2012

Developments in enzyme immobilization and near-infrared Raman spectroscopy with downstream renewable energy applications

Jason S. Lupoi
Iowa State University

Developments in enzyme immobilization and near-infrared Raman spectroscopy with downstream renewable energy applications

by

Jason S. Lupoi

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Analytical Chemistry

Program of Study Committee:

Emily A. Smith, Major Professor

Robert S. Houk

Joseph Burnett

Young-Jin Lee

Brian Trewyn

Iowa State University

Ames, Iowa

2012

For my wife, Kristin, with love

TABLE OF CONTENTS

APPENDIX: RAMAN SPECTROSCOPY MEASUREMENTS OF GLUCOSE AND XYLOSE IN HYDROLYSATE: ROLE OF CORN STOVER PRETREATMENT AND ENZYME COMPOSITION

(A paper published in *Bioresource Technology*, **2011**, 102, 5169-5176)

ACKNOWLEDGEMENTS

The end has arrived. Some compare graduate school to a roller coaster, filled with ups and downs, death-defying twists and turns, self-image distorting mirrors. I have reached the end of this stretch of track, and am eager to see what further the carnival of life has to offer.

Going through the process of graduate school has strengthened me emotionally and mentally. While there were times that I felt nothing would ever work again, the times when things were successfully ensured moments of elation. However fleeting they may have been, reflection upon these achievements reiterated in my mind the original reasons I chose to further myself and my education.

I would like to extend gratitude to my committee for their assistance in this process. I would also like to express my indebtedness to my colleagues in the Smith Group for listening to my research conundrums and making our work environment feel like a family. The help I received from my colleagues helped hone the path to my personal successes. I will always reflect upon this, and do my best to help others reach their potentials. The lasting friendships I was lucky enough to form, and the advice and guidance received, especially from Kris, Cherry, and Deepak will always be cherished. Without you guys, life would definitely have been quite vapid.

My family was quite instrumental in my success, both financially and emotionally. A kind ear listening to my rants and achievements really went a long way, and knowing that I had my family's support back on the home-front helped ease my vexation at myself, or share in the euphoria of my accomplishments.

Lastly, I would like to thank my wife, Kristin, for believing from the commencement of this journey, that I could achieve whatever I set out to. Without the comfort of our home sanctuary, I could not have retained my focus and drive. It was for my wife that I originally set sail on the course. What an adventurous and fascinating existence we have to look

forward to together, exploring our fullest potentials on an endless sea of knowledge, in our self-helmed schooner…

CHAPTER 1.

GENERAL INTRODUCTION

RESEARCH MOTIVATION AND OBJECTIVES

The use of renewable forms of energy has been proposed to be one long-term pathway to reducing the world's dependence on fossil fuels, despite the current requirement of fossil fuel input to generate renewable fuels. Branches of this pathway include bioethanol, solar, wind, geothermal, and tidal technologies, to name a few. Bioethanol is produced from the hydrolysis of cellulose to glucose, and subsequent or simultaneous fermentation of glucose to ethanol. Forest and agricultural wastes comprise a substantial proportion of available biomass that would help alleviate public concern over using food crops for fuel.[1] Additional sources of biomass include municipal solid waste, wastepaper, and dedicated bioethanol crops. Grasses such as *miscanthus* and switchgrass, as well as crop residues, such as sugarcane bagasse and corn stover, have shown promise as sources for bioethanol production.[1-2] In order for the bioethanol industry to supplant fossil fuel production, the conversion of lignocellulosic materials to fuel must be economically viable. This economic practicality includes the techniques used to screen and select from the diverse range of biomass, the cost of pretreatment reagents and organisms used in hydrolyzing and fermenting cellulose and glucose, respectively, and the neutralization of pretreatment effluents. Hemicellulosic and lignin wastes and side-reaction products must also be developed into useful bio-products. An instrumental technique is needed that can rapidly screen biomass *in situ*, or with little to no sample preparation, that is not inhibited by the presence of water, and that can provide both qualitative and quantitative information.

The first objective, in this thesis research, was to study how immobilized cellulase compared to free cellulase in the ability to convert cellulose to glucose at sub-optimal reaction conditions. Immobilization has been shown to provide enhanced enzyme stability as well as recyclability, which inherently lowers production costs of the end product. The second objective was to develop near-infrared (NIR) Raman spectroscopy applications for the characterization of lignocellulosic biomass. NIR Raman spectroscopy can meet current

needs in the rapid screening of biomass for biofuel production, given its non-invasive and non-destructive nature, and its amenability to samples containing water.

IMMOBILIZATION OF CELLULASE

One of the most costly parameters of bioethanol production is the enzymes used to degrade cellulose, e.g., cellulase.[3] To work towards alleviating this economically detrimental factor, enzymes can be immobilized on solid supports such as silica, polystyrene, etc. The use of immobilized enzymes, as opposed to enzymes free in solution, has been shown to provide greater enzyme stability and flexibility when choosing reaction conditions such as pH, temperature, and organic solvent content.[4-5] The recyclability of costly enzymes for use in subsequent reactions bestows another benefit to immobilization techniques. While enzymes used in solution can also be recycled, these methods are more labor-intensive, involving separation and purification steps.[6-7] Removing immobilized enzymes from the reaction matrix can be achieved simply, for example, via centrifugation or through removing metallic particles with a magnet.

Enzymes can be immobilized onto solid supports using, for example, electrostatic interactions between the protein and carrier (physisorption), or by covalently linking functional groups on the periphery of the protein to functionalities grafted on the carrier surface (chemisorption). Both techniques have their advantages and disadvantages.[4, 8] Physisorption is the simpler technique, where the enzyme is incubated with the carrier for a specific duration, and then non-immobilized enzyme is removed, often through centrifugation. The electrostatic interactions between enzyme and carrier depend on the differences in hydrophobicity/hydrophilicity. A hydrophilic protein, for example, will not have a high affinity for a hydrophobic carrier such as polystyrene. In this case, the protein can restructure such that buried hydrophobic regions will be exposed to the hydrophobic carrier.[9-10] Matching hydrophobic/hydrophilic proteins with like surfaces ensures better affinity and stability.

Physisorption can also result in the enzyme leaching or desorbing from the carrier surface. This can be overcome by chemisorption conjugation techniques. Often, lysine

residues in the enzyme will be linked to a particle that has been grafted with a specific functionality, for example, a carboxylic acid, such that a stable amide bond is created between the enzyme and carrier.[4] Chemisorbed enzymes have also conferred higher flexibility in non-optimal reaction conditions, compared to physisorbed enzymes.[4] In order to maximize chemisorption techniques, knowledge of the protein crystal structure is necessary, such that essential amino acids involved in catalysis are not hindered by or bound to the carrier. For some enzymes, the structure may not be available. Chemisorption can change the structure of the enzyme, which can cause inactivation.[4, 11]

HYDROLYSIS OF MICROCRYSTALLINE CELLULOSE

The enzymatic hydrolysis of microcrystalline cellulose is enabled by enzymes collectively called cellulase. This collection includes predominantly endoglucanases, exoglucanases (cellobiohydrolases), and β-glucosidases.[3, 12-14] Varying proportions of these specific enzymes have been isolated from organisms that produce cellulase. For example, Beldman et al. isolated 10 enzymes from *Trichoderma viride* (6 endoglucanases, 3 exoglucanases, and 1 β-glucosidase).[12] The *Trichoderma* cellulases lack large quantities of β-glucosidase, a key enzyme for the efficient breakdown to cellulose to glucose [15-16], although some commercial preparations such as Accellerase 1500 (Danisco, Rochester, NY) have supplemented β-glucosidase into the enzyme mixture.

A schematic depicting the degradation of cellulose to glucose is shown in Figure 1. Endoglucanases randomly cleave β-1,4 glucosidic bonds, with a preference for the amorphous regions of cellulose.[3, 13-14] This produces new chain ends. Cellobiohydrolases cleave cellobiose units from the ends of the cellulose chains.[3, 13-14] Glucose is predominantly produced when β-glucosidase cleaves cellobiose. Endoglucanases and cellobiohydrolases work synergistically, rather than in sequence.[14] Cellobiose and glucose formation causes strong inhibition of cellobiohydrolase and β-glucosidase, respectively.[17]

SIMULTANEOUS SACCHARIFICATION AND FERMENTATION

The fermentation of glucose to ethanol is typically achieved via the microorganism yeast. Simultaneous saccharification and fermentation (SSF) can improve upon separated hydrolysis and fermentation steps.[18] In SSF, cellulase and yeast are added to the reaction matrix, as well as a yeast extract/peptone solution, and the substrate (Figure 2). Therefore, as cellulase enzymes produce glucose from cellulose, it is quickly fermented to ethanol, thereby reducing the inhibitory effect of glucose.

The principle drawbacks of using SSF reactions are the buildup of ethanol in the reaction matrix, which can be detrimental to cellulase activity, and the difference in optimal reaction conditions required by cellulase and the yeast. Typically, temperatures of 45-50°C are used in the hydrolysis step, while the yeast requires much lower temperatures, typically below 37°C.[19-20] This dissimilarity in optimal temperature has led to one line of research to find or develop thermotolerant yeast strains.[21-22]

Chromatographic analytical techniques are the most commonly used for the detection of ethanol in reaction matrices. These include high performance liquid chromatography with refractive index detection [23-24], and gas chromatography/mass spectrometry (GCMS).[25-26] NIR spectroscopy can also be used when coupled with chemometrics to determine ethanol concentrations.[27] These techniques are not without their shortcomings, however. Chromatographic techniques require sample clean-up to prevent solids from clogging the columns. This is detrimental to the rapidity of the analysis, as is the retention time on long columns such as the Aminex HPX-87H (Bio-Rad, Hercules, CA, USA). NIR spectra must be used in conjunction with chemometrics such that subtle differences can be elucidated from what otherwise appears to be considerable spectral overlap.

LIGNOCELLULOSIC BIOMASS

Plant cell walls are primarily composed of cellulose (40-60%), hemicellulose (20-40%), and lignin (10-25%), in addition to what are termed extractives.[28-31] Cellulose and lignin rank first and second, respectively, in biopolymer abundance. Cellulose is a

homopolymer of β-glucopyranose units covalently linked via 1,4 glucosidic bonds. Glucose is the main hydrolysis product of cellulose, if complete degradation is achieved. Both crystalline and amorphous regions have been elucidated in the structure of cellulose, but the proportions depend upon the sample. Most cellulose is located in the secondary wall of plant cell, and due to the high degree of crystallinity and hydrogen bonding between individual fibers, provides tensile strength to the plant.[32-33] Hemicellulose is a heteropolymer consisting of a branched conglomeration of glucose, mannose, xylose, arabinose, rhamnose, as well as glucuronic and galacturonic acids. These components are readily released from plant cell walls via mild acid hydrolysis.[32-33] Hemicellulose functions as a support material within the plant cell wall, like cellulose.[33] The composition of hemicellulose depends upon plant source (i.e., hardwoods, softwoods, herbaceous plants), and also the plant organ (stem, root, leaves, bark, etc.).[28, 31, 33]

Lignin is a three-dimensional, heteropolymer of phenylpropanoid units (Figure 3). These units are derived from hydroxycinnamyl alcohols and to a lesser degree, cinnamaldehydes. The phenyl moieties (Figure 4) are named p-hydroxyphenol (H), guaiacyl (G), or syringyl (S) depending on whether there are 0, 1, or 2 methoxy functionalities on the phenyl ring, respectively.[34-35] The presence and proportion of these moieties allow lignin to be classified according to gymnosperm (predominantly G), angiosperm (G and S), and herbaceous (G, S, and H) lignins.[34-36] The primary functions of lignin are to strengthen cell walls by forming lignin-carbohydrate complexes, to provide resistance against microbial attack, and lignin plays a paramount role in controlling water transport by decreasing the permeability of the cell wall.[34-36]

Herbaceous plant cell walls contain a higher proportion of "extractives", compared to gymnosperms and angiosperms.[28-29, 31] These include substances such as tannins, terpenes, stilbenes, lignans, flavanoids, resin, waxes, chlorophyll, and fatty acids.[37-39] These molecules, many of which are phenolic structures, can be readily removed by performing a Soxhlet extraction using solvents such as ethanol or acetone.[40]

Lignocellulosic biomass has presented a possible solution to relinquishing the world's dependence on fossil fuels. A study done in 2005 estimated the annual sum of biomass generated from forest and agricultural wastes at 1.5 billion tons.[1] This is not including other sources such as dedicated bioethanol crops or municipal solid wastes, such as newsprint. In order for a biorefinery to be economically viable, hemicellulosic and lignin wastes generated from the hydrolysis of cellulose to glucose, and subsequent or simultaneous fermentation to ethanol, must also be developed into useful bio-products, such as fuels, adhesives, and organic acids.[41] While feedstocks such as sugarcane (bagasse) and the starch from corn grain are currently utilized in commercial ethanol plants, other potentially useful crops for bioethanol need to be found. Ideally, these crops would have high cellulose and low lignin contents as lignin has been shown to irreversibly bind cellulase enzymes reducing the efficiency of the hydrolysis.[42-44] Lignin also inhibits enzymes from reaching the cellulose interior due to its matrix like presence in the cell wall.[32, 45] Rapid growing perennial grasses, such as switchgrass or *miscanthus*, are common sources for biofuel research, as they can grow to 3.5 m in a single growing season, and can provide more tonnage per square acre than currently used feedstocks.[2]

The ratio of syringyl to guaiacyl units provides a measure of how susceptible plant lignin may be to degradation. Common techniques for quantifying this relationship include wet chemistry, such as acidolysis[46], thioacidolysis[47], and nitrobenzene oxidation[48], and instrumental techniques, such as pyrolysis gas chromatography/mass spectrometry.[49-50] These techniques are destructive and/or laborious.

The biomass used for thioacidolysis must be extractive-free to prevent detection of phenolics in the extractives. The mechanism of thioacidolysis has been outlined by Rolando et al.[51] After extracting phenolics from lignin, the samples are silylated for GC analysis. Thioacidolysis results in β-O-4 cleavage, which have been said to represent approximately 30% of the linkages found in softwood and herbaceous lignin, and greater than 50% in hardwoods. This technique cannot account for monomers linked via carbon-carbon or biphenyl ether linkages, therefore, does not provide a complete representation of monomer

content.[52] Nonetheless, thioacidolysis is one of the most widely accepted lignin monomer quantitation techniques.

RAMAN SPECTROSCOPY

A necessity in bioethanol research is a high-throughput, non-destructive, instrumental analysis method capable of (1) screening different feedstocks, (2) quantifying useful parameters such as lignin S/G ratios without the use of significant volumes of toxic reagents that require remediation, (3) detecting and/or monitoring the efficient removal of extractives such as carotenoids, and (4) detecting ethanol in complex fermentation matrices without extensive sample clean-up. A technique possessing these qualities would permit a rapid, *in situ* assessment of potentially high ethanol yielding plants. The method could be employed by agronomists to compare the raw feedstock with genetically altered biomass. NIR Raman spectroscopy is one promising analytical method for the bioenergy field.

The phenomenon of Raman scatter was first predicted by Adolf Smekal in 1921, and was later experimentally observed by Sir C.V. Raman and K.S. Krishnan, and independently by Grigory Landsberg and Leonid Mandelstam in 1928.[53-55] Raman spectroscopy is a vibrational, instrumental technique that measures the scattering of molecules when irradiated with an excitation source, conventionally a laser. The scattering generated from the excitation source that relaxes with an equivalent loss of energy is termed Rayleigh scattering, and is the most intense scattering process (Figure 5). About 1 in every 1 million photons is inelastically scattered with a shift in energy relative to the incident source that is characteristic of specific vibrational modes in a molecule.[56-57] This process is called either Stokes or anti-Stokes scattering, dependent upon whether the excited molecule has absorbed some of the energy, and has been promoted to a higher vibrational level (Stokes), or if molecules originally at higher vibrational levels relax to the ground state (anti-Stokes).[56-57] These processes are depicted in Figure 5. Anti-Stokes lines are much weaker than Stokes lines due to the lower probability of molecules populating excited vibrational levels at ambient conditions. Stokes scattering is red-shifted relative to the incident line, since the molecule is absorbing light, and the resulting photon is lower in energy. Anti-stokes

scattering occurs at higher energies (blue-shifts) since the molecule is losing energy by relaxing to the ground state.[56-57] Raman spectroscopy typically measures Stokes lines. Anti-stokes lines may be measured if the Stokes lines are inhibited by fluorescence.

In order for a molecule to be Raman "active", there must be a change in the polarizability of the electron cloud. This is a major difference when comparing Raman to infrared spectroscopy, where an infrared "active" molecule will have a change in dipole moment when excited. Another difference between infrared and Raman spectroscopy is the requirement that the exciting energy match the difference in energy from the ground to first excited vibrational level, such that the molecule is actually promoted to a higher, real vibration level in infrared spectroscopy. In Raman spectroscopy, molecules are promoted to short-lived, "virtual" vibrational levels, and therefore, do not necessitate matching the excitation energy to the difference between the ground and first excited vibrational level, unless performing resonance Raman spectroscopy, where the excitation wavelength resonates with a molecular electronic state transition.[56-57] Symmetric vibrations typically have the largest change in polarizability, and therefore, are the strongest vibrational modes in a Raman spectrum. In infrared spectroscopy, asymmetric vibrations lead to the largest change in dipole, and are therefore, the most intense vibrational modes in an infrared spectrum. Due to these specific selection rules, the two techniques are said to be complementary to one another, and the resultant spectra can be used in combination to provide the most complete structural information.

Raman spectroscopy provides a versatile analytical tool for the measurement of lignocellulosic biomass due to its non-destructiveness, little to no sample preparation, field-portability, and most importantly, the capability measuring analytes in an aqueous environment or complex matrices. When considering the use of Raman spectroscopy for the measurement of biomass, several parameters must be taken into consideration. The choice of a suitable excitation wavelength that minimizes spectral contributions from fluorescence is crucial. NIR lasers, while providing lower frequencies, and therefore less intense Raman spectra, are often selected for this purpose. Additionally, the low energy photons emitted by NIR lasers can prevent photo-damage to the sample. Another important parameter is the

selection of a detector that provides a high quantum efficiency for the excitation wavelength chosen. Charge-coupled devices (CCDs) are often employed at wavelengths below 800 nm, but suffer from a low response at wavelengths higher than 1.1 μm.[58] This limits their applicability to techniques employing 1064 nm excitation sources. The advent of photodiode detectors capable of efficient operating in this region, such as InGaAs detectors, have overcome this limitation.[58]

Raman spectra of lignocellulosic materials contain peaks characteristic of polysaccharide and lignin content. Fourier Transform (FT)-Raman has been routinely used in the characterization of biomass, but can suffer from low throughput due to longer acquisition times, and a decreased spectral resolution as the acquisition time is increased.[59-63] More recently, ultraviolet resonance Raman spectroscopy has been used to study lignin due to its strong absorption of light in this region.[64-65] Only vibrations in resonance are enhanced, and spectral contributions from polysaccharides have been reported to be present, but minimal. Precautions must be taken to prevent sample damage due to the higher energy of the excitation light, and often, samples must be rotated or attenuating filters must be inserted into the beam path to reduce the laser power.

NIR Raman spectroscopy can also be used to quantify glucose and ethanol yields from complex hydrolysates.[25, 66] In this work, both glucose and ethanol were accurately quantified in hydrolysate and fermentation liquors using 785 nm excitation. The detection limits of glucose and ethanol were 4 and 6 mg/mL, respectively. The detection limit of glucose was shown to be directly related to the choice of extraction and pretreatment solvents. The only required sample clean-up was the removal of unreacted, residual solids via centrifugation.

Chemometric techniques such as principal component analysis (PCA), principal component regression (PCR), and partial least squares (PLS) provided a way to extract useful information from the Raman spectra that other techniques such as using peak intensities or areas to generate calibration curves, or multipeak fitting algorithms could not do. The techniques work best when the data matrix size is reduced considerably by removing all non-

informative spectral regions, and of the remaining data, using spectral regions that represent the component(s) being measured.

Chemometric treatment of Raman spectra have been used frequently to elucidate both qualitative, structural information, and quantitative prediction of lignocellulosic biomass composition. Ona and co-workers have used PLS to determine cell morphology such as length, width, and thickness of vessels and fibers, as well as the quantitative determination of hemicellulosic sugar composition and S/G ratios from FT-Raman spectra.[67-71] PLS was also used to deconvolute spectra of lignin model monomers, and isolate characteristic spectral features for each monomer to aid in the interpretation of UV resonance Raman spectra.[72] Nuopponen et al. used PCA to identify key spectral characteristics of tropical hardwoods from FT-IR spectra.[64] Lastly, features characteristic of different types of wood were elucidated from Raman profiles by genetic algorithms, and then clustered via pattern recognition, such that the Raman spectra of 98 soft, hard, and tropical woods could be classified by PCA.[73]

DISSERTATION OVERVIEW

This dissertation focuses on techniques for (1) increasing ethanol yields from saccharification and fermentation of cellulose using immobilized cellulase, and (2) the characterization and classification of lignocellulosic feedstocks, and quantification of useful parameters such as the syringyl/guaiacyl (S/G) lignin monomer content using 1064 nm dispersive multichannel Raman spectroscopy and chemometrics. Chapter 1 has outlined the motivation and objectives for this research, and has provided an introduction into the immobilization of enzymes on solid supports, the composition and utility of lignocellulosic biomass and the mechanisms and importance of the biodegradation of lignin. The principles and advantages of NIR Raman spectroscopy for biomass analysis were also summarized. Chapters 2-5 contain published or preliminary manuscripts. In Chapter 2, the immobilization of cellulase onto silica nanoparticles is compared and contrasted to using solution cellulase in SSF reactions. The comparison of immobilized to solution cellulase is investigated using non-optimal conditions in reaction pH and temperature, and with added ethanol to simulate

the SSF matrix. For the work reported in this dissertation, only physisorption was employed (Figure 6). Simulated hydrolysis reactions, or incubations were used to remove weakly bound cellulase. After performing the first incubation, and centrifuging to isolate the supernatant containing desorbed cellulase, fresh buffer was added, and the incubation performed again. The reaction conditions used in the incubations were identical to the range of optimal and non-optimal conditions studied. The desorbed enzyme was quantified and the catalytic potential of the enzyme was determined by adding cellulose. The immobilized cellulase was found to degrade cellulose more efficiently at temperatures below the optimum for the free enzyme. This resulted in an overlap between hydrolysis and fermentation conditions, and presents another alternative to meshing reaction parameters, and increasing ethanol yields. SSF reactions were performed to compare ethanol yields between free and immobilized cellulase, where the immobilized enzyme led to a ~2.3-2.5 fold increase in ethanol yield.

In this work, the objective was to compare and contrast immobilized to solution cellulase. Desorbed enzyme, therefore, needed to be removed such that a more accurate measurement of glucose produced from immobilized enzyme could be obtained. If the objective is to maximize ethanol yield, desorbed enzyme would not need to be removed and quantified. Any enzyme that desorbed from the nanoparticles would presumably function as solution cellulase, provided no irreversible structural rearrangement has occurred. At the conclusion of the reaction, enzyme that was still immobilized to silica could be isolated via centrifugation for further reactions.

In Chapter 3, the use of a home-built, 1064 nm, dispersive Raman spectrometer (Figure 7) is illustrated for the chemical analysis of plant lignin, and is compared with a 785 nm, as well as a Fourier-Transform Raman spectrometer. The S/G/H lignin content of a commercial sugarcane lignin is predicted, both qualitatively using first derivative Raman spectra, and quantitatively using PLS.

In Chapter 4, the utility of this same NIR instrument is extended to the characterization/classification of raw, ground lignocellulosic biomass using spectral

differences and principal component analysis. The emphasis in this research was on herbaceous plants, namely switchgrass, *miscanthus*, alfalfa, orchard grass, red clover, and pampas grass. The quantitation of lignin S/G content using PCR is also described.

The home-built, Raman instrument used in this research is shown in Figure 7. The instrument utilizes a 1064 nm Nd:YVO_4 laser, with a maximum output power of 1W. The sample stage consists of an optimally designed aluminum block with a hole cut into the center for NMR tubes. Multiple sample holders were designed to encompass differences in NMR tube diameter, or the use of microscope slides. The scattered light is then collected at 90°, and focused onto the entrance slit of the f/4.1 spectrometer. The collected scattering is sent to a liquid nitrogen cooled, InGaAs 1024 x 1, multichannel detector. Wavelengths were calibrated with a 50:50 solution of acetonitrile and toluene for peaks occurring from 785-1600 cm^{-1}, or with solid 4-acetamidophenol for spectral regions from 700-3000 cm^{-1}.

Using the home-built instrument described above, Raman spectra of biomass could be generated in 5 minutes. When analyzing solutions of lignin model monomers (ferulic, sinapic, or coumaric acids in dioxane) or soluble lignin, a blank was also measured containing the solvent to account for any matrix spectral contributions. When the raw or extractive-free biomass was measured, there was no need for a blank spectral subtraction, as the NMR tubes were not found to generate signal in the regions of interest thereby increasing the throughput for this method. In order to compare lignin contributions to the spectra, the spectra were normalized to the 1094 cm^{-1} of cellulose. The only other spectral processing was the baseline correction of all spectra to a specific wavenumber location.

Soxhlet extractions using 95% ethanol were performed on all samples studied. The Raman spectra of the raw and extractive-free samples were compared. Generally, woody plants contain very little extractive content, while herbaceous samples can contain upwards of 25% extractive material.[31] Therefore, if woody plants are being investigated by Raman spectroscopy, it should not be necessary to perform extractions, as the Raman spectra generated from before and after extraction were virtually identical. When measuring herbaceous plants, however, the high degree of extractives, especially phenolic extractives,

will strongly contribute to the Raman spectrum, obscuring accurate qualitative and quantitative data analysis of lignin.

When coupled with Raman spectra to create the training set for subsequent S and G prediction, thioacidolysis would need to be performed once, with samples covering the range of possible S and G content. Provided the calibration model works well for these samples, the S and G proportions can be predicted by adding the spectral data of future samples into the unknown data matrix. In this research, the spectra used in the training set were measured with 570 mW excitation power. Some samples were measured at lower laser powers to prevent sample charring. Higher laser powers were also employed to test the robustness of the calibration model, by adding variance not contained in the training set. The model predicted these samples accurately at the 95% confidence interval.

Chapter 5 describes the determination of ethanol detection limits using 1064 nm NIR Raman spectroscopy. Ethanol was spiked into a SSF matrix that contained no yeast, and previously denatured enzyme. Microcrystalline cellulose and kenaf bast were used as substrates in the fermentation liquors.

In chapter 6, the novelty of this research is explored, and the impact this work could have in the attempts to increase ethanol yields both enzymatically, as well as through the rapid, spectroscopic screening of viable biomass feedstocks is described.

The appendix describes the use of NIR Raman spectroscopy for glucose and xylose detection in complex, hydrolysis, and extraction liquors. Glucose detection limits using 785 nm excitation were determined, after a variety of extractions and/or pretreatments were performed on corn stover. Multipeak curve fitting was used to deconvolute the spectra of samples containing both glucose and xylose.

FIGURES

Figure 1. The pathway of cellulose degradation by cellulase enzymes (endoglucanase, exoglucanase, and cellobiase)

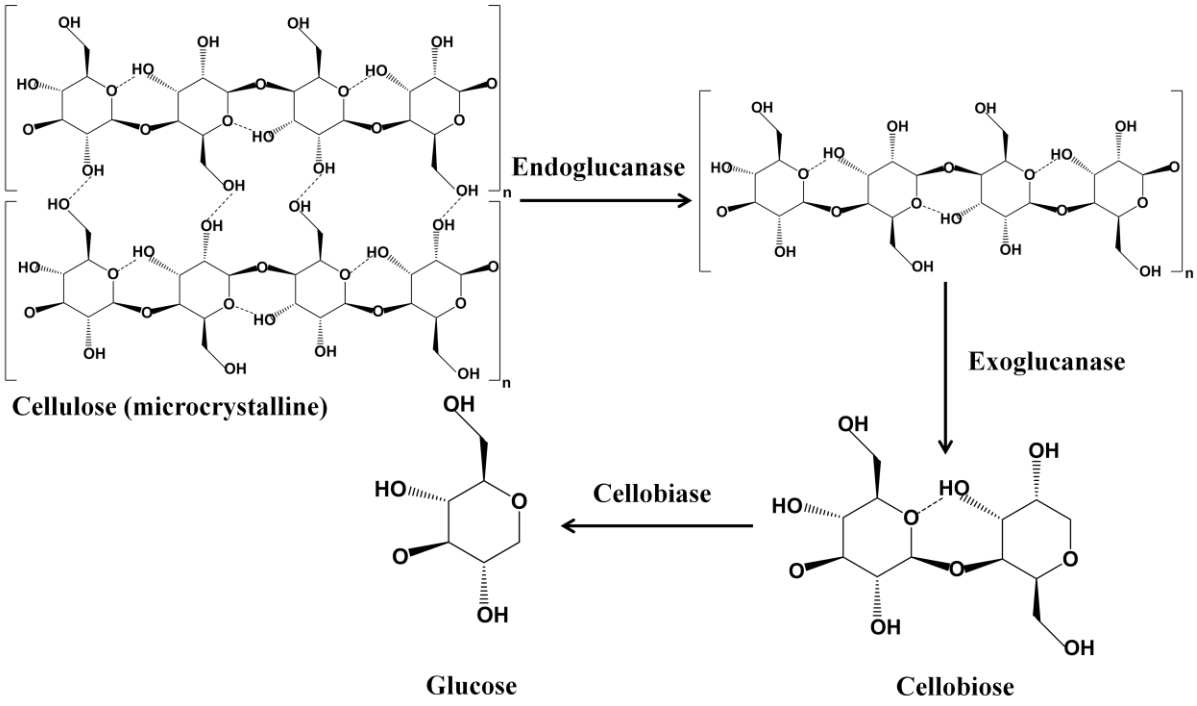


Figure 2. Simultaneous Saccharification and Fermentation

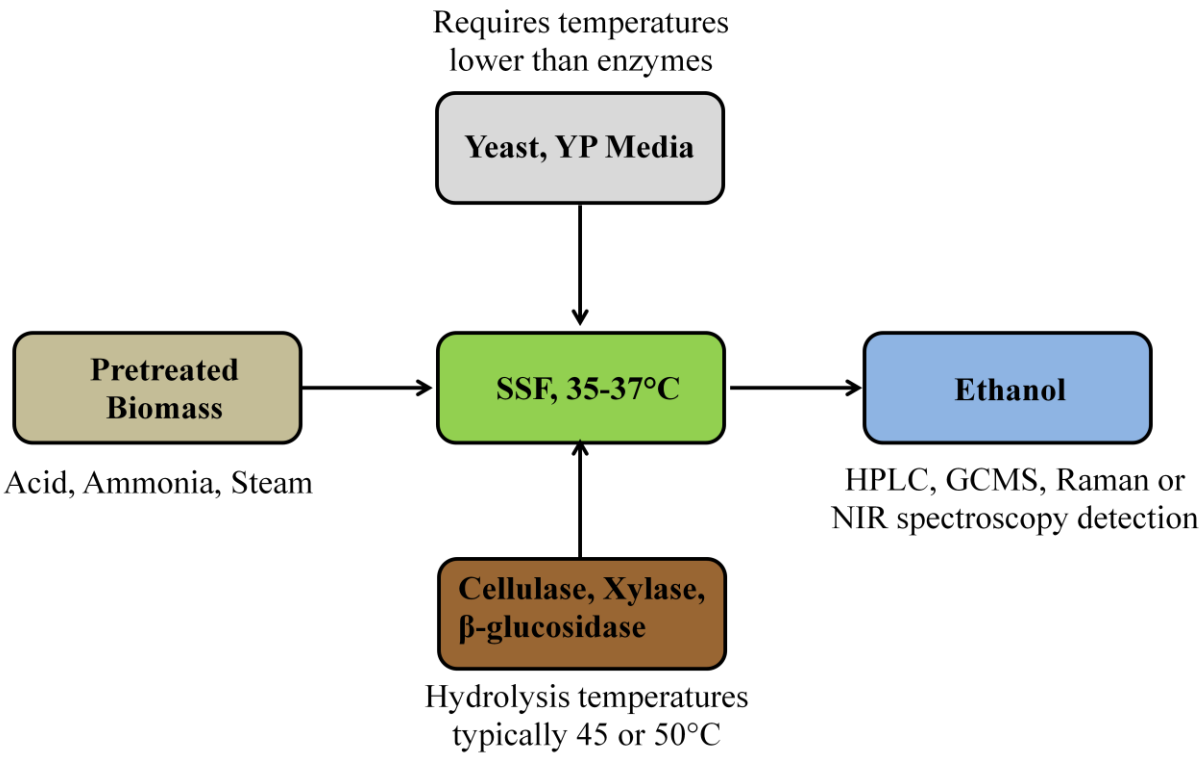


Figure 3. The structural representation of lignin.

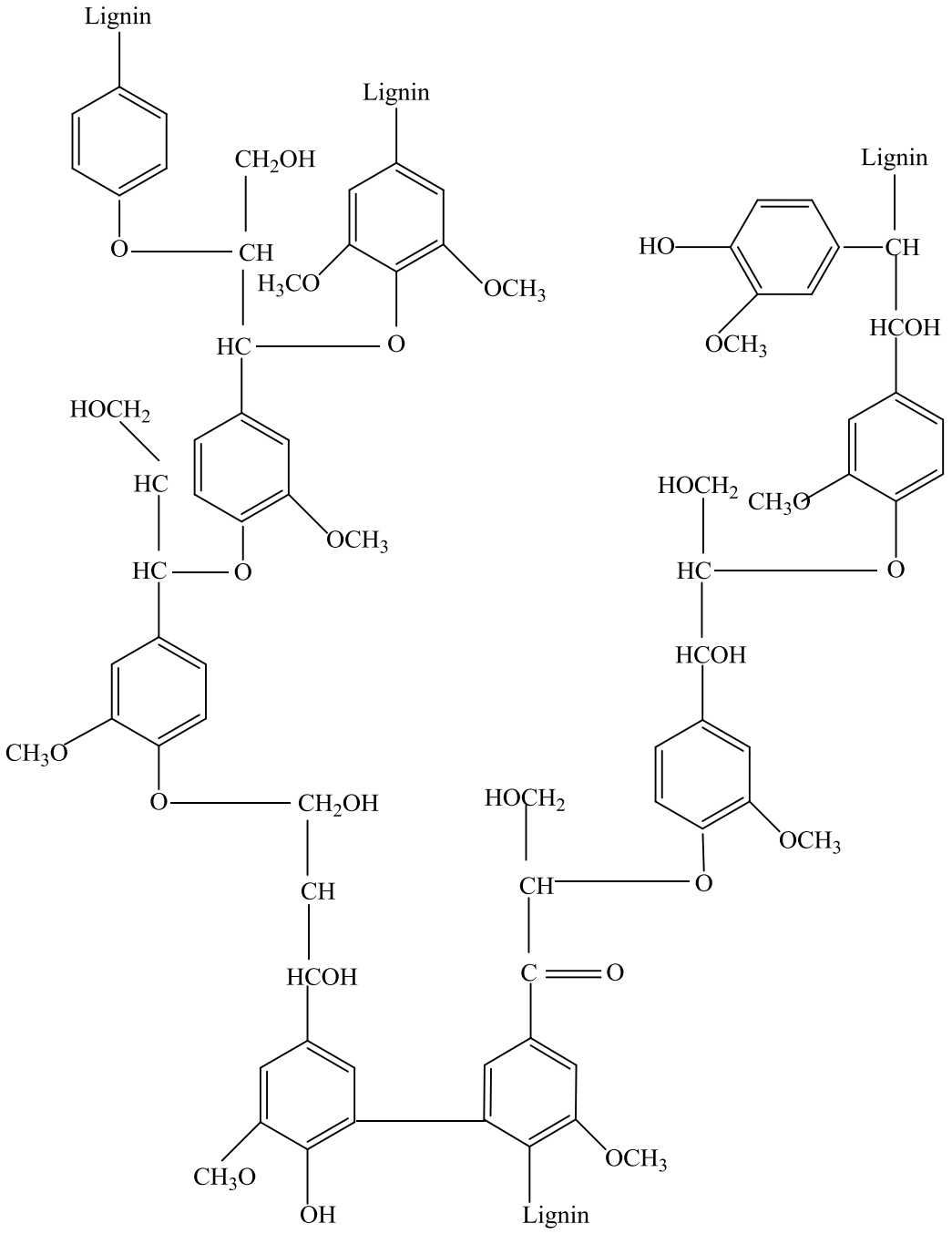


Figure 4. The structure of lignin model monomers sinapic (black), ferulic (red), and coumaric acids (blue).

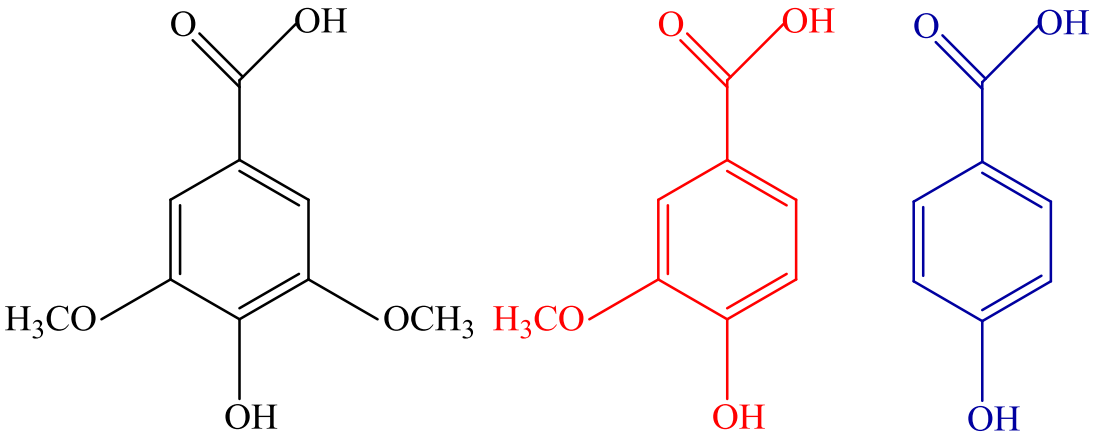


Figure 5. Energy diagrams of Stokes, Rayleigh, and Anti-Stokes Raman scattering. The photons from the excitation source collide with the molecule, and the molecule is promoted to a short-lived virtual state, whereby it immediately emits back to the lowest vibrational state with identical frequency to the incident light (Rayleigh), has a net increase in energy (Stokes), or has lost energy (Anti-Stokes).

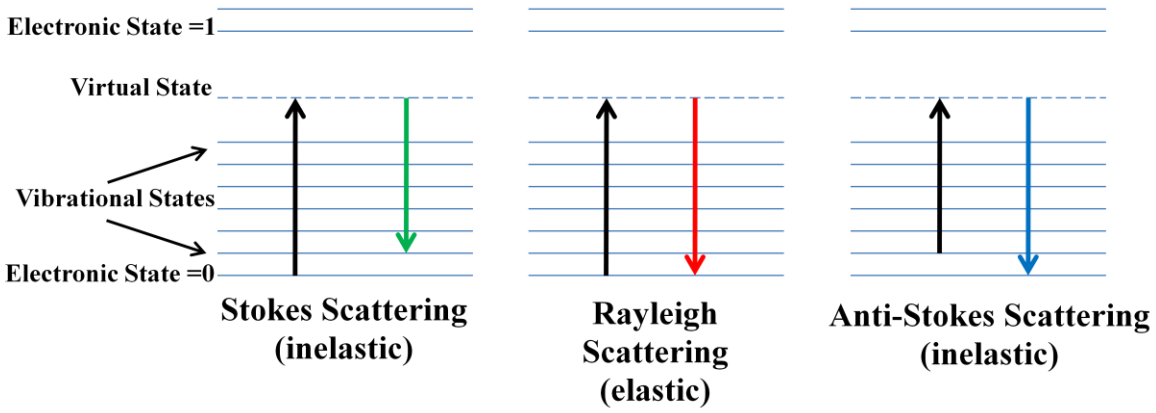


Figure 6. Schematic depicting the physisorption of cellulase to silica nanoparticles.

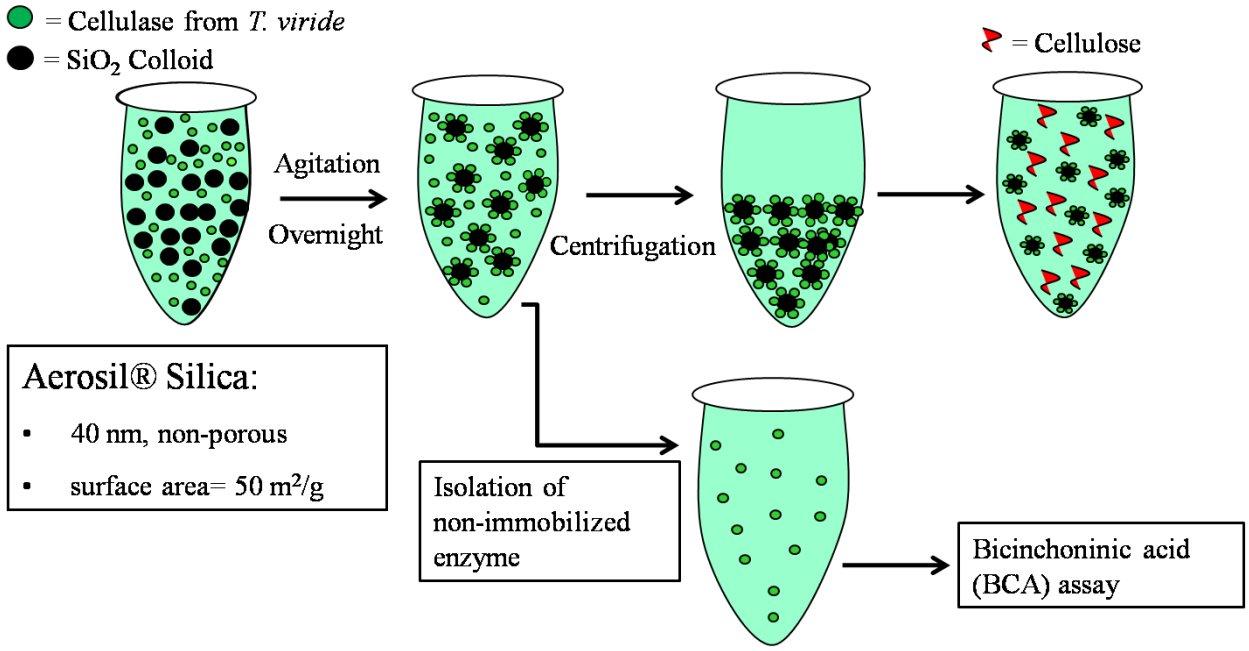
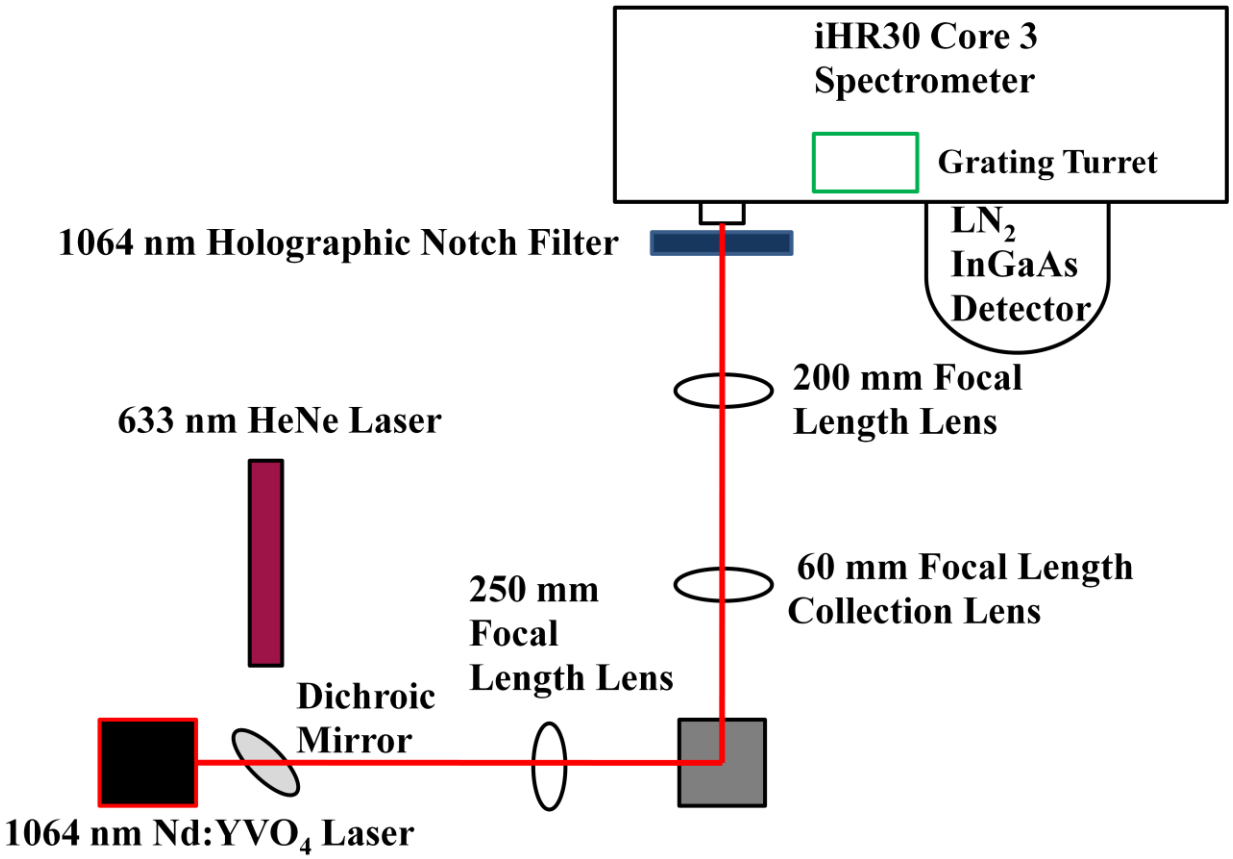


Figure 7. Instrument schematic of home-built 1064 nm dispersive, multichannel Raman spectrometer.



