00	141.40	133.48	7.92	77.20	72.71	4.49	42.28	40.62	1.66

Batch saccharification-shake flasks — **Cellulase Activity:** — **Enzyme** Spezyme — **Activity** 28.40 — **Protein mg/n** 106.60 — **FPU/mg protein** 0.27 — Enhanced BCA protocol

Experiment #9

Corn stover PS030312CS — Glucan **53.20** — Solids **0.401** — corn stover

Shake Flask	Initial Sample Size g	Initial Insoluble Solids wt%	Initial Dry Corn Stover g	Initial Wet Corn Stover g	Initial Glucan g	Initial Cellulase Loading mg protein/g c	Background glucose (g/kg)	Cellulase Loading FPU/g cellulo	Initial Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	DI Water (mL)	dens = 1.07 glucose (250 g/L) (mL)
1	100.00	25.0%	25.0	62.33	13.30	60	0	16.0	7.49	5.00	25.19	0.00
2	100.00	25.0%	25.0	62.33	13.30	60	0	16.0	7.49	5.00	25.19	0.00
3	100.00	27.0%	27.0	67.31	14.36	60	0	16.0	8.08	5.00	19.60	0.00
4	100.00	27.0%	27.0	67.31	14.36	60	0	16.0	8.08	5.00	19.60	0.00
5	100.00	30.0%	30.0	74.79	15.96	20	0	5.3	2.99	5.00	17.21	0.00
6	100.00	30.0%	30.0	74.79	15.96	20	0	5.3	2.99	5.00	17.21	0.00
7	100.00	30.0%	30.0	74.79	15.96	40	0	10.7	5.99	5.00	14.22	0.00
8	100.00	30.0%	30.0	74.79	15.96	40	0	10.7	5.99	5.00	14.22	0.00
9	100.00	32.0%	32.0	79.78	17.02	20	0	5.3	3.19	5.00	12.03	0.00
10	100.00	32.0%	32.0	79.78	17.02	20	0	5.3	3.19	5.00	12.03	0.00
11	100.00	32.0%	32.0	79.78	17.02	40	0	10.7	6.39	5.00	8.83	0.00
12	100.00	32.0%	32.0	79.78	17.02	40	0	10.7	6.39	5.00	8.83	0.00
13	100.00	10.0%	10.0	24.93	5.32	40	0	10.7	2.00	5.00	68.07	0.00
14	100.00	10.0%	10.0	24.93	5.32	40	0	10.7	2.00	5.00	68.07	0.00
15	100.00	10.0%	10.0	24.93	5.32	20	150	5.3	1.00	5.00	4.65	64.42
16	100.00	10.0%	10.0	24.93	5.32	20	150	5.3	1.00	5.00	4.65	64.42
total	1300.00		362.00	902.52				138.54	70.27	65.00	262.21	

Batch SSF-shake flasks — **Cellulase Activity:** — **Enzyme** Spezyme — **Activity** 28.40 — **Protein mg/n** 106.60 — **FPU/mg protein** 0.27 — Enhanced BCA protocol

Experiment #10

Corn stover PS030312CS — Glucan **53.20** — Solids **0.392** Dried corn stover, **0.28** Undried corn stover

Shake Flask	Sample Size g	Insoluble Solids wt%	Dry Corn Stover g	Wet Corn Stover g	Initial Glucan g	Initial Cellulase Loading mg protein/g c	Cellulase Loading FPU/g cellulo	Initial Undiluted Cellulase (mL)	1 M Citrate Buffer (mL)	10 X YP stock (mL)	inoculum (mL)	DI Water (mL)	Reaction Type	substrate drying
1	100.00	25.0%	25.0	63.76	13.30	20	5.3	2.50	5.00	10.00	10.00	8.75	SSF	Dried
2	100.00	25.0%	25.0	63.76	13.30	20	5.3	2.50	5.00	10.00	10.00	8.75	SSF	Dried
3	100.00	25.0%	25.0	63.76	13.30	40	10.7	4.99	5.00	10.00	10.00	6.25	SSF	Dried
4	100.00	25.0%	25.0	63.76	13.30	40	10.7	4.99	5.00	10.00	10.00	6.25	SSF	Dried
5	100.00	20.0%	20.0	51.01	10.64	20	5.3	2.00	5.00	10.00	10.00	22.00	SSF	Dried
6	100.00	20.0%	20.0	51.01	10.64	20	5.3	2.00	5.00	10.00	10.00	22.00	SSF	Dried
7	100.00	10.0%	10.0	25.50	5.32	20	5.3	1.00	5.00	10.00	10.00	48.50	SSF	Dried
8	100.00	10.0%	10.0	25.50	5.32	20	5.3	1.00	5.00	10.00	10.00	48.50	SSF	Dried
9	100.00	10.0%	10.0	25.50	5.32	20	5.3	1.00	5.00	10.00	10.00	48.50	Hydrolysis	Dried
10	100.00	10.0%	10.0	25.50	5.32	20	5.3	1.00	5.00	10.00	10.00	48.50	Hydrolysis	Dried
11	100.00	10.0%	10.0	36.30	5.32	20	5.3	1.00	5.00	10.00	10.00	37.70	Hydrolysis	Undried
12	100.00	10.0%	10.0	36.30	5.32	20	5.3	1.00	5.00	10.00	10.00	37.70	Hydrolysis	Undried
13	100.00	25.0%	25.0	63.76	13.30	20	5.3	2.50	5.00	10.00	10.00	8.75	Hydrolysis	Dried
14	100.00	25.0%	25.0	63.76	13.30	20	5.3	2.50	5.00	10.00	10.00	8.75	Hydrolysis	Dried
15	100.00	25.0%	25.0	63.76	13.30	40	10.7	4.99	5.00	10.00	10.00	6.25	Hydrolysis	Dried
16	100.00	25.0%	25.0	63.76	13.30	40	10.7	4.99	5.00	10.00	10.00	6.25	Hydrolysis	Dried
17	100.00	0%	0.0	0.0	5.0	0.0	0.0	0.0	5.00	10.00	10.00	75.00	Fermentation	
18	100.00	0%	0.0	0.0	5.0	0.0	0.0	0.0	5.00	10.00	10.00	75.00	Fermentation	
19	100.00	0%	0.0	0.0	13.0	0.0	0.0	0.0	5.00	10.00	10.00	75.00	Fermentation	
20	100.00	0%	0.0	0.0	13.0	0.0	0.0	0.0	5.00	10.00	10.00	75.00	Fermentation	
total	1600.00		300.00	531.66			74.60	24.95	80.00	120.00	120.00	343.39		

Fed-batch saccharification-Biofio 3000 | **Cellulase Activity:** | Enzyme | Activity | Protein mg/n | FPU/mg protein |

Experiment #11, 8/19/04-8/29/04 | Spezyme | 28.40 | 106.60 | 0.27 | Enhanced BCA protocol

Corn stover PS030312CS

				Glucan	Solids								
				53.20	**0.459**	com stover							
	Initial Sample Size	Final Sample Size	Initial Insoluble Solids	Initial Dry Corn Stover	Initial Wet Corn Stover	Initial Glucan	Initial Cellulase Loading	Cellulase Loading	Initial Undiluted Cellulase	1 M Citrate Buffer	DI Water	Final equiv Insoluble Solids	Temperature
Reactor	g	g	wt%	g	g	g	mg protein/g c	FPU/g cellulo	(mL)	(mL)	(mL)	wt%	(deg C)
BF 1	2555.00	4500.00	12.0%	306.6	667.39	163.11	40	10.7	61.20	225.00	1601.40	25.0%	45
BF 2	2838.00	4500.00	15.0%	425.7	926.64	226.47	40	10.7	84.98	225.00	1601.38	25.0%	45
SF 1	100.00	100.00	25.0%	25.0	54.42	13.30	40	10.7	4.99	5.00	35.59	25.0%	45
SF 2	100.00	100.00	25.0%	25.0	54.42	13.30	40	10.7	4.99	5.00	35.59	25.0%	45
SF 3	100.00	100.00	25.0%	25.0	54.42	13.30	40	10.7	4.99	5.00	35.59	25.0%	45
total	5693.00	9300.00		807.30	1757.29				161.16	465.00	3309.55		

	Feed #1 t = 14 hr total mass	Feed #1 Wet Corn Stover	Feed #1 Undiluted Cellulase	Feed #2 t = 24 hr total mass	Feed #2 Wet Corn Stover	Feed #2 Undiluted Cellulase	Feed #3 t = 48 hr total mass	Feed #3 Wet Corn Stover	Feed #3 Undiluted Cellulase	Feed #4 t = 72 hr total mass	Feed #4 Wet Corn Stover	Feed #4 Undiluted Cellulase		
Reactor	(g)	(g)	(mL)	(g)	(g)	(mL)	(g)	(g)	(mL)	(g)	(g)	(mL)	total PCS	total enzyme
1	327.82	300.26	27.54	183.05	167.68	15.36	305.076084	279.45	25.63	203.99	186.86	17.14	934.27	85.68
2	413.50	378.77	34.74	140.76	128.95	11.83	351.349647	321.83	29.51	307.36	281.54	25.82	1111.10	101.90
Total	741.32	679.05	62.27	140.78	296.63	27.20	351.35	601.28	55.14	511.35	468.40	42.96		

	Feed #5 t = 96 hr total mass	Feed #5 Wet Corn Stover	Feed #5 Undiluted Cellulase	Feed #6 t =120 hr total mass	Feed #6 Wet Corn Stover	Feed #6 Undiluted Cellulase	Feed #7 t =144 hr total mass	Feed #7 Wet Corn Stover	Feed #7 Undiluted Cellulase	Feed #8 t =168 hr total mass	Feed #8 Wet Corn Stover	Feed #8 Undiluted Cellulase		
Reactor	(g)	(g)	(mL)	(g)	(g)	(mL)	(g)	(g)	(mL)	(g)	(g)	(mL)	total PCS	total enzyme
1	204.92	187.70	17.21	174.256189	159.62	14.64	151.26	138.55	12.71	132.86	121.70	11.16	607.58	55.72
2	207.18	189.77	17.40	242.085015	221.75	20.34		0.00	0.00		0.00	0.00	411.52	37.74
Total	412.09	377.48	34.62	416.34	381.37	34.97	151.26	138.55	12.71	132.86	121.70	11.16		

	Feed #9 t = 192 hr total mass	Feed #9 Wet Corn Stover	Feed #9 Undiluted Cellulase	Feed #10 t = 216 hr total mass	Feed #10 Wet Corn Stover	Feed #10 Undiluted Cellulase		
Reactor	(g)	(g)	(mL)	(g)	(g)	(mL)	total PCS	total enzyme
1	118.81	108.83	9.98	142.32	130.36	11.96	239.19	21.94
2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	118.81	108.83	9.98	142.32	130.36	11.96		

Fed-batch saccharification-HSBR				Cellualse Activity:		Enzyme	Activity	Protein mg/n	FPU/mg protein	
Experiment #12, 9/14/04-9/24/04						Spezyme	28.40	106.60	0.27	Enhanced BCA protocol
Corn stover PS030312CS			Glucan	Solids						
			53.20	0.343	corn stover					
	Initial	Initial	Initial	Initial		Initial		Initial		
	Sample	Insoluble	Dry	Wet	Initial	Cellulase	Cellulase	Undiluted	1 M Citrate	
Reactor	Size	Solids	Corn Stover	Corn Stover	Glucan	Loading	Loading	Cellulase	Buffer	DI Water
type	g	wt%	g	g	g	mg protein/g (	FPU/g cellulo	(mL)	(mL)	(mL)
Reactor	2200.00	25.0%	550.0	1601.63	292.60	40	10.7	109.79	110.00	378.58
SF 1	100.00	25.0%	25.0	72.80	13.30	40	10.7	4.99	5.00	17.21
SF 2	100.00	25.0%	25.0	72.80	13.30	40	10.7	4.99	5.00	17.21
SF 3	100.00	25.0%	25.0	72.80	13.30	40	10.7	4.99	5.00	17.21
total	2500.00		625.00	1820.03				124.77	125.00	430.20

Fed-batch saccharification-HSBR				Cellualse Activity:		Enzyme	Activity	Protein mg/n	FPU/mg protein	
Experiment #13, 9/30/04-10/8/04						Spezyme	28.40	106.60	0.27	Enhanced BCA protocol
Corn stover PS030312CS			Glucan	Solids						
			53.20	0.336	corn stover					
	Initial	Initial	Initial	Initial		Initial		Initial		
	Sample	Insoluble	Dry	Wet	Initial	Cellulase	Cellulase	Undiluted	1 M Citrate	
Reactor	Size	Solids	Corn Stover	Corn Stover	Glucan	Loading	Loading	Cellulase	Buffer	DI Water
type	g	wt%	g	g	g	mg protein/g (	FPU/g cellulo	(mL)	(mL)	(mL)
Reactor	2500.00	25.0%	625.0	1859.01	332.50	40	10.7	124.77	125.00	391.22
SF 1	100.00	25.0%	25.0	74.36	13.30	40	10.7	4.99	5.00	15.65
SF 2	100.00	25.0%	25.0	74.36	13.30	40	10.7	4.99	5.00	15.65
SF 3	100.00	20.0%	20.0	59.49	10.64	40	10.7	3.99	5.00	31.52
SF 4	100.00	20.0%	20.0	59.49	10.64	40	10.7	3.99	5.00	31.52
SF 5	100.00	17.5%	17.5	52.05	9.31	40	10.7	3.49	5.00	39.45
SF 6	100.00	17.5%	17.5	52.05	9.31	40	10.7	3.49	5.00	39.45
SF 7	100.00	15.0%	15.0	44.62	7.98	40	10.7	2.99	5.00	47.39
SF 8	100.00	15.0%	15.0	44.62	7.98	40	10.7	2.99	5.00	47.39
total	3300.00			2320.05				155.71	165.00	659.25

Fed-batch saccharification-HSBR				Cellualse Activity:		Enzyme	Activity	Protein mg/n FPU/mg protein		
Experiment #14, 10/13/04-10/25/04						Spezyme	28.40	106.60	0.27	Enhanced BCA protocol
Corn stover PS030312CS			Glucan	Solids						
			53.20	0.314	corn stover					
	Initial	Initial	Initial	Initial		Initial		Initial		
	Sample	Insoluble	Dry	Wet	Initial	Cellulase	Cellulase	Undiluted	1 M Citrate	
Reactor	Size	Solids	Corn Stover	Corn Stover	Glucan	Loading	Loading	Cellulase	Buffer	DI Water
type	g	wt%	g	g	g	mg protein/g ı	FPU/g cellulo	(mL)	(mL)	(mL)
Reactor	2500.00	25.0%	625.0	1990.45	332.50	20	5.3	62.38	125.00	322.17
SF 1	100.00	25.0%	25.0	79.62	13.30	20	5.3	2.50	5.00	12.89
SF 2	100.00	25.0%	25.0	79.62	13.30	20	5.3	2.50	5.00	12.89
SF 3	100.00	25.0%	25.0	79.62	13.30	20	5.3	2.50	5.00	12.89
total	2800.00			2229.30				69.87	140.00	360.83