REFERENCES

1. R. D. Perlack, Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J., Erbach, D.C., "Biomass as a Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply, DOE/GO-102005-2135", (Oak Ridge National Laboratory, 2005).
2. A. Karp and I. Shield, The New phytologist 179, 1, 15 (2008).
3. Y. H. P. Zhang, M. E. Himmel, and J. R. Mielenz, Biotechnol. Adv. 24, 5, 452 (2006).
4. L. Cao, "Carrier-Bound Immobilized Enzymes", (Wiley-VCH Verlag, Weinheim, 2005), p. 53.
5. I. Chibata, Immobilized Enzymes. Research and Development (Kodansha LTD., Tokyo, 1978).
6. L. P. Ramos and J. N. Saddler, Applied Biochemistry and Biotechnology 45-46, 193 (1994).
7. M. Tu, X. Zhang, M. Paice, P. MacFarlane, and J. N. Saddler, Bioresource Technology 100, 24, 6407 (2009).
8. S. L. Hirsh, M. M. M. Bilek, N. J. Nosworthy, A. Kondyurin, C. G. dos Remedios, and D. R. McKenzie, Langmuir 26, 17, 14380 (2010).
9. C. A. Haynes and W. Norde, Colloids Surf., B 2, 6, 517 (1994).
10. W. Norde, Colloids and Interfaces in Life Sciences (Marcel Dekker, Inc., New York, 2003).
11. I. R. M. Tebeka, A. G. L. Silva, and D. F. S. Petri, Langmuir 25, 3, 1582 (2009).
12. G. Beldman, M. F. Searle-Van Leeuwen, F. M. Rombouts, and F. G. J. Voragen, European Journal of Biochemistry 146, 2, 301 (1985).
13. A. Goyal, B. Ghosh, and D. Eveleigh, Bioresource Technology 36, 1, 37 (1991).
14. Y.-H. P. Zhang and L. R. Lynd, Biotechnology and Bioengineering 88, 7, 797 (2004).
15. M. P. Coughlan, Biotechnology & Genetic Engineering Reviews 3, 39 (1985).
16. J. Woodward and A. Wiseman, Enzyme and Microbial Technology 4, 2, 73 (1982).
17. M. Holtzapple, M. Cognata, Y. Shu, and C. Hendrickson, Biotechnology and Bioengineering 36, 3, 275 (1990).

18. G. P. Philippidis, T. K. Smith, and C. E. Wyman, Biotechnology and Bioengineering 41, 9, 846 (1993).
19. K. Olofsson, M. Bertilsson, and G. Liden, Biotechnology for biofuels 1, 1, 7 (2008).
20. G. P. Philippidis, Handbook on Bioethanol, 253 (1996).
21. S. H. Krishna, T. J. Reddy, and G. V. Chowdary, Bioresource Technology 77, 2, 193 (2001).
22. J. Szczodrak and Z. Targonski, Biotechnology and Bioengineering 31, 4, 300 (1988).
23. K. Hoyer, M. Galbe, and G. Zacchi, Journal of Chemical Technology and Biotechnology 84, 4, 570 (2009).
24. H. Li, N.-J. Kim, M. Jiang, J. W. Kang, and H. N. Chang, Bioresource Technology 100, 13, 3245 (2009).
25. C.-J. Shih and E. A. Smith, Anal. Chim. Acta 653, 2, 200 (2009).
26. D. W. Templeton, "Determination of Ethanol Concentration in Biomass to Ethanol Fermentation Supernatants by Gas Chromatography, Laboratory Analytical Procedure #011", (U.S. D.O.E. National Renewable Energy Laboratory, 1994).
27. M. Blanco, A. C. Peinado, and J. Mas, Biotechnology and Bioengineering 88, 4, 536 (2004).
28. R. W. Bailey, "Structural Carbohydrates". In: G.W. Butler, R.W. Bailey, editors. Chemistry and Biochemistry of Herbage (London, UK: Academic Press, 1973. Pp. 157-211., 1973).
29. B. L. Browning, "Chemical Composition of Wood". In: B.L. Browning, editor. The Chemistry of Wood. (New York, NY: Interscience Publishers (A Division of John Wiley and Sons), 1963. Pp. 58-101.
30. R. C. Pettersen, Advances in Chemistry Series 207, Chem. Solid Wood, 57 (1984).
31. B. P. U.S Department of Energy, "Biomass Feedstock Composition and Property Database".
32. J. D. McMillan, ACS Symposium Series 566, Enzymatic Conversion of Biomass for Fuels Production, 292 (1994).
33. E. Sjostrom, "Wood Polysaccharides". In: Wood Chemistry: Fundamentals and Applications (New York, NY: Academic Press, 1981. Pp. 49-67., 1981).

34. B. L. Browning, "Wood Lignins". In: B.L. Browning, editor. The Chemistry of Wood. (New York, NY: Interscience Publishers (A Division of John Wiley and Sons), 1963. Pp. 249-311., 1963).
35. K. V. Sarkanen, Hergert, H.L., "Classification and Distribution". In: K.V. Sarkanen, C.H. Ludwig, editors. Lignins: Occurrence and Formation, Structure, Chemical and Macromolecular Properties, and Utilization. (New York, NY: John Wiley and Sons, 1971. Pp. 43-94., 1971).
36. T. Higuchi, "Biosynthesis of Lignin". In: T. Higuchi, editor Biosynthesis and Biodegradation of Wood Components. (Orlando, FL: Academic Press, Inc., 1985, Pp: 141-160., 1985).
37. M. A. Buchanan, "Extraneous Compounds of Wood". In: B.L. Browning, The Chemistry of Wood. (New York, NY: Interscience Publishers (A Division of John Wiley and Sons), 1963. Pp. 314-368., 1963).
38. H. Schulz and M. Baranska, Vibrational Spectroscopy 43, 1, 13 (2007).
39. E. Sjostrom, "Extractives". In: Wood Chemistry. (New York, NY: Academic Press, 1981. Pp.83-97., 1981).
40. K. Thammasouk, D. Tandjo, and M. H. Penner, Journal of Agricultural and Food Chemistry 45, 2, 437 (1997).
41. B. Kamm and M. Kamm, Applied Microbiology and Biotechnology 64, 2, 137 (2004).
42. R. Kumar and C. E. Wyman, Biotechnology and Bioengineering 103, 2, 252 (2009).
43. S. Nakagame, R. P. Chandra, and J. N. Saddler, Biotechnology and Bioengineering 105, 5, 871 (2009).
44. T. B. Vinzant, C. I. Ehrman, W. S. Adney, S. R. Thomas, and M. E. Himmel, Applied Biochemistry and Biotechnology 62, 1, 99 (1997).
45. Y.-H. P. Zhang, S.-Y. Ding, J. R. Mielenz, J.-B. Cui, R. T. Elander, M. Laser, M. E. Himmel, J. R. McMillan, and L. R. Lynd, Biotechnology and Bioengineering 97, 2, 214 (2007).
46. J. M. Pepper, P. E. T. Baylis, and E. Adler, Canadian Journal of Chemistry 37, 1241 (1959).

47. C. Lapierre, B. Monties, and C. Rolando, Journal of Wood Chemistry and Technology 5, 2, 277 (1985).
48. C.-L. Chen, "Nitrobenzene and Cupric Oxide Oxidations". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. (Berlin, Germany: Springer-Verlag, 1992. Pp 301-320, 1992).
49. F. F. Lopes, F. O. Silverio, D. C. F. Baffa, M. E. Loureiro, and M. H. P. Barbosa, Journal of Wood Chemistry and Technology 31, 4, 309 (2011).
50. T. Sonoda, T. Ona, H. Yokoi, Y. Ishida, H. Ohtani, and S. Tsuge, Analytical Chemistry 73, 22, 5429 (2001).
51. C. Rolando, Monties, B., LaPierre, C, "Thioacidolysis". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. (Berlin, Germany: Springer-Verlag, 1992, Pp.334-349., 1992).
52. C. Lapierre, "Determining Lignin Structure by Chemical Degradations". In: C. Heitner, D.R. Dimmel, J.A. Schmidt, editors. Lignin and Lignans: Advances in Chemistry (Boca Raton, FL: CRC Press, 2010. Pp. 11-48., 2010).
53. G. Landsberg and L. Mandelstam, Naturwissenschaften 16, 557 (1928).
54. C. V. Raman and K. S. Krishnan, Nature (London, United Kingdom) 121, 501 (1928).
55. A. Smekal, Naturwissenschaften 11, 873 (1923).
56. P. R. Carey, Molecular Biology: Biochemical Applications of Raman and Resonance Raman Spectroscopies (New York, NY: Academic Press, 1982. Pp. 1-47., 1982).
57. E. Smith, Modern Raman Spectroscopy (Chichester, UK: John Wiley and Sons, 2005. Pp. 1-30, 71-92., 2004).
58. N. Q. Dao, "Dispersive Raman Spectroscopy, Current Instrumental Designs". In: R.A. Meyers, editor. The Encyclopedia of Analytical Chemistry, Volume 15. (Chichester, UK: John Wiley and Sons, 2000. Pp. 13024-13058., 2000).
59. U. P. Agarwal, Advances in Lignocellulosics Characterization, 201 (1999).
60. U. P. Agarwal and S. A. Ralph, Applied Spectroscopy 51, 11, 1648 (1997).

61. R. H. Atalla, Agarwal, U.P., Bond, J.S., "Raman Spectroscopy". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. (Berlin, Germany: Springer-Verlag, 1992, Pp. 162-176).
62. M. W. Meyer, J. S. Lupoi, and E. A. Smith, Analytica Chimica Acta 706, 1, 164 (2011).
63. D. Stewart, H. M. Wilson, P. J. Hendra, and I. M. Morrison, Journal of Agricultural and Food Chemistry 43, 8, 2219 (1995).
64. M. H. Nuopponen, H. I. Wikberg, G. M. Birch, A.-S. Jaaskelainen, S. L. Maunu, T. Vuorinen, and D. Stewart, Journal of Applied Polymer Science 102, 1, 810 (2006).
65. A.-M. Saariaho, A.-S. Jaaskelainen, M. Nuopponen, and T. Vuorinen, Applied Spectroscopy 57, 1, 58 (2003).
66. C.-J. Shih, J. S. Lupoi, and E. A. Smith, Bioresource Technology 102, 8, 5169 (2011).
67. T. Ona, T. Sonoda, K. Ito, M. Shibata, T. Katayama, T. Kato, and Y. Ootake, Journal of Wood Chemistry and Technology 18, 1, 43 (1998).
68. T. Ona, T. Sonoda, K. Ito, M. Shibata, T. Kato, and Y. Ootake, Journal of Wood Chemistry and Technology 18, 1, 27 (1998).
69. T. Ona, T. Sonoda, K. Ito, M. Shibata, Y. Ootake, J. Ohshima, S. Yokota, and N. Yoshizawa, Applied Spectroscopy 53, 9, 1078 (1999).
70. T. Ona, T. Sonoda, K. Ito, M. Shibata, Y. Ootake, J. Ohshima, S. Yokota, and N. Yoshizawa, Analyst (Cambridge, United Kingdom) 124, 10, 1477 (1999).
71. T. Ona, T. Sonoda, J. Ohshima, S. Yokota, and N. Yoshizawa, Journal of Pulp and Paper Science 29, 1, 6 (2003).
72. A.-M. Saariaho, D. S. Argyropoulos, A.-S. Jaeaeskelaeinen, and T. Vuorinen, Vibrational Spectroscopy 37, 1, 111 (2005).
73. B. K. Lavine, C. E. Davidson, A. J. Moores, and P. R. Griffiths, Applied Spectroscopy 55, 8, 960 (2001).

CHAPTER 2.
EVALUATION OF NANOPARTICLE-IMMOBILIZED CELLULASE FOR IMPROVED ETHANOL YIELD IN SIMULTANEOUS SACCHARIFICATION AND FERMENTATION REACTIONS

(A paper published in *Biotechnology and Bioengineering*, (108), 2011, 2835-2843)
Jason S. Lupoi and Emily A. Smith

Abstract

Ethanol yields were 2.1 (p=0.06) to 2.3 (p=0.01) times higher in simultaneous saccharification and fermentation (SSF) reactions of microcrystalline cellulose when cellulase was physisorbed on silica nanoparticles compared to enzyme in solution. In SSF reactions, cellulose is hydrolyzed to glucose by cellulase while yeast simultaneously ferments glucose to ethanol. The 35°C temperature and the presence of ethanol in SSF reactions are not optimal conditions for cellulase. Immobilization onto solid supports can stabilize the enzyme and promote activity at non-optimum reaction conditions. Mock SSF reactions that did not contain yeast were used to measure saccharification products and identify the mechanism for the improved ethanol yield using immobilized cellulase. Cellulase adsorbed to 40 nm silica nanoparticles produced 1.6 times (p=0.01) more glucose than cellulase in solution in 96 hours at pH 4.8 and 35°C. There was no significant concentration (<250 µg) of soluble cellooligomers in either the solution or immobilized enzyme reactions. This suggests the mechanism for the immobilized enzyme's improved glucose yield compared to solution enzyme is the increased conversion of insoluble cellulose hydrolysis products to soluble cellooligomers at 35°C and in the presence of ethanol. The results show that silica-immobilized cellulase can be used to produce increased ethanol yields in the conversion of lignocellulosic materials by SSF.

Keywords: Enzyme Immobilization; Biofuels; Enzymatic Hydrolysis; Cellulase; Cellulose; Simultaneous Saccharification and Fermentation

1. Introduction

The production of portable liquid fuel from biomass can be achieved by the enzymatic hydrolysis of cellulose to glucose, and the subsequent or simultaneous fermentation of glucose to ethanol. Cellulase is a mixture of enzymes with endoglucanase, exoglucanase, and β-glucosidase activity, which catalyzes the multi-step hydrolysis of cellulose to glucose. Since enzyme cost is a significant portion of the total expense in obtaining ethanol from cellulose, efforts have focused on improving enzyme recyclability and efficiency (Cao 2005; Tebeka et al. 2009; Zhang et al. 2006). Despite recent advances in solution enzyme recyclability (Tu et al. 2009; Zhu et al. 2009), reuse of cellulase is still hindered by the intensive methods required to remove the enzyme from the reaction liquor.

The pioneering work on enzyme immobilization in the 1980s has demonstrated that attaching enzymes onto solid supports facilitates catalyst removal from the reaction liquor, enabling more convenient enzyme recycling (Jain and Wilkins 1987; Jones and Vasudevan 2010; Tebeka et al. 2009; Woodward et al. 1984). Tebeka et al. (2009) demonstrated this recyclability, reporting consistent catalytic efficiency after six hydrolysis cycles when using cellulase from *Trichoderma reesei* physisorbed or covalently attached onto silicon wafers.

In addition to facilitating recyclability, immobilizing enzymes to solid supports increases protein stability at non-optimal reaction conditions, including temperature, solvent composition, and pH. Immobilization of an endoglucanase from the hyperthermophilic *Pyrococcus furiosus* was found to stabilize the enzyme against denaturation with increasing reaction temperature (Koutsopoulos et al. 2004). Gole and co-authors (2001) reported increased enzymatic stability against shifts in pH and temperature for an immobilized fungal endoglucanase. Enzyme immobilization has been postulated to provide protection from protein degradation and aggregation (Gole et al. 2001). Furthermore, the immobilization of multiple enzymes on the same support can favorably position co-enzymes in close proximity to improve enzymatic efficiency (El-Zahab et al. 2004).

Cellulase and its individual constituents have previously been immobilized on a variety of solid supports. Calsavara et al. (2001) covalently attached β-glucosidase to controlled pore glass and found a shift in the optimal reaction pH by one unit and thermal stability at higher temperatures for the immobilized compared to solution enzyme. The adsorption of *Trichoderma reesei* cellobiohydrolase I on silanized silica was investigated by Suvajittanont et al. (2000a). They found that enzyme adsorbed to hydrophobic silica was partially desorbed by rinsing with buffer. Endoglucanase from *Thermomonospora fusca* or *T. reesei* immobilized to polystyrene showed a 20% decrease in enzyme activity when the adsorption time was increased from 30 minutes to 24 hours (Kongruang et al. 2003). A comparison between crude cellulase physisorbed or chemisorbed to polystyrene showed that plasma immersion ion implantation produced a distributed layer of enzyme that retained its native conformation; whereas, physisorbed cellulase unfolded and tended to form aggregates on the untreated polystyrene (Hirsh et al. 2010).

Materials that have been used as solid supports for enzyme immobilization include: polystyrene, Eupergit®, Sepharose, nylon, iron oxide, and others (Calsavara et al. 2001; Chim-Anage et al. 1986; Hirsh et al. 2010; Ho et al. 2008; Jain and Wilkins 1987; Karagulyan et al. 2008; Kongruang et al. 2003; Sundstrom et al. 1981; Suvajittanont et al. 2000a; Suvajittanont et al. 2000b; Tebeka et al. 2009; Woodward et al. 1984). It has been shown that the support used to immobilize the enzyme, as well as the immobilization method, will affect the function of the protein (Cao 2005; Engasser and Horvath 1976; Haynes and Norde 1994; Norde 2003).

In contrast to separate hydrolysis and fermentation reactions, simultaneous saccharification and fermentation (SSF) can result in higher ethanol yields, lower product inhibition of the enzyme, and lower production costs (Wyman and Editor 1996). A disadvantage of SSF reactions is the incompatibility of optimal reaction conditions for the hydrolysis enzymes and the yeast used for fermentation. The optimum pH and reaction temperature for *Trichoderma viride* cellulase is reported to be 4.5-5.2 (White et al. 1997) and 50°C, respectively. Fermentation temperatures typically range from 28-38°C for ideal yeast

activity (Szczodrak and Fiedurek 1996; Wyman and Editor 1996). Additionally, the ethanol produced during the fermentation can denature cellulase.

Herein, the physisorption of a commercial cellulase from *T. viride* onto non-porous, silica nanoparticles is reported. Ethanol yields for SSF and glucose yields for hydrolysis reactions have been determined for solution and silica-immobilized cellulase. In order to determine the mechanism for increased product yields, reaction products were measured at optimal and non-optimal solution enzyme conditions. Optimal hydrolysis conditions were altered by (i) adding ethanol into the reaction matrix to simulate SSF conditions, (ii) changing the hydrolysis buffer pH, and (iii) varying the hydrolysis temperature. To our knowledge, this is the first report of improved ethanol yields using immobilized cellulase for SSF reactions.

2. Materials and Methods

2.1 Materials

Cellulase from *T. viride*, crude powder (Sigma-Aldrich, St. Louis, MO, USA), microcrystalline cellulose (Sigma-Aldrich), hydroxyethyl cellulose (HEC) (Sigma-Aldrich), Aerosil OX-50 (Evonik, Essen, Germany), glucose-hexokinase (HK) assay kit (Sigma-Aldrich), bicinchoninic acid (BCA) assay kit (Pierce, Rockford, IL, USA), sulfuric acid, p-anisaldehyde (Acros Organics, Geel, Belgium), yeast extract (U.S. Biological, Swampscott, MA, USA), peptone (U.S. Biological), ethanol TT yeast (SPL International, Cheshire, United Kingdom), and ethanol were used as received. All other chemicals were from Fisher Scientific (Pittsburgh, PA, USA). Deionized water was from an ultrapure water system (Barnstead International, Dubuque, IA, USA) with a resistivity of 18.2 MΩ.

2.2 Cellulase Adsorption Isotherm

The enzyme chosen for this study was available purified and in large quantities, and had a lower activity than some enzymes that have been reported in the literature (Shih and Smith 2009). Cellulase dissolved in 500 µL 10 mM acetate buffer, pH 4.8 with 0.01% (w/v) sodium azide was added to 28.0 ± 0.5 mg Aerosil OX-50 silica (40 nm average diameter particles, total surface area 1.4 m^2). Initial cellulase concentrations ranged from 0 to 9.5

mg/mL. Samples were incubated overnight at ambient conditions with continuous agitation. Following incubation, the samples were centrifuged at 12,400 g for 10 minutes and the supernatant was collected. A 250 µL aliquot of acetate buffer was added to each tube to rinse the cellulase-coated silica of residual non-immobilized enzyme, and the sample was centrifuged at 12,400 g for 2 minutes. The rinse was combined with the original supernatant and then re-centrifuged at 16,900 g for 90 minutes to remove residual silica, which interfered with subsequent BCA assays (Smith et al. 1985) of non-immobilized enzyme concentration. Immobilized enzyme concentrations were calculated by subtracting enzyme in the supernatant from the initial enzyme concentration. The adsorption isotherm was best-fit to a BET isotherm using user-defined curve fitting functions in IGOR Pro 6.1 (WaveMetrics Inc., Portland, OR, USA). The BET isotherm equation is presented in the Supplemental Information.

2.3 Batch Immobilization of Cellulase on Aerosil OX-50

Silica (5.6 g) was added to 600 mg cellulase in 100 mL 10 mM acetate buffer, pH 4.8 in a polypropylene bottle. The mixture was incubated overnight at ambient conditions with continuous agitation. The protein-coated silica was collected by centrifugation and then incubated for 48 hours in medium of the same composition used for subsequent experiments. This rinse step was required to ensure the amount of protein that desorbed from the silica during the hydrolysis reaction was minimal for most reaction conditions studied. The amount of immobilized cellulase on the nanoparticles was quantified using BCA assays as described above.

2.4 Simultaneous Saccharification and Fermentation (SSF) Reactions

The SSF reactions were performed according to the National Renewable Energy Lab protocol NREL/TP-510-42630 with some modifications. All buffers, cellulose, glassware, and tubing were autoclaved prior to use. The solution or immobilized enzyme was added to 600 mg cellulose, 1.0 mL of 10× YP (10.0 g yeast extract and 20.0 g peptone in 100 mL deionized water), 0.02 g of ethanol TT yeast, and 50 mM citrate buffer, pH 4.8 to yield a final volume of 10 mL. Tygon™ tubing connected the reaction vial to a flask of glycerol

through which carbon dioxide escaped. The entire assembly was then placed in an incubated shaker set at 35 °C, 150 rpm for 96 h. All samples were centrifuged to remove solids prior to ethanol quantitation by headspace gas chromatography/mass spectrometry (Shih and Smith 2009). Two-way analysis of variance (ANOVA) was used to compare data sets using JMP 8.0 (SAS). Control reactions were performed without (i) enzyme, (ii) cellulose or (iii) silica to ensure the reaction matrix did not interfere with ethanol quantitation.

2.5 Enzymatic Hydrolysis of Cellulose

All reactions were performed in glass vials previously coated with 1 mg/mL bovine serum albumin and rinsed with buffer to prevent non-specific cellulase binding to the glass. The hydrolysis samples contained 5 mL 10 mM acetate buffer, pH 4.8, 4.1 ± 0.3 mg silica-immobilized (equivalent to ~140 mg dry silica) *or* solution cellulase, 40.1 ± 0.1 mg of microcrystalline cellulose, and were incubated at a specified temperature for a duration indicated in the text. The activity of the commercial enzyme used in this study varied between lots and with enzyme age. All data in each figure were collected using enzyme of the same age and lot.

For the pH studies, the reactions were performed as outlined above, except 5 mL of 10 mM citrate buffer, ranging from pH 3.3 to 6.3 in intervals of 0.5, was used in place of the acetate buffer. Similarly, for the temperature analyses, the reaction temperature was varied from 25 to 60°C in 5°C intervals. To simulate SSF conditions, ethanol (0, 8, 32, 78, 118, or 158 mg/mL) was added to 10 mM acetate buffer, pH 4.8 for a total volume of 5 mL. Cellulase recyclability at 35°C was measured by adding fresh substrate equivalent to the amount of digested cellulose (calculated using the ratio 1 g cellulose/1.1 g glucose) to the reaction and continuing the hydrolysis. Where noted in the text, hydrolysis reactions containing 0, 8 and 32 mg/mL ethanol were performed with 40.1 ± 0.1 mg hydroxyethyl cellulose, which is soluble in aqueous buffer, in place of insoluble cellulose. All reactions were performed in triplicate.

The amount of enzyme that desorbed from the silica during the hydrolysis reaction was quantified by BCA assays of solutions that were identical to the hydrolysis reactions, except no cellulose was added. The amount of glucose that could be produced by the

desorbed enzyme was measured in separate hydrolysis reactions. Further details of desorbed protein quantification and hydrolysis yield are described in the Supplemental Information.

2.6 Analysis of Hydrolysis Reaction Products

Glucose produced from the hydrolysis reaction was quantified using the glucose-HK assay (Bondar and Mead 1974). Absorbance values were measured at 340 nm using an Agilent 8453 spectrophotometer. A calibration curve was constructed using glucose standard solutions (Supporting Information Figure S1).

The heat-inactivated (75°C) hydrolysis liquors were analyzed for glucose, cellobiose, cellotriose, cellotetraose, and cellopentaose using thin layer chromatography (TLC, see Supplemental Information) and high performance liquid chromatography (HPLC). Hydrolysis samples (20 µL injection volume) were analyzed for cellooligosaccharides using a Varian HPLC with a refractive index (Prostar 355 RI, Palo Alto, CA, USA) detector. Saccharides were eluted using a 0.6 mL/min flow rate of water mobile phase through a HPX-87P Bio-Rad Aminex column heated to 75°C (Bio-Rad, Hercules, CA). The refractive index detector was heated to 50°C. The chromatograms were fit to an exponentially modified Gaussian function using IGOR Pro 6.1 (WaveMetrics Inc.) to obtain peak areas. Calibration curves were constructed using the peak areas from standard solutions (Supplemental Information Figure S2). Detection limits were calculated as 3 times the standard deviation in the peak area for samples with concentrations near the detection limit divided by the slope of the best-fit line to the calibration data.

3. Results and Discussion

3.1 Characterization of Cellulase Adsorption to Silica Nanoparticles

The purpose of this study was to evaluate the ethanol yields for SSF reactions using silica-immobilized cellulase, and compare the yields to those for cellulase in solution under identical conditions. Increased enzyme stability upon immobilization to a solid support should result in increased yields for SSF reactions where non-optimal solution enzyme conditions are required. The chosen solid support for the enzyme immobilization was non-porous silica nanoparticles. Silica was selected due to its well-characterized and utilized

surface chemistry. Nanoparticles were used for enzyme immobilization because they have large surface areas compared to planar substrates. Non-porous particles were chosen, despite lower surface areas than porous materials, to ensure that issues involving the transport of insoluble cellulose to the pore interior were not a factor in comparing solution versus immobilized cellulase performance. Physisorption was chosen as the adsorption mechanism for simplicity in catalyst preparation, and the low cost advantage over covalent attachment techniques. Physisorption is typically performed in mild conditions without chemically modifying the enzyme (Cao 2005; Chibata 1978; Karagulyan et al. 2008).

An adsorption isotherm was constructed to measure the amount of silica surface coated by the enzyme and to ensure that the conditions used in this study produced less than a monolayer of enzyme (Figure 1). The data were best-fit to a B.E.T. type II isotherm model (chi-squared=0.055), which indicates the possibility of forming protein multi-layers (Brunauer 1943). For enzyme immobilization onto silica nanoparticles, a solution enzyme concentration of 6.4 ± 1.6 mg/mL produces a monolayer of cellulase on the silica (1.2 ± 0.3 mg/m^2). Above this concentration, multi-layers of enzyme on the silica will result in buried enzyme that is not capable of participating in the hydrolysis reaction. For all subsequent experiments, immobilized enzyme surface coverage after performing the final rinse step was approximately 0.6 ± 0.2 mg/m^2. The exact surface coverage was measured for every reaction reported herein.

3.2 Ethanol and Saccharide Yields in Simultaneous Saccharification and Fermentation Reactions

The ethanol yields detected by GC/MS for the immobilized and solution cellulase SSF reactions are shown in Figure 2. The reactions were performed at 35°C using pH 4.8 or 5.3 medium. Compared to enzyme in solution, the ethanol yield from the SSF reactions were 2.3 (p=0.01) and 2.1 (p=0.06) times higher for the immobilized enzyme at pH 4.8 and 5.3, respectively. HPLC analysis of the SSF liquors' saccharide composition revealed that no glucose (detection limit 150 µg), cellobiose (detection limit 150 µg) or cellotriose (detection limit 200 µg) were measured in either the solution or immobilized enzyme SSF reactions

(data not shown). Cellobiose and cellotriose are soluble intermediates in the hydrolysis of insoluble cellulose to glucose.

The higher ethanol yield for the immobilized enzyme reaction is likely due to higher glucose yields as a result of increased enzymatic stability at the non-optimal enzyme conditions required for SSF. In order to test this hypothesis, mock SSF reactions without added yeast were performed in the same medium used for SSF reactions (10 mM acetate buffer, pH 4.8 and 35°C). Without yeast in the reaction vessel, the saccharides will accumulate and can be measured by HPLC (Figure 3), glucose HK assays (Supplemental Table S1), and TLC (Supplemental Figure S4). Significantly, the immobilized cellulase produced 1.6 times more glucose ($p=0.01$) than the comparable solution enzyme reactions after 96 hours.

A mechanism for the improved glucose yields using immobilized enzyme and these reaction conditions was determined by measuring soluble cellooligomeric species in the reaction matrix (Figure 3). Approximately 250 µg cellobiose and no higher order soluble cellooligomers were detected by HPLC in both the solution and immobilized cellulase mock SSF samples. The data suggest that the improved glucose yield for immobilized cellulase is the result of the enhanced conversion of insoluble cellulose hydrolysis products to soluble cellooligomers. Accumulation of a reaction intermediate would have indicated a limiting step in the cellulose hydrolysis, and soluble species did not accumulate at significant concentrations in either the immobilized or solution enzyme reaction. Cellulase is a mixture of enzymes with synergistic activities. The immobilization of cellulase proteins in close proximity on the silica support may facilitate multiple, simultaneous enzyme-cellulose contacts and the conversion of insoluble cellulose to soluble cellooligomers. Control reactions containing only silica and no enzyme did not produce detectable glucose (Figure 3), indicating that the silica support is not capable of hydrolyzing cellulose in the absence of enzyme. This does not, however, preclude a role for silica-cellulose interactions in the immobilized enzyme reaction. For example, cellulose may interact with the silica surface and favorably position the substrate in proximity to the enzyme.

It is informative to quantitatively compare the glucose yields in the mock SSF (no yeast) reactions to the ethanol yields in the SSF reactions. Glucose yields from the mock

SSF reactions were 23 ± 2 mg and 14 ± 1 mg for the immobilized and solution cellulase (Figure 3), respectively. Using the stoichiometric conversion factor of 1 g glucose converts to 0.51 g ethanol, the expected ethanol yields would be 10.6 ± 0.9 mg using immobilized cellulase and 6.8 ± 0.7 mg using cellulase in solution. The expected ethanol yield was obtained for the immobilized enzyme (10.9 ± 5.5 mg) SSF reaction, but less than the expected yield was obtained for the SSF reaction with solution enzyme (3.8 ± 3.8 mg, Figure 2). The enzyme in solution may have been destabilized by prolonged exposure to ethanol at the non-optimum 35°C reaction temperature, since the main difference between the SSF and mock SSF reaction is the presence of ethanol.

The immobilized enzyme also offers the advantage of recyclability in SSF reactions, which would be impractical using solution enzyme. For the immobilized enzyme, the glucose yield following the first hydrolysis was 40 ± 20% and 17 ± 8% for the second and third hydrolyses, respectively. The decrease in yield for subsequent cycles is through enzyme inactivation mechanisms and not from the desorption of enzyme from the support, as confirmed by the lack of soluble protein in the reaction matrix. This is consistent with previous enzyme recyclability studies that have shown a significant decrease in product yield from the first to second cycle, followed by a tapering off in total yield after subsequent cycles (Jones and Vasudevan 2010; Tebeka et al. 2009).

3.3 Mechanism of Improved Glucose Yield for Immobilized Cellulase

In order to determine if SSF reaction temperature, pH, ethanol content or a combination of factors explain the higher yields for immobilized compared to solution cellulase, and to see if further optimization of SSF product yield may be obtained, the effects of altering each reaction variable were investigated in a series of hydrolysis reactions. Desorption of cellulase from silica during the hydrolysis reaction would influence the comparison between immobilized and solution cellulase efficiency since the "immobilized" enzyme would be a mixture of solution and adsorbed species. For all reaction conditions included in this study, 2.18 ± 0.05 to 25.8 ± 0.2 % of the enzyme desorbs from the silica during the hydrolysis reaction, with the exception of the 30°C hydrolysis (40.9 ± 0.7%). The maximum amount of glucose produced from the cellulase that desorbed from the silica

during the hydrolysis reaction was measured for every reaction as described in the Supplemental Information and is represented by the dark gray portion of the data bars in Figures 4, 5 and 6. The dark gray portion of the data bar is the maximum possible amount contributed by the desorbed enzyme because the measurement assumes that all desorbed enzyme can contribute in the enzymatic reaction and that desorption occurs at the start of the hydrolysis. This assumption may not hold if the desorbed enzyme is denatured during the adsorption and/or desorption process, rendering it enzymatically inactive. The magnitude of the desorbed enzyme's contribution to total product yield is related to: (i) the solution enzyme activity at a given set of conditions; and (ii) the amount of enzyme that desorbs during the reaction.

Figure 4 shows a plot of glucose generated in 48 hours at 50 °C using the same concentration of immobilized and solution cellulase as a function of hydrolysis medium pH. For all pH values, the solution enzyme reactions have higher glucose yields than the immobilized enzyme reactions. This is in contrast to the mock SSF data discussed above, where reaction temperature (35°C) and time (96 hours) are different. At the optimum reaction conditions for the solution enzyme, the glucose yield of the immobilized enzyme is expected to be lower than the glucose yield for the solution enzyme. This is because the optimum conditions for the immobilized enzyme reaction may not coincide with those for the solution enzyme. Also, the same mass of enzyme was used in the solution and immobilized enzyme reactions, but this does not mean that that same amount of enzyme is available to participate in the catalysis. Enzyme immobilization results in a random orientation of the physisorbed enzyme on the silica substrate, and this leads to some inactive enzyme due to buried or inaccessible active sites (Koutsopoulos et al. 2004). There may also be steric constraints on the enzyme's active site by neighboring enzyme, but this should be limited in this data by the use of ~0.5 monolayer of protein, as described above.

Aside from the above mentioned overall lower glucose yields for the immobilized enzyme reactions, the plots of glucose yield versus reaction pH have similar trends for the immobilized and solution enzyme reactions (Figure 4). Both curves have an optimum reaction pH near 5.3 for high glucose yields, and substantial decreases in yield at higher and lower pH values. The optimum reaction pH is close to the 5.1 average of all pI values for the

cellulase proteins (Beldman et al. 1985), and is close to the previously reported range of 4.0 to 5.2 (White et al. 1997). The decrease in enzyme activity at non-optimum pH values is a result of variations in the ionization state of amino acids with a role in maintaining the protein structure or at the active site (Bisswanger 2008). The fact that the immobilized and solution enzyme's optimum pH is the same, within the data points included in this study, suggests that there is no or a small shift in the proteins' pI at the interface, and that plausible pH shifts at the local environment of the charged silica surface do not affect ionization states in structural or active site amino acids. The ratio of glucose yield for solution and immobilized enzyme reactions reveals conditions where immobilization confers stability on the enzyme. This ratio is relatively constant for pH values between 4.3 and 5.8, indicating that the immobilized enzyme is not stabilized relative to the solution enzyme against shifting pH values in the range that may be used for SSF reactions.

The glucose yield, as a function of hydrolysis temperature for immobilized and solution enzyme reactions, is shown in Figure 5. The plot of glucose yield versus reaction temperature for the solution enzyme shows an optimum temperature of 50°C. There was a rapid decrease in solution cellulase glucose yield at temperatures higher than the optimum, and there were small decreases in solution enzyme glucose yields at lower reaction temperatures. There are two competing factors that control enzyme activity as a function of reaction temperature: the rate of enzymatic activity increases at higher temperatures generating more glucose, and the rate of protein inactivation (e.g., denaturation) increases at higher temperatures generating less glucose. For the immobilized enzyme reaction, the same general shape of the glucose yield versus reaction temperature plot is obtained; however, the optimum reaction temperature shifts from approximately 45°C to 30°C. A shift of approximately 5°C to lower temperatures in the activity versus reaction temperature curve has previously been measured for β-galactosidase covalently immobilized onto silica, as well as crude cellulase physisorbed to controlled pore glass (Huang et al. 1996; Rogalski et al. 1985). For the immobilized cellulase reaction there is a concomitant shift in the activation energies of catalysis and enzyme inactivation with temperature compared to the solution enzyme reaction. It is not possible to say at this time if these thermal properties are altered for all or a subset of the cellulase proteins. Immobilization does not prevent heat inactivation

of the cellulase at higher temperatures, but more importantly it does confer higher glucose yields at the optimum temperatures (28-38°C) for the yeast that ferment glucose to ethanol (Szczodrak and Fiedurek 1996; Wyman and Editor 1996). This proves that the increased ethanol yield in the SSF reaction is at least in part due to the immobilized enzyme's higher activity compared to solution enzyme at 35°C. This trend should hold at temperatures down to at least 25°C.

Figure 6 shows the immobilized and solution cellulase glucose yields as a function of increasing ethanol concentration in the reaction liquor at pH 4.8 and 50°C. The ethanol concentrations chosen for these studies represent realistic ethanol yields produced in SSF reactions (Li et al. 2009; Tomas-Pejo et al. 2008). As the ethanol concentration increased, both solution and immobilized cellulase glucose yields increased with a maximum yield at 32 mg/mL ethanol. The increase in glucose yield is higher for the immobilized cellulase reaction (523 ± 125%) than the solution cellulase reaction (223 ± 74%). A 20.4, 40.7, and 9.3% increase in enzyme activity with the addition of up to 9% v/v ethanol has been previously reported for β-glucosidase after a 72 hour incubation at 30, 40, and 50°C, respectively (Chen and Jin 2006).

A possible explanation for the increase in glucose yield at low ethanol concentrations is increased cellulose solubility or enhanced cellulase accessibility to the cellulose. This was tested by performing the hydrolysis reactions using immobilized or solution enzyme and a soluble cellulose substrate: hydroxyethyl cellulose (HEC). When no ethanol was in the hydrolysis liquor, the total glucose yields using HEC were 0.11 ± 0.02 and 0.06 ± 0.02 mg glucose/mg cellulase for soluble and immobilized cellulase, respectively (Figure 6 inset). The glucose yields are considerably lower for HEC compared to microcrystalline cellulose hydrolysis. This could be due to the inability of cellobiohydrolase to catalyze the side chains of HEC (Vrsanska and Biely 1992; Zhang et al. 2006). Cellobiohydrolases comprise roughly 80% of *Trichoderma* cellulase (Goyal et al. 1991; Gritzali and Brown 1979; Nidetzky and Claeyssens 1994; Zhang and Lynd 2004). There was no statistically significant change in glucose yields with 8 and 32 mg/mL ethanol for the immobilized cellulase reactions (Figure 6 inset). With 32 mg/mL ethanol, there was a statistically significant decrease in glucose yield for the solution enzyme reaction. This decrease is likely due to enzyme denaturation by

the ethanol. The contrary results between reactions using soluble (Figure 6 inset) and insoluble (Figure 6) cellulose substrates support the notion that there are competing effects on microcrystalline cellulose solubility/accessibility and enzyme denaturation with increasing ethanol concentrations up to 32 mg/mL. At higher ethanol concentrations, protein denaturation dominates and less glucose is produced. Overall, the smaller decrease in immobilized enzyme glucose yield compared to solution enzyme with increasing ethanol concentration suggests that the immobilization process confers stability against protein inactivation by organic solvent. This in part explains the higher SSF ethanol yields reported above for the immobilized cellulase reactions.

4. Conclusions

Enzyme immobilization can be used to increase ethanol yields in the SSF reaction of cellulose. This is the result of increased glucose yield from enzymatic stabilization upon immobilization to the silica support at non-optimum reaction temperatures and in ethanol. The mechanism is likely higher conversion rates of insoluble cellulose and cellulose hydrolysis products to soluble cellooligomers. Considering the possibility of enzyme recyclability by easy removal of the nanoparticles from the reaction matrix, total ethanol yields are potentially greater than reported herein. Overall this provides strong evidence for an advantage of immobilized enzymes in SSF reactions of biomass. Further improvements may be made with strategies to orient the enzyme's active site toward the substrate, using porous supports with higher surface area and by varying the enzyme surface coverage.

Supplemental Information

Supplemental experimental details, glucose HK, HPLC, and GC/MS calibration curves, and supporting saccharide quantification data are available in a separate document.

Acknowledgements

Work at the Ames Laboratory was supported by the Department of Energy-Basic Energy Sciences under Contract No. DE-AC02-07CH11358. Assistance in the GC/MS measurement of fermentation samples was provided by Steve Veysey (Iowa State University,

Department of Chemistry, Chemical Instrumentation Facility). The authors thank Professor Pohl (Iowa State University) for the use of HPLC instrumentation.

References

Beldman G, Searle-Van Leeuwen MF, Rombouts FM, Voragen FGJ. 1985. The cellulase of Trichoderma viride. Purification, characterization and comparison of all detectable endoglucanases, exoglucanases and beta -glucosidases. Eur. J. Biochem. 146(2):301-8.

Bisswanger H. 2008. Enzyme Kinetics: Principles and Methods; Second, Revised and Updated Edition. Weinheim: Wiley-VCH. 301 p.

Bondar RJL, Mead DC. 1974. Evaluation of glucose 6-phosphate dehydrogenase from Leuconostoc mesenteroides in the hexokinase method for determining glucose in serum. Clin. Chem. (Washington, DC, United States) 20(5):586-90.

Brunauer S. 1943. The Adsorption of Gases and Vapors. Princeton: Princeton University Press.

Calsavara LPV, De Moraes FF, Zanin GM. 2001. Comparison of catalytic properties of free and immobilized cellobiase Novozym 188. Appl. Biochem. Biotechnol. 91-93(Symposium on Biotechnology for Fuels and Chemicals, 2000):615-626.

Cao L. 2005. Carrier-Bound Immobilized Enzymes. Weinheim: Wiley-VCH Verlag. p 53-168.

Chen H, Jin S. 2006. Effect of ethanol and yeast on cellulase activity and hydrolysis of crystalline cellulose. Enzyme Microb. Technol. 39(7):1430-1432.

Chibata I. 1978. Immobilized Enzymes. Research and Development. Tokyo: Kodansha LTD. 384 p.

Chim-Anage P, Kashiwagi Y, Magae Y, Ohta T, Sasaki T. 1986. Properties of cellulase immobilized on agarose gel with spacer. Biotechnol. Bioeng. 28(12):1876-8.

El-Zahab B, Jia H, Wang P. 2004. Enabling multienzyme biocatalysis using nanoporous materials. Biotechnol. Bioeng. 87(2):178-183.

Engasser JM, Horvath C. 1976. Diffusion and kinetics with immobilized enzymes. Appl. Biochem. Bioeng. 1(Immobilized Enzyme Princ.):127-220.

Gole A, Vyas S, Sainkar SR, Lachke A, Sastry M. 2001. Enhanced Temperature and pH Stability of Fatty Amine-Endoglucanase Composites: Fabrication, Substrate Protection, and Biological Activity. Langmuir 17(19):5964-5970.

Goyal A, Ghosh B, Eveleigh D. 1991. Characteristics of fungal cellulases. Bioresour. Technol. 36(1):37-50.

Gritzali M, Brown RD, Jr. 1979. The cellulase system of Trichoderma. Relationships between purified extracellular enzymes from induced or cellulose-grown cells. Adv. Chem. Ser. 181(Hydrolysis Cellul.: Mech. Enzym. Acid Catal.):237-60.

Haynes CA, Norde W. 1994. Globular proteins at solid/liquid interfaces. Colloids Surf., B 2(6):517-66.

Hirsh SL, Bilek MMM, Nosworthy NJ, Kondyurin A, dos Remedios CG, McKenzie DR. 2010. A Comparison of Covalent Immobilization and Physical Adsorption of a Cellulase Enzyme Mixture. Langmuir 26(17):14380-14388.

Ho KM, Mao X, Gu L, Li P. 2008. Facile Route to Enzyme Immobilization: Core-Shell Nanoenzyme Particles Consisting of Well-Defined Poly(methyl methacrylate) Cores and Cellulase Shells. Langmuir 24(19):11036-11042.

Huang XL, Walsh MK, Swaisgood HE. 1996. Simultaneous isolation and immobilization of streptavidin-beta -galactosidase: some kinetic characteristics of the immobilized enzyme and regeneration of bioreactors. Enzyme Microb. Technol. 19(5):378-383.

Jain P, Wilkins ES. 1987. Cellulase immobilized on modified nylon for saccharification of cellulose. Biotechnol. Bioeng. 30(9):1057-62.

Jones PO, Vasudevan PT. 2010. Cellulose hydrolysis by immobilized Trichoderma reesei cellulase. Biotechnol. Lett. 32(1):103-106.

Karagulyan HK, Gasparyan VK, Decker SR. 2008. Immobilization of fungal beta -glucosidase on silica gel and kaolin carriers. Appl. Biochem. Biotechnol. 146(1-3):39-47.

Kongruang S, Bothwell MK, McGuire J, Zhou M, Haugland RP. 2003. Assaying the activities of Thermomonospora fusca E5 and Trichoderma reesei CBHI cellulase bound to polystyrene. Enzyme Microb. Technol. 32(5):539-545.

Koutsopoulos S, van der Oost J, Norde W. 2004. Adsorption of an Endoglucanase from the Hyperthermophilic Pyrococcus furiosus on Hydrophobic (Polystyrene) and Hydrophilic (Silica) Surfaces Increases Protein Heat Stability. Langmuir 20(15):6401-6406.

Li H, Kim N-J, Jiang M, Kang JW, Chang HN. 2009. Simultaneous saccharification and fermentation of lignocellulosic residues pretreated with phosphoric acid-acetone for bioethanol production. Bioresour. Technol. 100(13):3245-3251.

Nidetzky B, Claeyssens M. 1994. Specific quantification of Trichoderma reesei cellulases in reconstituted mixtures and its application to cellulase-cellulose binding studies. Biotechnol. Bioeng. 44(8):961-6.

Norde W. 2003. Colloids and Interfaces in Life Sciences. New York: Marcel Dekker, Inc. 433 pp.

Rogalski J, Szczodrak J, Dawidowicz A, Ilczuk Z, Leonowicz A. 1985. Immobilization of cellulase and D-xylanase complexes from Aspergillus terreus F-413 on controlled porosity glasses. Enzyme Microb. Technol. 7(8):395-400.

Shih C-J, Smith EA. 2009. Determination of glucose and ethanol after enzymatic hydrolysis and fermentation of biomass using Raman spectroscopy. Anal. Chim. Acta 653(2):200-206.

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. 150(1):76-85.

Sundstrom DW, Klei HE, Coughlin RW, Biederman GJ, Brouwer CA. 1981. Enzymic hydrolysis of cellulose to glucose using immobilized beta -glucosidase. Biotechnol. Bioeng. 23(3):473-85.

Suvajittanont W, Bothwell MK, McGuire J. 2000a. Adsorption of Trichoderma reesei CBHI cellulase on silanized silica. Biotechnol. Bioeng. 69(6):688-692.

Suvajittanont W, McGuire J, Bothwell MK. 2000b. Adsorption of Thermomonospora fusca E5 cellulase on silanized silica. Biotechnol. Bioeng. 67(1):12-18.

Szczodrak J, Fiedurek J. 1996. Technology for conversion of lignocellulosic biomass to ethanol. Biomass Bioenergy 10(5/6):367-375.

Tebeka IRM, Silva AGL, Petri DFS. 2009. Hydrolytic Activity of Free and Immobilized Cellulase. Langmuir 25(3):1582-1587.

Tomas-Pejo E, Oliva JM, Ballesteros M, Olsson L. 2008. Comparison of SHF and SSF processes from steam-exploded wheat straw for ethanol production by xylose-fermenting and robust glucose-fermenting Saccharomyces cerevisiae strains. Biotechnol. Bioeng. 100(6):1122-1131.

Tu M, Zhang X, Paice M, MacFarlane P, Saddler JN. 2009. The potential of enzyme recycling during the hydrolysis of a mixed softwood feedstock. Bioresour. Technol. 100(24):6407-6415.

Vrsanska M, Biely P. 1992. The cellobiohydrolase I from Trichoderma reesei QM 9414: action on cello-oligosaccharides. Carbohydr. Res. 227:19-27.

White JS, White DC, Editors. 1997. Source Book of Enzymes. Boca Raton: CRC-Press. 1008 p.

Woodward J, Hillman SK, Vaughen BK. 1984. An immobilization method for the recovery of cellulase used in cellulose conversion to ethanol. Energy Biomass Wastes 8:1091-1111.

Wyman CE, Editor. 1996. Handbook on Bioethanol: Production and Utilization. Washington D.C.: Taylor & Francis. 424 p.

Zhang Y-HP, Lynd LR. 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. Biotechnol. Bioeng. 88(7):797-824.

Zhang YHP, Himmel ME, Mielenz JR. 2006. Outlook for cellulase improvement: screening and selection strategies. Biotechnol. Adv. 24(5):452-481.

Zhu Z, Sathitsuksanoh N, Percival Zhang YH. 2009. Direct quantitative determination of adsorbed cellulase on lignocellulosic biomass with its application to study cellulase desorption for potential recycling. Analyst (Cambridge, U. K.) 134(11):2267-2272.

Figure 1. Isotherm for the adsorption of cellulase to 40 nm average diameter nonporous silica nanoparticles (circles). The data are fit to a BET type II isotherm (solid line). Monolayer coverage was calculated to be 1.2 ± 0.3 mg/m^2. Error bars represent the standard deviation of 6 replicate measurements.

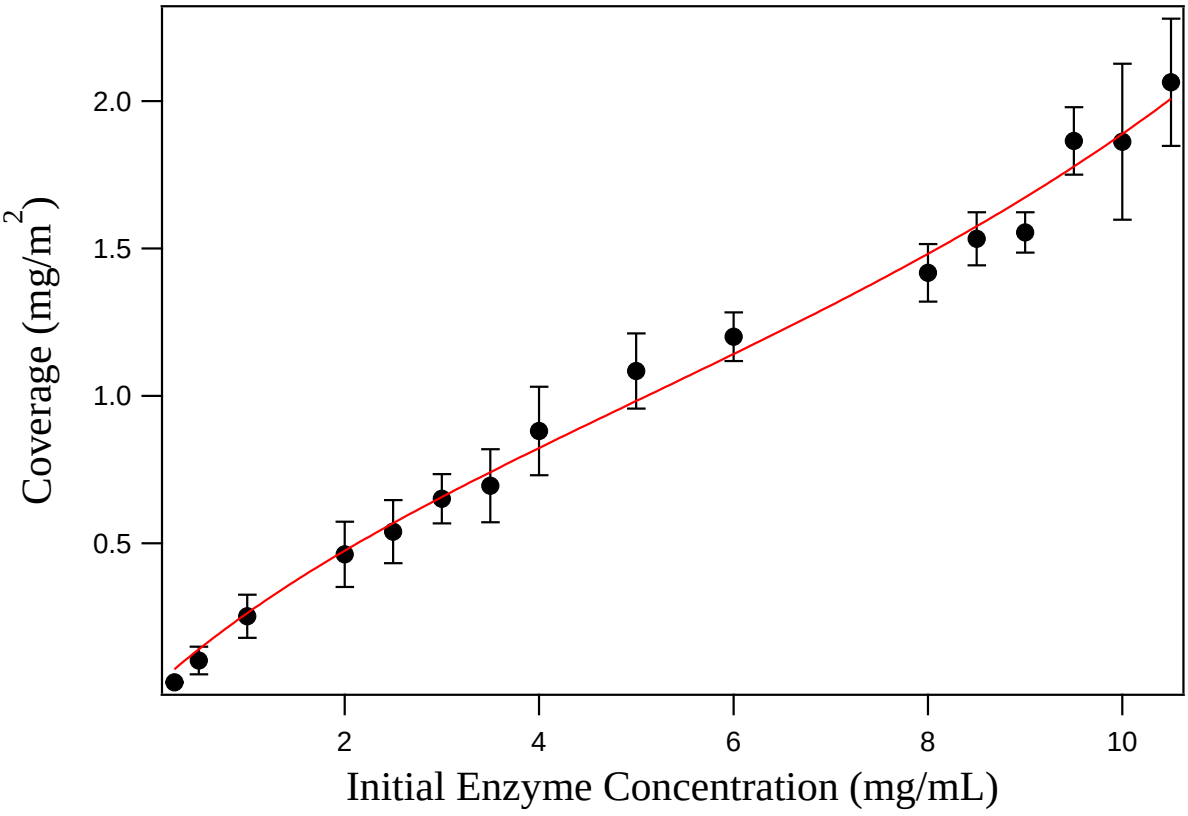


Figure 2. Ethanol produced during SSF reactions using silica-immobilized (black) or solution (light gray) cellulase at pH 4.8 and 5.3. The reaction was carried out for 96 hours, and the temperature was 35°C. P-values were calculated using the Student's t test and indicate statistical differences between the immobilized and solution reactions performed at the same pH.

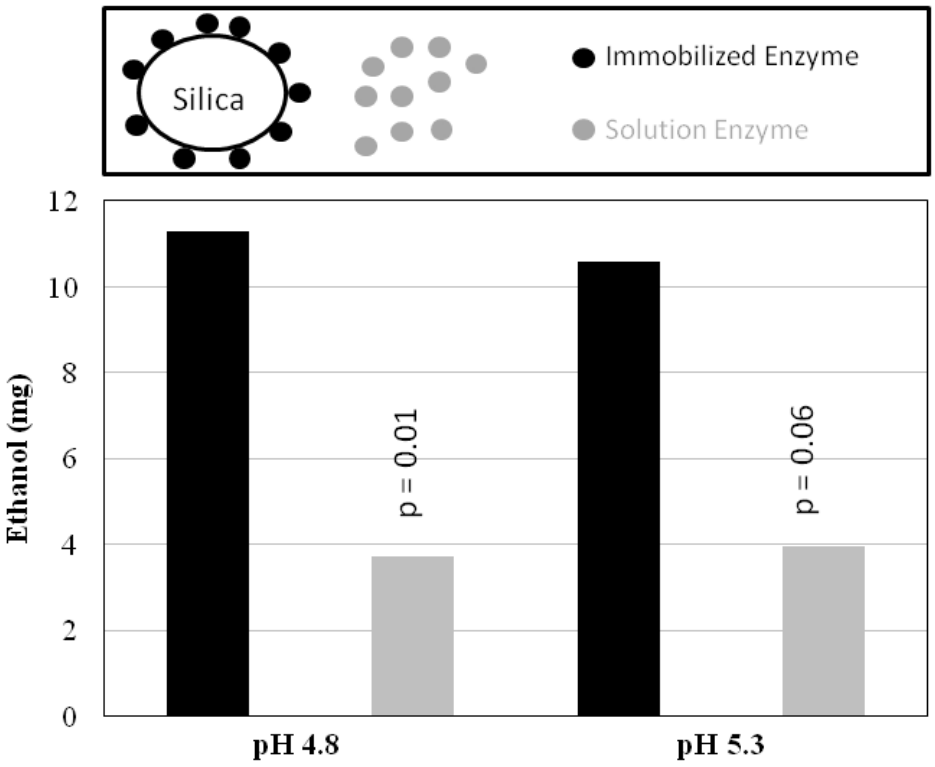


Figure 3. HPLC analysis of (A) 5.0 mg/mL glucose standard; (B) 5.0 mg/mL cellobiose standard; (C) 400 µg/mL cellotriose standard; (D) liquor from immobilized cellulase mock SSF reaction (no yeast, 96 hour, pH 4.8, 35°C); (E) liquor from solution cellulase mock SSF reaction (no yeast, 96 hour, pH 4.8, 35°C); (F) a control containing cellulose and no enzyme; and (G) a control containing cellulose and silica without enzyme. (D and E) The chromatograms for the mock SSF reactions were used for the quantitative analysis of glucose (average ± 1 standard deviation), cellobiose (below the limit of quantitation) and cellotriose (not detected). The HPLC calibration curves are shown in Supplemental Information Figure S2.

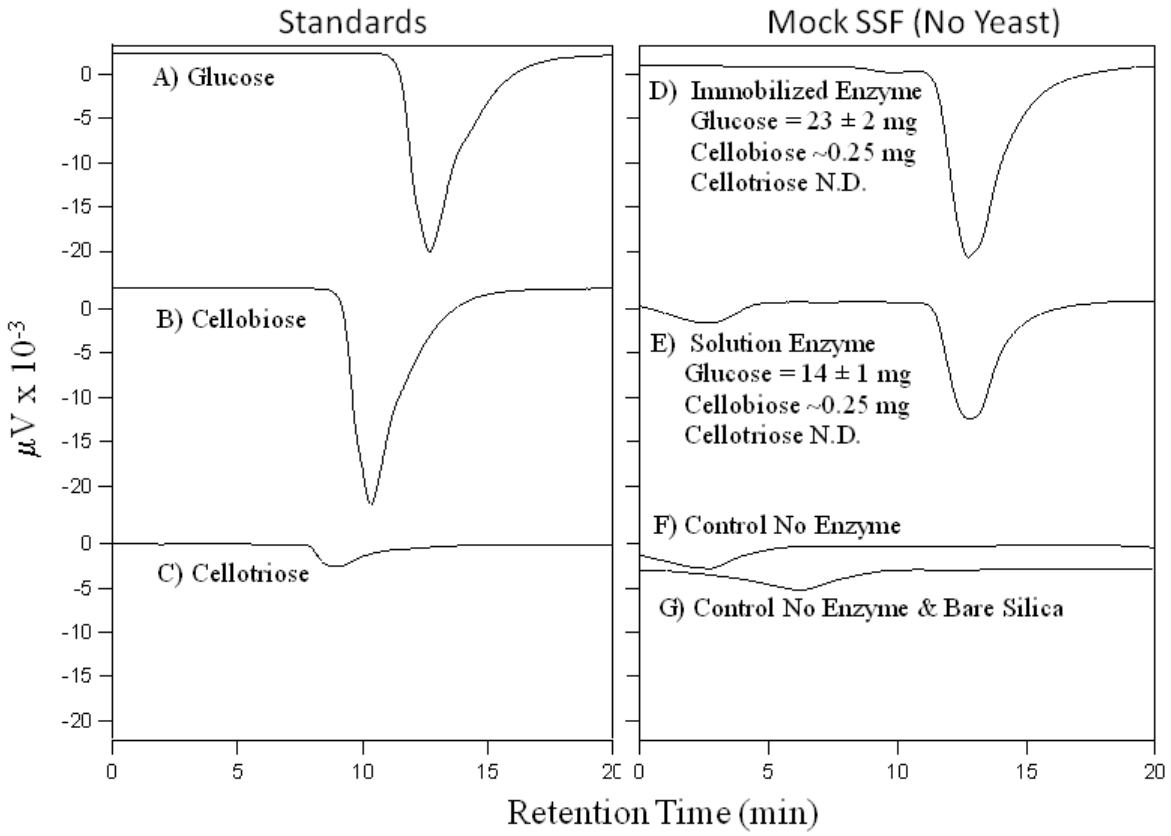


Figure 4. Glucose generated per mg enzyme in a 48 hour hydrolysis (50°C) reaction using solution cellulase (black), immobilized cellulase (green), and the maximum glucose that could be produced by cellulase desorbed from silica during the immobilized cellulase hydrolysis reaction (blue) as a function of hydrolysis medium pH. An expanded view between 0 and 0.25 mg glucose is shown for clarity. Error bars represent one standard deviation in three replicate measurements.

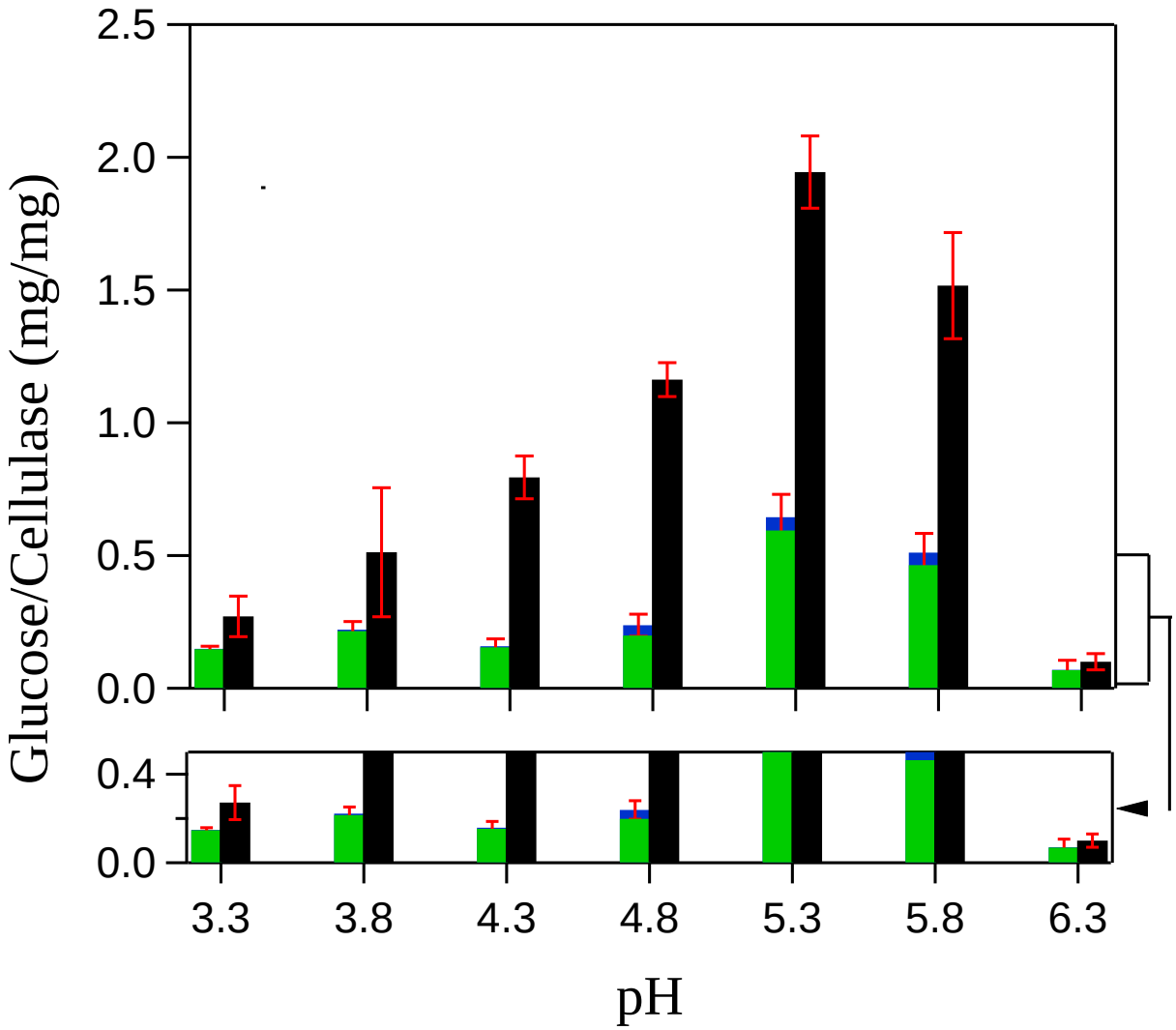


Figure 5. Glucose generated per mg enzyme in a 48 hour hydrolysis reaction (pH 4.8) using solution cellulase (black), immobilized cellulase (green), and the maximum glucose that could be produced by cellulase desorbed from silica during the immobilized cellulase hydrolysis reaction (blue) as a function of reaction temperature. Error bars represent one standard deviation in three replicate measurements.

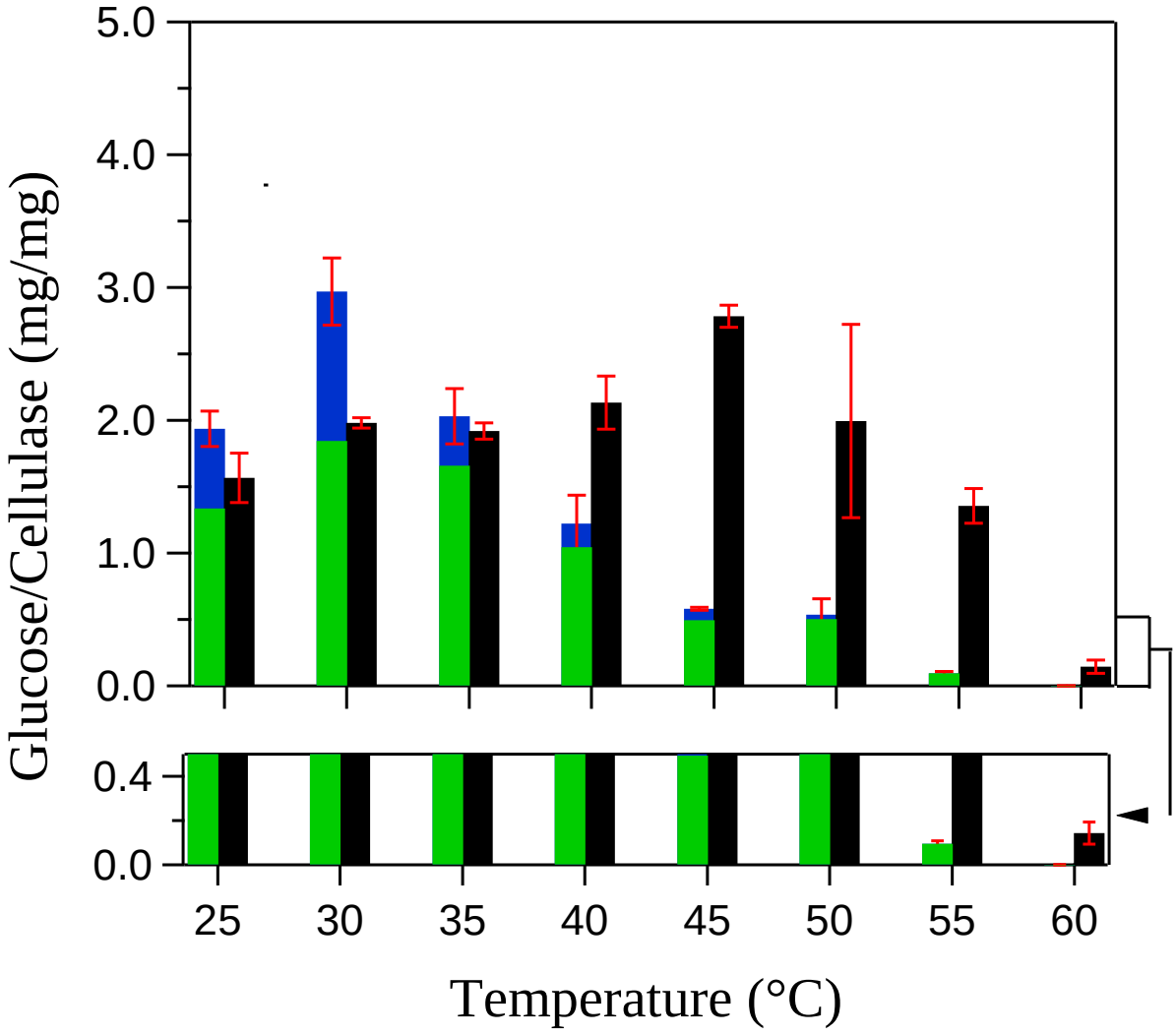
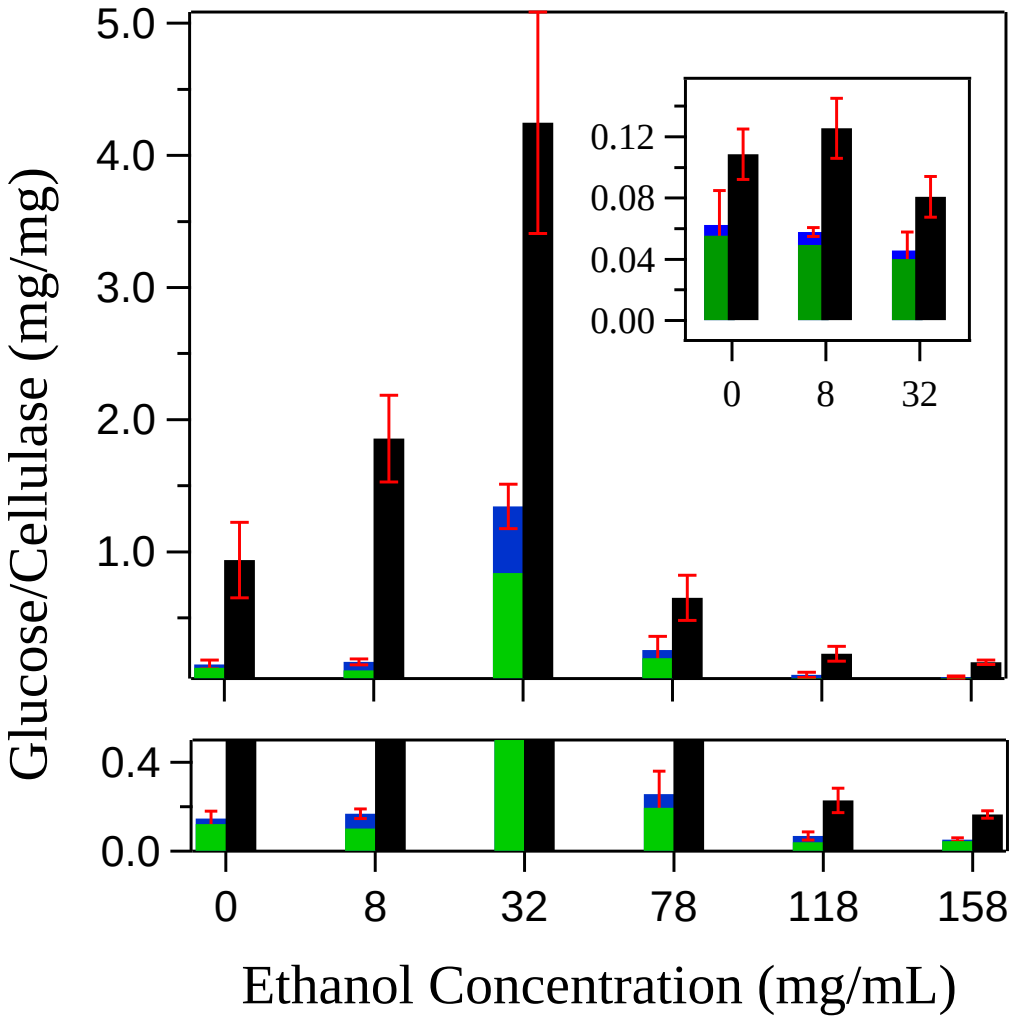


Figure 6. Glucose generated per mg enzyme in a 48 hour hydrolysis reaction (50°C, pH 4.8) using soluble cellulase (black), immobilized cellulase (green), and the maximum glucose that could be produced by cellulase desorbed from silica during the immobilized cellulase hydrolysis reaction (blue) as a function of ethanol concentration in the hydrolysis medium. The substrate was cellulose or (figure inset) soluble hydroxyethyl cellulose. An expanded view between 0 and 0.3 mg glucose is shown for clarity. Error bars represent one standard deviation in three replicate measurements.



SUPPLEMENTAL INFORMATION

EVALUATION OF NANOPARTICLE-IMMOBILIZED CELLULASE FOR IMPROVED ETHANOL YIELD IN SIMULTANEOUS SACCHARIFICATION AND FERMENTATION REACTIONS

Jason S. Lupoi and Emily A. Smith*

Ames Laboratory, U.S. Department of Energy, Ames, Iowa 50011-3111, U.S.A.

Department of Chemistry, Iowa State University, Ames, Iowa 50011-3111, U.S.A.

*Corresponding author. Tel.: 1-515-294-1424; fax: 1-515-294-0105; e-*mail address*: esmith1@iastate.edu; Department of Chemistry, 1605 Gilman Hall, Ames, IA 50010

Supplemental Methods

Quantitation of Adsorbed and Desorbed Cellulase

The adsorption isotherm was fit to a BET type II isotherm: $v=(v_m*c*p)/(p_0-p)\{1+(c-1)*(p/p_0)\}$; where V_m is the cellulase monolayer coverage on silica (mg/m^2), p is the initial enzyme concentration (mg/mL), p_0 is the saturation enzyme concentration (mg/mL), and *c* is a constant relating to the equilibrium adsorption constant.

Incubations were performed under all reaction conditions used in this study, where cellulose was omitted from the reaction mixture. At the end of the incubation, the cellulase-coated silica was removed from the solution by centrifugation at 2,500 g for 15 minutes. The amount of protein remaining in the supernatant corresponded to desorbed enzyme, and was quantified using BCA assays (Smith et al. 1985). Glucose was found to interfere with the BCA measurement of protein; therefore, quantitation of desorbed enzyme was not directly measured on the same sample were cellulose hydrolysis was performed. There was no trend in the amount of enzyme that desorbed from the silica surface with variations in the solvent composition, pH, or reaction temperature. In all cases, approximately 2.18 ± 0.05 to 25.8 ± 0.2 % (with the exception of the 30°C hydrolysis, 40.9 ± 0.7%) of the cellulase desorbed from the silica during the hydrolysis reaction. The cellulase that desorbed from the silica was likely due either to enzyme weakly associated with the silica surface (possibility due to the formation of protein-protein interactions rather than protein-silica interactions) that was not removed when the particles were rinsed in the same buffer used for subsequent incubations, or enzyme electrostatically repulsed by neighboring adsorbed enzyme molecules (Suvajittanont et al. 2000).

The cellulase that desorbed from the silica may possess catalytic activity towards cellulose degradation. It is also possible that part or all of the desorbed cellulase is inactive as a result of the adsorption/desorption process. In order to account for the maximum contribution of the desorbed cellulase to the immobilized cellulase reaction (i.e., assuming all the desorbed enzyme is active and that the desorption occurs at the beginning of the hydrolysis reaction), a hydrolysis reaction was set using enzyme in solution, where the amount of enzyme equaled the BCA quantified amount that desorbed from the silica, and a

parallel incubation was performed. The reaction conditions were the same as the immobilized enzyme reaction conditions the values were being compared to. The amount of glucose produced from these reactions was measured using the Glucose HK assay as discussed in section of 2.4 of the text. These values are plotted as the dark gray portion of the data bars in Figures 4, 5 and 6 for the immobilized cellulase reaction.

Thin Layer Chromatography (TLC)

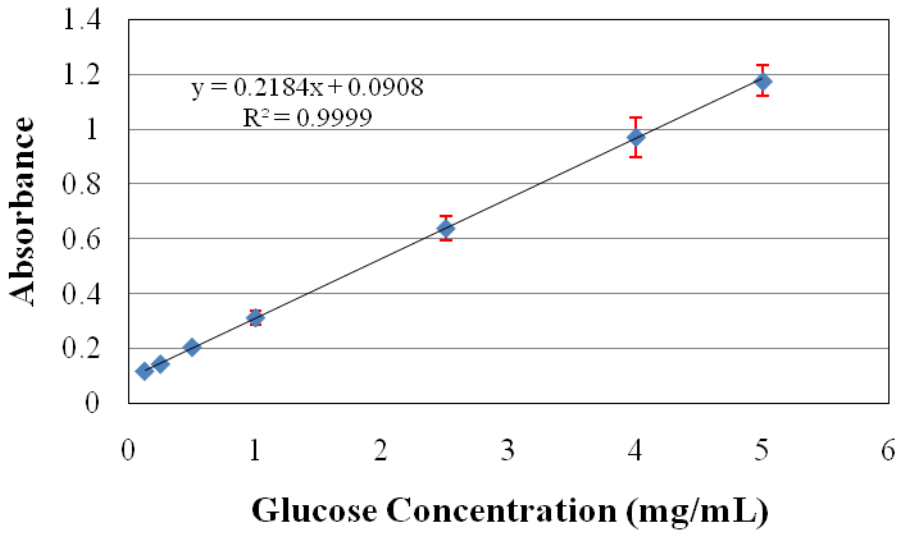
The TLC analysis of hydrolysis samples was previously described by Ke and Quinn (Ke and Quinn 1965), and was slightly modified. Briefly, Kieselguhr G plates (Analtech, Inc., Newark, DE) were activated by heating to 105°C for 30 minutes, and allowed to cool to room temperature. The plates were then heated to 50°C to facilitate spot drying. Standards of glucose (0.5-5 mg/mL), cellobiose (0.5-5 mg/mL), and cellotriose (1 mg/mL) were spotted, along with the immobilized and solution cellulase hydrolyses, and a control that lacked enzyme. For all standards and samples, 1 µL was spotted onto the plate 6 times (6 µL/spot). The mobile phase consisted of 65% sec-butanol, 15% 190 proof ethanol, and 20% water. After the solvent front had travelled approximately ¾ of the plate, the plates were removed from the developing chambers and allowed to air dry. The plates were then stained using a solution composed of 5 mL sulfuric acid, 90 mL 190 proof ethanol, 1 mL glacial acetic acid, and 0.5 mL p-anisaldehyde, followed by heating at 105°C until all spots were visible.

Ke YH, Quinn LY. 1965. Thin-layer chromatography of cello-oligosaccharides; its application to the detection of intermediates of cellulose digestion by bacteria. Iowa State Coll. J. Sci. 40(1):27-34.

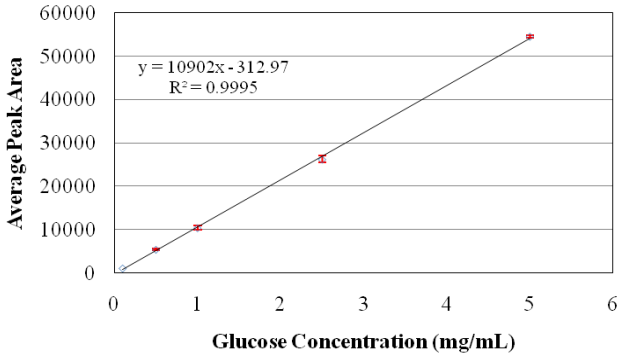
Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. 150(1):76-85.

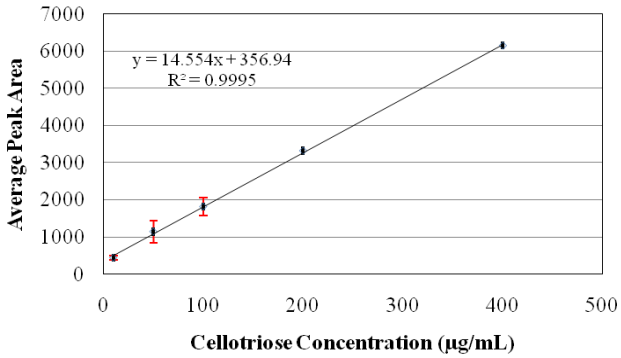
Suvajittanont W, Bothwell MK, McGuire J. 2000a. Adsorption of Trichoderma reesei CBHI cellulase on silanized silica. Biotechnol. Bioeng. 69(6):688-692.

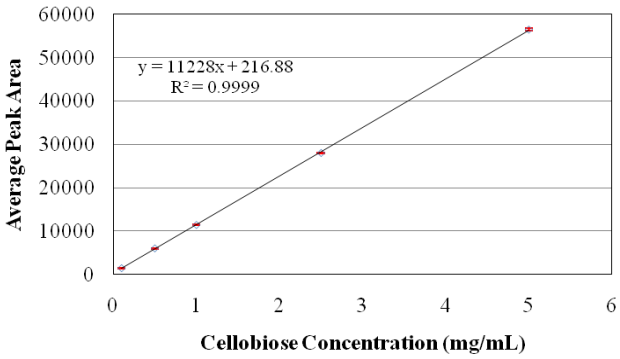
Supplemental Figure S1. Calibration curve produced using glucose standard solutions analyzed with glucose-HK assays using absorbance at 340 nm. Three replicate measurements were made for all concentrations. Error bars represent one standard deviation from three replicate measurements.



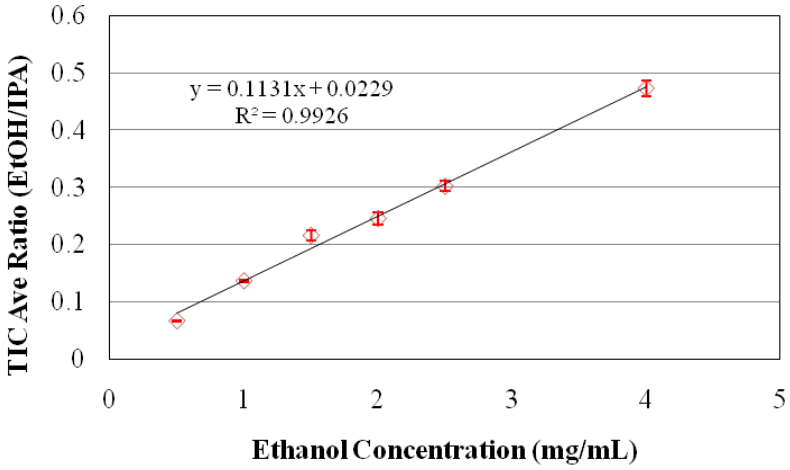
Supplemental Figure S2. Calibration curves from the HPLC analysis of glucose, cellobiose, and cellotriose standard solutions. The column temperature was 75°C, and the refractive index detector was heated to 50°C. Error bars represent one standard deviation from the replicate measurements. The calibration curves were used to measure the saccharide content in samples from mock SSF reactions that did not contain yeast.







Supplemental Figure S3. Gas Chromatography/Mass Spectrometry calibration curve generated using the total ion current (TIC) ratio of ethanol to isopropyl alcohol (internal standard) of standard solutions. Error bars represent one standard deviation from the replicate measurements. The calibration curve was used to determine ethanol concentration in SSF samples containing yeast.



Supplemental Figure S4. Thin layer chromatography (TLC) detection of glucose and a trace amount of cellobiose after a 96 hour hydrolysis using (A) immobilized, and (B) solution cellulase. Arrows represent the R_f of standard solutions, which were run on separate TLC plates.

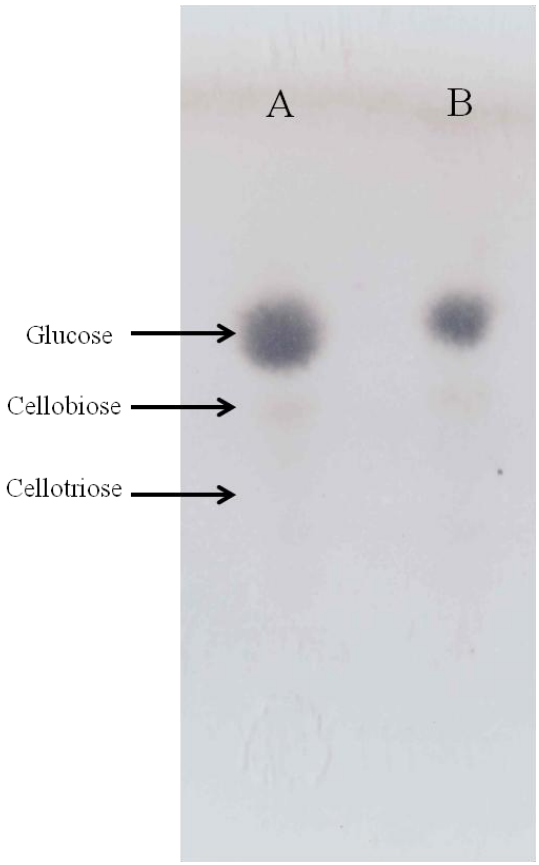


Table S1: Glucose measured by glucose HK assay or HPLC for 96 hour hydrolysis reaction (pH 4.8, 35°C) using solution cellulase or cellulase adsorbed to 40 nm average diameter silica nanoparticles. HPLC data is also shown in Figure 3 of the main manuscript.

Sample	Glucose (mg); Glucose HK Assay	Glucose (mg); HPLC Analysis	Cellobiose (mg); HPLC Analysis	Cellooligomers (mg); HPLC Analysis
Immobilized Cellulase	**19 ± 1**[1]	**23 ± 2**[1]	**~0.25**[3]	**ND**[4]
Solution Cellulase	**13 ± 1**[1] **(p=0.001)**[2]	**14 ± 1**[1] **(p=0.01)**[2]	**~0.25**[3]	**ND**[4]

1 Average ± standard deviation of three replicate experiments.

2 p-value from Student's t-test to compare immobilized and solution cellulase glucose yield

3 Below the limit of quantitation

4 Not detected

CHAPTER 3.

1064 NM DISPERSIVE MULTICHANNEL RAMAN SPECTROSCOPY FOR THE ANALYSIS OF PLANT LIGNIN

(A paper published in *Analytica Chimica Acta*, **2011**, 706, 164-170)

Matthew W. Meyer[1], Jason S. Lupoi[1], and Emily A. Smith

[1] Both authors contributed equally to this work

Abstract

The mixed phenylpropanoid polymer lignin is one of the most abundant biopolymers on the planet and is used in the paper, pulp and biorenewable industries. For many downstream applications, the lignin monomeric composition is required, but traditional methods for performing this analysis do not necessarily represent the lignin composition as it existed in the plant. Herein, it is shown that Raman spectroscopy can be used to measure the lignin monomer composition. The use of 1064 nm excitation is needed for lignin analyses since high fluorescence backgrounds are measured at wavelengths as long as 785 nm. The instrument used for these measurements is a 1064 nm dispersive multichannel Raman spectrometer that is suitable for applications outside of the laboratory, for example in-field or in-line analyses and using remote sensing fiber optics. This spectrometer has the capability of acquiring toluene/acetonitrile spectra with 800 cm^{-1} spectral coverage, 6.5 cm^{-1} spectral resolution and 54 S/N ratio in 10 seconds using 280 mW incident laser powers. The 1135 to 1350 cm^{-1} and 1560 to 1650 cm^{-1} regions of the lignin spectrum can be used to distinguish among the three primary model lignin monomers: coumaric, ferulic and sinapic acids. Mixtures of the three model monomers and first derivative spectra or partial least squares analysis of the phenyl ring breathing modes around 1600 cm^{-1} are used to determine sugarcane lignin monomer composition. Lignin extracted from sugarcane is shown to have a predominant dimethoxylated and monomethoxylated phenylpropanoid content with a lesser

amount of non-methoxylated phenol, which is consistent with sugarcane's classification as a non-woody angiosperm. The location of the phenyl ring breathing mode peaks do not shift in ethanol, methanol, isopropanol, 1,4 dioxane or acetone.

Keywords**:** 1064 nm dispersive multichannel Raman spectrometer; lignin monomer composition; p-hydroxyphenyl; guaiacyl; syringyl; partial least squares

1. Introduction

Lignin is one of three primary biopolymers in plant cell walls [1, 2]. It is a heteropolymer composed of phenylpropanoid monomers with the aromatic portions: p-hydroxyphenyl, guaiacyl, and syringyl. These monomers are often referred to by the one letter designation H, G, and S, respectively. Lignin is classified as heterogeneous not only because of its complex composition, but also because lignin from different species, tissue types, plant age and growing conditions can be quite different. Lignin is classified as predominately G (gymnosperms), G/S (angiosperms) and G/S/H (non-woody angiosperms), although many exceptions to these classifications are known [3-5]. Lignin has many commercial uses in the pulp and paper industry and as the raw material for biorenewable fuels and commodity chemicals. For these applications it is desirable to know the lignin purity and composition. For example, the resulting complex chemical mixture generated from lignin pyrolysis (i.e., bio-oil) is dependent on the starting monomer composition [6]. Additionally, the degradation of lignin in the environment is slow, and the monomer composition is a biomarker that can be used to identify the source of sediments and soils containing decaying plant materials [6]. Reported lignin monomer compositions in the literature have large variations due to the commonly utilized extraction and analysis procedures, which can alter phenylpropanoid structure [5, 7-9].