Fed-batch saccharification-HSBR				Cellualse Activity:		Enzyme	Activity	Protein mg/n FPU/mg protein		
Experiment #15, 1/17/05-1/27/05						Spezyme	28.40	106.60	0.27	Enhanced BCA protocol
Corn stover PS030312CS			Glucan	Solids						
			53.20	0.378	corn stover					
	Initial	Initial	Initial	Initial		Initial		Initial		
	Sample	Insoluble	Dry	Wet	Initial	Cellulase	Cellulase	Undiluted	1 M Citrate	
Reactor	Size	Solids	Corn Stover	Corn Stover	Glucan	Loading	Loading	Cellulase	Buffer	DI Water
type	g	wt%	g	g	g	mg protein/g ı	FPU/g cellulo	(mL)	(mL)	(mL)
Reactor	2500.00	27.0%	675.0	1786.66	359.10	40	10.7	134.75	125.00	453.59
SF 1	100.00	27.0%	27.0	71.47	14.36	40	10.7	5.39	5.00	18.14
SF 2	100.00	27.0%	27.0	71.47	14.36	40	10.7	5.39	5.00	18.14
SF 3	100.00	27.0%	27.0	71.47	14.36	40	10.7	5.39	5.00	18.14
total	2800.00			2001.06				150.92	140.00	508.02

B. Unwashed Solids Experiments

3/2/2004

Exp 3 Saccharification Work sheet:

							Cellulase Activity:		Enzyme	Activity	Protein mg/mL	FPU/mg protein			
									Spezyme	28.40	106.60	0.27			
Corn stover P030312CS		Glucan	52.00	Solids	0.420		total								
					0.34		corn stover		Sugar solution, %		0.0		0.039		
		Total											Sugars in hydrolyzate		
	Sample	Insoluble	Dry	Wet			Soluble	Cellulase		Sugar	Total	Total			
Flask/	Size	Solids	Corn Stover	Corn Stover	Glucan		Solids	Loading	Hydrolyzate	Solution	Solids	Carbohydrates	Flask/	Hydrolyzate	Hydrolyzate
Fermentor	(g)	(wt%)	(g)	(g)	(g)		(g)	mg protein/g	(mL)	(mL)	(g)	(g)	Fermentor	strength	conditioning
Flask 1	100	5.0	5.0	14.71	2.60		1.2	20	0.0	0.0	6.2	5.3	Flask 1	0	None
Flask 2	100	5.0	5.0	14.71	2.60		1.2	20	0.0	0.0	6.2	5.3	Flask 2	0	None
Flask 3	100	5.0	5.0	14.71	2.60		1.2	20	0.0	0.0	6.2	5.3	Flask 3	0	None
Fermentor 1	3500	5.0	175.0	514.71	91.00		41.2	20	0.0	0.0	216.2	186.7	Fermentor 1	0	None
Fermentor 2	3500	5.0	175.0	514.71	91.00		41.2	20	0.0	0.0	216.2	186.7	Fermentor 2	0	None
					1073.5										

			2.0	1M							Actual			
			1:2 Dilute	Citrate			Actual	Error	Weight		Wet	Actual		Total
Flask/		Cellulase	Cellulase	Buffer	Water	Final	Water	in Water	After	Water	Corn Stover	Glucan	Flask/	charge
Fermentor		(mL)	(mL)	(mL)	(mL)	Hydrolysate mL	Added (g)	Volume (+/-)(g)	Autoclave (g)	Correction (g)	Added (g)	Added (g)	Fermentor	g
Flask 1		0.49	0.98	0.0	84.8	-84.81		-84.81		0.00		2.60	Flask 1	100.5
Flask 2		0.49	0.98	0.0	84.8	-84.81		-84.81		0.00		2.60	Flask 2	100.5
Flask 3		0.49	0.98	0.0	84.8	-84.81		-84.81		0.00		2.60	Flask 3	100.5
250	Fermentor 1	17.07	34.15	0.0	2968.2	-2968.22		-2968.22		0.00		91.00	Fermentor 1	3517.1
100	Fermentor 2	17.07	34.15	0.0	2968.2	-2968.22		-2968.22		0.00		91.00	Fermentor 2	3517.1

3/10/2004

Exp 4 Saccharification Work sheet:

							Cellulase Activity:		Enzyme	Activity	Protein mg/mL	FPU/mg protein			
									Spezyme	28.40	106.60	0.27			
Corn stover P030312CS		Glucan	52.00	Solids	0.420		total								
					0.34		corn stover		Sugar solution, %		0.0		0.039		
		Total											Sugars in hydrolyzate		
	Sample	Insoluble	Dry	Wet			Soluble	Cellulase		Sugar	Total	Total			
Flask/	Size	Solids	Corn Stover	Corn Stover	Glucan		Solids	Loading	Hydrolyzate	Solution	Solids	Carbohydrates	Flask/	Hydrolyzate	
Fermentor	(g)	(wt%)	(g)	(g)	(g)		(g)	mg protein/g	(mL)	(mL)	(g)	(g)	Fermentor	conditioning	
Flask 1	100	10.0	10.0	29.41	5.20		2.4	20	0.0	0.0	12.4	10.7	Flask 1	None	
Flask 2	100	10.0	10.0	29.41	5.20		2.4	20	0.0	0.0	12.4	10.7	Flask 2	None	
Flask 3	100	10.0	10.0	29.41	5.20		2.4	20	0.0	0.0	12.4	10.7	Flask 3	None	
Fermentor 1	3500	10.0	350.0	1029.41	182.00		82.4	20	0.0	0.0	432.4	373.3	Fermentor 1	None	
Fermentor 2	3500	10.0	350.0	1029.41	182.00		82.4	20	0.0	0.0	432.4	373.3	Fermentor 2	None	
		10.00			2147.1										

		2.0	1M							Actual			
		pH adjustment	Citrate			Actual	Error	Weight		Wet	Actual		Total
Flask/	Cellulase	NaOH (4N)	Buffer	Water	Final Hydrolysate	Water	in Water	After	Water	Corn Stover	Glucan	Flask/	charge
Fermentor	(mL)	(mL)	(mL)	(mL)	mL	Added (g)	Volume (+/-)(g)	Autoclave (g)	Correction (g)	Added (g)	Added (g)	Fermentor	g
Flask 1	0.98	1.95	5.0	62.7	-62.66		-62.66		0.00		5.20	Flask 1	100.0
Flask 2	0.98	1.95	5.0	62.7	-62.66		-62.66		0.00		5.20	Flask 2	100.0
Flask 3	0.98	1.95	5.0	62.7	-62.66		-62.66		0.00		5.20	Flask 3	100.0
Fermentor 1	34.15	68.29	175.0	2193.1	-2193.15		-2193.15		0.00		182.00	Fermentor 1	3500.0
Fermentor 2	34.15	68.29	175.0	2193.1	-2193.15		-2193.15		0.00		182.00	Fermentor 2	3500.0
	71.2	142.4	365.0	4574.3									

3/22/2004

Exp 5 Saccharification Work sheet:

Cellulase Activity:

Enzyme	Activity	Protein mg/n	FPU/mg prot
Spezyme	28.40	106.60	0.27

Corn stover P030312CS Glucan **52.00** Solids **0.429** total **0.34** corn stover

Flask/ Fermentor	Sample Size (g)	Total Insoluble Solids (wt%)	Total Soluble Solids (g/kg)	Dry Corn Stover (g)	Wet Corn Stover (g)	Glucan (g)	Soluble Solids (g)	Cellulase Loading mg protein/g
Flask 1	100	13.5	37.2	13.5	**40.17**	7.02	3.7	20
Flask 2	100	13.5	37.2	13.5	**40.17**	7.02	3.7	20
Flask 3	100	13.5	37.2	13.5	**40.17**	7.02	3.7	20
Fermentor 1	3500	13.5	37.2	472.5	**1405.83**	245.70	130.3	20
Fermentor 2	3500	13.5	37.2	472.5	**1405.83**	245.70	130.3	20
					2932.2			

4N NaOH mL/ 100 g at 13.5%:

Flask/ Fermentor	tetracycline mL	Cellulase (mL)	1.75 pH adjustment NaOH (4N) (mL)	1M Citrate Buffer (mL)	Water (mL)	Total charge g
Flask 1	**0.4**	**1.32**	**1.75**	**5.0**	**51.4**	100.0
Flask 2	**0.4**	**1.32**	**1.75**	**5.0**	**51.4**	100.0
Flask 3	**0.4**	**1.32**	**1.75**	**5.0**	**51.4**	100.0
Fermentor 1	**14**	**46.10**	**61.25**	**175.0**	**1797.8**	3500.0
Fermentor 2	**14**	**46.10**	**61.25**	**175.0**	**1797.8**	3500.0
	29.2	96.1	127.8	365.0	3749.7	

4/6/2004

Exp 6 Saccharification Work sheet:

Cellulase Activity:

Enzyme	Activity	Protein mg/m	FPU/mg protein
Spezyme	28.40	106.60	0.27

Corn stover P030312CS Glucan **52.00** Solids **0.439** total **0.34** corn stover

Flask/ Fermentor	Sample Size (g)	Total Insoluble Solids (wt%)	Total Soluble Solids (g/kg)	Dry Corn Stover (g)	Wet Corn Stover (g)	Glucan (g)	Soluble Solids (g)	Cellulase Loading mg protein/g
Flask 1	100	15.0	42.5	15.0	**43.83**	7.80	4.3	20
Flask 2	100	15.0	42.5	15.0	**43.83**	7.80	4.3	20
Flask 3	100	15.0	42.5	15.0	**43.83**	7.80	4.3	20
Fermentor 1	3818.6	15.0	42.5	572.8	**1673.61**	297.85	162.4	20
Fermentor 2	3818.6	15.0	42.5	572.8	**1673.61**	297.85	162.4	20
					3478.7			

4N NaOH mL/ 100 g at 13.5%:

Flask/ Fermentor	tetracycline mL	Cellulase (mL)	1.75 pH adjustment NaOH (4N) (mL)	1M Citrate Buffer (mL)	Water (mL)	Glucan after 192 h (g)	Cellulase Loading at 200 h (mg/g)	Cellulase at 200 h (mL)
Flask 1	**0.4**	**1.46**	**1.75**	**5.0**	**47.6**	1.87	20	**0.35**
Flask 2	**0.4**	**1.46**	**1.75**	**5.0**	**47.6**	1.92	20	**0.36**
Flask 3	**0.4**	**1.46**	**1.75**	**5.0**	**47.6**			
Fermentor 1	**15.27**	**55.88**	**66.83**	**190.9**	**1816.1**	73.11	20	**13.72**
Fermentor 2	**15.27**	**55.88**	**66.83**	**190.9**	**1816.1**	80.94	20	**15.19**
	31.7	116.2	138.9	396.9	3774.9			29.6

4/26/2004 **Cellulase Activity:** **Enzyme** **Activity** **Protein mg/n** **FPU/mg prot**

Exp #7 Saccharification Work sheet: Spezyme 28.40 106.60 0.27

Corn stover P030312CS Glucan **52.00** Solids **40.90%** total
32.91% corn stover

Flask/ Fermentor	Sample Size (g)	Total Insoluble Solids (wt%)	Dry Corn Stover (g)	Wet Corn Stover (g)	Glucan (g)	Soluble Solids (g)	Cellulase Loading mg protein/g
Flask 1	100	10.0	10.0	**30.39**	5.20	2.4	20
Flask 2	100	10.0	10.0	**30.39**	5.20	2.4	20
Flask 3	100	10.0	10.0	**30.39**	5.20	2.4	20
Fermentor 1	13500	10.0	1350.0	**4102.10**	702.00	327.8	20
Fermentor 2	3500	10.0	350.0	**1063.51**	182.00	85.0	20
				5256.8			

Flask/ Fermentor	mL 4.0 N NaOH/100 mL = tetracycline mL	Cellulase (mL)	1.40 pH adjustmen NaOH (4N) (mL)	1M Citrate Buffer (mL)	Water (mL)	Total charge g
Flask 1	**0.4**	**0.98**	**1.40**	**5.0**	**61.8**	100.0
Flask 2	**0.4**	**0.98**	**1.40**	**5.0**	**61.8**	100.0
Flask 3	**0.4**	**0.98**	**1.40**	**5.0**	**61.8**	100.0
Fermentor 1	**54**	**131.71**	**189.00**	**675.0**	**8348.2**	13500.0
Fermentor 2	**14**	**34.15**	**49.00**	**175.0**	**2164.3**	3500.0
	69.2	168.8	242.2	865.0	10698.1	

5/19/2004 **Cellulase Activity:**

Enzyme	Activity	Protein mg/n	FPU/mg prot
Spezyme	28.40	106.60	0.27

Exp #8 Saccharification Worksheet:

Corn stover P030312CS Glucan **53.20** Solids **40.90%** total
32.69% insoluble

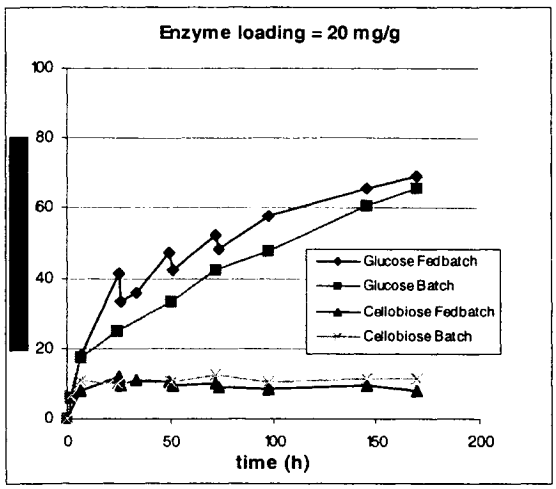
Flask/ Fermentor	Sample Size (g)	Total Insoluble Solids (wt%)	Dry Corn Stover (g)	Wet Corn Stover (g)	Glucan (g)	Soluble Solids (g)	Cellulase Loading mg protein/g
Flask 1	100	15.0	15.0	**45.89**	7.98	3.8	20
Flask 2	100	15.0	15.0	**45.89**	7.98	3.8	20
Flask 3	100	15.0	15.0	**45.89**	7.98	3.8	20
Fermentor 1	3750	15.0	562.5	**1720.71**	299.25	141.3	20
Fermentor 2	3750	15.0	562.5	**1720.71**	299.25	141.3	20
				3579.1			

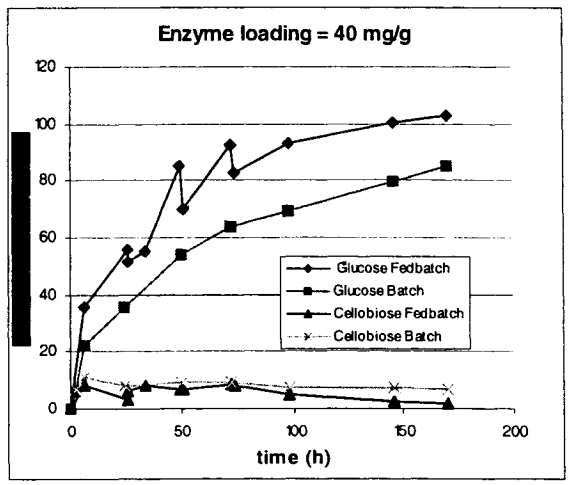
Flask/ Fermentor	mL 4.0 N NaOH/100 mL = tetracycline mL	Cellulase (mL)	2.00 pH adjustmen NaOH (4N) (mL)	1M Citrate Buffer (mL)	Water (mL)	Total charge g
Flask 1	**0.4**	**1.50**	**2.00**	**5.0**	**45.2**	100.0
Flask 2	**0.4**	**1.50**	**2.00**	**5.0**	**45.2**	100.0
Flask 3	**0.4**	**1.50**	**2.00**	**5.0**	**45.2**	100.0
Fermentor 1	**15**	**56.14**	**75.00**	**187.5**	**1695.6**	3750.0
Fermentor 2	**15**	**56.14**	**75.00**	**187.5**	**1695.6**	3750.0
	31.2	116.8	156.0	390.0	3526.9	