Raman spectroscopy has been used to measure chemical content in a wide-range of samples [10-14]. The appropriate selection of excitation wavelength is one of the most important factors for Raman spectroscopy experiments. Near infrared (commonly 785 to 1064 nm) excitation is often used for detecting biological materials that can exhibit fluorescence backgrounds using visible wavelengths [15]. The intensity of Raman scatter is

proportional to the incident frequency to the fourth power. Using the same laser radiance, excitation with 785 nm will generate 3.8-times more scattered photons compared to 1064 nm excitation. Despite the low energy of the photons, and even with low fluorescence quantum yield samples, a fluorescence background may render the Raman spectrum undetectable or increase detection limits using 785 nm excitation [16]. In these cases, the lower background obtained with 1064 nm excitation often outweighs the reduced scattering intensity [17]; and higher laser radiances may be used to compensate for the lower Raman cross section above one micron. The use of up to 2.3 W incident 1064 nm laser powers have been reported in the literature for Raman measurements [16, 18].

Reports of multichannel dispersive 1064 nm Raman spectrometers date back to 1986 [19], yet only a handful of references using a dispersive 1064 nm instrument are found in the recent literature. Fourier Transform (FT) Raman spectroscopy remains the primary instrument format for obtaining fluorescence-free Raman spectra. Yet, dispersive systems are better suited for many applications where robust, low cost instruments and remote sensing fibers are desirable [20-22]. Recently, portable Raman spectrometers have been made commercially available [23]. However, many of these systems are limited in performance characteristics, such as spectral range and spectral resolution.

The main limitation of 1064 nm dispersive Raman spectrometers is the availability of suitable low-noise detectors. Single channel detectors require minutes to hours to collect a 1064 nm Raman spectrum [18, 24-26]. Multichannel dispersive Raman instruments with excitation wavelengths below 800 nm commonly utilize charge coupled devices (CCD). CCDs have a low response beyond 1.1 µm, limiting their use in 1064 nm Raman spectroscopy [27]. Hamaguchi *et al.* overcame this limitation using a transfer electron InP-InGaAs photocathode image intensifier CCD for dispersive 1064 nm Raman spectroscopy for the analysis of DNA and lung tissue [16, 26, 28]. This detector has a ~4% quantum efficiency between 1 and 1.4 microns (0 to 2250 cm^{-1} Raman shift) and required lengthy acquisition times. CCDs have also been used to collect dispersive 1064 nm excitation anti-stokes Raman spectra [29].

Indium gallium arsenide (InGaAs) detectors with quantum efficiencies above 80% have high sensitivity between 900 and 1,700 nm. This enables detection out to the 3,100 cm^{-}

[1] C-H stretching region using 1064 nm excitation. These quantum efficiencies are substantially higher than those of the competing germanium detectors [30]. Dispersive multichannel 1064 nm Raman spectrometers equipped with 128 or 256 element Ge or InGaAs arrays have been used to measure strong Raman scattering solids and liquids (e.g., toluene, anthracene) [26, 31-33]. These multichannel InGaAs detectors had high dark and read noise and provide limited 100-500 cm^{-1}spectral coverage in a single acquisition. The development of improved multichannel InGaAs detectors with 1024 or more elements and low read noise has opened the possibility of improved 1064 nm dispersive multichannel Raman instruments with wide spectral coverage that compete with FT Raman instruments.

Raman spectroscopy is an ideal non-invasive screening method to measure lignin composition. Lignin has been previously analyzed by Kerr gated resonance Raman spectroscopy using a benzoquinone/laccase/mediator treatment to generate a charge transfer state that extends lignin's absorption out to 500 nm [34]. Ultraviolet RR spectroscopy has also been used to measure lignin since the fluorescence signal is spectrally separated from the Raman spectrum at these wavelengths [35]. FT Raman and surface enhance FT Raman spectroscopies using a 600 mW 1064 nm laser have been used to measure and assign lignin spectral peaks [36]. The presence of lignin in plant tissues has also been measured by Raman spectroscopy [37, 38]. Finally, the spectral differences of three model lignin monomers were identified experimentally and using DFT calculations [39], but were not extended to real plant lignin samples or monomer mixtures. The reported differences in the monomer spectra reveal the possibility of measuring lignin compositions using Raman spectroscopy.

Herein, we describe a dispersive 1024-multichannel 1064 nm Raman spectrometer with wide spectral coverage, S/N values that exceed those of a FT Raman instrument for most acquisition times, and background values orders of magnitude lower than a dispersive 785 nm Raman instrument for the analysis of lignin and model lignin monomers. The results show that the new generation multichannel InGaAs array detectors produce high quality spectra with ~280 mW laser radiance using relatively fast acquisition times, and will enable facile measurements in laboratory and other settings. It is demonstrated that Raman spectroscopy can be used to quantitatively measure the lignin monomer composition using partial least squares analysis and qualitatively using first derivative spectra.

2. Experimental

2.1 Materials

All chemicals were of analytical grade and purchased from Sigma-Aldrich (St. Louis, MO). Hydrolytic lignin (Aldrich, CAS 8072-93-3) was extracted from sugarcane. Aqueous solutions were prepared using deionized water (18.2 MΩ cm^{-1}) from an Ultrapure II water system (Thermo Scientific, Waltham, MA).

2.2 Dispersive Multichannel 1064 nm Instrumentation

A home-built 1064 nm dispersive multichannel Raman instrument is shown in Figure 1. A 1064 nm Nd:YVO_4 solid state laser (Blue Sky Research, Milpitas, CA) with a maximum output power of 1 W was used as the excitation source. The beam was focused at the sample with a plano-convex lens with a focal length of 250 mm (Newport Optics, Irvine, CA). Liquid samples were placed in a nuclear magnetic resonance tube (Sigma-Aldrich), and mounted on a home-built sample stage. Raman scatter was collected in a 90 degree collection geometry. The light was collected with a plano-convex lens with a focal length of 60 mm (f/1.2). The light was focused with a plano-convex lens with a focal length of 200 mm (f/3.9). The optical system was designed to slightly overfill the 150 µm entrance slit of the spectrometer to maximize signal. A holographic super notch filter (HNF) centered at 1064 nm (Kaiser Optical Systems, Ann Arbor, MI) was placed before the spectrometer. To aid in laser alignment, a helium neon laser was made co-linear with the 1064 nm laser using a dichroic mirror (Omega Optical, Brattleboro, VT) to reflect the 633 nm light and transmit the 1064 nm light. The collected signal was sent to a iHR320 spectrometer, fitted with a 1024 element InGaAs detector (Horiba Jobin Yvon, Edison, NJ). The InGaAs detector had a read noise of 462.1 e^- rms and a dark current of 3.7 ke^- $pixel^{-1}$ $second^{-1}$ when cooled to 170 K. Among the 300, 600 grooves/mm classically ruled and 950 grooves/mm blazed holographic gratings in the spectrometer, the 600 groove/mm classically ruled grating provided the highest efficiency and the best trade-off between spectral coverage and spectral resolution. Spectra were collected with 280 mW incident laser power, which produces less than 0.8 ± 0.1 ºC sample heating in 30 minutes for aqueous samples. Other instrument details are listed in Supporting Information Table S1.

2.3 Other Instrumentation

A home-built 785 nm Raman microscope has been previously reported [40]. Briefly, a 785 nm diode laser was directed onto the rear port of a Nikon inverted microscope with a 40x magnification, 1.0 numerical aperture objective. The collected signal was sent to a Holospec VPH spectrometer (Kaiser Optical System, Ann Arbor, MI) and a near IR enhanced PIXIS CCD camera (Princeton Instruments, Trenton, NJ). FT Raman measurements were collected with a commercially available Nicolet NXR 9650 FT Raman spectrometer (Thermo Scientific).

2.4 Raman Measurements

Spectra of a 1:1 (v/v) acetonitrile/toluene mixture were used both for wavelength calibration and to optimize the intensity of the collected Raman scatter. Laser powers were 189 mW (785 nm), 280 mW (1064 nm dispersive) and 480 mW (1064 nm FT). The slit widths were set to 100 micron (785 nm) and 150 micron (dispersive 1064 nm). Model lignin monomer or mixture (100 mg mL^{-1}) measurements were obtained with 60 second integrations and 3 accumulations (1064 nm dispersive) or 300 scans (1064 nm FT). Lignin (50 mg mL^{-1}) measurements were acquired with 600 second integrations and 3 accumulations using the 1064 nm dispersive instrument. Sinapic and ferulic acid in 1,4-dioxane or methanol were heated to 37 °C in a water bath to ensure dissolution. All Raman measurements were acquired at room temperature.

2.5 Data Analysis

Raman spectra were collected using the software Winspec32 (Princeton Instruments) for the 785 nm dispersive instrument and SynerJY (Horiba Jobin Yvon) for the 1064 nm dispersive instrument. FT Raman measurements were collected using the software Omnic (Thermo Scientific). All spectra were imported into Grams AI V 8.0 (Thermo Scientific) and background subtracted using a blank containing all matrix components except the analyte. The spectra were not intensity corrected for the instrument's varying throughput and quantum efficiency across the spectral region. Intensity correction of the spectra from the

1064 nm dispersive instrument can be performed using the curve shown in Supplemental Information Figure S1.

The spectra were imported into Igor Pro V 6.1 (WaveMetrics Inc. Lake Oswego, OR) for further analysis. The raw spectral intensity was divided by the total integration time to plot the spectra in counts per second. Signal-to-noise (S/N) ratios were calculated using the background subtracted peak intensity for the peak with highest intensity in the spectrum divided by the root mean square noise measured from 940 to 980 cm^{-1}. The multipeak fitting package was used to deconvolute three overlapping spectral peaks in the 1560 to 1650 cm^{-1} region. The procedure minimizes the residual between the experimental values and the fit peaks' location, width and height. The peak shape was set to a Gaussian profile. Igor Pro V6.1 was also used to obtain first derivative spectra.

The partial least squares (PLS) analysis was performed using The Unscrambler X (Camo Inc., Oslo, Norway). A data set was comprised of 36 rows, representing 36 standard monomer mixtures, and 261 columns, representing the spectral regions of interest (1138-1231 cm^{-1}, 1257-1286 cm^{-1}, 1325-1350 cm^{-1}, and 1580-1614 cm^{-1}). The data was mean-centered, and the model was randomly cross-validated. The NIPALS algorithm was used to calculate PLS factors. This model was used to predict the score coordinates for randomly selected monomer mixtures in order to test the model, and to determine the monomer composition of lignin extracted from sugarcane.

3. Results and Discussion

3.1 Comparison of 1064 and 785 nm Excitation for the Analysis of Lignin

The goal of this work is to demonstrate the development and use of a 1064 nm dispersive multichannel Raman spectrometer for the analysis of the heterogeneous biopolymer lignin, and to show that it can be used to measure lignin monomer composition. The impact of this work will be the capability to take Raman measurements in diverse locations using robust instrumentation and the identification of lignin compositions using Raman spectra for several downstream applications.

In order to determine the most suitable wavelength for the analysis of lignin, the Raman spectra of an extracted lignin sample were collected using 785 and 1064 nm

excitation (Figure 2A). Since the Raman scatter scales with the excitation frequency to the fourth power, the highest possible laser frequency that produces minimal fluorescence background should be used. The 785 nm spectrum with 60 second integration and 3 accumulations shows no discernible Raman peaks and a broad background that is 160 times higher than the background of the 1064 nm Raman spectrum with 600 second integration and 3 accumulations. The fluorescence spectrum of the same lignin sample, when excited with 785 nm, shows a low intensity peak that mimics the 785 nm Raman spectrum background (Supplemental Information Figure S2). This fluorescence is either intrinsic to the native lignin or a result of the extraction process.

The low energy photons of the 1064 nm laser produce a flat background in the Raman spectrum and discernible peaks that are characteristic of lignin (Figure 2B). The peak assignments are shown in Supplemental Information Table S2. The most intense peaks are centered around ~1600 cm^{-1} and are assigned to the aromatic ring breathing modes (Wilson notation v_{8a} and v_{8b}) and substituent carbon-carbon double bonds. The results show that 1064 nm excitation produces Raman spectra of extracted plant lignin with a low background.

3.2 1024-Channel Dispersive 1064 nm Raman Spectrometer

The multichannel dispersive 1064 nm Raman spectrometer (Figure 1) leverages recent developments in InGaAs detectors with 1024 sensing elements that are 25 x 250 microns with relatively low read and dark noise compared to previous generation near IR detectors. The S/N ratios and spectral resolution of the 1064 nm dispersive Raman spectrometer are compared to a commercial FT Raman instrument using a series of toluene-acetonitrile spectra (Supplemental Information Figure S3). The S/N values were calculated and graphed in Figure 3 as a function of total analysis time, which accounts for the amount of time required to average multiple spectra when appropriate. The FT Raman S/N curve is fairly constant at all analysis times included in this study. The dispersive instrument's S/N has been measured for: varying integration times with 1 accumulation, and set integration times (60 or 150 seconds) with varying number of accumulations. The dispersive S/N data are independently fit to a curve with a dependence on the square root of the total analysis time. All three fit curves nearly overlap, and are consistent with a detector that is dark noise

limited [30]. At total analysis times below 15 seconds, the FT spectra show a 2 to 8-fold higher S/N ratio compared to the dispersive instrument. However, at total analysis times greater than 15 seconds, the dispersive instrument has a higher S/N ratio than the FT instrument. The spectral resolution is 6.5 cm^{-1} for the dispersive spectra and 8 cm^{-1} for the FT spectra, although higher spectral resolution is possible with both instruments.

In order to investigate the possibility of using Raman spectroscopy to measure the source of extracted lignin and materials derived from lignin, model phenylpropanoid monomers were studied. Lignin is primarily composed of three phenyl monomers with hydroxy- and methoxy- substituents. Three carboxylic acids were used as model lignin monomers: ferulic (guaiacyl, G); coumaric (p-hydroxyphenyl, H) and sinapic (syringyl, S) acids. The Raman spectra of the model lignin monomers obtained using both the 1064 nm dispersive and FT Raman spectrometers are shown in Figure 4. Under the conditions used in this study, similar spectra are observed with both instruments and differences in the monomer spectra are identifiable in the 1135 to 1350 cm^{-1} and 1560 to 1650 cm^{-1} spectral regions. Peaks at 1155.2 ± 0.6 and 1169.3 ± 0.4 cm^{-1} are characteristic of sinapic and coumaric acid, respectively. The dispersive spectra shown in Figure 4 were collected without tiling together different spectral regions, and demonstrate that the dispersive Raman spectra have sufficient spectral coverage to encompass both of the critical regions for lignin monomer identification. This has been problematic with previously reported 1064 nm dispersive multichannel Raman instruments [31].

The phenyl ring breathing modes at ~1600 cm^{-1} are composed of two partially resolved peaks in coumaric and ferulic acids. Both vibrations are degenerate in sinapic acid. A procedure for determining peak maxima when there is peak overlap is to use the zero crossing points in first derivative spectra (Figure 4, inset). There are identifiable peak maxima for all three monomers (1587.8 ± 0.4 H, 1591.0 ± 0.5 G, 1594.0 ± 0.2 S, 1603.0 ± 0.9 G, 1606.2 ± 0.5 H cm^{-1}). The additive intensity of the v_{8a} and v_{8b} ring breathing mode peaks is sensitive to substitution for many phenyl compounds [41]. The intensity of the two peaks increases with increasing conjugation (i.e., benzene to styrene) and to a lesser extent with electron donating or electron withdrawing substituents on the phenyl ring. Peak deconvolution was used to determine the peak maxima and intensities in the 1560 to 1650

cm^{-1} spectral region for all three model lignin monomers (Supplemental Information Table S3). The intensity corrected [42] additive v_{8a} and v_{8b} normalized peak intensities for the three monomers are 2.2 for G, 1.0 for S, and 3.5 for H. The multipeak fitting and first derivative spectral results suggest it may be possible to evaluate the lignin monomer composition using Raman spectroscopy.

3.3 Analysis of Lignin Monomer Composition with First Derivative Spectra and Chemometrics

The Raman spectrum of a commercially obtained lignin sample from the non-woody angiosperm sugarcane is plotted in Figure 5 along with the spectra of three model monomer mixtures with different compositions (G-S-H%: 5-90-5, 51-40-9 and 25-25-50). Visual inspection of the spectra indicates that the lignin most closely matches the 51-40-9 model lignin monomer mixture, which is consistent with a non-woody angiosperm. There is a peak maximum at ~1606 cm^{-1} in both the monomer mixture and lignin spectra. The peak at 1633 cm^{-1} is assigned to substituent carbon-carbon double bonds, and it has a lower intensity in the lignin spectrum than is measured for the model lignin monomers. This suggests that the phenyl ring's substituent aromaticity is lower in the lignin sample than the model monomers. This is not surprising since the model monomers were chosen solely for the substitution pattern of the hydroxyl- and methoxy-groups.

In order to develop a quantitative model for the lignin monomer composition, partial least squares analysis was performed on the spectra from 36 monomer mixtures (Figure 6A). The PLS scores plot forms a triangle with the pure monomers at the corners. Mixtures with high G content are at the top of the triangle, mixtures high in S are to the left and those high in H are at the right side of the triangle. The majority of the monomer compositions were chosen based on real plant compositions. Since H content is not high in most plants, there are fewer mixtures represented in this part of the scores plot. Using the developed model, the sugarcane hydrolytic lignin was located on the scores plot closest to the standards with the following composition (G-S-H%): 51-40-9 and 50-25-25. This is consistent with the visual inspection of the Raman spectrum (Figure 5) and also for a non-woody angiosperm. For sugarcane lignin, the PLS model reported a 33-41-33 monomer composition. The uncertainty

is as large as the measurement due to noise in the lignin spectrum. Higher laser powers could be used without sample damage to increase the S/N ratio. A referenced literature value for sugarcane lignin composition is 39-33-28, which was obtained using alkaline nitrobenzene oxidation-gas chromatography.[43]

The first derivative Raman spectra of the 36 monomer mixtures were analyzed as a qualitative method for describing the monomer composition (Figure 6B). The spectra have been color coded based on the grouping shown in Figure 6A. It can be seen that the location of the monomer mixture on the scores plot corresponds to the peak maxima located using first derivative spectra. When S is greater than 70% of the monomer composition, there is a single peak maximum for the phenyl ring breathing mode at 1593.8 ± 0.4 cm^{-1} and a second, smaller peak at 1155.0 ± 0.9 cm^{-1}. When G or S/G dominates the monomer composition there is a peak maximum at 1603 ± 1 cm^{-1} when H is less than 20%, and 1604.7 ± 0.3 cm^{-1} when H is between 21 and 30% of the monomer mixture. When all three monomers are present at roughly equal proportions, there is a peak maximum at 1605.8 ± 0.2 cm^{-1} and when H is greater than 25% there is a small peak at 1168.4 ± 0.7 cm^{-1}. The first derivative spectra of the monomer compositions 49-51-0, 56-40-4 and 19-55-26 showed characteristics in between mixtures with predominant G or S compositions, and a clear assignment into one of the groups was not possible. The above information enables the lignin monomer composition to be qualitatively determined for most mixtures with minimal data analysis.

The additive v_{8a} and v_{8b} peak intensity for the 36 monomer mixtures shown in Figure 6B did not quantitatively correlate with the monomer composition (data not shown), contrary to the results for the pure monomer spectra. Overall, additional information is not gained from peak deconvolution of the phenyl ring breathing modes compared to the first derivative spectra and PLS analysis.

The peaks in the lignin spectrum may be sensitive to solvent effects, and if performed in plant tissues, the nature of the surrounding environment. In order to measure possible shifts in the phenyl ring breathing modes and to see if additional information about the monomer mixture can be measured, lignin spectra were collected in different non-aromatic solvents for which lignin has an appreciable solubility (Figure 7). The S/N ratio varied from solvent to solvent, likely due to differences in lignin solubility. When the spectra are

normalized to the most intense peak in the spectrum, few differences are observed in the spectra with the exception of lignin in 100 mM NaOH. As indicated by the first derivative spectra, all peak maximums occurred at 1605.0 ± 0.5 cm^{-1} except for the aqueous spectrum. Peak deconvolution also showed no shifts within the uncertainty in the phenyl ring breathing modes for 1,4-dioxane, ethanol, methanol, acetone and isopropanol. Lignin in dilute alkaline solutions is readily oxidized, and this is the likely explanation for the differences in the Raman spectrum of lignin in 100 mM NaOH [44]. This has been confirmed by a bathochromic shift and hyperchromic effect observed in the ultraviolet-visible spectrum of the alkaline lignin compared to that for the other solvents (data not shown) [8]. Given the fact that the spectra in different solvents do not provide additional details regarding the structure or composition of the lignin, it is desirable to use a nonaromatic solvent with the highest S/N ratio for Raman measurements, which includes methanol, ethanol or 1,4 dioxane.

4. Conclusions

The 1064 nm dispersive multichannel Raman spectrometer has performance characteristics that meet or exceed those of FT Raman spectroscopy for the analysis of lignin and model lignin monomers. The described instrument enables the collection of high quality spectra with simultaneous high spectral resolution and wide spectral coverage using a single acquisition. The described instrument will outperform alternative Raman spectrometers for the analysis of plant materials and other moderately fluorescent biological materials when robust instruments are needed, for example before plant harvest or at the biofuel processing plant. The origin of extracted lignin samples can be determined by Raman spectroscopy, and the monomer composition may be measured qualitatively using first derivative spectra and quantitatively using chemometrics. The partial least squares analysis with Raman spectroscopy offers the ability to determine more than one monomer component at a time. Current efforts are underway to use the partial least squares method for the analysis of lignin samples from diverse sources.

Acknowledgments

This research is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences through the Ames Laboratory. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358. The authors thank Kevin Langenwalter and the Department of Chemistry, Wichita State University for the use of the FT Raman instrument.

References

[1] S.-Y. Ding, M.E. Himmel, Journal of Agricultural and Food Chemistry, 54 (2006) 597.

[2] Y.-H.P. Zhang, L.R. Lynd, Biotechnology and Bioengineering, 88 (2004) 797.

[3] A.T. Martinez, M. Speranza, F.J. Ruiz-Duenas, P. Ferreira, S. Camarero, F. Guillen, M.J. Martinez, A. Gutierrez, J.C. del Rio, International Microbiology, 8 (2005) 195.

[4] N.G. Lewis, E. Yamamoto, Annual Review of Plant Physiology and Plant Molecular Biology, 41 (1990) 455.

[5] D.R.D. Cyril Heitner, and John A. Schmidt (Ed.), Lignin and Lignans, CRC Press 2010, 2010.

[6] C. Saiz-Jimenez, J.W. De Leeuw, Organic Geochemistry, 10 (1986) 869.

[7] D.N. Thompson, B.R. Hames, C.A. Reddy, H.E. Grethlein, Biotechnol. Bioeng., 57 (1998) 704.

[8] S.Y. Lin, Methods in Lignin Chemistry, Spring-Verlag, 1992.

[9] K. Saito, T. Kishimoto, Y. Matsushita, T. Imai, K. Fukushima, Surf. Interface Anal., 43 (2011) 281.

[10] C. Otto, C.J. De Grauw, J.J. Duindam, N.M. Sijtsema, J. Greve, Journal of Raman Spectroscopy, 28 (1997) 143.

[11] E.A. Cutmore, P.W. Skett, Spectrochim. Acta, Part A, 49A (1993) 809.

[12] C. Cheng, T.E. Kirkbride, D.N. Batchelder, R.J. Lacey, T.G. Sheldon, J. Forensic Sci., 40 (1995) 31.

[13] J.M. Chalmers, Everall, N. J., in: J.M. Chalmers, Griffiths, P. R. (Ed.), Handbook of Vibrational Spectroscopy, Wiley & Sons: Chichester, New York, 2002, p. 2389.

[14] Y.Y. Huang, C.M. Beal, W.W. Cai, R.S. Ruoff, E.M. Terentjev, Biotechnology and Bioengineering, 105 (2009) 889.

[15] Y. Wang, R.L. McCreery, Anal. Chem., 61 (1989) 2647.

[16] Y.-K. Min, T. Yamamoto, E. Kohda, T. Ito, H.-o. Hamaguchi, Journal of Raman Spectroscopy, 36 (2005) 73.

[17] V.M. Hallmark, C.G. Zimba, J.D. Swalen, J.F. Rabolt, Spectroscopy (Eugene, Oreg.), 2 (1987) 40.

[18] D.R. Porterfield, A. Campion, Journal of the American Chemical Society, 110 (1988) 408.

[19] M. Fujiwara, H. Hamaguchi, M. Tasumi, Applied Spectroscopy, 40 (1986) 137.

[20] E.N. Lewis, P.J. Treado, I.W. Levin, Applied Spectroscopy, 47 (1993) 539.

[21] G.J. Gervasio, M.J. Pelletier, At-Process, 3 (1997) 7.

[22] B. Marquardt, Micro Instrumentation (2007) 211.

[23] J. Hammock, P.J. Cong, E. Bergles, W. Yang, Spectroscopy (2010) 13.

[24] C. Engert, V. Deckert, W. Kiefer, S. Umapathy, H. Hamaguchi, Applied Spectroscopy, 48 (1994) 933.

[25] C. Engert, T. Michelis, W. Kiefer, Applied Spectroscopy, 45 (1991) 1333.

[26] K. Yuzaki, H.-o. Hamaguchi, Journal of Raman Spectroscopy, 35 (2004) 1013.

[27] S.C. Denson, C.J.S. Pommier, M.B. Denton, J. Chem. Educ., 84 (2006) 67.

[28] S. Kaminaka, T. Ito, H. Yamazaki, E. Kohda, H.-o. Hamaguchi, Journal of Raman Spectroscopy, 33 (2002) 498.

[29] M.L. Lewis, I.R. Lewis, P.R. Griffiths, Applied Spectroscopy, 58 (2004) 420.

[30] J.R. Gilchrist, J. Rebello, D. Lanzisera, J. Noonan, Laser Focus World, 36 (2000) 149.

[31] B. Chase, Y. Talmi, Applied Spectroscopy, 45 (1991) 929.

[32] J. Barbillat, E. Da Silva, J.L. Hallaert, Journal of Raman Spectroscopy, 24 (1993) 53.

[33] J. Barbillat, E. Da Silva, Spectrochim. Acta, Part A, 53A (1997) 2411.

[34] S. Barsberg, P. Matousek, M. Towrie, Macromolecular Bioscience, 5 (2005) 743.

[35] A.-M. Saariaho, A.-S. Jaaskelainen, M. Nuopponen, T. Vuorinen, Applied Spectroscopy, 57 (2003) 58.

[36] U.P. Agarwal, R.S. Reiner, Journal of Raman Spectroscopy, 40 (2009) 1527.

[37] U.P. Agarwal, Planta, 224 (2006) 1141.

[38] M. Schmidt, A.M. Schwartzberg, A. Carroll, A. Chaibang, P.D. Adams, P.J. Schuck, Biochemical and Biophysical Research Communications, 395 (2010) 521.

[39] K.L. Larsen, S. Barsberg, Journal of Physical Chemistry B, 114 (2010) 8009.

[40] C.-J. Shih, E.A. Smith, Anal. Chim. Acta, 653 (2009) 200.

[41] E.D. Schmid, R.D. Topsom, Journal of the American Chemical Society, 103 (1981) 1628.

[42] H.J. Bernstein, G. Allen, Journal of the Optical Society of America, 45 (1955) 237.

[43] X.-F. Sun, H. Wang, G. Zhang, P. Fowler, M. Rajaratnam, Journal of Applied Polymer Science, 120 (2011) 3587.

[44] F.E. Brauns, The Chemistry of Lignin Supplement Volume, Academic Press Inc., London, 1960.

Figure 1. Instrument schematic of the 1064 nm dispersive multichannel Raman spectrometer. The 1064 nm laser is focused with a plano-convex lens (L1) onto a sample. The Raman scatter is then collected with a plano-convex lens (L2) and focused onto the spectrometer slit with a plano-convex lens (L3). A holographic notch filter (HNF) was used to filter out Rayleigh scattering. The spectrometer is equipped with a 1024-multichannel InGaAs detector. The helium-neon laser is aligned co-linearly with the 1064 nm laser using a dichroic mirror to aid in instrument alignment.

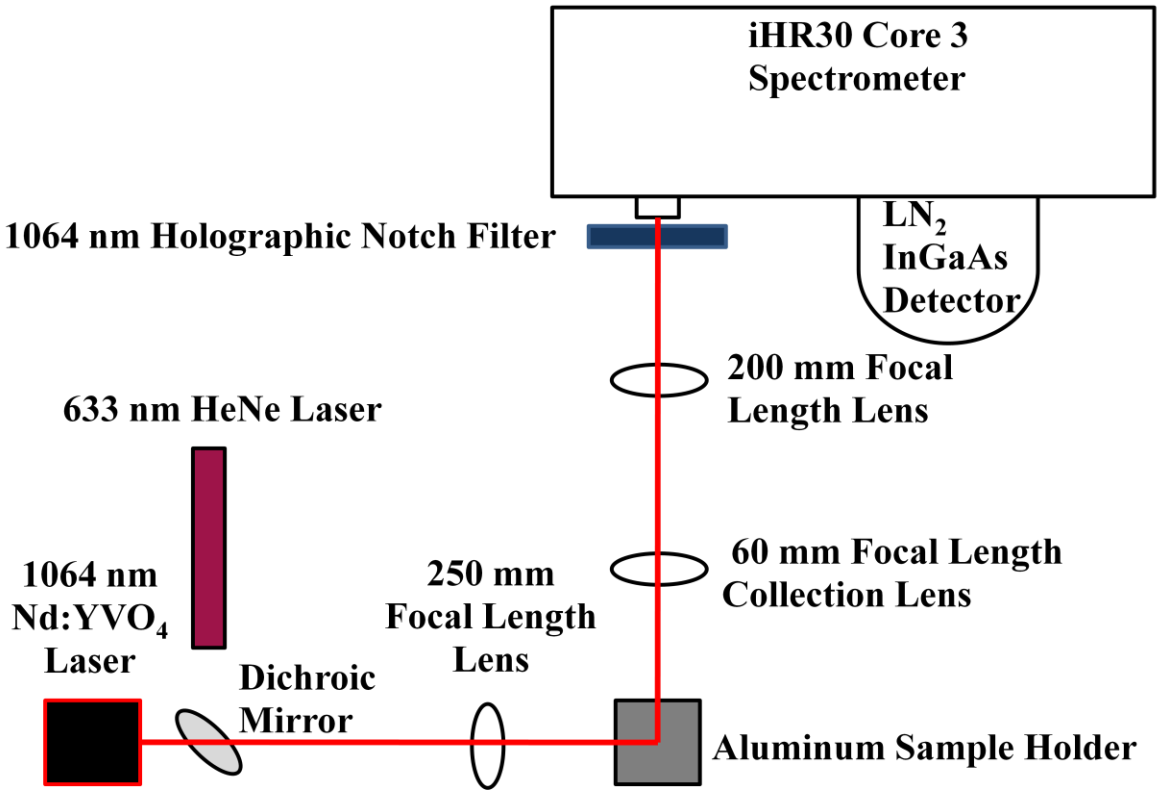


Figure 2. (A) Background subtracted Raman spectra of 50 mg mL^{-1} lignin in methanol obtained using a dispersive 785 nm (black) or 1064 nm (green) spectrometer. The 785 nm excitation spectrum has been divided by 60. (B) 1064 nm excitation lignin spectrum shown in Figure 2A plotted on a smaller scale. For other experimental parameters see Experimental Section and Supporting Information Table S1.

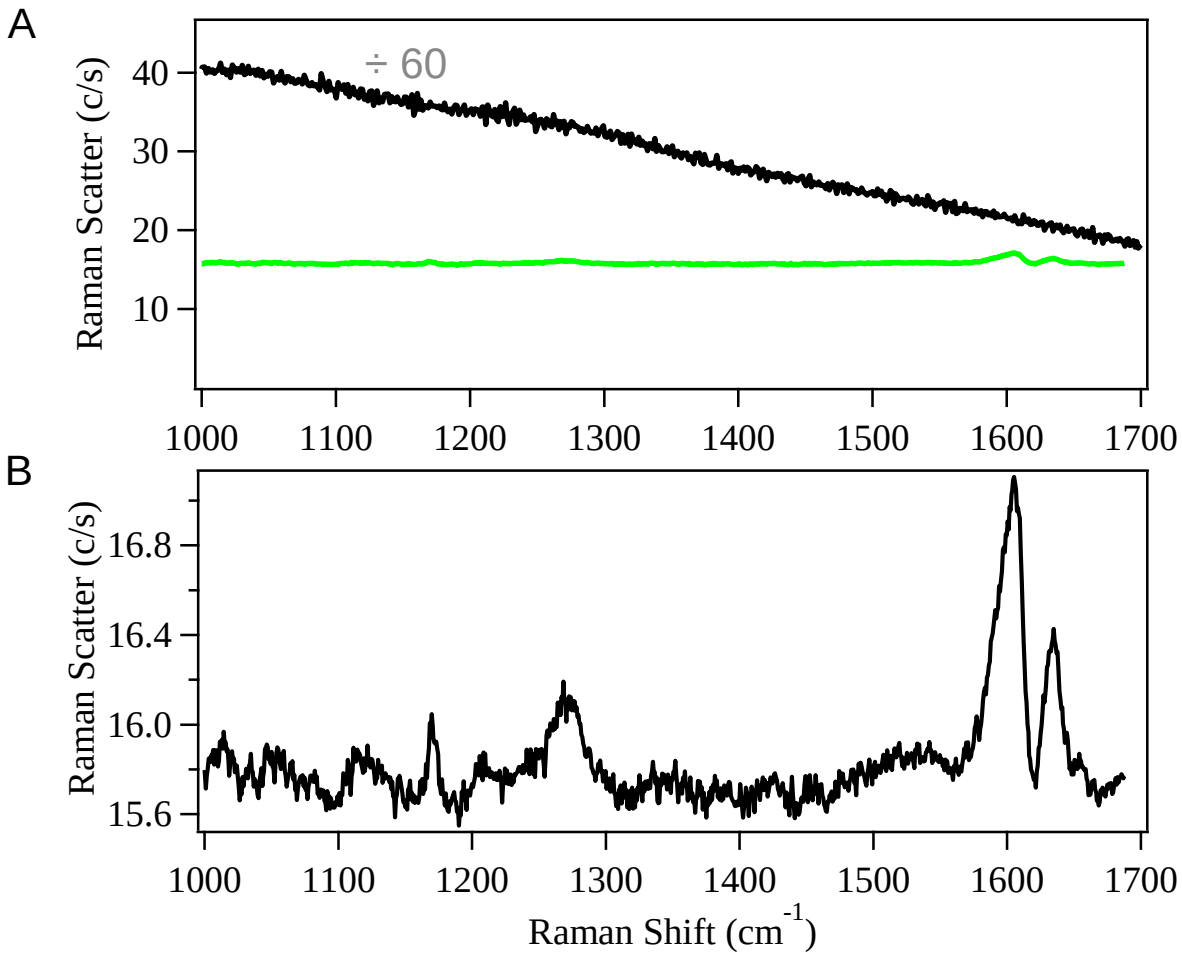


Figure 3. Plots of the measured S/N ratios as a function of total analysis time for the FT Raman instrument (gray symbols, 480 mw) or the dispersive 1064 nm Raman instrument (black symbols, triangle varying integration times with 1 accumulation, diamond 60 second integration time and varying number of accumulations, and circle 150 second integration time with varying number of accumulations, 280 mW). The data are fit (solid line) to a curve with total analysis time$^{1/2}$ dependence for the dispersive 1064 nm Raman data or a linear fit for the FT Raman data to guide the eye. Inset shows an expanded view at short total analysis times and linear fits for all data sets.

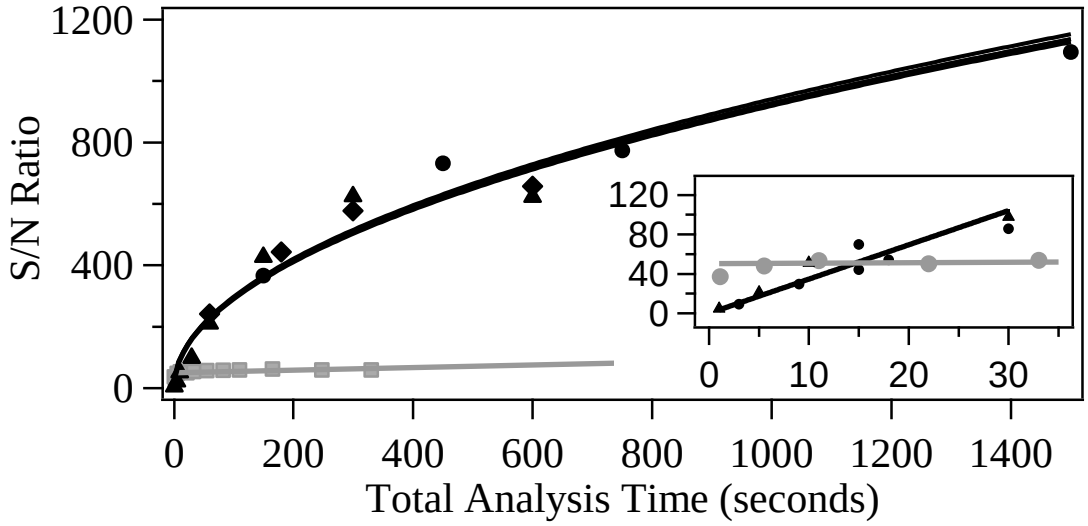
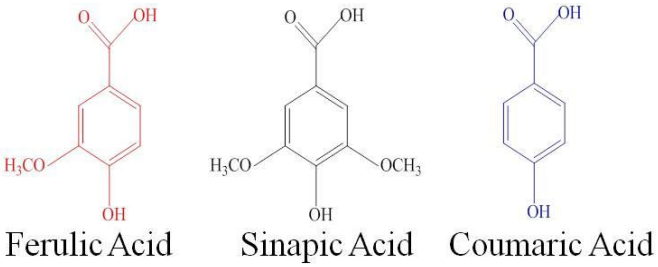


Figure 4. Raman spectra of model plant cell wall lignin monomers ferulic acid (G), sinapic acid (S) and coumaric acid (H) at 50 mg mL^{-1} in 1,4-dioxane obtained with the dispersive 1064 nm Raman (solid spectra, signal divided by 4) and FT Raman (dotted spectra) instrument. Spectra have been offset for clarity. The figure insets show an expanded view of the 1580 to 1640 cm^{-1} first derivative spectra with all spectra overlaid (G black, S light gray, H dark gray).



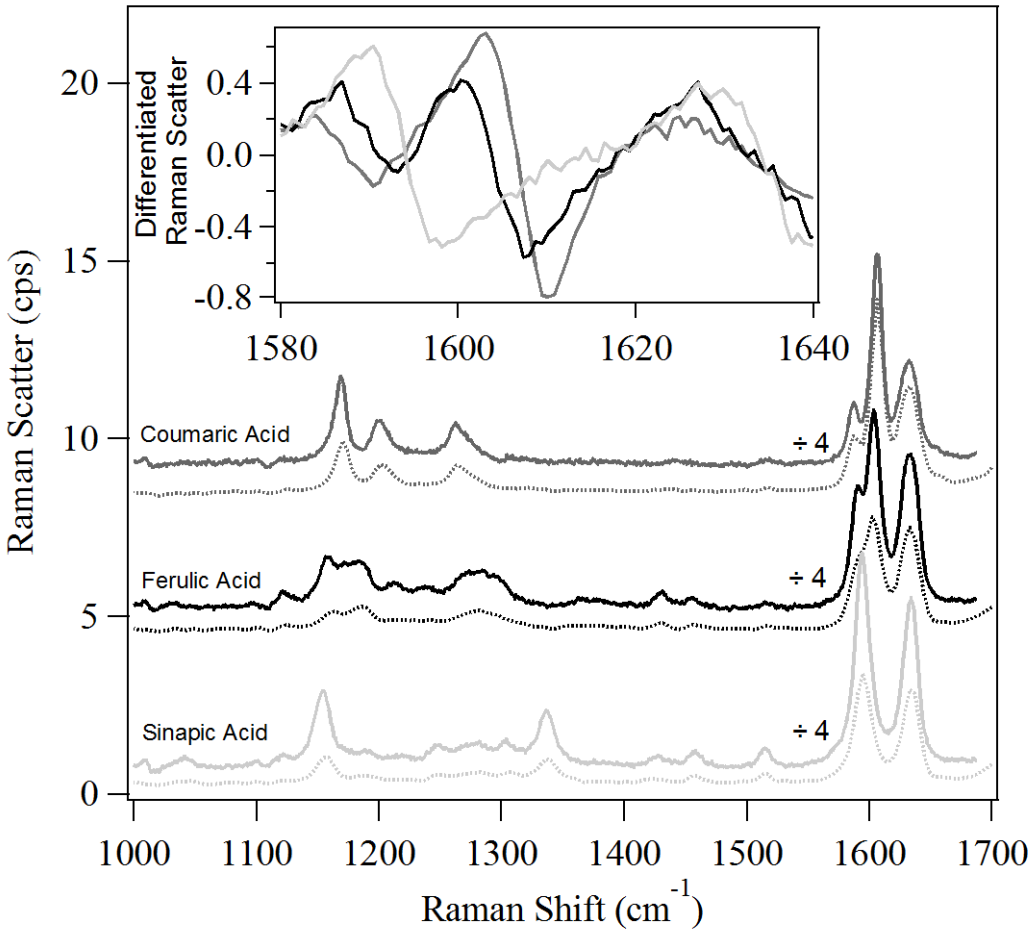


Figure 5. Raman spectra of model lignin monomers mixtures (G-S-H%: 5-90-5 blue; 51-40-9 green; 25-25-50 red) overlaid with the Raman spectrum of 50 mg mL^{-1} lignin in 1,4-dioxane (thick black) obtained with the dispersive 1064 nm Raman instrument.

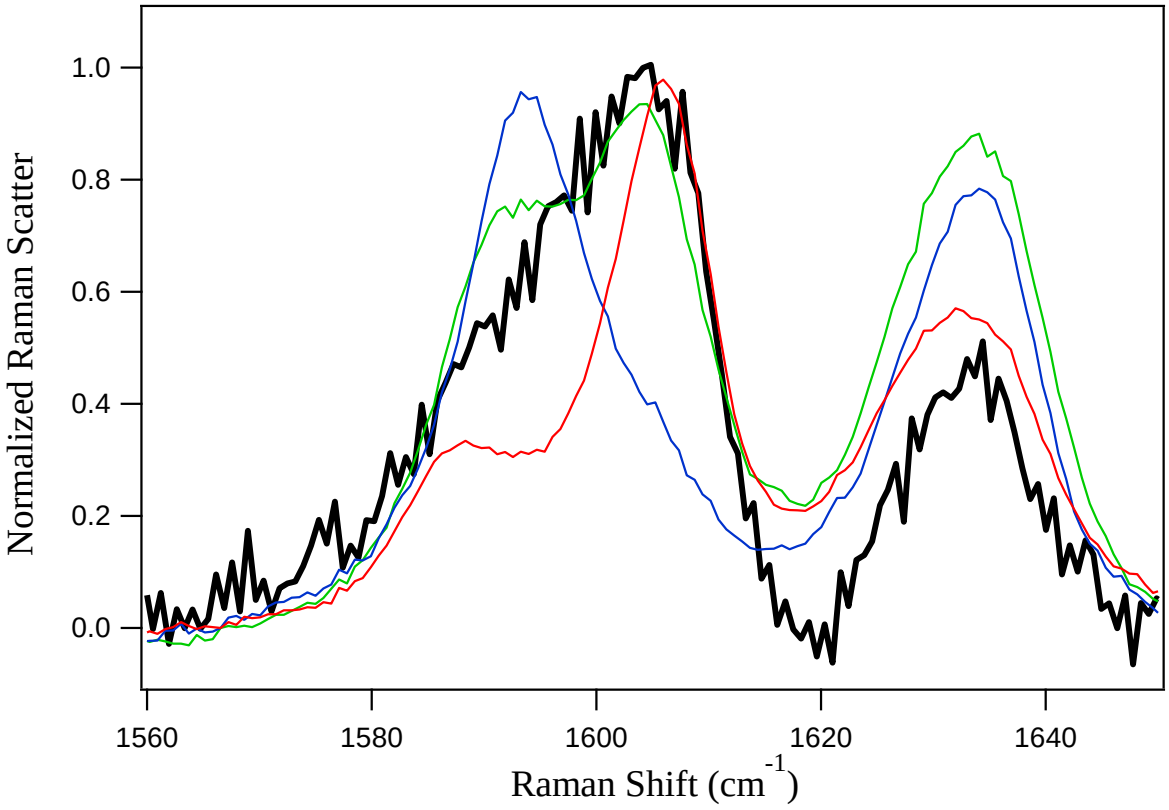


Figure 6. (A) PLS scores plot of 36 model lignin monomers, monomer mixtures and hydrolytic lignin extracted from sugarcane (red). The monomer composition is listed G-S-H. The indicated grouping is used to color code the first derivative spectra in (B). The sugarcane lignin falls in the region comprised by plant samples high in G, S and H. (B) First derivative Raman spectra of 33 model lignin monomers mixtures. The monomer compositions are listed from top to bottom, at 1600 cm^{-1}, in Supplemental Information Table S4. The three monomer mixtures not circled in the PLS scores plot are omitted in Figure 6B.