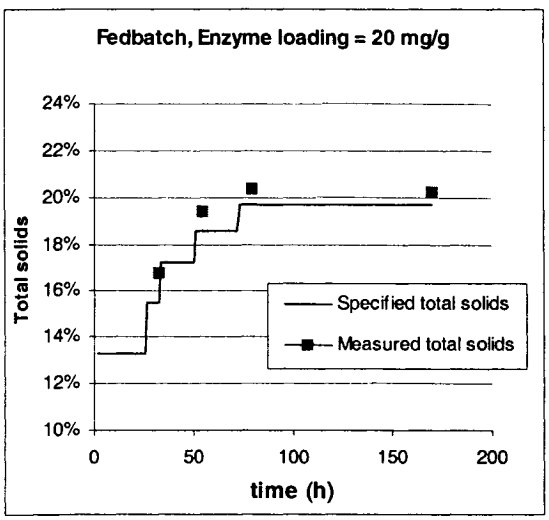
APPENDIX 4: SACCHARIFICATION EXPERIMENTAL DATA

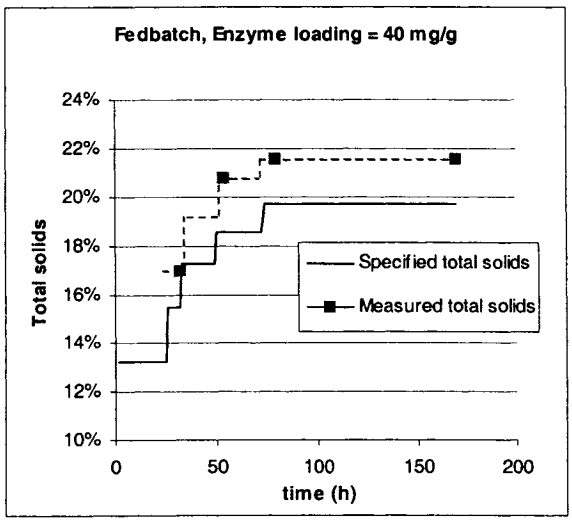
A. Washed Solids Experiments

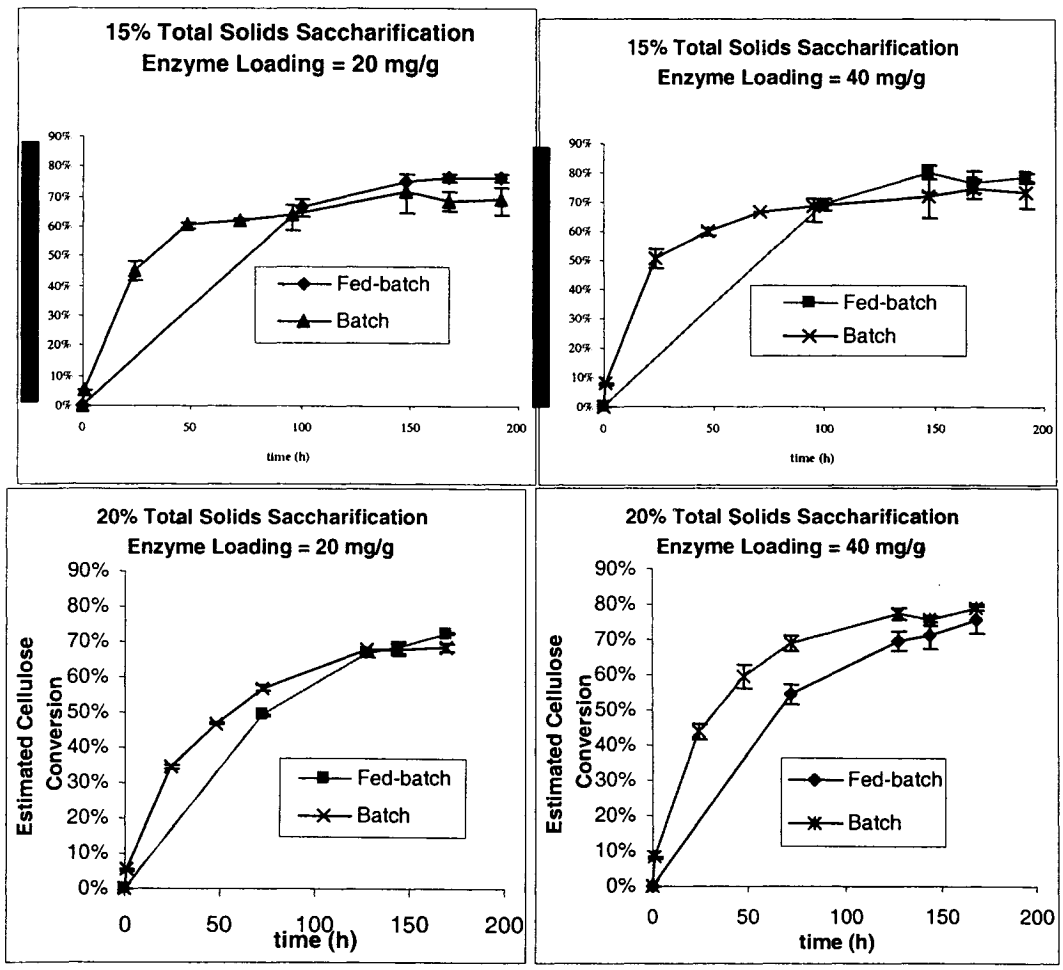
Experiment #1

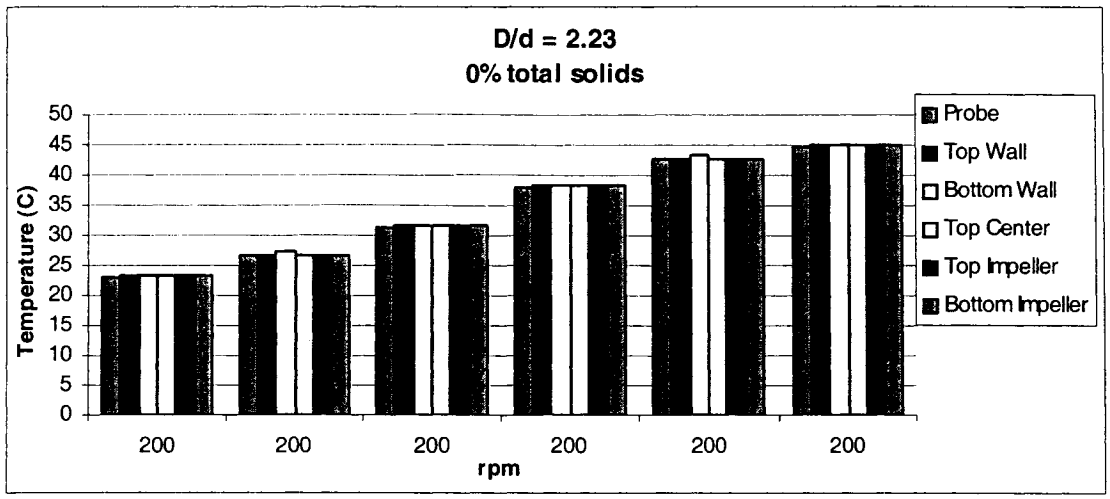


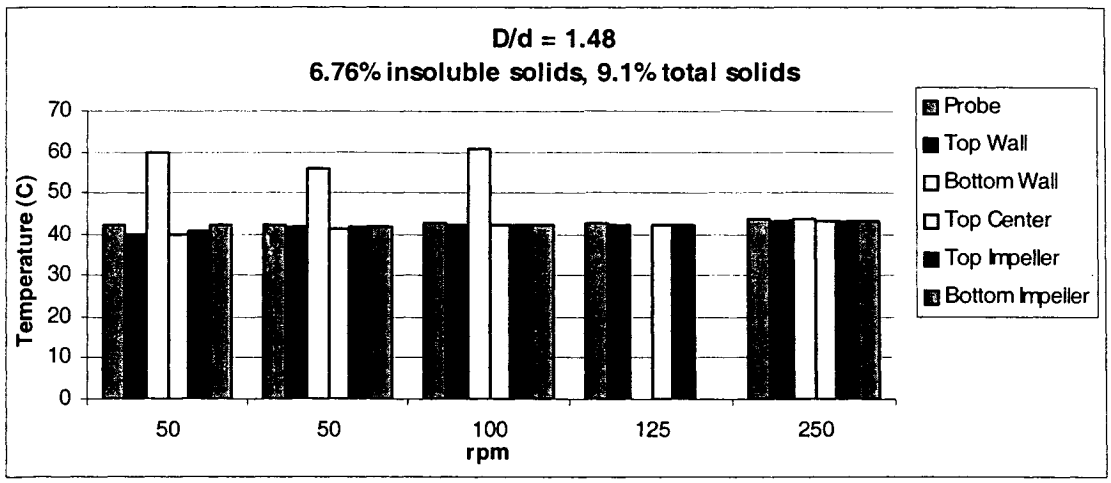


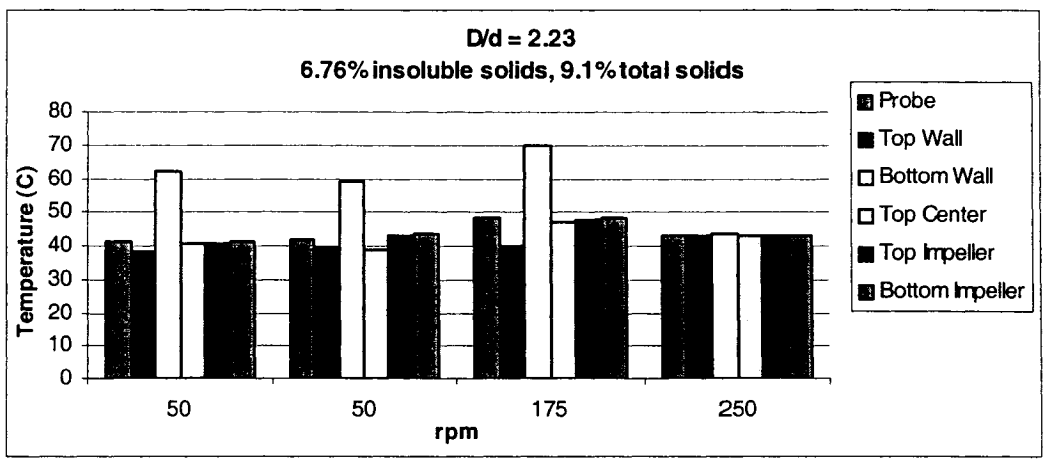


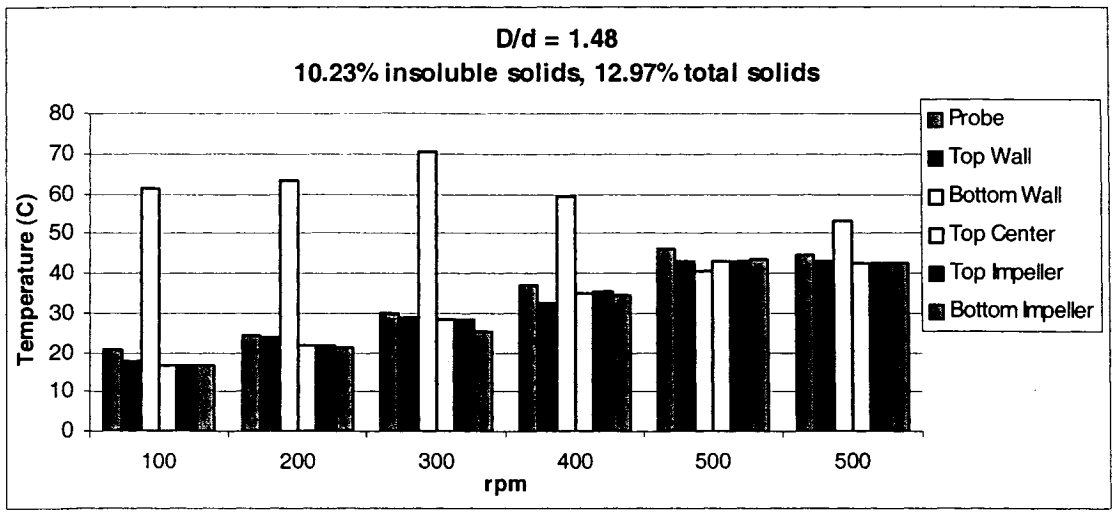


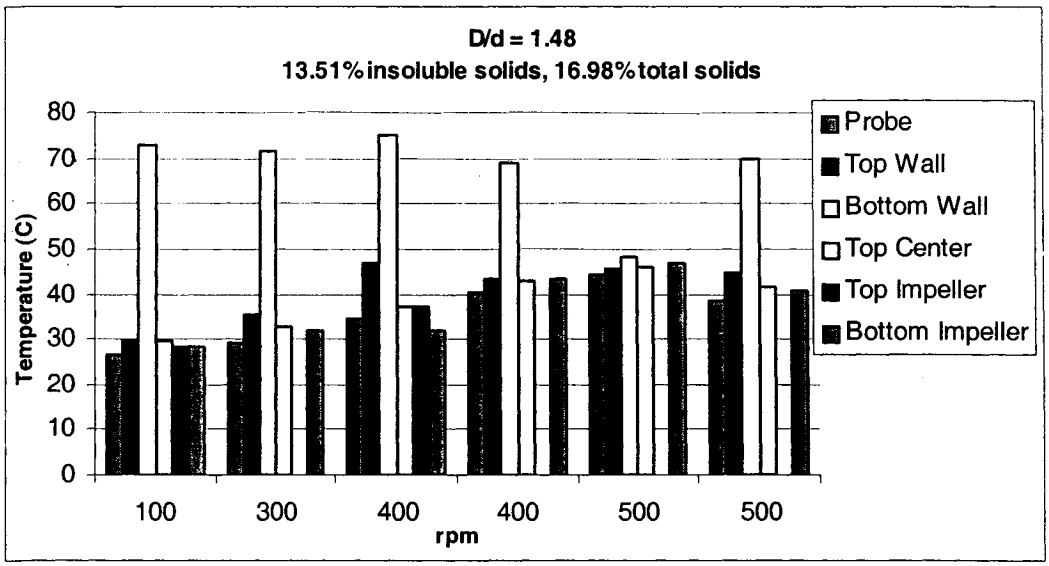


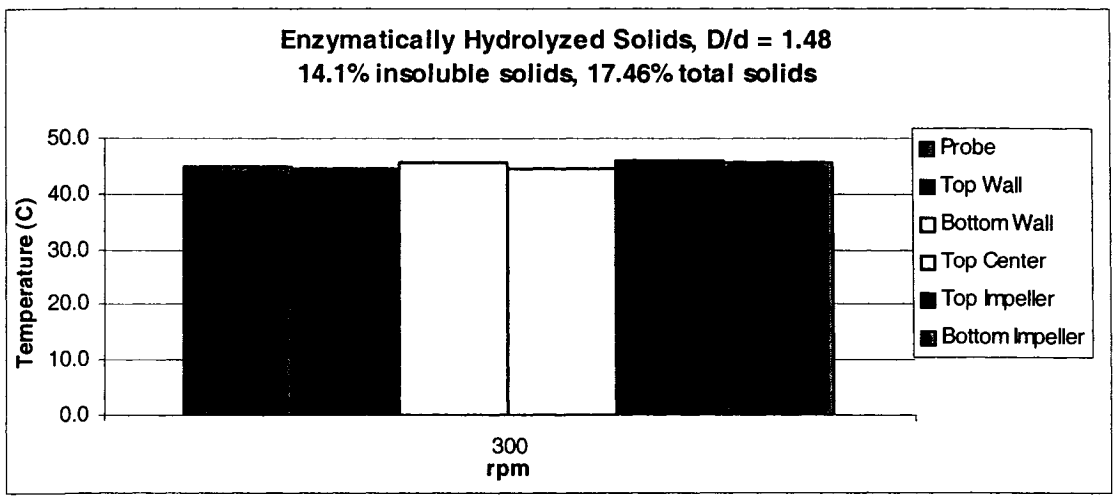


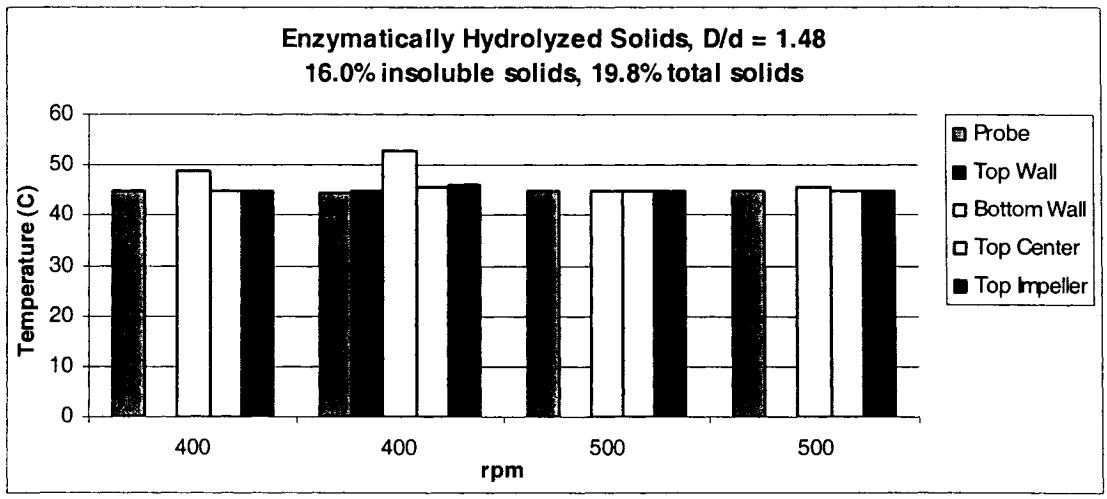


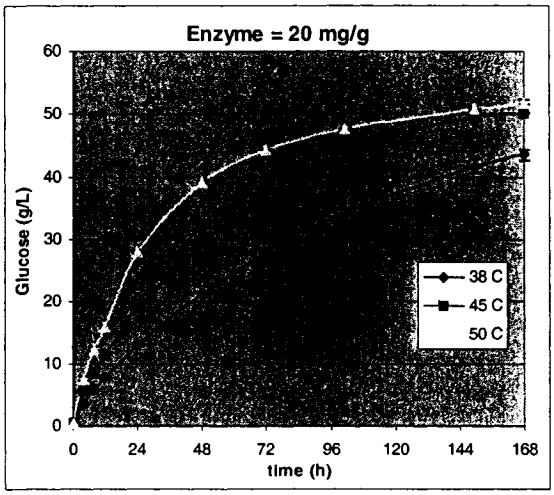


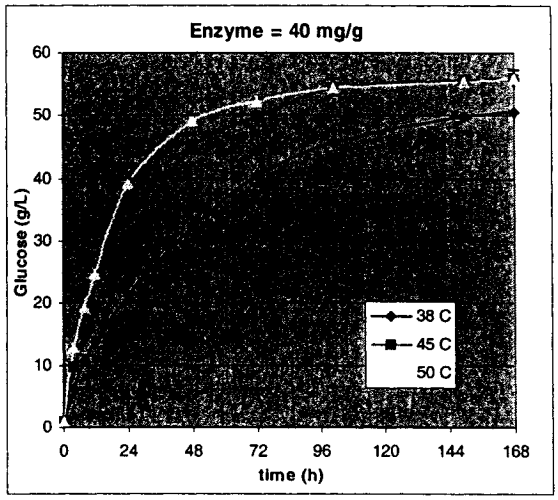


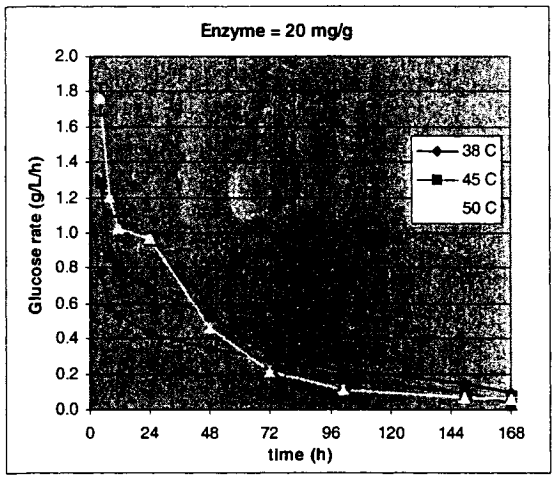


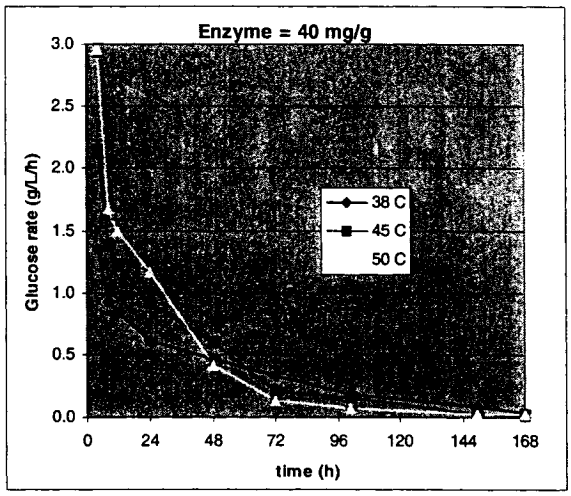


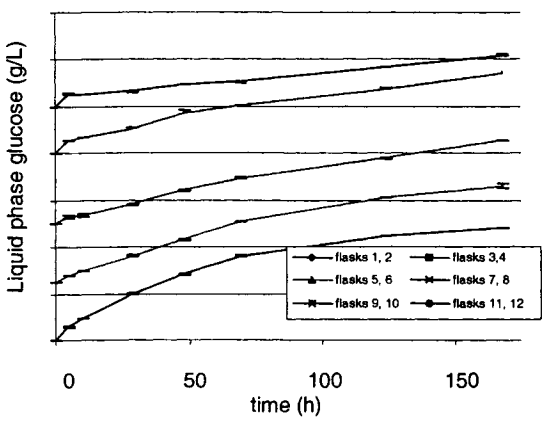


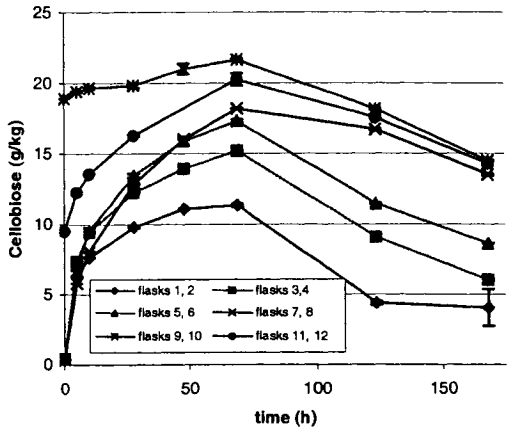


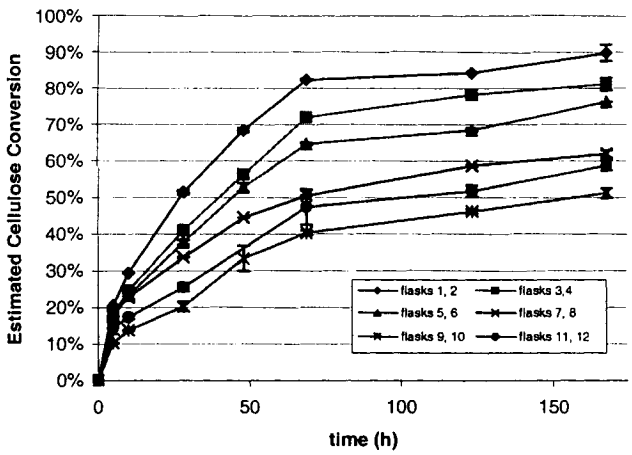


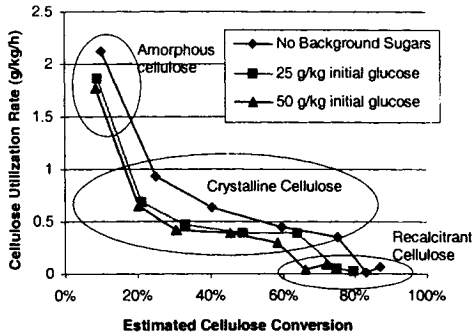




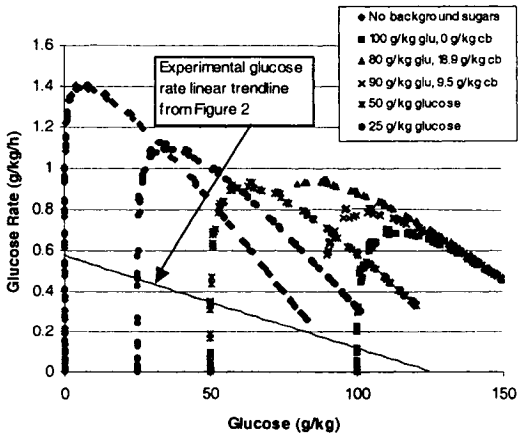




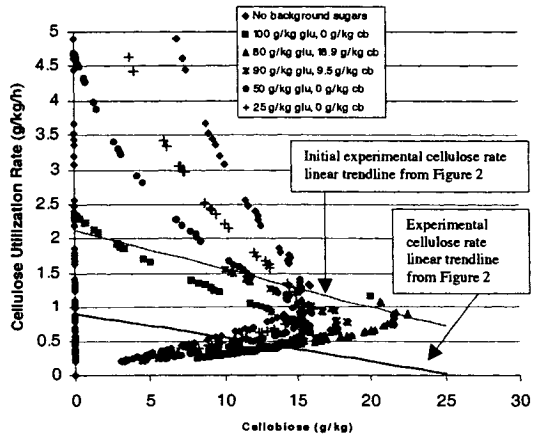




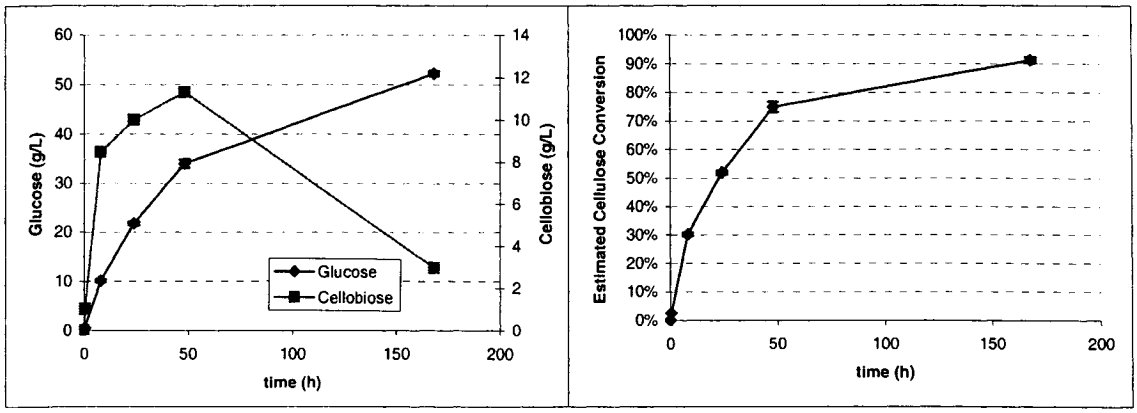
Model Predictions for Glucose Rates at 10% Initial Insoluble Solids, 20 mg/g Enzyme Loading

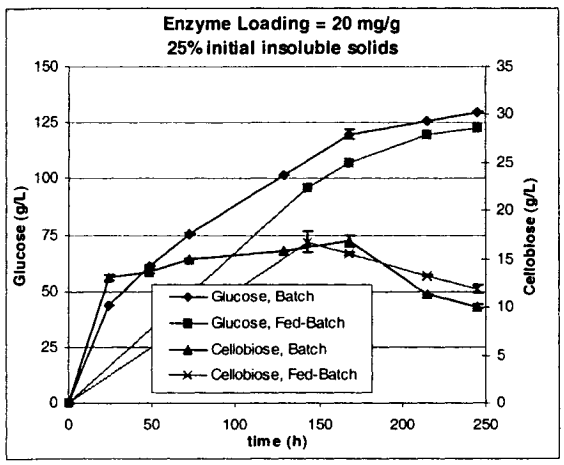


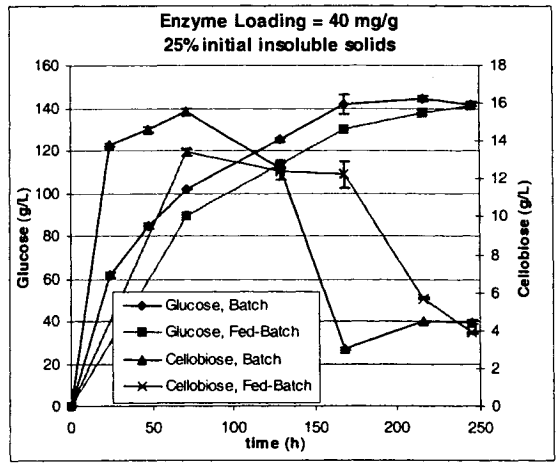
Model Predictions for Cellulose Rates at 10% Initial Insoluble Solids, 20 mg/g Enzyme Loading