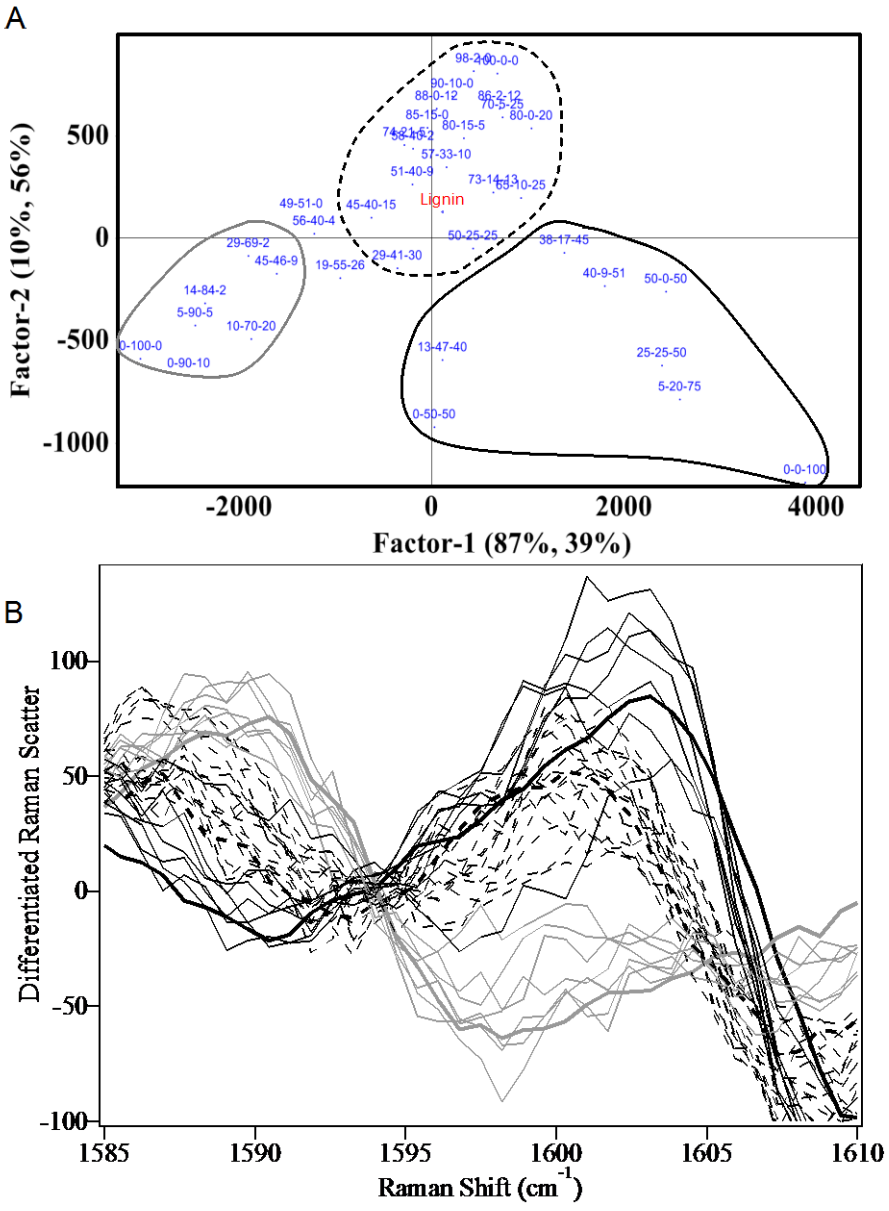
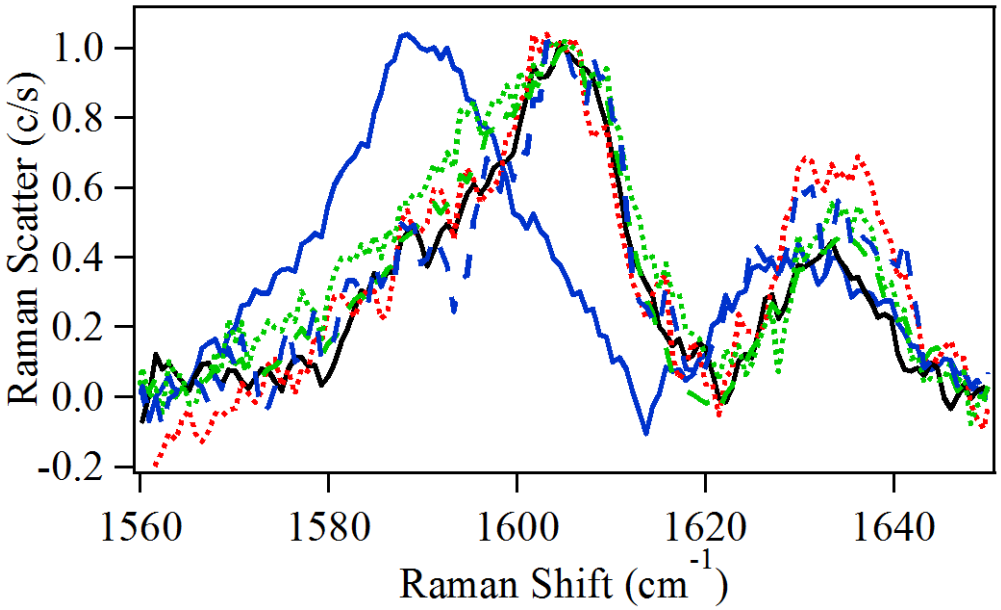


Figure 7. Raman spectra of sugarcane lignin in different solvents: 100 mM NaOH (solid blue), ethanol (solid black), 1,4-dioxane (dotted red), acetone (dotted green), methanol (dashed green), isopropanol (dashed blue). Spectra are normalized to the highest intensity in the shown spectral region. All spectra exhibit nearly complete overlap except the spectrum in 100 mM NaOH.



SUPPLEMENTARY DATA

1064 NM DISPERSIVE MULTICHANNEL RAMAN SPECTROSCOPY FOR THE ANALYSIS OF PLANT LIGNIN

Matthew W. Meyer,[1] Jason S. Lupoi,[1] Emily A. Smith

[1] Both authors contributed equally to this work.

Supplemental Table S1. Specifications and experimental parameters for the 1064 nm dispersive multichannel Raman instrument.

Detector Quantum Efficiency (%)[a]	82-85
Detector Geometry	1024 x 1 elements
Detector Element Size (µm)	25 x 250
Grating Efficiency (%)[b] 300 grooves/mm grating 600 grooves/mm grating 950 grooves/mm grating	 68 - 55 70 - 60 67 - 48
Spectral Resolution[c] / Spectral Coverage (cm^{-1}) 300 grooves/mm grating 600 grooves/mm grating 950 grooves/mm grating	 13 / 1715 7 / 792 4 / 411
Excitation Power at Sample (mW)	280
Beam Area (μm^2)	95

a 1064 to 1350 nm (0 to 1990 cm^{-1} Raman Shift) spectral range

b For spectral range: 600 grooves/mm grating 578 to 1370 cm^{-1}; 300 grooves/mm grating 144 to 1859 cm^{-1}; 950 grooves/mm grating 784 to 1195 cm^{-1}

c 150 µm slit

Supplemental Table S2. Raman peak assignments for sinapic acid (SA), ferulic acid (FA), coumaric acid (CA), and hydrolytic lignin in 1,4 dioxane.

Peak (cm^{-1})	Assignment
SA	
1154	OCH_3 Out of Phase Rock, Aromatic In Phase Bend
1337	Aromatic COC Stretch
1515	Aromatic In Phase Bend, CO Stretch
1594	Aromatic Ring Stretch (v_{8a} and v_{8b})[a]
1634	Substituent C=C
FA	
1158	OCH_3 Out of Phase Rock, Aromatic In Phase Bend
1589	Aromatic Ring Stretch (v_{8a})
1604	Aromatic Ring Stretch (v_{8b})
1633	Substituent C=C
CA	
1168	Hydroxyl COH Bend, Aromatic Ring Stretch
1200	Methoxy In Phase Wagging
1263	Aromatic Ring Bend, Methoxy CO Stretch
1588	Aromatic Ring Stretch (v_{8a})
1606	Aromatic Ring Stretch (v_{8b})
1632	Substituent C=C
Lignin	
1170	Hydroxyl COH Bend, Aromatic Ring Stretch
1270	Aromatic Ring Bend, Methoxy CO Stretch
1591	Aromatic Ring Stretch (v_{8a})
1604	Aromatic Ring Stretch (v_{8b})
1634	Substituent C=C

a Wilson numbering system

Supplemental Table S3. Location of Raman peak maxima as determined by a multipeak fit of spectra for the lignin model monomers in 1,4-dioxane or lignin in the indicated solvent.

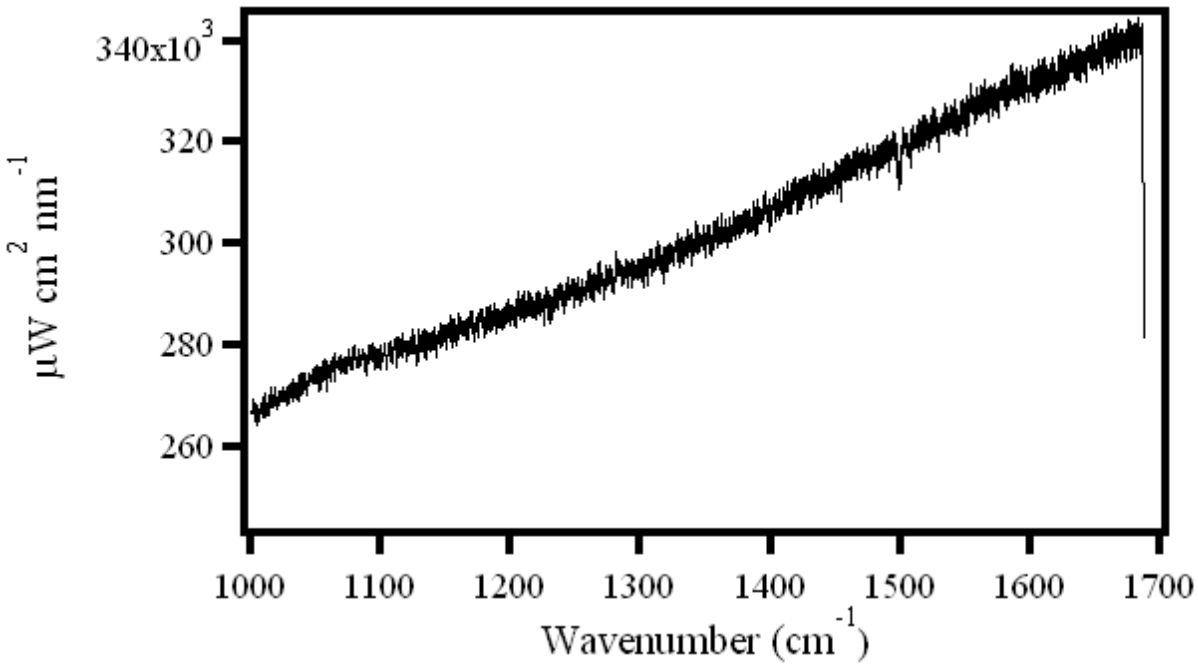
Lignin Model Monomer/Monomer Mixture	**Phenyl CC Stretch ($8a^a$, cm^{-1})[b]**	**Phenyl CC Stretch ($8b^a$, cm^{-1})[b]**	**Substituent C=C**	**Other Substituent**
Sinapic Acid (S)	**1594.5 ± 0.1**	**1594.5 ± 0.1**	1633.5 ± 0.1	1154.1 ± 0.3
Coumaric Acid (H)	1588.6 ± 0.2	**1606.10 ± 0.05**	1632.3 ± 0.1	1168.4 ± 0.1
Ferulic Acid (G)	1589.1 ± 0.2	**1603.9 ± 0.1**	1632.90 ± 0.07	
Lignin in indicated solvent				
1,4-dioxane	1591 ± 5	**1605.5 ±0.8**	1633.9 ± 0.4	
Ethanol	1590 ± 1	**1605.2 ± 0.5**	1632.8 ± 0.3	1168.4 ± 0.3
Methanol	1595 ± 2	**1606.3 ±0.2**	1634.2 ± 0.6	1169.7 ± 0.5
Acetone	1590 ± 10	**1605.5 ± 0.8**	1634.8 ± 0.6	
Isopropanol	1590 ± 3	**1605.9 ±0.6**	1632.6 ±0.5	
200mM NaOH aqueous	**1589.8 ± 0.3**	**1589.8 ± 0.3**	1631.3 ± 0.6	1167.2 ± 0.4

a Wilson numbering system notation
b Bold indicates most intense peak between v_{8a} and v_{8b}

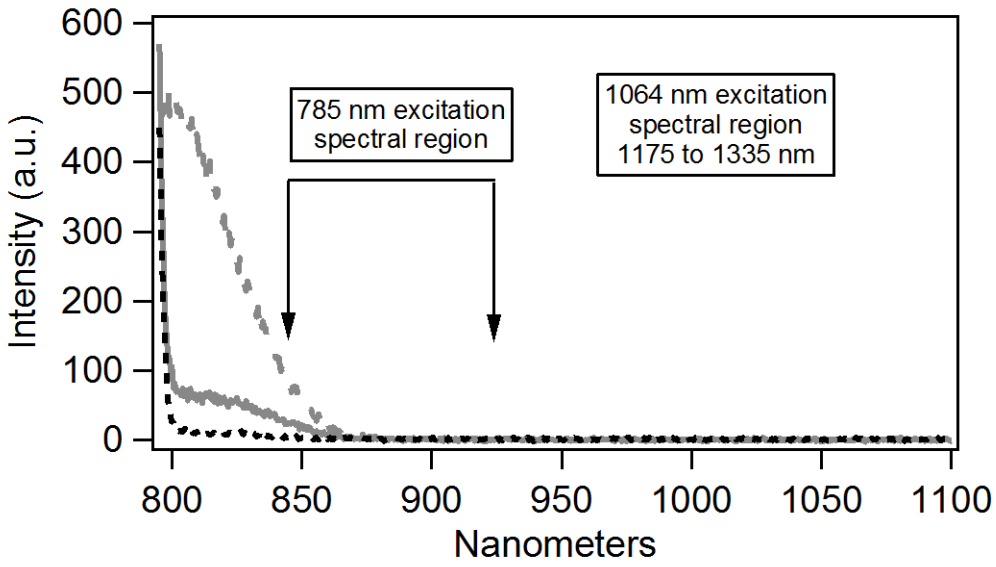
Supplemental Table S4. Monomer composition of the first derivative Raman spectra from top to bottom (at 1600 cm^{-1}). Numbers are percentage listed G, S, H.

Grey	Black	Dashed Black
45, 46, 9	25, 25, 50	57, 33, 10
29, 69, 2	40, 9, 51	90, 10, 0
10,70, 20	50, 0, 50	80, 0, 20
5, 90, 5	38, 17, 45	98, 2, 0
14, 84, 2	5, 20, 75	100-0-0
0, 90, 10	0, 50, 50	80, 15, 5
0-100-0	13, 47, 40	70, 5, 25
	0-0-100	65, 10, 25
		88, 12, 0
		85, 15, 0
		73, 14, 13
		74, 21, 5
		50, 40, 2
		86, 2, 12
		50, 25, 25
		51, 40, 9
		29, 41, 30
		45, 40, 15

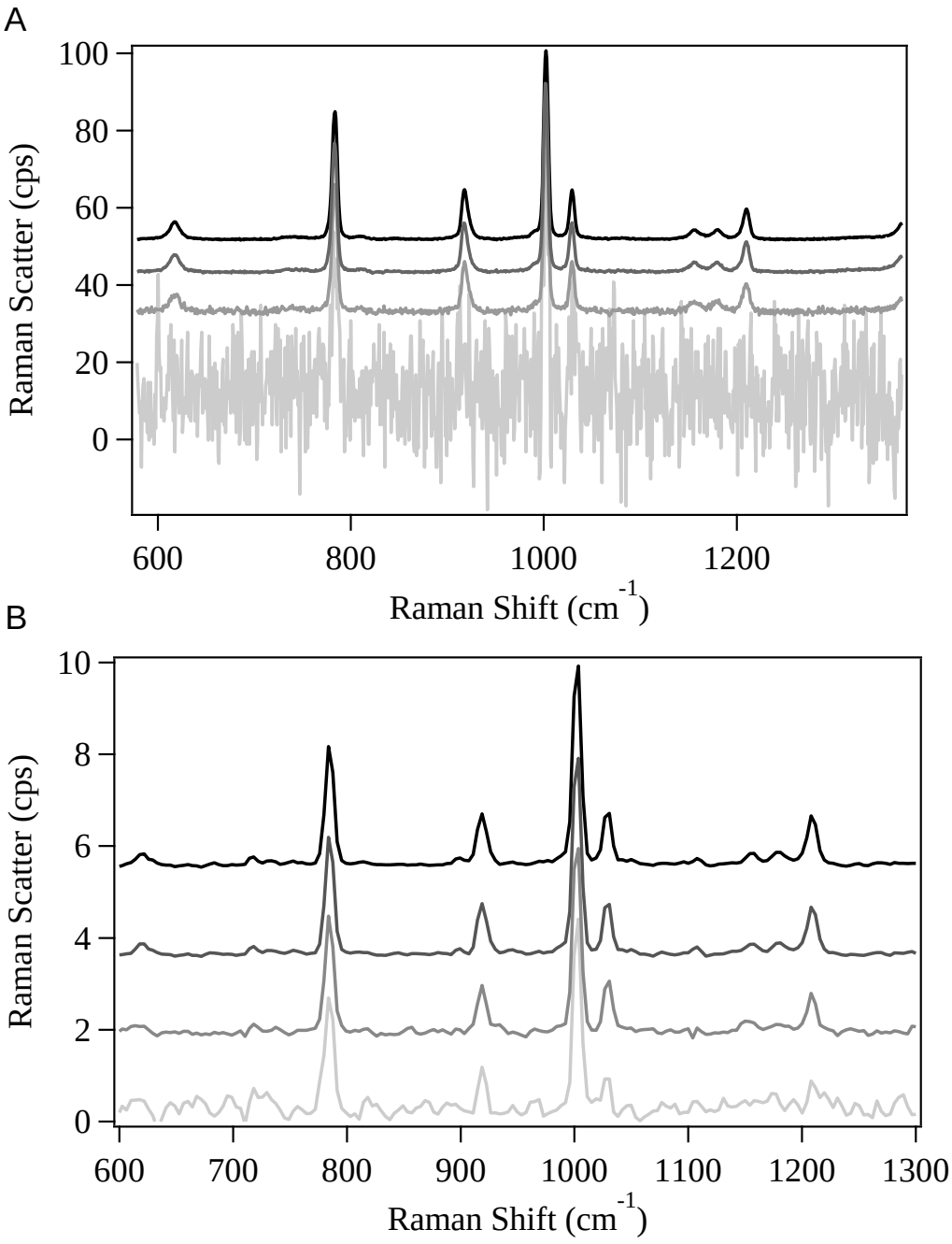
Supplemental Figure S1. Spectrum of a tungsten calibration lamp collected using the 1064 nm dispersive multichannel Raman spectrometer. The spectrum has been corrected for the varying lamp output in this spectral region, and can be used to intensity correct the Raman spectra for varying instrument throughputs at different wavenumbers.



Supplemental Figure S2. Fluorescence spectrum of 200 mg mL^{-1} AlexaFlour 750 (solid black) and 50 mg mL^{-1} lignin (solid gray) and a blank spectrum of solvent (water, dotted black). Excitation wavelength: 785 nm.



Supplemental Figure S3. Raman spectra of a toluene/acetonitrile mixture (1:1 v/v) using (A) 1064 nm dispersive and (B) 1064 nm FT Raman spectrometers. From bottom to top the spectra were collected with: (A) 1 second integration, 30 second integration, 150 second integration, and 300 second integration all with 1 accumulation; (B) a single scan, 10 scans, 150 scans, and 300 scans.



CHAPTER 4
CHARACTERIZATION OF WOODY AND HERBACEOUS BIOMASSES LIGNIN COMPOSITION WITH 1064 NM DISPERSIVE MULTICHANNEL RAMAN SPECTROSCOPY

(A paper published in *Applied Spectroscopy*, 66(8): 901-908 (2012) DOI: 10.1366/12-06621)

Jason S. Lupoi and Emily A. Smith

Abstract

Biomass representing different classes of bioenergy feedstocks, including woody and herbaceous species, was measured with 1064 nm Raman spectroscopy. Pine, oak, poplar, kenaf, *miscanthus*, pampas grass, switchgrass, alfalfa, orchard grass and red clover were included in this study. Spectral differences have been identified with an emphasis on lignin guaiacyl and syringyl monomer content, and carotenoid compounds. The interpretation of the Raman spectra was correlated with ^{13}C-nuclear magnetic resonance cross-polarization/magic-angle spinning spectra of select biomass samples. Thioacidolysis quantification of guaiacyl and syringyl monomer composition and the library of Raman spectra were used as a training set to develop a principal component analysis model for classifying plant samples and a principal component regression model for quantifying lignin guaiacyl and syringyl composition. Raman spectroscopy with 1064 nm excitation offers advantages over alternative techniques for biomass characterization including low spectral backgrounds, higher spectral resolution, short analysis times, and non-destructive analyses.

Index Headings

Near Infrared Raman Spectroscopy; Plant Cell Wall; Guaiacyl and Syringyl Lignin; Principle Component Analysis; Principle Component Regression

Introduction

Lignocellulosic biomass is widely considered to be one solution to relinquishing the world's dependence upon petroleum-based fuels and chemicals. Recent estimates indicate that the annual biomass surplus available from forest and agricultural crops is approximately 1.5 billion tons.[1] In order to use this biomass as a feedstock for several downstream applications, it must be ground and its composition and structure characterized. Common methods used to characterize biomass composition are time consuming and laborious. Current methodologies include wet chemistry techniques such as acidolysis,[2] thioacidolysis,[3] and nitrobenzene oxidation;[4] and instrumental techniques such as high performance liquid chromatography,[5] pyrolysis gas chromatography/mass spectrometry,[6-8] Fourier-transform (FT) infrared spectroscopy,[9-10] Raman spectroscopy[11-19] and ultraviolet resonance Raman spectroscopy.[20-21] Except for the listed spectral techniques, these methods are destructive and laborious. A simple, rapid instrumental technique that does not require extraction or other wet chemical procedures, and subsequent procedures for analyzing the acquired data are needed for characterizing a variety of lignocellulosic materials.

The cell walls of lignocellulosic plants contain predominantly cellulose, hemicellulose, and lignin. Cellulose is a linear polymer of 1,4-linked β-D-glucopyranosyl units, and is the most abundant biopolymer on Earth. Hemicellulose is a heteropolymer of mixed carbohydrate monomers. The saccharification of cellulose to glucose for subsequent

or simultaneous fermentation to ethanol generates large quantities of hemicellulosic and lignin-based wastes that can be utilized in a wide array of products such as fuels, organic acids, and adhesives.[22]

Lignin is the second most abundant biopolymer on the planet. The primary functions of lignin are to strengthen the plant cell wall by forming lignin-carbohydrate complexes and to resist attack by microorganisms, and lignin plays a key role in water transport by decreasing cell wall permeability.[23-25] Lignin is a three-dimensional network of phenylpropanoid units linked via dehydrogenation reactions.[23-25] The phenylpropanoid units are derived from the hydroxycinnamyl alcohols (coniferyl, sinapyl, and p-coumaryl), and to a lesser degree cinnamaldehydes (coniferaldehyde, sinapaldehyde, and p-coumaraldehyde).[23-25] The phenyl moieties are named guaiacyl (G), syringyl (S), and p-coumaryl (H), respectively. Lignins are often categorized into three classes: gymnosperm (i.e., softwoods), angiosperm (i.e., hardwoods), and herbaceous, partly based on their monomer composition. Gymnosperms consist of predominantly G, angiosperms mostly G and S, and herbaceous plants contain G, S, and H.[23-25] Exceptions to these classifications do exist. The S/G ratio is a key parameter for characterizing the potential of delignification reactions, chemical reactivity and the amount of energy necessary for pulping and bleaching feedstocks.[26-27]

Raman spectroscopy is non-invasive, requires little to no sample preparation, can provide high spectral resolution and does not suffer from broad water adsorption bands, as is the case for infrared spectroscopy. Previous Raman investigations of biomass have focused primarily on woody biomass and FT-Raman instrumentation. The measurement of a variety of hardwoods (e.g., oak, balsa), softwoods (e.g., redwood, pine) and other paper materials

have been measured using FT-Raman spectroscopy.[15] Hard and softwoods have been differentiated by FT-Raman and FT-infrared spectroscopy using variations in spectral peak intensities and unique peak locations.[9] The determination of structural changes brought about by biological and chemical treatments on oak wood and barley straw[16] as well as milled wood lignins[11] were measured by FT-Raman spectroscopy. The authors report a lack of spectral resolution in the case of barley straw attributed to fluorescence from substituted cinnamic acids known to be more prevalent in grasses than woods.[16] Lavine et al. developed algorithms for pattern recognition that enabled the classification of 98 different FT- Raman spectra representing hardwood, softwood, and tropical wood.[18] In these reports no attempts were undertaken to quantify lignin composition from the Raman spectra.

Sun and co-workers have developed a technique for measuring the S/G ratio of a variety of plant materials including eucalyptus, switchgrass, maize and sorghum using FT-Raman spectroscopy and chemometrics.[17] The authors determined spectral regions unique to G, S, or H monomers from the spectra of model lignin compounds. The spectral regions that were used to quantify the lignin S/G ratio also had spectral contributions from polysaccharides. This may explain why the ratios measured by Raman spectroscopy were much higher than those measured by pyrolysis-GC/MS.

Ultraviolet resonance Raman (UVRR) spectroscopy coupled with partial least squares has been employed to analyze lignin model compounds and determine characteristic spectral regions representing G, S, and H.[20-21] Nuopponen et al. characterized 25 tropical hardwoods using UVRR and other analysis techniques.[28] The possibility of photodegradation of the sample with UV excitation is high compared to excitation with infrared wavelengths. Low

laser power and mounting the sample on a rotational stage are precautions taken to limit sample damage. Only vibrations in resonance are enhanced in UVRR, and polysaccharides are reported to contribute little to the spectra.

Raman spectroscopy has rarely been used to measure herbaceous biomass such as perennial grasses, presumably due to the greater complexity of the plant cell walls and a higher extractive content (e.g., chlorophyll, waxes, terpenes, aliphatic acids, etc.) that increases the spectral background. In a previous study, the G/S/H ratio was determined using the Raman spectrum of isolated lignin from sugarcane and partial least squares analysis.[29] It was found that 1064 nm excitation was required to analyze extracted lignin due to the fluorescence generated when these samples were excited with 785 nm light. Since Raman spectral intensities are proportional to the photon frequency to the fourth power, the use of NIR excitation results in less Raman scatter. In the case of biomass analyses, the reduction in the fluorescence background with 1064 nm excitation is required to measure Raman spectra.[30-31]

In the present study, a library of ground plant materials, with emphasis on perennial grasses and other herbaceous feedstocks, was characterized using 1064 nm dispersive multichannel Raman spectroscopy. Differences in Raman spectra for each class of biomass were identified and validated using ^{13}C nuclear magnetic resonance (NMR) cross polarization/magic angle spinning (CP MAS) spectroscopy. G and S lignin was quantified using thioacidolysis, and the data combined with the Raman spectra to develop a principal component analysis (PCA) model for classifying feedstocks and a principal component regression (PCR) model for measuring G and S lignin in ground biomass.

Materials and Methods

Materials. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA), including extracted sugarcane lignin (product #371076). The biomass measured in this study were pine (*Pinus sp.*), oak (*Quercus sp.*), and poplar (*Populus sp.*) woods, *miscanthus* (*Miscanthus giganteus*), kenaf (*Hibiscus cannabinus*) bast and core, pampas grass (*Cortaderia selloana*), orchard grass (*Dactylis glomerata*), red clover (*Trifolium pretense*), switchgrass (*Panicum virgatum*), and alfalfa (*Medicago sativa*). All samples were ground with a Wiley mill and passed through a 1 mm mesh screen, with the exception of oak, pine, and poplar woods, which were analyzed as sawdust. *Miscanthus* was grown in the Hinds research farm in Ames, IA. Switchgrass samples were grown at the South Reynolds research farm in Boone, IA. Kenaf, orchard grass, red clover, and alfalfa were all grown at the Sorenson research farm in Ames, IA. Pampas grass was grown at a private residence in Ames, IA and samples of pine, poplar, and oak were obtained from a lumber yard in Ames, IA.

Pretreatment of Biomass. The pine and oak samples were pretreated using 1% sulfuric acid, and subjected to an enzymatic hydrolysis using a mixture of Accellerase 1500® and Accellerase XY® (Danisco US Inc., Genencor Division, Rochester, NY, USA), and Novozyme 188 (Sigma-Aldrich).[32] Some samples designated extractive-free were analyzed after a Soxhlet extraction using 95% ethanol.[32] All other plant materials were used without pretreatment.

Raman Spectroscopy. The home-built 1064 nm dispersive multichannel Raman instrument used for all Raman measurements was previously described.[29] A 300 groove/mm classically ruled grating provided the highest efficiency and the best trade-off between spectral coverage and spectral resolution for these measurements. Raman spectra of biomass, cellulose and lignin were acquired using 570 mW laser power for 5 minutes, unless otherwise noted below, and were baseline corrected using Grams AI V 8.0 (Thermo Scientific, Waltham, MA, USA). The spectra were imported into IGOR Pro V 6.1 (WaveMetrics Inc., Lake Oswego, OR, USA) for further analysis. The raw biomass spectra were 5-point smoothed using the Savitzky-Golay algorithm. The spectra showing the CH stretching region were acquired for 10 minutes using 570 mW excitation power. To compare the spectra for untreated versus extracted orchard grass and red clover, the raw biomass spectra were collected for 5 minutes using 570 mW excitation power and the extractive-free samples were analyzed at a reduced 450 mW laser power for 5 minutes to prevent sample charring. The spectra of extractive-free red clover and orchard grass have been multiplied by 1.3 to account for differences in laser power.

^{13}C NMR CP MAS. Spectra were acquired on a Bruker Avance II 600 MHz spectrometer (Bruker, Billerica, MA, USA) operating at 150 MHz for carbon. A 4 mm triple resonance probe operating in ^{1}H-^{13}C double resonance mode was used for all acquisitions. The samples were spun at 10 kHz, using a 2.6 μs excitation pulse for ^{1}H, 2 ms CP contact time, and dipolar decoupling at 96 kHz. In order to obtain a sufficient signal to noise ratio, 4000 (*miscanthus* and red clover) or 6000 (pine) scans were acquired for each sample. NMR spectra were baseline corrected using the TopSpin 3.0 software (Bruker). The spectral

regions selected for peak area integration were based upon standard literature ranges.[28] The total carbon signal intensity was determined by summing the values calculated for each integral. Relative intensities were determined by dividing the intensity of each spectral region by the total carbon intensity.

Thioacidolysis and Gas-Chromatography/Mass Spectrometry (GC/MS). The modified thioacidolysis technique of Robinson and Mansfield was employed with some modifications.[33] The procedure is outlined in the Supplemental Information. One µL of each thioacidolysis sample was analyzed using an Agilent 6890 gas chromatograph (Santa Clara, CA, USA) fitted with a 30-m, 0.25 mm diameter, J&W DB-5ms column (Agilent). A 10:1 split ratio was used for injections using helium as the carrier gas (1 mL min^{-1} flow rate). The inlet temperature was 260°C. The oven temperature profile was as described in Robinson and Mansfield.[33] A Waters Micromass® GCT TOF (Milford, MA, USA) run in electron ionization mode was used for mass detection.

The peaks indicative of G or S lignin degradation products were described by Lapierre et al.[3] These peaks were quantified using the summed total ion current (TIC) of the respective G and S peaks, adjusted to the response factor of the internal standard. H lignin peaks were much weaker in intensity, and showed an abundance of co-eluted reaction products, requiring the use of only m/z 239 for specificity to H lignin. This decreased the total counts, and prohibited measurements of H lignin for some samples. The G and S percentages used in this study were obtained from the ratio of the individual monomer TIC (G or S) to the total summed TIC (G + S).

Chemometrics. Principle component analysis (PCA) and principle component regression (PCR) were performed using The Unscrambler X (Camo Inc., Oslo, Norway). The data were mean-centered, and the model was randomly cross-validated. The NIPALS algorithm was used to calculate the principle components. For PCA, the training set had 11 rows representing the 11 biomass samples, and 76 columns representing the spectral region of interest (1581 to 1702 cm^{-1}, 74 spectral coordinates) plus the G and S values quantified by thioacidolysis. One sample was then removed from the data matrix to be treated as an unknown. The remaining 10 samples were used in the calibration set to classify the unknown. This step was performed for each sample to ensure the validity of the model. For PCR, the training set contained 10 rows representing all biomass spectra except red clover, which was rejected as an outlier, and 17 columns representing the spectral region from 1590-1614 cm^{-1} plus the G and S values. The calibration model was used to quantify G and S lignin from an unknown data matrix that contained a separate set of spectra for each feedstock.

Results and Discussion

The purpose of this study is to use Raman spectroscopy to characterize the lignin content in woody and herbaceous biomass, and compare the lignin monomer composition measured with Raman spectra and a laborious but accepted analysis method. A 1064 nm excitation wavelength was used for all Raman studies in order to limit background that swamps the Raman signal at lower wavelengths. The types of biomass included in this study are shown in Table I, along with representative literature compositions of the total dry matter

for each. Reported biomass compositions vary depending upon plant age and analysis method, among other factors.

Characterization of cell wall biopolymers with Raman spectra requires knowledge of spectral regions relatively free from interference of other biopolymers. Figure 1 shows the spectra of microcrystalline cellulose and a commercial lignin extracted from sugarcane. The peaks with the highest intensities in the cellulose spectrum are due to C-C and C-O stretches (1094 and 1119 cm^{-1}).[34] The peak at 896 cm^{-1} has previously been assigned to H-C-C and H-C-O bending, and has also been reported to correspond to amorphous cellulose content.[35] These cellulose peaks are relatively free of lignin spectral interference; however, the peaks for the mixed polysaccharide hemicellulose will have significant overlap with those shown for cellulose.[12] Cellulose does not contribute to the Raman spectrum in the region where the dominant lignin aromatic skeletal vibrations (~1600 cm^{-1}) or substituent C=C stretch (~1634 cm^{-1})[11, 14] are assigned. The spectra of model lignin monomers: ferulic (G), sinapic (S), and coumaric (H) acid, show unambiguous peak maxima (Supporting Information Figure S1). The aromatic skeletal vibrations (Wilson notation v_{8a} and v_{8b}) are composed of two partially resolved peaks in ferulic (1591 and 1601 cm^{-1}) and coumaric (1588 and 1606 cm^{-1}) acids, while the two modes are degenerate in sinapic acid (1596 cm^{-1}). Other lignin and cellulose spectral assignments are listed in Table II.

The Raman spectra of several biomass samples representing woody and herbaceous species are plotted in Figure 2A-C. The averages of all spectra shown in Figure 2A, B or C are plotted in panel D. These spectra show characteristic Raman peaks for polysaccharides and lignin. Other species that contribute to the Raman spectra are proteins, extractable

compounds and ash. The biomass spectra have been normalized to the 1094 cm^{-1} polysaccharide peak, and grouped according to spectral similarities. Normalizing the spectra is required since differences in the biomass particle size (Supporting Information Figure S2), despite similar grinding procedures, affect the Raman signal. Spectral differences and an interpretation of these differences amongst woody and herbaceous samples are outlined below.

Characterization of Woody Species by Raman Spectroscopy. Pine, oak, poplar and kenaf represent the woody plants measured in this study, and their Raman spectra are shown in Figure 2A. Averages of the reported compositions for these plant species are ~47% cellulose, ~26% hemicellulose and ~21% lignin by dry mass. Pine is a gymnosperm and oak and poplar are angiosperms. Kenaf is considered an herbaceous angiosperm; however, the bast or bark, and the woody base of the plant are similar to hardwoods. Pine has primarily G lignin,[24, 36] and its aromatic skeletal vibration is located at 1604 cm^{-1}. Oak and poplar woods have G and S lignin,[24, 36] and exhibit aromatic skeletal vibrations shifted to lower energies. Coniferyl and/or sinapyl alcohols are assigned to the peak at 1656 cm^{-1}. The 1270 cm^{-1} ring deformation and C-O stretch is reported to be a marker for G content, with higher intensity for softwoods than hardwoods.[14] Similarly, the peak at 1338 cm^{-1} is often associated with S lignin.[17] However, this peak is present in pine, which has been shown to be lacking or low in S units.[37] Caution must be exercised when characterizing lignin spectral contributions below approximately 1600 cm^{-1} in biomass since cellulose also has vibrational modes in this region (Figure 1).[9, 15, 21, 38] For the biomass samples included in this study, the CH stretching region did not produce unique features (Figure 2 insets).

Characterization of Herbaceous Species with Low Extractable Content by Raman Spectroscopy. The spectra for *miscanthus*, pampas grass, and switchgrass are presented in Figure 2B. The Raman spectra of the untreated and extractive-free biomass were essentially identical, as was measured for the woody biomass (data not shown). Therefore, extractable compounds have negligible contributions to the Raman spectra. The average reported cellulose, hemicellulose and lignin compositions for these plants (~42%, ~24%, ~19%, respectively) are also similar to those reported for the woody biomass. These herbaceous angiosperms are reported to contain p-coumaryl units in addition to G and S precursors.[24, 36] The pampas grass, switchgrass and *miscanthus* spectra show a pronounced shift from pine, oak, kenaf and poplar in the substituent C=C stretching region, which appears at ~1630 cm^{-1} for the former and ~1660 cm^{-1} for the latter. Agarwal et al. have assigned the ~1630 cm^{-1} peak to coniferaldehyde/sinapaldehyde and the ~1660 cm^{-1} peak to coniferyl alcohol/sinapyl alcohol.[11] The spectra in Figure 2B show a broad 1702 cm^{-1} carbonyl peak, which is much lower in intensity for the woody species. This is assigned to proteins and acyl groups in lignin.

Characterization of Herbaceous Species with High Extractable Content by Raman Spectroscopy. The Raman spectra of red clover, alfalfa, and orchard grass are presented in Figure 2C. Compared to the other biomass discussed above, the average reported cellulose (~23%), hemicellulose (~17%) and lignin (~9%) compositions by dry mass reported for red clover, alfalfa and orchard grass are significantly lower. Comparing their spectra to those in Figure 2A and 2B, three significant differences in the Figure 2C spectra are an intense peak at 1150 cm^{-1}, a unique peak at 1526 cm^{-1} and more complex lignin peaks in the 1600 cm^{-1}

region. The Raman spectra of orchard grass and red clover obtained after extracting the biomass with 95% ethanol reveals that the 1526 cm^{-1} peak is a measure of the amount of ethanol extractable compounds in the biomass (Figure 3). The peak intensity at 1150 cm^{-1} decreases by 2 to 3.5x after extraction. However, this peak also has contributions from non-extractable components. Vibrational modes between 1500-1550 and 1150-1170 cm^{-1} have been assigned to C=C and C-C stretches of tetraterpenes, such as carotenoids.[39] Another unique spectral feature shown in Figure 2C is the detection of three distinct peaks between 1555-1685 cm^{-1}, with peak maxima that decrease up to 4x after ethanol extraction. The peaks at 1630 and 1660 cm^{-1} indicate that these grasses have substantial coniferaldehyde/sinapaldehyde and coniferyl alcohol/sinapyl alcohol content; and these grasses are also known to contain cinnamic acid esters that will contribute to these peaks.[40-41]

NMR Characterization of Pine, Miscanthus, and Red Clover. ^{13}C NMR CP MAS spectra of one biomass sample shown in Figure 2A, B or C were collected to validate interpretations of the Raman spectra. Figure 4 shows the ^{13}C NMR spectra of pine (A), *miscanthus* (B) and red clover (C). In contrast to the 5 minutes required to collect a Raman spectrum, each NMR spectrum required 3 to 12 hours to collect. ^{13}C NMR provides semi-quantitative chemical compositions. The integrated NMR peak area is a measure of the relative carbon abundance for each carbon type (Figure 4D). Red clover has 8.5% carbon assigned to aliphatic C-H, indicative of high carotenoid concentrations compared to 1.3% or 0.5% for pine or *miscanthus*, respectively. The higher carotenoid content of red clover measured by NMR is consistent with the Raman spectroscopy measurement discussed above. Based on the NMR data, pine has the highest lignin abundance. Comparing the intensity of

the ~1605 cm^{-1} Raman peak is not necessarily indicative of the biomass' lignin abundance. Raman spectra of model lignin monomers show that the intensity of the ~1605 cm^{-1} aromatic ring breathing modes varies for G, S and H (Supplemental Information Figure S1). Therefore, the intensity of the ~1605 cm^{-1} lignin peak is determined by the amount of lignin present in the sample and its monomer composition.

Classification of Lignin by Raman Spectroscopy and Principle Component Analysis (PCA). A qualitative lignin description can be obtained using Raman spectra and PCA. An example of the PCA analysis using only spectral regions with predominant lignin peaks and all spectra in the training set is shown in Figure 5A. Of all the samples included in this study, pine is unique in that it is predominantly composed of G lignin. Not surprisingly, the PCA plot generated using all biomass spectra in the calibration set shows pine in the lower right hand corner away from the other samples. Diagonally from pine are plants high in S lignin such as kenaf, oak, and poplar. The grasses are clustered in the center to top-right region of the plot. Figure 5B shows the projection of each biomass sample when it is systematically removed from the training set. All feedstocks were accurately classified by the model, except pine. This is expected given pine's unique lignin composition among the samples used in this study.

Quantification of Lignin Monomer Composition by Raman Spectroscopy and Principle Component Regression (PCR). The quantification of G and S lignin was determined by thioacidolysis followed by GC/MS analysis of the products (Table III). The measured values were used to generate the calibration required to develop a PCR model for the quantification of lignin composition by Raman spectroscopy. Signal from H lignin was

below the GC/MS detection limit for many samples, and was not included in the PCR analysis. Analysis of the calibration data for outliers led to rejecting red clover from the training set. The root mean standard error obtained using the PCR model was between 2 to 4% (Supplemental Information, Table SI). A comparison of the thioacidolysis quantification values with those predicted using Raman spectra and PCR is shown in Table III. Overall, the PCR model accurately predicted the G and S percentages at the 95% confidence level in a variety of biomass samples. An exception was orchard grass. The G value for orchard grass measured by Raman spectroscopy was 1.8x higher than that determined by thioacidolysis. The rejection of red clover from the training set and inaccurate results obtained for orchard grass leads to the conclusion that the PCR model is not suitable for biomass with a significant composition of extractable compounds. The ratio of the 1605 cm^{-1}/1525 cm^{-1} peak intensity must be greater than 4.5 in order to obtain accurate results with the PCR model. In theory, another PCR model can be developed using a variety of biomass samples with high concentrations of extractable compounds. As discussed above, the negative S lignin value obtain for pine may be the result of its unique lignin composition among the data set. The PCR model was highly robust, as G and S percentages obtained from Raman spectra collected at different laser powers (450-600 mW) were statistically similar.

Conclusions

Raman spectral differences can be used to characterize woody and herbaceous biomass including its lignin monomer composition. The classification of biomass and quantification of G and S lignin was determined using PCA and PCR, respectively. Raman spectroscopy and chemometrics offer several benefits over other analysis techniques: little to

no sample pretreatment or extraction, non-destructive analysis, and a robust model capable of accurately quantifying G and S lignin in a variety of biomass feedstocks. Work is underway to further characterize the Raman spectral features associated with different cell wall polysaccharide components.

Acknowledgements

This research is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences through the Ames Laboratory. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358. Additional funding was provided to JSL by the GAANN fellowship through the Department of Chemistry, Iowa State University. The authors thank Drs. Emily Heaton and Kenneth Moore (Iowa State University, Department of Agronomy) for supplying switchgrass, *miscanthus*, alfalfa, kenaf, red clover, and orchard grass. The authors thank Dr. Sarah Cady and Steve Veysey (Iowa State University, Department of Chemistry, Chemical Instrumentation Facility) for assistance with ^{13}C-NMR CP MAS and GC/MS.

Supplemental Information

The supplemental information includes the thioacidolysis protocol, Raman spectra of lignin monomer standards, optical images of biomass samples, and parameters for the developed PCR model. The supplemental material is available in the online version of the journal, at http://www.s-a-s.org.