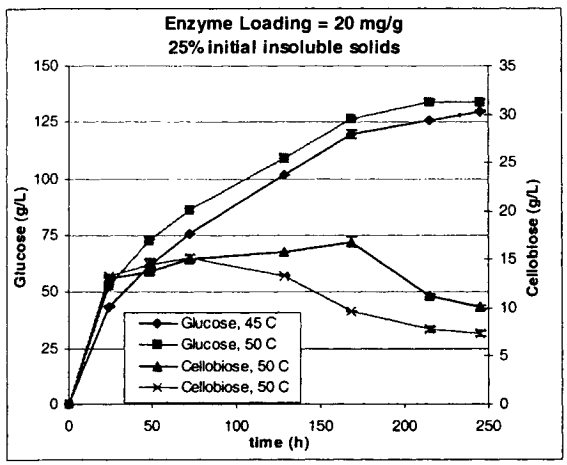


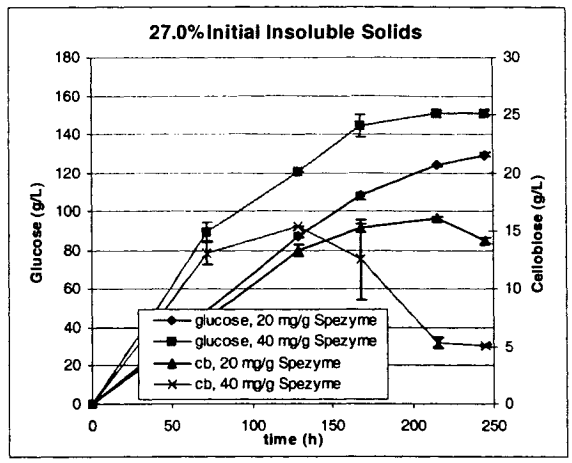
Experiment #7



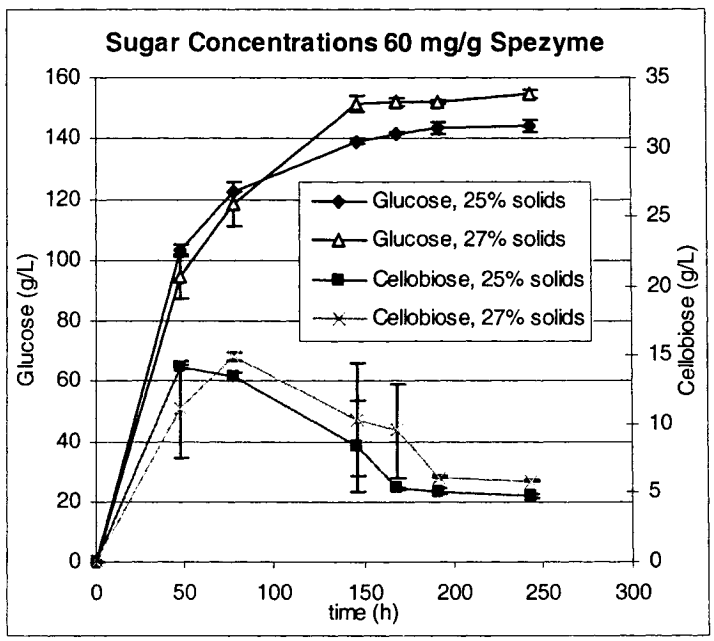


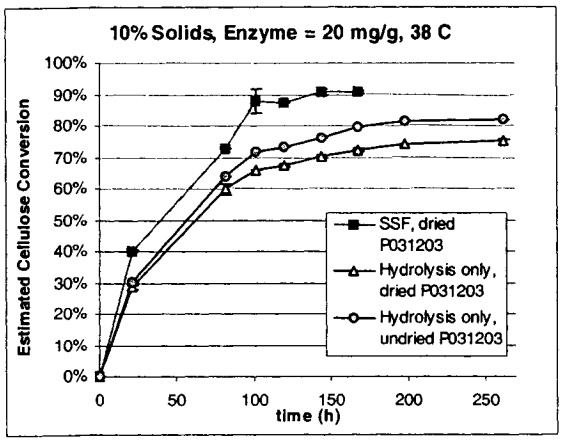


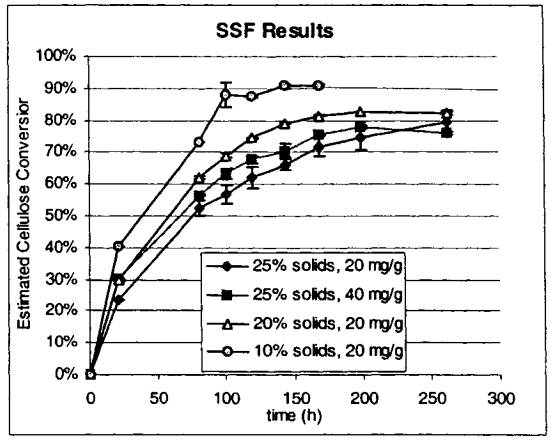


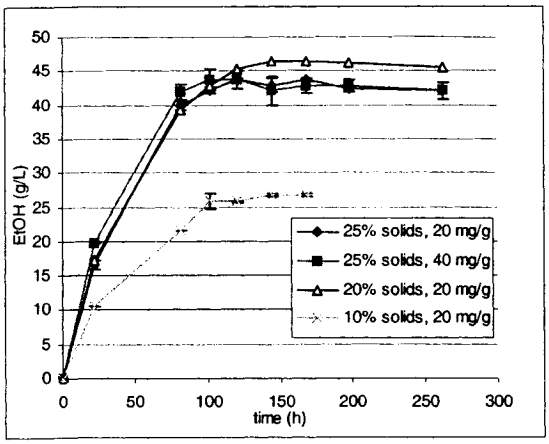


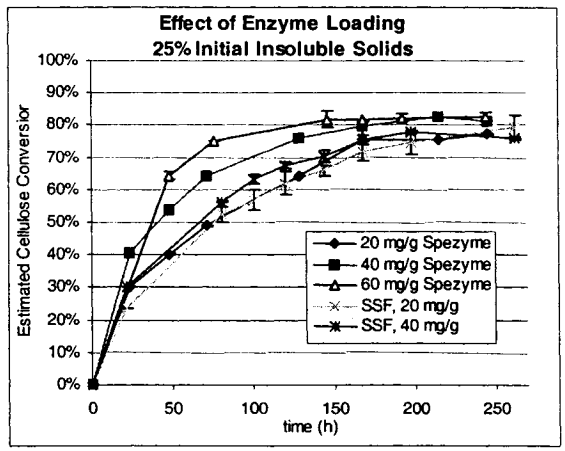
Experiment #9



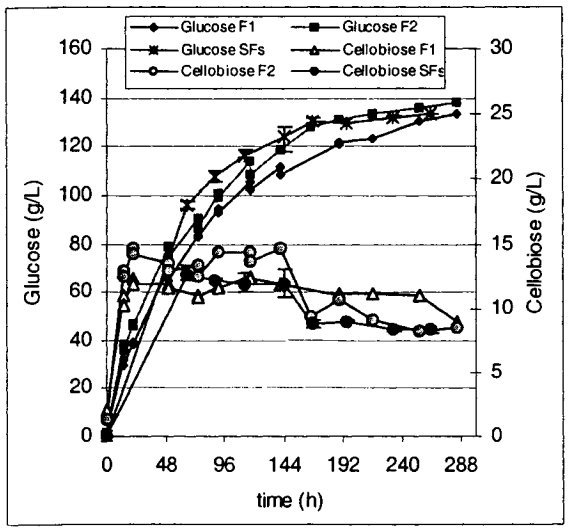


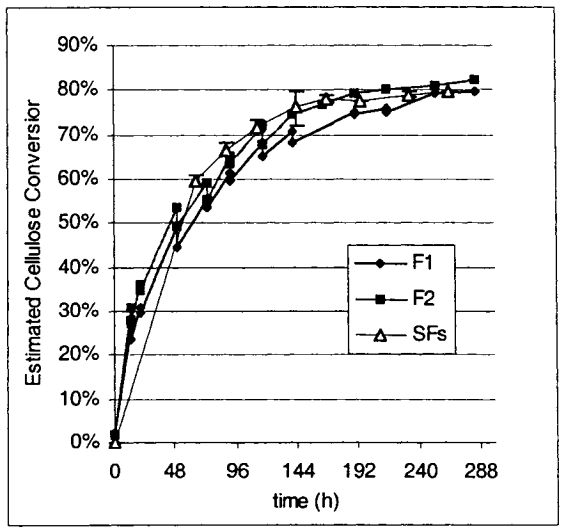


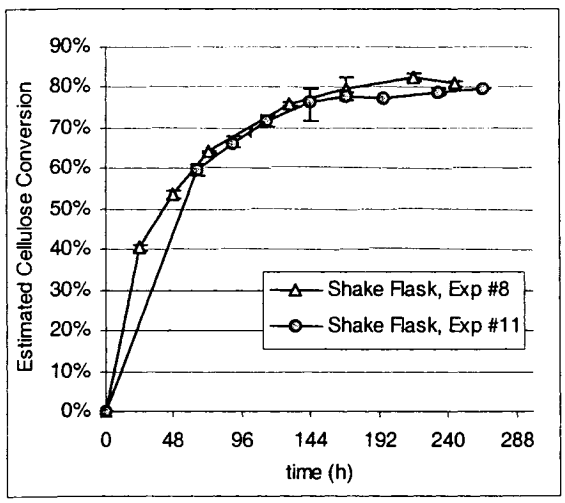


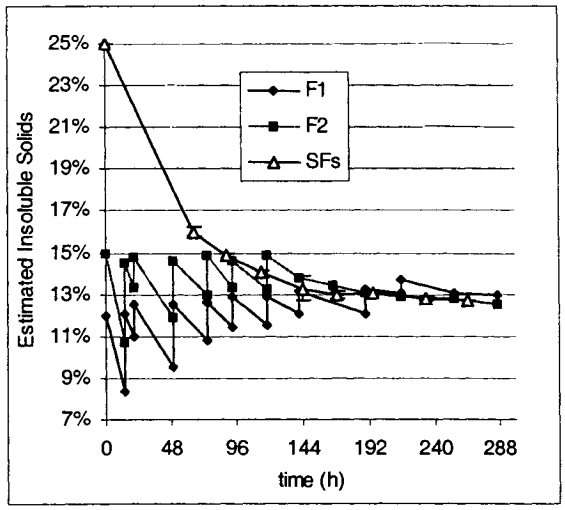


Experiment #11

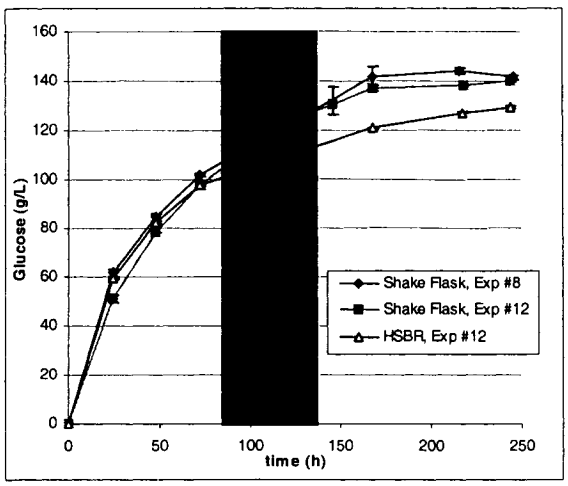


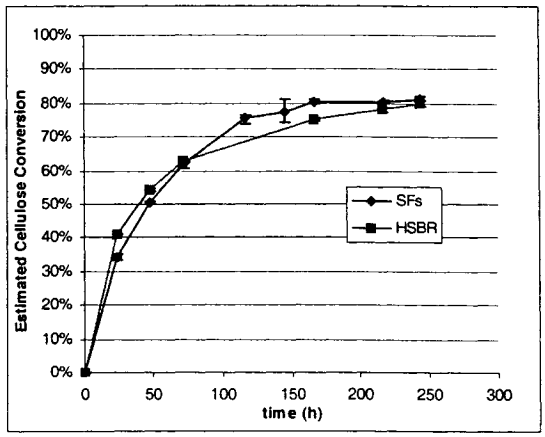


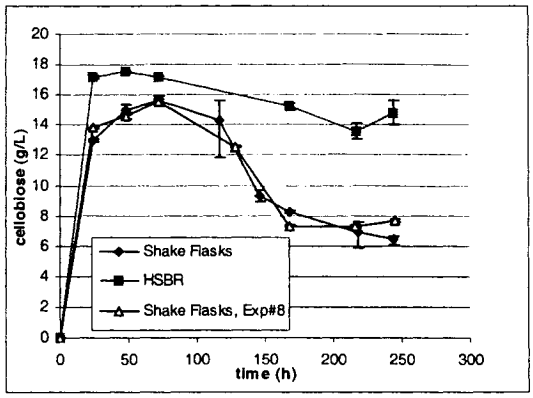


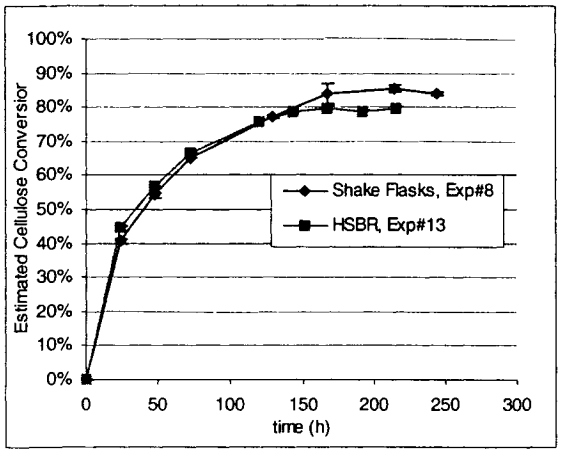


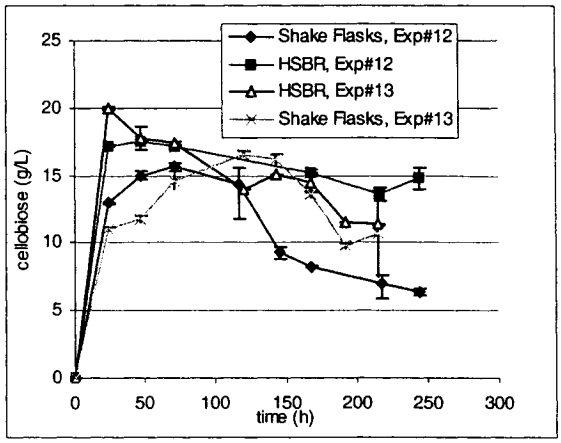
Experiment #12

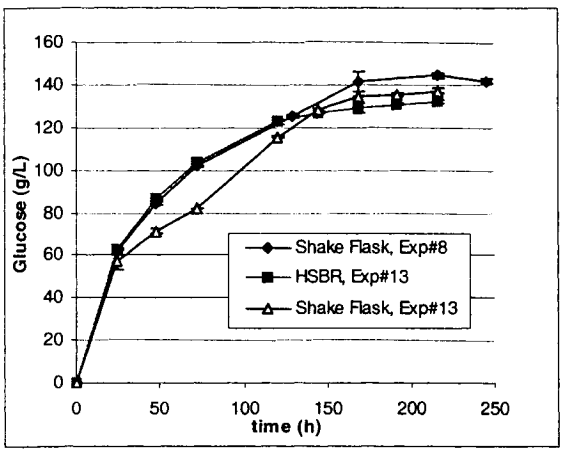


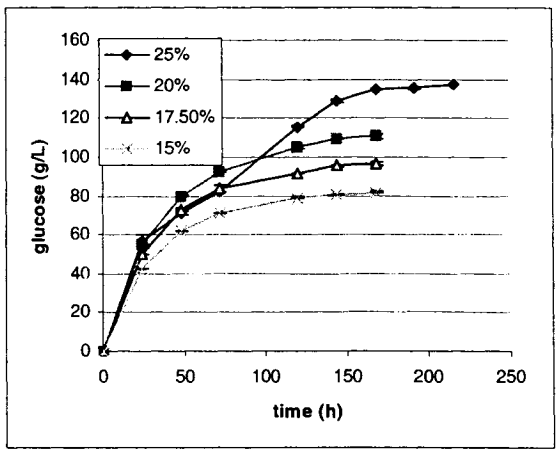


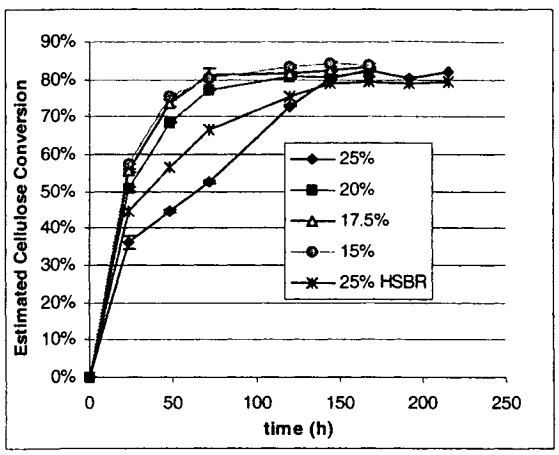


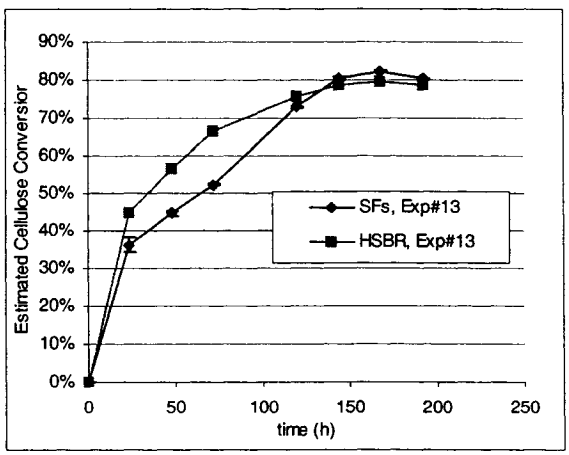




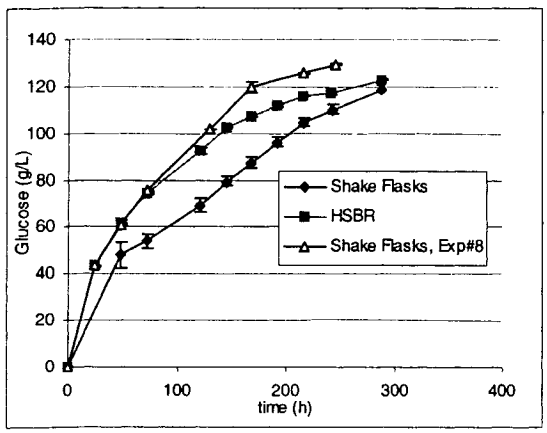


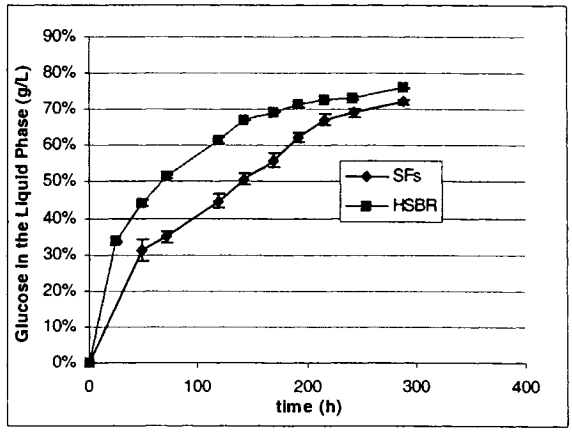




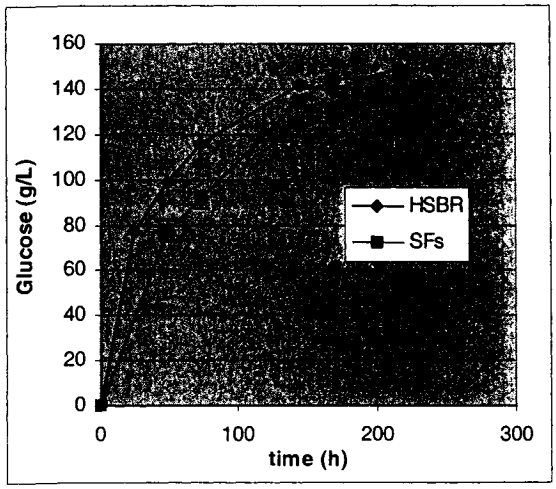


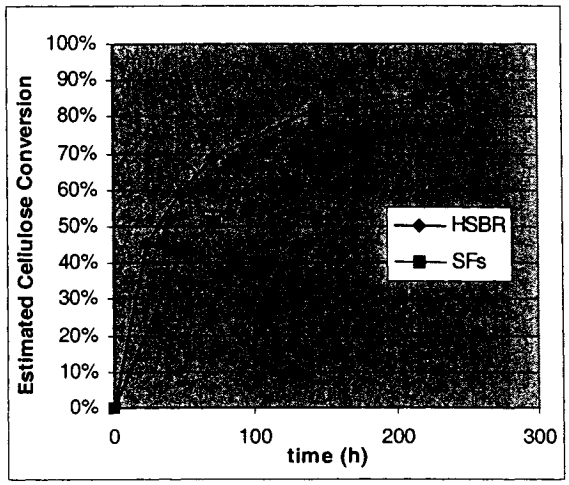
Experiment #14





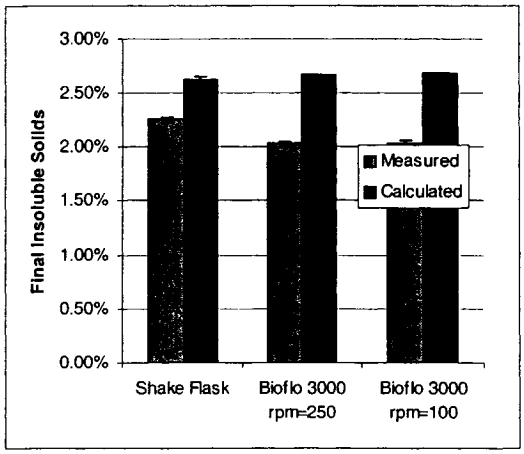
Experiment #15

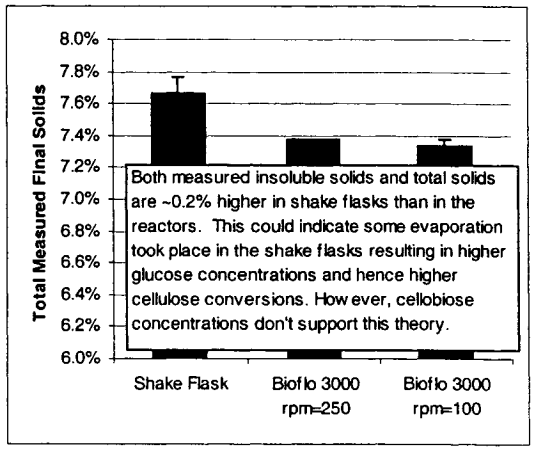


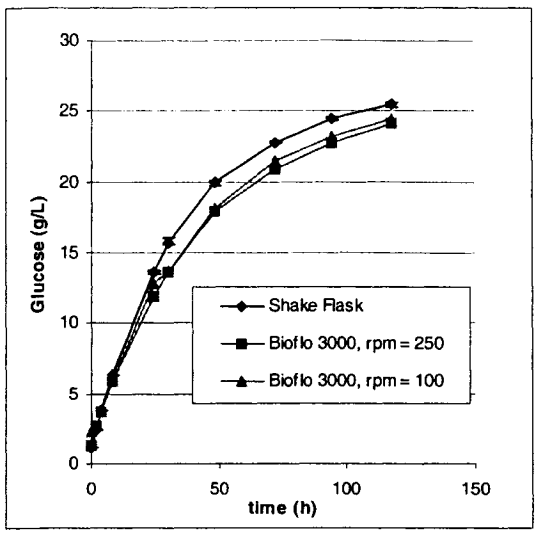


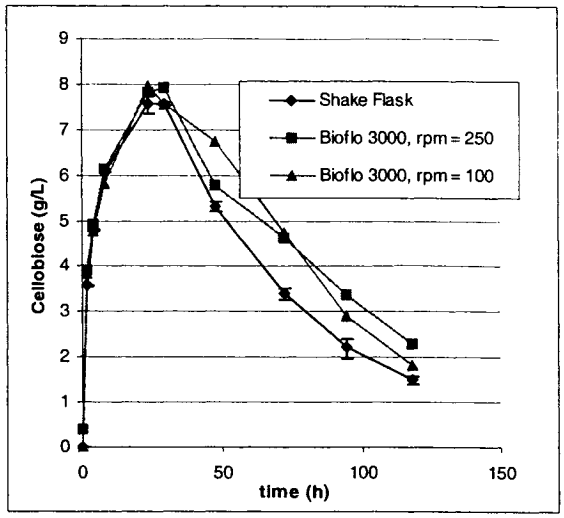
B. Unwashed Solids Experiments

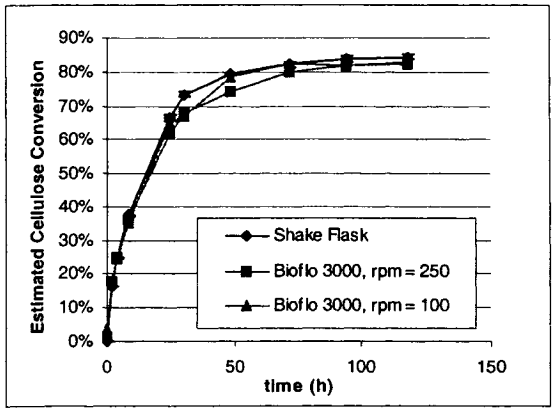
Experiment #3

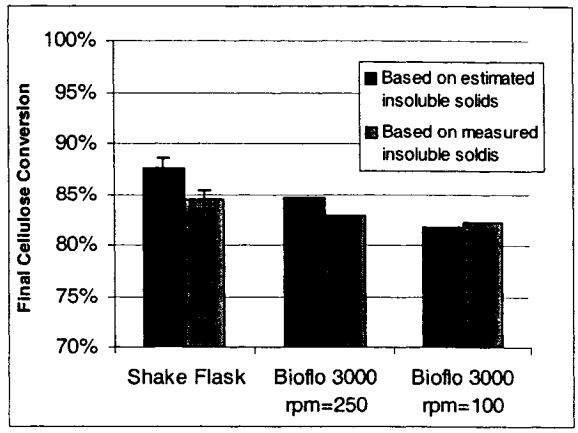




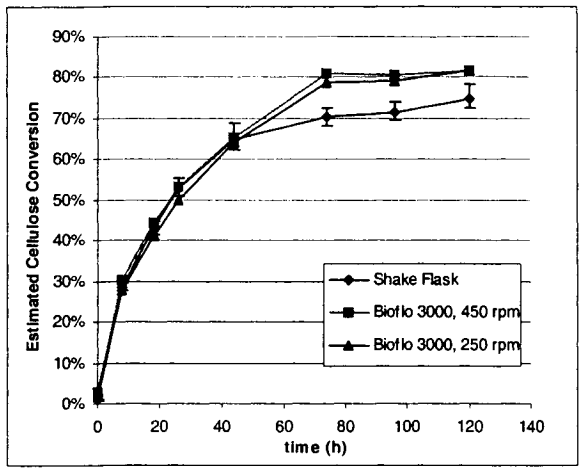


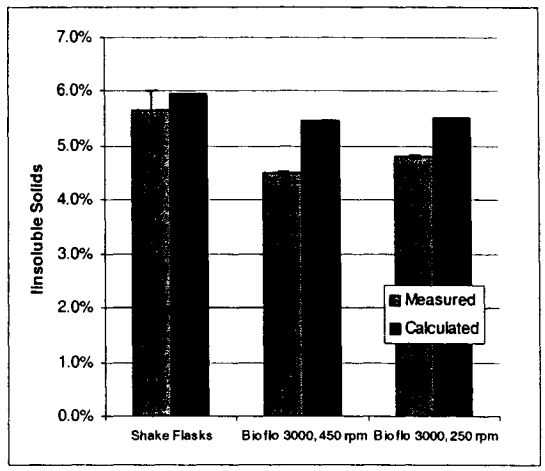


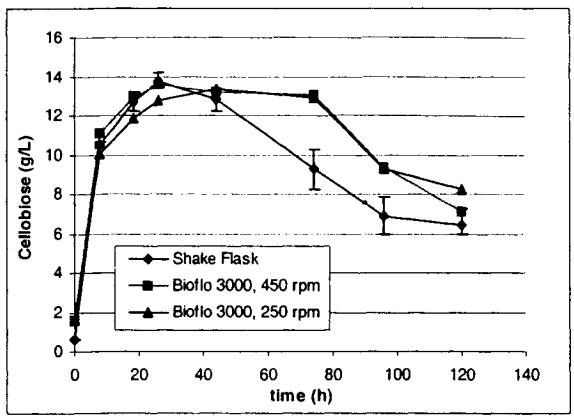


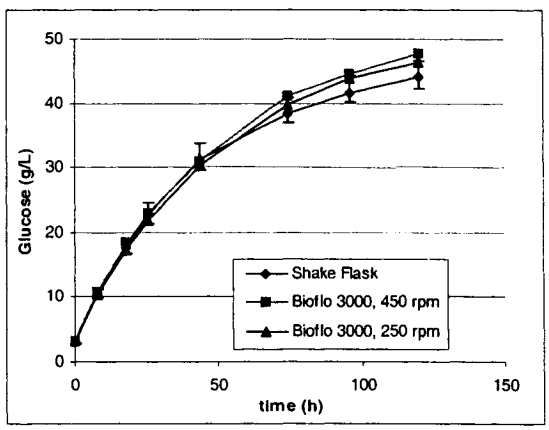


Experiment #4

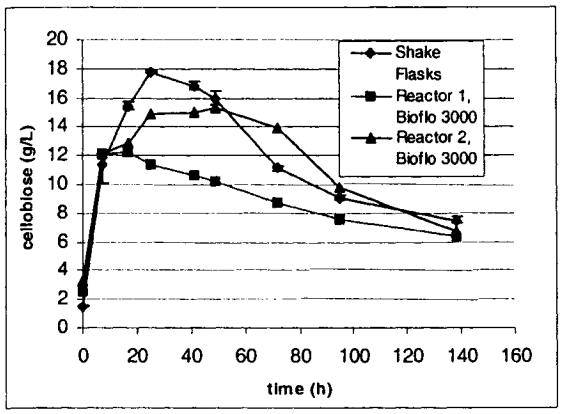


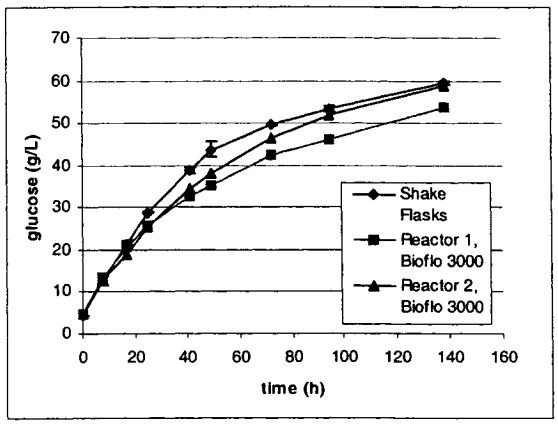


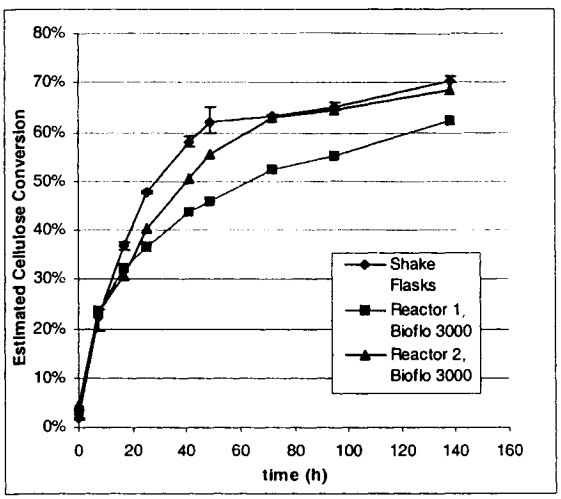


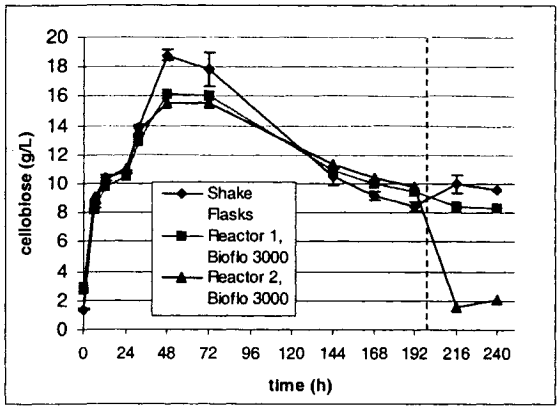


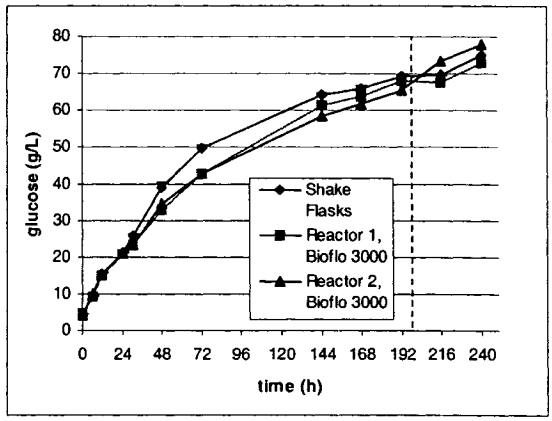
Experiment #5

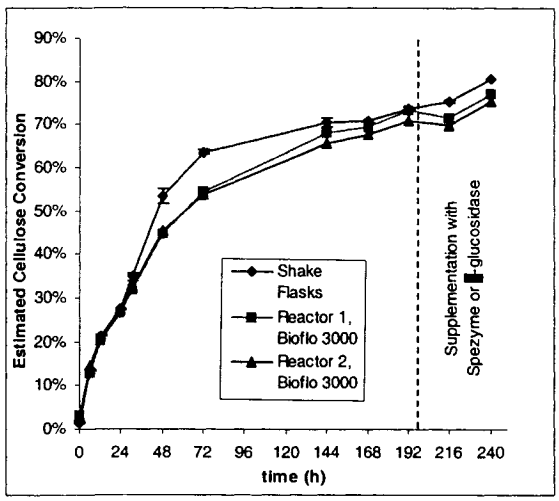


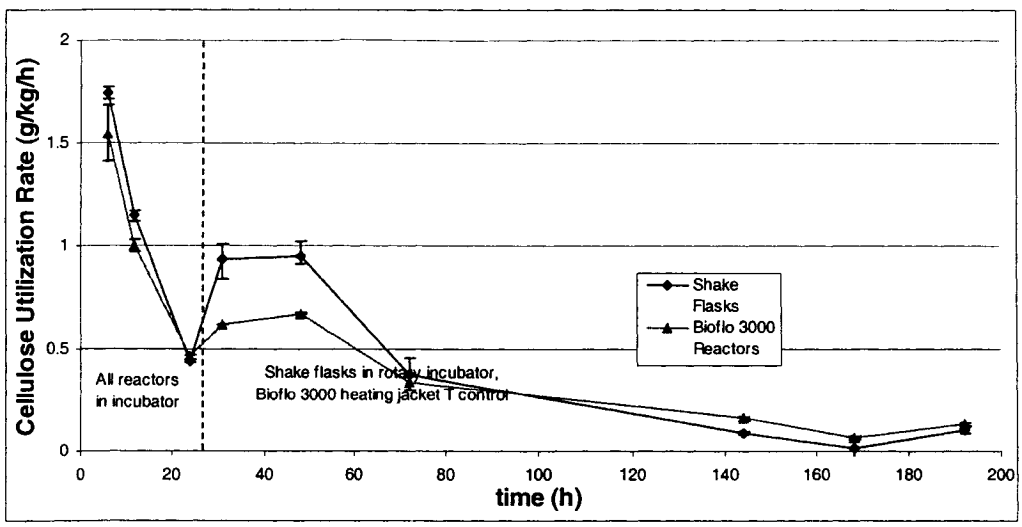




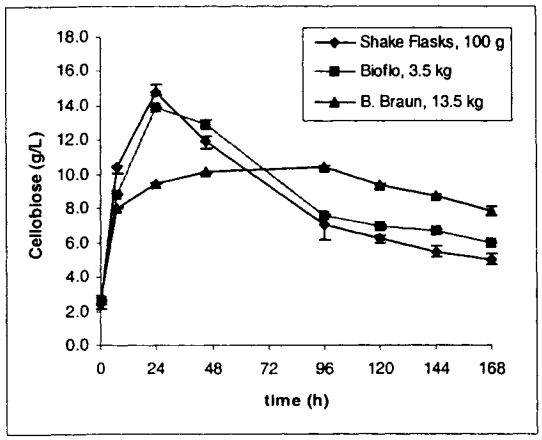


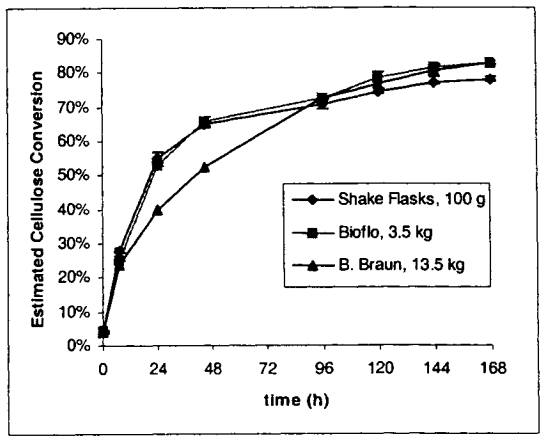


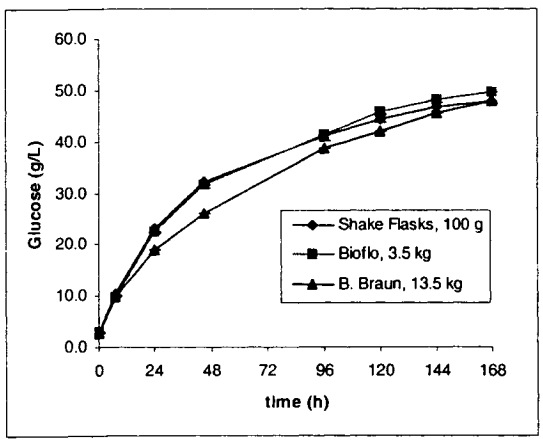


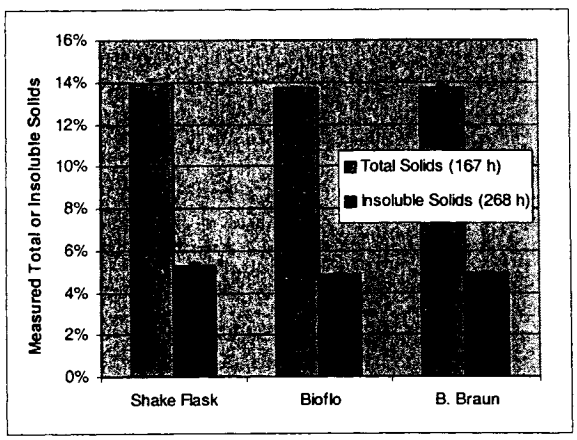


Experiment #7

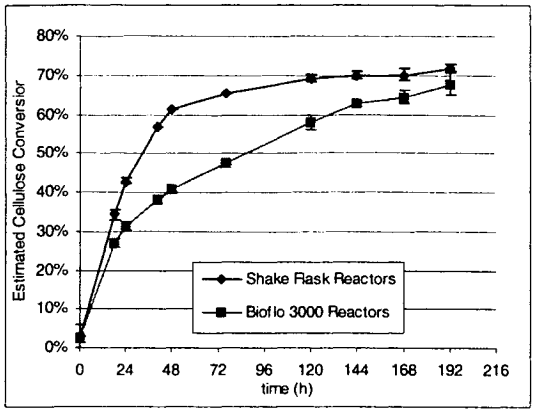


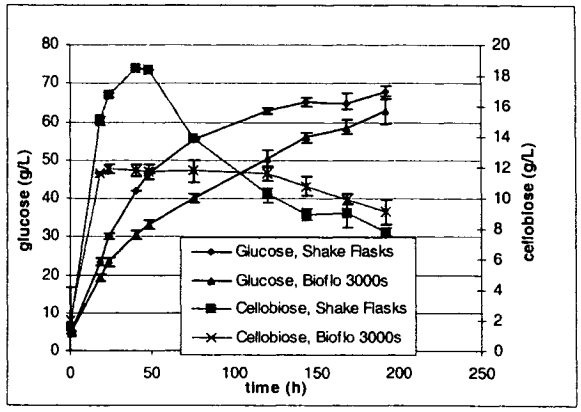






Experiment #8





APPENDIX 5: PROTEIN SPOT VOLUME DATA

Normalized Averages

Tentative ID	Gel → spot	G8A	G11	G15A	G19A	D	F1	F2	G8B	G15B	G19B
	1	0.05481229	0.08569907	0.1185265	0.06269989	0.05722198	0.05029423	0.16802856	0.05770377	0.11230209	0.03226645
	2	0.02055856	0.05170457	0.04318866	0.01704495	0.04797614	0.01061719	0	0.01015075	0.01778859	0.02375241
	3	0	0.02151305	0	0.00684725	0	0.02013302	0	0.02164316	0	0
	4	0.06252997	0.06277113	0.1046598	0.08929544	0.05995145	0.07008132	0.12126651	0.0658298	0.03451486	0.09822746
	5	0.01318162	0.0286713	0.04679313	0.01635604	0.0268343	0.04354582	0.03719331	0.01387699	0.01472379	0.03051776
	6	0.02390143	0.02482045	0.030525	0.0196298	0.03483105	0.02943943	0.03511459	0.02516299	0.04214632	0.03466387
	7	0	0.03218937	0.01300429	0.00709935	0.00744602	0.01116133	0.00542642	0	0	0
	8	0.05033769	0.04425545	0.03739177	0.04071177	0.01928254	0.04320176	0.04230852	0.05299301	0.02063694	0.05325984
ADH 1	9	0.04912443	0.01858271	0.0227664	0.04438285	0.02905268	0.04205552	0.04079783	0.05171517	0.0485654	0.06581668
	10	0.03945604	0.03105818	0.05264421	0.03387033	0.03004402	0.06005314	0.04334788	0.04153726	0.04272927	0.03603578
GAP	11	0	0.0180261	0.02658764	0.02356808	0	0.03081192	0.02498387	0	0.02416651	0.04592932
GAP	12	0.01412473	0.03782021	0.04603756	0.03903074	0.0360537	0.03922826	0.04881358	0.01487101	0.04960648	0.03951644
groES	13	0.01270581	0.03477256	0.04332491	0.01135594	0.02108456	0.00369302	0.00895087	0.0133759	0.02175902	0.01464433
PYK	14	0.02924855	0.03599114	0.04795901	0.03959767	0.03851441	0.02571275	0.03449219	0.03079173	0.06161205	0.06008444
PYK	15	0.01255248	0.02482045	0.02852458	0.04447579	0.02726167	0.02383352	0.01845314	0.01321566	0.02845696	0.02281501
	16	0	0	0.00467746	0.00976669	0.00536422	0.00552363	0.00238388	0	0.00129072	0.0049802
GDT	17	0	0.00628681	0.00650292	0.01158248	0.00987148	0.00994216	0.00672712	0	0.00687987	0.00933597
GDT	18	0	0	0.00779731	0.00504657	0.01040239	0.00813025	0.010241	0	0.01044263	0.00486777
GDT	19	0	0	0	0	0.00805183	0.00388001	0.00479948	0	0	0
PGK	20	0	0.00581757	0.0150078	0.01163011	0.02263986	0.02480025	0.01366424	0	0.02993492	0.02642786
TKT	21	0.0303888	0.00553747	0	0.00084458	0	0.00420911	0.01773556	0.03199217	0.00122167	0
TKT	22	0.01936355	0.00215466	0	0.00700757	0	0.00301986	0.0066108	0.02038569	0.00198654	0
TKT	23	0.01034256	0.00166627	0	0.00404864	0	0.00150339	0.00659116	0.01152229	0	0
ENO	24	0.06843444	0.04577687	0.06224995	0.06215968	0.02828164	0.03367284	0.02975767	0.05763723	0.03740836	0.08168099
ENO	24B	0.07863219	0.03889634	0.09722325	0.06615488	0.10956439	0.05050926	0.04279043	0.08278121	0.12505926	0.08548375
groEL	25	0.0082665	0.01030524	0.03581869	0.01741089	0.02271035	0.00646044	0.0084886	0.00870316	0.01598928	0.03064386
	26	0.00788926	0.00500718	0.02067773	0.01407788	0.0165773	0.00373976	0.01046609	0.00893945	0	0.0075341
	27	0.00764709	0.00852526	0.00378563	0.00688327	0	0.00433252	0.01165652	0.00717003	0	0
	28	0.02393794	0.00940986	0.01063691	0.01029644	0.00882948	0.00548436	0.00902036	0.02520101	0	0
dnaK	29	0.01001034	0.01115873	0.02230501	0.01110152	0	0.00322929	0.00678151	0.01053777	0	0.01213296
PDC	30	0.03909583	0.03814221	0.00764248	0.01410111	0.01720514	0.01777136	0.01359474	0.04115839	0.00200779	0.00264051
PDC	31	0.03166535	0.01098516	0.01176099	0.0168672	0.02308706	0.01543027	0.01604811	0.03333519	0.01089412	0.00243388
PDC	32	0.01496075	0.00526095	0	0.01763975	0.02151855	0.00602289	0.01917072	0.01575097	0.00539527	0.00324518
PDC	33	0	0	0	0.0062083	0.0159825	0.00276743	0.01109908	0	0.00091625	0.00212699
	34	0	0	0	0	0.01195989	0	0	0	0	0
	35	0	0	0	0	0.00559993	0	0	0	0	0
	36	0	0.00478453	0	0.00876644	0	0.01213367	0.00583884	0	0	0.00341382
	37	0	0.00599593	0	0.00409976	0.00449184	0.00790212	0.00298664	0	0	0
	38	0	0.00265861	0	0	0.00429137	0.00985615	0.00451245	0	0	0
	39	0	0.00299856	0	0	0.00274489	0.00211858	0	0	0	0
	40	0	0.00204573	0	0	0.0049809	0.00857154	0	0	0	0
	41	0	0.00459301	0	0	0.01510792	0.01186066	0.00487501	0	0	0
	42	0	0.00223127	0	0	0.0053488	0.00769831	0.00435231	0	0	0
	43	0	0.00528609	0.00812555	0.00628614	0.0089198	0.01658025	0.00463783	0	0	0
	44	0.00823243	0.00464927	0.01176099	0.0048421	0.01291597	0.02723484	0.01005519	0	0.04741411	0
	45	0.0133167	0.01390352	0	0.01337619	0.00496107	0.00665678	0	0.01401957	0	0.0088969
	46	0.00741345	0.00948887	0	0.00679381	0	0.00688117	0.00811243	0.00780419	0	0
	47	0.00250685	0	0	0	0	0.00450829	0.01355396	0.00565726	0	0
	48	0.00500395	0.00354202	0	0	0.00361066	0.00258418	0.00983916	0.00550517	0	0
	49	0.01120414	0.00377065	0	0.01292544	0.01203479	0.0038987	0.00422692	0.01581615	0	0

Tentative ID	Gel → spot	Normalized Averages G8A	G11	G15A	G19A	D	F1	F2	G8B	G15B	G19B
	50	0.00056587	0.00152262	0	0	0.00764208	0.0008695	0.01356453	0.00568578	0	0
	51	0	0	0	0	0.01224407	0.00663995	0	0	0	0
	52	0	0	0.02595128	0.08929544	0.02142162	0.02395506	0.03663133	0	0.04333213	0.02602525
XI	53	0.04022878	0.01747666	0.00639763	0.01106551	0	0.02929919	0.02564254	0.04235068	0.04700644	0.01057418
XK	54	0.01103377	0.01735935	0.00842128	0.00408117	0	0.02090715	0.01423226	0.01161599	0.04371988	0.01191115
	55	0.01035108	0.00304525	0	0.01287432	0.02243498	0.01226456	0.01724912	0.01779334	0.02362472	0.04643068
	56	0	0	0	0	0	0	0	0	0	0
	57	0.02329906	0.00864855	0	0	0.01299308	0.00106396	0	0.00672597	0	0
	58	0.0063888	0	0	0	0	0	0	0.02452882	0	0
	59	0	0	0	0	0	0.00975518	0	0	0	0
	60	0.00868877	0.00491621	0	0	0.00499632	0.00809846	0.01762075	0.00914722	0.0103125	0.01303541
	61	0.01495224	0.02184103	0	0	0.01465852	0.00712238	0.01246021	0.01227867	0.00717333	0.00205558
	62	0.00935443	0.01128801	0	0	0.00201571	0.00973087	0.02411522	0.00984792	0	0.00288055
	63	0.00757286	0.01946612	0	0	0	0.00668109	0.01233936	0.0059791	0.00636198	0.00518986
ADH 2	64	0.05021113	0.01634427	0.0361779	0.00933336	0.00834483	0.00459991	0	0.03478277	0.015814	0.01487222
	65	0.01748585	0.01675245	0	0	0	0.00688117	0	0.0184085	0.02941969	0.02885566
	66	0.00273928	0.00454393	0	0	0.00948155	0.00420911	0.00536146	0.0062534	0.00627434	0.00604065
	67	0	0.00608331	0.00857611	0.01223305	0.00411734	0.00863325	0.00458798	0	0	0
	68	0	0.00689251	0	0	0.00516595	0.00846309	0.00630564	0	0	0
	69	0	0	0	0	0	0.00352099	0	0	0	0
	70	0	0	0.00795988	0.01116309	0	0.00505055	0	0	0	0
	71	0.01628232	0.02024539	0	0	0	0.0099646	0.01376545	0.01959128	0	0
	72	0	0	0	0	0.00456234	0.00121729	0	0	0	0
	73	0	0	0.00962277	0.01753287	0.00540387	0	0.0059038	0	0	0
	74	0	0	0	0	0	0	0	0	0.0159999	0.0173608
	75	0	0	0	0	0.01022836	0	0	0	0	0
	76	0	0	0	0	0.00312821	0	0	0	0	0
	77	0	0	0	0	0.00264576	0	0	0	0	0
	78	0	0	0	0	0.01381919	0	0	0	0	0
	79	0	0	0	0	0.02165072	0	0.00845385	0	0	0
	80	0	0	0	0	0	0	0	0	0	0.00365994
	81	0	0	0	0	0	0	0	0	0	0