References

1. R. D. Perlack, L.L. Wright, A.F. Turhollow, R.L. Graham, B.J. Stokes, D.C. Erbach. "Biomass as a Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply, DOE/GO-102005-2135". Oak Ridge National Laboratory. 2005.
2. J. M. Pepper, P. E. T. Baylis, E. Adler. "Isolation and properties of lignins obtained by the acidolysis of spruce and aspen woods in dioxane-water medium". Can. J. Chem. 1959. 37: 1241-1248.
3. C. Lapierre, B. Monties, C. Rolando, "Thioacidolysis of lignin: comparison with acidolysis". J. Wood Chem. Technol. 1985. 5(2): 277-292.
4. C.-L. Chen. "Nitrobenzene and Cupric Oxide Oxidations". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. Berlin, Germany: Springer-Verlag, 1992. Chapter 6.2, Pp. 301-320.
5. A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, D. Crocker. "Determination of Structural Carbohydrates and Lignin in Biomass". 2011. Golden, CO: National Renewable Energy Lab, NREL/TP-510-42618.
6. F. F. Lopes, F. O. Silverio, D. C. F. Baffa, M. E. Loureiro, M. H. P. Barbosa. "Determination of Sugarcane Bagasse Lignin S/G/H Ratio by Pyrolysis GC/MS ". J. Wood Chem. Technol. 2011. 31(4): 309-323.
7. T. Sonoda, T. Ona, H. Yokoi, Y. Ishida, H. Ohtani, S. Tsuge. "Quantitative analysis of detailed lignin monomer composition by pyrolysis-gas chromatography combined

with preliminary acetylation of the samples". Anal. Chem. (Washington, DC, U. S.). 2001. 73(22): 5429-5435.

8. D. Meier, O. Faix. "Pyrolysis-Gas Chromatography-Mass Spectrometry". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. Berlin, Germany: Springer-Verlag, 1992. Chapter 4.7, Pp. 177-199.

9. P. A. Evans. "Differentiating "hard" from "soft" woods using Fourier transform infrared and Fourier transform Raman spectroscopy". Spectrochim. Acta, Part A. 1991. 47A(9-10): 1441-1447.

10. O. Faix. "Fourier Transform Infrared Spectroscopy". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. Berlin, Germany: Springer-Verlag, 1992, Chapter 4.1, Pp. 83-108.

11. U. P. Agarwal, J. D. McSweeny, S. A. Ralph. "FT-Raman Investigation of Milled-Wood Lignins: Softwood, Hardwood, and Chemically Modified Black Spruce Lignins". J. Wood Chem. Technol. 2011. 31(4): 324-344.

12. U. P. Agarwal, S. A. Ralph. "FT-Raman spectroscopy of wood: identifying contributions of lignin and carbohydrate polymers in the spectrum of black spruce (Picea mariana)". Appl. Spectrosc. 1997. 51(11): 1648-1655.

13. O. Faix. "Investigation of lignin polymer models (DHP's) by FTIR spectroscopy". Holzforschung. 1986. 40(5): 273-280.

14. K. L. Larsen, S. Barsberg. "Theoretical and Raman Spectroscopic Studies of Phenolic Lignin Model Monomers". J. Phys. Chem. B. 2010. 114(23): 8009-8021.

15. R. C. Kenton, R. L. Rubinovitz. "FT-Raman investigations of forest products". Appl. Spectrosc. 1990. 44(8): 1377-1380.

16. D. Stewart, H. M. Wilson, P. J. Hendra, I. M. Morrison. "Fourier-Transform Infrared and Raman Spectroscopic Study of Biochemical and Chemical Treatments of Oak Wood (Quercus rubra) and Barley (Hordeum vulgare) Straw". J. Agric. Food Chem. 1995. 43(8): 2219-2225.

17. L. Sun, P. Varanasi, F. Yang, D. Loque, B. A. Simmons, S. Singh. "Rapid determination of syringyl:guaiacyl ratios using FT-Raman spectroscopy". Biotechnol. Bioeng. 2012. 109(3): 647-656.

18. B. K. Lavine, C. E. Davidson, A. J. Moores, P. R. Griffiths. "Raman spectroscopy and genetic algorithms for the classification of wood types". Appl. Spectrosc. 2001. 55(8): 960-966.

19. R. H. Atalla, U.P. Agarwal, J.S. Bond. "Raman Spectroscopy". In: S.Y. Lin and C.W. Dence, editors. Methods in Lignin Chemistry. Berlin, Germany: Springer-Verlag, 1992, Chapter 4.6, Pp. 162-176.

20. A.-M. Saariaho, D. S. Argyropoulos, A.-S. Jaeaeskelaeinen, T. Vuorinen. "Development of the partial least squares models for the interpretation of the UV resonance Raman spectra of lignin model compounds". Vib. Spectrosc. 2005. 37(1): 111-121.

21. A.-M. Saariaho, A.-S. Jaaskelainen, M. Nuopponen, T. Vuorinen. "Ultraviolet resonance Raman spectroscopy in lignin analysis: determination of characteristic

vibrations of p-hydroxyphenyl, guaiacyl, and syringyl lignin structures". Appl. Spectrosc. 2003. 57(1): 58-66.

22. B. Kamm, M. Kamm. "Principles of biorefineries". Appl. Microbiol. Biotechnol. 2004. 64(2): 137-145.

23. B. L. Browning. "Wood Lignins". In: B.L. Browning, editor. The Chemistry of Wood. New York, NY: Interscience Publishers (A Division of John Wiley and Sons), 1963. Chapter 6, Pp. 249-311.

24. T. Higuchi. "Biosynthesis of Lignin". In: T. Higuchi, editor Biosynthesis and Biodegradation of Wood Components. Orlando, FL: Academic Press, Inc., 1985. Chapter 7, Pp. 141-160.

25. K. V. Sarkanen, H.L. Hergert. "Classification and Distribution". In: K.V. Sarkanen, C.H. Ludwig, editors. Lignins: Occurrence and Formation, Structure, Chemical and Macromolecular Properties, and Utilization. New York, NY: John Wiley and Sons, 1971. Chapter 3, Pp. 43-94.

26. B. H. Davison, S. R. Drescher, G. A. Tuskan, M. F. Davis, N. P. Nghiem. "Variation of S/G ratio and lignin content in a Populus family influences the release of xylose by dilute acid hydrolysis". Appl. Biochem. Biotechnol. 2006. 129-132: 427-435.

27. Y. Tsutsumi, R. Kondo, K. Sakai, H. Imamura. "The difference of reactivity between syringyl lignin and guaiacyl lignin in alkaline systems". Holzforschung. 1995. 49(5): 423-428.

28. M. H. Nuopponen, H. I. Wikberg, G. M. Birch, A.-S. Jaaskelainen, S. L. Maunu, T. Vuorinen, D. Stewart. "Characterization of 25 tropical hardwoods with Fourier

transform infrared, ultraviolet resonance Raman, and 13C-NMR cross-polarization/magic-angle spinning spectroscopy". J. Appl. Polym. Sci. 2006. 102(1): 810-819.

29. M. W. Meyer, J. S. Lupoi, E. A. Smith. "1064 nm dispersive multichannel Raman spectroscopy for the analysis of plant lignin". Anal. Chim. Acta. 2011. 706(1): 164-170.

30. M. Fujiwara, H. Hamaguchi, M. Tasumi. "Measurements of spontaneous Raman scattering with neodymium:YAG 1064-nm laser light". Appl. Spectrosc. 1986. 40(2): 137-139.

31. P. Vitek, E.M.A. Ali, H.G.M. Edwards, J. Jehlicka, R. Cox, K. Page. "Evaluation of portable Raman spectrometer with 1064 nm excitation for geological and forensic applications". Spectrochim. Acta, Part A. 2011. 86: 320-327.

32. C.-J. Shih, J. S. Lupoi, E. A. Smith, "Raman spectroscopy measurements of glucose and xylose in hydrolysate: Role of corn stover pretreatment and enzyme composition". Bioresour. Technol. 2011. 102(8): 5169-5176.

33. A. R. Robinson, S. D. Mansfield. "Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling". Plant J. 2009. 58(4): 706-714.

34. J. H. Wiley, R. H. Atalla. "Band assignments in the Raman spectra of celluloses". Carbohydr. Res. 1987. 160: 113-129.

35. A. J. Michell. "Second derivative FTIR spectra of native celluloses". Carbohydr. Res. 1990. 197: 53-60.

36. C. Lapierre. "Determining Lignin Structure by Chemical Degradations". In: C. Heitner, D.R. Dimmel, J.A. Schmidt, editors. Lignin and Lignans: Advances in Chemistry. Boca Raton, FL: CRC Press, 2010. Chapter 2, Pp. 11-48.

37. J. R. Obst, L. L. Landucci. "The syringyl content of softwood lignin". J. Wood Chem. Technol. 1986. 6(3): 311-327.

38. I. R. Lewis, N. W. Daniel, Jr., N. C. Chaffin, P. R. Griffiths. "Raman spectrometry and neural networks for the classification of wood types". Spectrochim. Acta, Part A. 1994. 50A(11): 1943-1958.

39. H. Schulz, M. Baranska. "Identification and quantification of valuable plant substances by IR and Raman spectroscopy". Vib. Spectrosc. 2007. 43(1): 13-25.

40. R. Calheiros, N. F. L. Machado, S. M. Fiuza, A. Gaspar, J. Garrido, N. Milhazes, F. Borges, M. P. M. Marques. "Antioxidant phenolic esters with potential anticancer activity: a Raman spectroscopy study". J. Raman Spectrosc. 2008. 39(1): 95-107.

41. J. Ralph. "Hydroxycinnamates in lignification". Phytochem. Rev. 2010. 9(1): 65-83.

42. U.S Department of Energy, "Biomass Feedstock Composition and Property Database". Page last reviewed May 14 2004. http://www.afdc.energy.gov/biomass/progs/search1.cgi [accessed Mar 27 2012].

43. H. P. S. Abdul Khalil, A. F. I. Yusra, A. H. Bhat, M. Jawaid. "Cell wall ultrastructure, anatomy, lignin distribution, and chemical composition of Malaysian cultivated kenaf fiber". Ind. Crops Prod. 2009. 31(1): 113-121.

44. F. Marin, J. L. Sanchez, J. Arauzo, R. Fuertes, A. Gonzalo. "Semichemical pulping of Miscanthus giganteus. Effect of pulping conditions on some pulp and paper properties". Bioresour. Technol. 2009. 100(17): 3933-3940.

45. R. C. Pettersen. "The chemical composition of wood". In: R. Rowell, editor. Advances in Chemistry Series 207, The Chemistry of Solid Wood. Washington D.C., U.S.A.: American Chemical Society, 1984. Chapter 2, Pp. 57-126.

46. J. Nousiainen, M. Rinne, M. Hellaemaeki, P. Huhtanen. "Prediction of the digestibility of primary growth and regrowth grass silages from chemical composition, pepsin-cellulase solubility and indigestible cell wall content". Anim. Feed Sci. Technol. 2003. 110(1-4): 61-74.

47. B. C. Aston. "Report of Chemistry Section". Ann. Rept. New Zealand Dept. Agr. 1935. 60-65.

48. P. Aaman, E. Nordkvist. " Chemical composition and in-vitro degradability of major chemical constituents of red clover harvested at different stages of maturity". J. Sci. Food Agric. 1983. 34(11): 1185-1189.

49. U. P. Agarwal, "An Overview of Raman Spectroscopy as Applied to Lignocellulosic Materials". In: D.S. Argyropoulos, editor. Advances in Lignocellulosics Characterization. Atlanta, GA: TAPPI Press. 1999. Chapter 9, Pp. 201-225.

50. G. Socrates, Infrared and Raman Characteristic Group Frequencies: Tables and Charts, 3rd Edition. Chichester, UK: John Wiley & Sons. Chapter 1, Pp. 35-47.

Table I. Dry weight compositions for the types of plant material included in this study as reported in the literature.

Plant Source	% Cellulose	% Hemicellulose	% Lignin	% Extractives
Alfalfa[42]	25	19	13	20
Kenaf-Bast[43]	55	32	15	5.5
Kenaf-Core[43]	49	38	19	4.7
Miscanthus[44]	51	25	22	3.8
Oak[45]	46	22	24	6
Orchard Grass[46]	27	22	2.8	39
Pampas Grass[47]	39	19	18	N/A
Pine[42]	42	21	26	2.7
Poplar[42]	42	19	23	2.8
Red Clover[48]	16	11	10	31
Switchgrass[42]	37	27	18	13

Table II. Raman Spectral Peak Assignments.

Approximate Peak Location (cm^{-1})	Primary Assignment
780	CO stretch[14]; lignin aromatic skeletal vibrations[14, 49]
805	CO stretch; aryl symmetric CH bend[14], CH out-of-plane bend[14]
829	CH out-of-plane bend[14]
896	cellulose HCC and HCO bending[34]
942	lignin CCH wag, aromatic skeletal vibrations[11, 14]
977	lignin CCH and –HC=CH-deformation, methyl wagging[11, 14]
1002	cellulose CC and CO stretch[34]
1033	cellulose CC and CO stretching[34]; lignin CH_3 wagging, CH_3 out-of-plane rock, aromatic skeletal vibrations, methoxy vibrations[11, 14, 34]
1074	cellulose CC and CO stretch[34]
1091	cellulose CC and CO stretch[34]
1117	cellulose CC and CO stretch[34]; lignin methoxy vibrations, aryl CH bend[14]
1147	cellulose CC and CO stretch[34], HCC and HCO bend[34]; lignin methoxy vibrations, aromatic CCH bend[14]
1155	C-C in carotenoids[39]
1169	lignin hydroxyl COH bend, aromatic skeletal vibrations[14]
1185	lignin methoxy vibrations, COH in plane bend[14]
1202	lignin methoxy vibrations[14]
1268	lignin aromatic skeletal vibrations, methoxy vibrations[11, 49]; cellulose HCC and HCO bend[34]
1338	cellulose HCC and HCO bend[34]
1376	lignin symmetric CH deformation; cellulose HCC, HCO, and HOC bend[34]
1427	lignin methoxy deformation, methyl bending, aromatic skeletal vibrations[11, 49]
1460	lignin methoxy deformation; cellulose HCH and HOC bend[34]
1528	C=C in carotenoids[39]
1605	lignin aromatic skeletal vibrations[14, 49]
1632	lignin C=C stretch of coniferaldehyde, sinapaldehyde, phenolic esters[11, 40, 49]
1656	lignin C=C stretch of coniferyl alcohol and sinapyl alcohol[11, 49]
1704	carbonyl stretch[50]

Table III. Comparison of G and S lignin percentages obtained by thioacidolysis or Raman Spectroscopy plus PCR. Uncertainties represent the 95% confidence interval.

Sample	%G, Thioacidolysis	%G, Raman Spectroscopy	%S, Thioacidolysis	%S, Raman Spectroscopy
Alfalfa	70 ± 10	78 ± 5	30 ± 10	21 ± 5
Kenaf-Bast	11 ± 2	10 ± 5	89 ± 2	90 ± 5
Kenaf-Core	24 ± 2	19 ± 2	76 ± 2	81 ± 2
Miscanthus	61 ± 1	66 ± 2	39 ± 1	34 ± 2
Oak	28 ± 5	34 ± 5	72 ± 5	66 ± 5
Orchard Grass	68 ± 5	120 ± 10	31 ± 2	-20 ± 10
Pampas Grass	57 ± 2	58 ± 2	43 ± 2	41 ± 2
Pine	99 ± 1	110 ± 10	1 ± 1	-10 ± 10
Poplar	17 ± 2	18 ± 2	83 ± 2	82 ± 2
Switchgrass	83 ± 7	75 ± 2	17 ± 5	23 ± 2

Figure 1. Raman spectra of solid microcrystalline cellulose (green) and 50 mg mL^{-1} lignin in methanol (brown). The spectra were divided by the total analysis time, and the lignin spectrum was multiplied by 10.

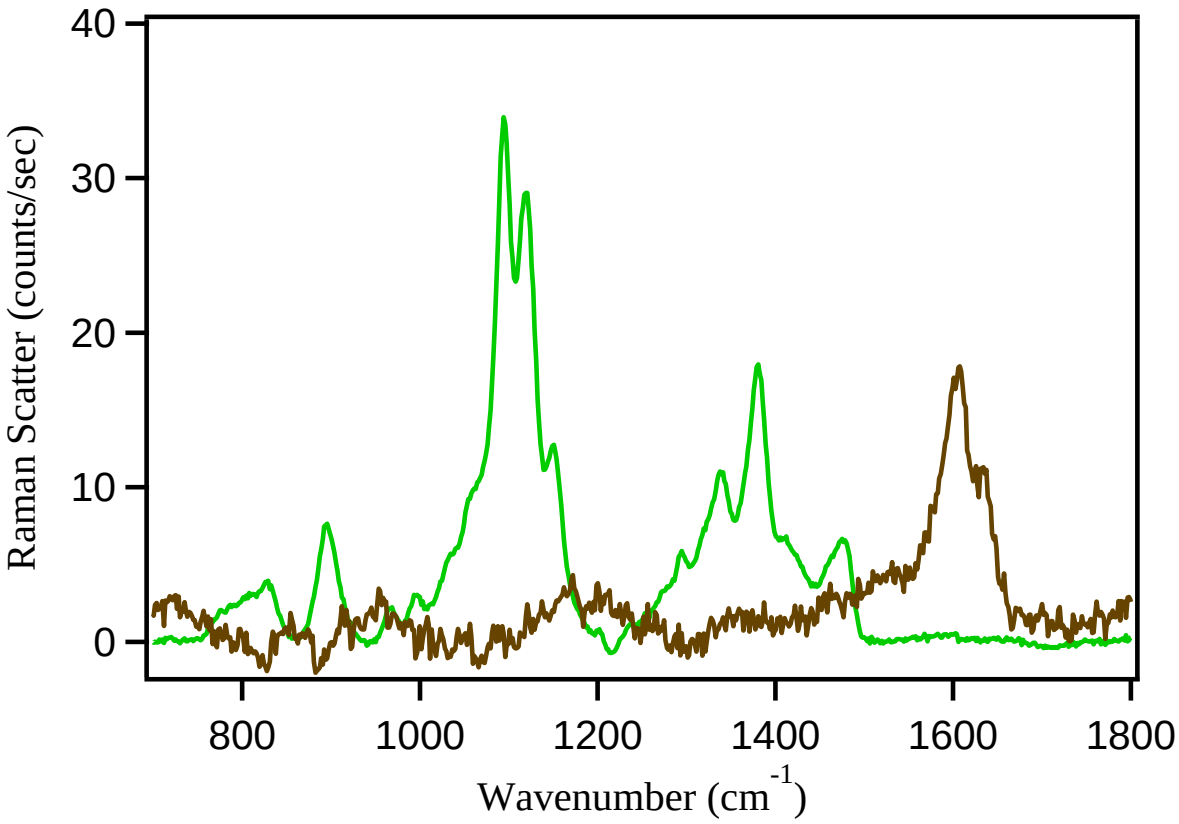


Figure 2. Raman spectra of solid biomass. Panel (A) shows the spectra of kenaf bast (blue), kenaf core (purple), oak (red), pine (black) and poplar (green). Panel (B) shows the spectra of *miscanthus* (red), pampas grass (black) and switchgrass (green). Panel (C) shows the spectra of alfalfa (blue), orchard grass (red) and red clover (green). Panel (D) represents the averages of the spectra in panels A (black), B (red), and C (blue).

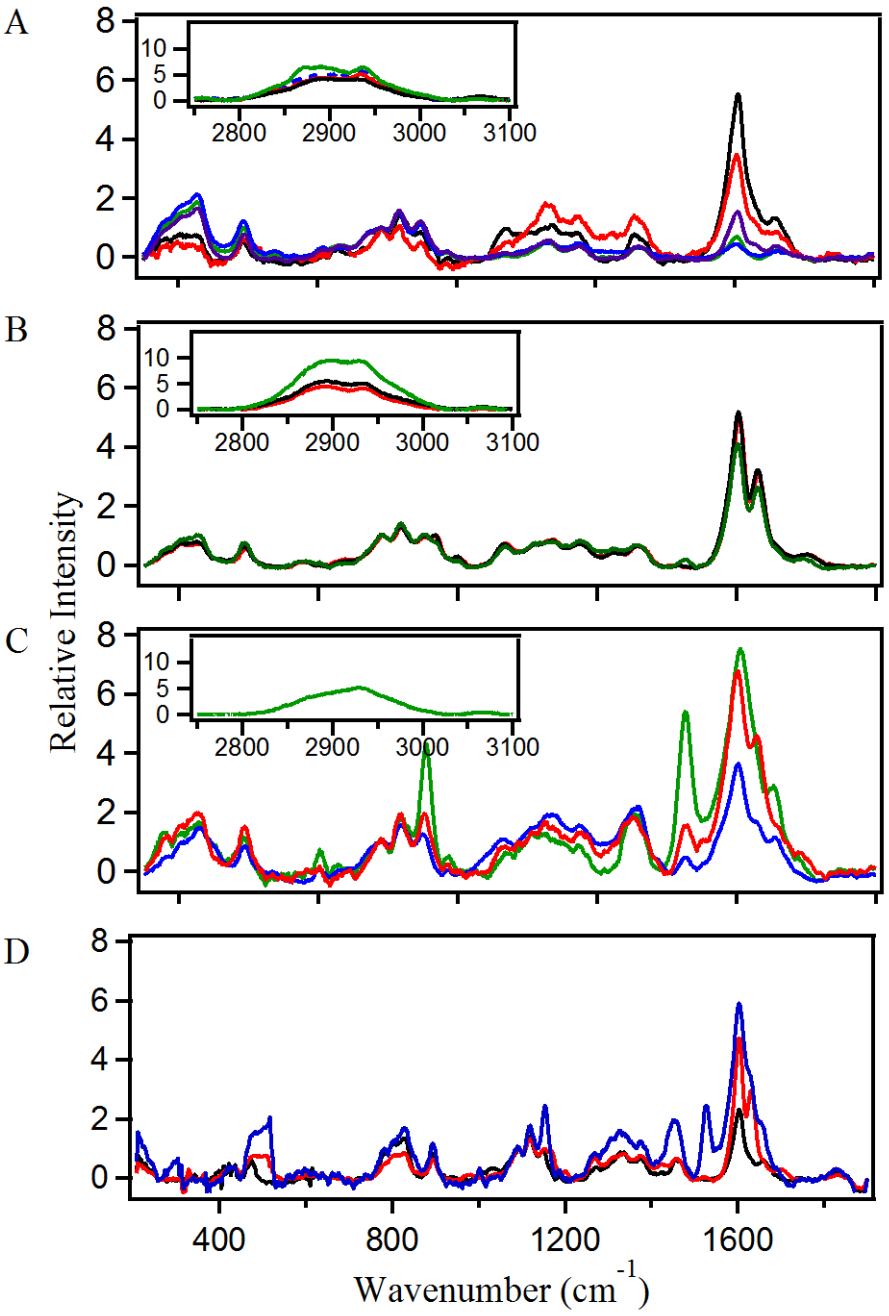


Figure 3. Raman spectra of orchard grass (A) and red clover (B) before (black) and after (green) Soxhlet extraction using 95% ethanol.

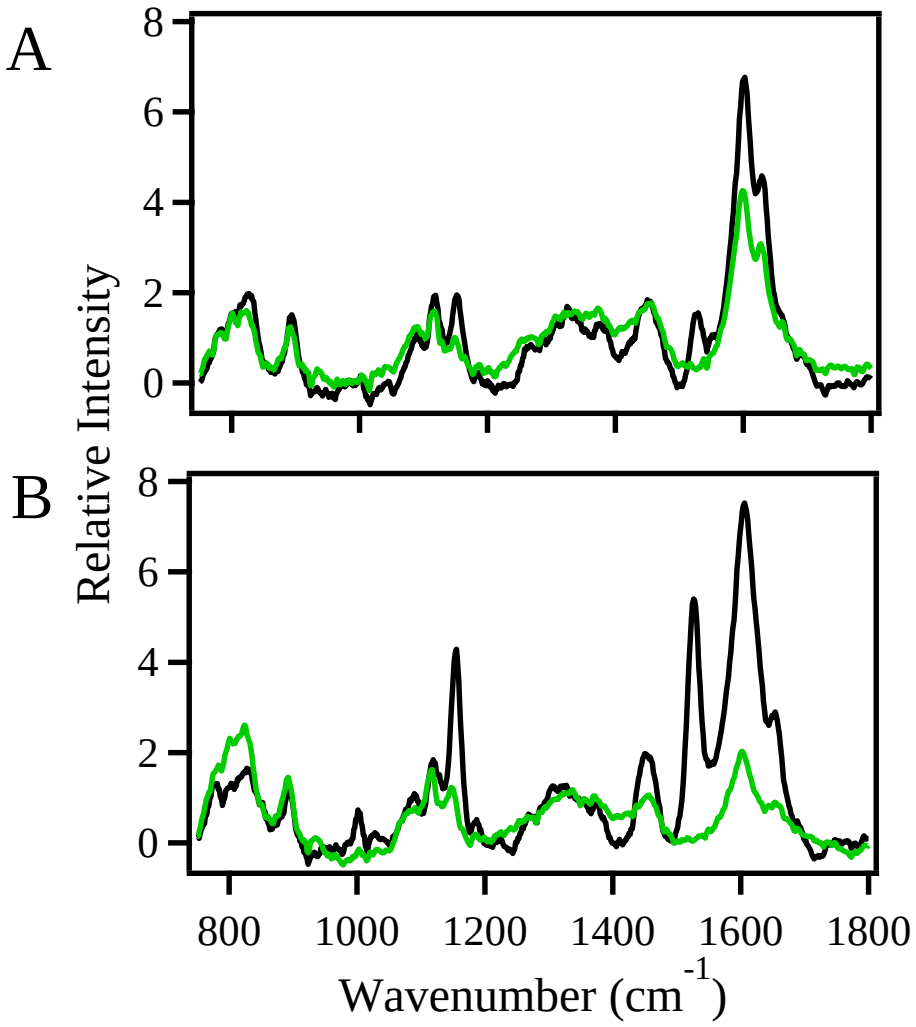


Figure 4. ^{13}C-NMR CP/MAS spectra of pine (A), *miscanthus* (B), and red clover (C). Panel (D) shows the relative carbon intensities calculated from the integration ranges listed below. Relative intensities were found by dividing the individual integrated peak height by the total carbon intensity (sum of individual intensities). The integration ranges used and assignments were as follows (in ppm): (1) 180-163, carboxylate carbon in hemicellulose, (2) 158-110, phenyl carbon in lignin, (3) 109.8-100, carbon 1 in cellulose, (4) 92-80.5, carbon 4 in cellulose, (5) 80-68.4, carbons 2, 3, and 5 in cellulose, (6) 68-60.1, carbon 6 in cellulose, (7) 59-52, methoxy carbon in lignin, hemicellulose, (8) 42-27.5, aliphatic C-H, (9) 24.1-18, methyl in hemicellulose.

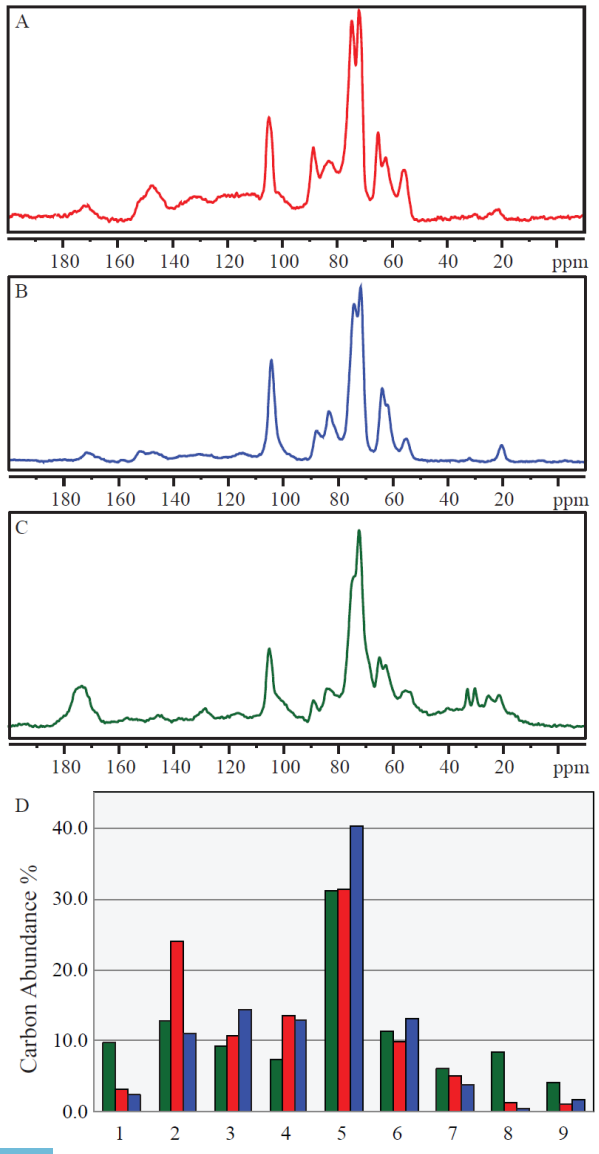
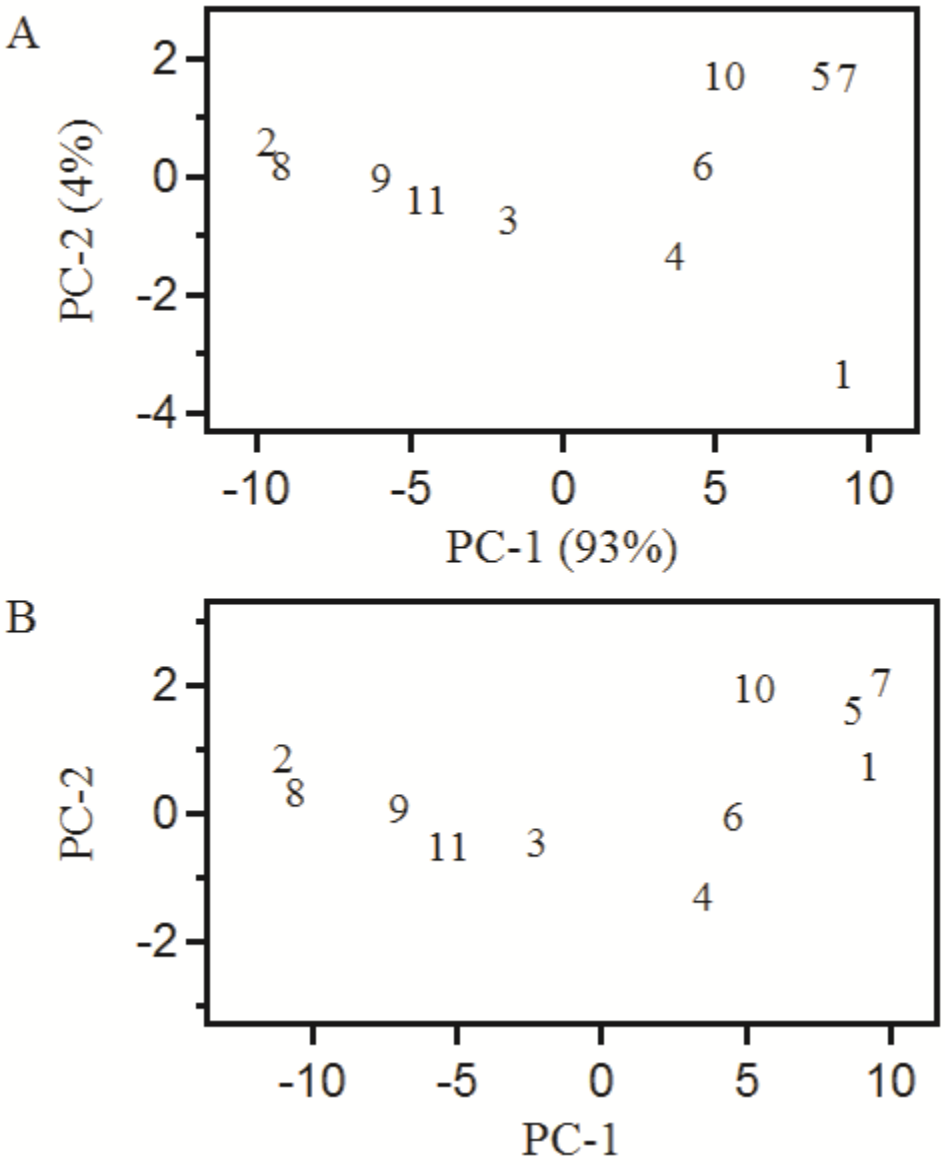


Figure 5. PCA scores plot of (A) the calibration set using all biomass samples, and (B) the projections using predicted scores when the biomass samples were sequentially left out of the calibration model. The numbers in the plots represent: 1 pine, 2 kenaf-bast, 3 red clover, 4 alfalfa, 5 *miscanthus*, 6 orchard grass, 7 pampas grass, 8 poplar, 9 kenaf-core, 10 switchgrass, and 11 oak.



SUPPLEMENTAL INFORMATION

CHARACTERIZATION OF WOODY AND HERBACEOUS BIOMASSES LIGNIN COMPOSITION WITH 1064 NM DISPERSIVE MULTICHANNEL RAMAN SPECTROSCOPY

Jason S. Lupoi,[1,2] and Emily A. Smith[1,2,*]

[1]U.S. Department of Energy, Ames Laboratory, Ames, Iowa, 50011

[2]Department of Chemistry, Iowa State University, 1605 Gilman Hall, Ames, Iowa 50011;

*telephone: 1-515-294-1424; fax: 1-515-294-0105; e-mail: esmith1@iastate.edu

Thioacidolysis. 10 mg of extractive-free biomass was put into a 5 mL glass vial with a Teflon® lined cap. 1 mL of fresh 2.5% v/v boron trifluoride etherate and 10% ethanethiol in dioxane was added to the biomass. The samples were flushed with nitrogen and heated for 4 hours at 100°C with periodic manual agitation. The reaction was terminated by placing the samples in the freezer for 5 minutes. 5 mg mL^{-1} tetracosane (50 µL) was added to each sample along with enough sodium bicarbonate (~400 µL) to bring the reaction mixture to pH 3-4 as measured with pH paper.

To extract the lignin degradation products, 2 mL of deionized water and 1 mL of methylene chloride were added to each sample. The samples were vortexed and allowed to phase-separate. Approximately 1.5 mL of the organic phase was removed and passed through a Pasteur pipette plugged with a small piece of lens paper and at least 50 mg of anhydrous sodium sulfate to remove residual water. The samples were dried to evaporation, and re-suspended in 1 mL of methylene chloride. Derivatization of the samples was achieved by combining 20 µL of the thioacidolysis products with 20 µL pyridine and 100 µL N,O-bis(trimethylsilyl) acetamide. The samples were incubated overnight.

Figure S1. Raman Spectra of 100 mg mL^{-1} ferulic (red), sinapic (black), and coumaric acids (blue) dissolved in dioxane. The spectra were collected using 570 mW 1064 nm excitation and 3 minutes acquisition time.

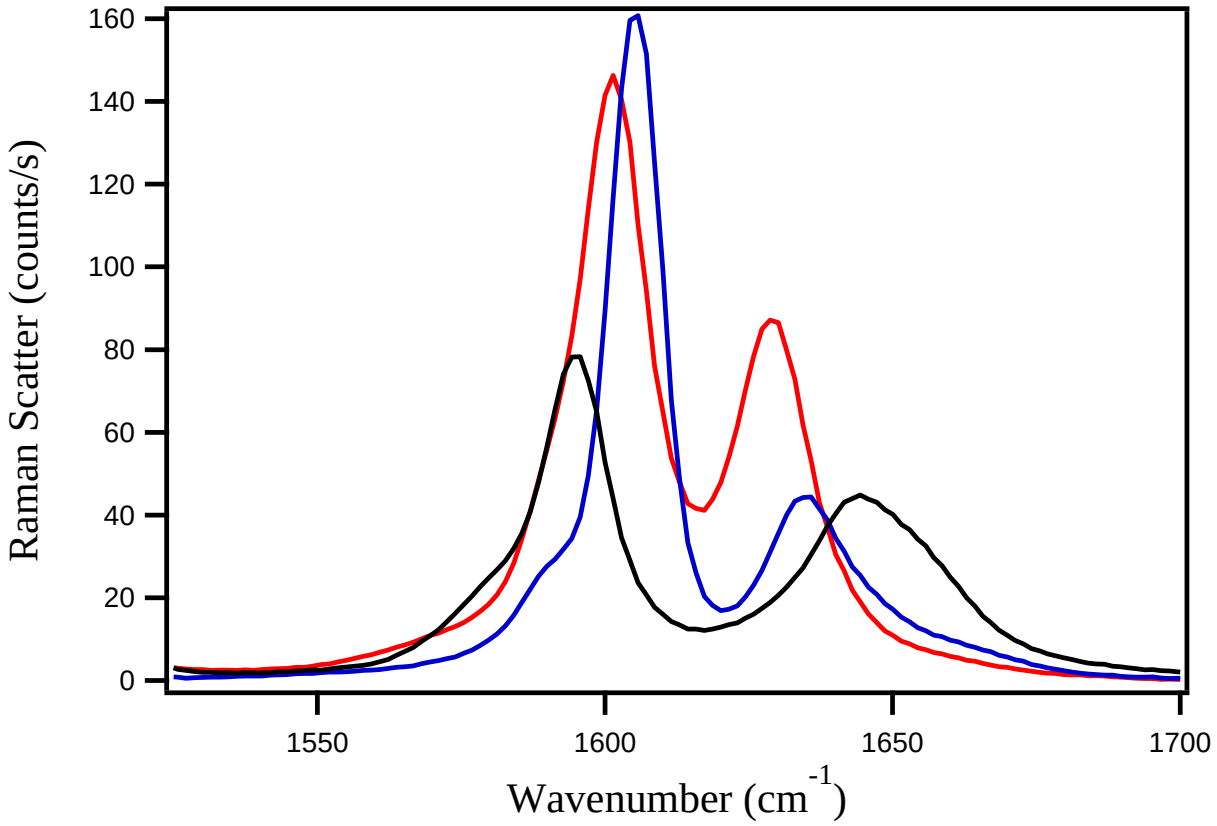


Figure S2. Bright field microscopy images showing the biomass particle size. Images were collected with a high resolution USB 2.0 CMOS camera (Thor Labs, Newton, NJ, USA). The detector element was 1280 x 1024 pixels. The camera was mounted on an inverted optical microscope. The Image J software package was used to determine the dimensions of the particles and adjust the contrast.

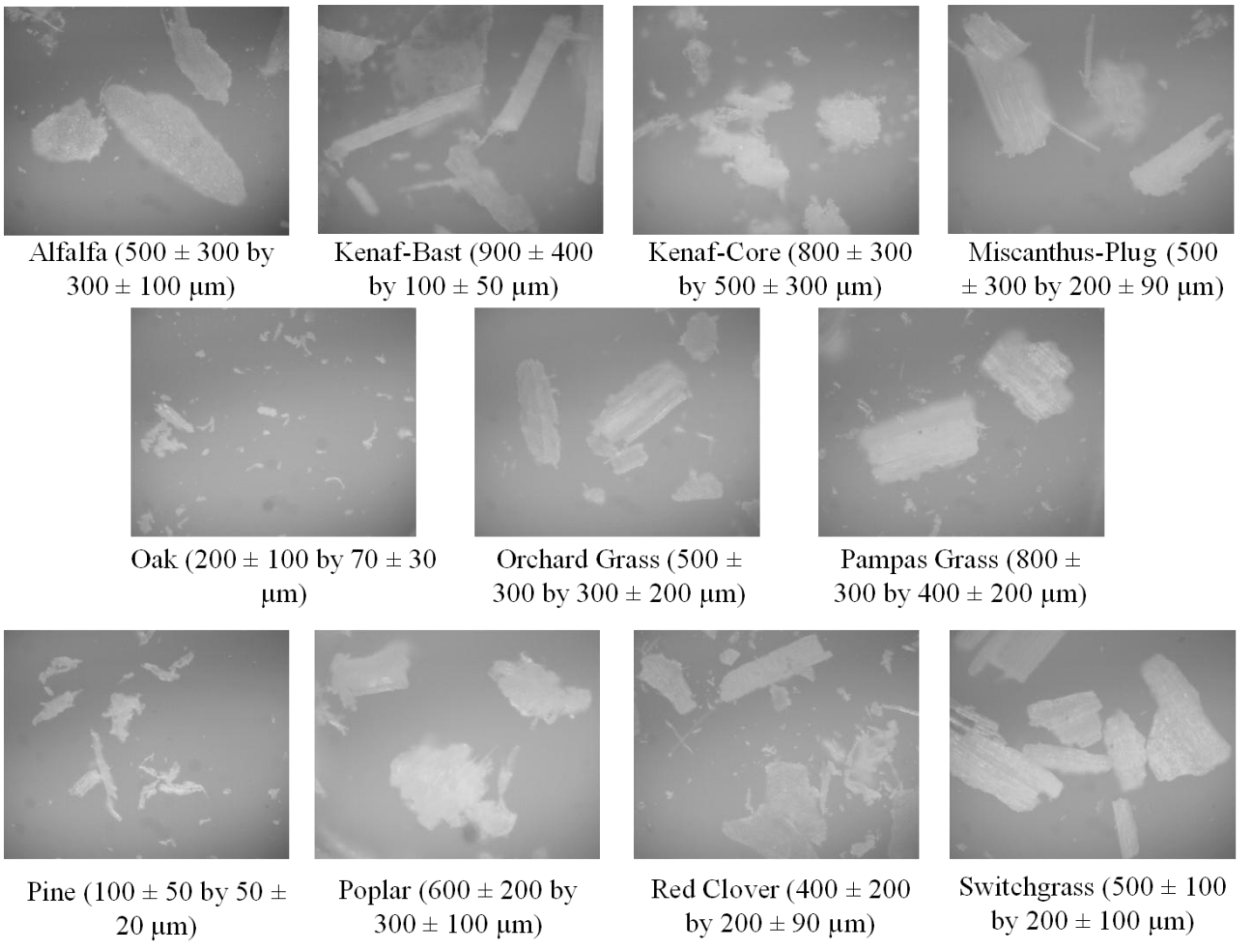


Table SI: Parameters for the PCR model used to quantify G and S lignin.

Variable	Slope	Root Mean Standard Error	R^2
G-Calibration	0.996	2%	0.997
G-Prediction	0.976	4%	0.985
S-Calibration	0.996	2%	0.996
S-Prediction	0.973	4%	0.986

CHAPTER 5

NEAR-INFRARED RAMAN SPECTROSCOPY MEASUREMENTS OF ETHANOL DETECTION LIMITS IN SIMULTANEOUS SACCHARIFICATION AND FERMENTATION LIQUORS

Jason S. Lupoi and Emily A. Smith

(This chapter represents the preliminary research into using 1064 nm Raman spectroscopy to determine ethanol produced during the simultaneous saccharification and fermentation of biomass)

Abstract

Ethanol detection limits were measured using 1064 nm near-infrared Raman spectroscopy. Ethanol was spiked into simultaneous saccharification and fermentation matrices using either microcrystalline cellulose or kenaf bast as substrates. The detection limit when microcrystalline cellulose was used was 3 mg ethanol/mL SSF matrix, while the detection limit calculated when kenaf bast was the substrate was 7 mg ethanol/mL SSF matrix. Detection limits were determined using the peak height calculated from a Gaussian curve fit to the most intense ethanol vibrational mode at 874 cm^{-1}. The ethanol detection limit in the kenaf bast versus the microcrystalline cellulose matrix is presumably lower due to interfering extraneous compounds extracted from the plant tissues during the reaction that increase spectral background.