	82	0	0	0	0.00451217	0	0	0	0	0	0
	83	0	0	0	0	0	0	0	0	0	0
	84	0	0	0	0	0	0	0	0	0	0
	85	0	0	0	0	0.01056541	0	0	0	0	0
	86	0	0	0	0	0.00709354	0	0	0	0	0
	87	0	0	0	0	0	0	0	0	0	0
	88	0	0	0	0	0	0	0	0	0	0
	89	0	0	0	0	0	0	0	0	0	0
	90	0	0	0	0	0	0	0	0	0	0
	91	0	0	0	0	0	0	0	0	0	0
	92	0	0	0	0	0	0	0	0	0	0
	93	0	0	0	0	0	0.01859037	0	0	0	0
	94	0	0	0.02751353	0.04472556	0	0	0	0.01593022	0.04741411	0

APPENDIX 6: Abbreviations and Symbols

A. Acronyms

2-DE	2-dimensional gel electrophoresis
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BFGS	Broyden-Fletcher-Goldfarb-Shanno quasi-newton optimization method
CBH	Cellobiohydrolase
CER	CO_2 evolution rate
CI	Crystallinity index
CSS	Closed system saccharification
DCW	Dry cell weight
DOE	Department of Energy
DOF	Degrees of freedom
DP	Degree of polymerization
ED	Entner-Douderoff
EG	Endoglucanase
EM	Embden-Meyerhoff
ESP	Enzyme Sugar Platform
EtOH	Ethanol
FPU	Filter paper units
HCA	Hierarchical cluster analysis
HHF	Hybrid hydrolysis and fermentation
HSP	Heat shock protein
HMF	Hydroxymethyl furfural
HPLC	High performance liquid chromatography
HSBR	High solids bioreactor
HSF	Hybrid saccharification and fermentation
IEF	Isoelectric focusing
IPG	Immobilized pH gradient
IUB-MB	International Union of Biochemistry and Molecular Biology
LAP	Laboratory analytical protocol
LS	Least squares
MPC	Model predictive control
MSE	Mean squared error
MTBE	Methyl tertiary butyl ether
NLP	Nonlinear programming
NMR	Nuclear magnetic resonance
NREL	National Renewable Energy Laboratory

OD	Optical density
ORF	Open read frame
PCA	Principal component analysis
PCS	Pretreated corn stover
PP	Pentose phosphate
PWS	Pretreated wheat straw
RNAP	RNA polymerase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHF	Separate hydrolysis and fermentation
SSCF	Simultaneous saccharification and cofermentation
SSF	Simultaneous saccharification and fermentation
STR	Stirred tank reactor
TCA	Tricarboxylic acid
PSS	Partial soluble sugars (glucose, cellobiose, and xylose by HPLC)
IS	Insoluble solids measurement
TS	Total solids measurement
ISC	Insoluble solids composition (analytical determination of solids composition)

B. Symbols

α	confidence level
B	linear model parameters
C	cellulose (g/kg)
C_0	initial cellulose concentration
cb	cellobiose in liquid fraction (g/L)
CB	cellobiose (g/kg)
D	dilution rate (h^{-1})
D	reactor tank diameter (cm)
Da	Damköhler number
Δ	change from initial conditions
ε	effectiveness factor
ξ	extent of cellulose conversion
E_1	enzyme 1 (all EGs and CBHs) mass concentration (g/kg)
E_2	enzyme 2 (β-glucosidase) mass concentration (g/kg)
$E_{1\ feed}$	enzyme 1 (all EGs and CBHs) mass concentration in the feed (g/kg)
$E_{2\ feed}$	enzyme 2 (β-glucosidase) mass concentration in the feed (g/kg)
f	any function
F	feed mass flow rate (kg/h)
F1	Feeding policy 1
F2	Feeding policy 2
g	glucose in liquid fraction (g/L)
G	glucose (g/kg)
G_{feed}	glucose mass concentration in the feed (g/kg)
G_{total}	total glucan in both phases (g glucose / kg)
H	reactor tank working height (cm)

k	discete time index
μ	mean
n	number of observations
ν	rate
ν_0	initial rate
ρ	liquid phase density (g/L)
Φ	NPL objective function
ϕ	discretized NPL objective function
r_1	rate of cellulose conversion to glucose (g/kg/h)
r_2	rate of cellulose conversion to cellobiose (g/kg/h)
r_3	rate of cellobiose conversion to glucose (g/kg/h)
r	desired control trajectory
r	radial dimension operator for reactor tank (cm)
R	reactor tank diameter (cm)
R	correlation coefficient
R_S	substrate reactivity
S_{equiv}	"equivalent" insoluble solids fed to reactor (kg)
S_I	Insoluble solids (g/g)
$S_{I,0}$	initial insoluble solids (g/g)
S_{IF}	insoluble solids in the feed stream (g/g)
$S_{I,SP}$	insoluble solids set point (g/g)
S_T	total solids in slurry (g/g)
$S_{T,0}$	initial total solids in slurry (g/g)
s^2	sample variance
t	t-statistic
t	time (h)
t_0	initial time (h)
t_f	final time (h)
T	temperature (°C)
u	controlled variable
V	working "volume" of the reactor (kg)
V_0	initial working "volume" of the reactor (kg)
x	vector of state variables
x_0	initial value of state variables
x_{cell}	cellulose fraction in the solid phase (g/g)
$x_{cell,0}$	initial cellulose fraction in the solid phase (g/g)
x	xylose in liquid fraction (g/L)
X	xylose mass concentration (g/kg)
$\underline{X}$	model variables or "factors"
$\overline{X}$	sample mean
X_{feed}	xylose mass concentration in the feed (g/kg)
$\hat{Y}$	estimated model response