Introduction

A 2008 U.S. D.O.E. report estimated that products derived from lignocellulosic biomass represent about 4% of U.S. energy consumption.[1] The conversion of plant matter to liquid fuels such as bioethanol has been cited as being one pathway to relinquishing the world's dependence on a dwindling fossil fuel supply. Dedicated bioethanol crops such as

switchgrass, *miscanthus*, or energy cane have shown promise in supplanting the current usage of corn for ethanol production.[2-3]

Ethanol is produced from the fermentation of glucose by microorganisms such as the yeast strain *Saccharomyces cerevisiae*. While separate, multi-step hydrolysis and fermentation reactions are still utilized, the use of simultaneous saccharification and fermentation (SSF) reactions has become more prevalent.[4-5] In a SSF reaction, both enzyme and yeast are present. As cellulose is hydrolyzed, the glucose formed is fermented to ethanol, reducing inhibition of cellulase by glucose permitting higher ethanol yields to be achieved.[4-5] The one-pot reaction technique used in SSF is less labor intensive and production costs have been reported to decrease when SSF is utilized.[4-5]

The detection of ethanol yields following fermentation reactions are typically achieved using chromatographic techniques such as high-performance liquid chromatography with refractive index detection[6-7], or gas chromatography/mass spectrometry.[8-9] Near-infrared (NIR) spectroscopy has also been used for ethanol detection, though, in order to quantify yields, chemometric analyses must be employed to extract useful features from the spectra.[10] Of these techniques, only NIR spectroscopy is capable of inline detection of ethanol production with probes in industrial reactors.[11]

Raman spectroscopy has been previously explored for the detection of ethanol following separate hydrolysis and fermentation reactions.[8] This research employed 785 nm excitation, and found that the ethanol detection limit was 6 mg/mL. The spectral background was high, however, and was also dependent upon the pretreatment selected for removing undesirable cell wall constituents. A follow-up study showed that extractable compounds in the plant were the major contributors to the high spectral background, when 785 nm excitation was used.[12]

The selection of wavelength is paramount to receiving quality Raman spectra. Raman scattering is highly dependent on the frequency of excitation (ν^4), therefore, by selecting a longer excitation wavelength, the energy interacting with the sample is decreased.[13] In the case of moving from 785 to 1064 nm, the energy is reduced by a factor of

3.8. Although the energy is decreased, S/N is often increased, since spectral backgrounds are greatly reduced when using 1064 nm excitation. This is especially true when analyzing biological samples that have a high degree of intrinsic fluorescence. Two (1064 and 785 nm) excitation wavelengths were compared for the analysis of lignin, and it was found that the spectral background from lignin was 160x higher when 785 nm excitation was employed.[14] Therefore, spectral features from constituents in fermentation liquors that have previously resulted in higher spectral background should be reduced when 1064 nm excitation is used.

Herein, the use of 1064 nm Raman spectroscopy is described for the determination of ethanol detection limits in fermentation liquors, when microcrystalline cellulose and kenaf bast have been used as substrates. Raman spectroscopy offers an alternative instrumental technique, capable of yielding qualitative and quantitative information, and is amenable to field portability, remote sensing, or online process chemistry.

Materials and Methods

Materials. Yeast extract and peptone were purchased at U.S. Biological (Swampscott, MA, USA). Accellerase 1500 was provided by Genencor (Rochester, NY, USA). Microcrystalline cellulose was purchased from Sigma-Aldrich (St. Louis, MO, USA). Kenaf bast was obtained from Dr. Kenneth Moore, Iowa State University, Department of Agronomy. The kenaf was grown at the Sorenson Research Farm in Ames, IA, ground using a Wiley mill, and passed through a 1 mm mesh screen.

Extraction and Pretreatment of Kenaf Bast. The kenaf was Soxhlet extracted using 95% ethanol for approximately 16 hours, oven-dried, and treated with 1% sulfuric acid solution at room temperature for 4 hours to remove the majority of hemicellulose from the cell wall. Following the dilute acid pretreatment, the kenaf was washed thoroughly until the rinse water pH was neutral as detected with pH paper. The biomass was not filtered to complete dryness to prevent collapse of the cell wall, as previously reported.[15]

Fermentation Liquor Preparation. Fermentation liquors were prepared using both cellulose and kenaf bast as substrates. For the cellulose fermentation liquor: 4 g

microcrystalline cellulose was combined with 5 mL 10X YP media, 2 mL Accellerase 1500 that had been previously denatured by heating in boiling water for 10 minutes, and 43 mL deionized water. No yeast was added to the reaction solution. The solution was incubated at 37°C for 24 hours with 150 rpm agitation. For the kenaf bast fermentation liquor, 20 g wet kenaf bast (~4 g dry) was combined with 5 mL 10X YP media, 1 mL previously denatured Accellerase 1500, and 44 mL deionized water. Ethanol standards were prepared such that the ethanol concentration was 10, 25, 40, 50, 75, or 100 µL/mL in the fermentation liquor.

Raman Spectroscopy. A home-built near-infrared Raman spectrometer has been previously described.[14] A 300 groove/mm classically ruled grating replica provided the maximum trade-off between spectral coverage and spectral resolution for these analyses. Raman spectra of ethanol in the fermentation liquors were acquired using 830 mW laser power for 5 minutes, and were baseline corrected using Grams AI V 8.0 (Thermo Scientific, Waltham, MA, USA). The spectra were imported into IGOR Pro V 6.1 (WaveMetrics Inc., Lake Oswego, OR, USA) for further analysis. Peak heights of the 874 cm^{-1} ethanol vibrational mode were determined using a multi-peak curve fit in IGOR Pro V 6.1. The 40 µL/mL sample was measured 7 times, and the concentration detection limit was calculated using the equation (t*s)/m, where t was 3.143, s was the standard deviation of the peak height, and m was the slope of the calibration curve.[16]

Results and Discussion

The Raman spectra of ethanol in a SSF matrix when microcrystalline cellulose was used are shown in Figure 1. A calibration curve was generated after determining the peak height of the 874 cm^{-1} ethanol vibrational mode (Figure 2). The detection limit was calculated to be 3 mg ethanol/mL SSF matrix (Table 1). Microcrystalline cellulose is pure cellulose, therefore, no prior extractions or pretreatments are required to remove extraneous compounds, such as terpenes, chlorophyll, flavanoids or hemicellulose that may impede the access of cellulose to cellulase.

Figure 3 shows the Raman spectra of ethanol when kenaf bast was used as a substrate. Figure 4 depicts the respective calibration curve generated using the 874 cm^{-1} ethanol

vibrational mode. The detection limit for ethanol in this matrix was calculated to be 7 mg ethanol/mL SSF matrix (Table 1). A Soxhlet extraction, using 95% ethanol, and a dilute acid (1% H_2SO_4) pretreatment were used to remove extractives and hemicellulose from the plant cell wall prior to incubating kenaf bast in the SSF matrix. While these treatments remove a majority of the compounds that increase spectral background, some "extractives" and hemicellulose remain in the biomass.[17] These molecules are suspected of leaching into solution during the incubation of the kenaf bast, creating a more complex SSF matrix. The spectral contributions of these undesirable components inhibit the detection of ethanol.

A previous study determined that the selection of pre-fermentation extraction and pretreatment is paramount to decreasing the detection limit of glucose.[12] These results would also apply for ethanol detection in complex matrices.

Conclusions

Near-infrared Raman spectroscopy can be employed for the rapid detection of ethanol in complex, fermentation liquors. The detection limit of ethanol is directly related to the complexity of the biomass substrate, as residual molecules not removed during extraction or pretreatment can inhibit the quantitation of low concentrations of ethanol. It is pertinent, therefore, to select pre-fermentation treatments that selectively remove a vast majority of these constituents such that ethanol detection limits can be lowered to that of pure cellulose. This is especially important when using near-infrared excitation, given the ν^4 Raman scattering dependence.

References

1. U. S. D.O.E., "Renewable Energy Annual 2008, ftp://ftp.eia.doe.gov/renewables/060308.pdf", (U.S. Energy Information Administration, Office of Coal, Nuclear, Electric and Alternate Fuel, U.S. Department of Energy, Washington D.C., 2010).
2. A. Karp and I. Shield, The New phytologist 179, 1, 15 (2008).

3. E. Lam, J. Shine, Jr., J. Da Silva, M. Lawton, S. Bonos, M. Calvino, H. Carrer, M. C. Silva-Filho, N. Glynn, Z. Helsel, J. Ma, E. Richard, Jr., G. M. Souza, and R. Ming, GCB Bioenergy 1, 3, 251 (2009).
4. K. Olofsson, M. Bertilsson, and G. Liden, Biotechnology for biofuels 1, 1, 7 (2008).
5. G. P. Philippidis, Handbook on Bioethanol, 253 (1996).
6. K. Hoyer, M. Galbe, and G. Zacchi, Journal of Chemical Technology and Biotechnology 84, 4, 570 (2009).
7. H. Li, N.-J. Kim, M. Jiang, J. W. Kang, and H. N. Chang, Bioresource Technology 100, 13, 3245 (2009).
8. C.-J. Shih and E. A. Smith, Anal. Chim. Acta 653, 2, 200 (2009).
9. D. W. Templeton, "Determination of Ethanol Concentration in Biomass to Ethanol Fermentation Supernatants by Gas Chromatography, Laboratory Analytical Procedure #011", (U.S. D.O.E. National Renewable Energy Laboratory, 1994).
10. M. Blanco, A. C. Peinado, and J. Mas, Biotechnology and Bioengineering 88, 4, 536 (2004).
11. J.W. Peterson and A.H. Ullman, "Process Analysis: Introduction". In: R.A. Meyers, editor. The Encyclopedia of Analytical Chemistry, Volume 15. (Chichester, UK: John Wiley and Sons, 2000. Pp. 8127-8145., 2000).
12. C.-J. Shih, J. S. Lupoi, and E. A. Smith, Bioresource Technology 102, 8, 5169 (2011).
13. E. Smith, Modern Raman Spectroscopy (Chichester, UK: John Wiley and Sons, 2005. Pp. 1-30, 71-92., 2004).
14. M. W. Meyer, J. S. Lupoi, and E. A. Smith, Analytica Chimica Acta 706, 1, 164 (2011).
15. M. Selig, Weiss, N, and Ji, Y, "Enzymatic Saccharification of Lignocellulosic Biomass ", (Golden, CO: National Renewable Energy Lab, 2008).
16. D. C. Harris, Quantitative Chemical Analysis (Bleyer, C, New York, 2007).
17. T. H. Kim, J. S. Kim, C. Sunwoo, and Y. Y. Lee, Bioresource Technology 90, 1, 39 (2003).

Table 1: Concentration detection limits of ethanol in SSF matrix.

Sample Matrix	Student's t-value	Standard Deviation	Slope	Detection Limit (mg/mL)
Cellulose	3.143	14.843	14.329	3
Kenaf Bast	3.143	28.437	13.416	7

Figure 1: The detection of ethanol spiked into a SSF matrix, where microcrystalline cellulose was used as the substrate. The ethanol concentrations, in µL ethanol/mL are as follows: 10-black, 25-red, 40-purple, 50-gray, 75-green, and 100-blue.

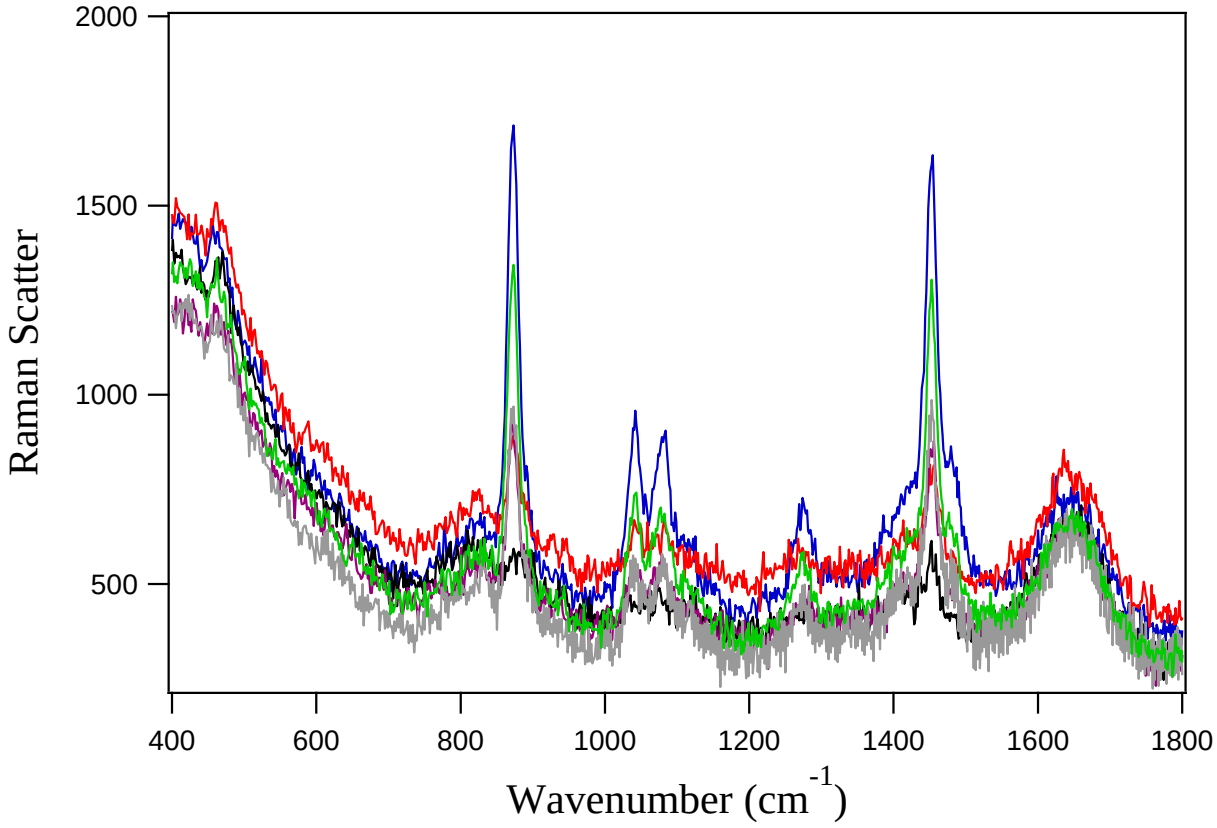


Figure 2: Calibration curve generated using multi-peak curve fitting for the detection of ethanol in SSF matrix with microcrystalline cellulose as the substrate.

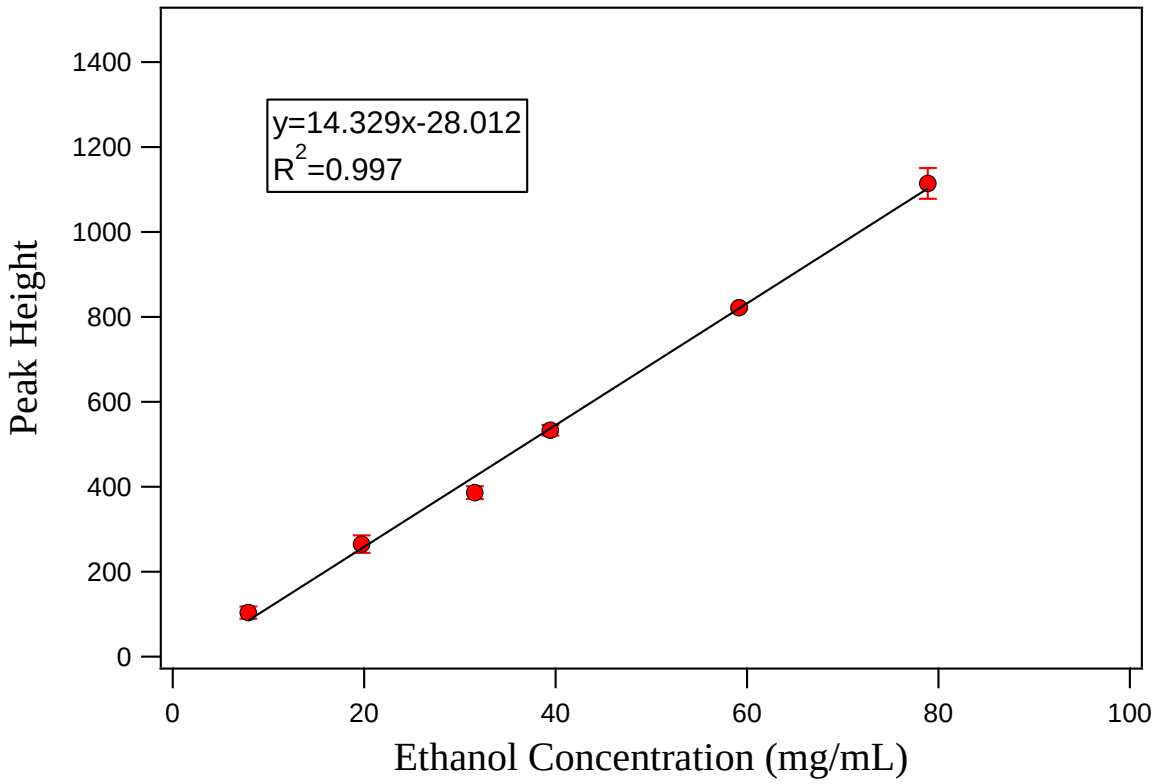


Figure 3: The detection of ethanol spiked into a SSF matrix, where microcrystalline cellulose was used as the substrate. The ethanol concentrations, in µL ethanol/mL are as follows: 10-black, 25-red, 40-purple, 50-gray, 75-green, and 100-blue.

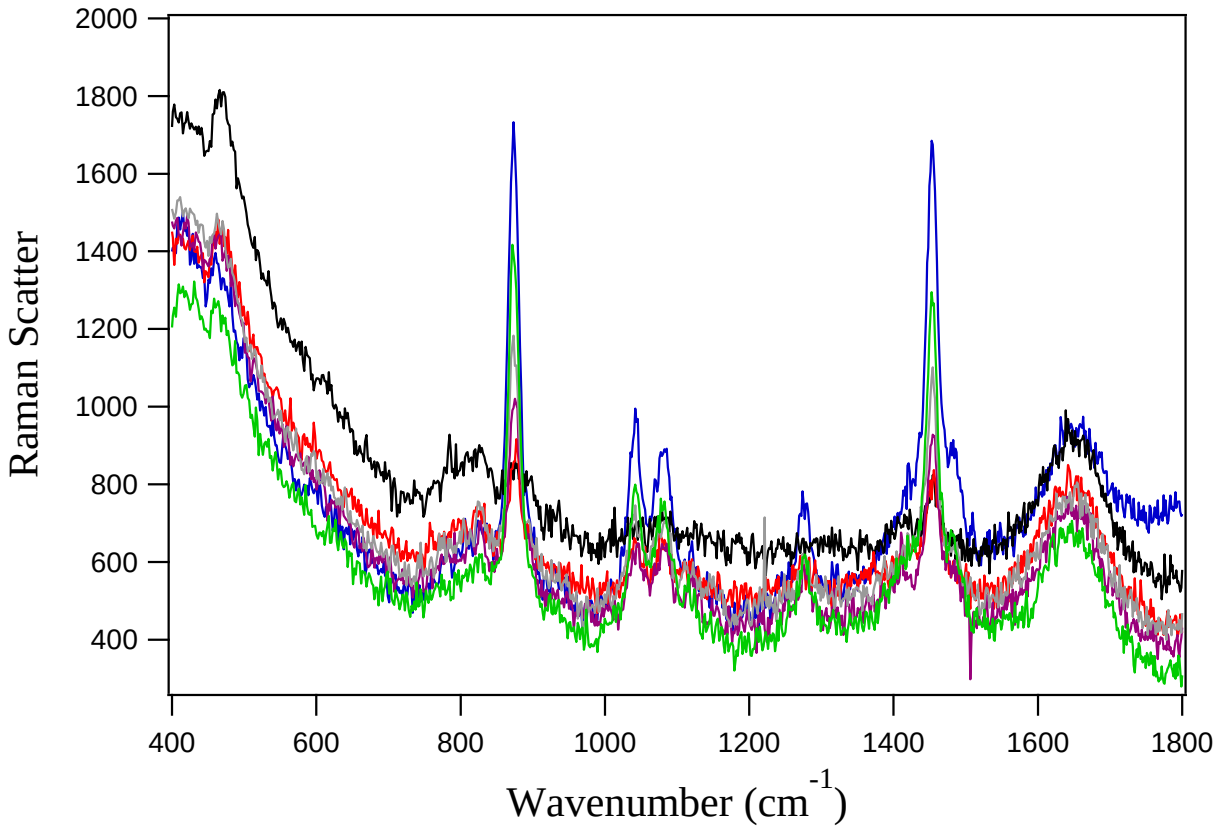
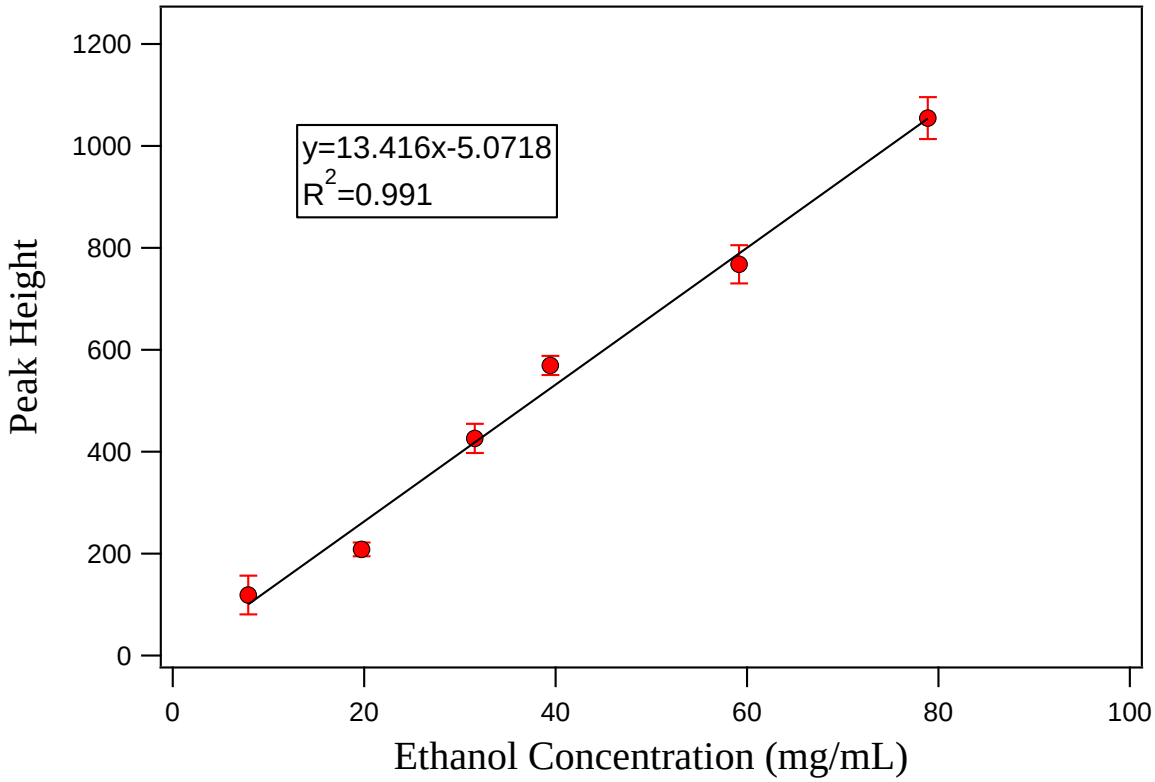


Figure 4: Calibration curve generated using multi-peak curve fitting for the detection of ethanol in SSF matrix with kenaf bast as the substrate.



CHAPTER 6

CONCLUSIONS

This thesis focuses upon: (1) the development of a cellulase immobilization technique as a way to increase ethanol yields from SSF reactions, and (2) the development and optimization of NIR Raman spectroscopy for the characterization of lignocellulosic biomass.

The use of immobilized enzymes allows the researcher to tailor reaction conditions to his or her needs, since the immobilization confers stability upon the enzyme at sub-optimal reaction conditions. The conditions used in hydrolysis and fermentation reactions differ in the optimal temperature range required by either cellulase or yeast. This has led researchers to explore genetically modified, thermotolerant yeasts. The physisorption of cellulase to silica nanoparticles resulted in a shift in the optimal hydrolysis temperature necessary for maximum glucose yield. The optimal temperature range for immobilized cellulase was lowered to between 25-35°C, which meshes with the most favorable yeast fermentation conditions. The solution cellulase functions optimally near 50°C. In this study, immobilized cellulase was shown to withstand SSF reaction conditions to a greater extent than the solution cellulase, as determined from total glucose yield. Therefore, during SSF reactions, increased ethanol yields should be recognized when using immobilized compared to free cellulase.

Despite advances in NIR dispersive Raman instrumentation, FT-Raman remains more prevalent in current literature for studying biomass. This research compared a commercial FT and a home-built 1064 nm dispersive Raman instrument, and found that at the longer acquisition times (above 15 seconds) necessary for obtaining lignin spectra, the dispersive instrument provided higher signal-to-noise.

The use of NIR, multichannel, dispersive, Raman spectroscopy has been described for characterizing a library of lignocellulosic feedstocks encompassing gymnosperms, angiosperms, and herbaceous plants. Excitation wavelength selection is paramount for elucidating valuable information from the Raman spectra of plants. In this research, using a 1064 nm laser was required for the detection of plant cell wall components such as lignin due

to the high background obtained at higher frequency excitation. The spectra of different feedstocks were visually classified according to woody, non-woody herbaceous, and grass nomenclature using spectral differences/similarities. Therefore, the use of NIR Raman spectroscopy makes possible the rapid screening of an array of lignocellulosic feedstocks. This can be especially realized when considering the field-portable instruments commercially available. For example, cellulose is the desirable plant cell wall constituent for production of biofuels. Feedstocks can be rapidly screened for cellulose comparison *in situ* such that only those possessing high cellulose content can be selected for downstream bioethanol applications.

Another useful, rapid, Raman spectroscopy screening method is the monitoring of the extraction of undesirable cellular components, such as carotenoids or non-lignin phenolics. By first recognizing which vibrational modes are from these molecules, their effective removal can be determined from the decline or disappearance of these spectral features. Soxhlet extractions are often performed for a multitude of hours; however, knowledge of the course of extractive removal could decrease the total processing time.

As agronomists explore the alteration of lignin to ensure more facilitated lignin degradation, the use of NIR Raman spectroscopy coupled with chemometrics will allow quantification of the lignin monomer content as a useful gauge to the success of the of the genetic modifications to produce more degradable lignin. Concurrently, the S/G ratio provides an indication of the loading requirements of toxic reagents (such as sodium chlorite) used in the bleaching of pulp. This research has shown the utility of using NIR Raman spectroscopy to quantify the S/G ratio using a PCR calibration model comprised of thioacidolysis S and G percentages coupled with Raman spectra. The knowledge of this key ratio can help reduce overloading with bleaching solvents that require remediation.

Other common analytical techniques for solid analysis such as NMR or GCMS require longer run times, greatly reducing instrumental throughput. The ubiquitous presence of water in biological materials also does not pose spectral problems in Raman spectroscopy, unlike infrared spectroscopy. Overall, Raman spectroscopy offers a high-throughput,

analytical tool, with amenities including the lack of tedious sample preparations, a non-destructive analysis, field portability, and the generation of both qualitative and quantitative information.

APPENDIX

RAMAN SPECTROSCOPY MEASUREMENTS OF GLUCOSE AND XYLOSE IN HYDROLYSATE: ROLE OF CORN STOVER PRETREATMENT AND ENZYME COMPOSITION

(A paper published in Bioresource Technology, 2011, Volume 102(8), 5169-5176)

*Chien-Ju Shih, Jason S. Lupoi and Emily A. Smith**

Ames Laboratory, U.S. Department of Energy, Ames, Iowa 50011-3111, U.S.A.

Department of Chemistry, Iowa State University, Ames, Iowa 50011-3111, U.S.A.

Abstract

The effect of corn stover pretreatment on glucose quantitation in hydrolysate using Raman spectroscopy is evaluated. Dilute sulfuric-acid pretreatment results in a 20 mg/mL glucose limit of detection in hydrolysate. Soaking in aqueous ammonia pretreatment produces a 5 mg/mL limit of detection. Water, ethanol or hexane extraction of corn stover reduces the spectral background that limits glucose detection in dilute acid hydrolysate. Additionally, a Raman spectroscopy multi-peak fitting method is presented to simultaneously measure glucose and xylose concentration in hydrolysate. This method yields a 6.1% average relative standard error at total saccharide concentrations above 45 mg/mL. When only cellulase is present, glucose and xylose yield were measured by Raman spectroscopy to be 32 ± 4 mg/mL and 7.0 ± 0.8 mg/mL, respectively. When both cellulase and hemicellulase were present, xylose yield increased to 18.0 ± 0.5 mg/mL. Enzymatic or colorimetric assays confirmed the validity of the Raman spectroscopy results.

Keywords: Raman Spectroscopy, Biomass Pretreatment, Glucose Quantitation, Xylose Quantitation, Enzymatic Hydrolysate, Spectral Multi-peak Fitting, Biofuels

1. Introduction

Lignocellulosic biomass, including corn stover, perennial grass, bagasse, wheat/rice straw and forestry/industry residue, is abundant and can be converted to portable biorenewable fuels without competing with food production (Knauf & Moniruzzaman, 2004). Plant cell walls primarily contain a phenolic polymer, lignin, and carbohydrate polymers, hemicellulose and cellulose (Koshijima & Watanabe, 2003). In lignocellulosic feedstocks, it is the carbohydrate polymers in the plant cell wall that can be chemically or biologically hydrolyzed to monosaccharides and then fermented to produce biofuels. Cellulose is a linear polymer of β-1,4-linked glucose units that is a key structural component in plants and can be hydrolyzed by cellulases. Hemicellulose is a heteropolysaccharide composed of several carbohydrate monomers with different linkages and substitutions on the primary branch (Pandey & Editor, 2009), and it can be hydrolyzed by hemicellulases. By dry mass, corn stover is ~38% cellulose and ~25% hemicellulose (Yang et al., 2010). Glucurono arabinoxylan is the most abundant hemicellulose in corn stover (Kim & Lee, 2007; Weiss et al., 2010). Lignin is a three-dimensional polymer consisting of syringyl-, guaiacyl-, and p-hydroxy- phenylpropanoid units (Ko et al., 2008). Since lignin in the cell wall interferes with the enzymatic hydrolysis of the polysaccharides, a biomass pretreatment step is required to disrupt the close association of the cell wall constituents for efficient hydrolysis (Méchin et al., 2005).

Many protocols for biomass pretreatment have been developed, and each protocol is associated with advantages and disadvantages. Dilute acid pretreatment breaks down the cell wall structure and exposes the cellulose for increased enzyme accessibility (Mosier et al., 2005). During dilute-acid pretreatment xylan is hydrolyzed directly into monosaccharides, which can be removed from the biomass with a rinse step. Approximately 70% of the lignin, 5% of the hemicellulose, and 95% of the cellulose remains after dilute acid pretreatment of corn stover (Kim et al., 2003). In contrast, soaking in aqueous ammonia pretreatment breaks

down the cell wall constituents and extracts a portion of the lignin from the cell wall (Kim & Lee, 2007). Approximately 15% of the lignin, 40% of the hemicellulose, and 90% of the cellulose remain after ammonia pretreatment of corn stover (Kim et al., 2003). On-going efforts aim to correlate efficient conversion methods with specific plant materials for the highest yield of fermentable sugars.

To date, many analytical methods have been developed to analyze monosaccharides in hydrolysate including high performance liquid chromatography (HPLC) (Agblevor et al., 2007; Lloyd & Wyman, 2005; Öhgren et al., 2007; Raymond & Ehrman, 1996; Sluiter et al., 2006), ultraviolet (UV)-visible spectrophotometry (Cara et al., 2008), capillary electrophoresis (Huber et al., 1994), gas chromatography/mass spectrometry (Kamm et al., 2006) and electrochemical methods (Tatsumi & Katano, 2004). Compared to the analytical techniques mentioned above, vibrational spectroscopy (i.e. Raman and infrared spectroscopy) provides intrinsic chemical content information and multiple analytes can be measured simultaneously and non-invasively. In contrast to infrared methods, Raman spectroscopy has advantages for the analysis of aqueous samples because water produces little scatter in the region of 0 to 3000 cm^{-1} (Brooker et al., 1989; Zhang et al., 2005). In general, compared to IR spectra, Raman spectra have narrower peak widths for aqueous samples and cover a wide spectral range, making it a better technique for quantitative analysis with complex matrices. Raman spectroscopy has been previously used to measure cell wall properties and biomass conversion in a number of species, including poplar (Gierlinger et al., 2006), black spruce (Agarwal, 2006), switchgrass (Li, 2009), corn stover (Shih et al., 2009) and *Eucalyptus globules* (Sun et al., 2011).

A previous study showed that the method used to pretreat corn stover prior to enzymatic hydrolysis affected the background in Raman spectra of the hydrolysate, and subsequently the quantitation of glucose in the hydrolysis broth (Shih & Smith, 2009). Pretreatment of corn stover with dilute sulfuric acid prevented measurements of glucose at concentrations generated in typical hydrolysis reactions of corn stover. On the other hand, soaking in aqueous ammonia pretreatment of corn stover enabled glucose quantitation at concentrations down to 8 mg mL^{-1}. It was hypothesized that the differing ability to measure

glucose was due to the amount of residual lignin that remained in the pretreated corn stover. The background may also result from intrinsic compounds that are extracted to different extents in the dilute acid and ammonia treated corn stover. Additionally, small molecules (e.g., furfural) may be generated in the pretreatment step that cause spectral interference in subsequent analysis steps (Mosier et al., 2005). Dilute acid and soaking in aqueous ammonia pretreatment are commonly utilized chemical pretreatment methods (Chen et al., 2009; Eggeman & Elander, 2005; Kumar & Wyman Charles, 2009; Wyman et al., 2005). Understanding the source(s) of the spectral background after different biomass pretreatments, as well as, quantitative measurements of additional saccharides could expand the utility of Raman spectroscopy for biofuel research and production.

Herein, the detection limit of glucose in corn stover hydrolysate for different pretreatment protocols is reported for Raman spectroscopy with 785 nm excitation. To date, much attention has been paid to the amount of glucose generated in the hydrolysate, but there has been less focus on the amount of xylose produced. This is due, in part, to the fact that commercial hydrolysis enzymes traditionally have had high cellulase activity and low hemicellulase activity. In order to significantly increase total sugar yields, enzyme mixtures containing or supplemented with hemicellulases can be used. Raman spectroscopy is suited to measure monosaccharide yield in these hydrolysates because multiple analytes can be measured simultaneously. A multi-peak fitting model was developed to measure glucose and xylose in the complex hydrolysis matrix.

2. Materials and methods

2.1 Materials

All chemicals were reagent grade. Sodium acetate trihydrate, glacial acetic acid, sulfuric acid, ethanol and ammonia were purchased from Fisher Scientific (Pittsburgh, PA). Cycloheximide, tetracycline (Fluka), D-glucose, microcrystalline cellulose, cellobiose, xylose, phloroglucinol and glucose HK reagent were purchased from Sigma-Aldrich (St. Louis, MO). Accellerase 1500® enzyme and Accellerase XY® hemicellulase enzyme complex were provided by Danisco US Inc., Genencor Division (Rochester, NY). All

solutions were prepared and rinses performed using ultrapure deionized (DI) water (18.2 $M\Omega cm^{-1}$) from an Easypure II water system (Barnstead Int.). The corn stover is Pioneer 34A20 grown in 2008 at the Sorenson Research Farm (Ames, IA).

2.2 Soxhlet extraction

The Soxhlet extraction of corn stover using water, ethanol, or hexane was performed as outlined in a National Renewable Energy Laboratory protocol (Sluiter et al., 2005). Approximately 5 g dry corn stover was loaded in a cotton cellulose thimble and placed in a glass Soxhlet extraction tube. A round-bottom flask containing 250 mL water, ethanol or hexane was connected to the extraction tube and refluxed for 24 hours.

2.3 Enzymatic hydrolysis reactions

Corn stover pretreatment by dilute acid or soaking in aqueous ammonia was performed as previously outlined (Shih & Smith, 2009). After soaking in aqueous ammonia, a neutralization step was performed by adding 5N HCl dropwise to the corn stover with mixing until the pH reached ~6. The biomass mixture was then centrifuged, the supernatant was removed and the corn stover was resuspended in water. The wash was repeated 6 times and the wet pretreated corn stover was used immediately for hydrolysis.

Enzymatic hydrolysis reactions were carried out in 0.1M acetate buffer, pH 4.8 with the addition of antibiotics. Approximately 7 g wet pretreated corn stover (corresponding to 1 g dry biomass) was placed in a 40 mL amber glass vial followed by the addition of 0.80 mL Accellerase 1500^{TM} enzyme, and/or 0.2 mL Accellerase XY enzyme, 200 µL cycloheximide, 30 µL tetracycline, and acetate buffer to bring the total volume to 13 mL. These vials were placed in an incubated shaker at 50 °C, 200 rpm for 48 hours. The samples were heated to 70 °C in a water bath for at least 10 minutes to terminate the enzymatic reaction at the end of the incubation period. Samples were then centrifuged at approximately 6200 g for 30 minutes and the supernatants were filtered using a 0.22 micron filter.

2.4 Raman spectroscopy measurements

A home-built Raman spectrometer consisting of a 785-nm diode laser, inverted optical microscope, spectrometer and an IR-enhanced CCD camera was used for all measurements as previously described (Shih & Smith, 2009). An 800 μm pinhole replaced the 100 μm pinhole used in previous studies. It is important to note that commercial Raman spectrometers with a 785 nm laser can be used to measure glucose or xylose yield in hydrolysate after calibration is performed. Near IR (e.g., 785 nm) excitation is critical to minimize sample fluorescence, which would render the Raman spectrum of hydrolysate undetectable at visible laser wavelengths. The Raman scatter was collected at room temperature using 10 second acquisitions with 20 accumulations for all measurements. The laser power for all measurements was approximately 130 mW at the sample.

2.5 Spectral analysis and statistical analysis

Spectra were collected using Winspec/32 version 2.5.21 (Roper Scientific, Germany) and were processed using GRAMS/AI 8.0 to perform background subtraction and in some cases baseline correction, as outlined in section 3.2. All spectra were background subtracted. The background resulted from spectral contributions from the instrument optics (measured with no sample in place), to some extent from the buffer and enzyme components (measured with sample matrix containing no biomass) and to a great extent from species that leach from the corn stover into the hydrolysates. The best background was found to be a Raman spectrum of a simulated hydrolysis reaction performed with heat inactivated enzyme but otherwise the same components used in a hydrolysis reaction (Supplemental Information Figure S1). All samples were measured in triplicate, unless otherwise noted.

2.6 Limit of glucose detection

The equivalent of 1 g dry biomass pretreated with dilute sulfuric acid (DA), soaking in aqueous ammonia (SAA), water extraction (E_w), ethanol extraction (E_e), hexane extraction (E_h), or in combination ($E_{w/e}$, $E_{w/e}$-DA, $E_{w/e}$-SAA) was used in a simulated hydrolysis experiment to create representative matrices for each pretreatment method. The simulated

hydrolysis was performed as described in section 2.3, except the enzyme was denatured prior to adding it to the reaction vials. After 48 hours, the mixtures were centrifuged at approximately 6200 g for 30 minutes and the supernatants were filtered using a 0.22 micron filter. These solutions were subsequently used to prepare standards with known glucose concentrations, and were measured by Raman spectroscopy.

The 1123 cm^{-1} spectral peak intensities for the series of standard solutions was used to construct the Raman spectroscopy calibration curves. The calibration data was fit using partial least square regression with standard deviation^{-2} weighting (Almeida et al., 2002). All calibration curves in this study had coefficients of determination (R^2) greater than 0.98, indicating that the concentration can be accurately predicted by the model. Detection limits were calculated by analyzing 7 replicate samples with concentrations near the detection limit and 1 order of magnitude higher, and computing 3 times the standard deviation of the replicate measurements divided by the slope of the calibration curve (Harris, 2007). Calculations for the uncertainty in the measurement of an unknown sample using the generated calibration curve is discussed in the supporting information.

2.7 Quantitative analysis of glucose and xylose

Standard solutions of xylose (10 to 100 mg mL^{-1}) or glucose (20 to 120 mg mL^{-1}) in hydrolysis broth were prepared by serial dilution and were measured by Raman spectroscopy in triplicate. The multi-peak fitting package 2.0 in Igor Pro 6.10A software (WaveMetrics. Inc.) was used to deconvolute glucose and xylose peaks. The peak shape was set to Gaussian. For each xylose or glucose standard, the location, intensity, and width of the peaks in the spectral region between 350 to 650 cm^{-1} were identified. Eight peaks were identified in this spectral region for glucose and six were identified for xylose. The intensity, as determined by deconvolution of the peaks at 519 cm^{-1} for glucose and 536 cm^{-1} for xylose, was used to generate a calibration curve.

Mixtures of glucose and xylose in hydrolysis broth were prepared in 2:1, 3:1 and 4:1 ratios at glucose concentrations between 15 and 100 mg mL^{-1} to test the calibration curves. These ratios represent the approximate amount of cellulose and hemicellulose in untreated

corn stover (2:1), and the composition after ammonia pretreatment (4:1). Raman spectra were collected for each mixture in triplicate, and the multi-peak fitting parameters were used to determine the glucose and xylose concentration. Specifically, fourteen peaks were used in the fit, and the Raman shift and width of each of the 14 peaks were set to the value obtained from the single component spectra. The peak intensities that minimized the residual between the data and the fit were identified. The calibration curves were then used to correlate the peak intensity with glucose or xylose concentration. This same method was used to measure the glucose and xylose concentration resulting from the cellulase or cellulase/xylanase hydrolysis of SAA pretreated corn stover.

2.8 UV-visible spectrophotometry measurements

Glucose HK assays were used as confirmatory tests of glucose concentration, (Bondar & Mead, 1974) and phloroglucinol-based colorimetric assays (Eberts et al., 1979; Johnson et al., 1984) were used to confirm xylose concentration. A calibration curve constructed with standard solutions was generated for each assay. A best fit line was generated by partial least square regression. All unknown samples were diluted to the linear range of the calibration curve before measurement, measured in triplicate and the average value ± 95% confidence interval was reported.

3. Results and discussion

3.1 The effect of biomass pretreatment on Raman spectral background for quantitative glucose measurements

Spectral background is the limiting factor for the quantitative measurement of glucose in corn stover hydrolysate by Raman spectroscopy using 785 nm excitation. The source of the spectral background in Raman measurements was systematically evaluated using several corn stover pretreatment protocols. Specifically, nine protocols were used in a simulated hydrolysis reaction that contained inactive enzyme (Figure 1), but was otherwise carried out using standard hydrolysis conditions. Soluble species present in the corn stover after pretreatment will leach into the hydrolysis broth. Insoluble components do not interfere with

the Raman spectral measurements since the samples are filtered prior to analysis. The filtered hydrolysis broth was used to prepare standard glucose samples to construct a Raman spectroscopy calibration curve, and the glucose limit of detection was calculated.

The pretreatment methods that were used in this study were dilute sulfuric acid (DA) and soaking in aqueous ammonia (SAA). In addition, Soxhlet extraction was utilized to remove water, ethanol or hexane extractives in the biomass, which represent approximately 3 to 6% of the total dry mass of corn stover (Sluiter et al., 2005). The Soxhlet extraction was performed to evaluate whether small molecule extractives contribute to the spectral background measured by Raman spectroscopy. Water extraction (E_w) removes inorganic materials, non-structural sugars, nitrogenous material and other minor components. Ethanol extraction (E_e) removes chlorophyll, waxes and other minor components. Hexane extraction (E_h) was used to remove nonpolar compounds. In addition, water and ethanol extraction ($E_{w/e}$) was used in combination. Extraction is not known to disrupt cell wall components or increase glucose yield, and is not commonly used as a pretreatment step in the bioethanol industry. Extraction with water and ethanol was also used in combination with dilute acid ($E_{w/e}$-DA) or soaking in aqueous ammonia pretreatment ($E_{w/e}$-SAA). Comparisons were performed with untreated corn stover.

The limit of glucose detection for each hydrolysate can be determined from the calibration curve constructed using the intensity of the 1123 cm^{-1} spectral peak (Table 1, Supporting Information Figure S1). This peak represents C-C and C-O stretching modes and is the most intense glucose peak in the Raman spectrum. The calibration curves and limits of detection depend on the spectral peak used to construct the curve. Since the relative standard deviation near the limit of detection is ~33%, only one significant figure is reported.

DA pretreated and untreated corn stover hydrolysis broths have the worst glucose detection limits (20 mg mL^{-1}) among the nine hydrolysates studied. Since the detection limits are the same for both sulfuric acid and untreated corn stover, this suggests that the most significant source of the spectral background that limits glucose quantitation is intrinsic to the biomass and not generated during acid pretreatment. The higher spectral background in

DA and untreated corn stover hydrolysate is also associated with the highest uncertainty in glucose quantitation by Raman spectroscopy.

Extracting water, ethanol or hexane soluble compounds from corn stover, without any further pretreatment, improves the glucose detection limit to 5-6 mg mL^{-1} and the uncertainty in glucose quantitation to 2-7 mg mL^{-1}, indicating that extractives contribute to the spectral background observed in the untreated biomass hydrolysis liquor. The glucose detection limit is similar after extraction with water, ethanol or hexane, suggesting that a single compound cannot be identified as the major spectral interferent. Rather, water, ethanol and hexane soluble extractives in combination contribute to the spectral background. Removing some of the interfering compounds is sufficient to reduce the glucose limit of detection to a value that is useful for measuring yield in corn stover hydrolysate.

It might be possible to further reduce spectral interference and the glucose detection limit by extraction using a combination of solvents. Extracting corn stover with water then ethanol does reduce the glucose detection limit to 4 mg mL^{-1}; however, this is not statistically different from the values obtained for extraction with ethanol or water alone. This suggests that the major spectral interferent is similarly removed by ethanol or water. Thammasouk *et al.* have analyzed the water and ethanol extraction products generated from corn stover (Thammasouk et al., 1997). They showed that the amount of extractives testing positive for Klason lignin, approximately 2.5% of the biomass dry weight, was similar using water or ethanol. Chen *et al.* have identified by reversed phase liquid chromatography a complex fraction of oligomeric species likely derived from phenolic glycosides in the water extract of corn stover (Chen et al., 1997). Compounds related to lignin with phenylpropanoid structure are likely to contribute to the Raman spectral background as they are strong Raman scatters and fluorescent, and the lignin spectrum obtained using similar conditions as used for these studies showed a very large background with no discernable peaks (Shih et al., 2009). It is reasonable that hexane can extract lignin or structurally related compounds from the corn stover, although this has not been quantitatively measured to our knowledge. Finally, aqueous 1% sulfuric acid used in the dilute acid pretreatment may affect the amount and

identity of species removed from the biomass compared to water extraction performed at neutral pH and elevated temperatures.

In order to confirm that extractives contribute to the Raman spectral background, a water and ethanol extraction was performed in combination with DA pretreatment. For $E_{w/e}$-DA pretreated corn stover the detection limit decreases to 6 mg mL^{-1}, which is of similar magnitude to the extracted corn stover alone. This indicates that glucose yields can be measured in DA pretreated hydrolysate by Raman spectroscopy if an extraction step is first performed. Ethanol extraction should minimally alter the hydrolysate's saccharide composition since ethanol extract is low in carbohydrate content (2.8% of the extract by dry weight) compared to water extract of corn stover (Thammasouk et al., 1997). An alternative to extracting interfering fluorescent compounds from the sample prior to Raman spectroscopy analysis is to illuminate the sample with an appropriate wavelength of light to cause photobleaching of the interferent (Macdonald & Wyeth, 2006). A potential drawback to this approach is unwanted sample photodegradation. If present, this can reduce the quality of the Raman spectrum through either increased background from char or degradation of the sugars.

SAA pretreatment, which removes approximately 85% of the lignin, produces the same detection limit (4 mg mL^{-1}) as that measured for water and ethanol extracted corn stover. This detection limit is lower than a previously reported value (8 mg mL^{-1}) due to a minor change in the Raman spectrometer (Section 2.6, Shih & Smith, 2009). A neutralization and rinse step after ammonia pretreatment was found to be crucial for lowering the detection limit, which was 16 mg mL^{-1} when neutralization was omitted. It is likely the extensive water rinse associated with the neutralization step reduces spectral interference; it is possible that the interferent is only removed near neutral pH values. The detection limit for corn stover that is $E_{w/e}$-SAA pretreated (6 mg mL^{-1}) and the uncertainty in glucose quantitation (6 mg mL^{-1}) is statistically similar to the value for SAA alone. SAA pretreatment is sufficient for glucose quantification by Raman spectroscopy, and prior extraction of the biomass is not required.

3.2 Simultaneous determination of glucose and xylose in hydrolysis liquor

Deconvoluting overlapping spectral peaks can be used to identify and quantify molecules that have overlapping spectral features. After glucose and cellobiose, xylose is the next most abundant hydrolysis product from corn stover (Yang et al., 2010). A multi-peak fitting model was developed using spectra of glucose and xylose standard solutions prepared in SAA hydrolysis broth at concentrations ranging from 10 to 100 mg mL^{-1}. The 350 to 650 cm^{-1} spectral region was chosen for deconvolution because xylose and glucose have relatively intense peaks with distinguishable peak maxima in this region. In order to perform the peak deconvolution, baseline correction is needed. This is most important at saccharide concentrations below 30 mg mL^{-1}. A linear baseline correction was performed for the 350 to 650 cm^{-1} spectral region. The linear correction has two numbers that need to be determined, the intercept and the slope. These parameters were set by ensuring the baseline correcting line passed through the data points at 378 and 628 cm^{-1}.

The spectra, multi-peak fits, and residuals between the experimental data and multi-peak fit for 100 mg mL^{-1} xylose or 120 mg mL^{-1} glucose solutions are shown in Figure 2A and B, respectively. In all cases, the residuals are of the same magnitude as the spectral noise, indicating the fit accurately models the experimental data. After deconvoluting the glucose or xylose spectra at all concentrations, the Raman shifts and the peak widths were averaged. These parameters are shown in Table 2. The Raman shift varied by no more than 1.1 cm^{-1}, and the peak width varied by no more than 1.2 cm^{-1} across all concentrations used to develop the average. Calibration curves were constructed for glucose (519 cm^{-1}) and xylose (536 cm^{-1}) using the peak intensity determined by spectral deconvolution (Supplemental Information Figure S3).

The calibration curves and multi-peak fit parameters were used to determine the glucose and xylose concentration in mixtures with known saccharide concentrations. As discussed above, the limit of detection will depend on the spectral peak used to construct the calibration curve. Using the multi-peak fitting algorithm on Raman spectra of SAA pretreated corn stover hydrosylate, the glucose limit of detection is 3 mg mL^{-1} with an

uncertainty of 2 mg mL^{-1} in glucose quantitation. The xylose detection limit is 1 mg mL^{-1} with a 1 mg mL^{-1} uncertainty in xylose quantitation.

The spectrum of a 60 mg mL^{-1} glucose and 50 mg mL^{-1} xylose mixture in SAA hydrolysis broth is shown in Figure 2C. A total of 14 peaks were used in the fit (8 peaks from glucose and 6 peaks from xylose). Using the set parameters listed in Table 2 and varying peak heights, a good fit is achieved between the data and the model, as shown by the residual. After baseline correction and peak deconvolution, the peak intensity and calibration curves were used to obtain the concentration of glucose and xylose. These were determined to be 64 ± 3 and 54 ± 3 mg mL^{-1}, respectively.

The multi-peak algorithm was then tested with mixtures of glucose and xylose at three ratios and glucose concentrations ranging from 15 to 100 mg mL^{-1}. These concentrations and ratios were selected based on expected yields in actual hydrolysis reactions where the cellulase is supplemented with hemicellulase. Using the procedure described above, the concentration of each sugar was determined and is reported in Table 3. The results reveal that for nearly every sample with a total saccharide concentration above 45 mg mL^{-1}, the measured glucose and xylose concentration agrees with the expected concentration at the 95% confidence interval with an average relative standard error of 6.1%. Samples with a total saccharide concentration below 40 mg mL^{-1} are measured at the expected concentration within the 95% confidence interval, with a higher average relative standard error of 22.1%.

The ability to use the multi-peak fit method to measure glucose and xylose yield in corn stover hydrolysis reactions was evaluated. The SAA-pretreated corn stover was chosen for this experiment because dilute acid pretreatment removes a majority of the hemicellulose, reducing the expected xylose yield. Three enzyme formulations were used for these studies: Accellerase 1500 (cellulase), Accellerase 1500/XY (hemicellulase), and XY. The accessory enzyme XY has high activity towards xylan. A sample containing all reaction components except active enzyme was used as a blank, as described in Section 2.5.

The saccharide spectra of SAA hydrolysate generated from the Accellerase 1500 enzyme and the Accellerase 1500/XY enzyme mixture were deconvoluted as described above. The hydrolysis yields are shown in Table 4. The sample containing only enzyme XY did not show interpretable spectral peaks, possibly because the xylan was not accessible to the xylanase in the presence of intact cellulose.

The glucose yield measured by Raman spectroscopy did not significantly differ when only cellulase was present (32 ± 4 mg mL^{-1}) or when both cellulase and hemicellulase were present (31 ± 4 mg mL^{-1}) in the reaction. The Raman spectroscopy results were confirmed using glucose-specific HK assays. The results obtained by UV- visible spectrophotometry agreed with the Raman spectroscopy results within the 95% confidence interval. In the literature, it has been reported that supplementation of cellulase with xylanase could improve glucose yield (Kumar & Wyman, 2009). The increase in glucose production depended on the concentration of the enzyme and the pretreatment method. Ammonia recycled percolation pretreatment produced the lowest increase in glucose with hemicellulase supplementation, and this is closest to the pretreatment method used in these studies. This is a possible explanation for why an increase in glucose was not measured by Raman spectroscopy or glucose HK assays when the cellulase was supplemented with hemicellulase. Additionally, the Accellerase 1500 enzyme has some reported hemicellulase activity, so additional supplementation with hemicellulase may not improve glucose yields.

The fact that the same glucose yield was measured by Raman spectroscopy and the glucose specific HK assay suggested that partial cellulose hydrolysis products (e.g., cellobiose) did not interfere with the Raman measurements and that cellobiose spectral contribution was negligible or not present. It is insightful to analyze the glucose and cellobiose Raman spectra (Supplemental Figure S4) to determine under what conditions cellobiose will interfere with glucose quantification using the multi-peak fitting algorithm. For the same concentration of glucose and cellobiose, the spectral intensity for cellobiose is 0.181 times the glucose spectral intensity at 519 cm^{-1}, the peak used to construct the glucose calibration curve. The overestimation of glucose yield due to the presence of cellobiose is at most: (cellobiose yield*0.181)/glucose yield. As an example, in a typical hydrolysis reaction

the glucose yield is 175 mg mL^{-1} and the cellobiose yield is 22 mg mL^{-1} (Yang et al., 2010). This represents a 2.3% error in the glucose concentration. Analysis of the glucose and cellobiose Raman spectra (Supporting Information Figure S4) reveals that it is possible to develop a spectral deconvolution method to measure cellobiose independently of glucose if the error is higher than can be tolerated.

As measured by Raman spectroscopy, the addition of hemicellulase to the hydrolysis reaction increased xylose yield by ~2.5 times compared to the reaction with only cellulase present. This increase is consistent with literature reports (Kumar & Wyman, 2009). The error in xylose determination associated with the presence of cellobiose can be determined in a manner similar to that described for glucose: (cellobiose yield*0.080)/xylose yield. A phloroglucinol colorimetric assay was used to confirm the concentration of xylose measured by Raman spectroscopy. This assay (Johnson et al., 1984) has little interference from glucose or proteins (e.g., cellulase and hemicellulase). The UV-visible spectrophotometry results agree with the Raman spectroscopic data within the 95% confidence interval, and suggest that Raman spectroscopy can be used to simultaneously measure the yield of glucose and xylose in corn stover hydrolysate.

In summary, the 1123 cm^{-1} Raman spectral peak intensity can be used to measure glucose in SAA pretreated corn stover with a 4 mg mL^{-1} limit of detection and 4 mg mL^{-1} uncertainty. Provided the corn stover is first extracted to remove spectrally interfering compounds, DA pretreated corn stover hydrolysate can be measured by the same method with a detection limit and uncertainty of 6 mg mL^{-1}. The results presented herein suggest that the amount of extractable compounds in the biomass at the start of the hydrolysis is the most important factor governing the glucose detection limit in the hydrolysate. It has been demonstrated that both glucose and xylose can be simultaneously measured in SAA corn stover hydrolysate using a multi-peak fitting procedure. Although this procedure requires additional time for data analysis compared to the use of the 1123 cm^{-1} peak intensity, additional information about xylose yield is gained. With the multi-peak fit method, the glucose limit of detection is 3 mg mL^{-1} with 2 mg mL^{-1} uncertainty and the xylose limit of detection is 1 mg mL^{-1} with 1 mg mL^{-1}uncertainty.

If glucose is the only analyte in the hydrolysates, a single peak analysis of the data will provide accurate, rapid results. However, if other saccharides are present in the hydrolysates, a single peak analysis can overestimate the glucose yield and the multi-peak fit should be used. The multi-peak fit will increase the analysis time to several minutes per sample since baseline correction and the fit have to be performed, although the possibility of automation exists. The discussion in Section 3.1 regarding the effects pretreatment and extraction have on Raman analyses of corn stover hydrolysate are generally valid for the multi-peak fitting procedure. The background at the spectral region used for the multi-peak fitting analysis (around 500 cm^{-1}) behaves similarly to the background around 1100 cm^{-1} where the single peak analysis is performed. The multi-peak fit parameters can also be used to measure glucose and xylose yield with E-DA pretreated corn stover; however, the expected xylose yield should first be estimated and compared with the detection limit.

4. Conclusions

Raman spectroscopy has several benefits over alternative analysis methods for hydrolysate: it requires minimal sample preparation and analysis time, it is non-invasive so the yield can be measured in the same sample over time, data collection/analysis can be automated and multiple sugars can be simultaneously measured. Current efforts are focused on extending the use of Raman spectroscopy to measure yields of other sugars found in lower abundance than glucose and xylose and to other biomass feedstocks.

Acknowledgements

Work at the Ames Laboratory was supported by the Department of Energy-Basic Energy Sciences under Contract No. DE-AC02-07CH11358. The authors are grateful to Dr. Kenneth J. Moore for providing corn stover and Danisco US Inc., Genencor Division, for providing the Accellerase enzymes.

Appendix A. Supplementary data

References

Agarwal, U.P., 2006. Raman imaging to investigate ultrastructure and composition of plan cell walls: distribution of lignin and cellulose in black spruce wood (Picea mariana). *Planta*, **224**, 141-1153.

Agblevor, F.A., Hames, B.R., Schell, D., Chum, H.L. 2007. Analysis of biomass sugars using a novel HPLC method. *Appl. Biochem. Biotechnol.*, **136**(3), 309-326.

Almeida, A.M., Castel-Branco, M.M., Falcao, A.C. 2002. Linear Regression for Calibration Lines Revisted: Weighted Schemes for Bioanalytical Methods. *J. Chromatography B*, **774**, 215-222.

Bondar, R.J., Mead, D.C. 1974. Evaluation of glucose-6-phosphate dehydrogenase from Leuconostoc mesenteroides in the hexokinase method for determining glucose in serum. *Clin. Chem.*, **20**(5), 586-90.

Brooker, M.H., Hancock, G., Rice, B.C., Shapter, J. 1989. Raman frequency and intensity studies of liquid water (H_2O, H_218O and D_2O). *J. Raman Spectrosc.*, **20**(10), 683-94.

Cara, C., Ruiz, E., Oliva, J.M., Sáez, F., Castro, E. 2008. Conversion of olive tree biomass into fermentable sugars by dilute acid pretreatment and enzymatic saccharification. *Bioresour. Technol.*, **99**(6), 1869-1876.

Chen, M., Zhao, J., Xia, L. 2009. Comparison of four different chemical pretreatments of corn stover for enhancing enzymatic digestibility. *Biomass Bioenergy*, **33**(10), 1381-1385.

Chen, S.-F. Mowery, R.A. Scarlata, C.J. Chambliss, C.K. 2007 Compositional Analysis of water-soluble materials in corn stover. *J. Agric. Food Chem*. **55** 5912-5918.

Eberts, T.J., Sample, R.H.B., Glick, M.R., Ellis, G.H. 1979. A simplified, colorimetric micromethod for xylose in serum or urine, with phloroglucinol. *Clin. Chem. (Winston-Salem, N. C.)*, **25**(8), 1440-3.

Eggeman, T., Elander, R.T. 2005. Process and economic analysis of pretreatment technologies. *Bioresour. Technol.*, **96**(18), 2019-2025.

Gierling, N., Schwanninger, M. 2006 Chemical imaging of poplar cell walls by confocal Raman microscopy. *Plant Physiology*, **140**, 1246-1254.

Harris, D.C. 2007. *Quantitative Chemical Analysis*. *7th ed*. Craig Bleyer, New York.

Huber, C., Grill, E., Oefner, P., Bobleter, O. 1994. Capillary electrophoretic determination of the component monosaccharides in hemicelluloses. *Fresenius' J. Anal. Chem.*, **348**(12), 825-31.

Johnson, S.L., Bliss, M., Mayersohn, M., Conrad, K.A. 1984. Phloroglucinol-based colorimetry of xylose in plasma and urine compared with a specific gas-chromatographic procedure. *Clin. Chem. (Winston-Salem, N. C.)*, **30**(9), 1571-4.

Kamm, B., Kamm, M., Schmidt, M., Starke, I., Kleinpeter, E. 2006. Chemical and biochemical generation of carbohydrates from lignocellulose-feedstock (Lupinus nootkatensis)-quantification of glucose. *Chemosphere*, **62**(1), 97-105.

Kim, T.H., Kim, J.S., Sunwoo, C., Lee, Y.Y. 2003. Pretreatment of corn stover by aqueous ammonia. *Bioresour. Technol.*, **90**(1), 39-47.

Kim, T.H., Lee, Y.Y. 2007. Pretreatment of corn stover by soaking in aqueous ammonia at moderate temperatures. *Appl. Biochem. Biotechnol.*, **137-140**, 81-92.

Knauf, M., Moniruzzaman, M. 2004. Lignocellulosic biomass processing: A perspective. *Int. Sugar J.*, **106**(1263), 147-150.

Ko, J.-J., Shimizu, Y., Ikeda, K., Kim, S.-K., Park, C.-H., Matsui, S. 2008. Biodegradation of high molecular weight lignin under sulfate reducing conditions: Lignin degradability and degradation by-products. *Bioresour. Technol.*, **100**(4), 1622-1627.

Koshijima, T., Watanabe, T. 2003. *Association Between Lignin and Carbohydrates in Wood*. Springer- Verlag Berlin Heidelberg, New York.

Kumar, R., Wyman, C.E. 2009. Effect of xylanase supplementation of cellulase on digestion of corn stover solids prepared by leading pretreatment technologies. *Bioresour. Technol.*, **100**(18), 4203-4213.

Kumar, R., Wyman Charles, E. 2009. Cellulase adsorption and relationship to features of corn stover solids produced by leading pretreatments. *Biotechnol Bioeng.*, **103**(2), 252-67.

Li, C., Knierim, B., Manisseri, C., Arora, R., Scheller, H.V., Auer, M., Vogel, K.P., Simmons, B.A., Singh, S. 2009. Comparison of dilute acid and ionic liquid pretreatment of switchgrass: Biomass recalcitrance, delinification and enzymatic saccharification. *Bioresource Technology.* **101**(13), 4900-4906.

Lloyd, T.A., Wyman, C.E. 2005. Combined sugar yields for dilute sulfuric acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids. *Bioresour. Technol.*, **96**(18), 1967-1977.

Macdonald, A.M., Wyeth, P. 2006. On the use of photobleaching to reduce fluorescence background in Raman spectroscopy to improve the reliability of pigment identification on painted textiles. *J. Raman Spectroscopy*, **37**, 830–835.

Méchin, V., Argillier, O., Rocher, F., Hébert, Y., Mila, I., Pollet, B., Barriére, Y., Lapierre, C. 2005. In search of a Maize ideotype for cell wall enzymatic degradability using histological and biochemical lignin characterization. *J. Agric. Food Chem.*, **53**(15), 5872-5881.

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., Ladisch, M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.*, **96**(6), 673-686.

Öhgren, K., Vehmaanperä, J., Siika-Aho, M., Galbe, M., Viikari, L., Zacchi, G. 2007. High temperature enzymatic prehydrolysis prior to simultaneous saccharification and fermentation of steam pretreated corn stover for ethanol production. *Enzyme Microb. Technol.*, **40**(4), 607-613.

Pandey, A., Editor. 2009. *Handbook of Plant-Based Biofuels*.

Raymond, R., Ehrman, T. 1996. HPLC Analysis of Liquid Fractions of Process Samples for Monomeric Sugars and Cellobiose. National Renewable Energy Laboratory.

Shih, C.-J., Smith, E.A. 2009. Determination of glucose and ethanol after enzymatic hydrolysis and fermentation of biomass using Raman spectroscopy. *Anal. Chim. Acta*, **653**(2), 200-206.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. 2006. Determination of sugars, byproducts, and degradation products in liquid fraction process samples. *NREL Laboratory Analytical Procedure (LAP)*(NREL/TP-510-42623).

Sluiter, A., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. 2005. Determination of extractives in biomass. *NREL Laboratory Analytical Procedure (LAP)*(NREL/TP-510-42619).

Sun, L., Simmons, B.A., Singh, S. 2011. Understanding tissue specific compositions of bioenergy feedstocks through hyperspectral Raman imaging. *Bioresource Techology*. **108**(2), 286-295.

Tatsumi, H., Katano, H. 2004. Kinetic analysis of enzymatic hydrolysis of raw starch by glucoamylase using an amperometric glucose sensor. *Chem. Lett.*, **33**(6), 692-693.

Thammasouk, K., Tandjo, D., Penner, M.H. 1997. Influence on the analysis of herbaceous biomass. *J. Agric. Food Chem.* **45**, 437-443.

Weiss, N.D., Farmer, J.D., Schell , D.J. 2010. Impact of corn stover composition on hemicellulose conversion during dilute acid pretreatment and enzymatic cellulose digestibility of the pretreated solids. *Bioresource Technology*, **101**, 674-678.

Wyman, C.E., Dale, B.E., Elander, R.T., Holtzapple, M., Ladisch, M.R., Lee, Y.Y. 2005. Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover. *Bioresour. Technol.*, **96**(18), 2026-2032.

Yang, M., Li, W., Liu , B., Li, Q., Xing, J. 2010. High-concentration sugars production from corn stover based on combined pretreatments and fed-batch process. *Bioresource Technology*, **101**, 4884-4888.

Zhang, X., Young, M.A., Lyandres, O., Van Duyne, R.P. 2005. Rapid detection of an anthrax biomarker by surface-enhanced Raman spectroscopy. *J. Am. Chem. Soc.*, **127**(12), 4484-4489.

Table 1: Glucose detection limits and average uncertainties for determining glucose concentration in corn stover hydrolysate generated using the indicated preprocessing protocols measured by Raman spectroscopy.

Preprocessing Protocol[a]	Limit of Detection (mg mL^{-1})	Uncertainty in Unknown Determination (mg mL^{-1})[b]
E_w	5	2
E_e	6	7
$E_{w/e}$	4	4
E_h	5	4
$E_{w/e}$ – DA	6	6
$E_{w/e}$ – SAA	6	6
SAA	4	4
DA	20	28
Untreated	20	22

a. Abbreviations are defined in Section 2.4

b. Assuming 3 replicate measurements of unknown, and an unknown concentration between 15 to 150 mg mL^{-1}. Equation used to calculate uncertainty is provided in the supplemental information.

Table 2: Raman shifts and peak widths calculated from glucose and xylose standards that were used for the quantitative analysis of hydrolysis products.

Peak	Raman Shift (cm^{-1})	Width (cm^{-1})	Analyte
0	408	16.5	Glucose
1	417	22.7	Xylose
2	423	7.59	Glucose
3	448	20.9	Glucose
4	471	16.2	Xylose
5	500	18.6	Glucose
6	508	10.6	Xylose
7	519	10.7	Glucose
8	536	13.6	Xylose
9	540	9.89	Glucose
10	558	15.8	Xylose
11	562	14.5	Glucose
12	589	13.8	Glucose
13	602	14.3	Xylose

Table 3: Glucose and xylose concentrations measured using multi-peak fitting on the Raman spectra of standard solutions prepared in SAA hydrolysis broth.

Mixture Composition	Measured Glucose (mg mL^{-1})[a]	Measured Xylose (mg mL^{-1})[a]
100 mg mL^{-1} glucose 50 mg mL^{-1} xylose	106 ± 9	51 ± 2
100 mg mL^{-1} glucose 25 mg mL^{-1} xylose	108 ± 3	25 ± 1
60 mg mL^{-1} glucose 30 mg mL^{-1} xylose	60 ± 3	28 ± 3
60 mg mL^{-1} glucose 20 mg mL^{-1} xylose	59 ± 3	19 ± 1
60 mg mL^{-1} glucose 15 mg mL^{-1} xylose	61 ± 4	14 ± 1
30 mg mL^{-1} glucose 15 mg mL^{-1} xylose	29 ± 3	15 ± 1
30 mg mL^{-1} glucose 10 mg mL^{-1} xylose	28 ± 3	9 ± 3
30 mg mL^{-1} glucose 7.5 mg mL^{-1} xylose	32 ± 2	7 ± 3
15 mg mL^{-1} glucose 7.5 mg mL^{-1} xylose	14 ± 4	7 ± 2
15 mg mL^{-1} glucose 5 mg mL^{-1} xylose	15 ± 2	5 ± 1
15 mg mL^{-1} glucose 3.8 mg mL^{-1} xylose	15 ± 2	4 ± 1

Table 4: Measured glucose and xylose hydrolysis yields for SAA-pretreated corn stover treated with cellulase (Accellerase 1500) or both cellulase and hemicellulase (Accellerase 1500 + XY).

Analysis Method Analyte	Accellerase 1500 Hydrolysis Yield (mg mL^{-1})[a]	Accellerase 1500 + XY Hydrolysis Yield (mg mL^{-1})[a]
Raman Spectroscopy glucose	32 ± 4	31 ± 4
UV-Vis Spectroscopy glucose	32 ± 4	30 ± 10
Raman Spectroscopy xylose	7.0 ± 0.8	18.0 ± 0.5
UV-Vis Spectroscopy xylose	6.6 ± 0.8	19 ± 6

a. All solutions were measured in triplicate. Average ± 95% confidence interval.

Figure 1: Overview of the measurements performed to evaluate Raman spectroscopy as an analysis method for glucose and xylose yield in corn stover hydrolysate after using different pretreatment protocols. The glucose detection limit and uncertainty was measured for 9 hydrolysates. A multi-peak fitting algorithm was developed to quantify glucose and xylose from the hydrolysate of aqueous ammonia-pretreated corn stover.

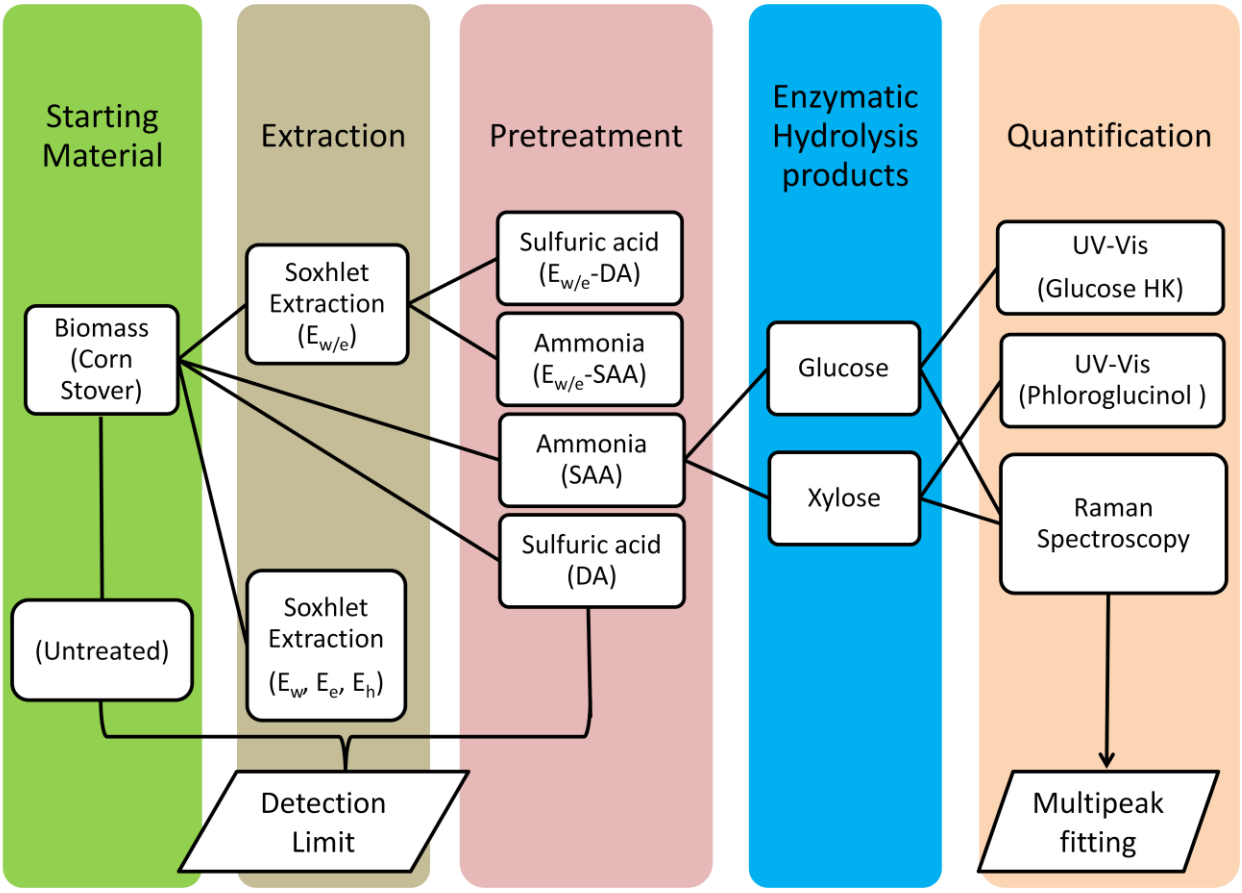
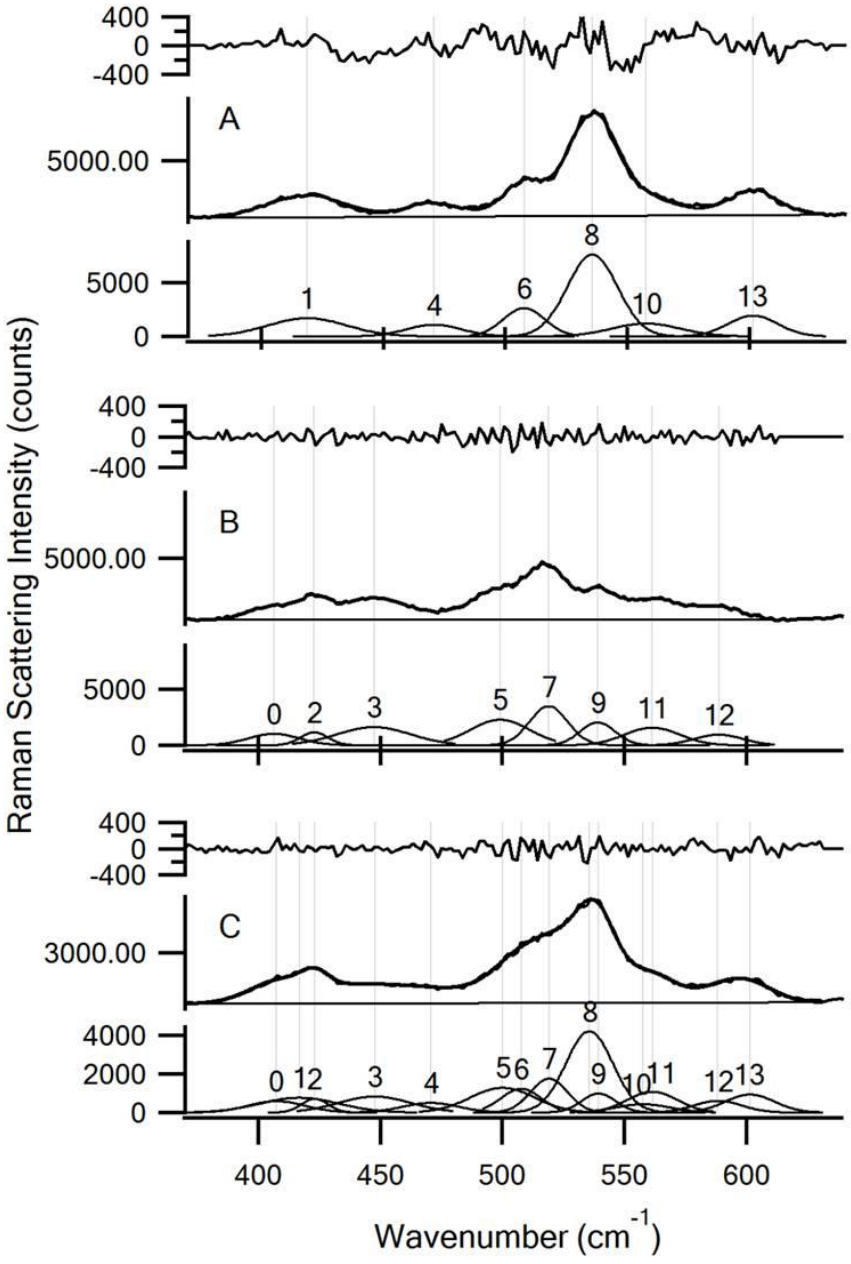


Figure 2: Multi-peak fitting results for (A) 100 mg mL^{-1} xylose; (B) 120 mg mL^{-1} glucose; and (C) a 60 mg mL^{-1} glucose and 50 mg mL^{-1} xylose mixture in SAA hydrolysis broth. The residual between the multi-peak fit and the experimental data is shown in the top panel. The residual is of the same order of magnitude as the spectral noise. The experimental spectra are shown in the middle panel (thicker line) and the fit results are shown in the bottom panel (thinner line). The indicated peak numbers correspond to the peaks identified in Table 2.



SUPPLEMENTAL INFORMATION

RAMAN SPECTROSCOPY MEASUREMENTS OF GLUCOSE AND XYLOSE IN HYDROLYSATE: ROLE OF CORN STOVER PRETREATMENT AND ENZYME COMPOSITION

*Chien-Ju Shih, Jason S. Lupoi and Emily A. Smith**

Ames Laboratory, U.S. Department of Energy, Ames, Iowa 50011-3111, U.S.A.

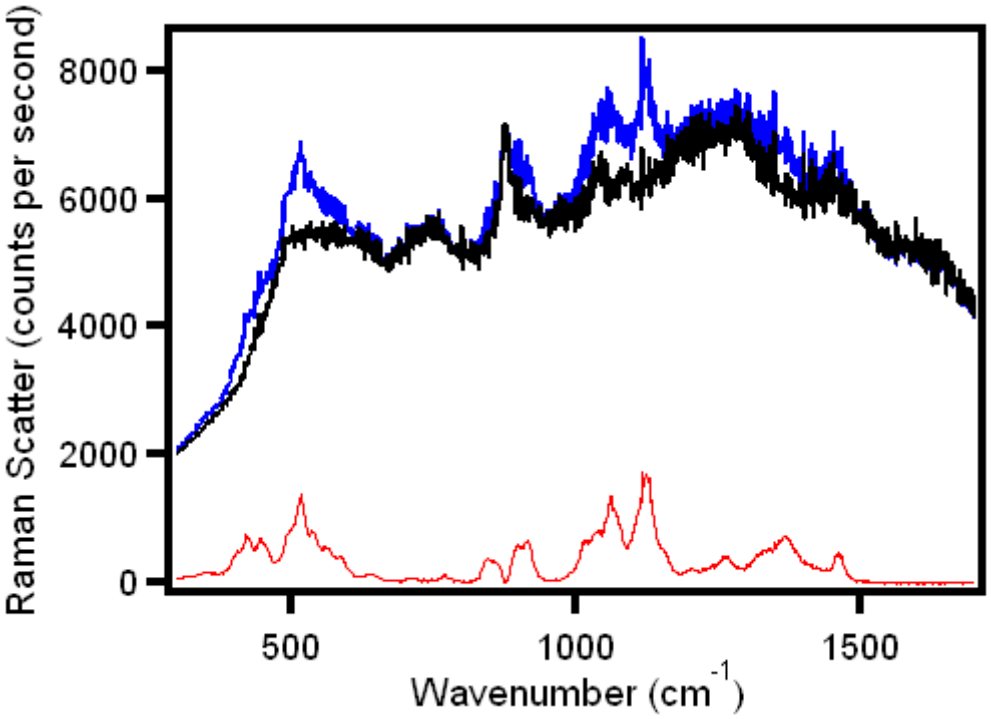
Department of Chemistry, Iowa State University, Ames, Iowa 50011-3111, U.S.A.

Supplemental Methods. The equation to calculate the uncertainty when using the Raman spectroscopy glucose calibration curves is:

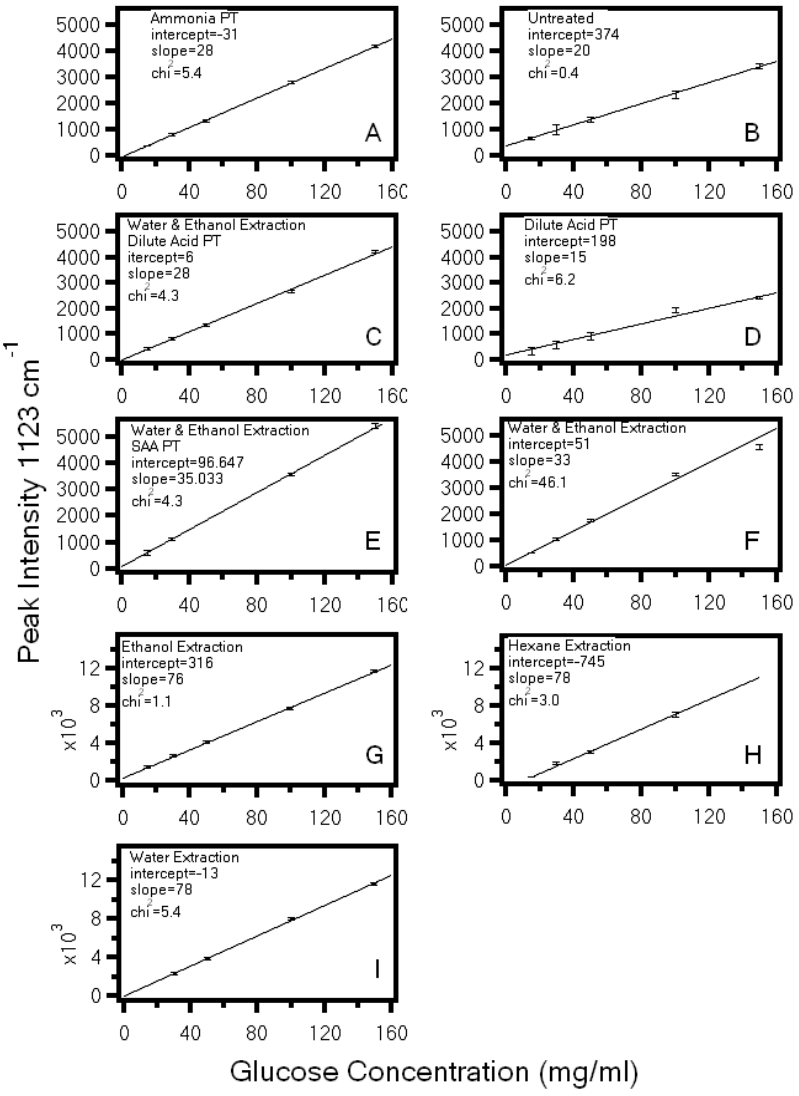
$$\text{Uncertainty} = \frac{s_{avg}}{|m|}\sqrt{\frac{1}{k}+\frac{1}{n}+\frac{(y-\bar{y})^2}{m^2\sum(x_i-\bar{x})^2}}$$

where s_{avg} is the average standard deviation in the Raman peak height; m is the slope of the weighted linear fit to the calibration data; n is the number of data points in the calibration curve; k is the number of replicate measurements taken for the unknown (assumed to be 3); y-bar is the average Raman peak height for all standard solutions used to construct the calibration curve; x_i are the glucose concentrations used to construct the calibration curve; and x-bar is the average of all the glucose concentrations used to construct the calibration curve.

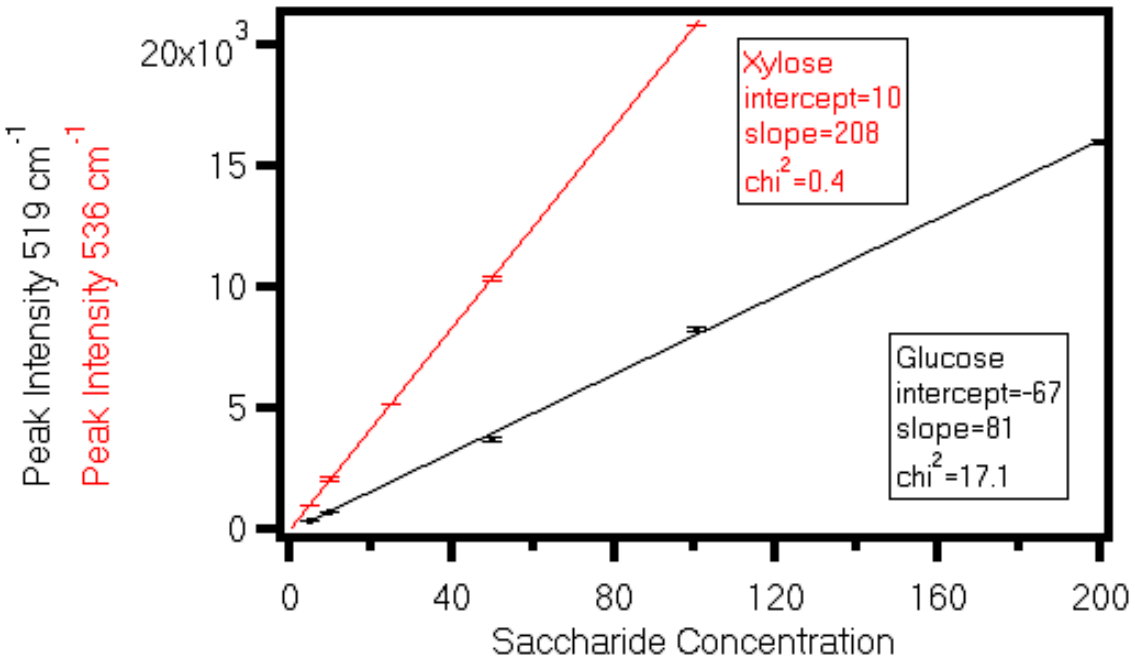
Supplemental Figure S1. Raman spectra of 'hydrolysate' obtained using heat inactivated enzyme and corn stover that had been previously extracted with hot water. The use of heat inactivated enzyme enabled control over the amount of glucose in the 'hydrolysate' using standard addition. 'Hydrolysate' containing: (blue) 150 mg ml^{-1} glucose and (black) no detectable glucose. The spectrum of 'hydrolysate' containing no added glucose (black) was used to subtract the background from the spectrum of glucose (or glucose plus xylose, as appropriate) containing 'hydrolysate' (blue) to generate the resultant spectrum (red) used to report values in the manuscript. All of the peaks in the background subtracted Raman spectrum can be assigned to glucose, indicating that the background subtraction protocol successfully removes spectral contributions from the sample matrix and instrument optics. If the background spectral intensity is equal to or greater than the spectral intensity from the glucose peaks, then glucose measurements cannot be obtained. Acquisition time 10 seconds with 20 accumulations; excitation wavelength 785 nm; excitation power 130 mW.



Supplemental Figure S2. Glucose calibration curves measured by Raman spectroscopy. The standard glucose solutions were prepared in hydrolysate using corn stover with the following pretreatment/extraction: (A) SAA pretreatment; (B) untreated; (C) $E_{w/e}$-DA; (D) DA PT; (E) $E_{w/e}$-SAA; (F) $E_{w/e}$; (G) E_e; (H) E_h; (I) E_w. Error bars represent 1 standard deviation in the 1123 cm^{-1} peak height for three replicate measurements. Weighted linear regression was used to find the best fit line. The calibration curves shown in (G), (H) and (I) were collected on separate days after altering the instrument alignment, which affects the collected signal and background, but not the detection limit or uncertainty. Further details can be found in the main text.



Supplemental Figure S3. Glucose (black) and xylose (red) calibration curves measured by Raman spectroscopy. The standard glucose or xylose solutions were prepared in SAA corn stover hydrolysate. Peak intensities were measured by deconvolution of the 350 to 650 cm^{-1} spectral region, as described in the text. Error bars represent 1 standard deviation in the 519 cm^{-1} or 536 cm^{-1}peak height for three replicate measurements. Weighted linear regression was used to find the best fit line.



Supplemental Figure S4. (A) Normalized, background subtracted Raman spectra of 50 mg mL^{-1} glucose (red) and 50 mg mL^{-1} cellobiose (black) in soaking in aqueous ammonia (SAA) pretreated corn stover hydrolysate. For the multi-peak fit analysis: the overestimation of glucose yield due to the presence of cellobiose is at most: (cellobiose yield*0.181)/glucose yield. **(B)** Same data shown in panel (A) except the spectra have been scaled by a constant factor to represent a 175 mg mL^{-1} glucose concentration and a 22 mg mL^{-1} cellobiose concentration, which are representative of concentrations that have been measured in corn stover hydrolysate. Acquisition time 10 seconds with 20 accumulations; excitation wavelength 785 nm; excitation power 130 mW. (Note: for the single peak analysis using the peak intensity at 1123 cm^{-1}, the overestimation of glucose yield due to the presence of cellobiose is at most (cellobiose yield*0.407)/glucose yield.)